Nuclear RNA forms an interconnected network of transcription-dependent and tunable microgels

Mattia Marenda[1,2*], Davide Michieletto[1,2*], Rafal Czapiewski[1], Jon Stocks[1], Sophie M Winterbourne[3], Jamilla Miles[1,3], Olivia CA Fleming[1,3], Elena Lazarova[1], Michael Chiang[2], Graeme R Grimes[1], Hannes Becher[1], Atlanta Cook[3], Davide Marenduzzo[2†], Ryu-Suke Nozawa[1,4†], Nick Gilbert[1†]

[1] MRC Human Genetics Unit, Institute of Genetics and Cancer, University of Edinburgh, Edinburgh, UK

[2] School of Physics and Astronomy, University of Edinburgh, Edinburgh, UK

[3] Institute for Biology, Biochemistry and Biotechnology, University of Edinburgh, UK

[4]Division of Experimental Pathology, Cancer Institute of the Japanese Foundation for Cancer Research (JFCR), Ariake 3-8-31 Koto-ku, 135-8550 Tokyo, Japan

* These authors contributed equally

[†]Correspondence to

Nick Gilbert: nick.gilbert@ed.ac.uk

Ryu-Suke Nozawa: ryusuke.nozawa@jfcr.or.jp

Davide Marenduzzo: davide.marenduzzo@ed.ac.uk

Highlights

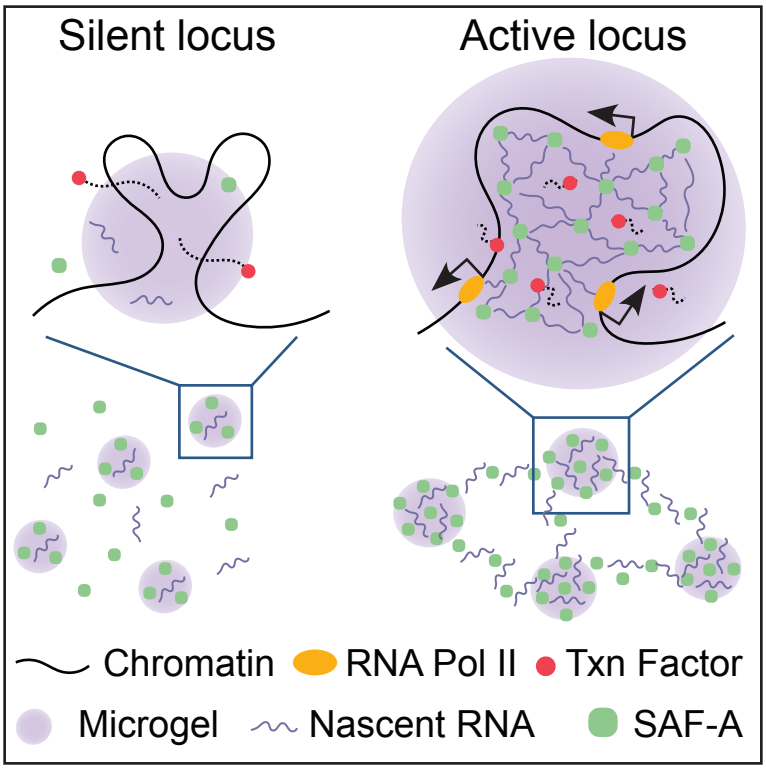
RNA and SAF-A interact to form clusters that form a nuclear-spanning network of microgels

Emergent microgel network requires transcription and XRN2 activity to undergo gelation and fluidization

Microgel network impacts nuclear protein mobility

Molecular dynamics modelling shows RNA/SAF-A microgels regulate chromatin decompaction by steric hinderance

Graphical Abstract



Summary

The human cell nucleus is comprised of proteins, chromatin and RNA, yet how they interact to form supramolecular structures and drive key biological processes remains unknown. Conflicting models have proposed either a fluid-like or solid-like nature for the intranuclear microenvironment. To reconcile this discrepancy, we investigated the 3D structure and properties of the nuclear interior using experiments and computer simulations. We reveal a novel mechanism where newly synthesized RNA interacts with SAF-A (scaffold attachment factor A, or HNRNPU), forming interconnected microgels degraded by the exonuclease XRN2, leading to dynamic cycles of gelation and fluidization. This emergent microgel network depends on transcription, and is disrupted by SAF-A depletion. It also decreases protein mobility and regulates chromatin compaction by modulating microphase separation, thereby opening transcriptionally active regions. This tunable intranuclear network exhibits scale-dependent fluid- and solid-like features, that we suggest may regulate transcription by controlling access to regulatory proteins and polymerases.

Keywords

Chromatin, nuclear architecture, transcription, SAF-A, HNRNPU, RNA, XRN2, bridging-induced attraction, microphase separation

Introduction

The nuclear interior is a complex microenvironment formed from chromatin, ribonucleoproteins, enzymes and nucleotides[1–4]. This intricate milieu facilitates key nuclear processes such as transcription and replication, but also controls access to regulatory factors. Although the individual components of the nuclear interior are well-characterized, its biophysical properties are currently debated, with two very different models suggesting that it either forms a solid-like structure or that it has fluid-like properties.

On one hand, we know from many classical studies that biochemical extraction of DNaseI-treated nuclei with high salt or detergent followed by cytological imaging reveals a complex network of 10-nm interacting filaments, indicating a nuclear matrix or 'scaffold'[5,6], consistent with a solid-like model. In this regime, the filaments in the nuclear matrix are decorated with heterogenous ribonucleoproteins (hnRNPs) and hard-to-extract pervasive RNAs, that are enriched in introns, repeat-rich and non-coding sequences[7,8].

On the other hand, although early studies of a nuclear matrix suggested it has a stable structure, these results belied the observation that many nuclear proteins are highly mobile[9–11], and that attaching fluorescent tags to nuclear proteins failed to reveal a pattern of staining reminiscent of the RNP-decorated filaments seen by electron microscopy. These observations instead suggest the nuclear microenvironment is fast responding to biochemical signals and highly dynamic, supporting a fluid-like model.

Biophysical extraction of the nucleus under harsh conditions has enabled the identification of key structural components of the elusive nuclear matrix, including matrins[12], SAF-B[13] and most notably hnRNP proteins, including a 90 kDa ubiquitously expressed polypeptide termed SAF-A (scaffold attachment factor A, also known as HNRNPU)[14,15]. hnRNPs are a family of heterogeneous RNA-binding proteins, with more than 20 different members identified in humans alone[16]. They have been implicated in a wide range of cellular processes, and their dysfunction has been linked to numerous human diseases, including cancer and neurological disorders[16,17]. One such protein, HNRNPA1, is composed of two RRM domains and RGG motifs[18,19], embedded in a intrinsically disordered region (IDR). Experiments *in vitro* show the IDR in the presence of RNA molecules can undergo phase separation into protein-rich droplets[20–22]. SAF-A is one of the most unique hnRNPs, it possesses a weak DNA binding domain, SPRY domain[23], ATP binding site, extensive IDRs, and RGG motifs that bind RNA[15,24,25]. It has been reported to be involved in splicing[26,27], X-inactivation[28,29], mitosis[30–33] and is important for regulating 3D chromatin structure[7,34,35] through interacting with RNA[2,35]. *Saf-a* knockout mouse models show embryonic lethality[36], whilst conditional deletion in the developing brain leads to rapid cell death of both postmitotic neurons and neural progenitors[37]. *Saf-a* haploinsufficiency also affects the expression of the neuropeptides arginine vasopressin (*Avp)* and vasoactive intestinal polypeptide (*Vip)* resulting in changes in metabolic activity[38] and cell type specific changes in transcription[39,40]. SAF-A is also clinically relevant, as patients with mutations in the gene have neurodevelopmental disorders[17,41–43]. In human iPS cells derived from a patient with

a SAF-A mutation, after neural differentiation there is aberrant chromatin reorganisation and altered transcription leading to a change in neuronal cell populations[44].

Cytologically, the most pronounced component of the nuclear interior is chromatin, a macromolecular complex of histone proteins and DNA. While understanding 3D chromatin structure is a key question in biology and biophysics due to its potential link to transcription, the relation between chromatin structure and the nuclear microenvironment is ill-understood. Recent microscopy-based studies in cells[45–47] and in vitro experiments[48,49] have examined chromatin mobility and come to contrasting conclusions. An analysis of FRAP recovery times indicate that nucleosomal domains are relatively immobile indicating that chromatin has a solid-like structure[45], whilst imaging of nucleosome pairs indicate that chromatin is condensed, with liquid-like characteristics[47], suggesting that chromatin at different scales has different physical characteristics. Interactions between chromatin and other nuclear components would also be expected to influence structure. For example, micromanipulation of a gene-locus using a magnet indicated that chromatin displacement is consistent with a rouse polymer model suggesting it has limited interactions with other structures, and indicative of the nucleus having fluid-like properties[50]. Similarly, studies suggest that in vitro reconstituted chromatin under most conditions can phase separate to form condensates which behave as liquid-like structures[48,49].

Transcriptionally active regions of chromatin often coalesce within the nucleus, forming 'A' compartments. Within these expansive regions, substructures emerge, giving rise to transcription hubs or factories. These can be visualised through staining nascent transcripts, or RNA polymerase II[51–53]. Typically, factories measure 90-nm in diameter, with a broad size distribution[52]. They are enriched in nascent mRNA[54], and previous studies have suggested a potential association with the nuclear matrix[55,56]. Transcription factories arise mechanistically due to bridging-induced attraction (BIA)[57,58], a general mechanism dependent on multivalency in chromatin-protein binding, leading to the formation of microphase separated clusters[59–62], in a process termed bridging-induced phase separation (BIPS). Despite this understanding, there are no existing biophysical models that have thoroughly examined the interplay between BIPS and the phase behaviour of the nuclear microenvironment.

In this study, we aimed to investigate the mechanistic and phase properties of the nuclear environment and their relationship with chromatin folding. Our investigation was built upon the hypothesis that SAF-A may engage with RNA molecules to create a nuclear mesh[2,35] and used stochastic optical reconstruction microscopy (STORM) to study the morphology and biophysical properties of this supramolecular structure[63]. Our analysis revealed that SAF-A and RNA interact to form clusters with a median size of approximately 120 nm, which likely encapsulate transcription factories originating through BIPS. Within the aqueous nuclear environment, the RNA/SAF-A mesh-like structure will undergo gelation[64,65], at least when bridging interactions, which have limited valence, dominate over non-specific interactions mediated by IDRs[66], to create microphase-separated RNA hydrogels, which we term 'microgels'. Microgels were observed to be interconnected

via RNA/SAF-A filaments, creating a percolating, space-spanning heterogeneous network at the nuclear scale. Importantly, this emergent, supramolecular intranuclear structure is both transcription-dependent and tunable, i.e. it can be adjusted in response to nuclear requirements. Depletion of SAF-A or inhibition of transcription resulted in a reduction in microgel size and diminished large-scale network connectivity. In contrast, modulating RNA turnover by controlling the level of the XRN2 enzyme, responsible for RNA degradation, promoted an increase in microgel size and connectivity.

Combining predictions from computer simulations and FISH (fluorescence in-situ hybridisation) microscopy, we found that the RNA/SAF-A nuclear network can influence large-scale chromatin folding. Our simulations suggest that microgels reduce bridging induced attraction of chromatin proteins through steric hindrance, stiffening and effectively swelling active chromatin. Furthermore, our modelling indicated that the degree of decompaction is modulated by transcriptional activity and RNA turnover, a hypothesis corroborated by targeted FISH experiments, designed to validate such predictions.

The interconnected nuclear microgel networks uncovered here may be a key component of what was traditionally understood as the nuclear matrix. Whilst the network exhibits solid-like characteristics at short timescales, it also demonstrates significant heterogeneity and renewability, resulting in liquid-like properties at long time scales. Furthermore, analysis of GFP multimer diffusion within this structure showed that small proteins diffuse freely. Consequently, the supramolecular RNA/SAF-A network exhibits complex viscoelastic behaviours that are time- and scale-dependent. This observation offers a potential reconciliation of the previously contradictory reports of fluid-like[47,50] and solid-like properties[45] within the nuclear interior.

We propose that the microgel network may constitute a self-organised, non-equilibrium biophysical structure capable of influencing nuclear processes involving DNA, such as transcription, and RNA processing, such as splicing or export, shedding light on the mechanisms that connect biophysical structures and pathological conditions.

SAF-A and newly synthesised RNA form clusters

To investigate intranuclear supramolecular structures under physiological conditions (as opposed to harsh extraction conditions used in classical studies), we used super-resolution direct Stochastic Optical Reconstruction Microscopy (dSTORM) in RPE1 cells to study the localisation of the nuclear protein SAF-A and of newly transcribed RNA (Fig 1; see Methods for details on dSTORM imaging).

dSTORM revealed that SAF-A had a punctate and clustered organisation, in agreement with our previous results[63] (Fig 1A). To quantify the imaging data, SAF-A clustering was analysed using the Ripley H function (Fig 1B), a normalised analysis method used to characterize spatial point patterns over a given area of interest, which showed a peak of about 40 nm, corresponding to a cluster diameter of approximately 100 nm (calculated assuming clusters are spherical); similarly, image analysis showed the maximal distance between any two points in a SAF-A cluster was 123$\pm$49 nm. By computing the average cluster density and accounting for the fact that the fraction of the imaging volume over the total nuclear volume, we estimated that an RPE1 nucleus contained 3700$\pm$1500 clusters.

Previously we demonstrated that SAF-A and RNA can interact[35], building upon this finding, the relationship between SAF-A clustering and the spatiotemporal organisation of RNA was explored. As a first step, the kinetics of RNA synthesis was analysed to determine the proportion of RNA retained in the nucleus, and that which is exported to the cytoplasm. To do so, cells were briefly pulsed with [5-^{3}H]uridine (Fig S1A) in the presence of low dose actinomycin D, to inhibit RNA Pol I and supress label incorporation into nucleolar RNA. As expected, levels of newly synthesised RNA increased in the cytoplasm, but after 1 h, 80% of non-ribosomal newly synthesised RNA remained in the nucleus (Fig S1B), in the form of high molecular weight ribonucleic acid species (Fig S1C). To investigate the intranuclear spatial localisation of RNA, RPE1 cells were treated with 5-ethynyluridine (5EU), a modified uracil base that is efficiently labelled using click-chemistry. Extensive 5EU labelling showed no adverse nuclear morphology and after overnight incubation there was clear accumulation of labelled RNA species in the cytoplasm (Fig S1D-E). To discern newly transcribed ribonucleoprotein structures, cells were pulse labelled with 5EU for 15 min, fixed at subsequent times (Fig S1F), then newly transcribed RNA was conjugated to CF647 and analysed by either confocal microscopy to quantify RNA levels (Fig S1G), or using dSTORM to quantify the spatial distribution of RNA (Fig 1C).

Nascent RNA (0 h) formed fine clusters (Fig 1C), with signal density showing a gradual increase over time (Fig 1D-E). Analysis of the Ripley H function for newly transcribed RNA revealed a clustering peak at ~20-30 nm (Fig 1F; 0 h), but at 0.5 h and 1 h there was a loss in clustering (no peak was visible in the Ripley analysis). As SAF-A is an RNA binding protein, we speculated that SAF-A clusters might form at sites of newly transcribed RNA. To test this idea, two-colour storm was used to visualise both RNA and SAF-A, which showed significant colocalization (Fig 1G), followed by RNA diffusion away from these clusters at 0.5 h and 1 h (Fig S1H).

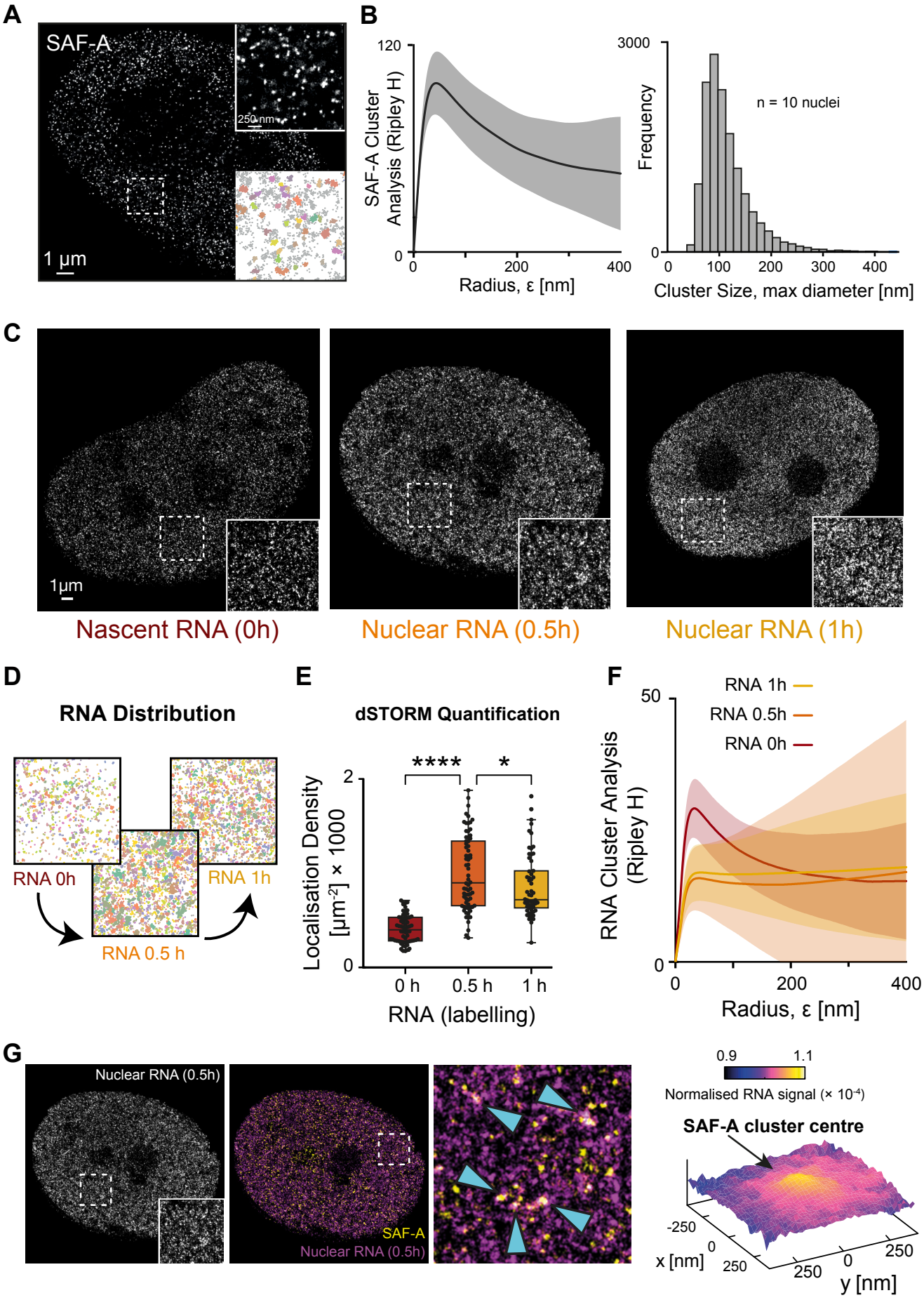

A
SAF-A
250 nm
1 µm
B
SAF-A Cluster Analysis (Ripley H)
120
0
0
200
400
Radius, ε [nm]
3000
Frequency
n = 10 nuclei
0
0
100
200
300
400
Cluster Size, max diameter [nm]
C
1µm
Nascent RNA (0h)
Nuclear RNA (0.5h)
Nuclear RNA (1h)
D
RNA Distribution
RNA 0h
RNA 1h
RNA 0.5 h
E
dSTORM Quantification
Localisation Density [µm⁻²] × 1000
2
0
****
*
0 h
0.5 h
1 h
RNA (labelling)
F
RNA 1h
RNA 0.5h
RNA 0h
RNA Cluster Analysis (Ripley H)
50
0
0
200
400
Radius, ε [nm]
G
Nuclear RNA (0.5h)
SAF-A
Nuclear RNA (0.5h)
0.9
1.1
Normalised RNA signal (× 10⁻⁴)
SAF-A cluster centre
-250
0
250
x [nm]
250
0
250
y [nm]

Figure 1

Figure 1. Organisation of SAF-A and nuclear RNA clusters

A. STORM imaging of SAF-A clusters in fixed RPE1 cells. Top inset, enlarged view of region shown in dashed white square. Bottom inset, X-Y plot of STORM localisations. Grey: not belonging to any cluster. Coloured: belonging to a cluster.

B. Left panel, Strength of SAF-A clusters determined using Ripley H algorithm (n = 10 nuclei). Right panel, Distribution of cluster size (n = 10 nuclei).

C. STORM imaging of nascent RNA and matured nuclear RNA. Cells were pulse labelled with 5EU then fixed at 0, 0.5 and 1 h. RNA was detected by attaching to CF647 dye using click chemistry . Insets are enlarged view of regions marked in dashed white square.

D. X-Y plot of STORM localisations shown in panel C insets, generated from the SuperStructure analysis using a fixed search radius. Different colours represent different detected clusters.

E. Quantification of RNA localisation density (n = 60 regions). P values for T-test (n = 60; *, $p < 0.05$; ****, $p < 0.0001$).

F. Strength of RNA clusters for each time point (0, 0.5, 1 h) determined using Ripley H algorithm (n = 10 nuclei).

G. Left, two colour STORM imaging of SAF-A (yellow, CF568 dye) and RNA (purple, CF647 dye), at 0.5 h after RNA labelling. For images showing RNA (left panels), inset corresponds to enlarged region marked by dashed box. For RNA/SAF-A images (middle panels), dashed box corresponds to enlarged view shown in right panel. Blue arrows correspond to regions of pronounced SAF-A and RNA overlap. Right, meta-analysis of RNA signal strength surrounding SAF-A clusters (n = 10 nuclei).

SAF-A SPRNK domain interacts with RNA

We hypothesized that the formation and biochemical basis of nuclear clusters might be mediated by SAF-A functioning as a bridging factor. Bridging could occur through self-association of its natively unstructured IDR via non-covalent cation-π interactions, or with RNA through RGG motifs[18] located in the protein's C-terminus[25,35], and additional RNA binding sites. Interactions through the IDR would be expected to under macrophase or liquid-liquid phase separation forming protein-rich droplets, as seen for HNRNPA1[20,22], inconsistent with our results (Fig 1A-B). Therefore, to assess the viability of the hypothesis that SAF-A binds to RNA, we analysed the central structured and highly conserved region of SAF-A (Fig 2A, Fig S2A) using homology modelling and AlphaFold. Despite there being two previously annotated domains in the centre of the protein, we revealed a single globular structure (Fig 2B) that encompassed a PNK-like P-type NTPase (Fig 2B), with a weak affinity for ATPγS (K_D = 806 μM) (Fig 2C). N-terminal to this domain is a SPRY domain with a conserved positively charged patch on its surface (Fig 2B), that we speculated might constitute an RNA binding site, as has been shown for the SAF-A homologue, HNRNPUL1[67]. Purified SPRY-PNK bound to homomeric oligomers of A, C or U, indicating that it binds RNA with little sequence specificity (Fig S2B-C). Interestingly, RNA binding was enhanced in the presence of ATP, but not ADP (Fig 2D), suggesting that RNA bridging might be dependent on local nucleotide concentration. Based on our structural analysis it appears that the two protein structures constitute a single entity that we term a SPRNK domain. To further test the RNA binding activity of the SPRNK domain four mutations were predicted based on surface electrostatics and conservation analysis of the AlphaFold model that might constitute an RNA binding surface. Mutation of these sites from positively charged to negatively charged amino acids (Fig 2E) significantly abrogated RNA binding (Fig 2F).

Thus, SAF-A has two binding sites, one located in the RGG domain[18,19,25,35] and one in the SPRNK domain, suggesting it could indeed function as an RNA bridging factor (Fig 2G). Concomitantly, transmission electron microscopy of immunoprecipitated cross-linked FLAG-SAF-A revealed filament shaped oligomers, that in the presence of RNaseA/T1 were disassembled (Fig S2D-E). Previous CLIP-seq studies have indicated that SAF-A binds extensively to exons and introns at levels consistent with the respective amounts of these RNA species[26,68–70]. However, these experiments did not differentiate between old and newly synthesised RNA, so to determine whether SAF-A also bound to newly synthesised RNA we used PAR-CLIP (photoactivatable ribonucleoside-enhanced crosslinking and immunoprecipitation). In PAR-CLIP newly synthesised RNA is labelled with 4sU making it a substrate for photo-crosslinking to bound proteins, which are then enriched by immuno-precipitation and analysed by sequencing. SAF-A was found to bind to all newly synthesised RNA classes, but frequently found bound to intronic RNAs (63%) (Fig S2F), consistent with other recent data[7].

As SAF-A and RNA are known to form a mesh-like structure in vitro[35] (Fig S2D-E), it is plausible to consider these clusters (Fig 1) as hydrated mesh-like structures, or hydrogels within the cell nucleus[64]. Additionally, our data suggests that these structures correspond to microphase separation, as evidenced by the self-limiting cluster size and the absence of coarsening in steady state[71]. Henceforth, we shall refer to these microphase-separated RNA hydrogels as microgels in the following text (Fig 2G).[53]

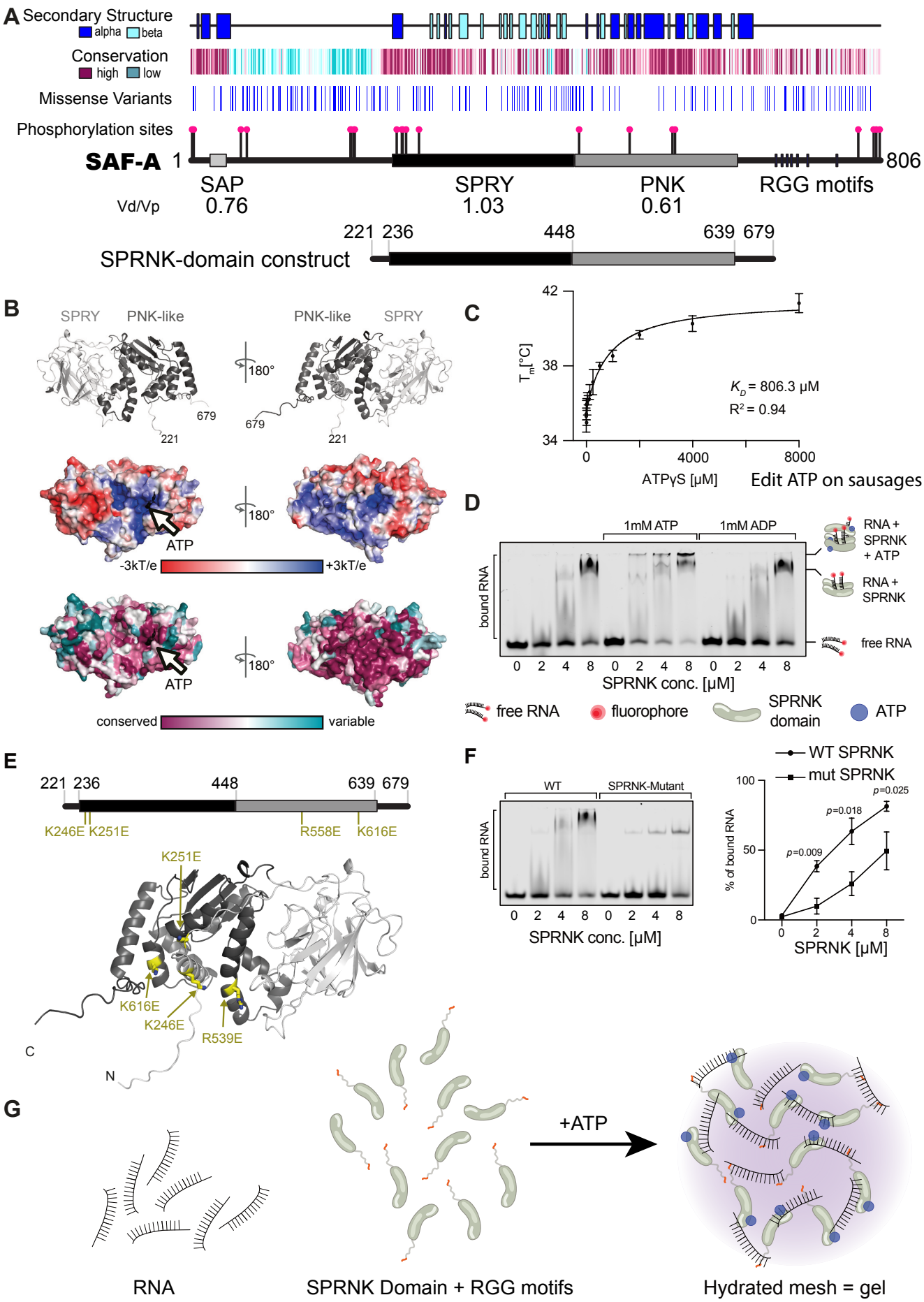

A
Secondary Structure
alpha
beta
Conservation
high
low
Missense Variants
Phosphorylation sites
SAF-A
1
806
SAP
SPRY
PNK
RGG motifs
Vd/Vp
0.76
1.03
0.61
221 236
448
639 679
SPRNK-domain construct
B
SPRY
PNK-like
PNK-like
SPRY
180°
679
221
ATP
-3kT/e
+3kT/e
conserved
variable
C
T_m [°C]
K_D = 806.3 µM
R^2 = 0.94
ATPγS [µM]
Edit ATP on sausages
D
1mM ATP
1mM ADP
bound RNA
RNA + SPRNK + ATP
RNA + SPRNK
free RNA
SPRNK conc. [µM]
free RNA
fluorophore
SPRNK domain
ATP
E
K246E K251E
R558E
K616E
K251E
K616E
K246E
R539E
C
N
F
WT
SPRNK-Mutant
bound RNA
SPRNK conc. [µM]
WT SPRNK
mut SPRNK
% of bound RNA
p=0.025
p=0.018
p=0.009
SPRNK [µM]
G
+ATP
RNA
SPRNK Domain + RGG motifs
Hydrated mesh = gel

Figure 2

Figure 2. SPRNK domain structure and RNA binding characteristics

A. Cartoon of the SAF-A protein depicting domains and their secondary protein structure (alpha helix, dark blue; beta strand, light blue). Lollipops indicate sites of post-translational modifications. Protein conservation (calculated by ConSurf) is shown from low (turquoise) to high (purple); positions of benign missense variants (from gnomAD) are dark blue lines. Vd/Vp ratios of missense variant depletion per domain compared to depletion over the full-length protein are shown below domains. The SPRNK domain construct (SPRY/PNK) used in binding experiments is shown.

B. AlphaFold2 models of SAF-A SPRNK domain showing secondary structure, electrostatic charge (APBS) and protein conservation (CONSURF).

C. Graph showing melting temperature of SPRNK domain in the presence of different concentrations of ATPγS.

D. Electrophoretic mobility shift assay (EMSA) with fluorescently labelled RNA and purified SPRNK domain in the presence or absence of nucleotides. Right, cartoons depicting the potential interactions between SAF-A, RNA and ATP.

E. Top, cartoon showing the SAF-A SPRNK domain with putative residues marked for RNA binding. Bottom, putative RNA binding residues superimposed on the secondary structure of SPRNK domain.

F. Left, EMSA with fluorescent RNA template with wild-type (WT) or putative SPRNK RNA-binding mutant. Right, quantification of EMSA. P-values are for a T-test, n=3.

G. Cartoon depicting how the SAF-A SPRNK domain and RGG motifs (red) can bridge between RNA molecules. In our model, RNA-mediated bridging of SAF-A, at structures such as transcription factories, results in the formation of microphase-separated RNA hydrogels, or “microgels”.

SAF-A acts as an RNA bridging factor to form a nuclear-spanning network of microgels

To understand how RNA/SAF-A microgels are positioned within the nuclear environment, we asked whether they are isolated from each other or interconnected. To do so, inter-connectivity between clusters (microgels) was measured using the SuperStructure algorithm[63].

In this algorithm, the decay rate of the number of detected clusters (using DBSCAN) was analysed while increasing the searching radius with cluster connectivity estimated as the slope of the exponential decay in the plot showing the number of clusters per localisation versus cluster radius (Fig S3A). As observed previously[63], SAF-A clusters showed moderate interconnectivity with a decay rate of 0.07 nm^{-1}, corresponding to a connection probability of 0.05 between clusters (Fig S3B-C). In contrast RNA clusters (Fig S3C) showed a lower inter-connectivity with a decay rate of 0.03 nm^{-1} for newly synthesised RNA, but a larger connectivity (steeper decay rate) over a period of 0.5 h and 1h (recall for these times RNA no longer forms defined clusters (Fig 1F)), similar to what is observed for SAF-A interconnectivity (Fig S1C). Together, these data can be interpreted by hypothesising that SAF-A acts as an RNA bridging protein which forms a nuclear-spanning, sparsely connected, network of microgels (Fig 3A). Newly synthesised RNA colocalises with the microgels, suggesting that these are associated with transcription, and may be transcription factories (see Discussion).

We reasoned that, if this hypothesis was correct, SAF-A depletion or transcription inhibition would affect the microscopic properties of the supramolecular network they form, both at the local and global level (respectively microgel properties and network interconnectivity) (Fig 3A). In agreement with this prediction, depletion of SAF-A reduced the density of newly synthesised RNA, indicative of increased dispersion, and reduced cluster connectivity (Fig 3B-C, Fig S3D), consistent with disrupting the RNA/SAF-A microgels.

Similarly, transcription inhibition using either triptolide or α-amanitin (data not shown), with concomitant rapid degradation of nuclear RNA levels (Fig S3E-F), caused a significant reduction in SAF-A cluster size, cluster density and connectivity (Fig 3D-E, Fig S3G-I). Fluorescence correlation spectroscopy (FCS; Fig S3J) detecting SNAP-SAF-A showed that transcription inhibition caused a concomitant increase in SAF-A mobility (Fig 3F, Fig S3K), supporting the notion that nuclear RNA and SAF-A dependently form and stabilise a RNA/SAF-A gel-like structure, inside which protein mobility might be reduced. To test this idea, we used FCS to examine the mobility of GFP oligomers, inducibly expressed from the ROSA26 safe harbour locus (Fig 3G; Fig S3L). As expected[72], a large GFP oligomer (n=5) was less mobile than a GFP monomer [7 vs 22 $\mu m^2/s$, respectively], but crucially the mobility of all GFP proteins increased when SAF-A was depleted (Fig 3G-H). This indicates that SAF-A is a key component of an interconnected nuclear network forming a mesh-like supramolecular structure. These results also allowed us to discriminate between different modes of SAF-A binding to RNA, as loss of RNA showed a reduction in cluster size consistent with it being a bridging factor in contrast to a self-associating RNA binder, such as FUS, which would have formed large protein aggregates in the absence of RNA[73]. Taken together, these observations show that the interconnected network of microphase separated RNA hydrogels formed by SAF-A and nascent transcripts is responsive in terms of its emergent morphological and biophysical properties.

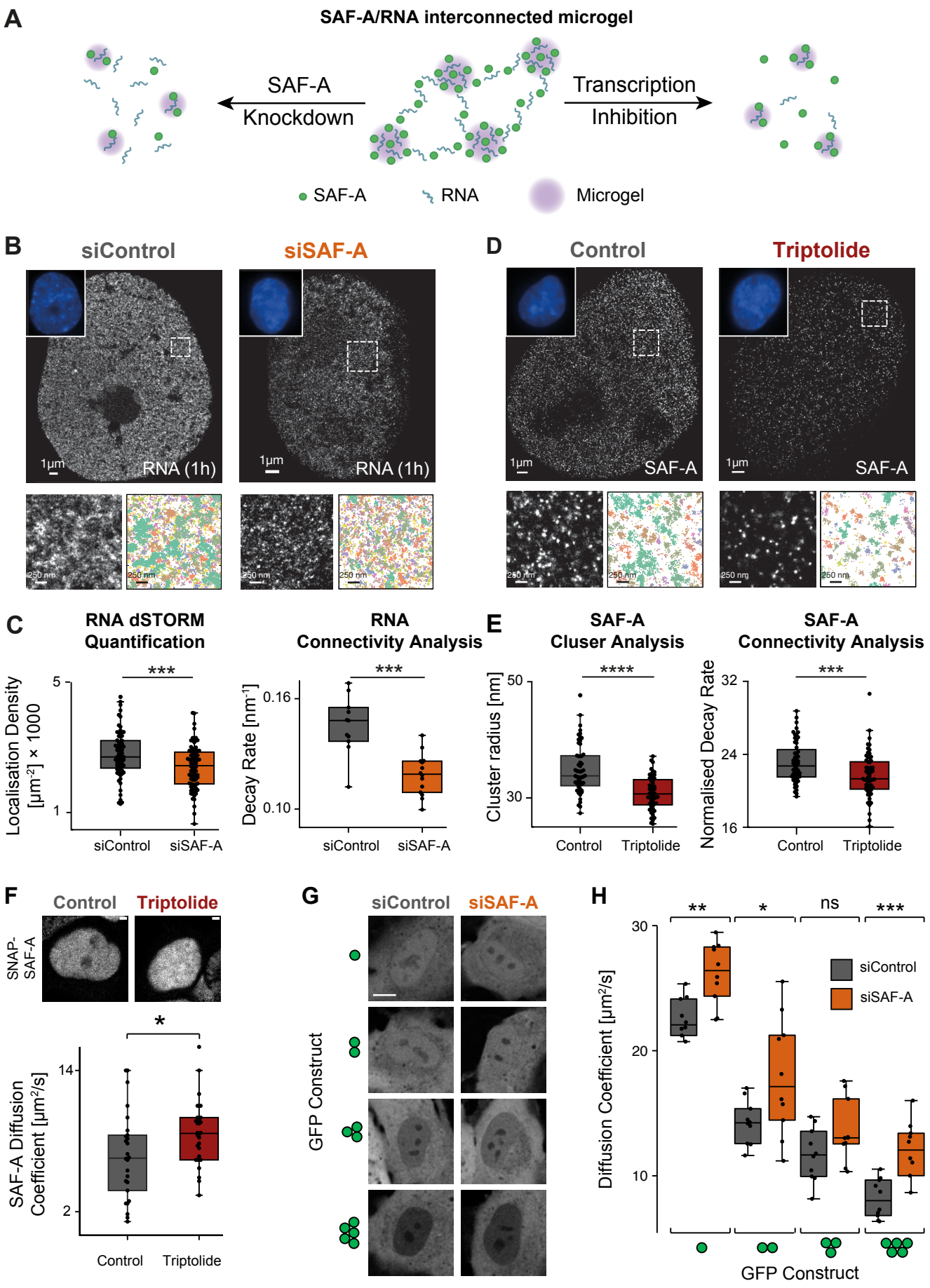

A
SAF-A/RNA interconnected microgel
SAF-A
Knockdown
Transcription
Inhibition
SAF-A
RNA
Microgel
B
siControl
siSAF-A
1µm
RNA (1h)
250 nm
D
Control
Triptolide
SAF-A
C
RNA dSTORM
Quantification
Localisation Density
[µm⁻²] × 1000
***
siControl
siSAF-A
RNA
Connectivity Analysis
Decay Rate [nm⁻¹]
E
SAF-A
Cluser Analysis
Cluster radius [nm]
****
Control
Triptolide
SAF-A
Connectivity Analysis
Normalised Decay Rate
F
SNAP-SAF-A
SAF-A Diffusion
Coefficient [µm²/s]
*
G
GFP Construct
H
Diffusion Coefficient [µm²/s]
**
ns
siControl
siSAF-A
GFP Construct

Figure 3

Figure 3. Interconnection and properties of RNA/SAF-A microgels.
A. Cartoon, hypothesising how a RNA/SAF-A interconnected microgel (middle) is affected by both SAF-A depletion (left) or transcription inhibition (right).
B. STORM of newly synthesised RNA labelled with CF647 dye in control (left panel) and after SAF-A depletion (right panel). DAPI staining is shown in the top left corner of each image. Lower left panel is enlarged view of regions marked in dashed white square, whilst lower right panel is X-Y plot of STORM localisations shown in left panel, generated from SuperStructure analysis for a fixed search radius. Colour represents detected clusters.
C. Left, boxplot showing RNA localisation density for control and SAF-A depleted cells (n > 75). Right, boxplot showing RNA connectivity calculated using the SuperStructure algorithm in control and SAF-A depleted cells (n=12 and 14, respectively).
D. STORM of SAF-A labelled with CF647 dye in control cells (left panel) and after transcription inhibition with Triptolide (right panel). DAPI staining is shown in the top left corner of each image. Lower left panel is enlarged view of regions marked in dashed white square, whilst lower right panel is X-Y plot of STORM localisations shown in left panel, generated from SuperStructure analysis for a fixed search radius. Colour represents detected clusters.
E. Left, boxplot showing SAF-A cluster size for control and triptolide treated cells (n = 58 regions). Right, boxplot showing SAF-A connectivity calculated using the SuperStructure algorithm in control and triptolide treated cells (n = 58 regions).
F. Top, Live cell imaging of control or triptolide treated cells expressing SNAP-SAF-A, visualised using TMR-STAR, for fluorescence correlation spectroscopy (FCS). Scale bar is 1 µm. Bottom, boxplot show SAF-A diffusion coefficient in cells above, determined using FCS (Control, n = 26 regions; Triptolide, n = 30 regions).
G. Live cell imaging of RPE1 cells expressing mono, di, tri and penta-GFP, treated with either siControl or siSAF-A. Scale bar is 5 µm.
H. Boxplot of GFP diffusion coefficients, measured by FCS, in control or SAF-A depleted cells expressing different GFP constructs, as described in G. For each condition, n = 10 regions.

P values are for a T-test; ns, not significant; *, < 0.05; **; p < 0.01; ***, p < 0.001; ****, p < 0.0001.

XRN2 degrades newly synthesised nuclear RNA

The nuclear distribution of newly synthesised RNA is rapidly remodelled (Fig 1C). We reasoned that proteins important for RNA turnover were likely to be spatially located in the vicinity of the nucleoplasmic network and microgels, so BioID[74] (proximity-dependent biotin identification) was used to identify proteins in the vicinity of the SAF-A, by tagging the protein on the N-terminus with BirA. As expected, SAF-A was observed to be in close proximity to its homologues HNRNPUL1 and HNRNPUL2 and other members of a nuclear scaffold: SAFB, SAFB2 and MATRIN3 (Fig S4A; Table S1), but, surprisingly, 5'-3' Exoribonuclease 2 (XRN2) was also enriched.

XRN2 is an exonuclease protein predominantly observed within the nucleus[75]. It is responsible for the degradation of RNA[8,76,77] and in the torpedo model of transcription termination[78,79]. As XRN2 interacted with SAF-A in an RNA-dependent manner (Fig 4A), we hypothesised that it may be responsible for degrading newly synthesised nuclear RNA within the SAF-A microgels (Fig 4B). To test this idea, the turnover of nuclear RNA was examined in control cells and in the absence of XRN2, XRN1 or an exosome subunit, EXOSC3. Cells were pulsed (15 min) with 5EU (Fig 4C), fixed, and newly synthesised RNA was analysed by attaching a fluorophore to 5EU by click chemistry. RNA accumulated over a 1 h period in the nucleus before being degraded (Fig 4C, Fig S4B), with a mean lifetime (τ) of 43 min. Depletion (Fig S4C-D) of either XRN1 or EXOSC3 had little effect on the persistence of newly synthesised RNA (Fig 4C, Fig S4E), but depletion of XRN2 with RNAi abrogated turnover of newly transcribed RNA and increased τ to 140 min, with additional RNA spreading to the cytoplasm at long time points (Fig 4C, Fig S4B). To confirm this result, cells with a degron-tagged version of XRN2[79], a gift from Steve West, were treated with auxin to trigger protein degradation (Fig S4F-G); results similarly showed that XRN2 depletion leads to an increase in nuclear RNA lifetime (Fig S4H). These results are consistent with the idea that XRN2 colocalises with SAF-A within the microgels, to degrade newly transcribed RNA and hence regulate the concentration of nascent transcripts (Fig 4B).

As XRN2 can degrade RNA, we reasoned that it may also act to depolymerise nucleoplasmic RNA/SAF-A microgels, such that XRN2 concentration may be effectively used to modulate newly transcribed RNA levels (Fig 4D), and hence tune the biophysical properties of the microgels. To test whether XRN2 had the capacity to alter SAF-A oligomerisation, newly synthesised RNA was labelled with 5EU, transcription was inhibited, and RNA was fixed in cells at different timepoints (Fig 4E). As expected, in siControl cells, newly synthesised RNA is rapidly degraded (Fig 4F, green outlined box; Fig 4D, model 1), whilst in siXRN2 cells RNA remains intact (Fig 4F, blue outlined box; Fig 4D, model 2). To test how this might affect RNA/SAF-A oligomerisation, SAF-A was crosslinked with BM-PEG2 and samples analysed on a denaturing protein gel (Fig 4G). Control cells (lane 1) showed high molecular weight SAF-A oligomers, whilst RNA degradation by XRN2 (lane 2) significantly reduced SAF-A oligomerisation. XRN2 knockdown and abrogation of RNA degradation instead enhanced SAF-A oligomerisation (lane 3, 4 in Fig 4G). Together this indicates that XRN2 can modulate RNA levels and concomitantly alter SAF-A oligomerisation (Fig 4H).

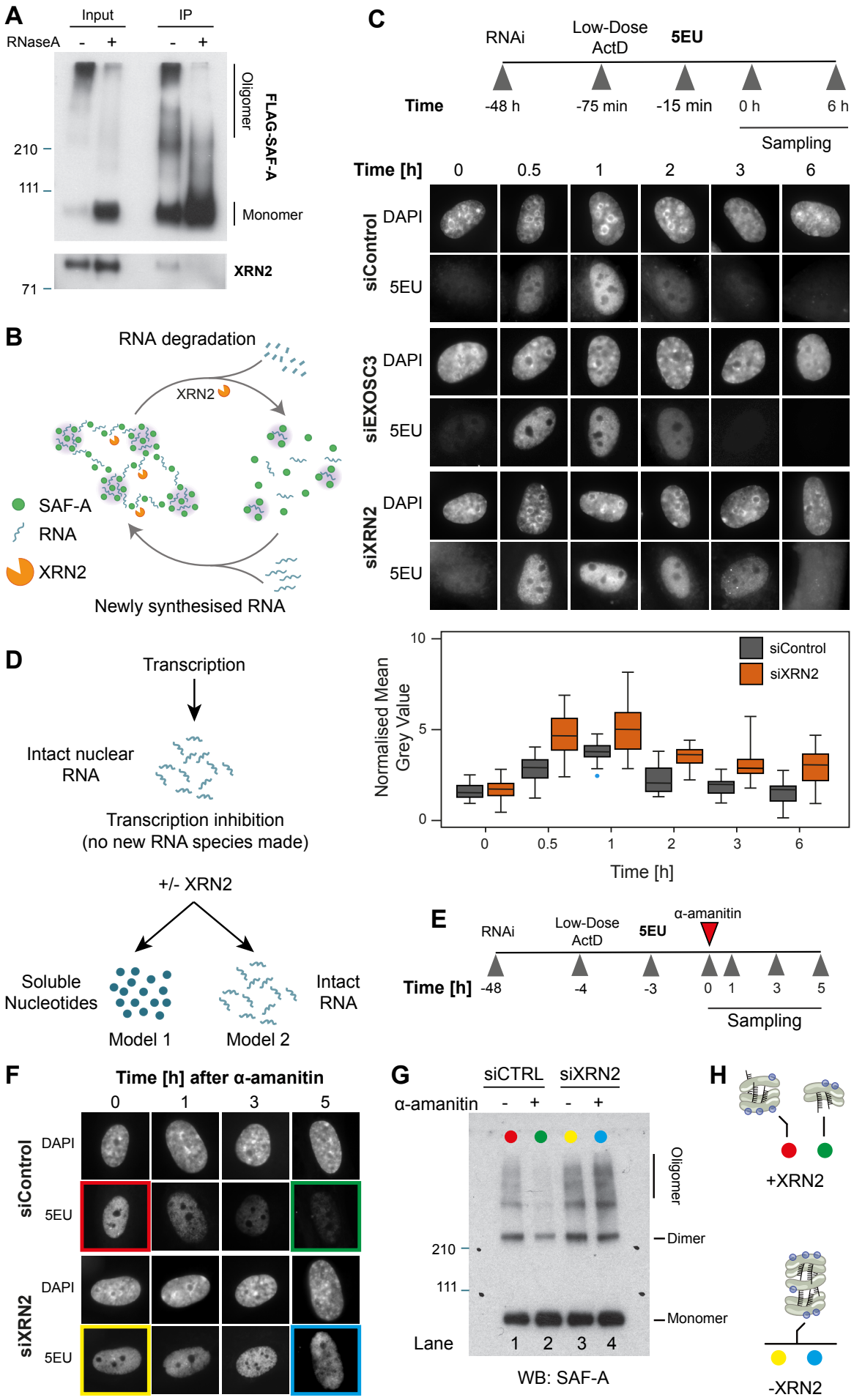

A
Input
IP
RNaseA
- +
- +
210
111
71
Oligomer
FLAG-SAF-A
Monomer
XRN2
B
RNA degradation
XRN2
SAF-A
RNA
XRN2
Newly synthesised RNA
C
RNAi
Low-Dose ActD
5EU
Time
-48 h
-75 min
-15 min
0 h
6 h
Sampling
Time [h]
0
0.5
1
2
3
6
siControl
siEXOSC3
siXRN2
DAPI
5EU
Normalised Mean Grey Value
siControl
siXRN2
Time [h]
D
Transcription
Intact nuclear RNA
Transcription inhibition
(no new RNA species made)
+/- XRN2
Soluble Nucleotides
Intact RNA
Model 1
Model 2
E
RNAi
Low-Dose ActD
5EU
α-amanitin
Time [h]
-48
-4
-3
0
1
3
5
Sampling
F
Time [h] after α-amanitin
0
1
3
5
siControl
siXRN2
DAPI
5EU
G
siCTRL
siXRN2
α-amanitin
- + - +
Oligomer
Dimer
Monomer
210
111
Lane
1 2 3 4
WB: SAF-A
H
+XRN2
-XRN2

Figure 4

Figure 4. XRN2-dependent turnover of RNA/SAF-A microgels

A. Western blots showing immunoprecipitation of oligomerised FLAG-SAF-A with XRN2.

B. Cartoon, depicting possible turnover of the RNA/SAF-A microgels.

C. Top, Diagram explaining how experiment was performed. Middle, fluorescent imaging of 5EU in newly synthesised RNA labelled by click chemistry with an azide-Alexa488 dye, in control cells or cells depleted for EXOSC3 (exosome) or XRN2. Bottom, quantification of 5EU in control or XRN2 depleted cells over time (n = 10 nuclei).

D. Diagram, discriminating between different roles of XRN2 in either turning over nuclear RNA to form soluble nucleotides (model 1), or having little effect on nuclear RNA (model 2).

E. Diagram explaining experiment to test for XRN2 function, as outlined in D.

F. Immunofluorescence to detect 5EU labelled RNA, conjugated to Alexa488 by Click chemistry, in control or XRN2 depleted cells, treated with α-amanitin to inhibit new RNA synthesis. Coloured boxes correspond to samples subsequently analysed in G.

G. Western blot for oligomeric SAF-A. Samples from Fig 4E-F were cross-linked and fractionated by SDS-PAGE. Coloured dots correspond to conditions shown in Fig 4F.

H. Interpretation of results. Bottom, in the absence of XRN2 SAF-A forms extensive oligomers independent of α-amanitin. Top, in the presence of XRN2 SAF-A oligomer is less extensive, and dependent on continuous RNA transcription.

XRN2 tunes size and connectivity of RNA/SAF-A microgels

We hypothesised that if XRN2 can degrade RNA then its depletion might be expected to alter microgel morphology (Fig 5A). To test this possibility, SAF-A was imaged using dSTORM in cells that were depleted of XRN2 in the presence or absence of the transcription inhibitor, triptolide (Fig 5B-C; S5A-B). SAF-A cluster size and connectivity increase when XRN2 is knocked down in both control and transcription inhibition conditions, which suggests that a non-equilibrium process exists between RNA production (polymerisation), and degradation (depolymerisation), that maintains the nuclear gel in a structural steady state (Fig 5A).

We reasoned that such a non-equilibrium process would endow the gel with a transient and dynamic nature. In order to test this hypothesis, FRAP (fluorescence recovery after photobleaching) was undertaken on SNAP-tagged SAF-A, labelled with TMR-Star in live cells (Fig 5D). To obtain the relevant recovery timescale without any *a priori* assumption, the FRAP curves were fitted with a double exponential function, which enabled the extraction of a fast and a slow population recovery time (Fig 5D, Fig S5C), as well as the protein immobile fraction (Fig S5D), which can be interpreted as the bound population. Control SAF-A displays a small immobile fraction (~15%) and mobile populations have a recovery time of ~2s and ~10s for the fast and slow populations, respectively. When transcription was inhibited (with concomitant RNA loss, Fig S3E-F and Fig 4F), both immobile fraction and fast population recovery times decreased (Fig 5D, Fig S5) suggesting that SAF-A is lacking a structural partner with which to form the microgel, so its overall diffusivity increased. Consequently, we speculated that loss of XRN2 may create a more expansive microgel and impact protein mobility. This, is partially observed in the increase of the FRAP recovery time and immobile fraction after loss of XRN2 when transcription is inhibited (Fig 5D, Fig S5D-E). To further test this idea, GFP oligomers were used as a probe for nuclear viscosity (Fig S3L-M). Loss of XRN2 significantly reduced protein mobility (Fig 5E) indicating that the nuclear environment is regulated by a non-equilibrium process involving production and degradation of RNA which effectively polymerises and depolymerises to form a nuclear interconnected microgel network and impact molecular mobility on the nuclear scale.

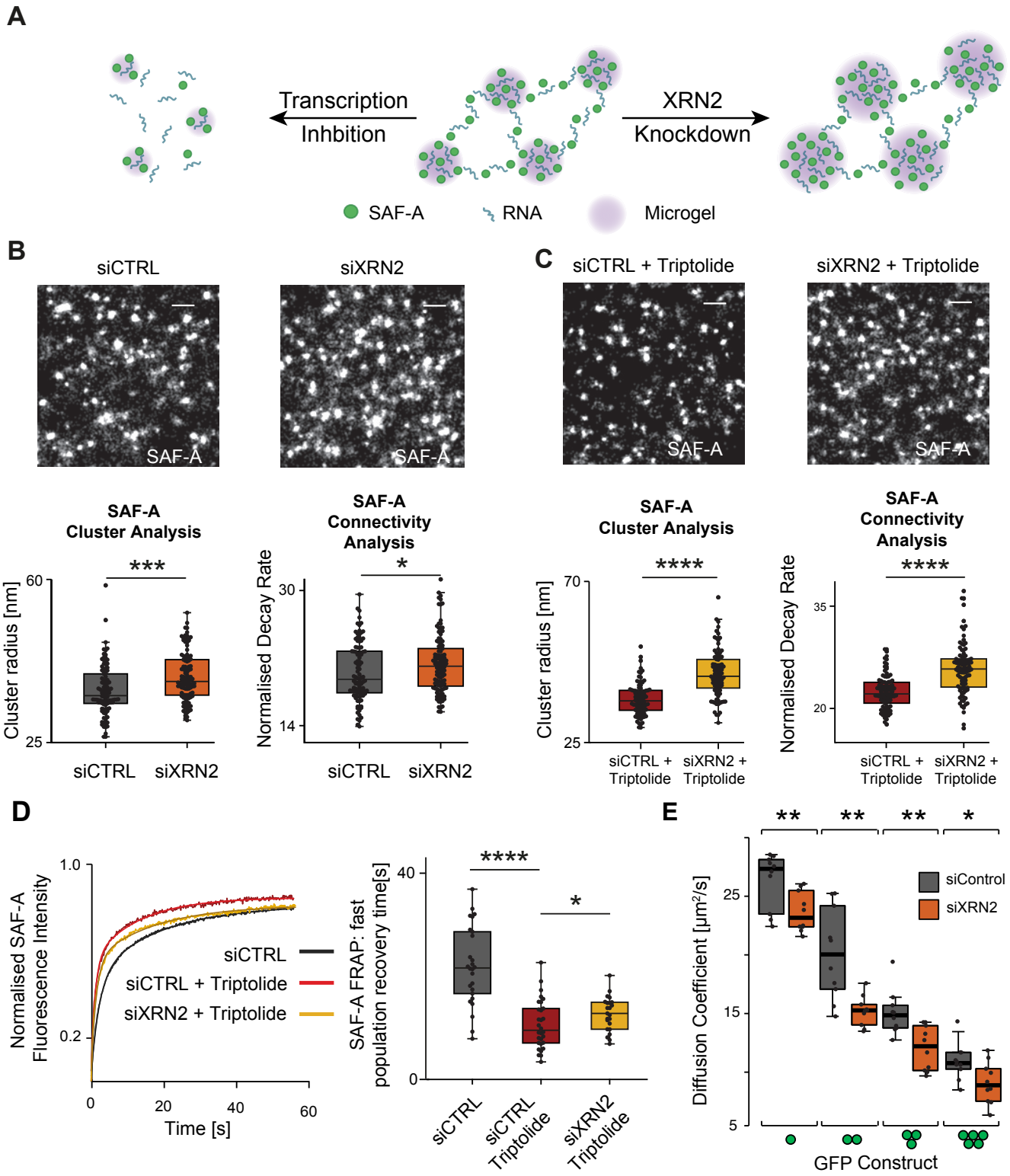

A
Transcription
Inhbition
XRN2
Knockdown
SAF-A
RNA
Microgel
B
siCTRL
siXRN2
SAF-A
SAF-A
Cluster Analysis
Cluster radius [nm]
60
25
***
siCTRL
siXRN2
SAF-A
Connectivity
Analysis
Normalised Decay Rate
30
14
*
siCTRL
siXRN2
C
siCTRL + Triptolide
siXRN2 + Triptolide
SAF-A
SAF-A
SAF-A
Cluster Analysis
Cluster radius [nm]
70
25
****
siCTRL +
Triptolide
siXRN2 +
Triptolide
SAF-A
Connectivity
Analysis
Normalised Decay Rate
35
20
****
siCTRL +
Triptolide
siXRN2 +
Triptolide
D
Normalised SAF-A
Fluorescence Intensity
1.0
0.2
0
20
40
60
Time [s]
siCTRL
siCTRL + Triptolide
siXRN2 + Triptolide
SAF-A FRAP: fast
population recovery time[s]
40
0
****
*
siCTRL
siCTRL
Triptolide
siXRN2
Triptolide
E
**
**
**
*
Diffusion Coefficient [µm²/s]
25
15
5
siControl
siXRN2
GFP Construct

Figure 5

Figure 5. Expansion of RNA microgels in absence of XRN2

A. Cartoon, outlining how a RNA/SAF-A interconnected microgel (middle) is influenced by transcription inhibition (left) or XRN2 depletion (right).

B. Top, STORM images of SAF-A in control and XRN2 depleted cells. Scale bar is 250 nm. Bottom, SAF-A cluster radius and normalised decay rate (connectivity) in control and after XRN2 depletion (n > 100 regions).

C. Top, SAF-A STORM in triptolide treated cells with or without XRN2 depletion. Scale bar is 250 nm. Bottom, SAF-A cluster radius and normalised decay rate (connectivity) in transcription inhibited control and XRN2 depleted cells (n > 100 regions).

D. Left, representative SAF-A FRAP curves and right, SAF-A fast population recovery times for control, triptolide treated or triptolide treated and XRN2 depleted cells (n > 20 regions).

E. Diffusion coefficient of different GFP constructs as measured by FCS in control or XRN2 depleted cells (n = 10 regions).

P values are for a T-test; *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$; ****, $p < 0.0001$.

Molecular dynamics modelling of chromatin folding near a RNA/SAF-A microgel

SAF-A and RNA contribute to regulating large-scale chromatin structure[7,34,35,63]. We therefore hypothesised that the nuclear microgel network we uncovered may interfere with chromatin folding and mediate chromatin decompaction as observed by microscopy[35], where SAF-A depletion leads to the global decompaction of gene-active human chromosome 19 (HSA19) and at specific gene loci. To quantitatively investigate the non-equilibrium dynamical interplay between transcription-induced RNA/SAF-A gelation and large-scale chromatin structure, we developed a simple coarse-grained polymer model where chromatin folds in the presence of an RNA-dependent SAF-A microgel.

In line with previous models[80,81], chromatin was described as a self-avoiding and semi-flexible polymer, coarse grained into 3 kbp beads. We modelled a 1 Mbp fibre (Fig 6A, Fig S6A), partitioned into an active region (euchromatin, 100 kbp in size) flanked by two larger inactive regions (heterochromatin, 450 kbp each), with moderate self-interactions between chromatin of the same type, and strong interaction between promoter beads, which we imagine to all be mediated by implicit chromatin-bridging proteins, such as transcription factors[81,82]. RNA was modelled as a density field with local concentration ρ, while SAF-A as spherical particles with attractive patches on its surface, which can self-assemble into a macroscopic polymeric gel. As SAF-A polymerisation is RNA-dependent (Fig 2, Fig 4G), attractive patches on SAF-A proteins were stochastically switched on at a rate $k_{on}(\rho)$, which increased (linearly) with the local RNA concentration. Patches switched off (became non-binding) at a constant rate k_{off}, reflective of bound nucleotide hydrolisation. These switching reactions were set to drive the system away from thermodynamic equilibrium.

The internal euchromatin region modelled an active gene locus. Every bead was decorated with a patchy particle describing the newly synthesised RNA, which serves as anchor for SAF-A to the chromatin. Furthermore, every chromatin bead within the locus produced RNA at rate k_{prod}. Newly synthesised RNA then diffused with coefficient D (set to match previous experiments[35,83]) and was degraded (by XRN2 and other enzymes) at a constant rate k_{deg} (Fig 6B). Due to the simultaneous presence of synthesis and degradation, the RNA density at a 3D distance $\vec{r}$ from a transcriptional source rapidly reaches a steady state given by $\rho_s(\vec{r}) = A \cdot \frac{e^{-\lambda|\vec{r}|}}{|\vec{r}|}$ where $A = \frac{k_{prod}}{4\pi D}$ measures transcriptional activity (i.e., the amount of RNA produced by the source), and $\lambda = \sqrt{\frac{D}{k_{deg}}}$ is the RNA decay length, which measures the distance that the RNA is able to explore by diffusion before degradation, on average. This steady state form was assumed to describe the RNA density field around each active bead. By varying the transcriptional activity of the gene (A) and the RNA degradation length-scale (λ), different biologically relevant regimes could be explored, corresponding to genes with different rates of transcriptional activity, and to wild-type or XNR2 knock-out phenotypes (for more details on the model including the equations for the force fields used, see Methods).

Simulations showed that RNA-dependent SAF-A gelation strongly influences chromatin folding, driving its decompaction at large scale. When the internal euchromatin was transcriptionally inactive ($A = 0$), SAF-A did not polymerise or gel, and heterochromatin collapsed into a globular domain with euchromatin on its surface (Fig S6B, left; Video 1). Instead, when the gene locus was active ($A > 0$), two important differences arose. First, there was a higher local concentration of oligomerising SAF-A in the vicinity of the whole gene (Fig 6C, Fig S6B, right; Video 2), similar to the SAF-A clusters observed by dSTORM (Fig 1). Second, the microgel which formed from oligomerising SAF-A particles effectively stiffened the active chromatin region, leading to a larger size of the active locus (Fig 6C, right). We argue that this effective stiffening is a consequence of the formation and swelling of a RNA/SAF-A microgel.

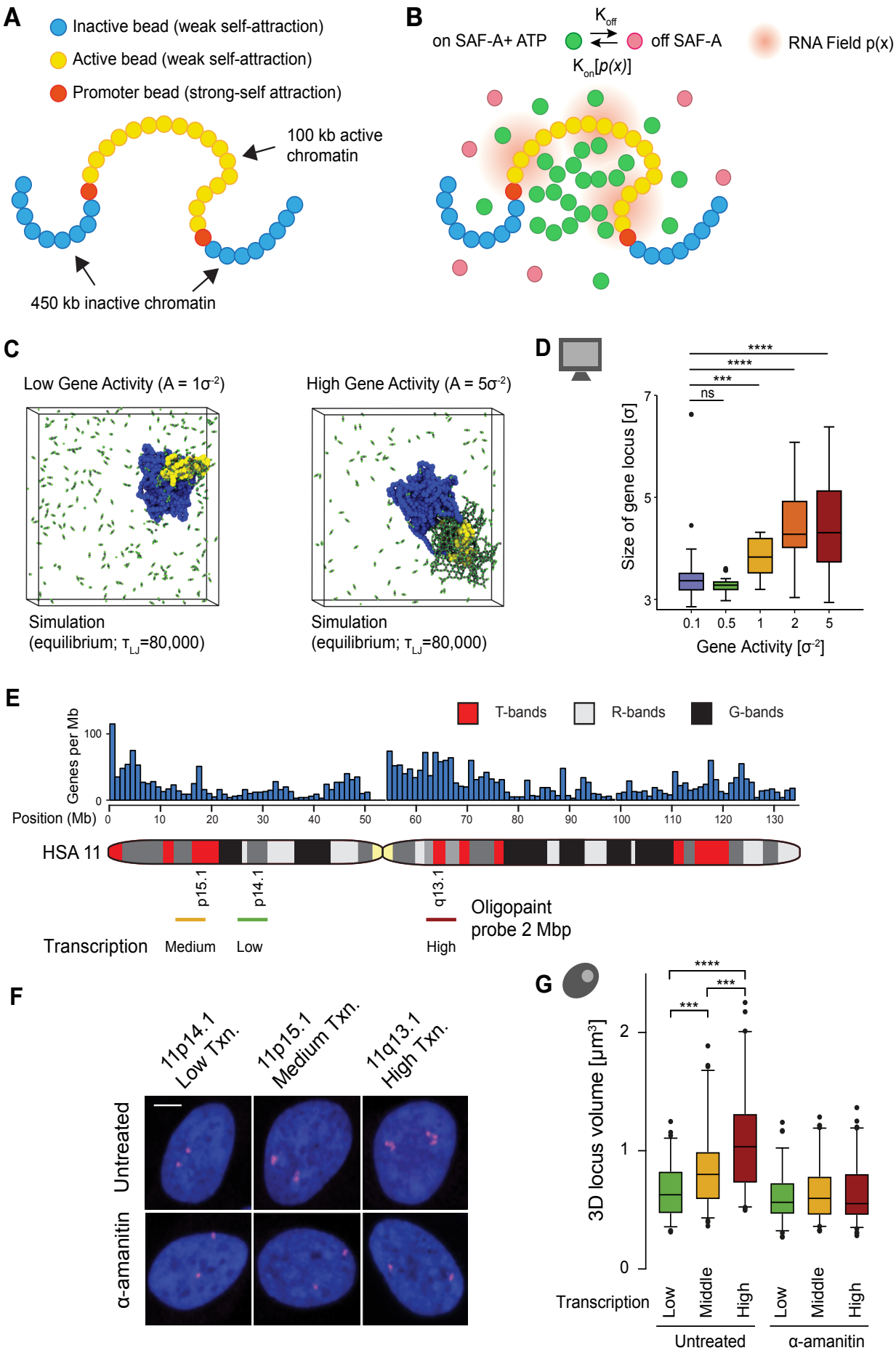

A
Inactive bead (weak self-attraction)
Active bead (weak self-attraction)
Promoter bead (strong-self attraction)
100 kb active chromatin
450 kb inactive chromatin
B
on SAF-A+ ATP
K_{off}
$K_{on}[p(x)]$
off SAF-A
RNA Field p(x)
C
Low Gene Activity (A = 1σ⁻²)
High Gene Activity (A = 5σ⁻²)
Simulation (equilibrium; τ_{LJ}=80,000)
Simulation (equilibrium; τ_{LJ}=80,000)
D
Size of gene locus [σ]
Gene Activity [σ⁻²]
ns
***
****
****
E
T-bands
R-bands
G-bands
Genes per Mb
Position (Mb)
HSA 11
p15.1
p14.1
q13.1
Oligopaint probe 2 Mbp
Transcription
Medium
Low
High
F
11p14.1 Low Txn.
11p15.1 Medium Txn.
11q13.1 High Txn.
Untreated
α-amanitin
G
3D locus volume [μm³]
Transcription
Low
Middle
High
Untreated
α-amanitin

Figure 6

Figure 6. Prediction and testing of polymer model describing how RNA/SAF-A microgels decompact chromatin

A. Cartoon (not to scale), of bead-spring polymer depicting active chromatin (yellow), flanked by promoter beads (red) and inactive chromatin (blue). In the simulation each bead is 3 kb in size.

B. Cartoon of chromatin shown in A, but now decorated with “off” SAF-A (red) and “on”, RNA binding competent, SAF-A (green). SAF-A RNA binding is proportional to the RNA field strength emanating from the active chromatin (yellow).

C. Left, snap-shot of chromatin polymer simulation showing compaction of inactive chromatin (blue) and weakly transcriptionally active chromatin ($A = 1\sigma^{-2}$, yellow), contrasted with, right, highly transcriptionally active chromatin ($A = 5\sigma^{-2}$, yellow). Active SAF-A is depicted as a grey bead, with two “sticky” green patches, the RNA field is not shown for clarity.

D. Boxplot showing the relationship between simulated gene activity and the size of the active gene locus (n = 25 simulations for each condition). P-values are for a T-test.

E. Cartoon showing gene density and chromosome bands across HSA11. Oligopaint probes spanning 2 Mb regions are marked below the chromosome.

F. Representative immunofluorescence images of control and α-amanitin treated cells hybridised to oligopaint probes for low, medium and highly transcriptionally active regions. Scale bar is 5 µm.

G. Quantification of the 3D locus volume measured from nuclei hybridised to oligopaint probes at differently transcriptionally active pain probes (n > 65). Boxplot colours correspond to the transcriptional activity of the probe regions, marked in E. P values are for a Wilcoxon test.

P values correspond to; ns, not significant; *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$; ****, $p < 0.0001$.

Microgel formation regulates large-scale chromatin structure

To characterise microgel-induced chromatin decompaction, the relationship between transcriptional activity (A) and active locus size (σ) was explored (Fig S6C). The size of the active locus was proportional to the transcriptional activity, with the behaviour suggestive of a transition or sharp crossover between a compacted and decompacted locus (Fig 6D, Fig S6B-C, Video1-2), and as expected gene activity was proportional to the amount of active SAF-A (Fig S6D). To test this prediction, DNA FISH was performed using fluorescent oligo probes "oligo paint" at three different 2 Mbp regions on HSA11 with "Low" (11p14.1), "Medium" (11p15.1) and "High" (11q13.1) gene activity (Fig 6E). The resulting volumes of the loci were in qualitative agreement with our simulations: the higher the gene density, and hence level of nascent transcripts, the bigger the volume occupied by the locus and the more enhanced its decompaction (Fig 6F-G).

The relation between SAF-A density and gene locus size was then analysed, again using molecular dynamic simulations. To do so, we considered a highly active gene ($A = 5\,\sigma^{-2}$) and gradually reduced the overall number of proteins in the simulation box to mimic experimental knock down of the SAF-A protein (Fig 7A, Fig S7A; Video 3-4). A 10-fold reduction in SAF-A protein concentration produced a 2-fold reduction in gene locus size, measured by radius of gyration (R_g, Fig 7B). To test this prediction experimentally, the volume of a 2 Mbp sized region on 11q13.1 was analysed using oligopaint DNA FISH in control and SAF-A depleted cells. In line with simulations, the volume of the locus in SAF-A depleted cells was dramatically reduced (Fig 7C).

To finally characterise how disruption of the SAF-A microgel by XRN2 modulates chromatin folding, simulations were performed starting from a condition of a highly active gene locus (Fig S7B) and a pre-formed SAF-A gel stabilising its decompacted state (Fig S7C). Over time the RNA activity was decreased (corresponding to transcription inhibition) for different RNA decay lengths (equivalent to different XRN2 activities) and the size of the gene locus monitored (Fig 7d, Fig S7C-D). For higher degradation rates (corresponding to smaller decay length, $\lambda = 10\,\sigma$) the decompacted locus collapsed precipitously; for lower degradation rates (corresponding to larger decay length, $\lambda = 20\,\sigma$), the collapse was less pronounced. These results suggested that XRN2 depletion should slow down, or stop altogether, locus compaction, after transcription inhibition. To test this prediction, oligopaint DNA FISH was utilised at the 11q13.1 locus (Fig 7E). After transcriptional inhibition there was a significant decrease in locus volume in control and EXOSC2 knockdown. Instead, as expected from simulations, XRN2 depletion abrogated chromatin compaction, consistent with the idea that once XRN2 can no longer degrade RNA efficiently, the microgel droplets grow, driving chromatin decompaction.

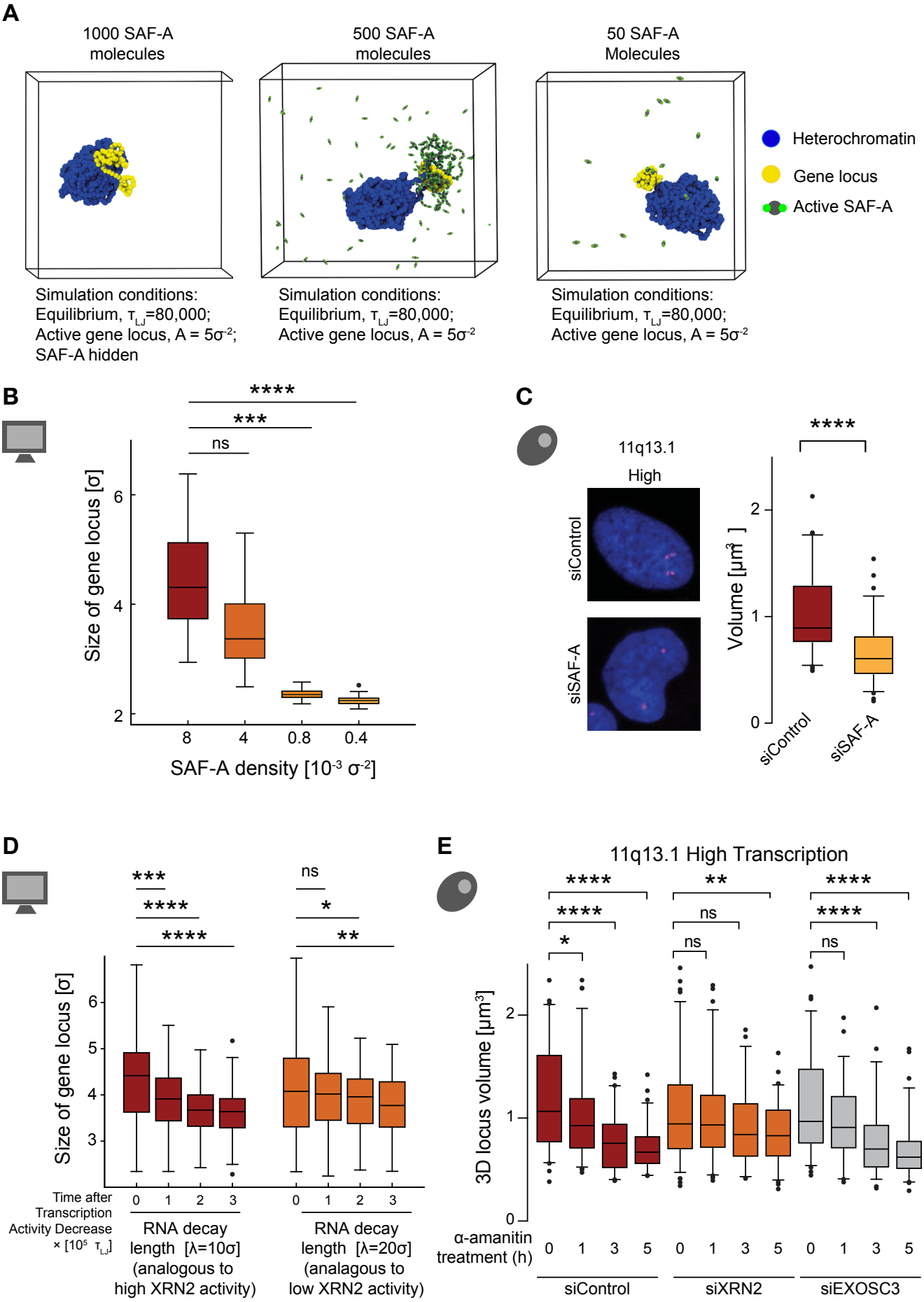

A
1000 SAF-A molecules
500 SAF-A molecules
50 SAF-A Molecules
Heterochromatin
Gene locus
Active SAF-A
Simulation conditions: Equilibrium, τLJ=80,000; Active gene locus, A = 5σ-2; SAF-A hidden
Simulation conditions: Equilibrium, τLJ=80,000; Active gene locus, A = 5σ-2
Simulation conditions: Equilibrium, τLJ=80,000; Active gene locus, A = 5σ-2
B
Size of gene locus [σ]
SAF-A density [10-3 σ-2]
8
4
0.8
0.4
ns
***
****
C
11q13.1
High
siControl
siSAF-A
Volume [μm3]
****
D
Size of gene locus [σ]
Time after Transcription Activity Decrease × [105 τLJ]
RNA decay length [λ=10σ] (analogous to high XRN2 activity)
RNA decay length [λ=20σ] (analagous to low XRN2 activity)
E
11q13.1 High Transcription
3D locus volume [μm3]
α-amanitin treatment (h)
siControl
siXRN2
siEXOSC3

Figure 7

Figure 7. Polymer model predicts effect of RNA/SAF-A on chromatin compaction

A. Molecular dynamic simulation snapshots of chromatin folding for inactive (blue) and highly active (yellow) chromatin regions, for different numbers of SAF-A molecules.

B. Boxplot of molecular dynamic simulation data showing the relationship between SAF-A density and the size of the active gene locus (n = 25 simulations per conditions). P values are for a T-test

C. Left, representative FISH images of oligoprobes hybridised to a highly transcribed region at 11q13.1 in control and SAF-A depleted cells. Right, boxplot showing volume of FISH probes (n > 65). P values are for a Wilcoxon test.

D. Boxplots showing the size of the gene locus determined from molecular dynamic simulations with different RNA decay lengths (analogous to XRN2 activity; red, high; orange, low) at different times after reducing transcription (n = 100 simulations per condition). P values are for a T-test.

E. Boxplots showing volume of oligopaint FISH probes at the highly transcribed 11q13.1 region for control, XRN2 depleted and EXOSC3 depleted cells, after transcription inhibition (n > 60). P values are for a Wilcoxon test.

P values correspond to; ns, not significant; *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$; ****, $p < 0.0001$.

Discussion

In this work, we have investigated the morphology and structural properties of the dynamic mesh formed by nuclear RNA and SAF-A, and its mechanistic role on modulating chromatin decompaction. Overall, our work suggests that SAF-A contributes to forming a supramolecular network of interconnected microphase separated RNA hydrogels, or microgels. We show that the physical parameters of this nuclear superstructure, such as the size of the microgels or connectivity between them, responds to biological inputs such as transcriptional activity, RNA turnover and SAF-A concentration. Specifically, the finding that RNA provides a glue to facilitate interconnection between microgels underscores the importance of RNA in structuring the organisation of nuclear interior, as also shown by other recent works[7,83]. Studying the perturbation of microgel size and connectivity between microgels when SAF-A is mutated will be a natural next step in this context, especially for mutations in the SPRNK domain, that have been identified in patients with SAF-A/HNRNPU neurodevelopmental disorders[3,41–43].

Through dSTORM, we detected ~4,000 SAF-A microgels per nucleus, with an average size 120 nm. By assuming that on average there are ~10,000 actively expressed genes in diploid RPE1 cells, and many more active enhancers, we can imagine that every microgel typically includes several regulatory elements and more than one gene. The average size and number of microgels also suggests that these are closely related to transcription factories, which are also about 90 nm in size and similar in number (although this varies 10-fold between different cell types[52]). Genome-wide simulations of chromatin structure via the “HiP-HoP” model[59,60] showed that bridging-induced phase separation (BIPS) provides a general mechanism to explain the formation of transcription factories, through a generic attraction arising between multivalent chromatin-binding proteins[57]. Mechanistically, we therefore hypothesise that transcription factories formed by BIPS nucleate the formation of RNA/SAF-A via local production of nascent RNA, with the recruitment of SAF-A. In our view, the resulting dynamic mesh might turn the scaffold into a suitable environment for the organisation of gene expression neighbourhoods[84,85]. This scaffold may help in the recruitment of co-activators, such as CBP/P300, or other nuclear proteins involved in RNA processing, such as SRSF1 (SF2) or SRSF2 (SC35). Investigating these issues may give new key insights on the link between RNA and nuclear protein organisation and gene expression.

From a functional perspective, we have shown that RNA/SAF-A microgels naturally link DNA transcription with chromatin organisation. In particular, we have shown by computer simulations that RNA/SAF-A microgel formation triggers a transition where the locus of sufficiently active genes expands dramatically, leading to decompaction of neighbouring chromatin regions as well, in good agreement with FISH experiments performed in chromosome regions with different gene density[35,86,87]. Within our model, chromatin decompaction depends on gelation close to the active locus, which sterically hinders the folding of chromatin in 3D, and effectively renders it locally stiffer. At genomic regions that are sufficiently transcriptionally active the chromatin fibre transitions from a rosette to an extended structure, see graphical abstract, with the surrounding nuclear environment

altering from a water-like to a gel-like phase. This observation may provide a mechanistic understanding for the observation of extreme decompaction, or gene melting, in some very active genes in the brain or thyroid[88,89], as well as an appealing way to reconcile the observation of this phenomenon with that of transcription factories.

Our data are consistent with a model where BIPS creates microgels of about 120 nm which broadly correspond to the size of transcription factories. However, euchromatin loops associated with a gene within the microgel can dramatically extend in 3D[86,90], if the gene activity increases beyond a critical threshold, as predicted by our model. We speculate this “gene popping” arises[86,90,91] as the number of active polymerases on a gene is sufficient to trigger the formation of enough RNA/SAF-A microgels to stiffen the chromatin such that it extends away from the factory of origin as observed in other studies[88,89], potentially leading to interconnections to other microgels or factories[92]. It is therefore no longer surprising that in the same cells some microgels correspond to conventional transcription factories, associated with chromatin looping, while others, following “gene popping”, due to a burst in activity, are associated with substantial chromatin decompaction.

Last, we come back to the nature of the nuclear interior[93,94], and to whether it is solid-like[45,50] or fluid-like[47]. One breakthrough of our work is that the RNA/SAF-A mesh is dynamic and its turnover (Fig 4A) is regulated by the continuous degradation of newly synthesised transcripts by XRN2. This process may also provide a pathway to RNA recycling, and could suggest the reuse of degraded nucleotides in the vicinity of a microgel, which itself is colocalised with an active gene. Degradation is also required to avoid the whole nucleus turning into a dense gel, which would hinder macromolecular diffusion and transport, and instead turn the RNA/SAF-A mesh into a “fluidised” structure that is continuously refreshed and controlled by RNA turnover over longer timescales. The net result is a viscoelastic heterogeneous environment, which shares fluid-like and solid-like properties[95] which are both length- and time-scale dependent and that are both probably essential for the regulation of transcriptional processes. To assess how this occurs mechanistically, we suggest it will be crucial to further investigate the material properties of the emergent microgel network we have described, and how they change by perturbing the RNA dynamics both *in vitro* and *in vivo*.

MATERIALS AND METHODS

Resource availability

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Nick Gilbert (nick.gilbert@ed.ac.uk).

Materials availability

All reagents generated in this study are available from the lead contact with a completed material transfer agreement if necessary.

Data and code availability

• All data are available in the manuscript or in the supplementary materials. All other data supporting the findings of this study are available from the lead contact upon request.
• This paper does not report original code.
• Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.
• For the purpose of open access, the authors have applied a Creative Commons Attribution (CC BY) licence to any author accepted manuscript version arising.
• The raw proteomics data for Fig S4A and Table S1 have been deposited to the ProteomeXchange PRIDE repository with the dataset identifier: XXXX.

Experimental model and study participant details

Cell culture

hTERT-RPE1 (ATCC, cat# ATCC-CRL-4000), HCT116, 293T and FlpIn T-REx 293 cells were cultured as described previously[35,96,97]. Transcription was blocked by adding α-amanitin (50 $\mu g.ml^{-1}$), triptolide (1 μM) and/or actinomycin D (50 $ng.ml^{-1}$) to cells for the times indicated. To label newly transcribed RNA, cells were incubated with 5-ethynyl uridine (5EU, 0.5 mM), 4-thiouridine (4SU, 0.5 mM) or [5'-^{3}H]uridine for the indicated time. 5EU was fluorescently labelled with Alexa 647 or Alexa 488 through Click labelling as previously described[35]. FlpIn T-Rex 293 stable cell line for doxycycline inducible expression of myc-BioID2-SAF-A was established as described[35]. hTERT-RPE1 stable cell lines for doxycycline inducible expression of GFP-multimers was established by electroporation with CRISPR-Cas9 and donor plasmids (see below for more details). Two days after electroporation, cells were selected with 2 $mg.ml^{-1}$ neomycin and single clones were isolated. The cells were treated with 1 $mg.ml^{-1}$ doxycycline for 24 hr for inducing expression of GFP oligomer proteins. To deplete XRN2 using the auxin-inducible degron system, auxin was added to cells at 500 nM for the indicated times.

Plasmids

For the BioID2 assay, the myc-BioID2 fragment from the myc-BioID2-MCS plasmid (addgene #74223) was replaced with 3xFLAG in the FlpIn 3xFLAG-SAF-A plasmid[35]. For FRAP and FCS experiments, the SNAP gene was inserted between the 3xFLAG tag and SAF-A gene in the FlpIn 3xFLAG-SAF-A plasmid (described Nozawa et al 2017). GFP-multimers fragment (gift from Karsten Rippe) were inserted into the Rosa26 donor plasmid (addgene #114699). For protein expression the SPRNK domain (amino acids 221-679) was synthesised as a G-block (IDT) and cloned into a T7 expression vector pEC_K_3C (gift from Atlanta Cook) by Gibson assembly. For the SPRNK RNA binding mutant the following mutations were introduced into the G-block: K246E, AAT-GAA; K251E, AAA-GAA; R539E, AGA-GAA; K616E, AAA-GAA. All plasmids were verified by DNA sequencing.

Antibodies

anti-SAF-A 3G6, 1:5000 for western blotting, Abcam, ab10297.
anti-SAF-A, 1:200 for immunofluorescence, Bethyl, A300-690A.
anti-XRN1, 1:3000 for WB, 1:100 for IF, Bethyl, A300-443A
anti-XRN2, 1:3000 for western blotting, 1:300 for immunostaining, Bethyl, A301-103A.
anti-EXOSC3, 1:1000 for WB, 1:200 for IF, SantaCruz, sc98776
anti-GAPDH, 1:5000 for WB, Cell Signaling Technology, D16H11
anti-FLAG, 1:10000 for WB, Sigma, M2
anti-myc, 1:1000 for WB, Cell Signaling Technology, 2276S

RNA interference

For siRNA treatment, cells (10%–20% confluent) were transfected with 10 nM Stealth RNA (ThermoFisher) using Lipofectamine RNAi MAX (ThermoFisher) for the time indicated. siRNA sequences and the identification numbers are CCUGGGAAUCGTGGCGGATATAATA HSS104917 for SAF-A, GCGAAAUGAUAACUUCUAATT s29016 for XRN1, GGAAAGUUGUGCAGUCGUATT s22412 for XRN2, UCAACUUUGAAUAUAUCUCCA s83102 for EXOSC3. The control RNA is Stealth RNAi siRNA Negative Control, Med GC (ThermoFisher).

Domain analysis and alpha fold structure prediction

Sequence comparisons of SAF-A/HNRNPU with paralogues HNRNPUL1 and HNRNPUL2 were generated by selecting orthologous and paralogues sequences from model vertebrates using UNIPROT. Sequences were aligned using MAFFT with default settings[98]. The multiple sequence alignment was visualised using Jalview[99]. AlphaFold2 models of SAF-A (downloaded from AlphaFoldDB 29.03.2023) were used for surface property analysis. Surface electrostatics were calculated using the adaptive Poisson-Boltzmann solver (APBS) plugin in PyMol with charge surfaces ranging from -3kT/e to +3kT/e[100]. Sequences for phylogenetic analysis of SAF-A were selected using the orthologous matrix database (OMA, group 116311)[101] then supplemented with additional sequences by using PSI-BLAST. Conservation scores were calculated using ConSurf

with the multiple sequence alignment and AlphaFold2 models as input[102]. Missense variants for the transcript ENST00000444376.2 were extracted from gnomAD (v2.1.1)[103] and filtered for benign variants. Vd/Vp values were then calculated using the 1D-to-3D suite[104]. Protein conservation, missense variant data, secondary structure elements and phosphorylation sites were displayed using a modified version of PlotProtein[105]. Candidate residues for generating point mutations to block RNA binding were selected by inspecting the AlphaFold2 models with respect to surface properties. Lysine and arginine residues (K246E, K251E, R539E and K616E), which contribute to positively-charged regions, were identified on the SPRNK domain surface that also corresponded to a region of high conservation across vertebrates.

Expression and Purification of recombinant proteins

Plasmids expressing wild-type and RNA-binding mutant SPRNK domain were transformed into the T7-express BL21 *E.coli*. Cells were cultured in 2x TY broth, and the expression of recombinant proteins was induced by 1 mM IPTG and cultured overnight at 20 °C. Cell pellets were lysed using a Continuous High Pressure Cell Disruptor at 4 °C and pressure of 25 kPSI. Soluble fractions of protein were purified by ion exchange chromatography on the ACTA Pure system using nickel pre-charged IMAC-HiTrap columns (Cytiva) and the peptides were size separated on a Superdex 200 column. Both, wild type and the RNA-binding mutant were the same size (54kDa) and similar yield.

Electrophoretic Mobility Shift Assays

To test the binding affinity of the wild-type and RNA-binding SPRNK domain to RNA oligos, and Electrophoretic Mobility Shift Assays (EMSA) were performed. Recombinant protein at increasing concentrations (0 - 8 μM) were incubated with 0.125 μM fluorophore labeled RNA oligomer (oligoU20-DY681, oligoU40-DY681, oligoA20-DY681, oligoC20-DY681) for 1h on ice in binding buffer (8 mM HEPES pH 7.5, 60 mM KOAc and 16 mM MgOAc) and the reaction products were separated by electrophoresis on 7% acrylamide gel and imaged on the Li-Cor Odyssey instrument. Free and bound RNA were quantified by Image Studio Lite v.5.2.2 software from Li-Cor. Three independent experiments were performed for each condition.

[5-^{3}H]uridine incorporation

Nascent transcription was measured as described previously[35,97].

PAR-CLIP

Photoactivatable Ribonucleoside-Enhanced Crosslinking and Immunoprecipitation (PAR-CLIP) was performed as previously described[106]. FlpIn T-REx 293 expressing FLAG-SAF-A[35] were grown overnight and treated with doxycycline for 24 h to induced FLAG-SAF-A expression. Cells were treated with 4sU to 0.5 mM final concentration for 15 min, washed with PBS, incubated for 1 h in complete medium (without 4sU), washed with PBS and exposed to 365 nm UV for 30 sec (0.15

J/cm^2). Cells were harvested by scrapping, and collected by centrifugation at 1300 rpm for 3 min in a microcentrifuge. The supernatant was discarded and cells snap frozen in dry ice and stored at -80°C. The remainder of the protocol was as described[106]. After sample processing sequencing libraries were prepared with a "NEB small RNA library kit". Samples were amplified using NEBNext index 1 and NEBNext SR primer and pair-end sequenced (75 bp). Adaptors were removed with cutadapt and samples were aligned using bowtie2, with an overall alignment rate of 56%. MACS was used for peak calling, and annotated using ChIPseeker. Scripts are available on request.

Oligopaint FISH

Cells were grown overnight or for two days on cover slips. Cover slips were rinsed with PBS and fixed in 2% para-formaldehyde for 10 min. Slides were rinsed with PBS and cells were permeabilized for 10 min with PBS supplemented with 0.2% Triton X-100. After rinsing, slides were stored in 70% Ethanol at 4°C. For processing, slides were dehydrated through ethanol and air dried. After denaturation in 70% formamide/2×SSC, pH 7.5, at 80°C for 30 min, slides were then transferred to 70% ethanol on ice, dehydrated in ethanol series, air-dried and hybridized with 300 ng of oligoprobes in 70 % formamide, 2×SSC, 10% dextran sulfate, 2 mg.ml-1 PureLink RNAse A (thermofisher) at 37°C overnight. Slides were washed with 2×SSC at 45°C for 20 min. Slides were counterstained with 0.5 µg.ml^{-1} DAPI.

myTags 20K immortal libraries (Arbor biosciences) for oligopaint FISH were synthesized to cover 2 Mb of genomic region. Gene density “High” probes (10882 probes) cover chr 11q13.1, 64500000-66500000, “Middle” probes (10801 probes) cover chr 11p15.1, 17000000-19000000, and “Low” probes (10830 probes) cover chr 11p14.1, 29500000-31500000. Oligopaint probes were prepared using the Cy3 labeled oligo CGTGGTCGCGTCTCA (Eurofin) according to the manufacturer’s instructions (Arbor biosciences).

Image capture and analysis

Oligopaint FISH slides were imaged using an LSM880 confocal laser scanning microscope (Zeiss) using 63x objective lens (0.4 NA Plan Apochromat Oil DIC M27) with immersion oil (Zeiss Immersol 518F; refractive index 1.518). Scale of x, y, was set at 0.052 µm/pixel, respectively. Eleven images per cell were acquired at 0.3-µm intervals and were Z-stacked with maximum intensity projection (Image J, version 2.1.0/1.53c). The volume of oligopaint FISH signals was quantified with ImageJ plugin 3D object Counter.

5-EU incorporation

Cells were treated with 1 mM 5-ethynyl uridine (Base Click) and 1 mM thymidine (Sigma-Aldrich) for the indicated time. Where stated cells were pretreated with 50 ng.ml^{-1} actinomycin D. Click

reaction was performed with Click-iT EdU Alexa Fluor 647 Imaging Kit or Click-iT Nascent RNA Capture Kit, according to the manufacturer's instructions (ThermoFisher).

Imaging and analysis of 5EU

5EU distribution in nuclei was imaged using a Photometrics Coolsnap HQ2 CCD camera and a Zeiss Axioplan II fluorescence microscope with Plan-neofluar/apochromat objective lenses (Carl Zeiss, Cambridge, UK), a Mercury Halide fluorescent light source (Exfo Excite 120, Excelitas Technologies) and Chroma #83000 triple band pass filter set (Chroma Technology Corp., Rockingham, VT) with the single excitation and emission filters installed in motorised filter wheels (Prior Scientific Instruments, Cambridge, UK). Image capture was performed using Micromanager (Version 1.4).

For the 5EU time-course (Fig S1G) images were captured using an Andor Dragonfly Spinning Disk Confocal with a 100x objective. 5EU was quantified as follows. Six images containing multiple nuclei were collected for each condition. Nuclei were segmented using the DAPI signal with a home-made script and the mean grey value of 5EU signal was extracted for each nucleus. The mean grey value was normalised to the average mean grey value at 0.5 h (peak value). The difference was calculated between normalised values.

dSTORM imaging and SuperStructure analysis

Immuno-staining for dSTORM imaging.

hTERT-RPE1 cells were grown in an 8-well Lab-Tek II Chambered Coverglass -- 1.5 borosilicate glass (Thermofisher scientific) at 37°C in order to have a final confluency of ~80%. Cells were fixed with 4% PFA (Sigma-Aldrich) in PBS for 10 minutes, followed by washing in PBS, permeabilisation with 0.2% Triton X-100 (Sigma-Aldrich) for 10 minutes, washed in PBS again and blocked with 1% BSA (Sigma-Aldrich) for 10 minutes.

In the case of SAF-A staining, immuno-fluorescence labelling was done by exposing the cells for 2 h to the primary SAF-A/HNRNPU polyclonal rabbit antibody (A300-690A, Bethyl Laboratories) at 10 $\mu g.ml^{-1}$ and washed. Cells were incubated for 1 hour with the secondary antibody. The secondary antibody was AffiniPure $F(ab')_2$ donkey anti-rabbit or donkey anti-mouse IgG (H+L) (711-006-152 and 715-007-003, Jackson ImmunoResearch Europe Ltd) conjugated to the organic fluorophore CF647 (MX647S100, Sigma-Aldrich) at a stoichiometric ratio of about 0.8 fluorophores per antibody molecule.

For RNA staining, 5EU was fluorescently labelled with Cy5 picolyl azide (1 µM) through Click labelling[35] for 30 min and then washed.

When performing two-colour dSTORM where both SAF-A and RNA were labelled, immunofluorescence for SAF-A was performed first and then click labelling for RNA. In this case,

the secondary antibody for labelling SAF-A was conjugated to the organic fluorophore CF568 (MX568S100, Sigma-Aldrich).

Oxygen scavenger imaging buffer for dSTORM was prepared fresh on the day of use as described[107]. The buffer was prepared as (i) 5.3 ml 200 mM Tris, 50 mM NaCl was mixed with (ii) 2 ml 40% glucose solution, (iii) 200 µl GLOX, (iv) 1.32 ml 1M 2ME (Sigma-Aldrich) and (v) 100ul 50 $\mu g.ml^{-1}$ DAPI solution (Sigma-Aldrich). The GLOX solution was prepared by mixing 160 µl 200 mM Tris/50 mM NaCl with 40 µl catalase from bovine liver (Sigma-Aldrich) and 18 mg glucose oxidase (Sigma-Aldrich).
The final volume was 8.9 ml which was sufficient to fill the chambers of the 8-well dish; a coverglass was seated on top of the dish to minimise exposure to air.

dSTORM Acquisition.
STORM acquisitions were performed at room temperature using a Nikon N-STORM system with Eclipse Ti-E inverted microscope with laser TIRF illuminator (Nikon UK Ltd, Kingston Upon Thames, UK). The microscope was equipped with a CFI SR HP Apo TIRF 100x objective lens (N.A. 1.49) with a 1.5x additional optical zoom. A cylindrical astigmatic lens was used to obtain elliptical shapes for emitters that reflect their z-position[108]. Laser light was provided via a Nikon LU-NV laser bed with 405, 488, 561 and 640 nm laser lines. CF647/Cy5 and CF568 fluorophores were stochastically excited using the 640 nm and 561 nm laser lines respectively, with an additional 405 nm weak pulse. A far-red and a multi-colour filter cube were used for single- and two-colour experiments, respectively. Images were acquired with an Andor iXon 897 EMCCD camera (Andor technologies, Belfast UK). The Z position was stabilised during the entire acquisition by the integrated perfect focus system (PFS). For each nucleus a stack of 20,000 frames of 10 or 20 ms exposure time were acquired using the Nikon NIS-Element software. For two-colour dSTORM the far-red (CF647) signal was collected first, and then the red (CF568) signal. Acquired images were 256 x 256 pixels with equal pixel being 106 nm. For every condition at least 10-15 nuclei (i.e. 10-15 independent datasets) were imaged.

Raw images and post-processing analysis.
The raw stack of frames was initially segmented based on a DAPI marker to mask out extra-nuclear signal. Frames were analysed using the Thunderstorm plugin[109] in FIJI. Firstly, frames were filtered by using the Wavelet functions to separate signal from noise. The B-Spline order was set to 3 and the B-Spline scale to 2 for localisations of around 5 pixels. In order to localise the emitter centroids, filtered images were thesholded (threshold value was set to 1.2 times the standard deviation of the 1st Wavelet function) and calculated the local maximum relative to the 8 nearest neighbours. Finally, the emitters signal distribution was fitted to elliptical gaussians (ellipses are necessary for z-position

reconstruction) using the weighted least square method and by setting 3 pixels as fitting radius and 2 pixels as sigma.
Localised data was then post-processed using the same plugin. (i) XY drift was corrected using a pair correlation analysis, (ii) filtered data with a position uncertainty < 40 nm, (iii) restricted the z-position to the interval [-100:100] nm and projected the data in a 2-dimensional plane, as the z-axis precision is around 100 nm.
Reconstructed images shown in the main text were created by using the average shifted histograms method of the Thunderstorm plugin with a 10x magnification (10.6 nm/pixel).

SuperStructure analysis for cluster connectivity.
dSTORM data was analysed with the SuperStructure algorithm to extract the connectivity between protein clusters[63]. For every nucleus, 5 to 8 local 1.5 µm circular ROIs with were defined and SuperStructure was run within such regions. SuperStructure curves were generated by spanning its spatial parameter $R \in [0 - 200\ nm]$ with a rate $dR = 10\ nm$. The super-cluster exponential regime was then visually identified in the averaged curve (as a linear decay in log scale). For both RNA and SAF-A such a regime is $R \in [20 - 70\ nm]$. SuperStructure regimes were fitted with the exponential function $f(R) = k \cdot e^{-R/\lambda}$ and the decay length λ was extracted. For every SAF-A simulation, we also calculated the normalised $\lambda^* = \lambda \cdot \sqrt{\rho_{cl}}$ to remove the cluster density ρ_{cl} contribution from the decay length. In the main text, the decay rate $\tau = 1/\lambda$ is shown for better visual representation, as higher τ reflects higher connectivity. The cluster and localisation density were calculated for the same ROIs defined above. Note that in the case of RNA where no evident cluster peak is observed in the Ripley analysis, the localisation density was used instead of the cluster size as observable. Note also that due to spatial inhomogeneity in the RNA signal the SuperStructure analysis was run for the entire nucleus instead of user-defined ROIs.

Ripley's analysis for clustering evaluation.
Ripley's analysis is typically used to verify if a distribution of points is clustered within a certain length-scale, i.e. if it deviates form a random distribution of points.
In order to do that, one needs to calculate the average local density of points for different length-scales and properly normalise it. This is the so-called Ripley's K function:

$$K(r) = \frac{1}{\rho} \sum_{i=1}^{n} \frac{N_p^i(r)}{n}$$

where $N_p^i(r)$ is the number of points found within a distance from a given point i, n is the total number of points in the system, and ρ is the average density of points in the system. The summation is performed for all the points in the system. In the case of a random distribution of points, the density

of points is constant and therefore $N^{i}_{p-rand}(r) = \rho \cdot \pi r^2$. As a result, $K_{rand}(r) = \pi \cdot r^2$. If $K(r) > K_{rand}(r)$, points show a clustering behaviour at that length-scale, otherwise if $K(r) < K_{rand}(r)$, points are over-dispersed.
The best way to visualise is through Ripley's H function:

$$H(r) = \sqrt{\frac{K(r)}{\pi}} - r$$

By using this function, $H_{rand}(r) = 0$ and therefore a clustering phenotype at a given length-scale is represented by a peak of the $H(r)$ function.
This approach has been used to check the clustering properties of RNA and SAF-A. We performed this analysis on 5 to 8 local square ROIs of $3\ \mu m$ size for every condition. The radius r spanned in the range $r \in [0 - 400\ nm]$.

Combine DBSCAN and Superstructure for cluster analysis.
Cluster analysis was performed by combining SuperStructure with DBSCAN.
By visually investigating SuperStructure curves, we identified the value of the spatial parameter R to fix for a cluster analysis with DBSCAN. The cluster analysis was performed for that R and by fixing the second DBSCAN parameter, i.e. minimum number of localisations to define a cluster, $N_{min} = 0$. Finally, the investigation of the distribution of cluster sizes allowed us to define a cut-off to remove clusters below a certain size. In the case of SAF-A, we identified $R = 20\ nm$ and minimum cluster size $30\ localisations$. No cluster analysis was performed for RNA as no evident peak of clustering was observed in the Ripley's analysis. In that case, the density of localisations was measured.

Co-localisation analysis of SAF-A with RNA signal.
Co-localisation analysis was performed in sequential steps. Firstly, DBSCAN was run to perform the cluster analysis and extract the centre of protein clusters. Then, a map of the density of RNA in the image was calculated by coarse graining the localisations in pixels of size $10nm$. By doing so a 2D matrix was obtained where each $10nm$ bin has the sum of the RNA localisations from STORM raw data. We do this procedure because RNA displays a continuous structure that cannot be easily and unambiguously broken up into clusters. To correlate the SAFA clusters with RNA we employ a method similar to the calculation of radial density function (RDF). While RDF typically computes the presence of a particle at a given distance r from another particle, here we compute the cumulative intensity (which we assume proportional to the probability) to find RNA at a given distance from the centre of a SAFA cluster. A 2D map with dimensions $1\mu m \times 1\mu m$ was compiled, divided in $10nm$ bins of the probability of finding RNA at a certain distance in the x and y directions from SAFA clusters (without assuming the system to be isotropic). This procedure was repeated for all the

clusters in a given nucleus and across $n = 10$ nuclei. Maps were then averaged to obtain the figures in the main and supplementary figures.

FRAP and FCS experiments

Cell preparation for FRAP and FCS

hTERT-RPE1 cells carrying an inducible copy of GFP multimers and FlpIn T-Rex 293 cells carrying an inducible copy of 3×FLAG-SNAP-SAF-A were grown in 8-well Lab-Tek II Chambered Coverglass -- 1.5 borosilicate glass (Thermofisher scientific).

To label SNAP-SAF-A protein, cells were incubated with SNAP TMR-STAR (NEB) at 1.2 μM for FRAP and 0.06 μM (NEB) for FCS for 30 min and then washed cells in PBS. For acquisition, cell media was replaced with Leibovitz media (ThermoFisher).

FRAP acquisitions and analyses

FRAP images were acquired using a Leica SP5 confocal (Leica Microsystems UK Ltd, Milton Keynes, UK) using LAS-AF software with appropriate FRAP wizard. The microscope was equipped with 405 nm diode, Argon and, 561 and 648 nm laser lines, three Photomultiplier tubes and one HyD GaSP detector. Images were scanned using a 40× air objective (N.A. 0.85). Confocal resolution and frequency were respectively set to 128×128bit and 1000Hz in order to obtain a spatial resolution of 179 nm and a time resolution of 140 ms in the Fly-mode of the wizard. Signal of TMR-STAR was excited by using the 561 nm laser at 8% power. FRAP bleaching was performed within a circular region of radius (R) = 2 μm by using all available lasers at 100% power. We acquired 10 pre-bleaching frames (1.4 s), 30 bleaching frames (4.2 s) and 400 post-bleaching frames (56 x). For every condition, 20 to 30 FRAP events were acquired on independent nuclei.

FRAP analysis was mainly performed using home-written scripts.

Firstly, cell movement was corrected by using the TurboReg plugin in FIJI[110], then the mean intensity of the bleaching region $I^B(t)$ was extracted, as well as that of the entire nucleus $I^C(t)$. A first normalisation was performed in order to correct the overall bleaching of the signal in time, as well as to normalise the signal in the range $[0,1]$, where 1 corresponds to the normalised pre-bleaching signal:

$$I^N(t) = \frac{I^B(t)/I^C(t)}{I^B_{pre}(t)/I^C_{pre}(t)}$$

where $I^B_{pre}(t)$ and $I^C_{pre}(t)$ are respectively the average pre-bleach signal in the bleached region and in the entire nucleus.

A second normalisation was performed in order to shift the minimum intensity of the signal to 0, while maintaining the signal in the $[0:1]$ range. This is necessary for visually comparing FRAP curves for

different conditions, as they typically have different minima in the intensity (bleaching is not 100% efficient). Note, that this second normalisation does not influence the fitting of the curves. The new normalised intensity is therefore:

$$I^{N2}(t) = \frac{I^N(t) - I^N_{min}}{1 - I^N_{min}}$$

Being I^N_{min} the minimum intensity of the FRAP curve.
The recovery of FRAP curves ($I^{N2}(t)$) was fitted with an exponential function by assuming the existence of two mobile populations (fast and slow population):

$$f(t) = c_1\left(1 - e^{-t/\tau_1}\right) + c_2\left(1 - e^{-t/\tau_2}\right)$$

Where τ_1 and τ_2 are respectively the recovery time of the fast and slow population. These two times can be though as inversely proportional to the diffusion time of that population: $D \sim R^2/\tau$.
The fraction of the mobile population is defined as $MF = c_1 + c_2$. As a consequence, one defines the fraction of signal that is not recovered as the immobile fraction $IF = 1 - MF$. Note that IF can be defined also independently from the fitting function, by measuring the fraction of signal that is not being recovered.

FCS acquisitions and analyses
For SNAP-SAF-A, FCS data was acquired using Leica SP5 Confocal with appropriate wizard using a PicoHarp 300 TCSPC module and APD detectors (Picoquant). Images were scanned using a 40X water objective. Confocal resolution and frequency were respectively set to 256x256bit and 600Hz. Signal of TMR-STAR was excited by using a 561 laser. FCS acquisition was performed in a very small volume and we measured the intensity signal for 100 s. For every condition, ~50 FCS measurements were acquired on independent nuclei.

For GFP multimers, data was acquired with a Zeiss LSM880 equipped with the FCS module (Carl Zeiss, Jena, Germany). GFP was excited with a 488-nm Argon laser line and scanned using a 63× (N.A. 1.2) water immersion objective (Zeiss). To calibrate the size of the point spread function, a 100 nM solution of Rhodamine 6G was prepared in PBS buffer and the beam waist of the resulting data was fitted.

The acquired data was processed using Symphotime 2.6 or Zen 2.3 SP1. The signal was restricted within an appropriate time-window where no bleaching was observed (10s in case of SNAP-SAF-A). We calculated the autocorrelation of the signal as function of the time lag τ:

$$G(\tau) = \frac{\langle I(t)I(t+\tau)\rangle_t}{\langle I(t)\rangle_t^2} - 1$$

Where $I(t)$ is the signal intensity at time t and $\langle\cdot\rangle_t$ is the sliding average over all values of t. We then fitted the autocorrelation curve with a single-population free diffusion model[111]:

$$f(\tau) = \frac{1}{N} \cdot \frac{1}{1+\left(\frac{\tau}{\tau_D}\right)} \cdot \sqrt{\frac{1}{1+\left(\frac{\tau}{\tau_D}\right)\left(\frac{\omega_0}{\omega_z}\right)^2}}$$

Where the model contains three fitting parameters: the average number of particles inside the confocal volume N, the structure parameter $\kappa = \omega_0/\omega_z$ that describes the shape of the effective detection volume and the proteins' diffusion time τ_D. The values of ω_0 and ω_z are determined from a calibration with free-diffusing dye with known diffusion coefficient. Finally, the diffusion coefficient of proteins is calculated as $D = \omega_0^2/4\tau_D$.

SAF-A oligomerization assay and detection

SAF-A oligomer formation was analysed as described previously[35].

BioID2 assay

FlpIn T-REx 293 cells carrying an inducible copy of myc-BioID2-SAF-A or myc-BioID2 as negative control were incubated in the presence of 50 µM biotin for 24 hours. After collecting cells with trypsin treatment and PBS wash 3 times to remove free biotin, the cell pellet (corresponding to 2 × 10^7 cells) was lysed in 500 µL of RIPA Lysis buffer (Santa Cruz Biotechnology, sc-24948) and sonicated to reduce sample viscosity from lysed DNA. Biotinylated proteins were purified using streptavidin magnetic beads and subjected to proteomics analysis by the IGC mass spectrometry facility. Samples were Strep Precipitated and processed on a Kingsfisher Duo robot. IPs, washes and on-bead digests were performed using a Thermo Kingfisher Duo, all steps are at 5°C unless otherwise stated. Beads were transferred into 500 µl of cleared lysate and incubated for 2 h with mixing. Beads were transferred for two washes in RIPA buffer and three washes in non-detergent lysis buffer. On-bead digest was performed by transferring the washed beads into 100 µl 2M urea, 100 mM Tris-HCl pH 7.5, 1 mM DTT containing 0.3 µg trypsin (Promega) per sample, beads were incubated at 27°C for 30 min with mixing to achieve limited proteolysis. The beads were then removed and tryptic digest of the released peptides was allowed to continue for 9 h at 37°C. Reduced cysteine residues were alkylated by adding iodoacetamide solution to a final concentration of 50 mM and incubated 30 min at room temperature, in the dark. Trypsin activity was inhibited by acidification of samples to a concentration of 1% TFA. Samples were desalted on

a C18 Stage tip and eluates were analysed by HPLC coupled to a Q-Exactive mass spectrometer as described[112].

Peptides and proteins were identified and quantified with the MaxQuant software package (1.5.3.8), and label-free quantification was performed by MaxLFQ[113]. The search included variable modifications for oxidation of methionine, protein N-terminal acetylation, and carbamidomethylation as fixed modification. The FDR, determined by searching a reverse database, was set at 0.01 for both peptides and proteins.

Electron microscopy

293T cells expressing FLAG-SAF-A were suspended in ATPase buffer in the presence or absence of RNaseA/T1 for 10 min at 37°C. Samples were treated with BM(PEG)2 protein cross-linker for 5 min at RT (as described previously[35]) and quenched using L-cysteine. Following sonication to break genomic DNA, samples were clarified by centrifugation. Supernatant was incubated with FLAG-agarose beads for 2 h at 4°C and SAF-A was eluted with 3 × FLAG peptide. For TEM, samples were absorbed on coated grids and negatively stained with 2% uranyl acetate and visualised using a Philips/FEI CM120 Biotwin microscope.

Molecular dynamics simulations for chromatin decompaction by the RNA/SAF-A mesh

The mechanism of chromatin decompaction by the RNA/SAF-A mesh of microgels is investigated by performing molecular dynamics simulations. These evolve the equation of motion of explicitly resolved chromatin (as a bead-and-spring polymer) and SAF-A (as patchy particles), within an implicit heat bath that mimics the viscous nuclear environment. In our polymer model, we consider a stretch of chromatin in which an active gene locus is flanked by two heterochromatin regions. The gene locus acts as a source for an RNA field that triggers SAF-A oligomerisation. Such oligomerisation creates a mesh-like structure (microgel) that allows the decompaction of the active chromatin regions, as shown in experiments.

Polymer model for the chromatin fibre

The chromatin is modelled as a bead-spring polymer chain where every bead has a nominal diameter σ and corresponds to $10\,nm$, in line with our previous coarse-grained models for chromatin[59]. In our chromatin model, we consider the inner/central part of the chain as the active genomic locus, which is delimited by two highly sticky beads (representing regulatory elements, such as a promoter and an enhancer) and flanked by two stretches of heterochromatic beads. The connectivity of the entire chain is provided by a harmonic potential between consecutive beads (i.e. a spring with large stiffness):

$$U_h(d) = k_h(d - d_0)^2$$

Where d is the distance between two consecutive beads, $k_h = 200k_BT/\sigma^2$ and $d_0 = 1.1\sigma$. Note that k_BT is the energy unit, where k_B is the Boltzmann constant and T the temperature.
The chromatin chain is also subject to a bending potential to introduce a persistence length (or bending stiffness) of the fibre:

$$U_b(\theta) = k_b(1 + cos\theta)\,,$$

where θ is the angle between two consecutive bonds and $k_b = 3k_BT$. Such parameters allow to mimic a persistence length $l_p \simeq 3\sigma = 30\ nm$.
An additional potential between chromatin beads is set to (implicitly) mimic protein bridges which drive folding and compaction of chromatin, as well as to introduce steric interaction between beads. This is a cut-and-shifted Lennard-Jones potential:

$$U_{LJ}(d) = 4\epsilon\left[\left(\frac{\sigma}{d}\right)^{12} - \left(\frac{\sigma}{d}\right)^{6} - \left(\frac{\sigma}{d_c}\right)^{12} + \left(\frac{\sigma}{d_c}\right)^{6}\right] \quad for\ d < d_c\,,$$

and 0 elsewhere. Our choice of $\epsilon = 0.9k_BT$ and $d_c = 1.8\,\sigma$ sets up a weak self-attraction potential between chromatin beads of the same type (either heterochromatin or active chromatin). In the case of promoter/enhancer beads, we set $\epsilon = 3.0k_BT$ in order to obtain a strongly self-attractive potential that enforce the presence of a chromatin loop. In case of chromatin beads of different types, we set the parameters $\epsilon = k_BT$ and $d_c = 2^{1/6}\,\sigma$ in order to take into account only steric interactions.
Finally, we add one patch on each bead along chromatin. The patches on the active chromatin beads are considered to be newly synthesised RNA and can bind to those on SAF-A (see below), whereas patches on heterochromatin and promoters are inert.

The equation for the RNA field

The beads of the active gene locus can produce newly transcribed RNA at a rate k_{prod} and with a probability of 10%, i.e. we assume that closeby RNA-PolII proteins have a typical distance along the fibre of $10\ \text{beads} \equiv 100\ nm$. Newly transcribed RNA then diffuses with a (very small) coefficient D, and is degraded at rate k_{deg}. Instead of explicitly simulate each RNA molecule in the system (which would be too expensive computationally), we solve the diffusion equation associated with the dynamics of the RNA density field. Given an RNA-producing source at location $\vec{r}_0$ (i.e. a bead of the gene locus), the density of RNA signal $\rho(\vec{r}, t)$ evolves in space and time as the following reaction-diffusion equation

$$\partial_t\rho(\vec{r}, t) = \ D\,\nabla^2\rho(\vec{r}, t)\ -\ k_{deg}\,\rho(\vec{r}, t)\ +\ k_{prod}\delta(\vec{r} - \vec{r}_0)\,,$$

where the terms on the right side of the equation are respectively the diffusion, the degradation and the (localised) production terms. The function $\delta(\vec{r} - \vec{r}_0)$ is the Dirac delta.
Since we perform simulations at equilibrium (or near equilibrium condition), we can consider the steady state form of the equation, i.e. $\partial_t \rho(\vec{r}, t) = 0$ and the resulting equation is also known as screened Poisson, or Yukawa, equation, and can be solved with the method of Green's functions. In other words, one needs to solve

$$[\nabla^2 - \lambda^{-2}]\rho(\vec{r}) = -f(\vec{r})$$

where $f(\vec{r}) = k_{on}\delta(\vec{r} - \vec{r_0})/D$ is the source term and $\lambda = \sqrt{D/k_{deg}}$ is the decay length, which measures the typical distance over which RNA diffuses before being degraded in our case. This equation is solved by first finding the associated Green's function $G(\vec{r})$ that is the solution of the equation

$$[\nabla^2 - \lambda^{-2}]G(\vec{r}) = -\delta(\vec{r})$$

which, in Fourier space, reads

$$G(\vec{k}) = \frac{1}{[k^2 + \lambda^{-2}]}$$

In 3D, the inverse transform gives

$$G(\vec{r}) = \iiint d\vec{k} \frac{e^{i\vec{k}\cdot\vec{r}}}{[k^2 + \lambda^{-2}]} = \frac{e^{-|\vec{r}|/\lambda}}{4\pi|\vec{r}|}$$

Now, the solution of the full problem is found as

$$\rho(\vec{r}) = \iiint G(\vec{r} - \vec{r}')\, f(\vec{r}') d\vec{r}' = \frac{k_{on}}{4\pi D} \iiint \frac{e^{-|\vec{r}-\vec{r}'|/\lambda}}{|\vec{r} - \vec{r}'|} \delta(\vec{r}' - \vec{r}_0)\, d\vec{r}' = A \frac{e^{-|\vec{r}-\vec{r}_0|/\lambda}}{|\vec{r} - \vec{r}_0|}$$

where $A = k_{prod}/4\pi D$ can be viewed as the gene activity. This equation shows that the contribution to the concentration of RNA at position $\vec{r}$ is given by a factor proportional to how much RNA is produced by the active locus (A), multiplied by a screening Yukawa function that depends on how much RNA is degraded before reaching $\vec{r}$ by diffusion (λ). In virtue of the superposition principle, if we have n independent sources of RNA, the contribution to location $\vec{r}$ is the linear sum of contributions of the n sources. Therefore, thanks to this equation we can calculate the instantaneous

(steady-state) local concentration of newly synthesised RNA given the position of the n sources of RNA. The number of RNA molecules in the locus may be found by integrating the RNA density over a volume enclosing the gene locus.

Polymer model for SAF-A proteins

SAF-A proteins are simulated as diffusing beads in the space. Steric interactions (with either chromatin beads or other proteins) are considered through a Lennard-Jones potential similar to that used between chromatin beads, with $\epsilon = k_B T$ and $d_c = 2^{1/6}\,\sigma$. In order to implement the RNA-binding properties of SAF-A, each protein bead is enriched with two additional patches. First, we introduced an attractive interaction between SAF-A patches and those on active chromatin (representing newly synthesised RNA) via a Morse potential:

$$U_{morse}(d) = D_0[e^{-2\alpha d} - 2e^{-\alpha d}] \quad \text{for } d < \sigma$$

where $D_0 = 40 k_B T$ and $\alpha = 16\,\sigma^{-1}$. The idea is that SAF-A coats the active chromatin region, binding to its beads via RNA. Some form of scaffolding to the chromatin polymer is required in order to ensure that the self-assembled structure is not entropically pushed away by the polymer.

Second, in order to implement the formation of the RNA/SAF-A mesh, we considered a two-state model for SAF-A beads. Thus, any of the SAF-A proteins can be either in an on (RNA-bound) or off (not RNA-bound) state. The conversion rate k_{on} from the off to the on state is set to be proportional to the local (known) concentration of RNA $\rho(\vec{r})$, whereas the inverse, from the on to the off state, is set to be a constant turnover rate k_{off}. Thus, at any given timestep of the simulation, there is a subset of SAFA proteins in the on state, and another in the off state. When the protein is in the on state, the two patches can bind to the patches of other on SAF-A proteins via the same Morse potential introduced above. The values of the parameters in the Morse potential between SAF-A patches are the same as those between SAF-A and chromatin patches: they lead to the creation of frequent oligomers made of 5-10 patchy-particles with about 1 branch in 5 to10 monomers. In spite of the large energy strength set for this potential, the self-assembled structure is still dynamic due to the non-equilibrium switching to the off state at rate k_{off}.

Molecular Dynamics Simulations

The dynamics of each chromatin and protein bead i with mass m evolves according to a Langevin equation at constant temperature T:

$$m\partial_{tt}\vec{r}_i(t) = -\nabla_i U_{tot}(r) - \gamma\partial_t\vec{r}_i(t) + \sqrt{2\gamma k_B T}\,\vec{\eta}_i(t)$$

where U_{tot} is the total potential energy resulting from the addition of the above interaction terms, ∇_i is the gradient taken respect to the coordinates of the ith bead, γ is the friction coefficient and $\vec{\eta}_i$ is

the delta-correlated stochastic white noise acting on every bead. Note that every SAF-A bead with its 2 patches is defined as a rigid body with mass $m_{prot} = 3m$, while every chromatin bead is defined as a rigid body with mass $m_{chr} = 2m$. We set the friction coefficient equal to $\gamma = m_{chr}/2\tau_{LJ}$ for each chromatin bead + patch and to $\gamma = m_{prot}/3\tau_{LJ}$ for each protein bead + 2 patches, where $\tau_{LJ} = \sigma\sqrt{m/k_BT}$ is the Lennard-Jones time. This choice ensures that both polymer and protein beads in isolation (i.e. not bound or joined into a chain) would have the same diffusion coefficient (even though they are formed by 2 and 3 particles, respectively). With this choice, the typical simulation time is $\tau_{LJ} = 6\pi\eta_s\sigma^3/k_BT \simeq 1\,ms$ for $\sigma = 10\,nm$ and with the viscosity of the nucleus set to be $\eta_s = 200\,cP$[114]. The equation of motion is integrated with timestep $\Delta t = 0.002\tau_{LJ}$ using the LAMMPS simulation package[115].

Simulation Parameters

In our standard simulations we consider a chain of 1000 chromatin beads, organised as follows: 450 heterochromatic beads, 1 promoter bead, 100 beads of active gene locus, 1 promoter bead, 450 heterochromatic beads. By assuming that each chromatin bead is around $10\,nm$, the active gene locus in our simulations is $\sim 100\,kbp$. We then consider 1000 SAF-A protein diffusing in proximity of the chain with an overall protein density of $0.008\,\sigma^{-3}$. For what regards RNA, we set $\lambda = 10\,\sigma$ and $A = 5\,\sigma^{-2}$ as standard values. These correspond to a turnover rate for RNA k_{deg} of order $1\,/min$, an RNA diffusion coefficient $D \simeq 150\,nm^2/s$ and the rate of RNA production $k_{prod} \simeq 100\,s^{-1}$. For the SAF-A modelling, the SAF-A de-oligomerisation rate is set to $k_{off} = 1/t_{off}$, with $t_{off} = 600\,\tau_{LJ} = 0.6\,s$, whereas the oligomerisation rate is $k_{on} = k_0 \min\{1, \rho\}$, with $k_0 = 0.01\,\tau_{LJ}^{-1}$, and we recall ρ denotes the RNA density field.

Simulations of different gene activities/degradation rates and SAF-A knockdown

Initially, we simulated different gene activities by varying A in the range $[0.1, 5]\,\sigma^{-2}$ and different RNA decay rates by varying λ in the range $[1, 100]\,\sigma$. We have also simulated the knockdown of SAF-A by spanning the density of SAF-A proteins in the range $[0.0004, 0.008]\,\sigma^{-3}$. For every simulation, we mainly measured the size of the gene locus as the gyration radius of the beads forming the locus:

$$R_g = \sqrt{\frac{1}{2N^2}\sum_{i,j}\left(\vec{r}_i - \vec{r}_j\right)}$$

Where N is the total number of beads in the active locus and the sum is over all these beads. Values in boxplots represent the average gyration radius of the latest 2×10^7 configurations. These simulations were first equilibrated and then run for $80000\,\tau_{LJ}$. We performed and averaged over between 10 and 25 replicas.

Simulations of XRN2 mechanism

Additional simulations were then carried out to understand the role of XRN2 in chromatin decompaction. In such simulations, transcription inhibition was simulated by smoothly and linearly decreasing the gene activity from $A = 5\,\sigma^{-2}$ to $1\,\sigma^{-2}$ in a time-range of $40000\,\tau_{LJ}$ and then the size of the gene locus in time was analysed (as the gyration radius of the active gene locus). Knockdown of XRN2 was simulated by doubling λ (i.e. reducing k_{deg} by a factor 4). These simulations were run for $160000\,\tau_{LJ}$. We performed and averaged over 100 replicas.

Data and Software Availability

Proteomics data are available via ProteomeXchange with identifier PXD053208. The GEO accession numbers for PAR-CLIP are GSE269702. The following high-level data and analysis have been deposited in the Edinburgh DataShare (https://doi.org/10.7488/ds/7751): dSTORM, FRAP, FCS, simulations. Simulation scripts and example initialization configurations which can be used to run the simulations using the LAMMPS software are available at https://git.ecdf.ed.ac.uk/chromatin-lab/nuclear_rna_gel. Scripts and raw data for the structural analysis of SAF-A (Figure 2) are available at https://git.ecdf.ed.ac.uk/cooklab/hnrnpu. All raw data and further analysis scripts are available upon request. Data analysis and simulation software used are all open-source packages.

Acknowledgements

We thank members of the Gilbert and Marenduzzo labs for helpful discussions. We are grateful to Stuart Aitken (University of Edinburgh) for bioinformatics support, Jimi Wills (University of Edinburgh) for help with mass-spec, metabolomics and BioID2 analysis, the Institute of Genetics and Cancer Advanced Imaging Resource (AIR) and Kayoko Suenaga and Akihiko Ichikawa (Carl Zeiss Co.,Ltd.) for supporting FCS analysis. Thanks to Steven West (University of Exeter) for supplying XRN2 degron HCT116 cells, Tomomi Kiyomitsu (OIST) for supplying the ROSA26 donor plasmid and Karsten Rippe (German Cancer Research Center) for the multimeric GFP plasmids. M.M. was supported by a cross-disciplinary fellowship funded by the University of Edinburgh and the UK Medical Research Council (MC_UU_00009/2). R.-S. N. is supported by the Leading Initiative for Excellent Young Researchers, MEXT, Japan. This work was also supported by JSPS KAKENHI grant numbers 24H02286, 24H01381, 22H05608, 22K19465, 20H03190 (to R.-S. N.). D. Mi. and D. Ma. were supported by the European Research Council (CoG 648050 THREEDCELLPHYSICS). N.G. was supported by a UK Medical Research Council senior non-clinical fellowship (MR/J00913X/1) and funding from the MRC (MC_UU_00007/13). A.G.C. was supported by a Wellcome Trust Senior Fellowship (200898/Z/16/Z). S.W. is supported by an BBSRC EastBIO doctoral studentship. N.G and D.Ma. are Wellcome Investigators (223097/Z/21).

Author Contributions

M.M., D.Mi., R-S.N., R.C., J.S., J.M., O.F., E.L. undertook lab-based experiments, M.M., D.Mi., M.C. carried out simulations, S.W., J.M., A.C., undertook protein structural analysis and modelling, S.A., G.G., H.B. did bioinformatic analysis, N.G., R-S.N., and D.Ma. conceived the project, and all authors contributed to writing the manuscript.

Declaration of Interests

The authors declare no competing financial interests.

References

1. Pederson, T. (2000). Half a century of “the nuclear matrix.” Mol.Biol.Cell *11*, 799–805.
2. Nozawa, R.-S., and Gilbert, N. (2019). RNA: Nuclear Glue for Folding the Genome. Trends Cell Biol *29*. https://doi.org/10.1016/j.tcb.2018.12.003.
3. Marenda, M., Lazarova, E., and Gilbert, N. (2022). The role of SAF-A/hnRNP U in regulating chromatin structure. Curr Opin Genet Dev *72*, 38–44. https://doi.org/10.1016/j.gde.2021.10.008.
4. Nickerson, J.A. (2022). The ribonucleoprotein network of the nucleus: a historical perspective. Curr Opin Genet Dev *75*, 101940. https://doi.org/10.1016/j.gde.2022.101940.
5. Nickerson, J. (2001). Experimental observations of a nuclear matrix. J Cell Sci *114*, 463–474.
6. Nickerson, J.A., Krockmalnic, G., Wan, K.M., and Penman, S. (1997). The nuclear matrix revealed by eluting chromatin from a cross-linked nucleus. Proceedings of the National Academy of Sciences USA *94*, 4446–4450.
7. Creamer, K.M., Kolpa, H.J., and Lawrence, J.B. (2021). Nascent RNA scaffolds contribute to chromosome territory architecture and counter chromatin compaction. Mol Cell *81*, 3509-3525.e5. https://doi.org/10.1016/j.molcel.2021.07.004.
8. Agostini, F., Zagalak, J., Attig, J., Ule, J., and Luscombe, N.M. (2021). Intergenic RNA mainly derives from nascent transcripts of known genes. Genome Biol *22*, 136. https://doi.org/10.1186/s13059-021-02350-x.
9. Festenstein, R., Pagakis, S.N., Hiragami, K., Lyon, D., Verreault, A., Sekkali, B., and Kioussis, D. (2003). Modulation of heterochromatin protein 1 dynamics in primary Mammalian cells. Science *299*, 719–721. https://doi.org/10.1126/science.1078694.
10. Phair, R.D., Scaffidi, P., Elbi, C., Vecerová, J., Dey, A., Ozato, K., Brown, D.T., Hager, G., Bustin, M., and Misteli, T. (2004). Global nature of dynamic protein-chromatin interactions in vivo: three-dimensional genome scanning and dynamic interaction networks of chromatin proteins. Mol Cell Biol *24*, 6393–6402. https://doi.org/10.1128/MCB.24.14.6393-6402.2004.
11. Cisse, I.I., Izeddin, I., Causse, S.Z., Boudarene, L., Senecal, A., Muresan, L., Dugast-Darzacq, C., Hajj, B., Dahan, M., and Darzacq, X. (2013). Real-Time Dynamics of RNA Polymerase II Clustering in Live Human Cells. Science (1979) *341*, 664–667. https://doi.org/10.1126/science.1239053.
12. Nakayasu, H., and Berezney, R. (1991). Nuclear matrins: identification of the major nuclear matrix proteins. Proc Natl Acad Sci U S A *88*, 10312–10316. https://doi.org/10.1073/pnas.88.22.10312.
13. Fackelmayer, F.O., and Richter, A. (1994). Purification of two isoforms of hnRNP-U and characterization of their nucleic acid binding activity. Biochemistry *33*, 10416–10422. https://doi.org/10.1021/bi00200a024.
14. Romig, H., Fackelmayer, F.O., Renz, A., Ramsperger, U., and Richter, A. (1992). Characterization of Saf-a, a Novel Nuclear-Dna Binding-Protein From Hela-Cells with High-Affinity for Nuclear Matrix Scaffold Attachment Dna Elements. EMBO Journal *11*, 3431–3440.

15. Kiledjian, M., and Dreyfuss, G. (1992). Primary structure and binding activity of the hnRNP U protein: binding RNA through RGG box. EMBO Journal *11*, 2655–2664.
16. Geuens, T., Bouhy, D., and Timmerman, V. (2016). The hnRNP family: insights into their role in health and disease. Hum Genet *135*, 851–867. https://doi.org/10.1007/s00439-016-1683-5.
17. Gillentine, M.A., Wang, T., Hoekzema, K., Rosenfeld, J., Liu, P., Guo, H., Kim, C.N., De Vries, B.B.A., Vissers, L.E.L.M., Nordenskjold, M., et al. (2021). Rare deleterious mutations of HNRNP genes result in shared neurodevelopmental disorders. Genome Med *13*, 63. https://doi.org/10.1186/s13073-021-00870-6.
18. Thandapani, P., O\textquoterightConnor, T.R., Bailey, T.L., and Richard, S. (2013). Defining the RGG/RG Motif. Mol Cell *50*, 613–623. https://doi.org/10.1016/j.molcel.2013.05.021.
19. Ozdilek, B.A., Thompson, V.F., Ahmed, N.S., White, C.I., Batey, R.T., and Schwartz, J.C. (2017). Intrinsically disordered RGG/RG domains mediate degenerate specificity in RNA binding. Nucleic Acids Res *45*, 7984–7996. https://doi.org/10.1093/nar/gkx460.
20. Ritsch, I., Lehmann, E., Emmanouilidis, L., Yulikov, M., Allain, F., and Jeschke, G. (2022). Phase Separation of Heterogeneous Nuclear Ribonucleoprotein A1 upon Specific RNA-Binding Observed by Magnetic Resonance. Angew Chem Int Ed Engl *61*, e202204311. https://doi.org/10.1002/anie.202204311.
21. Molliex, A., Temirov, J., Lee, J., Coughlin, M., Kanagaraj, A.P., Kim, H.J., Mittag, T., and Taylor, J.P. (2015). Phase separation by low complexity domains promotes stress granule assembly and drives pathological fibrillization. Cell *163*, 123–133. https://doi.org/10.1016/j.cell.2015.09.015.
22. Lin, Y., Protter, D.S.W., Rosen, M.K., and Parker, R. (2015). Formation and Maturation of Phase-Separated Liquid Droplets by RNA-Binding Proteins. Mol Cell *60*, 208–219. https://doi.org/10.1016/j.molcel.2015.08.018.
23. Ponting, C., Schultz, J., and Bork, P. (1997). SPRY domains in ryanodine receptors (Ca(2+)-release channels). Trends Biochem Sci *22*, 193–194.
24. Helbig, R., and Fackelmayer, F.O. (2003). Scaffold attachment factor A (SAF-A) is concentrated in inactive X chromosome territories through its RGG domain. Chromosoma *112*, 173–182. https://doi.org/10.1007/s00412-003-0258-0.
25. Kletzien, O.A., Wuttke, D.S., and Batey, R.T. (2024). The RNA-Binding Domain of hnRNP U Extends beyond the RGG/RG Motifs. Biochemistry. https://doi.org/10.1021/acs.biochem.3c00510.
26. Xiao, R., Tang, P., Yang, B., Huang, J., Zhou, Y., Shao, C., Li, H., Sun, H., Zhang, Y., and Fu, X.-D. (2012). Nuclear Matrix Factor hnRNP U/SAF-A Exerts a Global Control of Alternative Splicing by Regulating U2 snRNP Maturation. Mol Cell *45*, 656–668. https://doi.org/10.1016/j.molcel.2012.01.009.
27. Ye, J., Beetz, N., O'Keeffe, S., Tapia, J.C., Macpherson, L., Chen, W. V., Bassel-Duby, R., Olson, E.N., and Maniatis, T. (2015). HnRNP U protein is required for normal pre-mRNA splicing and postnatal heart development and function. Proc Natl Acad Sci U S A *112*, E3020–E3029. https://doi.org/10.1073/pnas.1508461112.
28. Hasegawa, Y., Brockdorff, N., Kawano, S., Tsutui, K., Tsutui, K., and Nakagawa, S. (2010). The matrix protein hnRNP U is required for chromosomal localization of xist RNA. Dev Cell *19*, 469–476. https://doi.org/10.1016/j.devcel.2010.08.006.
29. McHugh, C.A., Chen, C.K., Chow, A., Surka, C.F., Tran, C., McDonel, P., Pandya-Jones, A., Blanco, M., Burghard, C., Moradian, A., et al. (2015). The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3. Nature *521*, 232–236. https://doi.org/10.1038/nature14443.
30. Douglas, P., Ye, R., Morrice, N., Britton, S., Trinkle-Mulcahy, L., and Lees-Miller, S.P. (2015). Phosphorylation of SAF-A/hnRNP-U Serine 59 by Polo-Like Kinase 1 Is Required for Mitosis. Mol Cell Biol *35*, 2699–2713. https://doi.org/10.1128/mcb.01312-14.
31. Sharp, J.A., Perea-Resa, C., Wang, W., and Blower, M.D. (2020). Cell division requires RNA eviction from condensing chromosomes. J Cell Biol *219*. https://doi.org/10.1083/jcb.201910148.
32. Connolly, C., Takahashi, S., Miura, H., Hiratani, I., Gilbert, N., Donaldson, A.D., and Hiraga, S.-I. (2022). SAF-A promotes origin licensing and replication fork progression to ensure robust DNA replication. J Cell Sci *135*. https://doi.org/10.1242/jcs.258991.

33. Blower, M.D., Wang, W., and Sharp, J.A. (2023). Differential nuclear import regulates nuclear RNA inheritance following mitosis. Mol Biol Cell *34*, ar32. https://doi.org/10.1091/mbc.E23-01-0004.
34. Fan, H., Lv, P., Huo, X., Wu, J., Wang, Q., Cheng, L., Liu, Y., Tang, Q.Q., Zhang, L., Zhang, F., et al. (2018). The nuclear matrix protein HNRNPU maintains 3D genome architecture globally in mouse hepatocytes. Genome Res *28*, 192–202. https://doi.org/10.1101/gr.224576.117.
35. Nozawa, R.S., Boteva, L., Soares, D.C., Naughton, C., Dun, A.R., Buckle, A., Ramsahoye, B., Bruton, P.C., Saleeb, R.S., Arnedo, M., et al. (2017). SAF-A Regulates Interphase Chromosome Structure through Oligomerization with Chromatin-Associated RNAs. Cell *169*, 1214-1227.e18. https://doi.org/10.1016/j.cell.2017.05.029.
36. Roshon, M.J., and Ruley, H.E. (2005). Hypomorphic mutation in hnRNP U results in post-implantation lethality. Transgenic Res *14*, 179–192. https://doi.org/10.1007/s11248-004-8147-8.
37. Sapir, T., Kshirsagar, A., Gorelik, A., Olender, T., Porat, Z., Scheffer, I.E., Goldstein, D.B., Devinsky, O., and Reiner, O. (2022). Heterogeneous nuclear ribonucleoprotein U (HNRNPU) safeguards the developing mouse cortex. Nat Commun *13*, 4209. https://doi.org/10.1038/s41467-022-31752-z.
38. Lai, B., Zou, J., Lin, Z., Qu, Z., Song, A., Xu, Y., and Gao, X. (2018). Haploinsufficiency of hnRNP U Changes Activity Pattern and Metabolic Rhythms. Am J Pathol *188*, 173–183. https://doi.org/10.1016/j.ajpath.2017.09.017.
39. Dugger, S.A., Dhindsa, R.S., Sampaio, G.D.A., Ressler, A.K., Rafikian, E.E., Petri, S., Letts, V.A., Teoh, J., Ye, J., Colombo, S., et al. (2023). Neurodevelopmental deficits and cell-type-specific transcriptomic perturbations in a mouse model of HNRNPU haploinsufficiency. PLoS Genet *19*, e1010952. https://doi.org/10.1371/journal.pgen.1010952.
40. Ressler, A.K., Sampaio, G.L.A., Dugger, S.A., Sapir, T., Krizay, D., Boland, M.J., Reiner, O., and Goldstein, D.B. (2023). Evidence of shared transcriptomic dysregulation of HNRNPU-related disorder between human organoids and embryonic mice. iScience *26*, 105797. https://doi.org/10.1016/j.isci.2022.105797.
41. Yates, T.M., Vasudevan, P.C., Chandler, K.E., Donnelly, D.E., Stark, Z., Sadedin, S., Willoughby, J., and Balasubramanian, M. (2017). De novo mutations in HNRNPU result in a neurodevelopmental syndrome. Am J Med Genet A *173*, 3003–3012. https://doi.org/10.1002/ajmg.a.38492.
42. Bramswig, N.C., Lüdecke, H.J., Hamdan, F.F., Altmüller, J., Beleggia, F., Elcioglu, N.H., Freyer, C., Gerkes, E.H., Demirkol, Y.K., Knupp, K.G., et al. (2017). Heterozygous HNRNPU variants cause early onset epilepsy and severe intellectual disability. Hum Genet *136*, 821–834. https://doi.org/10.1007/s00439-017-1795-6.
43. Leduc, M.S., Chao, H.T., Qu, C., Walkiewicz, M., Xiao, R., Magoulas, P., Pan, S., Beuten, J., He, W., Bernstein, J.A., et al. (2017). Clinical and molecular characterization of de novo loss of function variants in HNRNPU. Am J Med Genet A *173*, 2680–2689. https://doi.org/10.1002/ajmg.a.38388.
44. Mastropasqua, F., Oksanen, M., Soldini, C., Alatar, S., Arora, A., Ballarino, R., Molinari, M., Agostini, F., Poulet, A., Watts, M., et al. (2023). Deficiency of the Heterogeneous Nuclear Ribonucleoprotein U locus leads to delayed hindbrain neurogenesis. Biol Open *12*. https://doi.org/10.1242/bio.060113.
45. Strickfaden, H., Tolsma, T.O., Sharma, A., Underhill, D.A., Hansen, J.C., and Hendzel, M.J. (2020). Condensed Chromatin Behaves like a Solid on the Mesoscale In Vitro and in Living Cells. Cell *183*, 1772-1784.e13. https://doi.org/10.1016/j.cell.2020.11.027.
46. Nozaki, T., Imai, R., Tanbo, M., Nagashima, R., Tamura, S., Tani, T., Joti, Y., Tomita, M., Hibino, K., Kanemaki, M.T., et al. (2017). Dynamic Organization of Chromatin Domains Revealed by Super-Resolution Live-Cell Imaging. Mol Cell *67*, 282-293.e7. https://doi.org/10.1016/j.molcel.2017.06.018.
47. Nozaki, T., Shinkai, S., Ide, S., Higashi, K., Tamura, S., Shimazoe, M.A., Nakagawa, M., Suzuki, Y., Okada, Y., Sasai, M., et al. (2023). Condensed but liquid-like domain organization of active chromatin regions in living human cells. Sci Adv *9*, eadf1488. https://doi.org/10.1126/sciadv.adf1488.

48. Gibson, B.A., Doolittle, L.K., Schneider, M.W.G., Jensen, L.E., Gamarra, N., Henry, L., Gerlich, D.W., Redding, S., and Rosen, M.K. (2019). Organization of Chromatin by Intrinsic and Regulated Phase Separation. Cell *179*, 470-484.e21. https://doi.org/10.1016/j.cell.2019.08.037.
49. Gibson, B.A., Blaukopf, C., Lou, T., Chen, L., Doolittle, L.K., Finkelstein, I., Narlikar, G.J., Gerlich, D.W., and Rosen, M.K. (2023). In diverse conditions, intrinsic chromatin condensates have liquid-like material properties. Proc Natl Acad Sci U S A *120*, e2218085120. https://doi.org/10.1073/pnas.2218085120.
50. Keizer, V.I.P., Grosse-Holz, S., Woringer, M., Zambon, L., Aizel, K., Bongaerts, M., Delille, F., Kolar-Znika, L., Scolari, V.F., Hoffmann, S., et al. (2022). Live-cell micromanipulation of a genomic locus reveals interphase chromatin mechanics. Science *377*, 489–495. https://doi.org/10.1126/science.abi9810.
51. Jackson, D.A., Hassan, A.B., Errington, R.J., and Cook, P.R. (1993). Visualization of focal sites of transcription within human nuclei. *1*, 1059–1065.
52. Papantonis, A., and Cook, P.R. (2013). Transcription factories: genome organization and gene regulation. Chem Rev *113*, 8683–8705. https://doi.org/10.1021/cr300513p.
53. Cook, P.R., and Marenduzzo, D. (2018). Transcription-driven genome organization: a model for chromosome structure and the regulation of gene expression tested through simulations. Nucleic Acids Res *46*, 9895–9906. https://doi.org/10.1093/nar/gky763.
54. Caudron-Herger, M., Cook, P.R., Rippe, K., and Papantonis, A. (2015). Dissecting the nascent human transcriptome by analysing the RNA content of transcription factories. Nucleic Acids Res *43*. https://doi.org/10.1093/nar/gkv390.
55. Hozák, P., Hassan, A.B., Jackson, D.A., and Cook, P.R. (1993). Visualization of replication factories attached to nucleoskeleton. Cell *73*, 361–373. https://doi.org/10.1016/0092-8674(93)90235-i.
56. Gavrilov, A.A., Zukher, I.S., Philonenko, E.S., Razin, S. V, and Iarovaia, O. V (2010). Mapping of the nuclear matrix-bound chromatin hubs by a new M3C experimental procedure. Nucleic Acids Res *38*, 8051–8060. https://doi.org/10.1093/nar/gkq712.
57. Brackley, C.A., Taylor, S., Papantonis, A., Cook, P.R., and Marenduzzo, D. (2013). Nonspecific bridging-induced attraction drives clustering of DNA-binding proteins and genome organization. Proceedings of the National Academy of Sciences USA *110*, E3605–E3611. https://doi.org/10.1073/pnas.1302950110.
58. Brackley, C.A., Johnson, J., Kelly, S., Cook, P.R., and Marenduzzo, D. (2016). Simulated binding of transcription factors to active and inactive regions folds human chromosomes into loops, rosettes and topological domains. Nucleic Acids Res *44*, 3503–3512. https://doi.org/10.1093/nar/gkw135.
59. Chiang, M., Brackley, C.A., Marenduzzo, D., and Gilbert, N. (2022). Predicting genome organisation and function with mechanistic modelling. Trends Genet *38*, 364–378. https://doi.org/10.1016/j.tig.2021.11.001.
60. Chiang, M., Brackley, C.A., Naughton, C., Nozawa, R.-S., Battaglia, C., Marenduzzo, D., and Gilbert, N. (2022). Gene structure heterogeneity drives transcription noise within human chromosomes. bioRxiv. https://doi.org/10.1101/2022.06.09.495447.
61. Brackley, C.A., and Marenduzzo, D. (2020). Bridging-induced microphase separation: photobleaching experiments, chromatin domains and the need for active reactions. Brief Funct Genomics *19*, 111–118. https://doi.org/10.1093/bfgp/elz032.
62. Brackley, C.A., Liebchen, B., Michieletto, D., Mouvet, F., Cook, P.R., and Marenduzzo, D. (2017). Ephemeral Protein Binding to DNA Shapes Stable Nuclear Bodies and Chromatin Domains. Biophys J *112*, 1085–1093. https://doi.org/10.1016/j.bpj.2017.01.025.
63. Marenda, M., Lazarova, E., van de Linde, S., Gilbert, N., and Michieletto, D. (2021). Parameter-free molecular super-structures quantification in single-molecule localization microscopy. J Cell Biol *220*. https://doi.org/10.1083/jcb.202010003.
64. Plamper, F.A., and Richtering, W. (2017). Functional Microgels and Microgel Systems. Acc Chem Res *50*, 131–140. https://doi.org/10.1021/acs.accounts.6b00544.
65. Gnan, N., Rovigatti, L., Bergman, M., and Zaccarelli, E. (2017). In Silico Synthesis of Microgel Particles. Macromolecules *50*, 8777–8786. https://doi.org/10.1021/acs.macromol.7b01600.

66. Sciortino, F., and Zaccarelli, E. (2011). Reversible gels of patchy particles. Curr Opin Solid State Mater Sci *15*, 246–253. https://doi.org/10.1016/j.cossms.2011.07.003.
67. Yonchev, I.D., Apostol, C. V, Griffith, L., Li, A., Butler, L., Patrick, H., Aguilar-Martinez, E., Whelan, A.G.R., Evans, C., Dickman, M.J., et al. (2023). hnRNPUL1 ensures efficient Integrator-mediated cleavage of snRNAs and is mutated in amyotrophic lateral sclerosis. bioRxiv. https://doi.org/10.1101/2023.07.22.550030.
68. Van Nostrand, E.L., Freese, P., Pratt, G.A., Wang, X., Wei, X., Xiao, R., Blue, S.M., Chen, J.-Y., Cody, N.A.L., Dominguez, D., et al. (2020). A large-scale binding and functional map of human RNA-binding proteins. Nature *583*, 711–719. https://doi.org/10.1038/s41586-020-2077-3.
69. Yugami, M., Okano, H., Nakanishi, A., and Yano, M. (2020). Analysis of the nucleocytoplasmic shuttling RNA-binding protein HNRNPU using optimized HITS-CLIP method. PLoS One *15*, e0231450. https://doi.org/10.1371/journal.pone.0231450.
70. Huelga, S.C., Vu, A.Q., Arnold, J.D., Liang, T.Y., Liu, P.P., Yan, B.Y., Donohue, J.P., Shiue, L., Hoon, S., Brenner, S., et al. (2012). Integrative genome-wide analysis reveals cooperative regulation of alternative splicing by hnRNP proteins. Cell Rep *1*, 167–178. https://doi.org/10.1016/j.celrep.2012.02.001.
71. Michieletto, D., and Gilbert, N. (2019). Role of nuclear RNA in regulating chromatin structure and transcription. Curr Opin Cell Biol *58*, 120–125. https://doi.org/10.1016/j.ceb.2019.03.007.
72. Hihara, S., Pack, C.-G., Kaizu, K., Tani, T., Hanafusa, T., Nozaki, T., Takemoto, S., Yoshimi, T., Yokota, H., Imamoto, N., et al. (2012). Local Nucleosome Dynamics Facilitate Chromatin Accessibility in Living Mammalian Cells. *2*, 1645–1656. https://doi.org/10.1016/j.celrep.2012.11.008.
73. Yang, L., Gal, J., Chen, J., and Zhu, H. (2014). Self-assembled FUS binds active chromatin and regulates gene transcription. Proc Natl Acad Sci U S A *111*, 17809–17814. https://doi.org/10.1073/pnas.1414004111.
74. Sears, R.M., May, D.G., and Roux, K.J. (2019). BioID as a Tool for Protein-Proximity Labeling in Living Cells. Methods Mol Biol *2012*, 299–313. https://doi.org/10.1007/978-1-4939-9546-2_15.
75. Miki, T.S., and Großhans, H. (2013). The multifunctional RNase XRN2. 825–830. https://doi.org/10.1042/BST20130001.
76. Davidson, L., Kerr, A., and West, S. (2012). Co-transcriptional degradation of aberrant pre-mRNA by Xrn2. EMBO J *31*, 2566–2578. https://doi.org/10.1038/emboj.2012.101.
77. Viegas, J.O., Azad, G.K., Lv, Y., Fishman, L., Paltiel, T., Pattabiraman, S., Park, J.E., Kaganovich, D., Sze, S.K., Rabani, M., et al. (2022). RNA degradation eliminates developmental transcripts during murine embryonic stem cell differentiation via CAPRIN1-XRN2. Dev Cell *57*, 2731-2744.e5. https://doi.org/10.1016/j.devcel.2022.11.014.
78. Fong, N., Brannan, K., Erickson, B., Kim, H., Cortazar, M.A., Sheridan, R.M., Nguyen, T., Karp, S., and Bentley, D.L. (2015). Effects of Transcription Elongation Rate and Xrn2 Exonuclease Activity on RNA Polymerase II Termination Suggest Widespread Kinetic Competition. Mol Cell *60*, 256–267. https://doi.org/10.1016/j.molcel.2015.09.026.
79. Eaton, J.D., Davidson, L., Bauer, D.L.V., Natsume, T., Kanemaki, M.T., and West, S. (2018). Xrn2 accelerates termination by RNA polymerase II, which is underpinned by CPSF73 activity. Genes Dev *32*, 127–139. https://doi.org/10.1101/gad.308528.117.
80. Michieletto, D., Orlandini, E., and Marenduzzo, D. (2016). Polymer model with epigenetic recoloring reveals a pathway for the de novo establishment and 3D organization of chromatin domains. Phys Rev X *6*, 41047. https://doi.org/10.1103/PhysRevX.6.041047.
81. Jost, D., Carrivain, P., Cavalli, G., and Vaillant, C. (2014). Modeling epigenome folding: Formation and dynamics of topologically associated chromatin domains. Nucleic Acids Res *42*, 9553–9561. https://doi.org/10.1093/nar/gku698.
82. Brackley, C.A., Brown, J.M., Waithe, D., Babbs, C., Davies, J., Hughes, J.R., Buckle, V.J., and Marenduzzo, D. (2016). Predicting the three-dimensional folding of cis-regulatory regions in mammalian genomes using bioinformatic data and polymer models. Genome Biol *17*, 59. https://doi.org/10.1186/s13059-016-0909-0.
83. Hall, L.L., Carone, D.M., Gomez, A. V., Kolpa, H.J., Byron, M., Mehta, N., Fackelmayer, F.O., and Lawrence, J.B. (2014). Stable C0T-1 repeat RNA is abundant and is associated

with euchromatic interphase chromosomes. Cell *156*, 907–919. https://doi.org/10.1016/j.cell.2014.01.042.
84. Oliver, B., Parisi, M., and Clark, D. (2002). Gene expression neighborhoods. J Biol *1*, 4. https://doi.org/10.1186/1475-4924-1-4.
85. Shopland, L.S., Johnson, C. V, Byron, M., McNeil, J., and Lawrence, J.B. (2003). Clustering of multiple specific genes and gene-rich R-bands around SC-35 domains: evidence for local euchromatic neighborhoods. Journal of Cell Biology *162*, 981–990. https://doi.org/10.1083/jcb.200303131.
86. Gilbert, N., Boyle, S., Fiegler, H., Woodfine, K., Carter, N.P., and Bickmore, W.A. (2004). Chromatin architecture of the human genome: Gene-rich domains are enriched in open chromatin fibers. Cell *118*, 555–566. https://doi.org/10.1016/j.cell.2004.08.011.
87. Naughton, C., Sproul, D., Hamilton, C., and Gilbert, N. (2010). Analysis of Active and Inactive X Chromosome Architecture Reveals the Independent Organization of 30 nm and Large-Scale Chromatin Structures. Mol Cell *40*, 397–409. https://doi.org/10.1016/j.molcel.2010.10.013.
88. Winick-Ng, W., Kukalev, A., Harabula, I., Zea-Redondo, L., Szabó, D., Meijer, M., Serebreni, L., Zhang, Y., Bianco, S., Chiariello, A.M., et al. (2021). Cell-type specialization is encoded by specific chromatin topologies. Nature *599*, 684–691. https://doi.org/10.1038/s41586-021-04081-2.
89. Leidescher, S., Ribisel, J., Ullrich, S., Feodorova, Y., Hildebrand, E., Galitsyna, A., Bultmann, S., Link, S., Thanisch, K., Mulholland, C., et al. (2022). Spatial organization of transcribed eukaryotic genes. Nat Cell Biol *24*, 327–339. https://doi.org/10.1038/s41556-022-00847-6.
90. Mahy, N.L., Perry, P.E., and Bickmore, W.A. (2002). Gene density and transcription influence the localization of chromatin outside of chromosome territories detectable by FISH. J Cell Biol *159*, 753–763. https://doi.org/10.1083/jcb.200207115.
91. Chambeyron, S., and Bickmore, W.A. (2004). Chromatin decondensation and nuclear reorganization of the HoxB locus upon induction of transcription. Genes Dev *18*, 1119–1130. https://doi.org/10.1101/gad.292104.
92. Larkin, J.D., Papantonis, A., Cook, P.R., and Marenduzzo, D. (2013). Space exploration by the promoter of a long human gene during one transcription cycle. Nucleic Acids Res *41*, 2216–2227. https://doi.org/10.1093/nar/gks1441.
93. Zidovska, A. (2020). Chromatin: Liquid or Solid? Cell *183*, 1737–1739. https://doi.org/10.1016/j.cell.2020.11.044.
94. Hertzog, M., and Erdel, F. (2023). The Material Properties of the Cell Nucleus: A Matter of Scale. Cells *12*. https://doi.org/10.3390/cells12151958.
95. Eshghi, I., Eaton, J.A., and Zidovska, A. (2021). Interphase Chromatin Undergoes a Local Sol-Gel Transition upon Cell Differentiation. Phys Rev Lett *126*, 228101. https://doi.org/10.1103/PhysRevLett.126.228101.
96. Boteva, L., Nozawa, R.S., Naughton, C., Samejima, K., Earnshaw, W.C., and Gilbert, N. (2020). Common Fragile Sites Are Characterized by Faulty Condensin Loading after Replication Stress. Cell Rep *32*, 108177. https://doi.org/10.1016/j.celrep.2020.108177.
97. Naughton, C., Avlonitis, N., Corless, S., Prendergast, J.G., Mati, I.K., Eijk, P.P., Cockroft, S.L., Bradley, M., Ylstra, B., and Gilbert, N. (2013). Transcription forms and remodels supercoiling domains unfolding large-scale chromatin structures. Nat Struct Mol Biol *20*, 387–395. https://doi.org/10.1038/nsmb.2509.
98. Katoh, K., and Standley, D.M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol *30*, 772–780. https://doi.org/10.1093/molbev/mst010.
99. Procter, J.B., Carstairs, G.M., Soares, B., Mourão, K., Ofoegbu, T.C., Barton, D., Lui, L., Menard, A., Sherstnev, N., Roldan-Martinez, D., et al. (2021). Alignment of Biological Sequences with Jalview. Methods Mol Biol *2231*, 203–224. https://doi.org/10.1007/978-1-0716-1036-7_13.
100. Baker, N.A., Sept, D., Joseph, S., Holst, M.J., and McCammon, J.A. (2001). Electrostatics of nanosystems: application to microtubules and the ribosome. Proc Natl Acad Sci U S A *98*, 10037–10041. https://doi.org/10.1073/pnas.181342398.

101. Altenhoff, A.M., Train, C.-M., Gilbert, K.J., Mediratta, I., Mendes de Farias, T., Moi, D., Nevers, Y., Radoykova, H.-S., Rossier, V., Warwick Vesztrocy, A., et al. (2021). OMA orthology in 2021: website overhaul, conserved isoforms, ancestral gene order and more. Nucleic Acids Res *49*, D373–D379. https://doi.org/10.1093/nar/gkaa1007.
102. Yariv, B., Yariv, E., Kessel, A., Masrati, G., Chorin, A. Ben, Martz, E., Mayrose, I., Pupko, T., and Ben-Tal, N. (2023). Using evolutionary data to make sense of macromolecules with a "face-lifted" ConSurf. Protein Sci *32*, e4582. https://doi.org/10.1002/pro.4582.
103. Karczewski, K.J., Francioli, L.C., Tiao, G., Cummings, B.B., Alföldi, J., Wang, Q., Collins, R.L., Laricchia, K.M., Ganna, A., Birnbaum, D.P., et al. (2020). The mutational constraint spectrum quantified from variation in 141,456 humans. Nature *581*, 434–443. https://doi.org/10.1038/s41586-020-2308-7.
104. Deák, G., and Cook, A.G. (2022). Missense Variants Reveal Functional Insights Into the Human ARID Family of Gene Regulators. J Mol Biol *434*, 167529. https://doi.org/10.1016/j.jmb.2022.167529.
105. Turner, T. (2013). Plot protein: visualization of mutations. J Clin Bioinforma *3*, 14. https://doi.org/10.1186/2043-9113-3-14.
106. Spitzer, J., Hafner, M., Landthaler, M., Ascano, M., Farazi, T., Wardle, G., Nusbaum, J., Khorshid, M., Burger, L., Zavolan, M., et al. (2014). PAR-CLIP (Photoactivatable Ribonucleoside-Enhanced Crosslinking and Immunoprecipitation): a step-by-step protocol to the transcriptome-wide identification of binding sites of RNA-binding proteins. Methods Enzymol *539*, 113–161. https://doi.org/10.1016/B978-0-12-420120-0.00008-6.
107. McSwiggen, D.T., Hansen, A.S., Teves, S.S., Marie-Nelly, H., Hao, Y., Heckert, A.B., Umemoto, K.K., Dugast-Darzacq, C., Tjian, R., and Darzacq, X. (2019). Evidence for DNA-mediated nuclear compartmentalization distinct from phase separation. Elife *8*, e47098. https://doi.org/10.7554/eLife.47098.
108. Huang, B., Wang, W., Bates, M., and Zhuang, X. (2008). Three-dimensional super-resolution imaging by stochastic optical reconstruction microscopy. Science *319*, 810–813. https://doi.org/10.1126/science.1153529.
109. Ovesný, M., Křížek, P., Borkovec, J., Svindrych, Z., and Hagen, G.M. (2014). ThunderSTORM: a comprehensive ImageJ plug-in for PALM and STORM data analysis and super-resolution imaging. Bioinformatics *30*, 2389–2390. https://doi.org/10.1093/bioinformatics/btu202.
110. Thévenaz, P., Ruttimann, U.E., and Unser, M. (1998). A pyramid approach to subpixel registration based on intensity. IEEE Trans Image Process *7*, 27–41. https://doi.org/10.1109/83.650848.
111. Perevoshchikova, I. V, Kotova, E.A., and Antonenko, Y.N. (2011). Fluorescence correlation spectroscopy in biology, chemistry, and medicine. Biochemistry (Mosc) *76*, 497–516. https://doi.org/10.1134/S0006297911050014.
112. Turriziani, B., Garcia-Munoz, A., Pilkington, R., Raso, C., Kolch, W., and von Kriegsheim, A. (2014). On-beads digestion in conjunction with data-dependent mass spectrometry: a shortcut to quantitative and dynamic interaction proteomics. Biology (Basel) *3*, 320–332. https://doi.org/10.3390/biology3020320.
113. Cox, J., Hein, M.Y., Luber, C.A., Paron, I., Nagaraj, N., and Mann, M. (2014). Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. Mol Cell Proteomics *13*, 2513–2526. https://doi.org/10.1074/mcp.M113.031591.
114. Baum, M., Erdel, F., Wachsmuth, M., and Rippe, K. (2014). Retrieving the intracellular topology from multi-scale protein mobility mapping in living cells. Nat Commun *5*, 4494. https://doi.org/10.1038/ncomms5494.
115. Plimpton, S. (1995). Fast Parallel Algorithms for Short-Range Molecular-Dynamics. J Comput Phys *117*, 1–19.