Article

# Enhancer–promoter contact formation requires RNAPII and antagonizes loop extrusion

Shu Zhang[1,2], Nadine Übelmesser[1,2], Mariano Barbieri[1] & Argyris Papantonis [1]✉

Homotypic chromatin interactions and loop extrusion are thought to be the two main drivers of mammalian chromosome folding. Here we tested the role of RNA polymerase II (RNAPII) across different scales of interphase chromatin organization in a cellular system allowing for its rapid, auxin-mediated degradation. We combined Micro-C and computational modeling to characterize subsets of loops differentially gained or lost upon RNAPII depletion. Gained loops, extrusion of which was antagonized by RNAPII, almost invariably formed by engaging new or rewired CTCF anchors. Lost loops selectively affected contacts between enhancers and promoters anchored by RNAPII, explaining the repression of most genes. Surprisingly, promoter–promoter interactions remained essentially unaffected by polymerase depletion, and cohesin occupancy was sustained. Together, our findings reconcile the role of RNAPII in transcription with its direct involvement in setting-up regulatory three-dimensional chromatin contacts genome wide, while also revealing an impact on cohesin loop extrusion.

Genomic functions like gene expression and DNA replication require a dynamic three-dimensional (3D) architecture of interphase chromatin[1,2]. Work in the last decade, combining genome wide chromosome conformation capture assays with the removal of different chromatin-binding proteins, has attributed key hallmarks of this 3D architecture to the interplay between the insulator factor CTCF and the ring-shaped cohesin complex[3,4]. Three-dimensional chromatin domains (from Mbp-sized 'topologically associating domains' or TADs[5] to kbp-sized 'loop domains[6]') are insulated from one another via CTCF-demarcated boundaries, while the chromatin in these domains is actively extruded into loops by cohesin[7–9]. Removing CTCF from chromatin leads to insulation loss at domain boundaries[10], while cohesin removal eliminates CTCF-anchored loops[6,11]. The physical interaction of an extruding cohesin complex with two convergently oriented CTCF-bound sites, and its stabilization together via STAG proteins, determines the length of loops genome wide[12–14].

Loops anchored at CTCF-bound sites appear as prominent dot-like features off the diagonal of high-resolution Hi-C contact maps[14]. These dots disappear in cells where the cohesin-loading factor, NIPBL, is eliminated[15] but multiply in cells lacking the cohesin-release factor, WAPL[16]. Thus, loop formation arises from regulated cohesin 'load-unload' cycles. However, recent live-cell imaging of the mouse *Fbn2* locus showed that full looping is rarely achieved and that, most of the time, cohesin-extruded loops within an active domain only partially form and do not bring the two CTCF anchors into proximity[17]. This can be explained by the notion that 3D genome architecture results from the antagonistic interplay of loop extrusion with homotypic (that is, active-to-active or inactive-to-inactive) compartmentalization of chromatin domains[18,19]. Then, one would predict that RNA polymerase II (RNAPII), a potent molecular motor capable of translocating and bridging DNA[20,21], influences 3D genome architecture via both physical interactions[22] and transcription[23].

Earlier studies in this direction have shown that pharmacological inhibition of transcription could neither dissolve nuclear subcompartments formed by active RNAPII[24] nor long-range contacts between genes and enhancers in select loci[25]. Similarly, TAD formation

[1]Institute of Pathology, University Medical Center Göttingen, Göttingen, Germany. [2]These authors contributed equally: Shu Zhang, Nadine Übelmesser. ✉e-mail: argyris.papantonis@med.uni-goettingen.de

in zygotes does not rely on transcriptional genome activation[26] and RNase A treatment of cell nuclei does not compromise TAD structure but does eliminate specific enhancer–promoter interactions[27,28]. On the other hand, Hi-C maps generated upon depletion of Mediator complex subunits or inhibition of RNAPII elongation while reexpressing cohesin in RAD21-depleted cells had no discernible effect on CTCF loop formation[6,29]. This also held true when depleting the basal transcription factor TAF12 or RNAPII and using promoter-capture Hi-C or Hi-C, respectively[30,31]. Moreover, haploid human cells depleted of Mediator could not sustain a transcription-permissive chromatin architecture[32]. However, even kbp-resolution Hi-C contact maps do not prominently feature loops other than those anchored by CTCF. This was remedied by the introduction of Micro-C, a Hi-C variant using micrococcal nuclease (MNase) to fragment chromatin and reveal tens of thousands of transcription-based loops along mammalian chromosomes[33,34]. In fact, its capture-based adaptation, Micro-Capture-C (MCC), allowed the mapping of 3D contacts between different *cis*-regulatory elements at near-base-pair resolution[35]. Still, pharmacological inhibition of transcription coupled to Micro-C did not affect looping and only reduced by ~1.25-fold the signal of 'stripes', which are thought to indicate asymmetric extrusion of loops[33].

This and other such data highlight the need for a conclusive dissection of whether and how core components of the transcriptional apparatus, like RNAPII and Mediator[36,37], contribute to the formation of 3D chromatin contacts. To this end, we applied Micro-C to a diploid human cell line allowing for the near-complete auxin-mediated degradation of the largest subunit of RNAPII, RPB1 (refs. [38,39]). We identified thousands of transcription-anchored and CTCF loops changing upon RNAPII depletion. We combined experiments with in silico models to interpret the interplay between cohesin loop extrusion and RNAPII-mediated looping.

## Results

### Transcription-based architecture of human chromosomes
Previously, we used in situ Hi-C to identify changes in 3D genome architecture upon RNAPII depletion from human diploid cells[39]. For these Hi-C experiments, we used asynchronous G1-sorted DLD-1 cells allowing for the quantitative degradation of RNAPII upon auxin addition for 14 h[37,39]. Effects at the level of TADs and compartments were small (as also seen after RNAPII depletion in asynchronous mESCs[30]), but we did identify ~800 CTCF loops that emerged in RNAPII-depleted cells and were substantially larger than the loops in untreated cells[39]. However, interactions between RNAPII-bound sites were scarce in that data, hence our incomplete understanding of how RNAPs contribute to 3D chromatin folding.

To address this and obtain a comprehensive view of the transcription-based architecture of human cells, we performed Micro-C. We generated contact maps containing >1.25 billion pairwise interactions (Supplementary Table 1), which revealed fine intradomain architecture compared to matching Hi-C data (of comparable sequencing depth; Fig. 1a,b). This detailed view of 3D chromatin folding allowed the detection of 31,913 loops, encompassing >80% of the loops detected by Hi-C (Fig. 1c). The anchors of these loops mapped predominantly to the A (active) compartment, suggesting that multiple RNAPII-anchored loops could be detected (Fig. 1d). Indeed, when we stratified these 31,913 loops by the presence or absence of RNAPII and/or H3K27ac or CTCF at their anchors, ~25% could be classified as transcription anchored. This increases to 40% once we consider anchors that have CTCF in addition to RNAPII/H3K27ac. Moreover, the 8,178 loops shared between Hi-C and Micro-C are featured more prominently in Micro-C contact maps (Fig. 1e). Finally, our highly resolved contact maps allowed the detection of thousands of cases where directional loop extrusion gives rise to stripes emanating by a CTCF and a transcriptional anchor (Fig. 1f). Thus, Micro-C allows us to study how transcription-anchored 3D interactions are remodeled upon RNAPII depletion.

## New CTCF-anchored loops emerge after RNAPII depletion
To ask how RNAPs affect 3D interactions, we used the 14-h auxin treatment determined in ref. [39], and generated Micro-C data in the presence or absence of RNAPII. Under these conditions, the vast majority of RNAPII was depleted from chromatin and led to decreased H3K27ac and cohesin levels genome wide (assessed using CUT&Tag; Extended Data Fig. 1a,b). We first used our Micro-C data to assess changes in nucleosome positioning after RNAPII depletion. For example, nucleosomes around CTCF sites became markedly more ordered; this did not come at the expense of CTCF binding to its cognate motifs but did result in a more focal ATAC-seq signal, that is, to locally constrained accessibility (Extended Data Fig. 1c). Decrease in chromatin accessibility was also observed at gene promoters and enhancers genome wide (Extended Data Fig. 1d). Notably, these changes were not a result of reduced availability of proteins like CTCF or cohesin as their levels on chromatin remained unaffected (Extended Data Fig. 1e), consistent with what we showed for this system before[39].

We next surveyed Micro-C contact maps to discover the widespread emergence of new and longer loops (Fig. 2a). These new loops typically arose in and around domains with active genes that became silenced upon RNAPII depletion. A total of 11,032 loops were gained or strengthened to surpass the detection threshold in RNAPII-depleted data. These were substantially longer than either transcription-or CTCF-anchored loops of control cells (Fig. 2b), and involved at least one CTCF-bound anchor in ~75% of cases, as well as increased local insulation (Fig. 2c–e). Looking into loops that have one 'CTCF only' anchor (for example, the left one) and one 'RNAPII only' (for example, the right one), we found that they specifically rewired the latter. From a total of 1,134 such loops, 74% rewired to a new anchor further downstream that almost invariably contained a CTCF-bound site (Fig. 2f). Rewiring often gave rise to nested loop structures (that is, 795 unchanged CTCF anchors gave rise to 1,134 new loops; Fig. 2a,f). The orientation of CTCF motifs in the new anchors was convergent with respect to that in the unchanged anchor (Fig. 2f). Interestingly, new CTCF anchors were disproportionately located at the TSSs and bodies of genes that were longer than average (Extended Data Fig. 1f) and became depleted of active RNAPs (Fig. 2g). In the absence of RNAPII and transcription, these anchors obtained more canonically spaced nucleosomes around them (Fig. 2h) but showed less reduction in SMC1A occupancy (Fig. 2i) than what was seen genome wide (Fig. 2a and Extended Data Fig. 1b,c,e).

Finally, we asked whether any loops gained after RNAPII depletion form via H3K27me3-mediated interactions (there were hints of ~200 such loops in our previous Hi-C data[39]). Despite no discernible changes of H3K27me3 levels in CUT&Tag (Extended Data Fig. 2a) and chromatin fractionation blots in control and auxin-treated cells (Extended Data Fig. 1e), 2,197 new loops with H3K27me3 peaks in at least one anchor arose. This increased to 3,184 when we considered anchors with H3K27me3 peaks in the next 5-kbp bin (Extended Data Fig. 2b,c). Such new loops typically emerged in bundles within facultative heterochromatin domains, often without CTCF association (Extended Data Fig. 2a). Therefore, our Micro-C data now explain that thousands of new and longer loops emerge after RNAPII depletion via CTCF- and Polycomb-driven interactions.

## RNAPII depletion leads to selective loss of enhancer loops
Next, we asked which loops and contacts are lost or weakened below the detection threshold upon RNAPII depletion. Lost loops were almost always found within CTCF loop-domain or TAD structures (Fig. 3a and Extended Data Fig. 3a). This is in line with transcription-anchored loops being the smallest (Fig. 3b); with the overall reduced contact frequency of loci separated by <1 Mbp upon RNAPII depletion (Extended Data Fig. 3b); and with gene regulatory domains being encompassed by CTCF loops[14]. The anchors of these 5,332 lost loops were substantially less likely to contain CTCF than those of unchanged loops (<25% have CTCF at both anchors; Fig. 3c). Following stratification
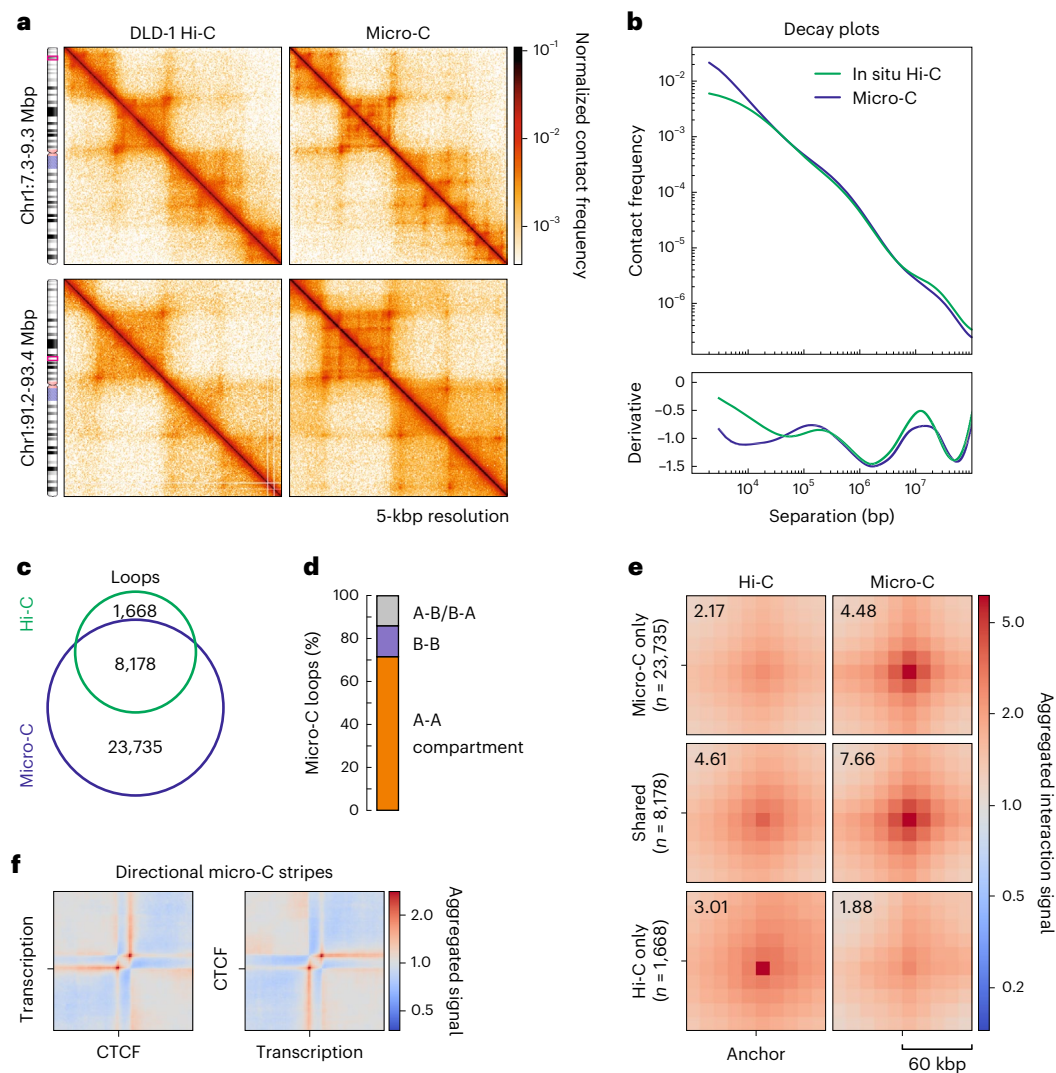
**Fig. 1 | Micro-C enhances high-resolution views of 3D genome folding.**
**a**, Comparison of exemplary Hi-C (left) and Micro-C 5-kbp resolution contact maps (right) from DLD1-mAID-RPB1 cells in two 2-Mbp segments from chr1. **b**, Hi-C and Micro-C interaction frequencies decaying as a function of genomic distance (top) and their first derivative (bottom). **c**, Venn diagram showing overlap of loops detected in Hi-C and Micro-C contact maps. **d**, Bar plot showing percent of Micro-C loops with anchors in the A- or B-compartment. **e**, Aggregate plots showing Hi-C (left) and Micro-C signal (right) around shared and unique loops. **f**, Aggregate Micro-C signal of stripes with one CTCF and one transcriptional anchor.

into promoter–promoter (P–P) and enhancer–promoter loops (E–P), we discovered that enhancer-anchored ones were most sensitive to RNAPII depletion (Fig. 3d). Moreover, E–P loops that did not contained CTCF in either anchor were almost fully lost, compared to those that did (Extended Data Fig. 3c). This selective loss was reflected in the reduced cohesin occupancy at E–P compared to P–P loop anchors (that was even more pronounced at the 590 superenhancers; Fig. 3e and Extended Data Fig. 3d). Surprisingly, much like their looping propensity, H3K27ac levels around promoters remained unaffected, while those at enhancers dropped by >50% (Extended Data Fig. 3e). In total, ~900 E–P loop domains (together spanning >700 Mbp) dissolved upon RNAPII depletion, which involved >40% (557/1,360) of all genes substantially downregulated upon auxin treatment ($|\log_2 FC| > 2$, $P_{adj} < 0.05$; RNA-sequencing (RNA-seq) data from ref. 39).

We also stratified loops not according to the type of elements in their anchors (that is, enhancers versus promoters), but on whether their anchors were marked by CTCF, transcription or both. Fifty-nine percent of loops lost upon RNAPII depletion were anchored solely by RNAPII/H3K27ac at both anchors, while CTCF-only loops were only 18% of the total number lost (Fig. 3f). For an additional 1,675 loops,

where only one anchor could be annotated, transcription anchors also predominated. Notably, CTCF-bound anchors displayed less reduction in cohesin occupancy than transcription-only anchors (that is, 28–35% compared to 48%; Fig. 3g). Together, we found that transcription-anchored loops are most sensitive to RNAPII depletion. This is selective for loops involving enhancers in at least one anchor and moderately influenced by CTCF presence.

Finally, even Micro-C features that at first appeared unaffected by auxin treatment did respond to RNAPII depletion. For instance, loop-like signal at the edges of stripes was enhanced in the absence of RNAPII (Extended Data Fig. 3f), likely due to the weakening of transcription in one of the anchors. Also, looking more carefully into loops that did not seem to change between the two conditions, we discovered that many rewired one anchor by <20 kbp (that is, by <4 bins in 5 kbp-resolution contact data). When we used control loop coordinates and auxin-treated Micro-C signal to plot aggregate plots, the signal appeared weaker. However, when we used matching auxin-treated coordinates (that is, shifted by <4 bins for the rewired anchor) and auxin-treated Micro-C signal, we noticed strengthening of these loops too (Extended Data Fig. 3g). This suggests that RNAPII
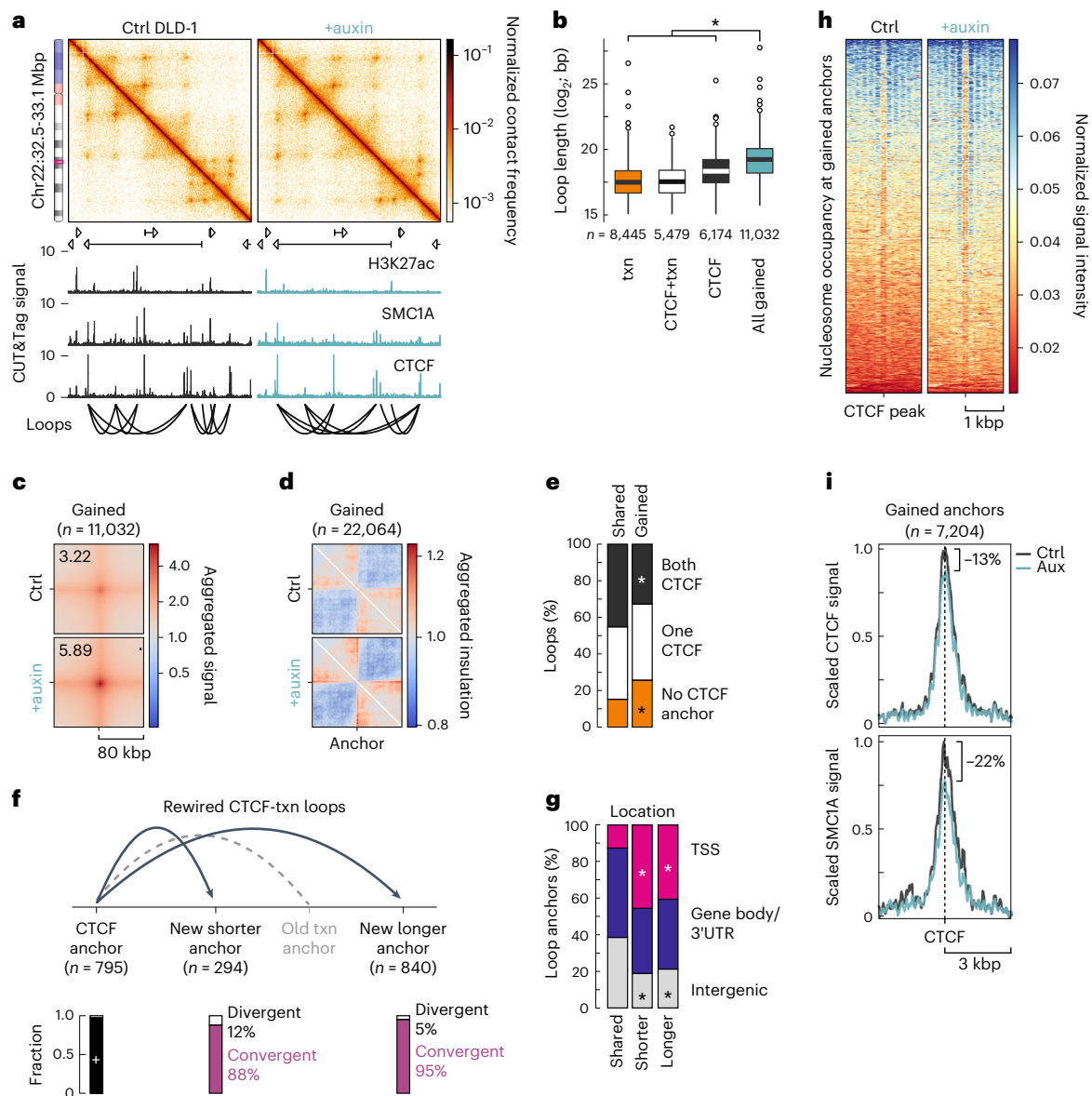
**Fig. 2 | Loops forming upon RNAPII degradation engage new CTCF anchors.**
**a**, Micro-C contact maps from control (left) and auxin-treated cells (right) showing emerging loop at 2-kbp resolution in an exemplary genomic region of chr22 aligned to H3K27ac, CTCF, and SMC1A CUT&Tag tracks. Gene models are depicted by arrows, and loops for each condition by spider plots (bottom). **b**, Box plots of the lengths of control loops with transcription-, CTCF- or transcription/CTCF anchors, and loops gained upon auxin treatment. In the plots, center lines represent the median value, box limits the 25th and 75th percentiles, and whiskers extend 1.5× the box's interquartile range. *$P < 0.01$, two-sided Wilcoxon–Mann–Whitney test. **c**, Aggregate plots for loops gained in auxin-treated cells.

**d**, Plots of mean insulation at the anchors of loops gained in auxin-treated cells. **e**, Bar plots showing percent of loops with none, one or two CTCF anchors that are shared or gained by auxin-treated cells. *$P < 0.01$, two-sided Fischer's exact test. **f**, Bar plots showing the fraction of convergent versus divergent CTCF motifs in gained loop anchors (*n*). **g**, Bar plots showing percent of gained loop anchors located in different genomic locations. *$P < 0.01$, two-sided Fischer's exact test. **h**, Heatmaps of nucleosome occupancy around gained CTCF anchors from **f**. **i**, Line plots of mean CTCF and SMC1A signal in the 6 kbp around gained loop anchors from **f**.

depletion from chromatin also allows for such fine-scale changes, presumably due to the absence of engaged polymerases that influence anchor selection.

## Modeling the interplay between loop extrusion and RNAPII

A key observation in our data was the overall reduced cohesin occupancy at RNAPII loop anchors (Fig. 3a,e,g), which nevertheless coincided with the emergence of new prominent CTCF-anchored loops (Fig. 2a–g). To test the interplay between RNAPII engaged to chromatin and cohesin loop extrusion, which would be challenging to do experimentally, we used computational modeling of 3D chromatin folding.

We first considered a synthetic 460 kbp-long polymer containing two genes transcribed in the sense direction, a cluster of eight enhancers between the genes, plus three CTCF binding sites (two encompassing the genes/enhancers and one located inside the downstream gene; Fig. 4a). Each bead in the polymer represented 2 kbp of chromatin, and we implemented two scenarios using established molecular dynamics approaches[40,41]. The first model approximated conditions in untreated DLD-1 cells (control), where RNAPs have specific affinity for promoters and enhancers, and can also transverse gene bodies at expected speeds to simulate transcription. In parallel, the model considers cohesin complexes able to bind the polymer and extrude loops at experimentally
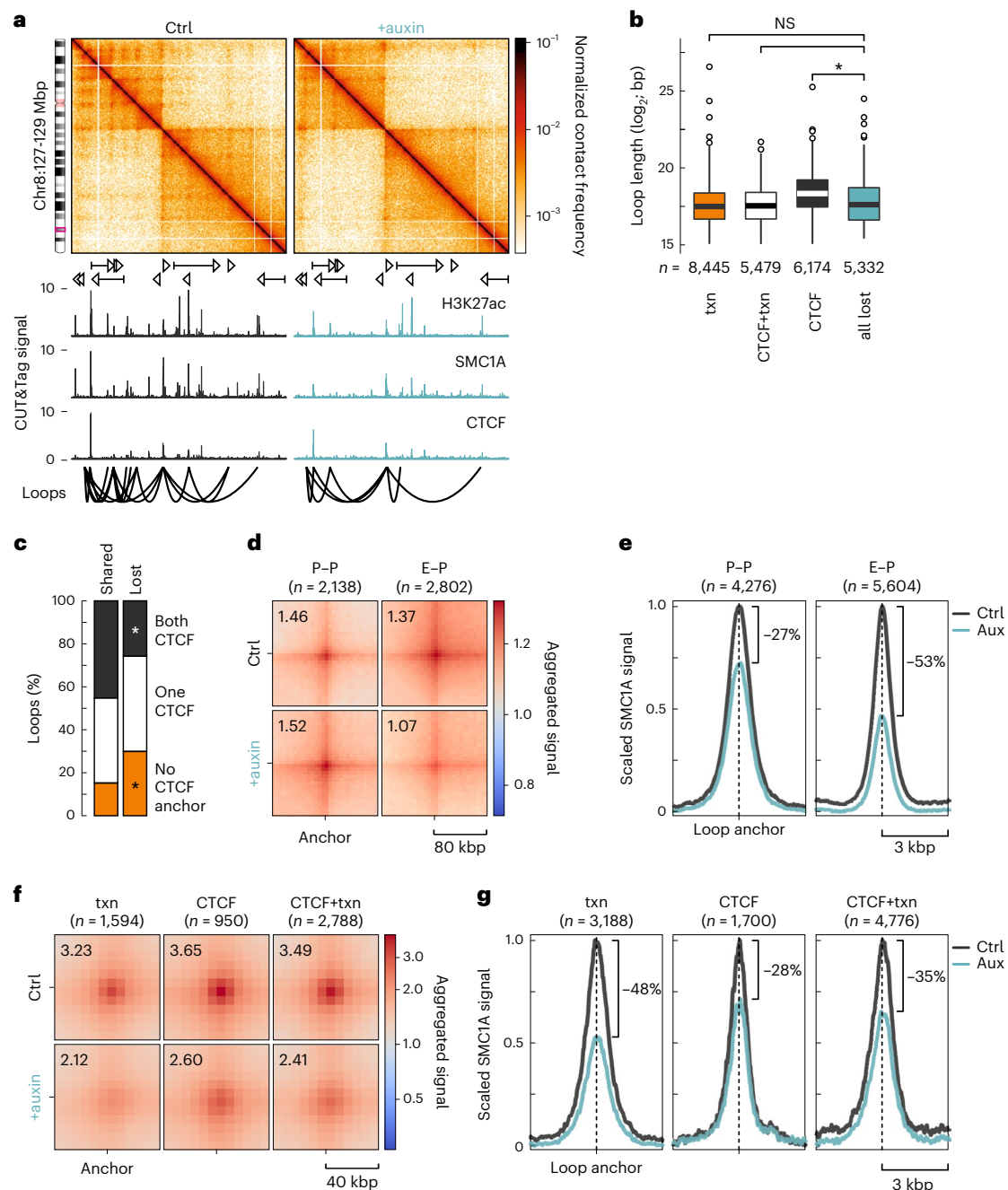
**Fig. 3 | RNAPII depletion selectively affects enhancer–promoter and enhancer–enhancer loops. a**, Micro-C contact maps from control (left) and auxin-treated cells (right) showing loop loss in an exemplary genomic region on chr8 at 4-kbp resolution aligned to H3K27ac, CTCF and SMC1A CUT&Tag tracks. Loops called for each condition are depicted by spider plots (bottom). **b**, Box plots of lengths of control loops with transcription-, CTCF- or transcription/CTCF anchors and of loops lost upon auxin treatment. In the plots, center lines represent the median value, box limits the 25th and 75th percentiles and whiskers extend 1.5× the box's interquartile range. *$P < 0.01$, two-sided Wilcoxon–Mann–Whitney test. **c**, Bar plot showing percent of loops with none, one or two CTCF anchors shared or lost by auxin-treated cells. *$P < 0.01$, two-sided Fischer's exact test. **d**, Aggregate plots for promoter–promoter (P–P) or enhancer–promoter loops (E–P) before (ctrl) and after RNAPII depletion (+auxin). **e**, Line plots of mean SMC1A signal around the loop anchors from **d**. **f**, As in **d**, but for loops with transcriptional (txn), CTCF or both anchors (CTCF + txn). **g**, As in **e**, but for loop anchors from **f**. NS, not significant.

deduced speeds (Methods; Supplementary Table 2). Based on the documented co-association of cohesin with RNAPII[39,42], we introduced a weak interaction potential between the two molecules. Thus, cohesin is allowed to bind any position in the polymer with a probability of 0.1 but binds promoters or enhancers with a 0.9 probability (as recently suggested experimentally[39,43–45]). In the model approximating RNAPII depletion (like in auxin-treated cells), all RNAPs are removed from the simulation and cohesin can now bind any position in the polymer with equal probability.

Control contact maps displayed a ~300 kbp-long CTCF loop that encompasses the two genes, and compartment-like interactions between the genes and enhancer cluster (Fig. 4a,b). Cohesin occupancy in this model showed the expected accumulation at the two distal, but not at the intragenic CTCF site (in line with transcribing RNAPs relocating cohesins[42,46,47]; Fig. 4a). Interestingly, despite loading being equally favored at promoters and enhancers, cohesin-loading rates are highest at the enhancer cluster (Fig. 4a, bottom). This agrees with what has been
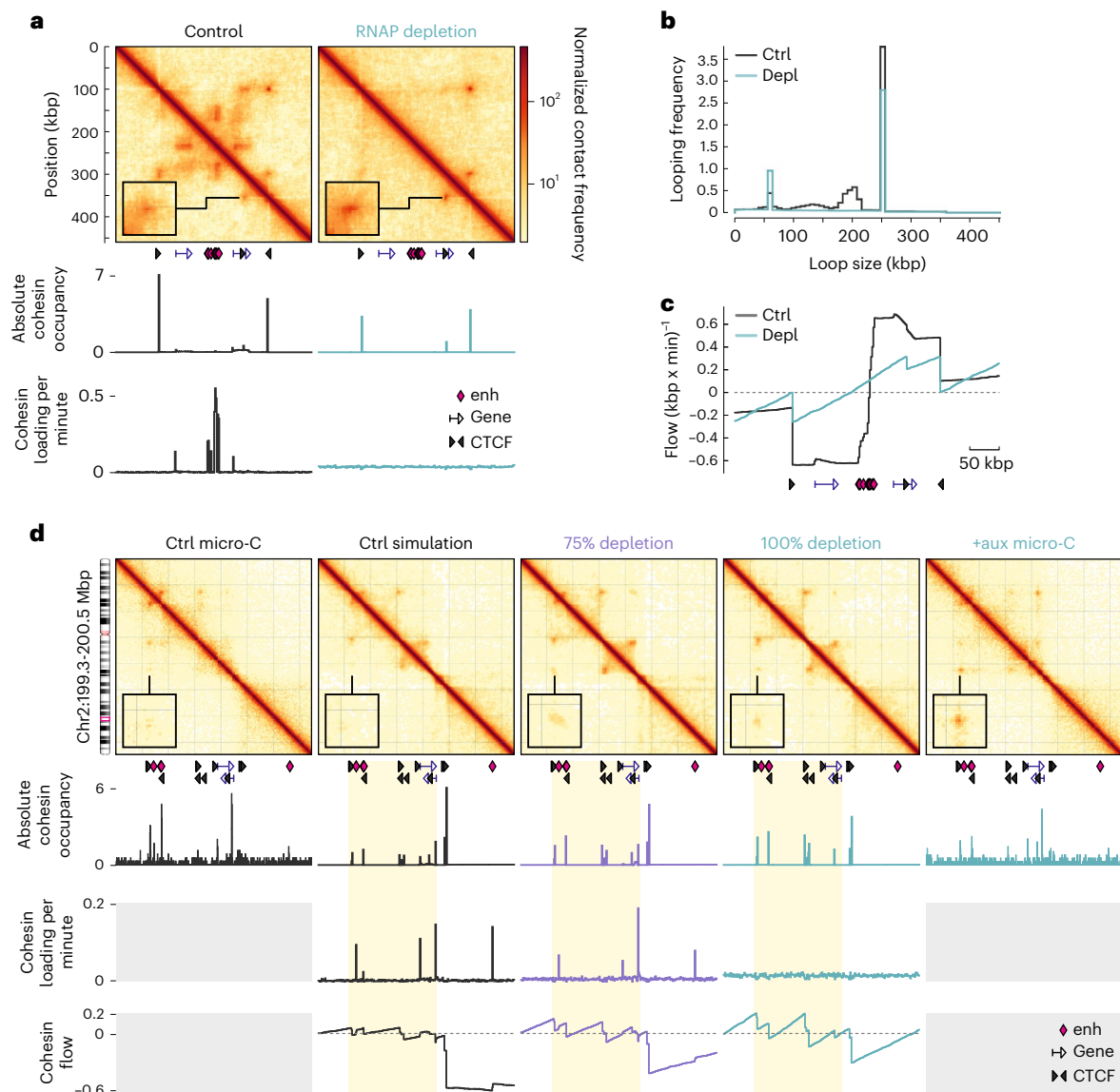
**Fig. 4 | Models of 3D chromatin folding in the presence or absence of RNAPII. a**, Average contact maps from 800 configurations of a toy model including (control) or not polymerases (RNAP depletion) are shown aligned to plots of absolute cohesin occupancy and loading. The positions of enhancers (diamonds), genes (arrows) and CTCF-bound site orientations are denoted (arrowheads). **b**, Plot of looping frequency per configuration as a function of distance in the models from **a. c**, Plot of the net flow of cohesin molecules along the polymer in each scenario. Positive and negative values represent extrusion in the sense and antisense direction, respectively. **d**, Average contact maps from 800 configurations of models simulating the presence (ctrl) or absence of RNAPs (75% or 100% depletion) in a 1.2-Mbp locus on chr2 compared to Micro-C data (far left and far right) aligned to plots of absolute cohesin occupancy, loading and flow.

recently experimentally deduced[43–45]. RNAP depletion in our model leads to the elimination of all contacts between genes and enhancers, and to the emergence of a new CTCF loop anchored at the intragenic CTCF site (Fig. 4a,b). This matches our Micro-C findings of intragenic CTCF sites being engaged in de novo looping following RNAPII depletion (Fig. 2e–g). These simulations also allow us to monitor the flow of cohesin along the polymer. We saw that RNAP presence on the polymer hindered cohesin-driven extrusion, especially around promoters. This was alleviated in the RNAP-depletion model (Fig. 4c; also suggested by recent preprints in refs. 48,49).

Finally, we used the same parameters to simulate the folding of a 1.2-Mbp locus on chr2 containing a number of enhancers and variously oriented CTCF sites, as well as two convergent active genes. In silico-generated contact maps showed good agreement with Micro-C data (Fig. 4d; SCC > 0.6) and allowed us to simulate different extents

of RNAP depletion. In the presence of RNAPs, the intragenic CTCF site overlapping the two gene bodies did not engage in loop formation in vivo or in silico. Upon depletion of 75% of RNAPs, diffuse interaction signal was observed, but full depletion was needed in order for a focal looping interaction to form (Fig. 4d). This was accompanied by increased cohesin occupancy at that CTCF anchor, and by reduced occupancy at active promoters. As before, the net flow of cohesin along the polymer was markedly less obstructed. Interestingly, under conditions of 75% RNAP depletion, the effects on cohesin loading and flow are not closer to the full depletion than to the control model (Fig. 4d, bottom). This suggests that the near-complete RNAP depletion from chromatin is required to obtain experimentally discernible changes. Taken together, the competition between active RNAPs and cohesin loaded preferentially, but not exclusively, at promoters and enhancers was what gave rise to most contact patterns seen by Micro-C.

## Discussion

In previous work, we established the necessity of RNAPII for reestablishing interphase chromatin folding after exiting from mitosis by a human cell line. We did not only show domain and compartment erosion in the absence of RNAPII, but also a dependency for cohesin loading onto chromatin[39]. However, in that same study, we could not identify 3D architecture changes of comparable magnitude by Hi-C applied to nonsynchronized cells. This remained perplexing until we obtained the Micro-C data analyzed here. Our high-resolution contact maps show that the transcription-based 3D architecture of interphase chromatin is markedly perturbed by RNAPII depletion. We could document both loss/weakening and gain/strengthening of specific loop-like interactions that beg the following questions.

First, how do longer and more pronounced CTCF loops arise in the absence of RNAPII? The data we present here, as well as recent preprinted work[49], see RNAPs as physical 'moving' barriers to loop extrusion. Along the same lines, recent studies found that transcribing RNAPII complexes can reposition cohesin complexes[42,46,47] or give rise to new spatial interactions[50]. Thus, depletion of the RNAPII barriers should allow for a more efficient extrusion of loops anchored at CTCF-bound sites, and facilitate the de novo engagement of CTCF anchors located inside of previously-active promoters and gene bodies (as also corroborated by our simulations). In parallel, a subset of loops that involved interactions between Polycomb-bound regions emerged in the absence of RNAPII. H3K27me3-marked regions are considered transcriptionally inert, but often bind 'poised' RNAPII[51], the removal of which might contribute directly or indirectly to the observed effects. Directly by a competitive interplay of RNAPII with Polycomb proteins and indirectly by RNAPs affecting cohesin loading (as cohesin depletion was shown to enhance contacts between Polycomb-bound regions[52]). A somewhat similar effect is now described in ref. 53 for promoter-proximal paused RNAPs maintaining local 3D chromatin architecture in erythrocytes.

Second, why did promoter- and enhancer-anchored interactions respond differently to RNAPII depletion? Here we observed two unforeseen results. H3K27ac levels dropped genome wide following RNAPII depletion but were substantially more reduced at enhancers compared to promoters. At the same time, enhancer-anchored loops were selectively weakened upon RNAPII depletion, but promoter–promoter ones remained essentially unaffected. This was independent of whether these interactions involved CTCF at their anchors and showed decreased cohesin occupancy in a pattern similar to H3K27ac. This was striking, given that promoters and enhancers are thought to be variants of a single class of cis-elements[54]. Nevertheless, our data suggest that spatial communication between promoters relies on a different set of factors (perhaps STAG1 versus −2 (ref. 55)) than that between enhancers and their target promoters. Identifying these factors and their attributes represents the next challenge in the field.

Third, how is cohesin chromatin occupancy affected by the absence of RNAPII? In the M-to-G1 transition, reduced cohesin loading correlated with the depletion of RNAPII from chromatin despite the fact that chromatin accessibility was not reduced (and, thus, could not be solely responsible for any reduction in cohesin loading[39]). Here we documented a similar reduction in chromatin-associated NIPBL and cohesin levels following RNAPII depletion from interphase cells. This was most apparent at enhancers–promoter loops and in line with (1) the binding of cohesin loaders at promoters[42] (although the specificity of some of this data is now debated[49]), (2) the fact that the loader NIPBL and unloader WAPL copurify with RNAPII complexes[39] and (3) recent work pointing to enhancers as cohesin-loading sites for the formation of 3D interactions[43–45]. Moreover, we should revisit studies where pioneer transcription factors (like OCT4 and SOX2 (ref. 56)) and chromatin remodelers (like SNF2h[57]) affected the loading/unloading cycles of cohesin onto chromatin. Still, distinguishing between cohesin loading or stalling at a given position remains challenging in vivo, due

to the processive nature of extruding complexes. However, our simulations show that, by introducing a weak interaction between RNAPII and cohesin, the latter is predominantly directed to active promoters and enhancers. Disfavoring cohesin loading at promoters (by competition with RNAPII) generates contacts rarely seen by Micro-C.

Cohesin, its loader NIPBL and the Mediator complex were proposed to co-associate to physically and functionally connect active enhancers and promoters[58]. Hi-C studies that followed this work found Mediator and RNAPII to be dispensable for 3D chromatin folding[29,30]. This is now challenged by our data, MCC data showing loss of enhancer–promoter contacts following acute depletion of Mediator[37] and decreased cohesin binding to RNAPII-transcribed genes following depletion of the yeast Med14 subunit[59]. For RNAPII specifically, a previous attempt to deplete it from chromatin via triptolide inhibition (for 45 min) only resulted in the reduction of 'stripes' genome wide and not in any loss of E–P loops[33]. This difference in our data can be attributed to the inefficiency of short-term triptolide treatments in removing most RNAPII from chromatin; in contrast, our 'degron' approach can achieve strong (>90%) depletion[39]. In summary, our observations put the still-debated role of RNAPII in 3D chromatin folding under a different light: they provide definitive evidence for the necessity of RNAPII in sustaining enhancer–promoter interactions, as well as for direct antagonism during the formation of CTCF loops. The former appears to require the presence of RNAPII on chromatin, whereas the latter likely also implicate ongoing transcription. Nonetheless, there remain aspects of polymerase-based 3D chromatin architecture to be elucidated, like the differential dependency of promoter- compared to enhancer-anchored interactions or the mechanistic details of RNAPII influence on cohesin loading.

## Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41588-023-01364-4.

## References

1.  Aboelnour, E. & Bonev, B. Decoding the organization, dynamics, and function of the 4D genome. *Dev. Cell* **56**, 1562–1573 (2021).
2.  Razin, S. V. & Kantidze, O. L. The twisted path of the 3D genome: where does it lead? *Trends Biochem. Sci.* **47**, 736–744 (2022).
3.  Xiang, J. F. & Corces, V. G. Regulation of 3D chromatin organization by CTCF. *Curr. Opin. Genet. Dev.* **67**, 33–40 (2021).
4.  van Ruiten, M. S. & Rowland, B. D. On the choreography of genome folding: a grand pas de deux of cohesin and CTCF. *Curr. Opin. Cell Biol.* **70**, 84–90 (2021).
5.  Beagan, J. A. & Phillips-Cremins, J. E. On the existence and functionality of topologically associating domains. *Nat. Genet.* **52**, 8–16 (2020).
6.  Rao, S. S. P. et al. Cohesin loss eliminates all loop domains. *Cell* **17**, 305–320 (2017).
7.  Fudenberg, G. et al. Formation of chromosomal domains by loop extrusion. *Cell Rep.* **15**, 2038–2049 (2016).
8.  Davidson, I. F. et al. DNA loop extrusion by human cohesin. *Science* **366**, 1338–1345 (2019).
9.  Kim, Y., Shi, Z., Zhang, H., Finkelstein, I. J. & Yu, H. Human cohesin compacts DNA by loop extrusion. *Science* **366**, 1345–1349 (2019).
10. Nora, E. P. et al. Targeted degradation of CTCF decouples local insulation of chromosome domains from genomic compartmentalization. *Cell* **169**, 930–944 (2017).
11. Wutz, G. et al. Topologically associating domains and chromatin loops depend on cohesin and are regulated by CTCF, WAPL, and PDS5 proteins. *EMBO J.* **36**, 3573–3599 (2017).

12. Wutz, G. et al. ESCO1 and CTCF enable formation of long chromatin loops by protecting cohesin$^{STAG1}$ from WAPL. *eLife* **9**, e52091 (2020).

13. Li, Y. et al. The structural basis for cohesin-CTCF-anchored loops. *Nature* **578**, 472–476 (2020).

14. Rao, S. S. et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **159**, 1665–1680 (2014).

15. Schwarzer, W. et al. Two independent modes of chromatin organization revealed by cohesin removal. *Nature* **551**, 51–56 (2017).

16. Haarhuis, J. H. I. et al. The cohesin release factor WAPL restricts chromatin loop extension. *Cell* **169**, 693–707 (2017).

17. Gabriele, M. et al. Dynamics of CTCF- and cohesin-mediated chromatin looping revealed by live-cell imaging. *Science* **376**, 496–501 (2022).

18. Nuebler, J., Fudenberg, G., Imakaev, M., Abdennur, N. & Mirny, L. A. Chromatin organization by an interplay of loop extrusion and compartmental segregation. *Proc. Natl Acad. Sci. USA* **115**, E6697–E6706 (2018).

19. Rada-Iglesias, A., Grosveld, F. G. & Papantonis, A. Forces driving the three-dimensional folding of eukaryotic genomes. *Mol. Syst. Biol.* **14**, e8214 (2018).

20. Papantonis, A. & Cook, P. R. Fixing the model for transcription: the DNA moves, not the polymerase. *Transcription* **2**, 41–44 (2011).

21. Lee, K. & Blobel, G. A. Chromatin architecture underpinning transcription elongation. *Nucleus* **7**, 1–8 (2016).

22. Cook, P. R. & Marenduzzo, D. Transcription-driven genome organization: a model for chromosome structure and the regulation of gene expression tested through simulations. *Nucleic Acids Res.* **46**, 9895–9906 (2018).

23. Racko, D., Benedetti, F., Dorier, J. & Stasiak, A. Transcription-induced supercoiling as the driving force of chromatin loop extrusion during formation of TADs in interphase chromosomes. *Nucleic Acids Res.* **46**, 1648–1660 (2018).

24. Mitchell, J. A. & Fraser, P. Transcription factories are nuclear subcompartments that remain in the absence of transcription. *Genes Dev.* **22**, 20–25 (2008).

25. Palstra, R. J. et al. Maintenance of long-range DNA interactions after inhibition of ongoing RNA polymerase II transcription. *PLoS ONE* **3**, e1661 (2008).

26. Ke, Y. et al. 3D chromatin structures of mature gametes and structural reprogramming during mammalian embryogenesis. *Cell* **170**, 367–381 (2017).

27. Barutcu, A. R., Blencowe, B. J. & Rinn, J. L. Differential contribution of steady-state RNA and active transcription in chromatin organization. *EMBO Rep.* **20**, e48068 (2019).

28. Brant, L. et al. Exploiting native forces to capture chromosome conformation in mammalian cell nuclei. *Mol. Syst. Biol.* **12**, 891 (2016).

29. El Khattabi, L. et al. A pliable Mediator acts as a functional rather than an architectural bridge between promoters and enhancers. *Cell* **178**, 1145–1158 (2019).

30. Jiang, Y. et al. Genome-wide analyses of chromatin interactions after the loss of Pol I, Pol II, and Pol III. *Genome Biol.* **21**, 158 (2020).

31. Sun, F. et al. The Pol II preinitiation complex (PIC) influences Mediator binding but not promoter-enhancer looping. *Genes Dev.* **35**, 1175–1189 (2021).

32. Haarhuis, J. H. I. et al. A Mediator-cohesin axis controls heterochromatin domain formation. *Nat. Commun.* **13**, 754 (2022).

33. Hsieh, T. S. et al. Resolving the 3D landscape of transcription-linked mammalian chromatin folding. *Mol. Cell* **78**, 539–553 (2020).

34. Krietenstein, N. et al. Ultrastructural details of mammalian chromosome architecture. *Mol. Cell* **78**, 554–565 (2020).

35. Hua, P. et al. Defining genome architecture at base-pair resolution. *Nature* **595**, 125–129 (2021).

36. Cramer, P. Organization and regulation of gene transcription. *Nature* **573**, 45–54 (2019).

37. Ramasamy, S. et al. The mediator complex regulates enhancer–promoter interactions. Preprint at *bioRxiv*, https://doi.org/10.1101/2022.06.15.496245 (2022).

38. Nagashima, R. et al. Single nucleosome imaging reveals loose genome chromatin networks via active RNA polymerase II. *J. Cell Biol.* **218**, 1511–1530 (2019).

39. Zhang, S. et al. RNA polymerase II is required for spatial chromatin reorganization following exit from mitosis. *Sci. Adv.* **7**, eabg8205 (2021).

40. Buckle, A., Brackley, C. A., Boyle, S., Marenduzzo, D. & Gilbert, N. Polymer simulations of heteromorphic chromatin predict the 3D folding of complex genomic loci. *Mol. Cell* **72**, 786–797 (2018).

41. Fiorillo, L. et al. Inference of chromosome 3D structures from GAM data by a physics computational approach. *Methods* **181–182**, 70–79 (2020).

42. Busslinger, G. A. et al. Cohesin is positioned in mammalian genomes by transcription, CTCF and Wapl. *Nature* **544**, 503–507 (2017).

43. Zhu, Y., Denholtz, M., Lu, H. & Murre, C. Calcium signaling instructs NIPBL recruitment at active enhancers and promoters via distinct mechanisms to reconstruct genome compartmentalization. *Genes Dev.* **35**, 65–81 (2021).

44. Rinaldi, L. et al. The glucocorticoid receptor associates with the cohesin loader NIPBL to promote long-range gene regulation. *Sci. Adv.* **8**, eabj8360 (2022).

45. Rinzema, N. J. et al. Building regulatory landscapes reveals that an enhancer can recruit cohesin to create contact domains, engage CTCF sites and activate distant genes. *Nat. Struct. Mol. Biol.* **29**, 563–574 (2022).

46. Heinz, S. et al. Transcription elongation can affect genome 3D structure. *Cell* **174**, 1522–1536 (2018).

47. Olan, I. et al. Transcription-dependent cohesin repositioning rewires chromatin loops in cellular senescence. *Nat. Commun.* **11**, 6049 (2020).

48. Valton, A. L. et al. A cohesin traffic pattern genetically linked to gene regulation. *Nat. Struct. Mol. Biol.* **29**, 1239–1251 (2022).

49. Banigan, E. J. et al. Transcription shapes 3D chromatin organization by interacting with loop extrusion. *Proc. Natl Acad. Sci. USA* **120**, e2210480120 (2023).

50. Rosencrance, C. D. et al. Chromatin hyperacetylation impacts chromosome folding by forming a nuclear subcompartment. *Mol. Cell* **78**, 112–126 (2020).

51. Ferrai, C. et al. RNA polymerase II primes Polycomb-repressed developmental genes throughout terminal neuronal differentiation. *Mol. Syst. Biol.* **13**, 946 (2017).

52. Rhodes, J. D. P. et al. Cohesin disrupts polycomb-dependent chromosome interactions in embryonic stem cells. *Cell Rep.* **30**, 820–835 (2020).

53. Penagos-Puig, A. et al. RNA polymerase II pausing contributes to maintain chromatin organization in erythrocytes. Preprint at *bioRxiv* https://doi.org/10.1101/2022.06.16.496295 (2022).

54. Andersson, R., Sandelin, A. & Danko, C. G. A unified architecture of transcriptional regulatory elements. *Trends Genet.* **31**, 426–433 (2015).

55. Casa, V. et al. Redundant and specific roles of cohesin STAG subunits in chromatin looping and transcriptional control. *Genome Res.* **30**, 515–527 (2020).

56. Liu, N. Q. et al. WAPL maintains a cohesin loading cycle to preserve cell-type-specific distal gene regulation. *Nat. Genet.* **53**, 100–109 (2021).
57. Hakimi, M. A. et al. A chromatin remodelling complex that loads cohesin onto human chromosomes. *Nature* **418**, 994–998 (2002).
58. Kagey, M. H. et al. Mediator and cohesin connect gene expression and chromatin architecture. *Nature* **467**, 430–435 (2010).
59. Mattingly, M. et al. Mediator recruits the cohesin loader Scc2 to RNA Pol II-transcribed genes and promotes sister chromatid cohesion. *Curr. Biol.* **32**, 2884–2896 (2022).

## Methods

### Cell synchronization and sorting

mAID-POLR2A(RPB1)-mClover DLD-1 cells[38] were grown in RPMI-1640 medium supplemented with 10% FBS under 5% $CO_2$. Inducible depletion of RPB1 initiated via treatment with doxycycline for 24 h to induce *TIR1* expression, before addition of 500 µM indole-3-acetic acid solution ('auxin'; Sigma-Aldrich) for 14 h to induce RPB1 degradation. Cells treated with auxin were collected, resuspended in 1 µg ml$^{-1}$ propidium iodide and sorted to isolate G1 cells on a FACS Canto II flow cytometer (Becton Dickinson; gating strategy in Supplementary Fig. 1).

### Micro-C and data analysis

Micro-C was performed using the Micro-C v1.0 kit of Cantata Bio (formerly Dovetail Genomics) as per the manufacturer's instructions. Micro-C libraries (at least three per each biological replicate) that passed QC criteria were pooled and paired-end sequenced on a NovaSeq6000 platform (Illumina) to >600 million read pairs per replicate (Supplementary Table 1). The Dovetail Genomics pipeline (https://micro-c.readthedocs.io/en/latest/fastq_to_bam.html) was used to produce Micro-C contact matrices. In brief, read pairs were mapped to human reference genome hg38 using BWA ver. 0.7.17, after which low mapping quality (<40) reads and PCR duplicates were filtered out. Next, ICE-balanced *.cool* files and KR-balanced *.hic* files were generated and visualized via HiGlass ver. 1.11.7, cooltools (https://cooltools.readthedocs.io/en/latest/notebooks/contacts_vs_distance.html) used to generate decay plots via, and sub-compartment analysis was performed using CALDER[60] considered 50 kbp-resolution Micro-C data. For loop calling, we used a multi-tool (HiCCUPS ver. 1.19.01, cooltools ver. 0.5.4, and mustache ver. 1.0.1) and a multi-resolution (5- and 10-kbp) approach[33,34]. Loop lists derived from each tool were merged using *pgltools* ver. 1.2.1 (ref. 61) as follows: dots from both 10- and 5-kbp resolution are retained if they are supported by >10 read counts, and kept at native resolution. To further annotate loops as CTCF- or transcription-anchored, using CTCF, H3K27ac CUT&Tag peaks (from this work), as well as RNAPII peaks and nascent RNA-seq signal (RPKM > 10; from ref. 39). All intersections were performed using *pgltools intersect1D* without any distance tolerance for CTCF anchors, and with a 10-kbp tolerance for enhancers and promoter anchors (annotated TSS ± 2 kbp) identified using *chipseeker* ver. 3.16 (ref. 62). Note that promoters of all gene isoforms were considered, and 'superenhancers' were called using our H3K27ac CUT&Tag data and the ROSE algorithm[63]. Finally, aggregate plots for loops and boundaries were generated using the *coolpup.py* tool[64]. All custom code used is available at: https://github.com/shuzhangcourage/Micro-C-CUT-tag/tree/v1.0.0.

### Cleavage under targets and tagmentation (CUT&Tag)

Following lifting from plates using Accutase and FACS sorting, 0.5 million G1-phase DLD-1 cells were processed using the CUT&Tag-IT Kit according to the manufacturer's instructions (Active Motif). Samples were paired-end sequenced to obtain at least $10^7$ reads, and processed according to a standard pipeline (https://yezhengstat.github.io/CUT-Tag_tutorial/). Briefly, paired-end reads were trimmed for adapter removal and mapped to human (hg38) and *Escherichia coli* reference genomes (ASM584v2) using Bowtie2 ver. 2.4.4 (ref. 65). *E. coli*-mapped reads were quantified and used for calibrating human-mapped reads. Peak calling was performed using a multi-FDR-tryout method (FDR < 0.01 to <0.1) and IgG controls for thresholding. Acceptable FDRs could vary between different datasets but were always kept the same for control and auxin-treated samples. Thus, for CTCF, an FDR < 0.1 was selected and, for additional stringency, we only considered a CUT&Tag peak as CTCF-bound if it encompassed a canonical CTCF motif (assessed using FIMO[66]). For H3K27ac and H3K27me3, peaks were selected on the basis of FDR < 0.025 and <0.01, respectively, while for SMC1A, an FDR < 0.1 was used. Heatmaps were generated using Deeptools ver. 2.0 (ref. 67).

### Chromatin fractionation and western blotting

For assessing protein abundance in different subcellular fractions, chromatin fractionation was applied[68]. Typically, 2 million cells are washed in ice-cold PBS, scraped off and spun for 5 min at 1,000*g*. Subsequently, the pellet is lysed in a lysis buffer (20 mM Tris–HCl pH 8, 1% NP-40, 150 mM NaCl, 1× protease inhibitor cocktail) for 20 min on ice. Approximately 1/3 of this lysate is kept as the 'total cell fraction'. The remaining lysate is centrifuged for 5 min at 1,000*g* at 4 °C, and the supernatant is transferred to a new tube, while the pellet is kept as the 'chromatin fraction'. The supernatant is centrifuged for 15 min >12,000*g* at 4 °C and what does not precipitate is kept as the 'soluble fraction'. The 'chromatin fraction' pellet is gently resuspended and washed 3× in lysis buffer, before last centrifugation for 5 min at 1,000*g* at 4 °C. The final pellet is resuspended in TBS/T, incubated with benzonase for 30 min at 37 °C, and sonicated (five cycles of 30 s 'on/off') on a Bioruptor (Diagenode). In preparation for western blotting, all lysates are boiled at 95 °C for 5 min in 1× Laemmli buffer, and protein concentration in each fraction is determined using the Pierce BCA Protein Assay Kit (Thermo Fisher Scientific). Following separation on precast SDS-PAGE gels (Bio-Rad), proteins were detected using antibodies against p-Ser5 RPB1 (Active Motif, 61085; 1:2,000), RPB1 (Abcam, ab817; 1:500), NIPBL (Bethyl, A301-779A; 1:10,000), MED24 (Affinity Biosciences, AF0346; 1:1,000), Lamin B1 (Abcam ab16048; 1:10,000), SMC1A (Abcam, ab9262; 1:4,000), CTCF (Active Motif, 61311; 1:2,000), H3K27me3 (Active Motif, 39155; 1:1,000) and HSC70 (Santa Cruz, sc-7298; 1:2,000), and visualized using the Pierce SuperSignal WestPico ECL kit (Thermo Fisher Scientific). Source data for our western blots are shown in Supplementary Fig. 2.

### Simulations of chromatin folding

We performed Molecular Dynamics simulations via the multi-purpose EspressoMD package ver. 4.3 (ref. 69). In our simulations, individual proteins are represented by 'beads' interacting via phenomenological force fields and moved according to the Langevin equation, and the chromatin fiber is represented as a chain of beads connected by bonds. The position of every bead in the system, either a protein or chromatin bead, evolves according to the Langevin differential equation that encodes Newton's laws in the case of thermal bath with the friction γ due to an implied solvent in presence of forces between beads encoded by energy potential functions $U$[40,70]. Langevin equations for all beads are simultaneously solved in EspressoMD using a standard velocity-Verlet numerical algorithm. The potential connecting $i$ and $i + 1$ beads of the fiber is a finitely extensible nonlinear elastic (FENE) spring that adds up to a steric repulsion potential between nonadjacent sites of the polymer, the Weeks-Chandler-Andersen potential:

$$U_{\text{FENE}} = -\frac{K_{\text{FENE}}R_0^2}{2} \ln \left[ 1 - \left( \frac{r_{i,i+1} - r\_0}{R_0} \right) \right] \quad (1)$$

$$U_{\text{WCA}} = \begin{cases} 4\epsilon \left[ \left( \frac{\sigma}{r} \right)^{12} - \left( \frac{\sigma}{r} \right)^{6} + \frac{1}{4} \right] & \text{if } r < \sigma 2^{1/6} \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

where $r_{i,i+1}$ is the distance between consecutive beads, and σ is where the interaction from repulsive becomes attractive and can be interpreted as the diameter of the particles. This value is a natural length scale of the system. In FENE we fix parameters to have an equilibrium distance of 1.6 σ with a maximum extension of 0.8 σ, and bond energy of $K_{\text{FENE}} = 30$ $k_B T$. Because our fiber is resolved at 2 kbp, chromatin rigidity cannot be neglected (that is, we are below the estimated persistence length). Bending rigidity of the polymer is introduced via the Kratky-Porod potential for every three adjacent chromatin beads where θ is the angle between three consecutive beads as given by:

$$U_{KP}(\theta) = K_{BEND}(1 - \cos(\theta)) \tag{3}$$

and $K_{BEND}$ is the bending energy. The persistence length in units of $\sigma$ is given by $L_p = K_{BEND}/k_B T$.

To model more complex aspects of transcription and loop extrusion, and the impact their interplay has on 3D chromatin organization, we encoded in the model as follows: (1) full 3D loop extrusion by the interplay of cohesin dimers and CTCF; (2) transcription by RNAPII particles. To simulate association between cohesin and RNAPII with the chromatin fiber, we employed a harmonic potential mimicking formation of a stable bond between two particles that fluctuate around an equilibrium distance $d_0$:

$$U_H = \frac{1}{2}K_H(r - d_0)^2 \tag{4}$$

To regulate the lifetime of the above interaction, we introduced mechanisms of bond formation and removal according to cutoff distance $cd$ below which a bond is formed with a certain probability rate of detachment in units of time $\tau_b = 2\tau$, $\tau$ the fundamental MD unit of time (see below). These are then set to approximate the experimentally observed range of RNAPII transcription and cohesin loop extrusion speeds and chromatin residence time. The above mechanics is added on top of the SBS model we previously employed in ref. 71. The model encodes the association tendency of RNAPII with promoters by means of the shifted, truncated Lennard–Jones (LJ) potential that allows spontaneous colocalization of beads with lifetime and stability properties depending on the depth of the energy well $\epsilon$:

$$U_{LJ} = \begin{cases} 4\epsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^6 + c_{shift}\right] & \text{if } r < r_{cut} \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

where $r_{cut} = 2.5\sigma$ for all LJ potentials in the simulations, $r$ is the separation of any two beads. This is a standard widely used in the field to simulate phenomenological course-grained affinities[40,70].

For RNAPII interactions and transcription, the polymerase is represented as a bead with LJ interaction with specific beads of the chromatin fiber representing promoters and enhancers with energy $\epsilon = 1.5$ and $2\,k_B T$, respectively. Such mild affinity helps to identify promoters as the correct sites where transcription initiation will take place (that is, RNAPII forming stable bonds with promoter beads) before the elongation process on the gene body starts. LJ interactions were also introduced among RNAPII beads ($\epsilon = 2.5\,k_B T$) to simulate their tendency to form condensates acting as transcription hubs, as well as between RNAPII and cohesin ($\epsilon = 3\,k_B T$) to simulate the latter preferential loading at promoter/enhancer beads. RNAPII transcription dynamics are simulated as a four-step process: attachment to a promoter in an exclusive manner, elongation starts, elongation proceeds through the gene body, detachment at TES. A bond is formed if the beads are less distant in space than the cutoff $2.7\,\sigma$. To simulate the tendency of RNAPII to reel in gene body beads, a secondary bond is formed with the next bead on the chromatin fiber in the direction of transcription ($i + 1$ bead, where $i$ is the promoter coordinate on the fiber, and if transcription occurs in the sense direction; $i - 1$ in the antisense direction). In the next step, RNAPII moves on the next site by forming new bonds with $i + 1$ site and dissolving the old ones with $i$. This happens at a given rate ($0.4\,\tau_b^{-1}$) and only if the beads are found within the cutoff distance $1.05\,\sigma$. These values are selected to obtain RNAPII transcription speed approximately in the range of 1–10 kbp min$^{-1}$ observed experimentally. Upon reaching the TES, RNAPII stops and becomes unbound with rate $0.2\,\tau_b^{-1}$. Upon binding with promoters RNAPII loses its LJ interaction with promoters, because this is substituted by the bond itself. On the contrary, RNAPII increases the LJ affinity with enhancers ($3\,k_B T$) to favor associations

between actively transcribed segments. This preserves the stability of condensates/hubs during the process of transcription.

CTCF interacts via LJ interactions ($\epsilon = 1.5\,k_B T$) with specific sites of the chromatin fiber representing the oriented cognate binding motifs. Once a bond is formed (with rate $0.8\,\tau_b^{-1}$) it is pair-exclusive (that is, other CTCF cannot bind that same site). The bond dissolves at the rate $2 \times 10^{-15}\,\tau_b^{-1}$ and CTCF is again free to diffuse and search for other binding sites. Cohesin dimers are represented as bead pairs connected by one bond ($r_0 = 1.6\sigma$ and $K = 8\,k_B T$). Extrusion has the following three steps: attachment, active extrusion and detachment. For attachment, each cohesin monomer forms a bond with the chromatin fiber. Bonds form when a cohesin monomer and a chromatin site come within the cutoff distance of $1.6\,\sigma$ with attachment rate of $0.1\,\tau_b^{-1}$. Only the case where both monomers simultaneously form bonds on adjacent chromatin beads is considered a successful attachment and the dimer is retained for the next step, otherwise bonds dissolve. If a promoter is already engaged in a bond with RNAPII, cohesin is forbidden to bind that promoter. Also, to favor cohesin loading in correspondence of active transcribing promoters, a 90% chance of binding has been introduced when a cohesin molecule is close to a promoter/enhancer and at the same time one RNAPII is close by as well (cutoff distance $1.5\,\sigma$), otherwise cohesin binding chance drops to 10%. The active extrusion and detachment steps follow the same mechanics as for RNAPII, with the difference that RNAPII can reel through cohesin-bound sites while the converse is not allowed. New bonds are formed if the distance is below $1.1\sigma$. Such value produces ranges of cohesin extruding speed of 15–30 kbp min$^{-1}$, which is within the range of experimentally observed values[6], as cohesin detachment occurs at the rate $10^{-4}\,\tau_b^{-1}$ to fit its known chromatin residence time of 20 min. Finally, CTCF 'loop anchors' are modeled so that cohesin cannot form new bonds with the next $i \pm 1$ site if the latter is already bound by CTCF, provided it has the binding motif in convergent orientation. This renders extrusion dependent on CTCF dynamics. Last, cohesin has LJ affinity with RNAPII both in the bound ($\epsilon = 3\,k_B T$) and unbound state ($\epsilon = 2\,k_B T$), higher affinity for bound RNAPII mimicking RNAPII suggested role in cohesin loading on chromatin[39].

RNAPII and LE dynamics are performed using a python script that drives the EspressoMD library. The polymer initializes as a random walk and its dynamics first evolves in the absence of extrusion and transcription to generate an equilibrium coil conformation. In the following step, both extrusion and transcription are switched on, and their dynamics evolve until a new steady-state conformation is obtained. Across all simulations, we used standard values for the friction coefficient ($\gamma = 0.5$) and the time step ($t = 0.01$), and we let the system evolve for up to $10^8$ steps. As in previous studies, to connect our in silico space–time units with real distances and times of the biological process, we assumed that the concentration of DNA in the 3D simulation space is the same as that in a human nucleus. If we use a total DNA amount of 6 Gbp and a nucleus radius of 5 µm, we obtain the rough estimation of $\sigma = 65$ nm. For time units, we consider the standard MD relation $\tau = \eta(6\,\pi\,\sigma^3/\epsilon)$. Assuming a viscosity ~0.25 P, the fundamental time unit is $\tau = 0.03$ s. By running simulations starting from independent configurations and by sampling periodically the system we obtain an ensemble of configurations up to $10^3$ for the measurement of the quantities shown. Concentrations of CTCF, cohesin and RNAPII are taken from physiological values and range from 10 to 50 nmol l$^{-1}$. The energy scale of the system is given by the Boltzmann factor $k_B$ multiplied by the temperature of the system $T = 310$ K.

### Statistical analyses
$P$ values derived from Fisher's exact test were calculated using the online GraphPad interface (https://www.graphpad.com/quickcalcs/), and those derived from the Wilcoxon–Mann–Whitney test using R. Unless otherwise stated, $P < 0.01$ were deemed significant. For comparing Micro-C and simulated contact matrices distance-corrected

Pearson's and stratum-adjusted correlation coefficients from HiCRep ver. 1.5.0 (ref. 72) with $h = 5$ were used as metrics.

## Reporting summary

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

## Data availability

NGS data generated in this study are available via the NCBI Gene Expression Omnibus repository under accession number GSE178593 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE178593). All other data used for analyses come from our previous study[39], and are available under accession number GSE160321 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE160321). Source data are provided with this paper.

## Code availability

All custom code used for Micro-C analysis is available at https://github.com/shuzhangcourage/Micro-C-CUT-tag/tree/v1.0.0 (ref. 73), and all the code used for our Molecular Dynamics simulations is available at https://zenodo.org/record/7674875#.Y_n5zXbMLBR (ref. 74).

## References

60. Liu, Y. et al. Systematic inference and comparison of multi-scale chromatin sub-compartments connects spatial organization to cell phenotypes. *Nat. Commun.* **12**, 2439 (2021).
61. Greenwald, W. W. et al. Pgltools: a genomic arithmetic tool suite for manipulation of Hi-C peak and other chromatin interaction data. *BMC Bioinformatics* **18**, 207 (2017).
62. Yu, G., Wang, L. G. & He, Q. Y. ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics* **31**, 2382–2383 (2015).
63. Whyte, W. A. et al. Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* **153**, 307–319 (2013).
64. Flyamer, I. M., Illingworth, R. S. & Bickmore, W. A. Coolpup. py: versatile pile-up analysis of Hi-C data. *Bioinformatics* **36**, 2980–2985 (2020).
65. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
66. Grant, C. E., Bailey, T. L. & Noble, W. S. FIMO: scanning for occurrences of a given motif. *Bioinformatics* **27**, 1017–1018 (2011).
67. Ramírez, F., Dündar, F., Diehl, S., Grüning, B. A. & Manke, T. DeepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res.* **42**, W187–W191 (2014).
68. Watrin, E. et al. Human Scc4 is required for cohesin binding to chromatin, sister-chromatid cohesion, and mitotic progression. *Curr. Biol.* **16**, 863–874 (2006).
69. Reynwar, B. J. et al. Aggregation and vesiculation of membrane proteins by curvature-mediated interactions. *Nature* **447**, 461–464 (2007).
70. Chiariello, A. M., Annunziatella, C., Bianco, S., Esposito, A. & Nicodemi, M. Polymer physics of chromosome large-scale 3D organisation. *Sci. Rep.* **6**, 29775 (2016).
71. Barbieri, M. et al. Active and poised promoter states drive folding of the extended HoxB locus in mouse embryonic stem cells. *Nat. Struct. Mol. Biol.* **24**, 515–524 (2017).
72. Yang, T. et al. HiCRep: assessing the reproducibility of Hi-C data using a stratum-adjusted correlation coefficient. *Genome Res.* **27**, 1939–1949 (2017).
73. Zhang, S. Code used for Micro-C and CUT&Tag data analysis. *Zenodo* https://doi.org/10.5281/zenodo.7656606 (2023).
74. Barbieri, M. Code used for MD simulations. *Zenodo* https://doi.org/10.5281/zenodo.7674875 (2023).

## Author contributions

S.Z. performed all bioinformatics analyses; N.Ü. performed all experiments; M.B. performed the computational modeling; and A.P. conceived and supervised the study and compiled the manuscript with input from all co-authors.

## Competing interests

All authors declare no competing interest.

## Additional information

**Extended data** is available for this paper at https://doi.org/10.1038/s41588-023-01364-4.

**Supplementary information** The online version contains supplementary material available at https://doi.org/10.1038/s41588-023-01364-4.

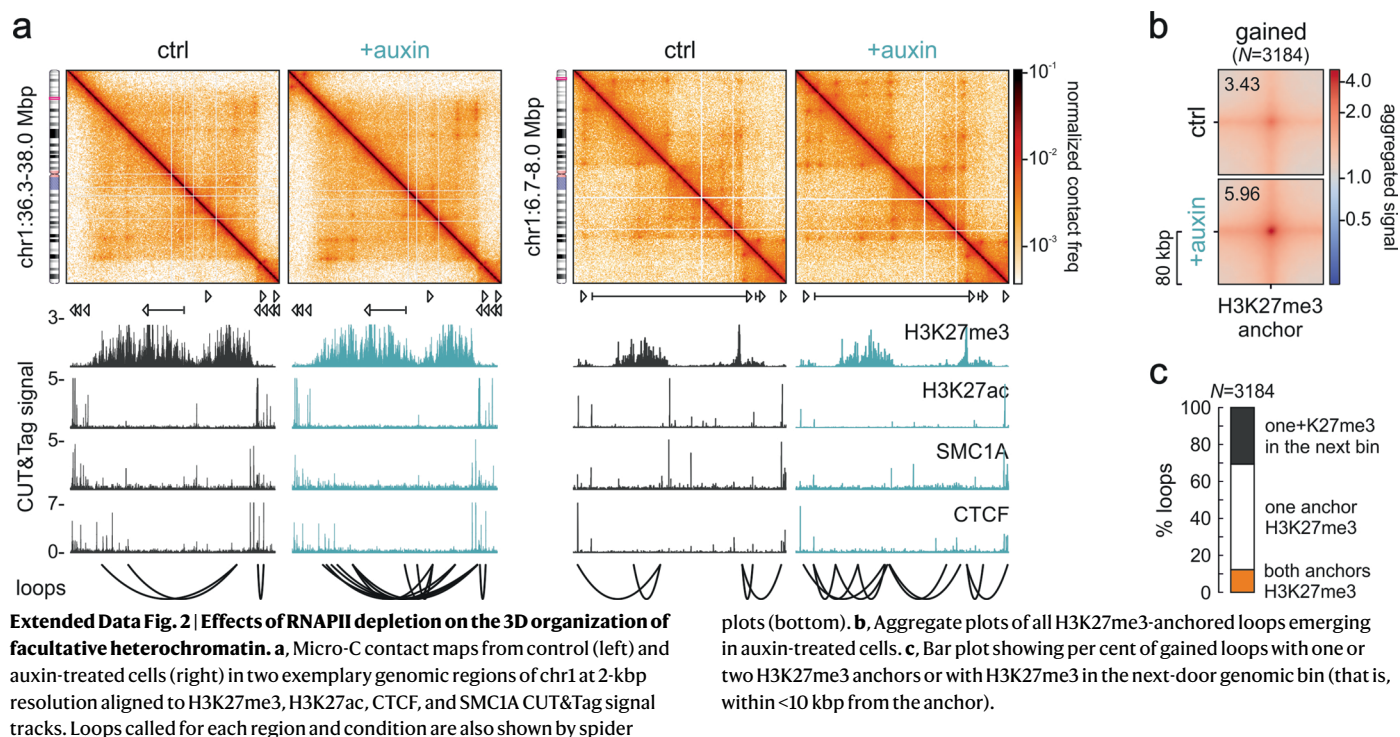**Correspondence and requests for materials** should be addressed to Argyris Papantonis.

**Peer review information** *Nature Genetics* thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.
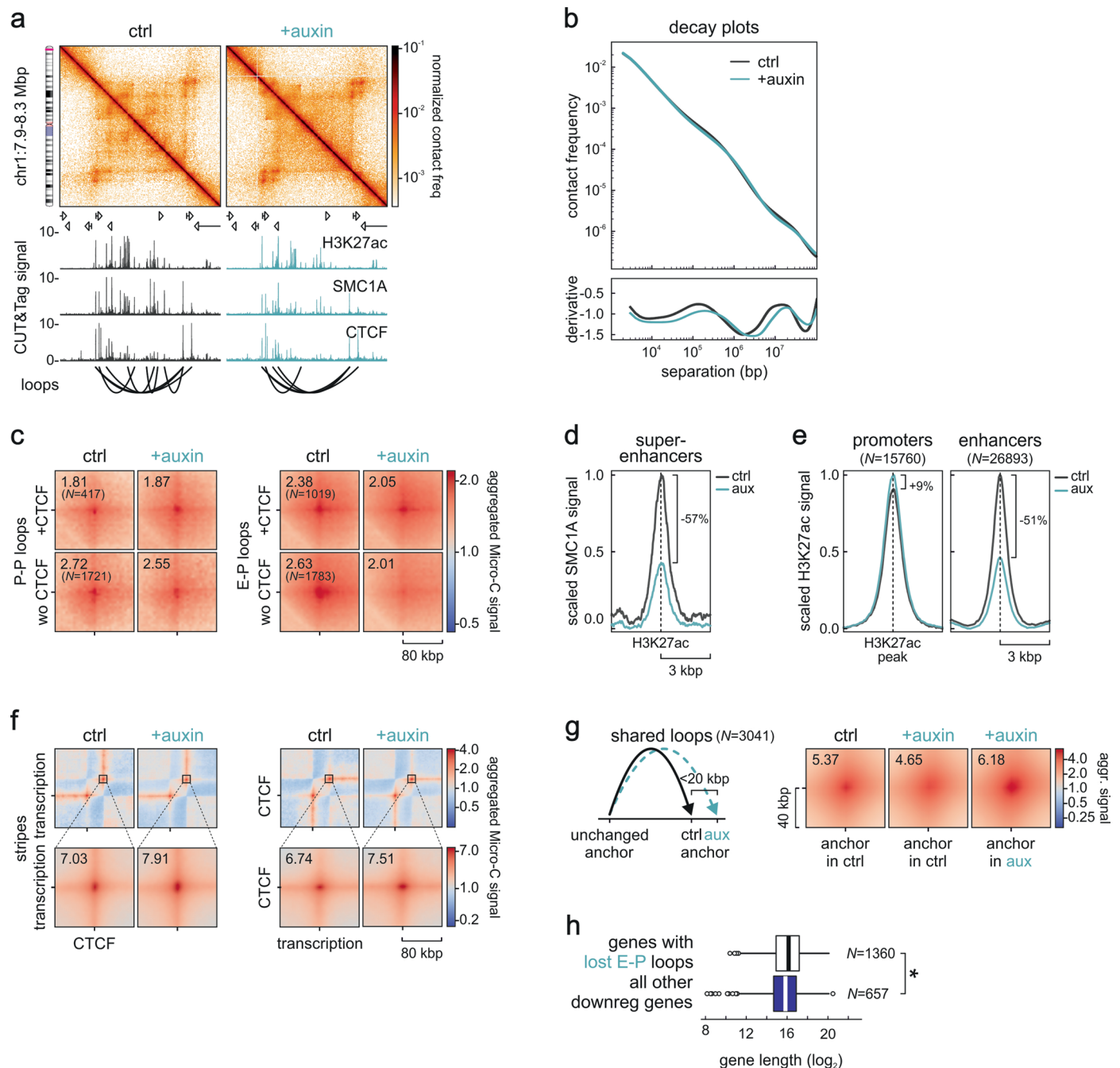
**Reprints and permissions information** is available at www.nature.com/reprints.
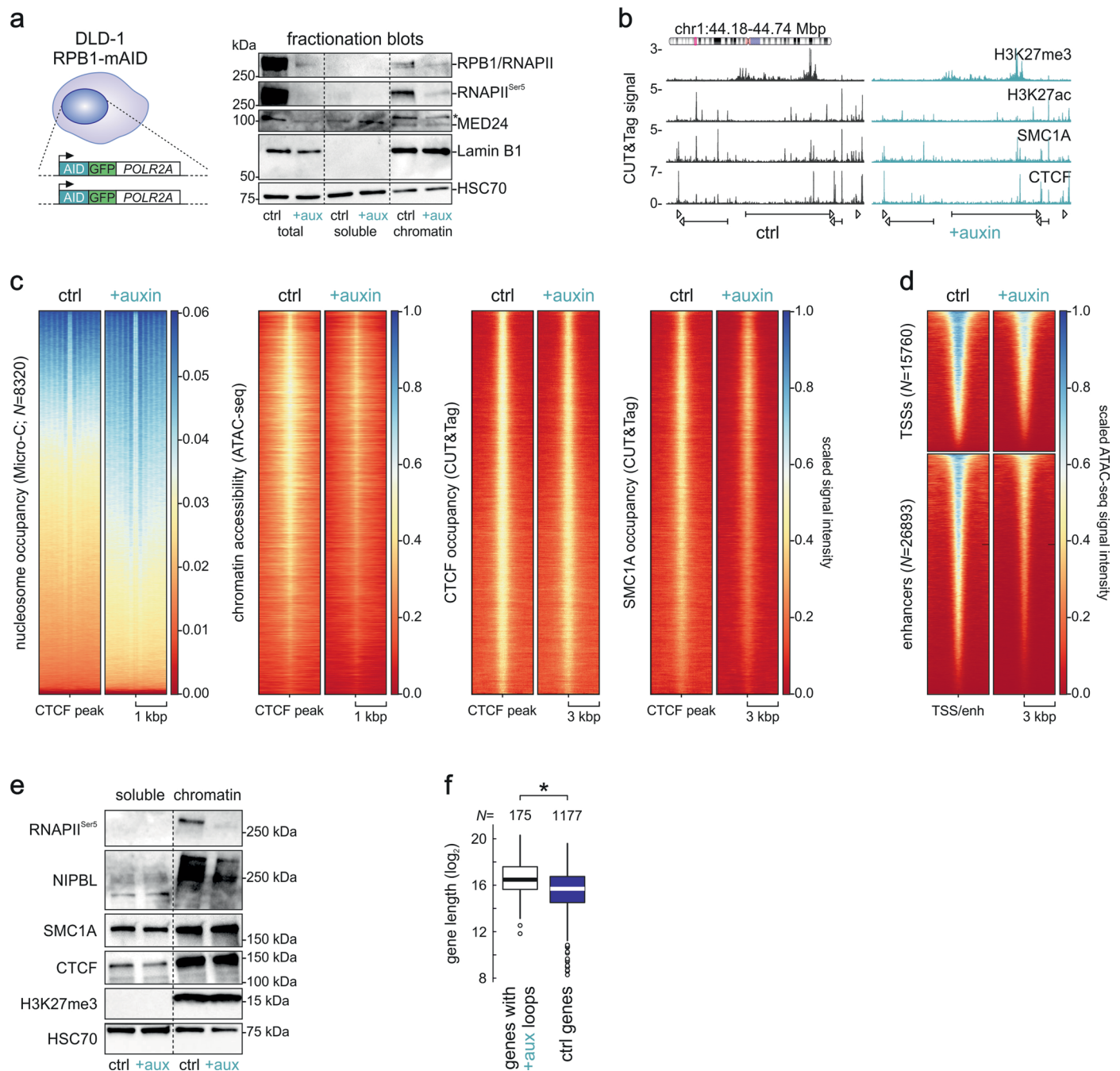
**Extended Data Fig. 1 | Effects of RNAPII depletion on chromatin organization and protein levels. a**, Left: Schematic of the biallelic tagging strategy in the endogenous *POLR2A* loci. Right: Fractionation blots showing the levels of RPB1 and Ser5-phosphorylated RNAPII, Mediator subunit 24, and Lamin B1 from DLD1-mAID-RBP1 cells treated or not with auxin to deplete RNAPII. HSC70 levels provide a control. Blots have been replicated at least twice. **b**, Representative tracks of CUT&Tag signal for H3K27me3, H3K27ac, SMC1A, and CTCF from control (left) and auxin-treated DLD1-mAID-RBP1 cells (right) along 0.55 Mbp of chr1. **c**, Heatmaps of nucleosome occupancy deduced from Micro-C data, of chromatin accessibility deduced from ATAC-seq, and of CTCF and SMC1A

occupancy deduced from CUT&Tag around CTCF loop anchors before (ctrl) and after RNAPII degradation (+auxin). **d**, As in panel c, but showing scaled ATAC-seq signal around gene promoters and enhancers. **e**, As in panel a, but for Ser5-phosphorylated RNAPII, NIPBL, SMC1A, CTCF, and H3K27me3 levels in the soluble and chromatin fractions of DLD-1 cells. HSC70 levels provide a control. Blots have been replicated at least twice. **f**, Boxplots depicting the distribution of genes containing (genes with +aux loops) or not (ctrl genes) gained loop anchors upon RNAPII depletion. In the plots, center lines represent the median value, box-limits the 25th and 75th percentiles, and whiskers extend 1.5x each box's interquartile range. *$P < 0.01$, two-sided Wilcoxon-Mann-Whitney test.

**Extended Data Fig. 2 | Effects of RNAPII depletion on the 3D organization of facultative heterochromatin. a**, Micro-C contact maps from control (left) and auxin-treated cells (right) in two exemplary genomic regions of chr1 at 2-kbp resolution aligned to H3K27me3, H3K27ac, CTCF, and SMC1A CUT&Tag signal tracks. Loops called for each region and condition are also shown by spider plots (bottom). **b**, Aggregate plots of all H3K27me3-anchored loops emerging in auxin-treated cells. **c**, Bar plot showing per cent of gained loops with one or two H3K27me3 anchors or with H3K27me3 in the next-door genomic bin (that is, within <10 kbp from the anchor).

**Extended Data Fig. 3 | Changes in loops and stripes following RNAPII depletion.** **a**, Micro-C contact maps from control (left) and auxin-treated cells (right) in an exemplary genomic region on chr1 at 4-kbp resolution aligned to H3K27ac, CTCF, and SMC1A CUT&Tag signal tracks. Loops called for each region and condition are also shown by spider plots (bottom). **b**, Plots of interaction frequency decay as a function of genomic distance from control and auxin-treated cells (top) and their first derivative (bottom). **c**, Aggregate plots of gene promoter-promoter (P-P) or enhancer-promoter loops (E-P) in control and auxin-treated cells that involve (+CTCF) or not CTCF (wo CTCF) in at least one anchor. **d**, Line plot showing mean SMC1A CUT&Tag signal from control and auxin-treated cells in the 6 kbp around

H3K27ac peaks from 590 super-enhancers. **e**, As in panel c, but for H3K27ac signal around active gene promoters or enhancers. **f**, Average plots showing mean signal of stripes with one CTCF and one transcriptional anchor before (ctrl) and after RNAPII depletion (+auxin). *Zoom-in*: Aggregate plots for loops at the end of the stripes. **g**, As in panel c, but for shared loops that rewire one anchor by <20 kbp (see cartoon). **h**, Boxplots depicting the lengths of genes downregulated upon RNAPII depletion that are linked (genes with lost E-P loops) or not (all other genes) to lost E-P loops. In the plots, center lines represent the medians, box-limits the 25th and 75th percentiles, and whiskers extend 1.5x each box's interquartile range. *: $P < 0.01$, two-sided Wilcoxon-Mann-Whitney test.

# nature portfolio

Corresponding author(s): A. Papantonis

Last updated by author(s): Feb 25, 2023

# Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

## Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

| n/a | Confirmed | |
|---|---|---|
| ☐ | ☒ | The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement |
| ☐ | ☒ | A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| ☐ | ☒ | The statistical test(s) used AND whether they are one- or two-sided<br>*Only common tests should be described solely by name; describe more complex techniques in the Methods section.* |
| ☒ | ☐ | A description of all covariates tested |
| ☒ | ☐ | A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| ☐ | ☒ | A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| ☐ | ☒ | For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted<br>*Give P values as exact values whenever suitable.* |
| ☒ | ☐ | For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| ☒ | ☐ | For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| ☒ | ☐ | Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated |

*Our web collection on statistics for biologists contains articles on many of the points above.*

## Software and code

Policy information about availability of computer code

| | |
|---|---|
| Data collection | No software was used for data collection |
| Data analysis | All code used for Micro-C data analysis in this study is available here: https://github.com/shuzhangcourage/HiC-data-analysis<br>For basic analysis the following standard tools were used:<br>- BWA ver. 0.7.17 (https://sourceforge.net/projects/bio-bwa/files/)<br>- HiCRep ver. 1.5.0 (https://github.com/TaoYang-dev/hicrep)<br>- HiCCUPS ver. 1.19.01 (https://github.com/aidenlab/juicer/wiki/CPU-HiCCUPS)<br>- HiGlass ver. 1.11.7 (https://github.com/higlass/higlass)<br>- Deeptools ver.2 (https://github.com/deeptools/deepTools/blob/develop/docs/index.rst)<br>- FIMO (https://meme-suite.org/meme/doc/fimo.html)<br>- Bowtie 2 ver. 2.4.4 (https://bowtie-bio.sourceforge.net/bowtie2/index.shtml)<br>- the CUT&Tag pipeline from the Henikoff lab (https://yezhengstat.github.io/CUTTag_tutorial/)<br>- coolpup.py (https://github.com/open2c/coolpuppy)<br>- the ROSE algorithm to detect super-enhancers (https://github.com/stjude/ROSE)<br>- chipseeker ver. 3.16 (https://bioconductor.org/packages/release/bioc/html/ChIPseeker.html)<br>- pgltools ver. 1.2.1 (https://github.com/billgreenwald/pgltools)<br>- cooltools ver. 0.5.4 (https://github.com/open2c/cooltools)<br>- the Dovetail Genomics pipeline for Micro-C data handling (https://github.com/open2c/cooltools)<br>- EspressoMD package ver. 4.3 (https://espressomd.github.io/doc/espressomd.html) |

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

## Data

NGS data generated in this study are available via the NCBI Gene Expression Omnibus repository under the accession number GSE178593. All other data used for analyses come from our previous study, Zhang et al., Science Advances 2021 (PMID: 34678064), and were already available under accession number GSE160321. The reference genome used for all data was hg38. Links are provided in our Data Availability statement.

## Human research participants

| | |
|---|---|
| Reporting on sex and gender | n/a |
| Population characteristics | n/a |
| Recruitment | n/a |
| Ethics oversight | n/a |

Note that full information on the approval of the study protocol must also be provided in the manuscript.

# Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences          ☐ Behavioural & social sciences          ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

# Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

| | |
|---|---|
| Sample size | Micro-C was performed in two replicates each consisting of 2-4 independent pooled libraries as is typical for such studies. In all other assays three independent replicates were produced (unless otherwise stated) as is the standard in the field. No sample size calculation was used. |
| Data exclusions | No data were excluded from any of our analyses. |
| Replication | Typically, three independently obtained replicates were used for all experiments, unless otherwise stated. All attempts at replication were consistently successful. |
| Randomization | Given the number of conditions tested and types of experiments used, randomization is not applicable. |
| Blinding | Micro-C datasets were analysed blindly and assigned "control" and "degron" etiquettes only at the very end. For all other experiments blinding was not applicable since all experiments were performed by one person (N.U.). |

# Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

## Materials & experimental systems

| n/a | Involved in the study |
|---|---|
| ☐ | ☒ Antibodies |
| ☐ | ☒ Eukaryotic cell lines |
| ☒ | ☐ Palaeontology and archaeology |
| ☒ | ☐ Animals and other organisms |
| ☒ | ☐ Clinical data |
| ☒ | ☐ Dual use research of concern |

## Methods

| n/a | Involved in the study |
|---|---|
| ☒ | ☐ ChIP-seq |
| ☐ | ☒ Flow cytometry |
| ☒ | ☐ MRI-based neuroimaging |

# Antibodies

**Antibodies used**
phospho-Ser5 RPB1 (Active Motif 61085; 1:2000), RPB1 (Abcam ab817; 1:500), NIPBL (Bethyl A301-779A; 1:10000), MED24 (Affinity Biosciences AF0346; 1:1000), Lamin B1 (Abcam ab16048; 1:10000), SMC1A (Abcam ab9262; 1:4000), CTCF (Active Motif 61311; 1:2000), H3K27me3 (Active Motif 39155; 1:1000) and HSC70 (Santa Cruz sc-7298; 1:2000).

**Validation**
All antibodies used are commercially validated via western blots against protein extracts from human cells, and our blotting profiles match those provided by the manufacturers. For updated information on the respective commercial validations please visit:
https://www.thermofisher.com/antibody/product/A301-779A.html (validated via WB)
https://www.affbiotech.com/goods-257-AF0346-MED24_Antibody.html (validated via WB)
https://www.abcam.com/smc1a-antibody-ab9262.html (validated via WB and ICH)
https://www.abcam.com/rna-polymerase-ii-ctd-repeat-ysptsps-antibody-chip-grade-ab26721.html (validated for WB, IF, IHC, ChIP)
https://www.abcam.com/lamin-b1-antibody-nuclear-envelope-marker-ab16048.html (validated for ICCIF, WB, IP)
https://www.activemotif.com/catalog/details/61085 (experimentally validated for ChIP-seq, WB, IP, IF, ICC)
https://www.activemotif.com/catalog/details/61311 (experimentally validated for ChIP-seq, CUT&Tag, WB, IP, IF, EMSA, ICC)
https://www.activemotif.com/catalog/details/39155 (experimentally validated for ChIP-seq, CUT&Tag, WB, IF, IHC, ICC)
https://www.scbt.com/p/hsc-70-antibody-b-6?productCanUrl=hsc-70-antibody-b-6&_requestid=2428416 (validated for WB, IHC, IF)

# Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

**Cell line source(s)**
DLD-1 mAID-RPB1 colon cancer cells were provided by the Maeshima lab.

**Authentication**
This is an ATCC commercially-validated line (via STR profiling and/or karyotyping).

**Mycoplasma contamination**
Cell stocks were tested regularly for mycoplasma and only used when negative (which was always the case here).

**Commonly misidentified lines** (See ICLAC register)
Our line is not found in the ICLAC register.

# Flow Cytometry

## Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

## Methodology

**Sample preparation**
DLD-1 cells (control or treated with auxin for 14 h) were used/ Inducible depletion of RPB1 initiated via treatment with doxycycline for 24 h to induce TIR1 expression, before addition of 500 μM indole-3-acetic acid solution ("auxin"; Sigma-Aldrich) for 14 h to induce RPB1 degradation. Cells treated with auxin were harvested, resuspended in 1 μg/ml propidium iodide (PI) before sorting out G1 cells.

**Instrument**
FACS Canto II flow cytometer (Becton Dickinson)

**Software**
No custom software was used, just the on-instrument interface by BD allowing standard PI-based cell cycle analysis.

**Cell population abundance**
Starting DLD-1 cell populations were always very similar in size and with no detectable cell death; end populations (G1-phase cells) were also comparable in size (approx. 51-58% of the starting material).

**Gating strategy**
Standard PI-based gating was used, whereby the amount of PI signal is used to discriminate between G1-, S- or G2/M-phase

Gating strategy

cells. Gating for either control or auxin-treated cells was consistently set to the sames values (see P1-P3 gates in Suppl. Fig. 1), since signal dispersion was almost identical between our conditions. Gating for PI-based cell cycle sorting is essentially automated on Canto II, and we abided to default parameters every time.

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.