This is a repository copy of *Autophagy receptor NDP52 alters DNA conformation to modulate RNA polymerase II transcription*.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/197927/

Version: Published Version

Article

Autophagy receptor NDP52 alters DNA conformation to modulate RNA polymerase II transcription

Check for updates

Ália dos Santos [1,6], Daniel E. Rollins[2], Yukti Hari-Gupta[3,7], Hannah McArthur [3], Mingxue Du[2], Sabrina Yong Zi Ru[3], Kseniia Pidlisna[3], Ane Stranger[3], Faeeza Lorgat[1], Danielle Lambert [1], Ian Brown [3], Kevin Howland [3], Jesse Aaron [4], Lin Wang [5], Peter J. I. Ellis [3], Teng-Leong Chew [4], Marisa Martin-Fernandez [5], Alice L. B. Pyne [2] & Christopher P. Toseland [1] ✉

NDP52 is an autophagy receptor involved in the recognition and degradation of invading pathogens and damaged organelles. Although NDP52 was first identified in the nucleus and is expressed throughout the cell, to date, there is no clear nuclear functions for NDP52. Here, we use a multidisciplinary approach to characterise the biochemical properties and nuclear roles of NDP52. We find that NDP52 clusters with RNA Polymerase II (RNAPII) at transcription initiation sites and that its overexpression promotes the formation of additional transcriptional clusters. We also show that depletion of NDP52 impacts overall gene expression levels in two model mammalian cells, and that transcription inhibition affects the spatial organisation and molecular dynamics of NDP52 in the nucleus. This directly links NDP52 to a role in RNAPII-dependent transcription. Furthermore, we also show that NDP52 binds specifically and with high affinity to double-stranded DNA (dsDNA) and that this interaction leads to changes in DNA structure in vitro. This, together with our proteomics data indicating enrichment for interactions with nucleosome remodelling proteins and DNA structure regulators, suggests a possible function for NDP52 in chromatin regulation. Overall, here we uncover nuclear roles for NDP52 in gene expression and DNA structure regulation.

NDP52/CALCOCO2, a 446 amino-acid autophagy receptor, was first identified in the nucleus, as a component of nuclear dots—multiprotein sub-compartments that respond to environmental stresses, such as viral infections[1]. However, later reports showed that the protein is distributed throughout the cell, with higher levels in the cytoplasm[2]. NDP52 has since been linked to cytoplasmic roles in autophagy and cell adhesion, where it is known to be required for pathogen-containing autophagosome maturation and membrane ruffle formation[3–5]; however, no nuclear function has been attributed to this protein.

NDP52 comprises a skeletal muscle and kidney enriched inositol phosphatase (SKIP) carboxyl homology (SKICH) domain, which facilitates membrane localisation[3]; a long coiled-coil (CC) region that includes a predicted leucine-zipper (LZ) domain, and two zinc finger domains at the C-terminal—ZF1 and ZF2 (Fig. 1A)[1]. The CC region of

[1]Department of Oncology and Metabolism, University of Sheffield, Sheffield S10 2RX, UK. [2]Department of Materials Science and Engineering, University of Sheffield, Sheffield S1 3JD, UK. [3]School of Biosciences, University of Kent, Canterbury CT2 7NJ, UK. [4]Advanced Imaging Center, HHMI Janelia Research Campus, Ashburn, VA 20147, USA. [5]Central Laser Facility, Research Complex at Harwell, Science and Technology Facilities Council, Rutherford Appleton Laboratory, Harwell, Didcot, Oxford OX11 0QX, UK. [6]Present address: MRC LMB, Francis Crick Avenue, Cambridge CB2 0QH, UK. [7]Present address: MRC LMCB, University College London, Gower Street, London WC1E 6BT, UK. ✉e-mail: c.toseland@sheffield.ac.uk
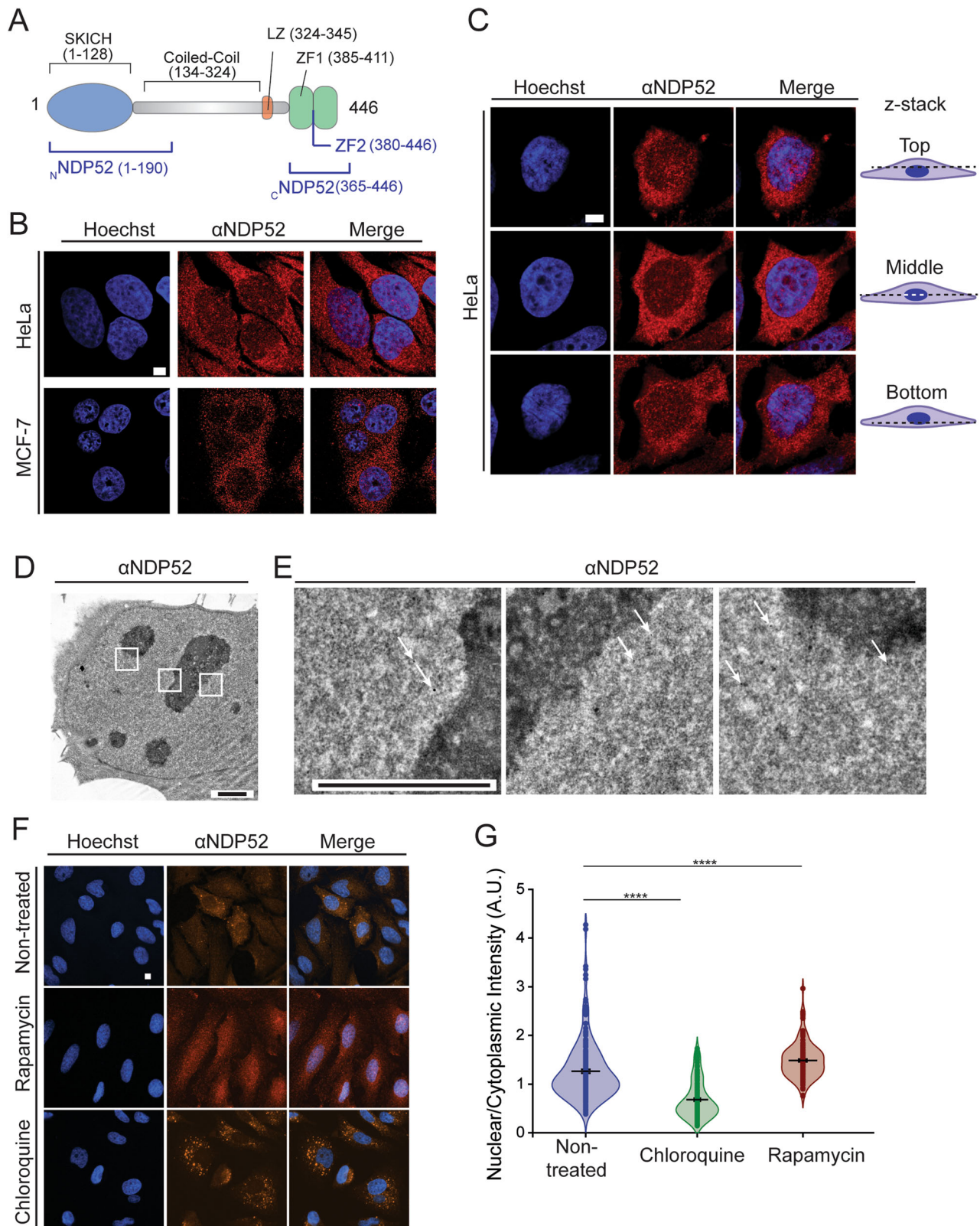
NDP52 has been identified as a potential homo-dimerisation domain for the protein[6]. At the C-terminal, ZF1 has been characterised as an unconventional dynamic zinc finger, whilst ZF2 is a canonical C_2H_2-type zinc finger[7]. The C-terminal domains of NDP52 are responsible for interactions with ubiquitin, which allows binding to ubiquitylated pathogens, as well as interactions with actin-based motor Myosin VI (MVI)[3,7–9]. In the cytoplasm, interactions between NDP52 and MVI allow

autophagosome maturation[5]. However, there is little information available regarding the biochemical and structural properties of the full-length protein, which limits our understanding of its functions.

NDP52 is a member of the Calcium-binding and coiled-coil domain containing (CALCOCO) family. Other members are TAX1BP1 and CoCoA. NDP52 shares high sequence homology with both TAX1BP1 and CoCoA, and all three proteins have similar domain

Fig. 1 | NDP52 is distributed throughout the nucleus. A Diagram of NDP52 displaying protein domains and key features, as well as recombinant constructs used in this study (in blue). **B** Confocal imaging of HeLa and MCF-7 cells labelled by immunofluorescence against NDP52 (red), with DNA staining shown in blue. Scale bar = 5 µm. **C** Confocal imaging of the intracellular distribution of NDP52 at different z points in HeLa. Scale bar = 5 µm. **D** Electron microscopy of HeLa cells following immune-gold labelling of NDP52. Scale bar = 2 µm. Approximate thickness = 70–80 nm. White boxes indicate enhanced sections shown in (**E**). Electron microscopy detail images of immuno-gold labelled NDP52 (black dots indicated by white arrows) in the nucleus. Scale bar = 500 nm. **F** Immunofluorescence staining of NDP52 in HeLa cells following treatment with chloroquine and rapamycin to inhibit and induce autophagy, respectively. Scale bar = 5 µm. **G** Quantification of nuclear:cytoplasmic fluorescence intensity for NDP52 following inhibition or stimulation of autophagy. AU arbitrary units. Line represents mean, $N = 50$ cells. ****$p < 0.0001$ by two-tailed t-test. Raw data are available in the Source Data file.

structure. Interestingly, whilst TAX1BP1 is also a known autophagy receptor[3,10], CoCoA is a well-characterised transcription coactivator[11]. Recently, CoCoA has also been linked to roles in autophagy, further highlighting potential functional similarities between these proteins[12,13]. Moreover, NDP52 binding partner MVI, has important nuclear roles in transcription, DNA repair and chromatin regulation[14]. Furthermore, a recent study by Fili et al. has revealed that the interaction between NDP52 and MVI enhances RNA polymerase II (RNAPII) transcriptional activity in vitro[15,16].

Here, we explore the spatial organisation of NDP52 in the nucleus, as well as its dynamic behaviour, and assess how perturbation of this protein affects gene expression in cells. We find that NDP52 forms clusters in the nucleus at RNAPII transcription initiation sites and that knockdown of NDP52 significantly affects gene expression in both HeLa and MCF-7 cells. Furthermore, our biochemical analysis shows that NDP52 binds to double-stranded DNA (dsDNA) with high affinity and atomic force microscopy (AFM) suggests this results in changes to DNA shape and structure in vitro. We have also explored the nuclear interactome of NDP52, which shows enrichment for proteins involved in DNA structure and nucleosome regulation. We suggest that NDP52 has a regulatory role in RNAPII-dependent transcription, and that this arises both from direct interactions with chromatin as well as from protein-protein interactions with chromatin regulators and transcription factors. Overall, this highlights a wider role of NDP52 across the cell and it remains to be determined if there are links between its cytoplasmic and nuclear functions.

Results

Nuclear organisation and dynamics of NDP52

To attribute a nuclear function to NDP52, we first assessed its nuclear localisation in two example mammalian cancer cell lines. Immunofluorescence staining of NDP52 in both HeLa and MCF-7 cells shows that NDP52 is distributed throughout the cytoplasm and nucleus (Fig. 1B). Confocal imaging of different focal planes also shows distribution of the protein throughout the organelle (Fig. 1C). To further confirm this, we also used electron microscopy with gold-immunolabelling of endogenous NDP52. Imaging of negative stained HeLa sections (c.a. 70 nm thickness) (Fig. 1D) clearly shows NDP52 particles in nuclear regions, which can be observed in zoomed-in sections in Fig. 1E. The presence of NDP52 in the nucleus is consistent with previous reports[1,2,15].

NDP52 is a well-established autophagy receptor[3–5]. To explore how induction or inhibition of autophagy impacts the nuclear localisation of NDP52, we used rapamycin and chloroquine, respectively (Fig. 1F). The induction of autophagy led to an increase in the nuclear:cytoplasmic ratio of NDP52 in HeLa cells, whilst autophagy inhibition led to a decrease (Fig. 1G and Supplementary Fig. 1). This suggests a link between the cytoplasmic autophagy functions of NDP52 and a role within the nucleus.

Within the nuclear region, NDP52 appears to cluster into small punctate regions of high fluorescence intensity (Fig. 2A). To explore the nuclear organisation of NDP52, we used Stochastic Optical Reconstruction Microscopy (STORM) in both HeLa and MCF-7 cells (Fig. 2B). STORM allows us to visualise with high spatial precision and quantify individual molecules of NDP52 within a specified region of interest (ROI) (Fig. 2C), in this case the nuclear region. Furthermore, in-depth analysis of STORM data can also provide information regarding the clustering behaviour of the protein[17]. Protein clustering is often related to the molecular function of a protein and is particularly important in the enhancement of enzymatic processes such as transcription, DNA repair and DNA replication[18–22]. Hence, as we investigate a nuclear function for NDP52, it is important to study its spatial organisation and how this might be linked to its nuclear role. To determine if NDP52 is randomly distributed or forms clusters (Fig. 2D), we used a linearised Ripley's K function[23]. In both cell lines we observe a high probability for nuclear clustering of NDP52, as the Ripley's K function deviates from zero towards positive values (Fig. 2E, F). To further understand the organisation of NDP52 clusters in the nucleus, we used ClusDoC software[23]. We defined NDP52 clusters, as regions where a minimum of 5 neighbouring molecules are spaced at a distance smaller than the mean value of localisation precision from STORM acquisition (described Methods). This allowed us to generate cluster maps for selected nuclear regions (Fig. 2G) and determine that ~45% (±6) and 79% (±3) of NDP52 molecules are clustered in HeLa and MCF-7 cells, respectively. This corresponds to an average of 1604 (±307) and 2285 (±478) clusters per cell in HeLa and MCF-7 cells, respectively, with an average size of 3513 nm² (±249) and 6190 nm² (±890), and 44 (±6) and 109 (±64) molecules of NDP52 per cluster (Fig. 2H–K). In both HeLa and MCF-7 cells, we observe large cell-to-cell variation for clustering data. Although STORM provides detailed information regarding the spatial organisation of molecules, it is also a low-throughput technique. Cell variability could be a result of cells not being synchronised; however, due to this limited throughput, it is also not possible to identify multiple subpopulations within the data.

To assess how the spatial distribution of NDP52 relates to its molecular dynamics in the nucleus, we transiently expressed Halo-NDP52 in HeLa cells. This allowed us to use fluorescence recovery after photobleaching (FRAP) to assess how dynamic NDP52 molecules are in the nucleus of live-cells (Fig. 3A). Our data show that NDP52 has a recovery half-time of 7.5 s (±0.8) (Fig. 3B and Supplementary Fig. 2A) and a mobile fraction of 0.65 (±0.02) (Fig. 3C). This is in agreement with molecular clustering data showing that ~45% of NDP52 molecules are clustered, and would therefore be expected to be less dynamic. To obtain more detailed information on the dynamic behaviour of nuclear NDP52, we used aberration-corrected multi-focal microscopy (acMFM). This technique allows us to simultaneously track single-molecules across nine focal planes in live-cells, covering 4 µm in the z axis and 20 × 20 µm in xy (Fig. 3D). We obtained 3D trajectories of Halo-NDP52 molecules in the nucleus of HeLa cells (Fig. 3E). Analysis of the different trajectories can then provide information on how confined or diffuse molecules are (Fig. 3F). By measuring the mean squared displacement (MSD) of each molecule (Supplementary Fig. 2B), we were also able to calculate diffusion coefficients (Fig. 3G) and anomalous diffusion constants (α) (Supplementary Fig. 2C) for each track. These were then plotted as the average diffusion coefficient or average α per cell (Fig. 3H, I). Under normal conditions, NDP52 nuclear diffusion is relatively slow ($D = 0.24$ (±0.008) µm²/s) and molecules are mostly confined, displaying an α value lower than 1 ($\alpha = 0.7$ (±0.006))
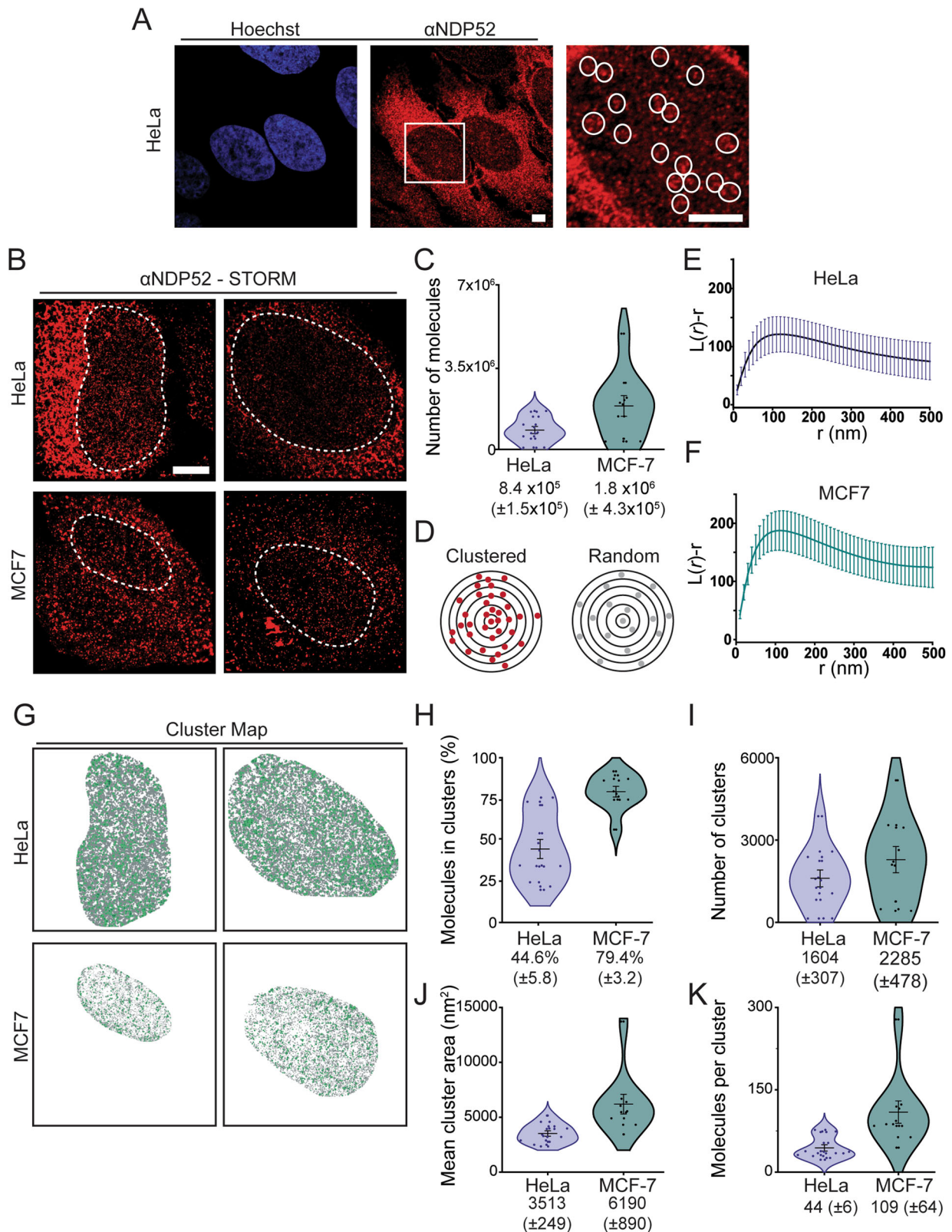
(Fig. 3H, I). Furthermore, acMFM data show that ~55% (±0.84) of nuclear NDP52 molecules are static ($D < 0.1\,\mu m^2/s$), which closely relates to the estimated percentage of clustered molecules calculated from STORM data.

Overall, the clustering behaviour and confined dynamics of NDP52 molecules in the nucleus support our hypothesis of a nuclear function for this protein.

NDP52 oligomerisation and structure

Having investigated the spatial organisation and molecular dynamics of nuclear NDP52, which suggests a nuclear function, we wanted to investigate the biochemical properties of the protein to understand its potential roles. For this, we used different recombinant NDP52 constructs, including the full-length protein (NDP52-FL), an N-terminal truncated region (NNDP52), which includes the SKICH domain and part

Fig. 2 | Spatial organisation of NDP52 in the nucleus. A Confocal image of NDP52 in HeLa cells showing detail of dense nuclear staining in white circles (zoomed-in right panel). Scale bar = 5 µm **B** Example STORM images of NDP52 in HeLa and MCF-7 cells. Dotted lines represent selected regions of interest (ROIs) for the nucleus. These regions were used for cluster analysis. Scale bar = 5 µm. **C** Total number of molecules detected within ROI $n = 12$ (HeLa), $n = 10$ (MCF-7). Errors bars are mean ± SEM. **D** Diagram depicting molecular clustering and random distribution. **E**, **F** Linearised Ripley's K Function, $L(r) - r$ (where r is the radius), calculated for selected ROIs from STORM images in HeLa and MCF-7. A value of zero in this plot signifies molecules are randomly distributed with respect to each other (as shown in **D**), whilst positive values indicate molecular clustering. Mean values are plotted ± SEM. $n = 12$ (HeLa), $n = 10$ (MCF-7). **G** Cluster maps generated for ROIs displayed in (**B**), using parameters specified in Methods. Clustered molecules are shown in green. Cluster analysis of NDP52 in the nucleus of HeLa and MCF-7 showing: **H** percentage of molecules in clusters; **I** number of clusters in ROIs; **J** mean cluster area in nm² and **K** number of molecules per clusters. Mean ± SEM values are shown. $n = 12$ (HeLa), $n = 10$ (MCF-7). Raw data are available in the Source Data file.

of the coiled-coil region (amino acid residues 1–190), a C-terminal region ($_C$NDP52), which includes a section of the coiled-coil and both zinc finger domains (amino acid residues 365–446). Lastly, the two zinc finger domains (ZFs−amino acid residues 380–446) (Fig. 4A and Supplementary Fig. 3). All the recombinant proteins presented stable secondary structure, as shown by circular dichroism and/or nano-differential scanning fluorimetry (nano-DSF) (Supplementary Fig. 4A–K). The individual first zinc finger domains of NDP52 were not selected for biochemical studies, as they lacked a stable secondary structure. This is in agreement with previous structural reports for this domain[7].

Previous work showed that full-length NDP52 is mainly a dimer in solution[6,8]. To confirm this, we used Size-Exclusion Chromatography with Multi-Angle Light Scattering (SEC-MALS). SEC-MALS allows us to obtain accurate molecular weight information from gel filtration elution profiles and to identify different oligomeric species in solution. Our SEC-MALS data show that the majority of NDP52 is present in the dimeric form (second peak average molecular weight = 117 kDa), but it also shows the presence of higher oligomeric forms, such as trimers and tetramers (first elution peak with an average molecular weight of 333 kDa) (Fig. 4B). Through titrations of RED-tris-NTA labelled NDP52-FL with unlabelled NDP52-FL, microscale thermophoresis (MST) confirms oligomerisation of the full-length protein, with an estimated K_D value of 0.21 µM (±0.006) (Fig. 4C) which supports the presence of a minimal dimer complex at the concentrations used in the SEC-MALS experiments. Mass photometry data enables mass determination at the single-molecule level. At 100 nM NDP52, we observed a mostly dimeric state, with a small population of trimers also present (Fig. 4D). This is consistent with the MST and SEC-MALS analysis and we conclude that NDP52 readily oligomerizes.

From the SEC profile, NDP52 appears to be an elongated protein, eluting at a much earlier elution volume than expected for a globular protein. The estimated radii of gyration from SEC-MALS data are 11.7 nm, for the second peak (corresponding to NDP52 dimers) and 14.8 nm for the higher-oligomeric forms (first elution peak) of NDP52 (Fig. 4E). This translates into an approximate end-to-end measurement of 23 nm for dimeric NDP52. To directly measure NDP52 particle size, we used dynamic light-scattering (DLS) which showed that NDP52-FL particles can be measured at a range of diameters between 10–43 nm, with a maximum at 15.5 nm (Fig. 4F). To obtain more information regarding the overall shape of NDP52-FL, we used small-angle X-ray scattering (SAXS). SEC-SAXS is a robust technique for the study of macromolecule conformation in solution. Our SAXS data estimate a radius of gyration for NDP52 between 9–15 nm (end-to-end value 18–30 nm) (Fig. 4G), consistent with a predicted rod-shape structure for the protein. Variability in radii measurements for NDP52 could be a direct consequence of its elongated shape, as measurements for different profiles of the protein will be more varied than in a globular protein. Using SAXS we were also able to generate an envelope model for NDP52-FL, showing the predicted elongated shape (Fig. 4H, I).

To directly visualise and measure protein shape and size, we used atomic force microscopy (AFM) imaging. In agreement with the biochemical data (Fig. 4B–H), AFM imaging of NDP52-FL shows a distribution of proteins of elongated shape (Fig. 4J). Moreover, we counted 15% of NDP52-FL molecules in an oligomeric state. The resolution of AFM imaging allows visualisation of the different domains of the NDP52-FL protein, with the larger SKICH domain and smaller C-terminal region distinguishable by height (Supplementary Fig. 5) and linked together by a thin, flexible linker (Fig. 4K). The length of the protein (maximum bounding size) has a wide distribution with a clear peak at 20 ± 12 nm (Fig. 4L), as expected from the SAXS data. This variability is driven by the thin, coiled-coil, flexible linker which can adopt a variety of conformations, allowing the protein to bend, and leading to variability in the protein length.

In agreement with the observed elongated shape for the protein, the width (minimum bounding size) of NDP52-FL was significantly less than the length with a peak at 13 ± 6 nm (Fig. 4L). The widths of NDP52-FL also occupy a narrower distribution compared to the lengths since the coiled-coil only allows for flexibility along the length of the protein (Fig. 4K). It is therefore likely that the width of NDP52-FL corresponds to the diameter of the globular domains at both ends (Fig. 4K). To probe this hypothesis, we measured the dimensions of a truncated version of the protein, $_C$NDP52. AFM imaging (Supplementary Fig. 6A–C) showed that the minimum and maximum bounding sizes for $_C$NDP52 largely overlap, with the peak in the probability distributions occurring at values of 13 ± 6 nm and 9 ± 3 nm respectively, indicating relatively globular conformations (Supplementary Fig. 6B). These measurements closely match the minimum bounding size of NDP52-FL (Supplementary Fig. 6C), showing that the width of NDP52-FL is determined by the size of its globular domains.

Although α-helical coiled-coil domains are often drivers of protein oligomerisation[24,25], AFM imaging also showed the protein's terminal domains acting as the interface for dimerisation of the protein (Fig. 4M). Dimers were observed as even longer elongated molecules, with two smaller globular domains linked by two thin linkers to one central globular domain, most likely formed of two terminal regions. To test which regions of NDP52 are capable of oligomerising, we used different truncated regions of the protein (Fig. 4A). SEC-MALS of $_C$NDP52 shows that this region is mostly present in dimeric and monomeric forms, although trimers could also be detected (Supplementary Fig. 6D). MST data confirms oligomerisation of this domain, generating a K_D of 0.05 µM (±0.006) (Supplementary Fig. 6E). These observations were confirmed by AFM imaging where we could identify monomeric, dimeric and trimeric forms of $_C$NDP52 (Supplementary Fig. 6A). Oligomerisation of $_C$NDP52, which lacks the presence of the coiled-coil region, suggests that these domains might also be important for interactions between monomers during dimerisation of the full-length protein.

We also investigated the ability of the ZFs domain to oligomerise, using SEC-MALS and MST. Our data show that this domain can also homo-oligomerise, presenting itself as a monomer, dimer and trimer in solution, with an oligomerisation K_D of 0.18 µM (±0.017) (Supplementary Fig. 6F, G). When testing oligomerisation of the N-terminal region of NDP52, $_N$NDP52, containing the SKICH domain and part of the coiled-coil region, we observe a clearer preference for the dimeric form (Supplementary Fig. 6H, I). Interestingly, we could also observe an interaction between $_C$NDP52 and $_N$NDP52 (Supplementary Fig. 6J). It is possible that these two opposing regions interact in the full-length protein, due to the presence of a relatively flexible central coiled-coil region, or between homo-oligomers of NDP52.
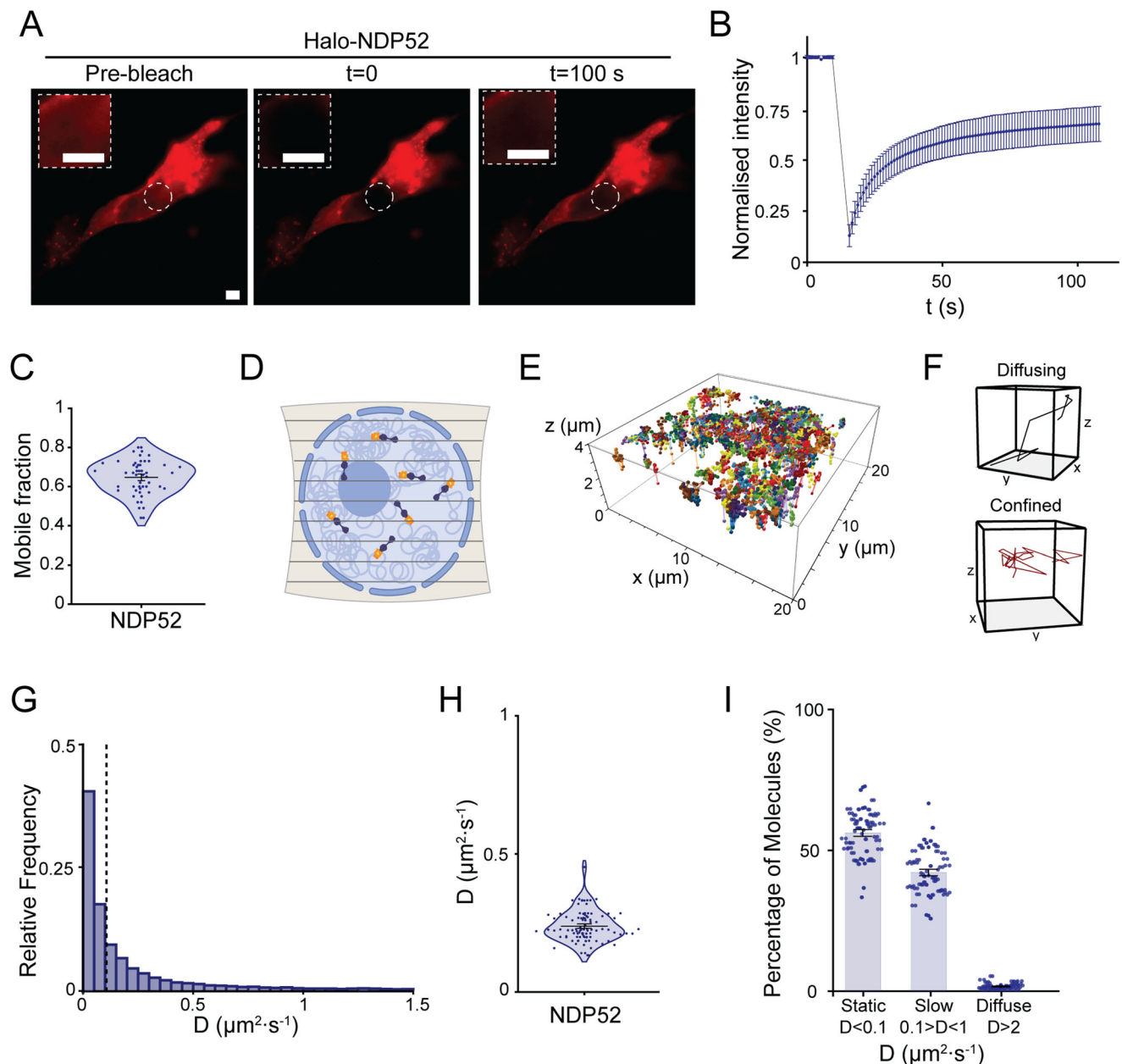
Fig. 3 | Molecular dynamics of NDP52 in the nucleus. A Example of fluorescence recovery after photobleaching (FRAP) image acquired in HeLa cells transiently expressing Halo-NDP52. Insets display zoomed-in detail of nuclear area selected for photobleaching. Scale bar = 5 μm. **B** Normalised fluorescence intensity profile, in function of time for FRAP experiments. Estimated value of fluorescence recovery $t_{1/2}$ is shown on the graph. Mean values ± SEM are shown. $n = 27$ cells. **C** Calculated mobile fraction from FRAP data. Mean values ± SEM are shown. $n = 27$ cells. **D** Diagram depicting simultaneous acquisition of nine focal planes (covering 4 μm in z and 20 μm × 20 μm in xy) for 3D single-molecule tracking of Halo-NDP52 in the nucleus. Created with BioRender.com. **E** Example of 3D reconstructed trajectories for a single nucleus over time. **F** Example of diffusive and confined trajectories over time. **G** Histogram of diffusion constants from the nucleus of HeLa cells transiently expressing Halo-NDP52. Dotted lines represent the applied threshold to differentiate between static and dynamic molecules (14,322 molecules from 51 cells). **H** Diffusion coefficient values for Halo-NDP52. Each data point represents the mean diffusion coefficient for a cell ($n = 51$ cells). Errors bars are mean ± SEM. **I** Percentage of molecules considered static ($D < 0.1\,\mu m^2/s$), slow moving ($0.1 < D < 1\,\mu m^2/s$) or diffuse ($D > 1\,\mu m^2/s$) per Cell. $n = 51$ cells. Errors bars are mean ± SEM. Raw data are available in the Source Data file.

As previously mentioned, NDP52 shares high sequence identity with its family member CoCoA—a protein with known nuclear functions in transcription co-activation[6]. However, very little is known regarding the oligomeric states of CoCoA, or how this may align with NDP52. To test if recombinant CoCoA (Fig. 4N) can also form dimers, we used SEC-MALS. Our data show that the main peak for CoCoA is a complex mixture of molecular weights, ranging from 148 kDa (equivalent to the dimeric form of CoCoA) to 300 kDa (Fig. 4O). Using MST, we further confirmed the ability of CoCoA to oligomerise, with a

calculated K_D of 0.085 μM (±0.002) (Fig. 4P). Essentially, CoCoA and NDP52 display similar biochemical properties.

NDP52 binds and oligomerises with double-stranded DNA

Having determined the oligomeric state of NDP52-FL and clarified its nuclear localisation, we decided to test NDP52 binding to DNA. Previously, Fili et al. showed that NDP52 can bind to double-stranded DNA (dsDNA) with high-affinity[15]. As we have established that NDP52 is present in confined clusters within the nucleus, DNA binding could
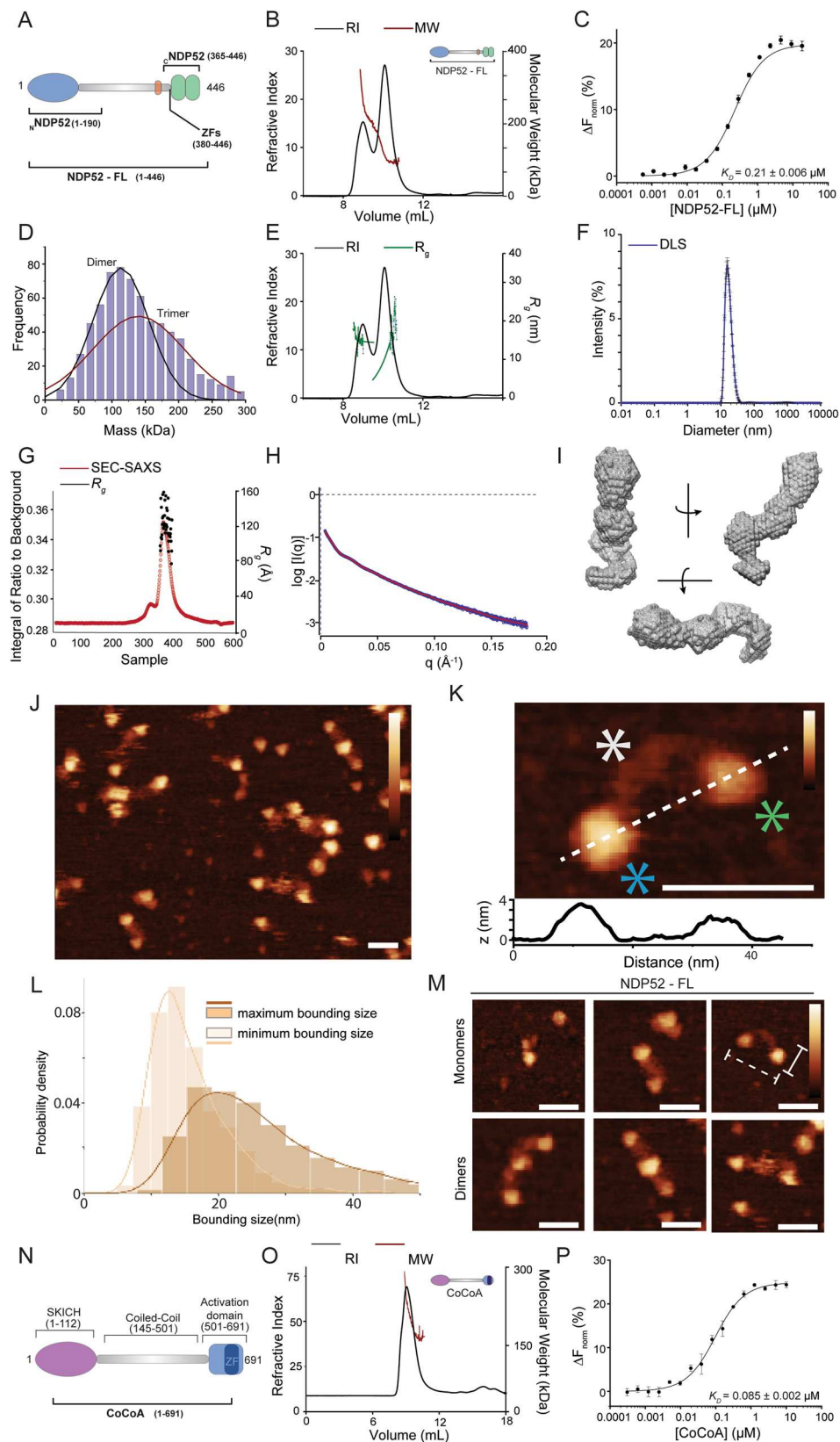
be an essential part of its nuclear role. Hence, we used an electro-phoretic mobility shift assay (EMSA) to investigate the formation of NDP52-dsDNA complexes. Using dsDNA 40 long (ds40) we show that NDP52 can form complexes with dsDNA, in vitro, evident by the formation of a higher band in the EMSA (Fig. 5A). To explore this interaction in a quantitative manner, we used fluorescence

spectroscopy. For this, NDP52-FL was titrated against two different lengths of FITC-labelled DNA−40 and 15 bp (ds40 and ds15, respec-tively). Our data confirmed a high-affinity interaction between NDP52 and DNA, with K_D values <100 nM for both DNA lengths (Fig. 5B).

To directly visualise this interaction, we employed AFM imaging. We used linearised dsDNA 339 bp long (ds339)−~115 nm long to

Fig. 4 | Oligomerisation and structure of NDP52. A Diagram of NDP52 showing recombinant constructs. **B** SEC-MALS profile for NDP52-FL. Refractive index (RI) trace is shown in black, as well as the calculated molecular weight (MW) values, across the peaks (in red). **C** Microscale thermophoresis, showing oligomerisation of NDP52. Calculated K_D is shown. **D** Histogram showing calculated mass of NDP52-FL from mass photometry. Mean values of Gaussian curves closely correspond to dimeric and trimeric molecular weights of NDP52 (Gaussian max values = 105 kDa and 140 kDa, respectively). **E** Radius of gyration (R_g) calculated from SEC-MALS data shown in (**B**). Estimated R_g for peak one was 14.84 nm and 11.7 nm for the second peak. RI trace shown again for NDP52-FL in black, and R_g across peaks shown in green. **F** DLS trace for NDP52-FL showing calculated diameter for the protein, with values between 10–43 nm and maximum at 15.5 nm. **G** SEC-SAXS for NDP52-FL showing R_g values across peaks. Radii values between 9–15 nm. **H** Experimental SAXS curve for NDP52-FL. **I** Bead model of NDP52-FL from SEC-SAXS data. **J** AFM image of NDP52-FL showing multiple molecules on PLL. For all AFM images, Scale bar = 25 nm. Height scale = 4.5 nm. **K** High-resolution AFM image

of an individual NDP52-FL molecule on PLL. Protein domains are indicated by asterisks, with SKICH in blue, coiled coil in grey and c-terminal domain in green. Below, a line profile taken from the AFM image of the NDP52-FL protein along the white dotted line from left to right is shown. **L** Histogram and kernel density estimate (KDE) plots for maximum and minimum bounding size measured from Topostats (Supplementary Fig. 6K). Peaks in KDE plots were used to determine particle size (KDE max ± SD) minimum = 13 ± 6 nm, maximum = 20 ± 12. N = 1365 particles. **M** AFM images of NDP52-FL showing the protein in monomeric and dimeric forms. The white lines indicate the maximum (dotted line) and minimum (solid line) bounding sizes. **N** Diagram depicting CoCoA domains and key features. **O** SEC-MALS trace for CoCoA. RI trace is shown in black and calculated molecular weight values are shown in red (MW values 140–300 kDa). **P** Microscale thermophoresis, showing oligomerisation of CoCoA. For all experiments, values represent average ± SEM of three individual experiments. Raw data are available in the Source Data file.

observe direct interactions between NDP52-FL and DNA (Fig. 5C). We can observe direct interactions between NDP52-FL and ds339. Furthermore, we can also observe that more than one molecule of NDP52-FL can interact with DNA (Fig. 5C). This agrees with mass photometry measurements that show that when incubated with ds40, the measured mass for NDP52-FL increases from its dimer/trimer values (112 and 157 kDa, calculated for NDP52-FL alone) to 1334 kDa (Fig. 5D). Similar to NDP52, when testing CoCoA for dsDNA binding, we also observe high-affinity interactions in fluorescence spectroscopy assays (Fig. 5E).

Zinc finger domains are well-known for their ability to bind DNA[26]. Therefore, we used the $_C$NDP52 and ZFs recombinant constructs in fluorescence spectroscopy assays. As expected, both constructs interact tightly with dsDNA, with K_D values in the low nM range (Fig. 5F, H). However, DNA binding curves for these constructs do not reach saturation at higher concentrations of protein, as they do for NDP52-FL. This could be explained by the clustering behaviour of $_C$NDP52 around DNA, that we observe by AFM imaging of this domain complexed with ds339 (Fig. 5G). Although some degree of interaction could be detected for $_N$NDP52-dsDNA, this presents much lower affinity than other domains (Fig. 5I). Experiments with single-stranded DNA (ss40) are highly variable and could not be fitted to a binding equation (Fig. 5J). Moreover, ssDNA could not outcompete NDP52 binding ds40 (Supplementary Fig. 7A). This suggests that NDP52-FL preferentially binds dsDNA.

Since we have established that NDP52 can bind to DNA in vitro, we hypothesised that NDP52 could also interact with genomic DNA in cells. To test this, we performed chromatin immunoprecipitation (ChIP), and detected the presence of NDP52 bound to chromatin-enriched cellular fractions (Supplementary Fig. 7B). We also tested different genomic loci for the presence of NDP52, through ChIP-qPCR, including genes regulated by nuclear receptors (previously linked to coactivator functions of CoCoA)[11], and inflammation-related targets, where NDP52 has been shown to have a role[27,28]. Our ChIP-qPCR data suggest that NDP52 is present throughout the gene body of different genes (Supplementary Fig. 7C). This supports our hypothesis that NDP52 can bind DNA and this could be a mechanism through which the protein could impact gene expression.

NDP52 can alter DNA shape
Having established that NDP52-FL can interact with DNA, we then investigated if these interactions could cause local changes to DNA shape or structure (Fig. 6). Using AFM imaging, we observed several instances where NDP52 appears to be able to loop or bridge individual strands of dsDNA and bridge between molecules (Fig. 6 and Supplementary Fig. 7D). This looping leads to a compaction of DNA molecules which we quantified by measuring the maximum bounding size (area) of each molecule in the presence and absence of NDP52-FL (Fig. 6C, D).

A shift towards lower bounding size was observed in the presence of NDP52-FL (Fig. 6D). Highly compacted molecules were found at the lower end of the population (Fig. 6E). Bridging between molecules was also observed in the presence of NDP52-FL (Fig. 6F). NDP52 was not always observed bound to DNA in these configurations. We suggest this arises from the dynamic interaction between DNA and the protein along with the four washing steps used to remove DNA and protein that was not bound to the mica. Despite this, the structural changes induced in the DNA remained. Indeed, these changes were dependent on the strength of absorption between the DNA and the surface which was modulated by the surface chemistry (Supplementary Fig. 7D).

To test if NDP52-FL has a preferential binding location, i.e., at the flexible DNA ends vs. the constrained dsDNA in the middle, we divided ds339 molecules into two regions. These two regions were the edge— accounting for 50% of the DNA molecule (25% at each end) and middle —accounting for the central 50% of ds339 (Fig. 6G diagram). We observed that NDP52-FL preferentially binds at the ends of linear dsDNA, ~20–25 nm into the ds339 molecule (corresponding to 19.39 ± 0.86% of ds339 length) (Fig. 6G). This bias may be due to the extra conformational flexibility around the ends of the DNA strands and the confinement of the centre of the molecule.

Interestingly, we could also observe changes in DNA structure when incubating ds339 with $_C$NDP52, with instances of DNA bridging and looping also observed (Supplementary Fig. 7E, F).

NDP52 in involved in RNAPII-dependent transcription
As previously mentioned, one of the most well-known binding partners of NDP52 is MVI. Previous work has linked the interaction between NDP52 and MVI to the enhancement of RNA polymerase II (RNAPII) activity[15]. Furthermore, colocalising foci of NDP52 and RNAPII have been previously observed in the nucleus[15], also shown in Fig. 7A.

To further explore the role of NDP52 in RNAPII-related transcription, we used STORM. STORM not only allows us to improve colocalisation estimates between NDP52 and RNAPII molecules, relative to conventional optical microscopy, but also allows us to measure colocalisation of clusters for both proteins. Here, we used phospho-Ser5-RNAPII immunofluorescence staining, which selects for the pool of RNAPII molecules involved in transcription initiation. STORM data show that, under normal conditions, ~28.7% (±1.7) of NDP52 is colocalised with RNAPII, and 20.2% (±2.8) of RNAPII is found colocalising with NDP52 (Fig. 7B, C). Colocalisation of clusters between NDP52 and RNAPII can also be observed in ClusDoC-generated heat maps and histograms (Fig. 7D and Supplementary Fig. 8A), with nuclear regions of high colocalisation density for each channel represented in red. Interestingly, our data show that NDP52 clusters that colocalise with RNAPII clusters are ~6.5 times larger than non-colocalised clusters, and RNAPII clusters colocalised with NDP52 ~12-fold larger (Fig. 7E). Although there are more non-colocalised clusters than colocalised
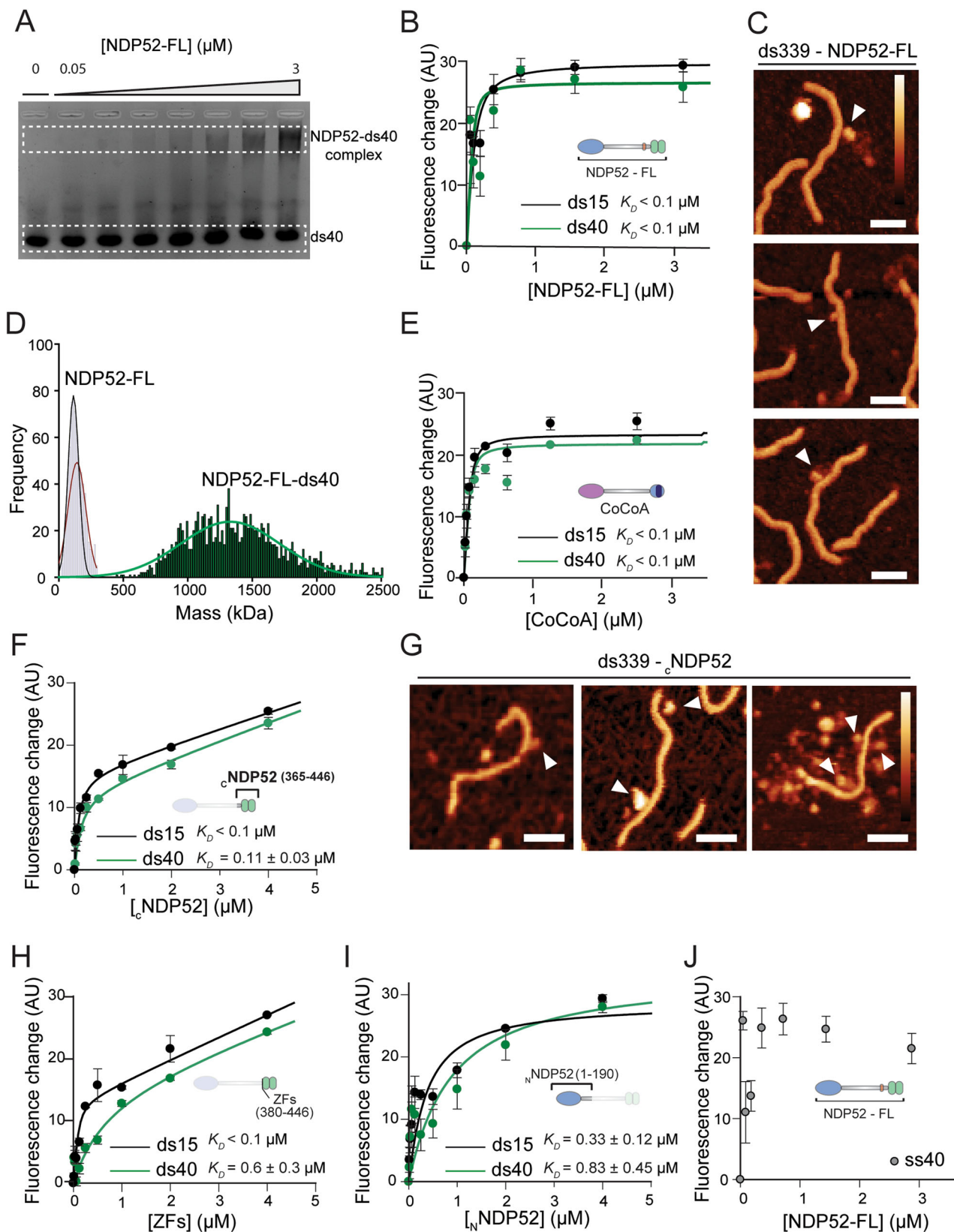
between NDP52 and RNAPII, colocalised clusters also have higher density of molecules (2.5 times higher density for NDP52 and 2 times higher for RNAPII) (Supplementary Fig. 8B–E) This further suggests a relationship between the nuclear organisation of NDP52 and transcription.

To test if NDP52 can affect the spatial organisation of RNAPII, we overexpressed the nuclear pool of NDP52. For this, we used a Halo-NLS-NDP52 construct (Supplementary Fig. 8F). We then used STORM and cluster analysis to quantify changes in the distribution of RNAPII in the nucleus (Fig. 7F). Overexpression of nuclear NDP52 did not have an

Fig. 5 | NDP52 binds to and oligomerises with double-stranded DNA through its C-terminal domain. A Electrophoretic-mobility shift assay (EMSA) for NDP52-FL with ds40. dsDNA was used at 250 nM, with increasing concentrations of NDP52-FL run in each well (0.05; 0.1; 0.2; 0.5; 1; 2 and 3 μM). Lower band represents free ds40 and top band represents DNA in complex with NDP52-FL. **B** Fluorescence spectroscopy titrations of NDP52-FL against 40 bp or 15 bp fluorescein amidite double-stranded DNA (ds40 and ds15, respectively). **C** AFM images showing direct visualisation of NDP52-FL binding to linear 339 bp DNA (ds339) adsorbed with magnesium buffer. Binding events are marked by white arrowheads. Scale bar = 25 nm. Height scale = 4.5 nm (scale bar inset in **C** and **G**). **D** Mass photometry histogram for NDP52 showing a large shift in detected mass when NDP52-FL is incubated with ds40. Histograms and Gaussian fittings for NDP52 alone are the same as the ones shown in Fig. 4D (in black and red). Histogram for NDP52-FL-ds40 (green) was also fitted to a Gaussian function. Mean value calculated as 1334 kDa). **E** Fluorescence spectroscopy titrations of CoCoA against ds40 and ds15DNA. **F** Fluorescence spectroscopy titrations of $_C$NDP52 against ds40 and ds15. Calculated K_D values are shown. **G** AFM images showing $_C$NDP52 binding and clustering around linear ds339 on PLL. Scale bar = 25 nm. Height scale = 4.5 nm (scale bar inset in **C**). **H** Fluorescence spectroscopy titrations of ZFs against ds40 and ds15 DNA. **I** Fluorescence spectroscopy titrations of $_N$NDP52 against ds40 and ds15 DNA. **J** Fluorescence spectroscopy titrations of NDP52-FL with single-stranded 40 bases DNA (ss40). For all protein-DNA fluorescent assays, DNA concentration was kept at 100 nM and K_D values represent mean ± SEM of $n = 3$ independent experiments. AU arbitrary units. Data fitting was performed as described in "Methods". Raw data are available in the Source Data file.

effect on the number of molecules of RNAPII, or the propensity for RNAPII to form clusters (Fig. 7G and Supplementary Fig. 8H–J). However, we did observe a significant increase in the number of RNAPII clusters in cells transiently expressing Halo-NLS-NDP52 (Fig. 7H). This suggests that overexpression of NDP52 might allow the formation of new transcription hubs in the cell, but the overall size of each cluster is not dependent on NDP52.

Based on NDP52's in vitro interactions with DNA, we then explored if the observed impact on RNAPII may be related to changes to chromatin. We performed a chromatin accessibility assay and found that expression of Halo-NLS-NDP52 led to an increase in DNA accessibility (Fig. 7I). This was irrespective of whether the region was typically "open" or "closed", represented by *GAPDH* and *HBB*, respectively.

Having determined that NDP52 can be found clustering at transcription initiation sites, we set out to explore how depletion of NDP52 would affect global gene expression. For this, we performed RNA-Seq in both HeLa and MCF-7 cells, following siRNA knockdown of NDP52 (Supplementary Figs. 9 and 10). Overall, we observed significant changes in gene expression levels for both cell lines, with 1420 genes and 1140 genes differentially expressed in HeLa and MCF-7, respectively ($-0.5 > \log_2 FC < 0.5$, and $p_{adj} < 0.05$) (Supplementary Figs. 9A and 10A). In both HeLa and MCF-7 datasets, more genes are downregulated than upregulated, also showing an overall negative impact on transcription caused by depletion of NDP52. Gene Ontology (GO) analysis for both up and downregulated genes was then performed for both cell lines (Supplementary Figs. 9B and 10B). For HeLa, genes involved in the "regulation of transcription", as well as "cell migration" and "tissue development" were significantly affected (Supplementary Fig. 9B). In MCF-7, NDP52 knockdown was shown to also affect the expression of genes involved in "cell migration", "tissue development", as in HeLa, but also "cell cycle", "DNA replication" and "chromosome segregation" (Supplementary Fig. 10B). This suggests that whilst some genes and processes are equally affected by NDP52 knockdown in both cell lines, others might be more susceptible depending on the unique characteristics of the cell line.

NDP52 nuclear interactome
Following our observation that NDP52 colocalises and clusters at RNAPII transcription initiation sites, and can drive the formation of additional RNAPII clusters when overexpressed, we decided to explore the nuclear interactome of NDP52. To identify partners of NDP52, we used label-free quantitative LC-MS/MS to analyse pull-downs of recombinant NDP52-FL, $_C$NDP52, CoCoA and ZFs from HeLa nuclear extracts (Supplementary Fig. 11A–C and Fig. 8A–D). Proteins were identified using Progenesis software (Waters), as described in "Methods". Following protein identification, we selected proteins enriched in NDP52-FL pull-downs, compared to control beads without recombinant protein. We used $\log_2 FC > 1$ and $p_{adj} < 0.05$ as a threshold to investigate the interactome of NDP52. Following this, we performed GO analysis to determine biological processes that could shed light on a nuclear role for NDP52. Interestingly, the top enriched biological

processes for NDP52 interactions were "DNA geometric change" and "DNA duplex unwinding", with "chromosome organisation" also scoring high (Supplementary Fig. 12A). Top enriched molecular functions also relate to "nucleosome-dependent ATPase activity", "DNA helicase activity" or DNA binding (Supplementary Fig. 12B). These data reinforce the concept that NDP52 has significant roles in DNA binding and structure, as indicated in our biochemical studies and AFM imaging, which could, in turn, impact gene expression. Although several known transcription factors were identified in proteomics (Fig. 8B), gene expression-related GO functions were not particularly enriched. This suggests that although NDP52 might directly interact with transcription-related proteins and also affect gene expression through these interactions, its largest contribution might, instead, arise through regulation of DNA structure.

We were able to map 16% of NDP52-FL interactions to the C-terminal of NDP52 ($_C$NDP52), also through LC-MS/MS with recombinant $_C$NDP52 (Fig. 8A). Figure 8B shows examples of top hits for NDP52-FL and the identification of some of these hits in $_C$NDP52 proteomics. Fold change values, are relative to control beads for pull-downs, with infinity fold-change (∞) indicating hits only present in pull-downs and not in control samples.

We also compared how the interactome of NDP52 relates to its close family member CoCoA (Fig. 8C). Similar to NDP52, the nuclear interactome of CoCoA also showed enrichment for "DNA duplex unwinding" and "DNA geometric change". However, "gene expression" was clearly enriched for CoCoA (Supplementary Fig. 13A, B). Furthermore, when comparing CoCoA and NDP52 interactions, a quarter of CoCoA hits were common to NDP52 (Fig. 8C), showing a degree of overlap between both interactomes, as expected for proteins with high homology. Importantly, our data show that whilst both proteins could have similar functions and overlapping interactomes, they do not appear to be redundant. Figure 8D shows some of the top hits for CoCoA and identification in NDP52-FL pull-downs.

As ZFs is highly conserved in both NDP52 and CoCoA, we also tested if some of the interactions in common between CoCoA and NDP52 could be mapped to this domain (Supplementary Fig. 11B). Interestingly, only 5% of common interactions between NDP52 and CoCoA occur independently of ZFs-binding, suggesting that this domain could account for similarities in the interactomes between both proteins. Furthermore, 86% of proteomics common hits between NDP52-FL and $_C$NDP52, could be mapped to ZFs. This, together with the fact that ZF1 is largely unstructured, could indicate that protein-protein interactions at the C-terminal of NDP52 are mostly sustained by ZF2. Examples of top hits for CoCoA and NDP52 are shown for ZFs proteomics in Supplementary Fig. 11C.

Changes to the nuclear organisation and dynamics of NDP52 following transcription inhibition
As we have established that NDP52 has a role in transcription, and can impact the organisation of RNAPII, we tested whether transcription inhibition would affect its nuclear organisation and dynamics. To
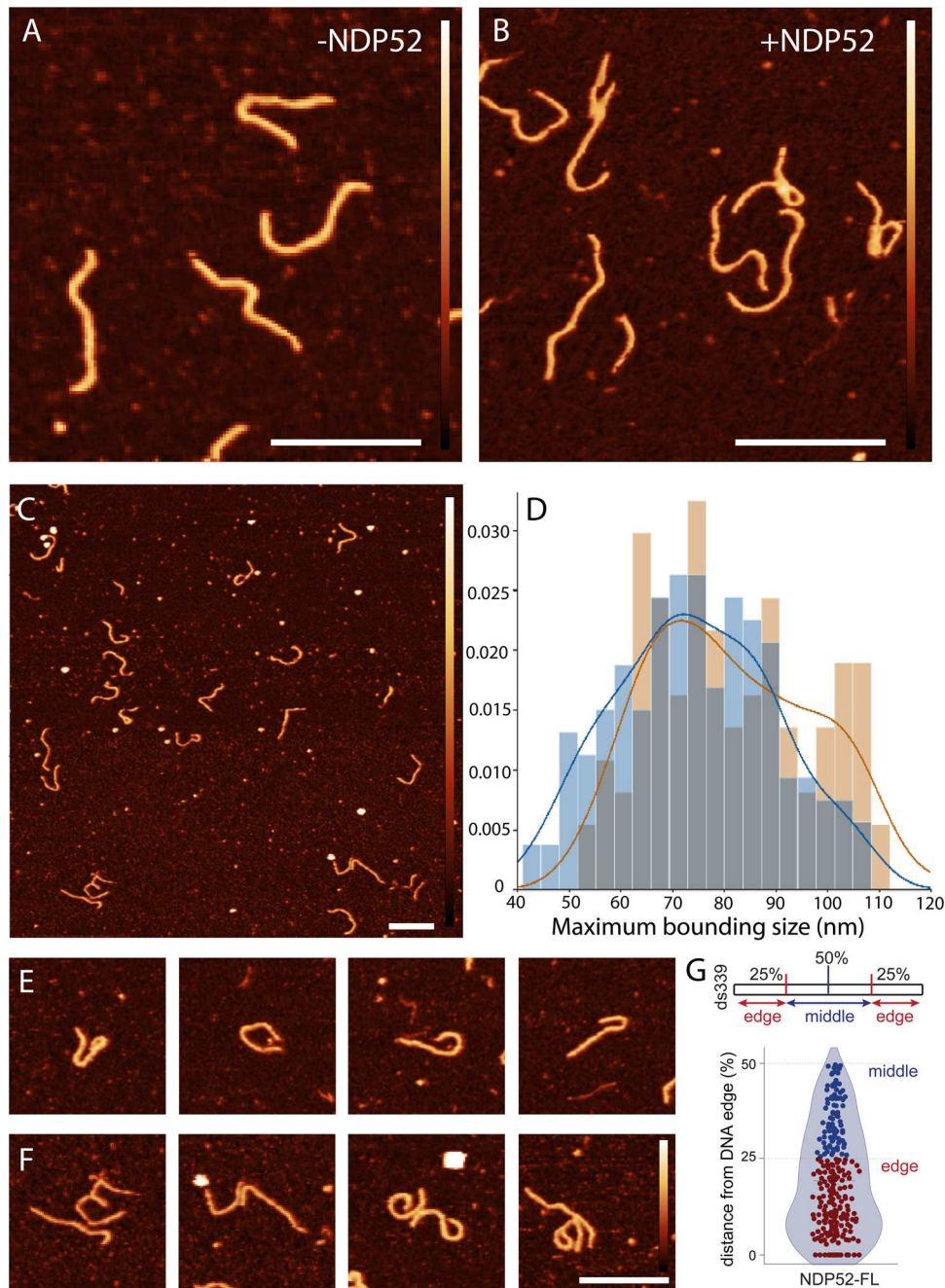
Fig. 6 | NDP52 changes DNA conformation in vitro. A AFM image of ds339 on PLO in the absence of NDP52. Scale bar = 100 nm. Height scale = 4.5 nm. **B** AFM image of ds339 on PLO with NDP52. Scale bar = 100 nm. Height scale = 4.5 nm. **C** AFM images of NDP52-FL bridging strands of linear (339 bp) dsDNA on poly-L-ornithine surface. Scale bar = 100 nm. Height scale = 4.5 nm. **D** Normalised histogram comparing the maximum bounding size of DNA with (light blue) and without NDP52 (orange); their distributions overlap (dark blue). **E** Cropped AFM images of DNA molecules in the presence of NDP52 on PLO. First two images are from between 50–60 nm maximum bounding size while the second two are between 80–90 nm. **F** AFM images showing bridging between DNA molecules in the presence of NDP52 on PLO. xy scale bars 100 nm. **G** Preference for NDP52-FL binding on ds339 molecule. Diagram shows ds339 edge and middle references on linear DNA. Violin plot shows %distance from DNA edge values for NDP52-FL binding. Mean ± SEM is shown for n = 270 binding events. Raw data are available in the Source Data file.

address this, we used α-amanitin, an irreversible RNAPII inhibitor that promotes degradation of RNAPII[29]. As expected, α-amanitin treatment leads to depletion not only of RNAPII molecules, but also of its clusters (Fig. 9A–D and Supplementary Fig. 14A–C).

We then used STORM to compare the spatial organisation of NDP52 in the nucleus of non-treated HeLa cells vs. cells treated with α-amanitin (Fig. 9E, F). The linearised Ripley's K function clearly shows that, compared to non-treated cells, there is a reduced probability for NDP52 clustering in the nucleus, following α-amanitin treatment (Fig. 9G). This is further confirmed through cluster analysis, which

shows a reduction from 44.6% (±5.8) to 20.8% (±4.2) in the percentage of NDP52 molecules forming clusters following α-amanitin treatment (Fig. 9H).

We also explored how reduced clustering of NDP52 in the nucleus, following transcription inhibition, would affect its molecular dynamics. As the number of NDP52 in clusters is markedly reduced, we expected an increase in the dynamic behaviour of the protein. To test this hypothesis, we used acMFM to determine diffusion coefficient and anomalous diffusion changes in cells transiently expressing Halo-NDP52, in non-treated and α-amanitin-treated cells. As expected, loss
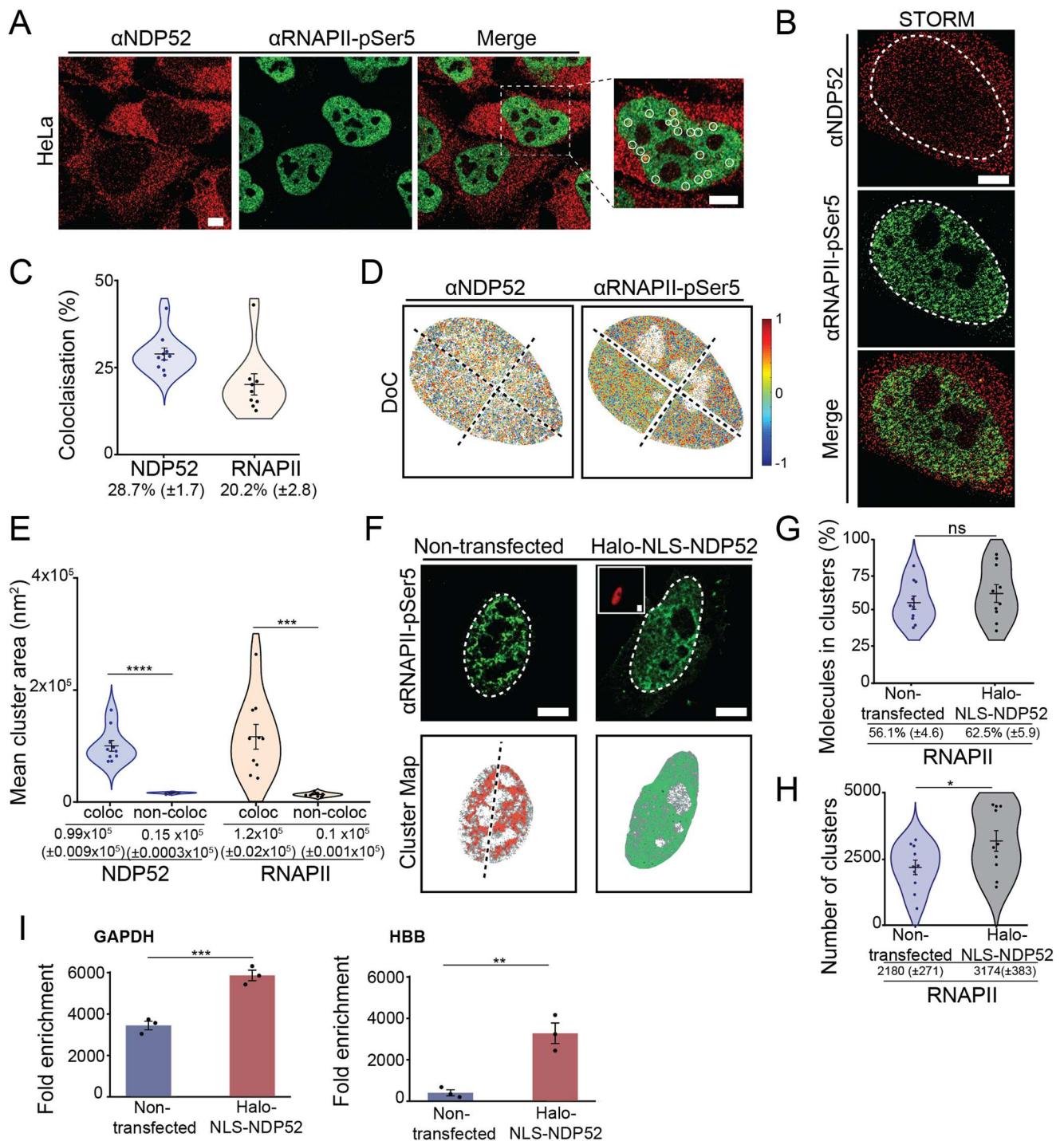
of NDP52 clusters, observed in STORM data, correlated with a significant increase in the diffusion coefficient and anomalous diffusion constant, α, for Halo-NDP52 in the nucleus of α-amanitin-treated cells (Fig. 9I–L and Supplementary Fig. 14D). A proportion of static NDP52 molecules was reduced following α-amanitin treatment (reduction from 56.3% (±1.2) in non-treated to 50.5% (±1.1) in α-amanitin-treated) (Fig. 9J). This was accompanied by a significant increase in molecules in slow diffusion (increase from 42.1% (±1.2) in non-treated to 47.4 (±1.0) in α-amanitin-treated) and a small, non-significant, increase in diffuse molecules (from 1.7% (±0.1) in non-treated to 2.1 (±0.2) in α-amanitin-treated). Overall, our data show that transcription inhibition by α-amanitin disrupts global nuclear NDP52 clustering in the nucleus, which correlates with higher molecular diffusion of the protein.

Having shown that the spatial organisation of nuclear NDP52 is altered following transcription inhibition, we also tested if this would also cause changes to the interactome of the protein. For this we used label-free quantitative LC-MS/MS of co-immunoprecipitation assays, for endogenous NDP52 from whole-cell HeLa extracts, with and without α-amanitin treatment (Fig. 9E). Co-immunoprecipitation assays were performed as six replicates and compared to protein A controls. The same $\log_2 FC > 1$ and $p_{adj} < 0.05$ threshold was used to identify enriched GO processes, whilst in non-treated cells we can observe "gene expression" as an enriched process, α-amanitin treatment disrupts this and appears to change NDP52 interactome to "regulation of DNA replication", "signal transduction in response to DNA damage", and "chromosome organisation" (Supplementary

Fig. 7 | Colocalisation of NDP52 with RNAPII-pSer5. A Immunofluorescence confocal image of NDP52 (red) and RNAPII-pSer5 (green) in HeLa cells, showing detail of colocalising foci in white circles (zoomed-in right panel). Scale bar = 5 μm. **B** Example STORM images of NDP52 and RNAPII-pSer5 in HeLa. The nuclear region (determined by RNAPII-pSer5 fluorescence) was used for ClusDoC analysis (shown as dotted white line). Scale bar = 5 μm. **C** Colocalisation analysis of NDP52 and RNAPII-pSer5 clusters (*n* = 10 cells Errors bars are mean ± SEM). **D** Cluster colocalisation heat maps for NDP52 and RNAPII-pSer5 generated from the STORM data shown in (**B**). DoC score of 1 represents perfect colocalisation between molecules, and DoC score −1 represents segregation. A DoC score of 0.4 was used as threshold for colocalisation. Due to high molecular density the nucleus was split into four ROIs for ClusDoC analysis. Axis of separation for the images are shown as dotted lines. **E** Mean cluster area is shown for colocalised and non-colocalised clusters of NDP52 and RNAPII-pSer5. *n* = 10 cells. Errors bars are mean ± SEM. **F** STORM rendering for RNAPII-pSer5 and generated cluster maps for HeLa cells non-transfected or transiently expressing a Halo-NLS-NDP52 construct. Inset in Halo-NLS-NDP52

panel shows wide-field channel Halo ligand-JF549 labelled cells. Scale bar = 5 μm. **G** Percentage of RNAPII-pSer5 molecules in clusters in non-transfected or transiently expressing Halo-NLS-NDP52 HeLa cells (*n* = 10 cells. Errors bars are mean ± SEM). ns = *p* > 0.05 by two-tailed *t*-test. **H** Number of RNAPII-pSer5 clusters in non-transfected or transiently expressing Halo-NLS-NDP52 HeLa cells. Mean ± SEM values are shown. Each point represents the average value per cell. *n* = 10 cells (non-transfected) *n* = 10 cells (Halo-NLS-NDP52). **p* < 0.05; ****p* < 0.001; *****p* < 0.0001 by two-tailed *t*-test. **I** Chromatin accessibility assay. qPCRs were performed against the stated genes on chromatin purified from HeLa cells, in the presence and absence of transfected Halo-NLS-NDP52. Chromatin was divided into nuclease and control samples where nuclease activity was determined by the difference between the two samples (Fold enrichment). *HBB* represents a repressed, inaccessible, control gene. Conversely, *GAPDH* is from an open, assessable region. Error bars represent SEM from three independent experiments. ***p* < 0.01; ****p* < 0.001; by two-tailed *t*-test. Raw data are available in the Source Data file.

Fig. 14E, F). Figure 8E, F shows examples of top hits for endogenous NDP52 pull-downs in non-treated vs. α-amanitin-treated cells. The shift observed in interacting partners suggests that NDP52 preferentially interacts with different proteins, depending on cell state, and could change its interactome in response to environmental stresses. Interestingly, in non-treated cells we observe different GO enrichments to those observed in recombinant protein pull-down assays (Fig. 8 and Supplementary Fig. 12). The high concentration of recombinant NDP52-FL used could have allowed the identification of different interactions, that under normal conditions and cellular levels of NDP52 are less enriched.

Overall, our data show NDP52 as a transcription regulator, with functions in DNA structure.

Discussion

In this study, we used a multidisciplinary approach to shed light on the nuclear role of autophagy receptor NDP52. Although NDP52 was first observed in the nucleus[1], until now, no clear nuclear function had been attributed to this protein. By investigating its nuclear organisation, dynamics, interactome and biochemical characteristics, we have been able to link its function to transcription and DNA regulation.

To enhance their activity and functional efficiency, many nuclear proteins involved in transcription and other nuclear processes form molecular clusters[18,19,30–32]. Here, we have determined that nuclear NDP52 clusters at regions of transcription initiation with RNAPII, and that its overexpression can increase the number of transcriptional clusters available in the nucleus. RNAPII clustering is directly related to transcription activity[33–35] and changes to its spatial organisation and clustering behaviour impact whole gene-expression levels[36]. This provides a direct association between NDP52 and transcriptional regulation. Furthermore, we have also shown that knockdown of NDP52 impacts gene expression in both HeLa and MCF-7 cells. Altogether, these data support a role for NDP52 in RNAPII-related transcription.

Although future studies will be necessary to determine a mechanism for the regulatory role of NDP52 in transcription, here we propose two different strategies: (1) through interactions with transcriptional machinery and regulatory factors at transcriptional sites and/or (2) through direct DNA structure regulation and interaction with chromatin remodellers (Fig. 10). In support of our first hypothesis, we show that NDP52 colocalises with RNAPII at transcription initiation sites and that clustering of NDP52 is abrogated following transcription inhibition. Although we observed more non-colocalised clusters of NDP52-Ser5 RNAPII than colocalised, it is important to note that not all Ser5-RNAPII represents actively transcribing complexes, and that part of this population is stalled/paused. Hence, it is possible that NDP52 preferentially localises with a subset of RNAPII clusters, for example, with molecules that are going through initiation into active elongation.

Following transcription inhibition with α-amanitin treatment, the observed increase in NDP52 molecular dynamics suggests a reduction in binding sites available for the protein in the nucleus. These data indicate that loss of RNAPII molecules, due to its degradation, directly affects the nuclear organisation of NDP52. Furthermore, our proteomics data, both with recombinant and endogenous NDP52, also shows interactions with different transcriptional regulators. These interactions are markedly reduced in endogenous NDP52 pull-downs when cells are pre-treated with α-amanitin (Fig. 8), which further supports the hypothesis that interactions between NDP52 and transcriptional regulators are an important part of the regulatory function of this protein.

Equally, we also show that NDP52 binds specifically and with high affinity to dsDNA, and we believe this interaction to be crucial for its observed role in transcription. Importantly, we have shown that NDP52 can be isolated in complex with different genomic *loci*, through ChIP-qPCR, and this could be a regulatory strategy for the protein. In fact, CoCoA, a gene paralog of NDP52 and known transcription co-activator, has been found present at different genes regulated by nuclear receptors, such as the *TFF1/PS2* gene[37]. Although our biochemical studies also show that CoCoA can directly bind DNA, previous studies have determined that interactions of this protein with histone methyltransferases and acetyltransferases at the gene body allows the recruitment of basal transcriptional machinery, thus promoting transcriptional activity[11,37,38]. It will be interesting, in future studies, to produce a more comprehensive analysis of NDP52-genomic DNA interactions, through ChIP-Seq. There are also no available ChIP-Seq data for CoCoA. It would be informative to determine how similar the genomic targets of these two proteins are, given their biochemical likeness.

Interestingly, when exploring NDP52 binding to DNA, we observed that NDP52 promotes changes in DNA structure in vitro– through bending, bridging and DNA looping. This, together with our proteomics data showing regulators of chromatin and DNA structure as possible binding partners of NDP52, suggests a role in chromatin regulation for NDP52. Chromatin conformation and structure are important determinants of accessibility to transcriptional machinery[39]; as a result, chromatin regulation is directly linked to transcriptional activity. In line with this, we found that overexpression of NDP52 in the nucleus leads to an increase in chromatin accessibility through driving decondensation processes. Our data suggest that, in addition to direct regulation at transcriptional sites, either through direct interactions with RNAPII or other transcription factors, NDP52 activity on DNA structure and links to chromatin organisation could also drive transcriptional regulation. This could also explain how overexpression of NDP52 leads to an increase in RNAPII clusters. In future studies, it will be important to address specifically the role of NDP52 in chromatin structure and regulation
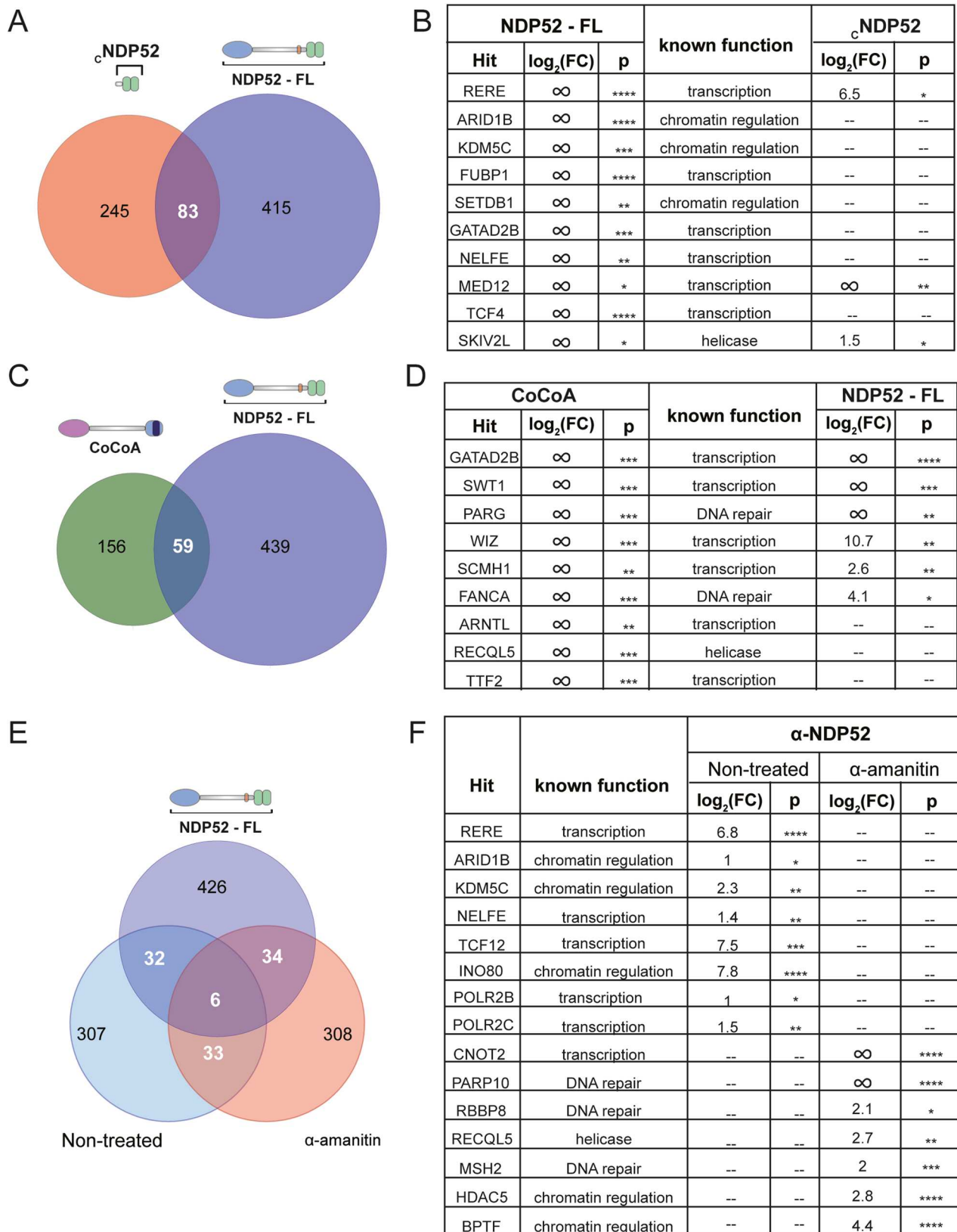
Fig. 8 | NDP52-FL, CoCoA and CNDP52 interactomes from HeLa nuclear extract.
A Venn diagram of hits found in NDP52-FL and $_c$NDP52. **B** Examples of top hits for NDP52, and their identification in $_c$NDP52 proteomics data. log$_2$FC is relative to beads control. **C** Venn diagram of hits found in NDP52-FL and CoCoA. **D** Examples of top hits for CoCoA, and their identification in NDP52-FL proteomics data. log$_2$FC is relative to beads control. **E** Venn diagram showing overlap of identified hits between recombinant NDP52-FL proteomics and co-immunoprecipitation of endogenous NDP52 for non-treated and α-amanitin-treated cells. **F** Examples of hits identified in non-treated and α-amanitin-treated cells. log$_2$FC is relative to beads control. *FDR < 0.05, **FDR < 0.01, ***FDR < 0.001, ****FDR < 0.0001 by ANOVA.
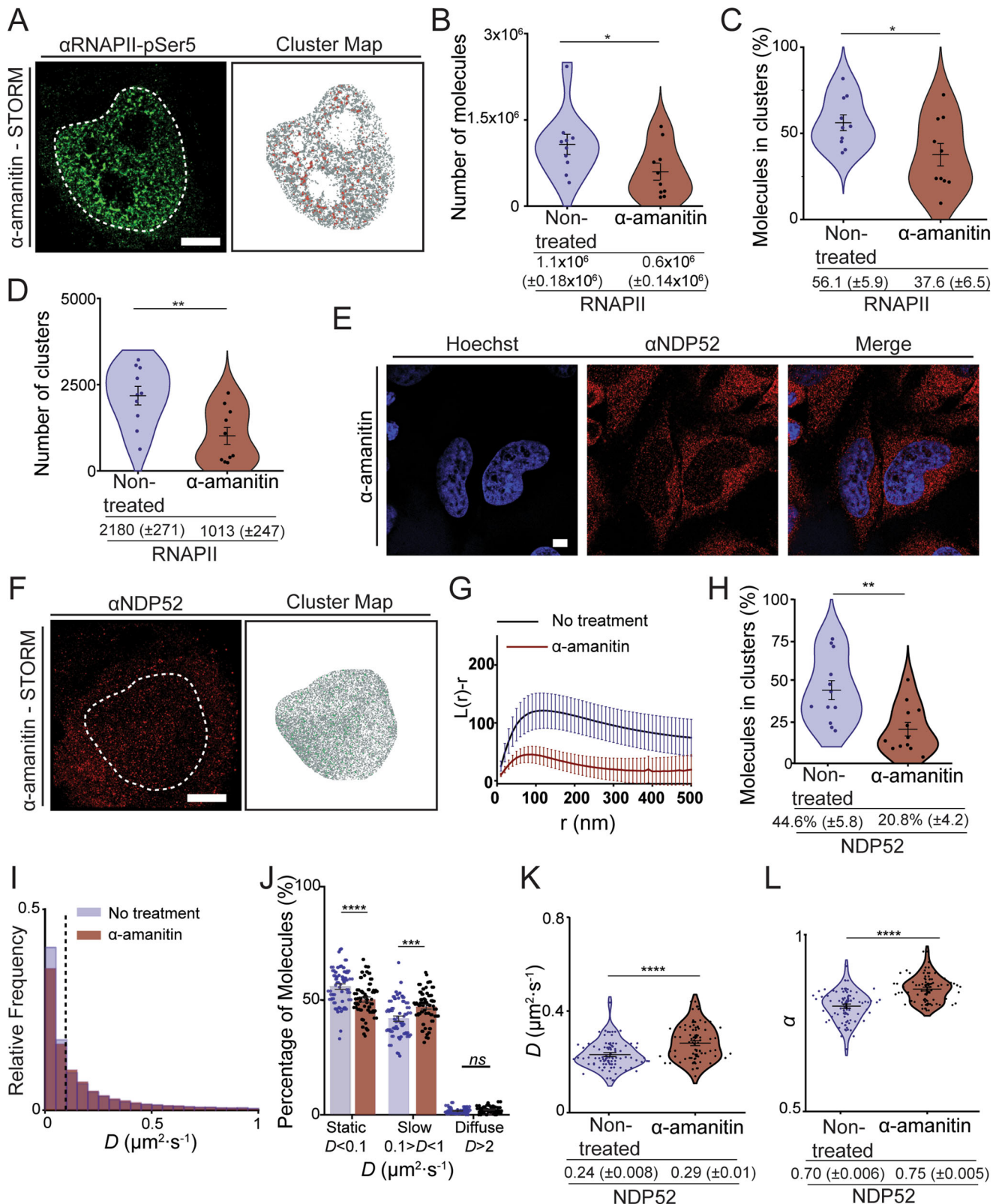
and explore some of the possible interactions with DNA binding proteins and regulators identified in our proteomics data.

Our work has also provided a detailed biochemical analysis of NDP52. We show that NDP52, in solution, is predominantly dimeric, as previously suggested[6]; although it can also associate into higher oligomeric forms, such as trimers and tetramers. We also found that, in addition to its coiled-coil region, the C-terminal and N-terminal

domains can also independently interact and form oligomeric structures. We also provide evidence that NDP52 interacts with DNA through its zinc finger domains and that both the full-length and C-terminal domains of NDP52 can oligomerise with and around DNA in vitro. Furthermore, we show that NDP52 can modify the local conformation of DNA, in vitro. In cells, this could provide a mechanism for a possible function of NDP52 in chromatin structure regulation, by

Fig. 9 | Organisation and dynamics of nuclear NDP52 following transcription inhibition. A Example of STORM rendering and cluster map, generated following DBSCAN analysis of RNAPII-pSer5 following α-amanitin treatment. Scale bar = 5 μm. **B** Calculated number of RNAPII-pSer5 molecules in the nucleus of non-treated vs. α-amanitin-treated cells ($n = 10$ cells). **C** Percentage of RNAPII-pSer5 molecules in clusters for non-treated vs. α-amanitin-treated cells ($n = 10$ cells). **D** Number of clusters in selected ROIs for RNAPII-pSer5 in non-treated vs. α-amanitin-treated cells. $n = 10$ cells (non-treated) $n = 10$ cells (α-amanitin). **E** Confocal image of NDP52 in HeLa cells following treatment with transcription inhibitor α-amanitin. Scale bar = 5 μm. Hoechst DNA stain is shown in blue. **F** Example STORM image of NDP52 in HeLa cells treated with α-amanitin and corresponding cluster map. Scale bar = 5 μm. Clustered molecules are shown in green. **G** Linearised Ripley's K Function, $L(r)−r$, calculated for selected ROIs from STORM images. Ripley's K values red = α-amanitin-treated cells and blue = non-treated cells (values are the same as shown in Fig. 2C). $n = 11$ (α-amanitin) $n = 12$ (non-treated). **H** Percentage of

molecules in clusters in non-treated HeLa cells compared to α-amanitin treatment. Values for non-treated cells are the same as shown in Fig. 2G. $n = 11$ (α-amanitin) $n = 12$ (non-treated). **I** Histogram of diffusion constants for Halo-NDP52 non-treated (blue) and α-amanitin-treated (red) HeLa cells. Dotted lines represent the applied threshold to differentiate between static and dynamic molecules (14,322 molecules from 51 cells for non-treated condition and 14,492 molecules from 50 cells). Non-treated cell values are the same as shown in Fig. 3G. **J** Percentage of molecules considered static ($D < 0.1 \, \mu m^2/s$), slow moving ($0.1 < D < 1 \, \mu m^2/s$) or diffuse ($D > 1 \, \mu m^2/s$) per cell. $n = 51$ cells. (non-treated−same as Fig. 3J) and $n = 50$ (α-amanitin). **K** Diffusion coefficient values and (**L**) Anomalous diffusion constants (α) for Halo-NDP52. Each data point represents the mean value for a cell. $n = 51$ (non-treated−same as Fig. 3H) and $n = 50$ (α-amanitin). For all experiments: errors bars are mean ± SEM. ns = $p > 0.05$; *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; ****$p < 0.0001$ by a two-tailed t-test. Raw data are available in the Source Data file.
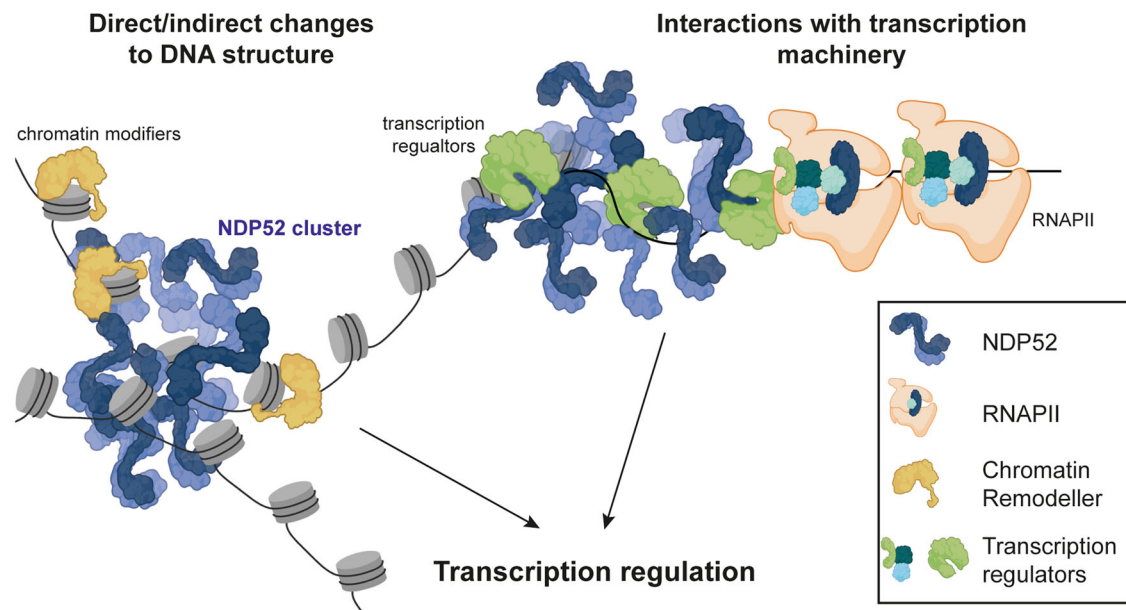

Fig. 10 | Model of possible mechanism for NDP52's activity in transcription. NDP52 could directly interact with DNA in the nucleus, or with chromatin modifiers (e.g. histone modifiers), to cause local changes to chromatin structure. Conversely, interactions with transcription factors/coactivators and transcription machinery could also modulate transcription activity of genes. Created with BioRender.com.

increasing local concentrations of the protein around DNA. Moreover, given that the C-terminal domain is the main region for interactions with DNA, it is possible that, in the cell, NDP52 binds to DNA through this region whilst simultaneously sustaining interactions with other proteins through its N-terminal and coiled-coil regions. AFM data suggests a variety of spatial orientations might be available during oligomerisation of the protein, but further studies will be necessary.

Whilst the C-terminal of NDP52 is crucial for its DNA binding activity, our proteomics data also indicate that many important regulatory protein-protein interactions also occur through this region. In fact, previous studies have identified the C-terminal domain of NDP52 as the main interacting region with MVI and ubiquitin[40]. Interestingly, the majority of common hits found in recombinant proteomics between full-length NDP52 and CoCoA arise from the C-terminal domain region of the protein. This is not surprising as this region displays high levels of amino acid homology in both proteins.

Interestingly, our proteomics data with endogenous NDP52 following treatment with α-amanitin, showed an enrichment for proteins involved in cell stress and DNA damage response. Although this was not explored in this study, it will be interesting to understand, in the future, how nuclear NDP52 responds to different cellular stresses. In the cytoplasm, NDP52 is known to be activated by certain cell stresses, namely in response to bacterial or viral infections or the presence of

damaged organelles[40–42]. Moreover, we demonstrated that the induction of autophagy increases the nuclear pool of NDP52. It is possible that NDP52 has a dual cytoplasmic-nuclear role in cells and that its nuclear activity is linked to its cytoplasmic function in autophagy. Following cellular infection, innate immunity and apoptotic pathways are activated in the cytoplasm that lead to the translocation of different proteins into the nucleus for their activity in transcription, chromatin, and DNA repair regulation[43–45]. Understanding the molecular role of NDP52 and its nuclear activity in context of its already known cytoplasmic function in autophagy will be important in future studies.

Overall, here, we provide evidence for NDP52 as a transcriptional regulator, with possible functions in chromatin structure and organisation.

Methods

Constructs

A list of constructs is provided in Supplementary Table 1.

Cell culture and transfection

HeLa (ECACC 93021013) and MCF-7 (ECACC 86012803) cells were cultured in MEM Alpha medium (Gibco), with GlutaMax (no nucleosides), supplemented with 10% Fetal Bovine Serum (Gibco), 100 μg/ml streptomycin (Gibco) and 100 units/ml penicillin, at 37 °C and 5% CO_2.

To inhibit transcription, cells were treated with 5 µg/ml α-amanitin for 4 h at 37 °C and 5% CO_2.

For transient transfection of Halo-NDP52 or Halo-NLS-NDP52, HeLa cells were cultured in Nunc LabTek dishes (Merck) or glass coverslips #1.5, were transfected using Lipofectamine 2000 (Invitrogen) for 24 h.

Immunofluorescence

Following nuclear staining using Hoechst 33342 (Thermo Scientific), HeLa and MCF-7 cells cultured on glass coverslips #1.5 were fixed for 15 min at room temperature, in 4%(w/v) paraformaldehyde (PFA) in Phosphate-buffered saline (PBS). Residual PFA was then quenched for 15 min using 50 mM Ammonium Chloride in PBS.

Cells were permeabilised and blocked in 2% (w/v) Bovine Serum Albumin (BSA), 0.1% (v/v) Triton X-100 in PBS for 30 min. Cells were then labelled against endogenous proteins for 1 h in 2% (w/v) BSA, with appropriate primary antibody and, subsequently, with appropriate fluorophore-conjugated secondary antibodies. When using anti-phospho antibodies, the immunofluorescence protocol was performed in Tris-buffered saline (TBS).

For endogenous NDP52, rabbit anti-NDP52 (1:200, Genetex GTX115378) antibodies were used. For RNA polymerase II, mouse anti-RNAPII phospho Ser 5 (1:500, Abcam, ab5408) was used. Secondary conjugated antibodies Donkey anti-rabbit Alexa 647 (1:500, Abcam, 181347) and Donkey anti-mouse Alexa 488 (1:500, Abcam, ab181289) were used.

Coverslips were then mounted on microscope slides in Mowiol solution (10%(w/v) mowiol 4–88, 25%(v/v) glycerol, 0.2 M Tris-HCl, pH 8.5) with 2.5%(w/v) DABCO (Sigma).

Autophagy assays

HeLa cells were cultured on #1.5 glass coverslips for 24 h to 50% confluency. Cells were then treated with 500 nM rapamycin, 50 µM chloroquine or DMSO for 18 h. Cells were then stained with Hoechst 33342 (Thermo Scientific) and fixed for 15 min at room temperature, in 4%(w/v) PFA in PBS. Residual PFA was then quenched for 15 min using 50 mM Ammonium Chloride in PBS. The immunofluorescence protocol is then followed.

Confocal microscopy

Fixed cells were imaged using a Zeiss LSM980, with a Plan-Achromat 63 × 1.4 NA oil immersion objective (Carl Zeiss, 420782-9900-000). Three laser lines: 405, 488 and 642 nm were used for excitation of Hoechst, Alexa-fluor 488 and Alexa-fluor 647 fluorophores. Built-in multi-band dichroic mirror MBS405/488/561 (Carl Zeiss, 1784-995) were used to reflect excitation laser beams onto samples. For fluorescence signal collection, the used emission spectral bands were: 410–524 nm (Hoechst), 493–578 nm (Alexa-fluor 488) and 564–697 nm (Alexa-fluor 647). The green channel (Alexa-fluor 488) was imaged using a 1 gallium arsenide phosphide (GaAsP) detector, while the blue (Hoechst) and red (Alexa-fluor 647) channels were imaged using two multi-anode photomultiplier tubes (MA-PMTs). For imaging acquisition and rendering, Zeiss ZEN Blue software (v2.3) was used. Confocal Images were deconvolved using the Zeiss Zen Blue software (v2.3), using the regularised inverse filter method.

Stochastic optical reconstruction microscopy (STORM)

No. 1.5, 25 mm round glass coverslips were cleaned by incubation with etch solution (5:1:1 ratio of H_2O:H_2O_2 (50% wt in H_2O stabilised, Fisher Scientific):NH_4OH (ACS reagent, 29–30% NH3 basis, Sigma) for 2 h in a 70 °C water bath. Cleaned coverslips were washed in filtered water and ethanol and allowed to dry before cell seeding.

Cells were fixed for 15 min in 4%(w/v) PFA in PBS and residual PFA was quenched with 50 mM ammonium chloride in PBS for 15 min. Immunofluorescence was performed using filtered TBS. Cells were first permeabilised and blocked for 30 min in 3% (w/v) BSA, 0.1% (v/v) Triton X-100. Cells were then incubated in primary antibody for 1 h, at the same dilution as for the normal immunofluorescence protocol. Cells were washed three times (10 min each wash) with 0.2% (w/v) BSA, 0.05% (v/v) Triton X-100 in TBS. Cells were subsequently incubated in an appropriate fluorophore-conjugated secondary antibody for 1 h, at a 1:250 dilution, in 3% (w/v) BSA, 0.1 % (v/v) Triton X-100. Cells were washed in TBS and PBS and fixed in 4% (w/v) PFA in PBS a second time. Cells were stored in PBS supplemented with 0.02% (w/v) NaN_3 in the dark until imaging.

Before imaging, coverslips were washed in filtered H_2O and assembled into Attofluor cell chambers (Invitrogen). Imaging was performed in STORM buffer—10% (w/v) glucose, 10 mM NaCl, 50 mM Tris-HCl pH8.0—supplemented with GLOX solution (5.6% (w/v) glucose oxidase and 3.4 mg/ml catalase in 50 mM NaCl, 10 mM Tris-HCl pH 8.0) and 0.1% (v/v) 2-mercaptoethanol.

STORM imaging was performed using a Zeiss Elyra PS.1 system. For sample illumination HR Diode 488 nm (100 mW) and HR Diode 642 nm (150 mW) lasers were used, where power density on the sample was 7–12 kW/cm^2 and 7–14 kW/cm^2, respectively. Built-in multi-band dichroic mirror MBS 405/488/642 (Carl Zeiss 1784-996) were used to reflect excitation laser beams onto samples. To reduce background fluorescence levels, imaging was performed using Highly Inclined and Laminated (HILO) illumination with a ×100 NA 1.46 oil immersion objective (Carl Zeiss alpha Plan-Apochromat, 420792-9800-000). For fluorescence signal collection, a BP 420-480/BP 495-550/LP 650 multi-bandpass emission filter (Carl Zeiss 1769-207) was used and a final image was acquired using an Andor iXon DU 897 EMCCD camera with 25msec exposure, for 25,000 frames.

Image processing was performed in Zeiss Zen Black software. For two-colour STORM images, channel alignment was performed following a calibration procedure using pre-mounted MultiSpec beads (Carl Zeiss, 2076-515). For the calibration procedure, the affine method was performed, to account for lateral stretching and tilting between the two channels. This was performed for each day of acquisitions. Blinking event detection was then performed in Zeiss Zen Black software using a 9-pixel mask with a signal-to-noise ratio of 6, accounting for overlap, to allow localisation of molecules in dense environments. Final molecule positions were then determined through fitting of a 2D Gaussian function. Molecule positions were subjected to model-based cross-correlation drift correction. For Alexa-fluor-647 and Alexa-fluor-488 labelled molecules, the typical mean value of localisation precision was 20 nm and 30 nm, respectively. Molecule localisation tables were exported as.txt files for further analysis using ClusDoC software[23].

ClusDoC

Following export of molecule positions from Zeiss ZEN Black software, STORM data were further analysed using ClusDoC software (https://github.com/PRNicovich/ClusDoC)[23]. The nucleus was selected as an ROI for cluster analysis. The Ripley K function was first calculated for the ROI selected, to identify the *r* max. This value was then used in DBSCAN analysis for single-channel images or ClusDoC analysis for two-channel images. Minimum cluster size was set to 5 molecules with a smoothing value set at 7 and an epsilon value set at the mean localisation precision value for the dye. Other parameters remained at default values.

Aberration corrected multi-focal microscopy

Cells transiently expressing Halo-NDP52 were labelled for 15 min with 10 nM Halo tag-JF549 ligand in cell culture medium at 37 °C, 5% CO_2. Cells were then washed three times in complete media and incubated for at least 30 min at 37 °C, 5% CO_2 before imaging. For imaging, cell media was replaced with FluoroBrite DMEM medium (Thermo Fisher Scientific).

Single-molecule tracking experiments were performed using an aberration-corrected multi-focal microscope (acMFM), described in Abrahamsson et al.[46]. Briefly, a custom optical system appended to the detection path of an optical Nikon Ti microscope was used. The detection path of the microscope included a diffractive multi-focal grating in a conjugate pupil plane, a chromatic correction grating, to reverse spectral dispersion, and a nine-faceted prism followed by the final imaging lens. A 561 nm laser was used for excitation, with a 4–6 kW/cm^2 power density at the back aperture of a 100 × 1.4NA objective (Nikon).

AcMFM imaging produces nine separate, simultaneous images, each representing a separate focal plane, with an axial separation of ca. 400 nm between them. Field of view is ca. 20 µm. The nine-image array was digitised via an EMCCD camera (iXon Du897, Andor), at up to 32 msec temporal resolution, with a typical duration of 30 s.

3D+t single-molecule images were reconstructed via a calibration procedure in Matlab (MathWorks) that accounts and calculates (1) inter-plane spacing, (2) affine transformation for the correct alignment of each focal plane in xy and (3) slight variations in the detection efficiency in each plane—typically less than 5–15% from the mean.

Reconstructed data were pre-processed, including background subtraction and deconvolution (3–5 Richardson-Lucy iterations) and/or Gaussian de-noising prior to 3D particle tracking using the MOSAIC software suite. Maximum particle displacement was set at 400 nm and a minimum 10 frames was required. Detected tracks were reconstructed and diffusion constants calculated through MSD analysis using custom Matlab software assuming an anomalous diffusion model.

Fluorescence recovery after photobleaching

Cells transiently expressing Halo-NDP52 were labelled for 15 min with 10 nM Halo tag-JF549 ligand in cell culture medium at 37 °C, 5% CO$_2$. Cells were then washed three times in CO$_2$-independent medium (Thermo Fisher) before imaging.

FRAP measurement was performed using a Zeiss LSM 880 system equipped with a 100× NA 1.46 oil immersion objective (Carl Zeiss alpha Plan-Apochromat, 420792-9800-000). For sample illumination a 20 mW 561 nm diode laser was used. Built-in multi-band dichroic mirror MBS 458/561 was used to reflect excitation laser beams onto samples. For fluorescence signal collection, the wavelengths from 566 to 685 nm were captured using a multi-anode photomultiplier tube (MA-PMT) with 0.96 µs pixel dwell time. The detector master gain was 900, and digital gain was 1.

Ten frames of confocal microscopy image under 8 mW 561 nm laser illumination were acquired before photobleaching. Selected regions of interest (ROIs) were exposed to full laser power, followed by 100 s of confocal microscopy image acquisition. The time course of fluorescence intensity from the selected ROIs was recorded by Zeiss ZEN 2.3 Blue software. Fluorescence intensity time traces from ROIs, whole-cell areas and background areas were exported as.txt files, and then were analysed using easyFRAP Software (https://easyfrap.vmnet.upatras.gr/?AspxAutoDetectCookieSupport=1)[47].

RNA-sequencing and analysis

Total RNA was extracted from three replicates of NDP52 KD (using CALCOCO2 siRNA, Ambion, 4392420) and scrambled siRNA (using control siRNA, Qiagen, 1027280) in MCF-7 and HeLa cells. Ice cold TRIzol reagent was added to each cell culture dish and homogenised. The mixture was incubated at room temperature for 10 min. Chloroform was then added to the mixture and incubated at room temperature for 5 min. The samples were centrifuged at 8000 × g. The top, aqueous layer was collected and isopropanol was added. The mixture was then centrifuged at 12,000 × g for 30 min and the supernatant discarded. The pellet was washed in 75% (v/v) ethanol and centrifuged at 7500 × g for 5 min. The pellet was air dried and resuspended in RNAse-free H$_2$O. RNA concentration and quality was then assessed by measuring absorbance at 260 nm and A260/A280 ratio. RNA samples were stored at −80 °C.

The following procedures were performed by GENEWIZ and Glasgow Polyomics. The RNA-seq libraries were prepared using Poly-A selection. Resulting libraries concentration, size distribution and quality were assessed on a Qubit fluorometer and on an Agilent 2100 bioanalyzer. Paired-end sequencing (2 × 150 bp) was then performed on an Illumina NovaSeq next generation sequencer for HeLa cells (GENEWIZ) and (2 × 75 bp) on a HiSeq sequencer for MCF7 cells (Glasgow Polyomics).

Sequence reads were trimmed to remove possible adapter sequences and nucleotides with poor quality using Trimmomatic v.0.36. The trimmed reads were mapped to the Homo sapiens GRCh38 reference genome available on ENSEMBL using the STAR aligner v.2.5.2b. BAM files were generated as a result of this step. Unique gene hit counts were calculated by using featureCounts from the Subread package v.1.5.2. After extraction of gene hit counts, the gene hit counts table was used for downstream differential expression analysis. Using DESeq2, a comparison of gene expression between the customer-defined groups of samples was performed. The Wald test was used to generate p values and log2 fold changes. Genes with an adjusted p value < 0.05 and absolute log2 fold change >1 were called as differentially expressed genes for each comparison.

Differentially expressed genes by at least 1.5-fold ($-0.5 \geq \log_2 FC \geq 0.5$) and adjust p value < 0.05 were subjected to Gene Ontology analysis, using iDEP93 (http://bioinformatics.sdstate.edu/idep93/)[48]. RNA-Seq data were deposited in the Gene Expression Omnibus (GEO) database under the accession number GSE188567.

Chromatin immunoprecipitation and qPCR

To identify NDP52-DNA interactions, ChIP was performed using anti-rabbit NDP52 antibody (Genetex GTX115378). HeLa cells were cross-linked by adding formaldehyde directly to the cell medium, to a final concentration of 0.75% (v/v). Cells were left to incubate with gentle rotation at room temperature for 10 min. The reaction was stopped by adding glycine to a final concentration of 125 mM and incubating the mixture for 5 min at room temperature with rotation. Cells were washed twice with cold PBS and scraped in cold PBS. All cells were collected by centrifugation at 1000 × g at 4 °C for 5 min. The pellet was resuspended in ChIP lysis buffer—50 mM HEPES-KOH pH 7.5, 140 mM NaCl, 1 mM EDTA pH 8.0, 1% (v/v) Triton X-100, 0.1% (m/v) Sodium Deoxycholate, 0.1% (m/v) SDS—supplemented with protease inhibitors, using 750 µl per 1 × 10^7 cells. Cells were sonicated using a diogenode bioruptor sonicator to shear DNA until an average DNA fragment size of 200–800 bp was achieved. Fragment size was determined using a 1.5% agarose gel. Cell debris were removed through centrifugation of samples at 8000 × g for 10 min at 4 °C. The supernatant, enriched for chromatin, was stored at −80 °C until used for immunoprecipitation experiments.

Chromatin fractions were diluted 1:10 in RIPA buffer—50 mM Tris-HCl pH 8.0, 150 mM NaCl, 2 mM EDTA pH 8.0, 1% (v/v) NP40, 0.5% (m/v) Sodium Deoxycholate, 0.1% (m/v) SDS—supplemented with protease inhibitors. Three samples were used for immunoprecipitation with NDP52 and three samples for no-antibody control (beads only). In total, 10% of total chromatin was removed as input sample and stored at −20 °C. All samples were pre-cleared using protein A magnetic beads (Thermo Fisher Scientific) for 30 min at 4 °C with end-to-end rotation. Immunoprecipitation replicates were incubated overnight with NDP52 antibody (1:50 dilution) at 4 °C with end-to-end rotation. The following day, 40 µl of protein A magnetic beads, pre-equilibrated in RIPA buffer, were added to each sample, including the no-antibody controls. Samples were incubated with end-to-end rotation at 4 °C for 1 h. Following this, beads were collected using a magnetic rack and washed twice in low-salt buffer (20 mM Tris-HCl pH 8.0, 150 mM NaCl, 0.1%

(m/v) SDS, 1% (v/v) Triton X-100, 2 mM EDTA) followed by a wash in high-salt (20 mM Tris-HCl pH 8.0, 500 mM NaCl, 0.1% (m/v) SDS, 1% (v/v) Triton X-100, 2 mM EDTA), a wash in LiCl buffer (10 mM Tris-HCl pH 8.0, 250 mM LiCl, 1% (m/v) Sodium deoxycholate, 1% (v/v) NP40, 1 mM EDTA pH 8.0) and, finally, in TE buffer (10 mM Tris-HCl pH 8.0; 1 mM EDTA pH 8.0).

DNA was eluted by incubating the beads with 120 μl elution buffer (1%(w/v) SDS; 100 mM NaHCO₃) at 30 °C, with shaking. To reverse crosslinking, eluted protein-DNA complexes and input samples were incubated overnight with 4.8 μl NaCl (5 M) and 2 μl RNAse A (10 mg/ml) at 65 °C with shaking. The following day, samples were incubated with Proteinase K for 1 h at 60 °C. The DNA was purified using phenol:chloroform extraction and samples analysed using QuantiNova SYBR Green qPCR kit (Qiagen). A list of qPCR primers for ChIP is supplied in Supplementary Table 2.

Chromatin accessibility assay

HeLa cell monolayers were transfected with Halo-NLS-NDP52 or lipofectamine control, for 24 h. Cells were trypsinized and washed in PBS. Chromatin was then isolated, digested and then DNA purified using the Chromatin Accessibility Assay Kit (Abcam ab185901) according to manufacturer's instructions. The samples were analysed by qPCR using primers in Table S2.

Protein expression and purification in *Escherichia coli*

Recombinant protein expression was performed in *E. coli* BL21 DE3 cell (Invitrogen) in Luria Bertani media. Proteins were purified by affinity chromatography, using HisTrap FF columns (GE Healthcare). Protein fractions were further purified using Size Exclusion Chromatography, using a Superdex 200 16/600 column (GE Healthcare).

Size exclusion chromatography and multi-angle light scattering

In total, 100 μl samples of recombinant proteins, at concentrations of 1 mg/ml (NDP52-FL and CoCoA) and 5 mg/ml ($_N$NDP52, $_C$NDP52 and ZFs) were loaded onto a Superdex 200 (30 × 1 cm) analytical gel filtration column (GE Healthcare), equilibrated in 50 mM Tris pH7.5, 150 mM NaCl, 1 mM DTT and controlled by Waters 626 HPLC at room temperature. A Viskotek SEC-MALS 9 and Viscotek RI detector VE3580 (Malvern Panalytical) were used for detection. Analysis was performed using Omnisec software (Malvern Panalytical).

Dynamic light scattering

Dynamic light scattering measurements were performed at 20 °C, using a Zetasizer Nano ZS DLS system (Malvern Panalytical). Before measuring light scattering intensity at 90° angle, samples were centrifuged at 20,000 × g for 10 min. Analysis was undertaken using the Zetasizer software.

Microscale thermophoresis

Recombinantly purified protein constructs were labelled with RED-tris NTA dye (NanoTemper Technologies GmbH) in PBS to a concentration of 100 nM. A 20 μM stock of non-labelled protein was also prepared. This stock was used in a 16-step serial dilution in PBS buffer. For oligomerisation studies, 10 μM of protein was used as the highest ligand concentration, with Red-tris-NTA labelled protein kept at a final concentration of 50 nM for all reactions.

Reactions were incubated for 15 min at room temperature in the dark and loaded into Monolith NT.115 Capillaries (NanoTemper Technologies GmbH). Microscale thermophoresis measurements were performed using a Monolith NT.115 (NanoTemper Technologies GmbH) with 20%(RED) LED and high MST power. Binding assays were run as three independent experiments and the data were fitted using a K_D model with ligand-induced initial fluorescence change, as described by Jerabek-Willemsen et al.[49].

Circular dichroism

Recombinantly purified constructs were prepared in 50 mM Tris-HCl pH7.5, 150 mM NaCl. Circular dichroism spectra were obtained from 200 μl samples in a 1-mm cuvette, in a J175 spectropolarimeter from Jasco, with data collected at 0.5 nm intervals with averaging of 16 scans. For thermostability data, spectra were collected between 20 and 90 °C, and mean residue ellipticity values at 222 or 215 nm wavelength were fitted to a simple sigmoidal curve.

Electrophoretic mobility shift assay

Reactions were performed in a final volume of 30 μl, with 250 nM ds40 and increasing concentrations of NDP52-FL (between 0.5 and 3 μM) in 50 mM Tris-HCl pH 7.5, 50 mM NaCl and 3 mM MgCl₂. After 5 min incubation, reactions were supplemented with 3 μl of 30% (v/v) glycerol, loaded on a 3% agarose gel and run in Tris-Borate buffer at 60 V. Gels were incubated for 30 min in ethidium bromide, washed for 20 min with H₂O and visualised under UV light.

Nano-differential scanning fluorimetry

Thermostability of recombinantly produced proteins NDP52-FL, $_C$NDP52 and ZFs was assessed using nano-DSF, at protein concentrations 50, 75 and 100 μM, respectively. Samples were prepared in 50 mM Tris-HCl pH 7.5, 150 mM NaCl and loaded in nanoDSF Grade Standard Capillaries (NanoTemper Technologies GmbH), for NDP52-FL, or nanoDSF Grade High Sensitivity Capillaries (NanoTemper Technologies GmbH), for $_C$NDP52 and ZFs. Data were acquired using a Prometheus NT.48 (NanoTemper Technologies GmbH). Thermal denaturation of proteins was detected with heating in a linear thermal ramp (2 °C/min⁻¹) between 20 and 90 °C, with an excitation power of 60–90%. Temperature unfolding was detected by following fluorescence emission at 350 and 330 nm wavelength. Melting temperatures were determined as the maxima of the first derivative of the ratio 350 nm/330 nm, using NanoTemper software (NanoTemper Technologies GmbH).

Mass photometry

Before measurements, samples were centrifuged at 20,000 × g for 10 min. The samples were then diluted to 10 nM immediately prior to measurements in 50 mM Tris-HCl pH7.5, 150 mM NaCl. Measurements were performed on clean glass coverslips and recorded on the OneMP mass photometer (Refeyn Ltd) for 60 s. Each measurement was repeated at least three times. The recorded videos were analysed using DiscoverMP (Refeyn Ltd) to quantify protein binding events. The molecular weight was obtained by contrast comparison with known mass standards (BSA, Urease and IgG) measured on the same day.

Sample preparation for AFM

Preparation of DNA and protein samples for imaging was carried out as described fully in a published protocol[50]. Linear 339 bp DNA molecules (ds339) (Supplementary Table 3), NDP52-FL and CNDP52 were adsorbed onto freshly cleaved mica disks (diameter 5 mm, Agar Scientific, UK), separately and in combination, at room temperature using the cationic polymers, poly-L-lysine (PLL), 0.01% solution, MW 150,000–300,000 (Sigma-Aldrich) or poly-L-ornithine (PLO), 0.1 mg ml⁻¹ solution, MW 30,000–70,000 (Sigma-Aldrich)[51]. Briefly, 20 μl of either PLL or PLO solution was pipetted onto freshly cleaved mica to adsorb for 1 min. The cationic polymer coated surface was washed in a stream of MilliQ® ultrapure water, resistivity >18.2 MΩ to remove non-adsorbed PLL. To immobilise DNA, protein and DNA-protein complexes 20 μl of buffer was pipetted onto the treated mica and the DNA and protein samples were added to this buffer (Supplementary Table 4). All samples were adsorbed for 10 min followed by four washes in the same buffer to remove unbound DNA or protein. Alternatively, the samples were adsorbed onto freshly cleaved mica with divalent cations. Here 20 μl of buffer containing either MgCl₂ or

NiCl$_2$ was placed onto the mica and protein and/or ds339 DNA was added then the sample incubated for 30 min before washing four times in the same with nickel chloride buffer in both cases (Supplementary Table 4). Nickel buffer was used while imaging samples adsorbed using either Mg^{2+} or Ni^{2+} ions for greater imaging stability.

AFM imaging
All AFM measurements were performed in liquid following a published protocol[50]. Experiments were carried out in Bruker Nanoscope software PeakForce Tapping mode on a FastScan Dimension XR AFM (Bruker) using FastScan D AFM probes (Bruker). Continuous force–distance curves were recorded over 40 nm (PeakForce Tapping amplitude of 20 nm), at a frequency of 8 kHz with the tip-sample feedback set by PeakForce setpoints in the range of 5–12 mV as referenced from the force baseline resulting in peak forces of 40–100 pN. Images were recorded at 512 × 512 pixels to ensure a resolution ≥1 nm/pixel at line rates of 1–4 Hz.

AFM image processing
A modified version of TopoStats[52,53], a package of python scripts, was used to automate the processing of the AFM data and the analysis of the DNA and protein molecules. The code is freely available at https://github.com/AFM-SPM/TopoStats. Raw AFM data were processed using the "pygwytracing.py" script which utilises the Gwyddion "pygwy" module for automated image correction, molecule identification and morphological analysis.

Basic image processing was performed by the "editfile" function, using various Gwyddion processes to align and level data as well as correcting imaging artefacts. A gaussian filter ($\sigma=1$) of 1.5 pixels (1–2 nm) was applied to remove pixel errors and high-frequency noise.

Protein molecules were identified in the images of NDP52-FL and $_C$NDP52 using a combination of Gwyddion's automated masking protocols so that masks defined the positions of individual molecules in the images. These masked molecules within a flattened AFM image were identified using the "mask_outliers2" function, which masks data points with height values that deviate from the mean by a customisable multiplier of σ (with 3σ corresponding to a standard gaussian). This multiplier was optimised to select features based on their height in order to correctly mask the protein molecules and oligomers, and the values were 0.7σ for $_C$NDP52, 0.57σ for NDP52-FL, and 1.5σ for the selective masking of terminal domains in NDP52-FL. Example masked data created using these parameters is shown in Supplementary Fig. 6K. The Gwyddion functions "grains_remove_touching_border" and "grains_remove_by_size" were used to remove molecules that were cut off by the edge of the image and those that were <50 nm^2 respectively. Large aggregates and small contaminants were removed using the "removelargeobjects" and "removesmallobjects" functions respectively, which used the function "find_median_pixel_area" to determine the median size of the masked molecules then removed objects outside of a customisable size range based on this. The morphological properties of individual masked molecules were calculated for each image using the "grainanalysis" function, which uses the Gwyddion function "grains_get_values" to obtain statistical properties, including the length and width of the masked molecules. These are referred to in Gwyddion as "GRAIN_VALUE_MINIMUM_BOUND_SIZE" and "GRAIN_VALUE_MAXIMUM_BOUND_SIZE" respectively, and measure the maximum and minimum bounding sizes of the 2D mask for each molecule.

Each grain's values were appended to an array [appended_data] for morphological analysis of individual molecules from all images in a single directory. This array was converted to a pandas dataframe[54] using the "getdataforallfiles" function and saved out using the "savestats" function as ".json" and ".csv" files with the name of the directory in the original path.

Statistical analysis and plotting were performed using the "Plotting.py" script. This script uses the "importfromjson" function to import the.json file exported by "savestats" in pygwytracing.py and calculates various morphological properties from the masked molecules, including the maximum and minimum bounding size. Distributions were generated and plotted for the maximum and minimum bound sizes using the matplotlib[55] and seaborn libraries (https://doi.org/10.5281/zenodo.883859) within the function "plotdist2var" (https://zenodo.org/record/12710#.YfhFffXP1lY). The comparison of the domain height distributions (Supplementary Fig. 5A) was plotted with the function "plotviolin2var".

The binding position preference for NDP52-FL on ds339 was measured manually by loading the processed images into ImageJ[56] and using the measurement tools to determine the distance from the nearest end that the protein binding occurred. The numbers of monomers and oligomers were counted manually and the percentage of molecules in oligomeric states calculated.

DNA binding assays
FITC-labelled and unlabelled oligonucleotides were purchased from Sigma-Aldrich. For dsDNA preparation, equimolar concentrations of complementary ssDNA oligonucleotides were mixed[57]. A list of oligonucleotides is provided in Supplementary Table 3.

Recombinant NDP52 (NDP52-FL, $_C$NDP52, $_N$NDP52 and ZFs) and CoCoA constructs were titrated into 100 nM of FITC-labelled ssDNA 40 bp or dsDNA, 15 or 40 bp, in 50 mM Tris-HCl pH7.5, 150 mM NaCl, 1 mM DTT. Measurements were performed using a ClarioStar microplate reader (BMG Labtech). Fluorescence excitation was performed at 495 nm wavelength and emission spectra were measured between 515 and 570 nm wavelength, with fluorescence intensity values taken at 520 nm. Change in fluorescence was plotted in function of protein concentration, using three independent replicates for each experiment. Titration curves for NDP52-FL, CoCoA and $_N$NDP52 were fitted to a binding quadratic equation:

$$\begin{aligned}
&[\text{Protein.DNA}] \\
&= \frac{([\text{Protein}]_t + [\text{DNA}]_t + K_d) - \sqrt{([\text{Protein}]_t + [\text{DNA}]_t + K_d)^2 - 4[\text{Protein}]_t[\text{DNA}]_t}}{2}
\end{aligned}$$

(1)

For $_C$NDP52 and ZFs, a modified quadratic equation, accounting for a linear portion of the curve was used:

$$\begin{aligned}
&[\text{Protein.DNA}] \\
&= \frac{([\text{Protein}]_t + [\text{DNA}]_t + K_d) - \sqrt{([\text{Protein}]_t + [\text{DNA}]_t + K_d)^2 - 4[\text{Protein}]_t[\text{DNA}]_t}}{2} \\
&+ (a + (b[\text{DNA}]_t))
\end{aligned}$$

(2)

Small-angle X-ray scattering and ab initio envelope calculation
Recombinantly expressed and purified NDP52-FL in SAXS buffer (50 mM Tris-HCl pH7.5, 150 mM NaCl, 1 mM DTT) was used for SEC-SAXS experiments at a concentration of 5 mg/ml. NDP52-FL was analysed using a Superdex 200 increase 3.2/300 column, at a flow rate of 0.075 ml/min (Cytiva Life Sciences), using an Agilent 1200 HPLC system (Agilent LC). SEC-SAXS experiments were performed at the B21 Beamline, Diamond Light Source UK, by core facility staff. For SEC-SAXS analysis and envelope generation, ScÅtter software (Version J) was used in combination with the ATSAS package[58].

Nuclear isolation and extract preparation
Nuclear isolation was performed as previously described and characterised[15,59,60]. Briefly, HeLa cells were collected and washed once

with ice-cold PBS, then washed with ice cold Hypotonic buffer N (10 mM Hepes pH 7.5, 2 mM MgCl$_2$, 25 mM KCl supplemented with 1 mM PMSF, 1 mM DTT and Protease Inhibitors (Thermo Fisher Scientific)). Cells were resuspended in cold Hypotonic buffer N and incubated on ice for 1 h. Following this, cells were homogenised on ice using a glass Dounce homogeniser (Wheaton) and cell lysate was supplemented with sucrose solution to a final concentration of 220 mM, before centrifugation. The pellet, corresponding to isolated nuclei, was washed with cold Buffer N (10 mM Hepes pH 7.5, 2 mM MgCl2, 25 mM KCl, 250 mM sucrose, 1 mM PMSF, 1 mM DTT, supplemented with Protease Inhibitor Cocktail.

For nuclear extract preparation, nuclei were incubated in nuclear "Hypotonic lysis buffer" (10 mM Hepes pH 7.5, 2 mM MgCl$_2$, 25 mM KCl, 0.1%(V/V) Triton X-100, 0.1%(V/V) NP-40, supplemented with 1 mM PMSF, 1 mM DTT and Protease Inhibitors) for 1 h on ice. Lysed nuclei were then used for recombinant protein pull-downs.

Immunoblotting

Cell pellets from HeLa and MCF-7 cells, following NDP52 KD or control siRNA, were heat-denatured and resolved by SDS-PAGE. Membranes were probed against actin (Abcam, ab6276) and NDP52 by incubation with rabbit polyclonal primary antibody (1:2000 dilution, GeneTex, GTX115378) and, subsequently, a goat anti-rabbit antibody, coupled to horseradish peroxidase (1:15,000 dilution, Abcam, ab6721). Bands were visualised with ECL Western Blotting detection reagents (Invitrogen) using a ChemiDoc gel imager (Bio-Rad). For Ponceau S staining, membranes were incubated for 5 min in Ponceau S reagent (Sigma), washed three times with water and then imaged.

Co-immunoprecipitation

HeLa cells (non-treated or following α-amanitin treatment for 4 h) were collected and centrifuged at $500 \times g$ for 5 min at 4 °C. In total, 1×10^6 cells were used per co-immunoprecipitation assay. Each pellet was resuspended in 200 µl of Lysis buffer (10 mM Hepes pH7.5, 2 mM MgCl$_2$, 25 mM KCl, 0.1 mM DTT, 0.01 mM PMSF, 0.1%(V/V) Triton X-100, 0.1%(V/V) NP40 and supplemented with protease inhibitors). Cells were left in Lysis buffer, on ice, for 1 h. All samples were ple-cleared using protein A magnetic beads (Thermo Fisher Scientific) for 30 min at 4 °C with end-to-end rotation. Immunoprecipitation replicates were incubated overnight with NDP52 antibody (1:100 dilution) at 4 °C with end-to-end rotation. The following day, 50 µl of protein A magnetic beads, pre-equilibrated in Lysis buffer, were added to each sample, including the no-antibody controls, and incubated at 4 °C with end-to-end rotation for 2 h. Following this, beads were collected using a magnetic rack and washed three times with PBS. After removing all PBS, 50 µl of loading buffer (NuPAGE LDS sample buffer supplemented with 50 mM DTT) were added and samples were incubated at 95 °C for 10 min. Samples were then loaded on SDS-PAGE gels for in-gel protein digestion for LC-MS/MS.

Recombinant protein pull-downs

Following recombinant protein purification, 2.5 mg of protein, per replicate, were incubated with nuclear extract (2×10^6 nuclei per pull-down) at 4 °C with end-to-end rotation, for 4 h. Following this, 100 µl of Ni^{2+} magnetic beads (HisPur Ni-NTA Thermo Fisher), pre-equilibrated in equilibration buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 20 mM imidazole, 1 mM DTT), were added to each sample, including the no-protein controls, and incubated at 4 °C with end-to-end rotation for 2 h. Following this, beads were collected using a magnetic rack and washed three times with "low imidazole buffer" (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 40 mM imidazole, 1 mM DTT). Following this step, three elutions were performed using "high imidazole buffer" (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 400 mM imidazole, 1 mM DTT). Eluted samples were loaded on SDS-PAGE gels for in-gel protein digestion for LC-MS/MS.

In-gel digestion and LC-MS/MS

Following co-immunoprecipitation or recombinant protein pull-downs, samples were run only within the stacking portion of SDS-PAGE gels. Following this, gels were then stained, leaving a single band in the stacking portion of the gel, with all the protein content of each sample. Gel bands for each replicate were extracted, cut into 1 × 1 mm squares and transferred into clean 1.5 ml tubes. Gel particles were incubated with 50 mM ammonium bicarbonate and acetonitrile in a 1:1 ratio at room temperature for 15 min and then centrifuged at $8000 \times g$ for 60 s and the supernatant discarded. Samples were then incubated in acetonitrile for 15 min and centrifuged to remove supernatant. Gel particles were then incubated in 10 mM DTT, 50 mM ammonium bicarbonate and incubated at 56 °C for 30 min. Following this, the samples were centrifuged, the supernatant removed, and samples briefly incubated in acetonitrile until gel pieces shrunk. Samples were centrifuged and 55 mM iodoacetamide in 50 mM ammonium bicarbonate solution was added so that all gel particles were submerged. Samples were incubated in the dark, at room temperature, for 20 min. Following centrifugation, the supernatant was removed and washed in 50 mM ammonium bicarbonate solution:acetonitrile (1:1) for 15 min, followed by 50 mM ammonium bicarbonate for 15 min and acetonitrile for 15 min (between each step samples were centrifuged and supernatant discarded). Samples were then centrifuged and all liquid removed.

For tryptic digestion, gel particles were incubated in digestion buffer (25 mM ammonium bicarbonate, 10%(V/V) acetonitrile and 10 ng/µl Trypsin (Sigma-Aldrich, EMS0006)) for 30 min on ice. Digestion buffer was replenished as needed during this process to ensure gel particles were covered in solution. After 30 min, excess digestion buffer was removed from each sample and replaced with 25 mM ammonium bicarbonate, 10%(V/V) acetonitrile solution. Samples were incubated overnight at room temperature.

The next day, acetonitrile was added to each tube and samples were sonicated in an ultrasound bath for 15 min. Samples were then centrifuged and the supernatant, containing digested protein for mass spectrometry, was transferred into clean 1.5 ml tubes. 50%(v/v) acetonitrile and 5%(v/v) formic acid solution was added to gel particles and these were sonicated again. The supernatant from this step was combined with the previously collected sample into the same tube. Extracted protein samples were vacuum dried and resuspended in 10% (v/v) acetonitrile, 0.1%(v/v) trifluoroacetic acid for nanoLC-MS.

Peptides were separated on a HSS T3 Acquity column (Waters) 75 µm i.d. × 15 cm (1.8 µm, 100 A) using an Acquity M-Class UPLC (Waters), elution was performed with a linear gradient from 3 to 40% B over 40 min (solvent A = 0.1% formic acid, solvent B = 0.1% formic acid, acetonitrile) and the eluate directed via a nanospray source to a Synapt G2-Si (Waters) with data collected in UDMSe mode. Mass spectrometry data were imported into the software package Progenesis QI (Non-Linear Dynamics) and searched against a protein database using an MSe Search algorithm with an FDR set to 4%. Progenesis QI software (Waters) provided quality control information and quantification of peptides. The peptides were assigned using the "human proteome including enolase v5 2017" from UNIPROT as a reference library, accounting for trypsin cleavage, carbamidomethyl modifications to cysteine residues and methionine oxidation. Maximum protein mass was set to 500 kDa with a maximum of one missed cleavage allowed.

For peptide and protein assignments, a minimum of three fragments per peptide was required and a minimum of five fragments per protein. All assigned proteins contained at least one unique peptide. Following PCA analysis, replicates that did not cluster were excluded. Hits with a log$_2$FC > 1 and ANOVA $p < 0.05$, compared to controls (protein A or Ni^{2+} magnetic beads-pull-downs) were considered for further analysis. Protein hits were submitted to Gene Ontology analysis using Gene Ontology Resource (https://geneontology.org).

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE[61] partner repository with the data identifier PXD030238.

Electron microscopy

Cells grown on Aclar membrane (Agar Scientific) were fixed in 2% (w/v) formaldehyde and 0.5% glutaraldehyde in CAB (100 mM Sodium cacodylate buffer pH7.2) for 2 h at room temperature. The sample was washed 2 × 10 min in CAB. Cells were dehydrated by incubation in an ethanol gradient, 50% ethanol for 10 min, 70% ethanol overnight, and 90% ethanol for 10 min followed by three 10-min washes in 100% dry ethanol. Cells were then suspended in LR White resin medium grade (London Resin Company) for 4 h and then in fresh LR White resin overnight. Following 2 × 4-h changes in fresh LR White resin samples were placed in sealed gelatine capsules and were polymerised upright at 60 °C for 20 h. Ultrathin sections were cut using a Leica EM UC7 ultramicrotome equipped with a diamond knife (DiATOME 45°). Sections (80 nm) were collected on uncoated 400-mesh gold grids.

Samples were blocked in a 20 μl drop of 2% BSA in TBST (20 mM Tris, 500 mM NaCl, 0.1% BSA and 0.05% Tween 20) at room temperature for 30 min. Grids were then transferred directly into a 20 μl drop of Rabbit anti-NDP52 (1:200, Genetex GTX115378) TBST and incubated for 1 h. Grids were washed in 6 × TBST. Grids were then moved into a drop of goat anti-rabbit IgG 5 nm gold (British Biocell International) diluted 1:50 and then moved to a fresh drop of the same antibody and incubated for 30 min. Excess antibody was removed by washing in 6 × 20 μl drops of TBST and 6 × 20 μl drops of milliQ water and dried.

Grids were stained for 15 min in 4.5% uranyl acetate in 1% acetic acid solution and then washed in 6 × 20 μl drops of milliQ water. Grids were then stained with Reynolds lead citrate for 3 min and washed in 6 × 20 μl drops of milliQ water. Electron microscopy was performed using a JEOL-1230 transmission electron microscope operated at an accelerating voltage of 80 kV equipped with a Gatan One View digital camera.

Reproducibility

All experiments were performed as three independent sets. Data were then averaged or combined to complete datasets.

Graphics

Unless stated, data fitting and plotting was performed using GraphPad Prism 9 and Grafit Version 5 (Erithacus Software Ltd). Cartoons were generated using the BioRender software.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The data that support this study are available from the corresponding author upon reasonable request. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium with the data identifier PXD030238. RNA-Seq data generated in the course of this study were deposited in the Gene Expression Omnibus (GEO) database under the accession number GSE188567. Source data are provided with this paper.

References

1. Korioth, F., Gieffers, C., Maul, G. G. & Frey, J. Molecular characterization of NDP52, a novel protein of the nuclear domain 10, which is redistributed upon virus infection and interferon treatment. *J. Cell Biol.* **130**, 1–13 (1995).
2. Sternsdorf, T., Jensen, K., Zuchner, D. & Will, H. Cellular localization, expression, and structure of the nuclear dot protein 52. *J. Cell Biol.* **138**, 435–448 (1997).
3. Morriswood, B. et al. T6BP and NDP52 are myosin VI binding partners with potential roles in cytokine signalling and cell adhesion. *J. Cell Sci.* **120**, 2574–2585 (2007).
4. Mostowy, S. et al. p62 and NDP52 proteins target intracytosolic Shigella and Listeria to different autophagy pathways. *J. Biol. Chem.* **286**, 26987–26995 (2011).
5. Tumbarello, D. A. et al. Autophagy receptors link myosin VI to autophagosomes to mediate Tom1-dependent autophagosome maturation and fusion with the lysosome. *Nat. Cell Biol.* **14**, 1024–1035 (2012).
6. Kim, B. W., Hong, S. B., Kim, J. H., Kwon, D. H. & Song, H. K. Structural basis for recognition of autophagic receptor NDP52 by the sugar receptor galectin-8. *Nat. Commun.* **4**, 1613 (2013).
7. Xie, X. et al. Molecular basis of ubiquitin recognition by the autophagy receptor CALCOCO2. *Autophagy* **11**, 1775–1789 (2015).
8. Dos Santos, A. et al. Binding partners regulate unfolding of myosin VI to activate the molecular motor. *Biochem. J.* **479**, 1409–1428 (2022).
9. Fili, N. & Toseland, C. P. Unconventional myosins: how regulation meets function. *Int. J. Mol. Sci.* **21** https://doi.org/10.3390/ijms21010067 (2019).
10. Tumbarello, D. A. et al. The autophagy receptor TAX1BP1 and the molecular motor myosin VI are required for clearance of salmonella typhimurium by autophagy. *PLoS Pathog.* **11**, e1005174 (2015).
11. Kim, J. H., Li, H. & Stallcup, M. R. CoCoA, a nuclear receptor coactivator which acts through an N-terminal activation domain of p160 coactivators. *Mol. Cell* **12**, 1537–1549 (2003).
12. Nthiga, T. M. et al. CALCOCO1 acts with VAMP-associated proteins to mediate ER-phagy. *EMBO J.* **39**, e103649 (2020).
13. Stefely, J. A. et al. Mass spectrometry proteomics reveals a function for mammalian CALCOCO1 in MTOR-regulated selective autophagy. *Autophagy* **16**, 2219–2237 (2020).
14. Cook, A. W., Gough, R. E. & Toseland, C. P. Nuclear myosins—roles for molecular transporters and anchors. *J. Cell Sci.* **133** https://doi.org/10.1242/jcs.242420 (2020).
15. Fili, N. et al. NDP52 activates nuclear myosin VI to enhance RNA polymerase II transcription. *Nat. Commun.* **8**, 1871 (2017).
16. Fili, N. et al. Competition between two high- and low-affinity protein-binding sites in myosin VI controls its cellular function. *J. Biol. Chem.* **295**, 337–347 (2020).
17. Dos Santos, A., Gough, R. E., Wang, L. & Toseland, C. P. Measuring nuclear organization of proteins with STORM imaging and cluster analysis. *Methods Mol. Biol.* **2476**, 293–309 (2022).
18. Carmo-Fonseca, M. The contribution of nuclear compartmentalization to gene regulation. *Cell* **108**, 513–521 (2002).
19. Cook, P. R. A model for all genomes: the role of transcription factories. *J. Mol. Biol.* **395**, 1–10 (2010).
20. Meister, P. et al. Nuclear factories for signalling and repairing DNA double strand breaks in living fission yeast. *Nucleic Acids Res.* **31**, 5064–5073 (2003).
21. Schrank, B. R. et al. Nuclear ARP2/3 drives DNA break clustering for homology-directed repair. *Nature* **559**, 61–66 (2018).
22. Sporbert, A., Gahl, A., Ankerhold, R., Leonhardt, H. & Cardoso, M. C. DNA polymerase clamp shows little turnover at established replication sites but sequential de novo assembly at adjacent origin clusters. *Mol. Cell* **10**, 1355–1365 (2002).
23. Pageon, S. V., Nicovich, P. R., Mollazade, M., Tabarin, T. & Gaus, K. Clus-DoC: a combined cluster detection and colocalization analysis for single-molecule localization microscopy data. *Mol. Biol. Cell* **27**, 3627–3636 (2016).
24. Ciani, B. et al. Molecular basis of coiled-coil oligomerization-state specificity. *Proc. Natl Acad. Sci. USA* **107**, 19850–19855 (2010).
25. Lupas, A. N. & Gruber, M. The structure of alpha-helical coiled coils. *Adv. Protein Chem.* **70**, 37–78 (2005).

26. Cassandri, M. et al. Zinc-finger proteins in health and disease. *Cell Death Discov.* **3**, 17071 (2017).

27. Ellinghaus, D. et al. Association between variants of PRDM1 and NDP52 and Crohn's disease, based on exome sequencing and functional studies. *Gastroenterology* **145**, 339–347 (2013).

28. Till, A. et al. Autophagy receptor CALCOCO2/NDP52 takes center stage in Crohn disease. *Autophagy* **9**, 1256–1257 (2013).

29. Bushnell, D. A., Cramer, P. & Kornberg, R. D. Structural basis of transcription: alpha-amanitin-RNA polymerase II cocrystal at 2.8 A resolution. *Proc. Natl Acad. Sci. USA* **99**, 1218–1222 (2002).

30. Cremer, T. et al. Chromosome territories—a functional nuclear landscape. *Curr. Opin. Cell Biol.* **18**, 307–316 (2006).

31. Leake, M. C. Transcription factors in eukaryotic cells can functionally regulate gene expression by acting in oligomeric assemblies formed from an intrinsically disordered protein phase transition enabled by molecular crowding. *Transcription* **9**, 298–306 (2018).

32. Wollman, A. J. et al. Transcription factor clusters regulate genes in eukaryotic cells. *eLife* **6**, e27451 (2017).

33. Cho, W. K. et al. Mediator and RNA polymerase II clusters associate in transcription-dependent condensates. *Science* **361**, 412–415 (2018).

34. Cho, W.-K. et al. RNA polymerase II cluster dynamics predict mRNA output in living cells. *eLife* **5**, e13617 (2016).

35. Cho, W.-K. et al. Super-resolution imaging of fluorescently labeled, endogenous RNA polymerase II in living cells with CRISPR/Cas9-mediated gene editing. *Sci. Rep.* **6**, 35949 (2016).

36. Hari-Gupta, Y. et al. Myosin VI regulates the spatial organisation of mammalian transcription initiation. *Nat. Commun.* **13**, 1346 (2022).

37. Kim, J. H. et al. CCAR1, a key regulator of mediator complex recruitment to nuclear receptor transcription complexes. *Mol. Cell* **31**, 510–519 (2008).

38. Kim, J. H., Yang, C. K. & Stallcup, M. R. Downstream signaling mechanism of the C-terminal activation domain of transcriptional coactivator CoCoA. *Nucleic Acids Res.* **34**, 2736–2750 (2006).

39. Imbalzano, A. N., Kwon, H., Green, M. R. & Kingston, R. E. Facilitated binding of TATA-binding protein to nucleosomal DNA. *Nature* **370**, 481–485 (1994).

40. Vargas, J. N. S. et al. Spatiotemporal control of ULK1 activation by NDP52 and TBK1 during selective autophagy. *Mol. Cell* **74**, 347–362.e6 (2019).

41. Thurston, T. L., Ryzhakov, G., Bloor, S., von Muhlinen, N. & Randow, F. The TBK1 adaptor and autophagy receptor NDP52 restricts the proliferation of ubiquitin-coated bacteria. *Nat. Immunol.* **10**, 1215–1221 (2009).

42. Verlhac, P. et al. Autophagy receptor NDP52 regulates pathogen-containing autophagosome maturation. *Cell Host Microbe* **17**, 515–525 (2015).

43. Liang, S. H. & Clarke, M. F. Regulation of p53 localization. *Eur. J. Biochem.* **268**, 2779–2783 (2001).

44. Mercurio, F. & Manning, A. M. NF-κB as a primary regulator of the stress response. *Oncogene* **18**, 6163–6171 (1999).

45. Shim, M. S., Nettesheim, A., Hirt, J. & Liton, P. B. The autophagic protein LC3 translocates to the nucleus and localizes in the nucleolus associated to NUFIP1 in response to cyclic mechanical stress. *Autophagy* **16**, 1248–1261 (2020).

46. Abrahamsson, S. et al. Fast multicolor 3D imaging using aberration-corrected multifocus microscopy. *Nat. Methods* **10**, 60–63 (2013).

47. Koulouras, G. et al. EasyFRAP-web: a web-based tool for the analysis of fluorescence recovery after photobleaching data. *Nucleic Acids Res.* **46**, W467–W472 (2018).

48. Ge, S. X., Son, E. W. & Yao, R. iDEP: an integrated web application for differential expression and pathway analysis of RNA-Seq data. *BMC Bioinformatics* **19**, 534 (2018).

49. Jerabek-Willemsen, M. et al. MicroScale thermophoresis: interaction analysis and beyond. *J. Mol. Struct.* **1077**, 101–113 (2014).

50. Haynes, P. J., Main, K. H. & Pyne, A. L. B. Atomic force microscopy of DNA and DNA-protein interactions. protocols.io. https://doi.org/10.17504/protocols.io.bncemate (2020).

51. Podesta, A. et al. Atomic force microscopy study of DNA deposited on poly L-ornithine-coated mica. *J. Microsc.* **215**, 236–240 (2004).

52. Beton, J. G. et al. TopoStats—a program for automated tracing of biomolecules from AFM images. *Methods* **193**, 68–79 (2021).

53. Du, M. & Pyne, A. L. P. TopoStats Software. ORCA. https://doi.org/10.15131/shef.data.21444687 (2022).

54. van der Walt, S., Colbert, S. & Varoquaux, G. The NumPy array: a structure for efficient numerical computation. *Comput. Sci. Eng.* **13** https://doi.org/10.1109/MCSE.2011.37 (2011).

55. Hunter, J. D. Matplotlib: a 2D graphics environment. *Comput. Sci. Eng.* **9** https://doi.org/10.1109/MCSE.2007.55 (2007).

56. Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH image to ImageJ: 25 years of image analysis. *Nat. Methods* **9**, 671–675 (2012).

57. Toseland, C. P. & Webb, M. R. Fluorescence tools to measure helicase activity in real time. *Methods* **51**, 259–268 (2010).

58. Manalastas-Cantos, K. et al. ATSAS 3.0: expanded functionality and new tools for small-angle scattering data analysis. *J. Appl. Crystallogr.* **54**, 343–355 (2021).

59. Lherbette, M. et al. Atomic force microscopy micro-rheology reveals large structural inhomogeneities in single cell-nuclei. *Sci. Rep.* **7**, 8116 (2017).

60. Dos Santos, A. et al. DNA damage alters nuclear mechanics through chromatin reorganization. *Nucleic Acids Res.* **49**, 340–353 (2021).

61. Perez-Riverol, Y. et al. The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res.* **47**, D442–D450 (2019).

Acknowledgements

Author contributions

A.d.S. and C.P.T. conceived the study. A.d.S., D.E.R., A.L.B.P., Y.H.-G. and C.P.T. designed experiments. A.d.S. and C.P.T. performed single-molecule imaging experiments. A.d.S., D.E.R. and A.L.B.P. carried out AFM imaging and A.d.S., D.E.R., M.D. and A.L.B.P. carried out AFM analysis. Imaging was supported by L.W., M.M.-F., T.-L.C. and J.A. A.d.S. and Y.H.-G. performed and analysed the genomics experiments. A.d.S., H.M., K.P., S.Y.Z.R. and F.L. expressed, purified and performed experiments with recombinant proteins. L.W., J.A., A.d.S. and C.P.T. contributed to single-molecule data analysis. A.d.S., A.S., P.J.I.E. and K.H. performed and analysed the mass spectrometry experiments. D.L. performed and analysed autophagy assays. I.B. prepared samples and

performed electron microscopy. C.P.T. supervised the study. A.d.S. and C.P.T. wrote the manuscript with comments from all authors.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-023-38572-9.

Correspondence and requests for materials should be addressed to Christopher P. Toseland.

Peer review information *Nature Communications* thanks Tokuko Haraguchi and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.