ARTICLE

Check for updates

CTCF and transcription influence chromatin structure re-configuration after mitosis

Haoyue Zhang[1,2 ✉], Jessica Lam [2,3], Di Zhang [2,3], Yemin Lan[3], Marit W. Vermunt [2], Cheryl A. Keller [4], Belinda Giardine [4], Ross C. Hardison[4] & Gerd A. Blobel [2,3 ✉]

During mitosis, transcription is globally attenuated and chromatin architecture is dramatically reconfigured. We exploited the M- to G1-phase progression to interrogate the contributions of the architectural factor CTCF and the process of transcription to genome re-sculpting in newborn nuclei. Depletion of CTCF during the M- to G1-phase transition alters short-range compartmentalization after mitosis. Chromatin domain boundary re-formation is impaired upon CTCF loss, but a subset of boundaries, characterized by transitions in chromatin states, is established normally. Without CTCF, structural loops fail to form, leading to illegitimate contacts between cis-regulatory elements (CREs). Transient CRE contacts that are normally resolved after telophase persist deeply into G1-phase in CTCF-depleted cells. CTCF loss-associated gains in transcription are often linked to increased, normally illegitimate enhancer-promoter contacts. In contrast, at genes whose expression declines upon CTCF loss, CTCF seems to function as a conventional transcription activator, independent of its architectural role. CTCF-anchored structural loops facilitate formation of CRE loops nested within them, especially those involving weak CREs. Transcription inhibition does not significantly affect global architecture or transcription start site-associated boundaries. However, ongoing transcription contributes considerably to the formation of gene domains, regions of enriched contacts along gene bodies. Notably, gene domains emerge in ana/telophase prior to completion of the first round of transcription, suggesting that epigenetic features in gene bodies contribute to genome reconfiguration prior to transcription. The focus on the de novo formation of nuclear architecture during G1 entry yields insights into the contributions of CTCF and transcription to chromatin architecture dynamics during the mitosis to G1-phase progression.

[1] Institute of Molecular Physiology, Shenzhen Bay Laboratory, Shenzhen, Guangdong, China. [2] Division of Hematology, The Children's Hospital of Philadelphia, Philadelphia, PA, USA. [3] Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA. [4] Department of Biochemistry and Molecular Biology, Pennsylvania State University, University Park, PA, USA. ✉email: zhang_adam@szbl.ac.cn; blobel@email.chop.edu

The mitotic phase of the cell cycle is characterized by rapid and extensive re-organization of chromatin architecture and global attenuation of transcription[1–5]. Studies of chromatin dynamics during entry into and exit from mitosis have informed the mechanistic basis underlying the hierarchical organization of chromatin. During mitotic exit, A/B compartmentalization is detectable as early as in ana/telophase but intensifies and expands thereafter[1–5]. Contacts between CREs such as promoters and enhancers are re-established with variable kinetics, some forming gradually and plateauing deeper into G1 while others are transient in nature being developed fully in ana/telophase only to fade upon G1-entry[1]. Resumption of transcription follows similarly variable characteristics: some genes display a spike in activity early in G1 and settle down at later stages whereas others are activated in a gradual fashion[1,6,7].

The multi-functional transcription factor CTCF frequently co-localizes with boundaries of contact domains, such as topologically associated domains (TADs), and is proposed to assist their formation in collaboration with the cohesin ring complex through a process termed "loop extrusion"[8–10]. Accordingly, acute depletion of CTCF or cohesin leads to widespread weakening of boundaries in interphase cells[11–15]. CTCF and cohesin are evicted from mitotic chromatin to varying extents[1,16–19], and measuring the rates by which they return to chromatin has enabled correlative assessments of their roles in post-mitotic genome folding and transcriptional activation. Upon mitotic exit, CTCF is immediately recruited back to chromatin prior to the formation of domain boundaries and architectural loops[1]. The rate-limiting step in the formation of these latter structures appears to be the accumulation of cohesin at CTCF-bound sites, which occurs more gradually as chromatid extrusion proceeds.

Another feature frequently associated with domain boundaries is transcription start sites (TSS)[20,21], but the role of transcription in boundary formation is still being debated. Inhibition of transcription compromises boundary strength in *Drosophila melanogaster* embryos[20,21]. Yet, neither genetic nor chemical inhibition of transcription elicited a significant impact on higher-order structures of mammalian genomes[22,23]. A recent gain-of-function study demonstrated that ectopic insertions of TSSs can lead to the formation of new domain-like structures spanning the lengths of the de novo transcripts[24].

Most studies on how CTCF depletion or transcription inhibition impact chromatin architecture were carried out in asynchronously growing cells and thus do not distinguish requirements for establishment versus maintenance of genome structure[11,13–15,20–22,25]. The transition from mitosis into G1-phase offers a chance to examine genome refolding in relation to CTCF binding and gene activation. Here, we interrogated the contributions of CTCF and the process of active transcription to the establishment of post-mitotic chromatin architecture by acutely depleting CTCF through the auxin-inducible degron (AID) system alone or jointly with chemical inhibition of transcription during the mitosis to G1-phase transition[1,11,26]. We demonstrate that CTCF loss alters short-range compartmental interactions after mitosis, suggesting a previously underappreciated role for CTCF in genome compartmentalization. CTCF is required for the proper re-formation of contacts among regulatory elements and normal transcription reactivation. Active transcription contributes minimally to higher-order chromatin re-organization. However, gene domains were shaped by the concerted action of transcription elongation and pre-existing epigenetic features along gene bodies. In sum, our findings elucidate how CTCF and transcription impact post-mitotic genome architectural dynamics.

Results

Cell cycle stage-specific degradation of CTCF. To explore the impact of CTCF loss specifically during the period when chromatin architecture is rebuilt, we employed a murine erythroblast line G1E-ER4 in which both CTCF alleles were engineered to contain a C-terminal fusion to the AID-mCherry domains (Supplementary Fig. 1a)[11,26]. A TIR-expressing construct was transduced into the cells to allow for rapid auxin-induced CTCF degradation (Supplementary Fig. 1a). CTCF became virtually undetectable after 1 h exposure to auxin (Supplementary Fig. 1b). The acute nature of AID-mediated degradation enabled the removal of CTCF at precisely chosen time points. We applied auxin during nocodazole-induced prometaphase-arrest/release (Fig. 1a). CTCF-depleted cells were enriched by FACS at defined time points during the prometaphase-to-G1 phase transition on the basis of mCherry fluorescence signal and DNA content staining (Supplementary Fig. 1c)[1,27]. To facilitate FACS purification of cells at ana/telophase, we adopted GFP fused to a mitotic-specific degron (MD) (Supplementary Fig. 1a). Purified cells were processed for in situ Hi-C to detect architectural features (Supplementary Fig. 1d; Supplementary Data 1). Short-term depletion of CTCF did not impede cell cycle progression, and accordingly, post-mitotic Hi-C contact decay curves were highly similar between auxin treated and control cells (Supplementary Fig. 1c; 2a), enabling pair-wise comparisons between CTCF depleted and replete cells at each post-mitotic cell cycle stage.

Local compartmentalization after mitosis requires CTCF. A/B compartment emergence in ana/telophase, as well as expansion and intensity gains, occurred at comparable rates in control and CTCF-depleted cells, suggesting that CTCF is dispensable for global compartmentalization after mitosis (Fig. 1b; Supplementary Fig. 2b, d–g). Since prior reports suggested that loop extrusion counteracts compartmentalization[12,28], we examined whether CTCF loss and the resulting extended extrusion process may affect short-range compartmental interactions. We analyzed the post-mitotic interaction frequency between consecutive B-type compartment domain pairs flanking a single A-type compartment domain and found it to be significantly increased upon CTCF depletion (Fig. 1c–e; Supplementary Fig. 3a). It is noteworthy that such increment of short-range B–B interactions grew as cells proceeded towards G1 (Fig. 1e; Supplementary Fig. 3h), in line with the progressive loading of cohesin after mitosis. Moreover, gains of B–B interactions only occurred locally and tapered off as they were further separated (Supplementary Fig. 3e–h), consistent with the limited residence time of cohesin on chromatin[29]. Intriguingly, interactions between short-range A-type compartment domains were diminished in CTCF-deficient cells (Fig. 1f–h; Supplementary Fig. 3b, e, g, i). Together, our data highlight a previously undescribed role of CTCF to regulate short-range compartmental interactions, likely by restricting cohesin-driven loop extruding.

CTCF dependent and independent mechanisms drive boundary reformation after mitosis. We next investigated post-mitotic boundary reformation upon CTCF depletion. We identified 6376 boundaries with high concordance among biological replicates (Supplementary Fig. 4a, b; Supplementary Data 2)[30]. *K*-means clustering yielded five groups of boundaries with distinct sensitivities to CTCF depletion (Fig. 2a). Only ~20% of boundaries (cluster1) were fully dependent on CTCF after mitosis (Fig. 2a; Supplementary Fig. 4c, f). ~31.9% (cluster2) were partially dependent, and ~27% (cluster3) were unaffected by CTCF loss (Fig. 2a; Supplementary Fig. 4d, e, g, h) indicative of CTCF-independent mechanisms. As expected, cluster3 boundaries
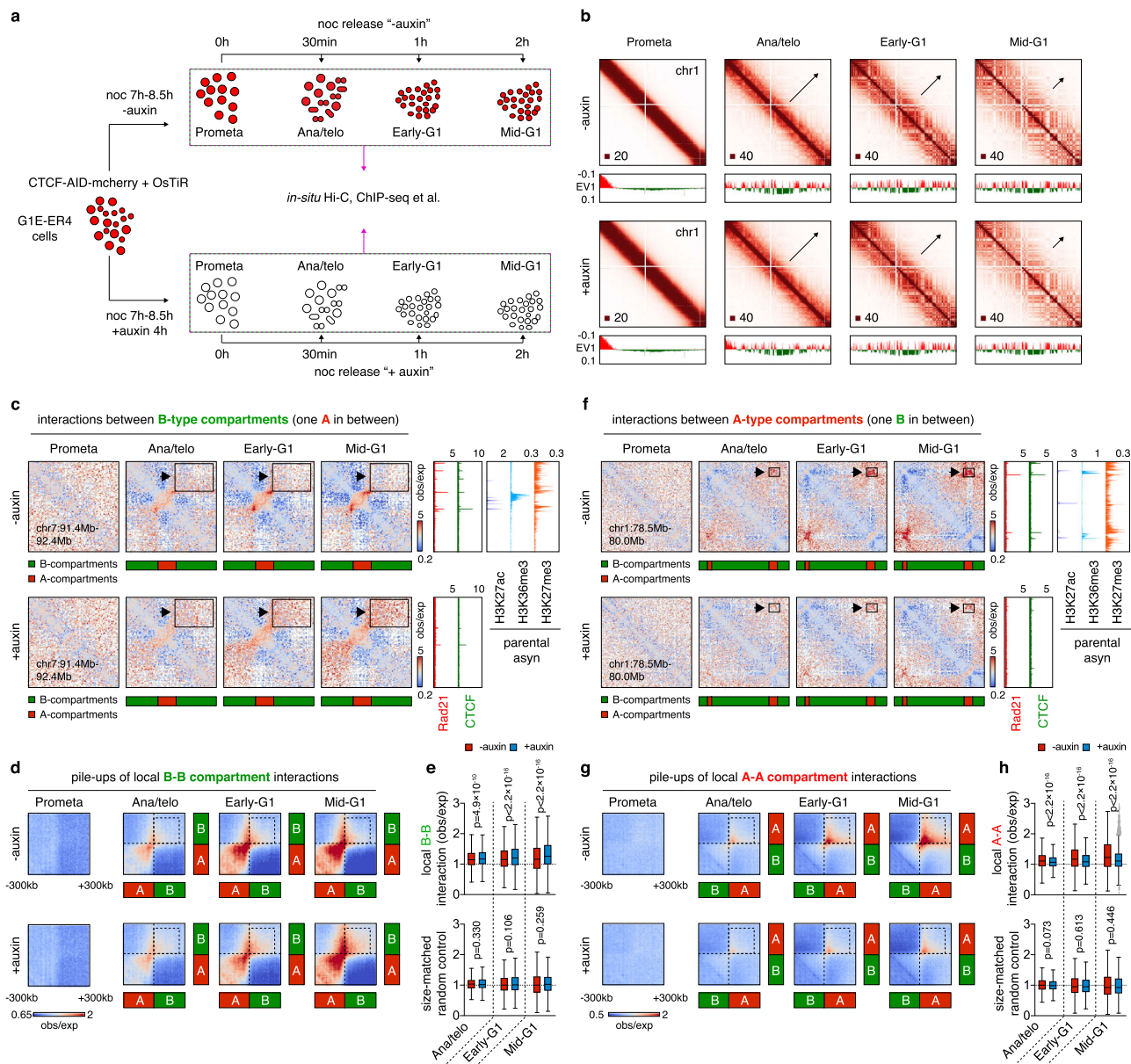
Fig. 1 Alteration of local compartmentalization upon CTCF removal. a Strategy for harvesting mitotic and post-mitotic populations with or without CTCF. **b** KR balanced Hi-C contact matrices showing global compartment reformation of chr1 in untreated and auxin-treated cells after mitosis. Bin size: 100 kb. Black arrows indicate the progressive spreading of compartments throughout the entire chromosome. Browser tracks with compartment PC1 values are shown for each contact map. **c** KR balanced Hi-C contact matrices showing representative local B–B interaction changes with or without CTCF depletion after mitosis. Bin size: 10 kb. Arrows and boxes highlight the increased local B–B interactions after CTCF depletion across cell cycle stages. Tracks of CTCF and Rad21 with or without auxin treatment, as well as histone, marks H3K27ac, H3K36me3, and H27me3 are from asynchronous G1E-ER4 cells. **d** Pile-up Hi-C matrices showing the increased local interactions between all consecutive (with one A-type in between) B-type compartment domains. Bin size: 10 kb. Dotted boxes indicate the increased local B–B interactions genome-wide. **e** Upper panel: Boxplots showing quantification of interactions in the dotted boxes (250 kb × 250 kb) in (**d**) ($n = 1604$ pairs of short-range B–B interactions). Lower panel: Boxplots showing the effect of CTCF depletion on the interactions between randomly selected genomic pairs ($n = 465$) that are distance-matched to the upper panel. Boxplots present upper and lower quartiles with the centerline as the median. Whiskers denote 1.5 × interquartile range (IQR). *P* values were calculated using a two-sided paired Wilcoxon signed-rank test. **f, g** Similar to **c** and **d**, showing examples and pile-ups of local consecutive (with one B-type in between) A–A interactions genome-wide. **h** Upper panel: Boxplots showing quantification of interactions in the dotted boxes (250 kb × 250 kb) in (**g**) ($n = 1612$ pairs of short-range A–A interactions). Lower panel: Boxplots showing the effect of CTCF depletion on the interactions between randomly selected genomic pairs ($n = 469$) that are distance-matched to the upper panel. Boxplots present upper and lower quartiles with the centerline as the median. Whiskers denote 1.5 × interquartile range (IQR). *P* values were calculated using a two-sided paired Wilcoxon signed-rank test.

displayed markedly lower CTCF/cohesin occupancy than cluster1 and 2 (Fig. 2b). Visual inspection suggested that cluster2 and 3 boundaries are frequently located at transitions between regions enriched for the repressive histone mark H3K27me3 and the transcription elongation mark H3K36me3 (Supplementary

Fig. 4i). We further quantified chromatin state transitions by principal component analysis (PCA) using 10 kb binned H3K27me3 and H3K36me3 ChIP-seq signals across the ±50 kb region of each boundary (Fig. 2c). We found that the extreme (top or bottom 20%) PC1 projections reliably predicted
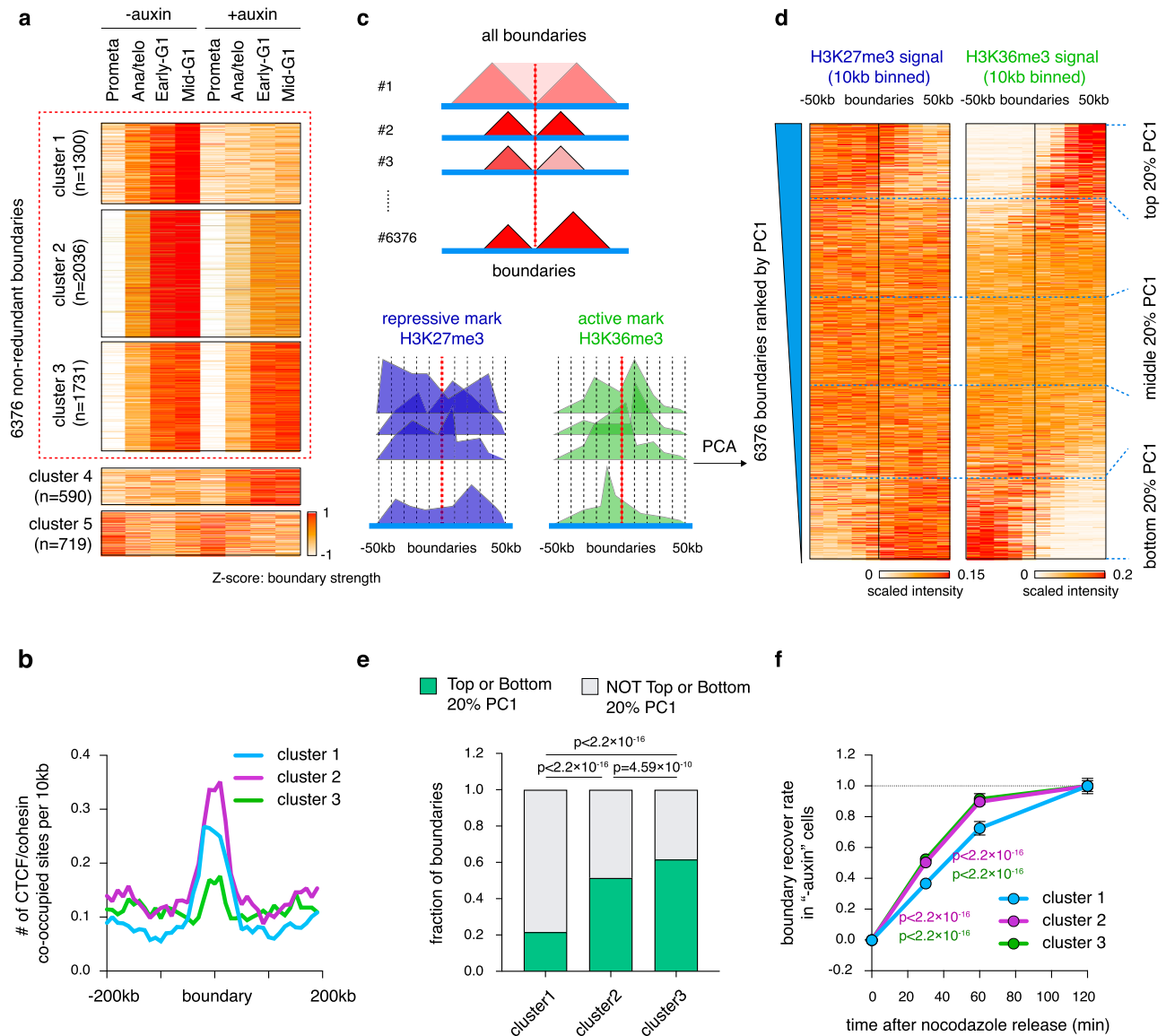
Fig. 2 Reformation of boundaries displays distinct responses to CTCF loss. a *k*-means clustering of boundaries depending on their sensitivity to CTCF depletion. The z-scores in prometaphase for cluster1 boundaries should not be interpreted as the absolute insulation intensity, because they are calculated as relative values across all time points (see the absolute insulation intensity in Supplementary Fig. 4f–h). **b** Average occupancy of CTCF/cohesin peaks per 10 kb for boundaries from cluster1–3. **c** Schematic of the principal component analysis (PCA)-based method using the H3k36me3 and H3K27me3 histone marks to assess boundaries as defined here as chromatin state transitions. **d** ChIP-seq signal intensities of H3K27me3 and H3K36me3 in a 100 kb window centered on boundaries. Boundaries were ranked by their PC1 projections in descending order. The top and bottom regions (20%) of the heatmap indicate the transition of chromatin state from 5′ inactive to 3′ active and 5′ active to 3′ inactive, respectively. **e** Bar graphs showing the fraction of boundaries from each cluster with top or bottom 20% PC1 values. *P* values were computed by two-sided Fisher's exact test. $n = 282, 1050, 1068$ for cluster1, 2, and 3 boundaries with top or bottom 20% PC1 respectively. $n = 1018, 986, 663$ for cluster1, 2, and 3 boundaries not with top or bottom 20% PC1, respectively. **f** Line graph showing the kinetics of boundary formation of cluster1–3 in untreated cells. Error bars denote 95% confidence interval. Purple and green colored *P* values are calculated from comparisons between cluster1 ($n = 1300$) and cluster2 ($n = 2036$) or cluster3 ($n = 1731$) boundaries, respectively. Two-sided Wilcoxon signed-rank test.

transitions of chromatin states at boundaries (Fig. 2d). Importantly, cluster2 and 3 boundaries are significantly more enriched with top and bottom 20% PC1s compared to cluster 1 (Fig. 2e). Therefore, cluster1 boundaries seem to be primarily driven by CTCF/cohesin mediated loop extrusion, Cluster3 boundaries may be formed through segregation of active and inactive chromatin, and cluster 2 boundaries by both mechanisms. Intriguingly, under normal conditions ("-auxin"), cluster2 and 3 boundaries were reformed significantly faster than cluster1 after mitosis (Fig. 2f), suggesting that chromatin segregation may mediate more rapid insulation than loop extrusion after mitosis. Our data

demonstrate that CTCF dependent and independent mechanisms can work separately or jointly to drive boundary formation after mitosis.

Variable requirement of CTCF for post-mitotic loop formation. CTCF is frequently found at chromatin domain boundaries and anchors of architectural loops where it can promote or inhibit loop contacts among regulatory regions, such as enhancers and promoters. We stratified chromatin loops based on the composition of their loop anchors and asked to what extent they
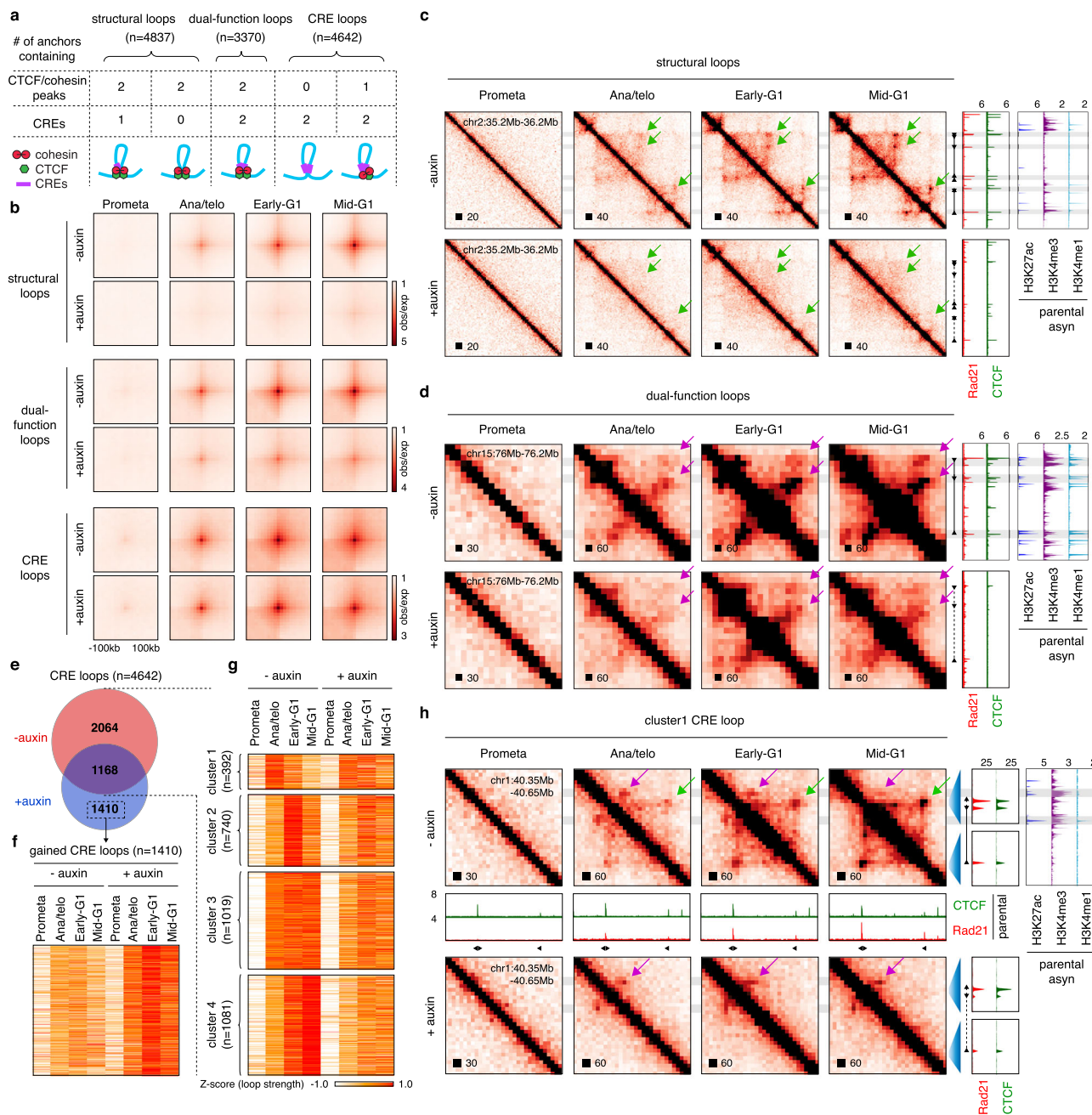
Fig. 3 CTCF loops constrain CRE contacts after mitosis. a Schematic showing the stratification of loops ("structural loops", "dual-function loops", and "CRE loops") based on the presence at their anchors of CTCF/cohesin co-occupied sites and CREs. CRE denotes cis-regulatory element. **b** APA plots showing the signals of loop categories before and after CTCF depletion across cell cycle stages. Bin size: 10 kb. **c** KR balanced Hi-C contact matrices of representative regions containing structural loops. Bin size: 10 kb. Tracks of CTCF and Rad21 with or without auxin treatment as well as H3K27ac, H3K4me3, and H3K4me1 were from asynchronous G1E-ER4 cells. **d** Similar to (**c**), KR balanced Hi-C contact matrices of representative regions containing dual-function loops. Bin size: 10 kb. **e** Venn diagram of CRE loops. **f** Heatmap displaying intensities of the 1410 newly gained loops after CTCF depletion. **g** Heatmap showing the result of k-means clustering on the 3232 CRE loops detected in untreated control samples. **h** Similar to **c**, **d**, KR balanced Hi-C contact matrices of a representative region containing a cluster1-P transient CRE loop. Additional tracks of CTCF and Rad21 from parental cells across designated cell cycle stages are shown[1].

are rebuilt upon CTCF depletion. A modified HICCUPS algorithm[1,31] identified a union of 16,370 loops across all time points and auxin treatment conditions with high concordance (Supplementary Fig. 5a, b; Supplementary Data 3). Newly called loops were also appreciated visually at each cell cycle stage, supporting the validity of our loop calling method (Supplementary Fig. 5c). Among all loops, 8207 (~50%) harbor CTCF/cohesin co-occupied sites at both anchors. These were further

sub-categorized into 4837 "structural loops" with one or no anchors containing CREs (as defined previously[1]), and 3370 "dual-function loops" with both anchors containing CREs (Fig. 3a). The post-mitotic reemergence of both structural and dual-function loops was severely disrupted upon CTCF loss as evidenced by aggregated peak analysis (APA) and PCA (Fig. 3b–d; Supplementary Fig. 5d). We also called 4642 "CRE loops" with both anchors containing CREs and only one or no

anchor harboring CTCF/cohesin peaks (Fig. 3a). While APA plots failed to reveal major changes in CRE loop establishment, ~44.5% and ~30.4% of CRE loops were lost and newly gained, respectively, after auxin treatment (Fig. 3b, e, f), uncovering a considerable shift in their reformation after CTCF depletion. Our results thus point to a critical role for CTCF in the formation of diverse loop categories after mitosis.

Transient post-mitotic CRE loops are terminated by interfering structural loops. Previously, we uncovered a group of transient CRE loops whose intensities spiked in ana/telophase and subsequently faded in G1[1]. The function of such transient contacts if any, and the causative role for CTCF in their disruption remained unknown. Our acute CTCF depletion system enabled us to test this question globally. Using k-means clustering, we identified 392 transient CRE loops (cluster1) in control cells (Fig. 3g). Strikingly, 228 of these CRE loops persisted deep into G1 after CTCF depletion (cluster1-P), implying that CTCF is capable of blocking a subset of early-established CRE contacts after mitosis (Supplementary Fig. 6a, b). 164 CRE loops maintained their transient nature in the absence of CTCF (cluster1-NP) (Supplementary Fig. 6a, b). Visual examination of a representative cluster1-P CRE loop revealed the emergence of an interfering structural loop in control but not in auxin-treated cells (Fig. 3h). A high fraction (144, ~63.1%) of cluster1-P loops is potentially interrupted by structural loops, whereas remaining clusters (1-NP, 2, 3, and 4) were less affected (Supplementary Fig. 6c). Notably, the 1410 newly gained CRE loops after CTCF loss also showed a high level of structural loop interruption, comparable to that of cluster1-P (Supplementary Fig. 6c). This observation held true after randomized matching of CRE loop sizes from different clusters (Supplementary Fig. 6d).

CTCF resumes full chromatin occupancy in ana/telophase while cohesin accumulation is delayed[1]. Interference of CRE loops by structural loops occurs at later cell cycle stages (Fig. 3h), suggesting that CTCF by itself is insufficient to block CRE interaction but requires engagement in cohesin-mediated looped contacts. To test this idea, we performed genome-wide enrichment analysis to compare the likelihood of CRE loops to be disrupted by structural loops versus "loop-free" CTCF/cohesin co-occupied sites (Supplementary Fig. 6e). In comparison to structural loops, "loop-free" CTCF/cohesin sites displayed a reduced tendency to disrupt cluster1-P and the gained CRE contacts (Supplementary Fig. 6f). Furthermore, CTCF peaks independent of cohesin and structural loops were evenly distributed across all CRE loop clusters (Supplementary Fig. 6f). Together, dissection of the cell cycle dynamics of different loop categories uncovered support for the notion that CTCF is more effective as an insulator when part of a looped structure. A caveat of this interpretation is that small structural loops are undetectable in our Hi-C data.

The interposition of structural loops does not always disrupt CRE contacts: ~37.5% of the CRE loops in cluster2–4 had interposed structural loops. The failure of these to break up CRE loops might be related to their relative positioning. We observed that structural loops that are capable of weakening cluster 1-P CRE loops tended to reside near (~50 kb) the CREs (Fig. 3h). To quantify this trait, we measured the distance of the influenced CRE to the most proximal structural loop anchor inside the CRE loop: s_i-min (Supplementary Fig. 7a). Remarkably, s_i-min was significantly shorter for cluster1-P and the gained CRE loops compared to the non-insulated cluster2–4 (Supplementary Fig. 7b). This observation holds true for size-matched CRE loops (Supplementary Fig. 7c). Moreover, we observed a significant negative correlation between insulation strength and s_i-min after

mitosis (Supplementary Fig. 7d). Together, our data suggest that the insulating, CRE loop disrupting the function of CTCF is linked not only to CTCF's ability to form a loop but also to the relative position of the insulating loop.

Structural loops can facilitate CRE connectivity after mitosis. CTCF depletion weakened a subset of CRE loops after mitosis (Fig. 3e), suggesting a supportive role of CTCF for certain CRE contact formation. APA plots demonstrated that the slower-forming cluster4 CRE loops were attenuated the most in the G1 phase when compared to the other clusters (Supplementary Fig. 8a, b). Interestingly, cluster4 CRE loops were significantly more likely to reside within structural loops compared to size-matched ones from other clusters, implying a supportive role of structural loops for CRE-contacts (Supplementary Fig. 8c). Accordingly, the strengthening effects on cluster4 increased with cell cycle progression, consistent with the progressive reformation of structural loops after mitosis (Supplementary Fig. 8d). It is noteworthy that anchors of cluster4 CRE loops displayed significantly weaker decoration with active histone marks H3K27ac, H3K4me1, and H3K4me3 than cluster3 (Supplementary Fig. 8e), suggesting that weak CRE contacts are especially reliant on encompassing structural loops.

Reduced enhancer–promoter interactions do not account for transcription loss upon CTCF depletion. Given that CTCF loss altered post-mitotic CRE loop reformation, we examined how these changes impact gene reactivation after mitosis. We generated PolII ChIP-seq datasets during the mitosis-to-G1 phase transition with or without auxin treatment with high concordance among biological replicates (Supplementary Fig. 9a). We identified 7238 active genes across all time points (Supplementary Data 4). ~52.0% of these genes showed post-mitotic transcriptional spiking in control cells (Supplementary Fig. 9b), as observed previously[1,6]. This spiking pattern was overwhelmingly maintained in the absence of CTCF (Supplementary Fig. 9b–d), suggesting that CTCF is not essential for the re-activation of many genes, and validating that CTCF-deficient cells progress normally from prometaphase to G1. Consistent with previous reports, only a small fraction of genes (426, ~5.7%) were differentially expressed in at least one post-mitotic time point ($q < 0.05$, fold change > 1.25 fold) with 203 up-regulated and 223 down-regulated after mitosis (Fig. 4a) (see the "Methods" section). Of note, most of these gene expression changes were already detectable in the early-G1 phase, when transcription initiates, suggesting an instant effect of CTCF on these genes (Fig. 4a). The genes most down-regulated upon CTCF loss displayed the highest CTCF occupancy at their TSS in CTCF replete cells (Fig. 4b, c; Supplementary Fig. 9e, f), which could be explained by CTCF functioning as a direct transcription activator, or by mediating contacts with distal enhancers.

To distinguish between these possibilities, we implemented the activity-by-contact (ABC) model to call high confidence enhancer–promoter (E–P) contacts[32] (Fig. 4d). Using a stringent ABC score threshold of 0.04 (see the "Methods" section), we identified 7725 E–P pairs associated with active genes. Unexpectedly, E–P pairs associated with down-regulated genes showed no significant reduction of contact intensity upon CTCF loss across all tested cell cycle stages compared to non-regulated genes (Fig. 4e; Supplementary Fig. 10a). This argues against CTCF-dependent looping as a predominant mechanism to normally activate these genes and suggests that CTCF functions as a transcriptional activator near the TSS.
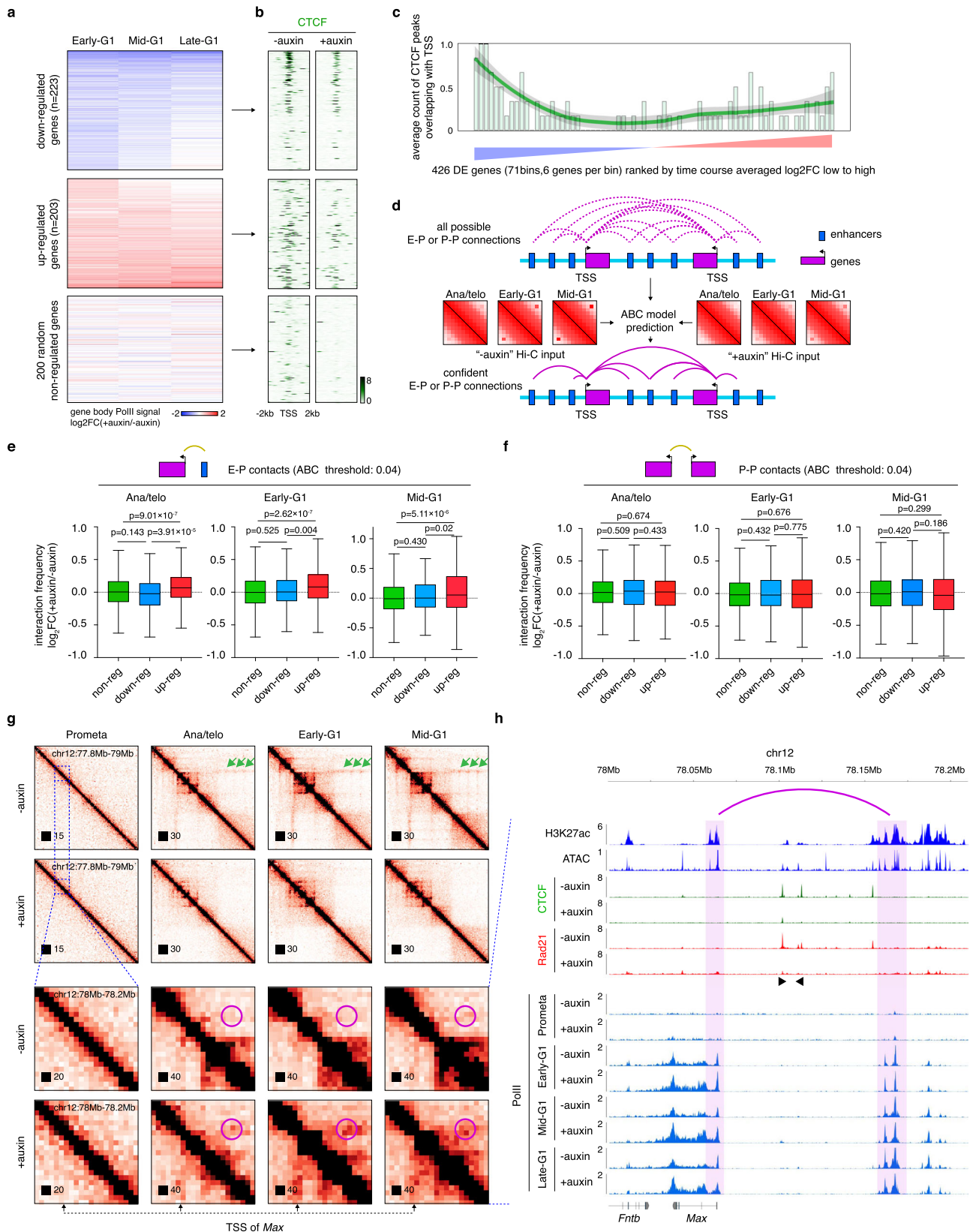
Gained enhancer–promoter interactions account for gene activation upon CTCF removal. We next explored how CTCF loss could lead to up-regulation of genes. We found that E–P pairs associated with up-regulated genes were significantly strengthened in the absence of CTCF after mitosis (Fig. 4e, g, h; Supplementary Fig. 10a). To quantify to what extent genes were regulated by distal enhancers, we set out to identify significantly altered E–P pairs (differential E–P interaction analysis) upon CTCF depletion. ~20.2% of up-regulated genes (e.g. Max) displayed significantly strengthened E–P interactions after CTCF loss, while only ~6% and ~0.8% of non-regulated and down-regulated genes respectively were associated with increased E–P

Fig. 4 CTCF loss alters transcription reactivation profiles after mitosis. a Heatmap displaying differentially expressed genes based on PolII ChIP-seq read counts over the gene bodies (+500 from TSS to TES), plotted as log2 fold-change (FC). **b** Meta-region plots of CTCF ChIP-seq signals from asynchronous cells before and after auxin treatment, centered on down-regulated, up-regulated, or 200 random non-regulated gene TSS. Plots were generated by Deeptools (2.5.4). **c** Quantification of (**b**) showing the number of CTCF peaks overlapping with TSS. The green line represents lowess smoothing of bar plots. Error band denotes 95% confidence interval. **d** Schematic showing the implementation of the ABC (activity by contact) model to predict confidently E–P (enhancer–promoter) and P–P (promoter–promoter) interactions using input asynchronous H3K27ac ChIP-seq and ATAC-seq data from G1E-ER4 cells as well as in-situ Hi-C datasets from this study. **e** Boxplots showing the $\log_2$ fold change upon CTCF depletion of interaction strength of E–P pairs (ABC score cutoff = 0.04) associated with either non-regulated ($n = 7211$ E–P pairs), down-regulated ($n = 188$ E–P pairs), or up-regulated ($n = 318$ E–P pairs) genes. $\log_2$ fold change of interaction strength was calculated using the LIMMA R package for each cell cycle stage. Boxplots present upper and lower quartiles with the centerline as the median. Whiskers denote 1.5 × interquartile range (IQR). *P* values were calculated using a two-sided Wilcoxon signed-rank test. **f** Similar to (**e**) showing the interaction changes of P–P pairs after CTCF depletion. $n = 11,117$, 349, and 291 P–P pairs for non-regulated, down-regulated, and up-regulated genes respectively. *P* values were calculated using a two-sided Wilcoxon signed-rank test. **g** KR balanced Hi-C contact matrices showing the *Max* locus across cell cycle stages in control and auxin-treated samples. Bin size: 10 kb. Green arrows indicate the structural loops that insulate the *Max* promoter from a nearby enhancer. Purple circles demarcate the increase of interactions between the *Max* promoter and a nearby enhancer upon CTCF depletion after mitosis. Note that the gain in interactions occurs at the earliest tested time point. **h** ChIP-seq genome browser tracks of the same region as that shown in the lower panel in (**g**). Note increased expression of *Max* after mitosis in auxin-treated samples. Purple arch annotates the elevated interaction between the *Max* promoter and the nearby enhancer. Black arrows indicate the motif orientation of CTCF-binding sites.

contacts (Supplementary Fig. 10b). This suggests that CTCF attenuates gene expression by interfering with E–P interactions. Accordingly, E–P pairs associated with up-regulated genes were more likely to be interrupted by structural loops, confirming an insulating role of structural loops in gene regulation (Supplementary Fig. 10c). The above observations were held across various ABC score thresholds (0.01–0.05), ruling out potential bias due to thresholding (Supplementary Fig. 10d, e). We also identified 11,766 promoter–promoter (P–P) pairs, which were essentially unchanged in all groups (up, down, non-reg) of genes (Fig. 4f; Supplementary Fig. 10a, e). This suggests that, in contrast to E–P contacts, P–P interactions contribute little to post-mitotic transcription reactivation. In sum, while overall CTCF depletion exerts modest effects on post-mitotic gene activation, proper regulation of some genes requires CTCF for their activation while others require it for shielding them from inappropriate enhancer influence.

Compartment and boundary reformation was independent of transcription. The role of transcription in chromatin architecture is being debated. It was previously proposed that inhibition of transcription does not compromise boundary strength, while others reported that induction of transcription may lead to boundary formation at TSS and compartmental interaction changes[25]. We next sought to investigate whether transcription facilitates post-mitotic compartment and boundary reformation. We treated cells with triptolide, a drug that inhibits transcription initiation, during the mitosis-to-G1 phase transition (Supplementary Fig. 11a)[33,34], followed by in situ Hi-C. Reformation of A/B compartments was ostensibly unperturbed in G1 upon transcription inhibition (Supplementary Fig. 2b–d, f). The CTCF independent cluster3 boundaries also remained intact after transcription inhibition (Supplementary Fig. 4h), ruling out active PolII complexes as underlying mechanisms. Moreover, insulation at TSS was unperturbed after transcription inhibition (Supplementary Fig. 11b). Consistent with this notion, insulation was progressively gained after mitosis at the TSS of the 100 most spiking genes even after their transcription was dialed down, suggesting that insulation, as it occurs at TSS, can be uncoupled from the process of transcription (Supplementary Fig. 11c–e). To examine whether transcription activity contributes to CTCF loss-induced chromatin changes, we removed CTCF and blocked transcription re-initiation simultaneously in mitosis and examined compartmentalization in the G1 phase (Supplementary Fig. 11a). CTCF loss-induced alterations of local compartmental

interactions (gain of B–B and loss of A–A) were also observed upon transcription inhibition (Supplementary Fig. 3c, d, j, k). CTCF loss-mediated increases in loop intensities for cluster1-P CRE loops were faithfully recapitulated in G1 upon transcription inhibition (Supplementary Fig. 6g). Furthermore, the weakening of boundaries (cluster1 and 2 boundaries) as a result of CTCF withdrawal was recapitulated in triptolide-treated cells (Supplementary Fig. 4f, g). Together, our results suggest that transcription is not a significant driving force for the formation of compartments and boundaries, or CTCF-dependent chromatin remodeling after mitosis.

Transcription dependent and independent mechanisms drive gene domain formation at G1 entry. Active genes can appear as distinct squares on Hi-C contact maps, which have been interpreted to reflect gene domains caused by the process of transcription[21,22]. We examined to what extent transcription reactivation may dynamically influence genome reconfiguration by performing integrative analysis of our Hi-C and post-mitotic PolII ChIP-seq datasets. A visually appreciable gene domain at the *Rfwd2* locus can be observed as early as in ana/telophase, even before PolII reached the transcription end site (TES) (Fig. 5a, b), suggesting that gene domains may appear quickly after mitosis prior to the completion of transcription elongation. To test this possibility genome-wide, we quantified the post-mitotic recovery rates of gene domains and gene-body PolII occupancy for all active genes. We found that small genes (30–50 kb) displayed comparable rates of recovery between gene domains and PolII occupancy (Fig. 5c). However, the recovery rate of PolII was markedly reduced as gene size increased beyond 50 kb and became significantly slower compared to that of gene domains (Fig. 5c, d; Supplementary Fig. 12a–g). Of note, CTCF removal had little impact on the recovery of gene domains after mitosis (Fig. 5c, d; Supplementary Fig. 12a–g). Transcription inhibition by triptolide diminished but did not abolish gene domain formation after mitosis (Supplementary Fig. 12h), indicating that active transcription accounts partially but not entirely for gene domain formation. Notably, gene domains were precisely decorated by H3K36me3. This mark delineates active gene bodies, but unlike the process of transcription itself, is stable throughout mitosis (Fig. 5b, d; Supplementary Fig. 12a–g)[34]. This suggests that the re-establishment of gene domains that precedes the onset transcription might be facilitated by this chromatin mark, nominating H3K36me3 as a potential mitotic bookmark.
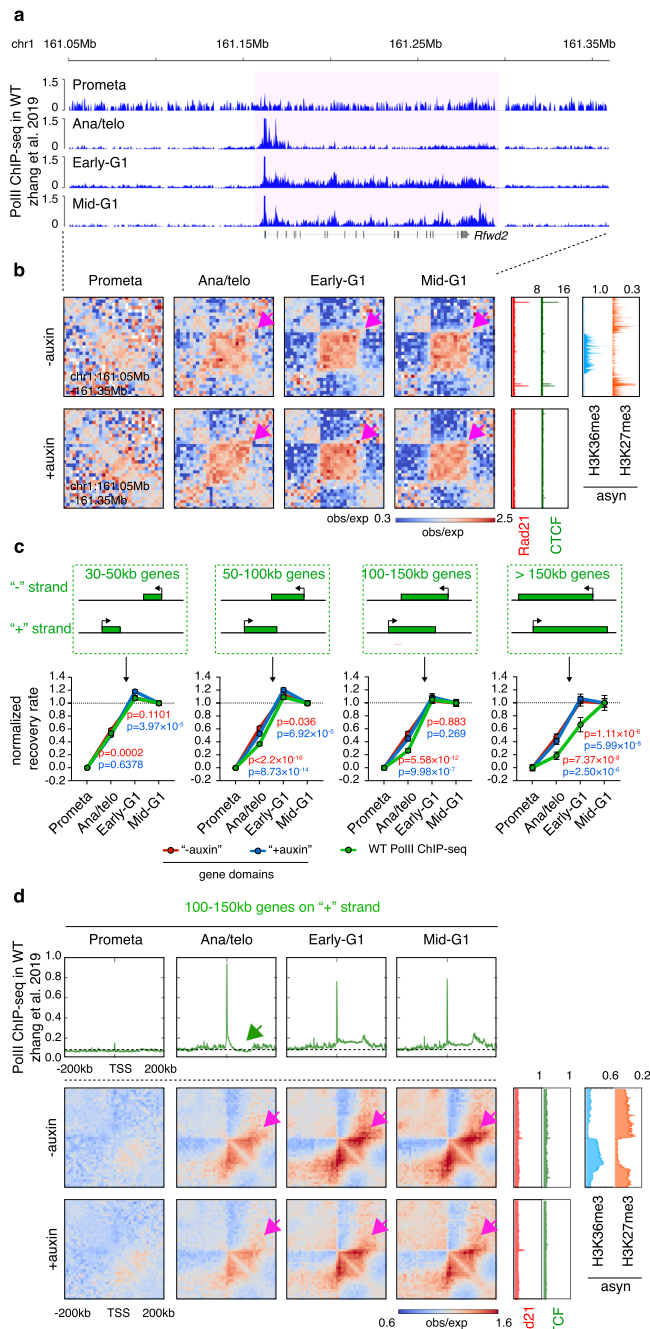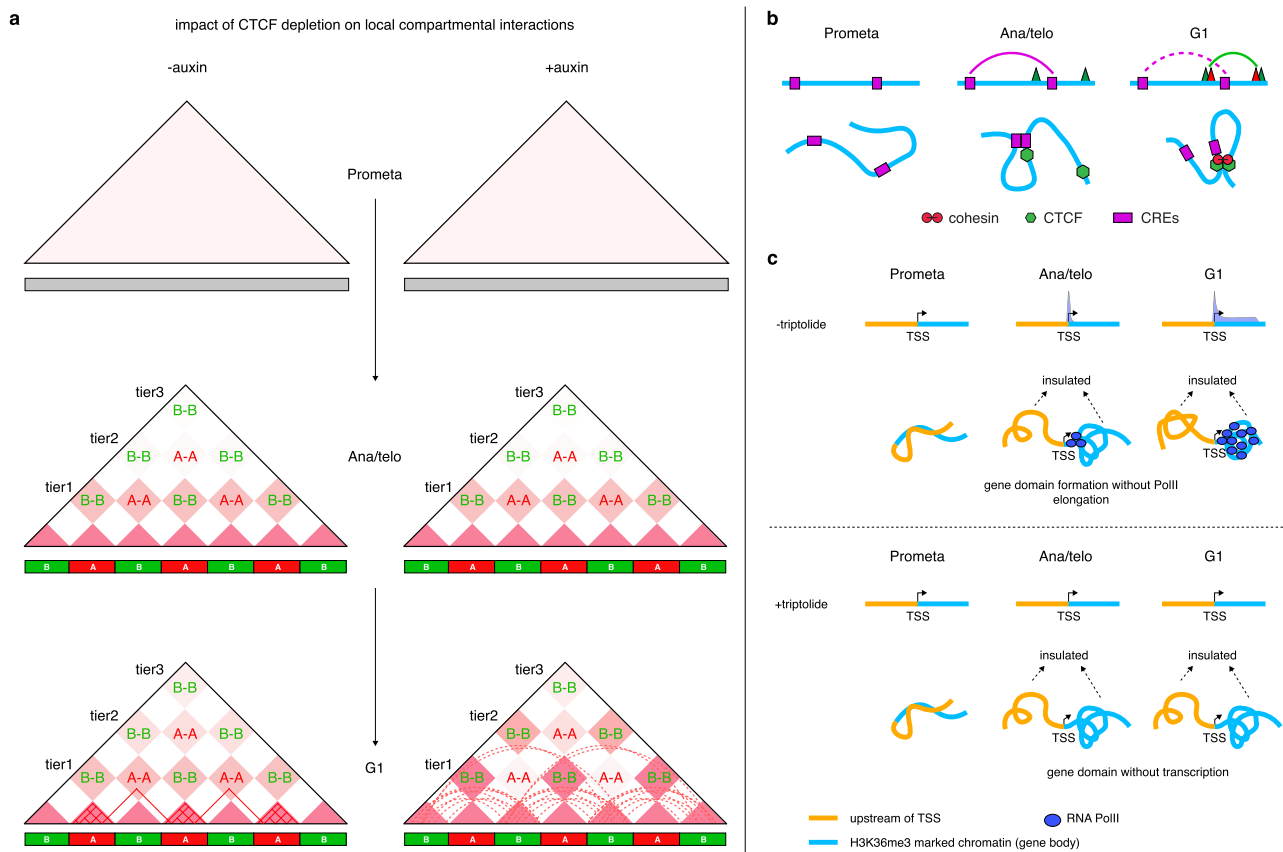
Fig. 5 Gene domains emerge prior to completion of the first round of transcription after mitosis. a PolII ChIP-seq genome browser tracks at the *Rfwd2* locus across cell cycle stages in parental cells. Note that in ana/telophase PolII is detected at the promoter region but the initial round of transcription has not been completed. **b** KR balanced Hi-C contact matrices of the same region as in (**a**) across cell cycle stages in control and auxin-treated samples. Bin size: 10 kb. Purple arrows indicate the domain of the *Rfwd2* gene in post-mitotic stages. Tracks of CTCF and Rad21 with or without auxin treatment as well as histone marks H3K36me3 and H27me3 are from asynchronously growing G1E-ER4 cells. **c** Upper panel: Schematic of genes with different sizes. Lower panel: Line graphs of recovery rates of gene domains in the control and auxin-treated samples and the recovery rate of PolII occupancy over the gene body. Genes corresponding to the size ranges in the upper panel were separately plotted. $n = 952, 846, 260,$ and 119 genes with size ranges of 30–50, 50–100, 100–150 kb, and over 150 kb, respectively. P values were calculated using a two-sided paired Wilcoxon signed-rank test. Red and blue asterisks represent comparisons between PolII and gene domains in untreated control or auxin-treated samples, respectively. Error bars denote SEM. **d** Upper panel: Meta-region pile-up plots of PolII ChIP-seq signals corresponding to the 100–150kb genes on the plus strand across cell cycle stages. Plots are centered on TSS. Lower panel: Pile-up Hi-C matrices showing the domains of the genes corresponding to the upper panel across cell cycle stages in untreated and auxin treated samples. Bin size: 10 kb. Plots are centered on TSS. Gene domains are labeled with purple arrows. Meta-region plots of CTCF and Rad21 with or without auxin treatment, as well as H3K36me3 and H3K27me3, are shown on the right.

Discussion

Examination of the earliest stages of transition from prometaphase into G1 phase affords a unique view into how chromatin is configured de novo in newborn nuclei. The AID protein degradation system enabled investigation into the role of CTCF in this process. A meaningful interpretation of the experiments in this study requires that CTCF degradation does not significantly impede cell cycle progression. This was demonstrated by (1) flow cytometry measuring DNA content, cell size, and GFP-MD levels, (2) the presence of highly comparable contact decay curves across all post-mitotic time points, and (3) relatively stable post-mitotic gene expression patterns, including widespread gene spiking. As observed previously, only prolonged depletion of CTCF (36 h) was found to delay cell cycle progression (Supplementary Fig. 1e), but this did not impact the present study.

Our data suggest that CTCF influences chromatin structure at several levels during G1 entry. First, CTCF-based structural loops constrain short-range B–B compartmental interactions while promoting local A–A compartmental interactions, revealing a previously underappreciated role for CTCF in chromatin compartmentalization. We speculate that these observations are driven by altered loop extrusion after CTCF loss. Removal of CTCF may allow cohesin to travel beyond CTCF-binding sites, thereby increasing loop sizes. Structural loops originating within A-type compartment domains may thus extend into flanking B-type compartment domains and increase their contact probability (Fig. 6a). The rising gains of B–B interactions during the progression into G1 in the absence of CTCF are consistent with the gradual loading and advancement of cohesin after mitosis. The spatial confinement in B–B interaction gains might be due to limitations of the loop extrusion process. Our findings thus reveal that loop extrusion can elicit both positive and negative effects on compartmentalization depending on the type and location of the compartment (Fig. 6a)[12,28,35]. Second, the transient nature of many post-mitotic CRE contacts might be explained by the disruptive nature of emerging nearby structural loops. Thus, considering the cell cycle dynamics of structural as well as CRE loops during G1 progression allowed for the inference that CTCF's ability to disrupt established CRE interactions, and hence function as an insulator, requires its engagement in loops (Fig. 6b). Third, we uncovered a previously underappreciated role of the genomic positioning of structural loop relative to CREs. Specifically, CRE contacts were most sensitive to disruption when the structural loop anchor was close to the CRE (small s_i-min). While the mechanism underlying this observation is unclear, it is possible that the distance sensitivity of CRE contacts to the disruptive effects of extruding structural loops might be a function of the cohesin complex reaching and being arrested at CTCF sites more frequently. Fourth, CTCF-anchored loops may facilitate interactions between CREs by providing structural support. Weaker

Fig. 6 Mechanistic models. a Schematic showing how CTCF removal can impact local but not distal interactions between the same type of compartments. Short-range B–B interactions were enhanced potentially due to increased extrusion loop size from A compartments after CTCF removal. Note, the effect was progressively observed in the G1 phase because of the gradual action of loop extrusion. Solid lines in the bottom panel represent structural loops formed and stabilized within A-type compartment domains in CTCF repleted conditions. Dotted lines represent actively extruding loops that are unleashed from A-type compartment domains into flanking B-type compartment domains due to CTCF depletion. **b** Schematic showing the rapid dissolution of established CRE loops as nearby disruptive structural loops emerge after mitosis. **c** Schematic showing that a gene domain is already partially established prior to full coverage of the gene body by PolII.

CREs appear to be more reliant on such "supportive" structural loops.

Up-regulation of genes caused by CTCF loss was associated with enhanced interactions between promoters and enhancers and was observed at the earliest measurements (1 h after mitosis), suggesting a tight temporal relationship of promoter–enhancer proximity and transcription. Additionally, we found that the down-regulated genes generally did not display measurable loss in E–P contacts, but instead appear to depend on CTCF binding at their TSS. This implies that at these genes CTCF might function as a transcriptional activator, independent of its role in chromatin looping. This observation diverges from previous reports that CTCF depletion can diminish E–P interactions and result in transcription loss[15,36–38]. However, we cannot rule out that disruption of shorter-range E–P loops, undetectable in our Hi-C experiments, might account for down-regulation gene transcription.

Prior studies reported a correlation between transcription activity and gene domains in asynchronously growing cells[21,22,39]. However, comparing the kinetics of transcription re-activation and gene domain reformation, and inhibiting transcription pharmacologically revealed that the process of transcription per se does not account for the entirety of post-mitotic gene domain formation (Fig. 6c). We speculate that additional mechanisms (e.g. H3K36me3 histone modification along gene body) may contribute to self-aggregation of genes and pre-configure genes for subsequent activation.

In summary, by leveraging the AID system and chemical transcription inhibition in the context of cell cycle dynamics, we were able to deepen our insights into the mechanisms by which CTCF and transcription reshape multiple facets of chromatin architecture from a randomly organized state in mitosis to the fully established structures in interphase.

Methods

prometaphase synchronization. To acquire post-mitotic populations, nocodazole-treated cells were pelleted at 1200 rpm for 3 min. Cells were then washed once and immediately re-suspended in a warm nocodazole-free medium containing 1 mM auxin for 60 min (early-G1) and 120 min (mid-G1), respectively. To enrich for cells at ana/telophase, G1E-ER4 CTCF-AID-mCherry cells overexpressing OsTiR-IRES-GFP-MD were first synchronized as above described and then released from nocodazole for 30 min before harvest.

Control samples underwent the exact same treatment except that no auxin was added.

Transcription inhibition. G1E-ER4 CTCF-AID-mCherry cells expressing OsTiR-IRES-GFP were arrested in prometaphase with nocodazole (200 ng/ml) as above. Triptolide (1 μM) was added to the cultures during the last hour of nocodazole exposure. For "+auxin" samples, auxin was added during the last 4 h of nocodazole treatment. Cells were released into the warm nocodazole-free medium with 1 μM triptolide and with or without 1 mM auxin for 2 h.

Cell sorting. Control and "+auxin" samples were acquired as described previously[1]. Briefly, for in situ Hi-C experiments, cells were pelleted at 1200 rpm for 3 min. Cells were then re-suspended in 1× PBS and crosslinked with 2% formaldehyde for 10 min at RT. Crosslinking was quenched with 1 M (final concentration) glycine for 5 min at RT. Cells were permeabilized by 0.1% Triton X-100 for 5 min at RT and stained with antibody against the mitosis-specific antigen pMPM2 (Millipore, 05-368, 0.2 μl/10 million cells) for 50 min at RT. Cells were then treated with APC-conjugated F(ab′)2-goat anti-mouse secondary antibody (Thermo Fisher Scientific, 17-4010-82, 2 μl/10 million cells) for 30 min at RT. Cells were pelleted and re-suspended in 1× FACS sorting buffer (1× PBS, 2% FBS, 2 mM EDTA, and 0.02% NaN₃) containing 20 ng/ml DAPI at a density of about 50–100 million cells/ml. Prometaphase cells were purified via FACS on the basis of pMPM2 signal (+) and DAPI signal (4N). To harvest populations after mitotic exit, cells were pelleted and crosslinked with 2% formaldehyde at designated time points (ana/telophase, early-G1, and mid-G1 phase). Crosslinking was halted with 1 M (final concentration) glycine, and cells were permeabilized with 0.1% Triton X-100. Finally, cells were re-suspended in 1× FACS sorting buffer and subjected to FACS sorting. Ana/telophase cells were sorted based on GFP signal (reduced) and DAPI signal (4N). Early-G1 and mid-G1 cells were sorted based on the DAPI signal (2N). In addition, mCherry positive and negative populations were gated to collect untreated control and auxin-treated samples, respectively, for all mitotic and post-mitotic time points. Sorted cells were snap-frozen and stored at −80 °C. FACS plots were generated by FlowJo software (version 10.4.0)

The exact same procedure was carried out for cells that had undergone triptolide treatment.

For PolII ChIP-seq: Cells with or without auxin treatment were harvested at 0 min (prometa), 60 min (early-G1), 120 min (mid-G1), or 240 min (late-G1) after nocodazole release. Cells were re-suspended in 1× PBS and crosslinked with 1% formaldehyde for 10 min at RT. Crosslinking was stopped with 1 M glycine followed by permeabilization with Triton X-100. All samples (±auxin at all cell cycle stages) were stained with anti-pMPM2 antibody (Millipore, 05-368, 0.2 μl/10 million cells) for 50 min at RT, followed by APC-conjugated F(ab′)2-goat anti-mouse secondary antibody (Thermo Fisher Scientific, 17-4010-82, 2 μl/10 million cells) for 30 min at RT. Cells were re-suspended in 1× FACS buffer containing DAPI and subjected to FACS sorting. Prometaphase cells were purified via FACS based on pMPM2 signal (+) and DAPI signal (4N). Early-, mid- and late-G1 cells were sorted based on DAPI signal (2N). mCherry positive and negative populations were gated to collect untreated control and auxin-treated samples for all mitotic and post-mitotic time points. Sorted cells were snap-frozen and stored at −80 °C. Note that protease inhibitor and PMSF were added to all buffers during the entire sample preparation procedure.

In-situ Hi-C. In-situ Hi-C experiments were performed as previously described[1]. Briefly, sorted cells (5 million for prometaphase and ana/telophase and 10 million for early- and mid-G1 phase) were lysed in 1 ml cold lysis buffer (10 mM Tris pH 8, 10 mM NaCl, 0.2% NP-40/Igepal) for 10 min on ice. Nuclei were pelleted at 4 °C and washed with 1.2× DpnII buffer. Nuclei were permeabilized with 0.3% SDS for 1 h at 37 °C and quenched with 1.8% Triton X-100 for 1 h at 37 °C. Chromatin was digested with 300U DpnII restriction enzyme (NEB, R0543M) in situ at 37 °C overnight with shaking. 300U DpnII restriction enzyme was added for an additional 4 h at 37 °C with shaking. Nuclei were incubated at 65 °C for 20 min to inactivate DpnII. After cool down, digested chromatin fragments were blunted with pCTP, pGTP, pTTP and Biotin-14-dATP (Thermal Fisher Scientific, 19524016) using 40U DNA Polymerase I, Large (Klenow) fragment (NEB, M0210). DNA was ligated in-situ with 4000U T4 DNA ligase (NEB, M0202M) for 4 h at 16°C followed by further incubation for 2 h at RT. Nuclei were then incubated in 10% SDS containing proteinase K (3115879 BMB) at 65°C overnight to reverse crosslinking. RNA was then digested with DNase-free RNase at 37 °C for 30 min. DNA was then extracted by phenol–chloroform extraction, precipitated, and dissolved in nuclease-free water. DNA was sonicated to 200–300 bp fragments (Epishear, Active Motif, 100% amplitude, 30 s ON and 30 s OFF, 25–30 min) and purified with AMPure XP beads (Beckman Coulter). Biotin-labeled DNA was purified by incubation with

100 μl Dynabeads MyOne Streptavidin C1 beads (Thermal Fisher Scientific, 65002) at RT for 15 min. DNA libraries were constructed using the NEBNext DNA Library Prep Master Mix Set for Illumina (NEB E6040, M0543L, E7335S). To elute DNA, streptavidin bead-bound DNA was incubated in 0.1% SDS at 98 °C for 10 min. DNA was purified with AMPure XP beads and index labeled with NEBNext multiplex oligos for six cycles on a thermal cycler, using the NEBNext Q5 Hot Start HIFI PCR master mix. Index labeled PCR products were then purified with AMPure XP beads and sequenced on an Illumina NextSeq 500 sequencer. Sequencing data was collected through NextSeq Control Software 2.2.0.

ChIP-seq. Chromatin immunoprecipitation (ChIP) was performed using anti-RNA Polymerase II antibody (Cell Signaling, 14958, 5 μg/IP) as described previously[1]. Briefly, following sorting, cells were re-suspended in 1 ml pre-cooled cell lysis buffer supplemented with protease inhibitors (PI) and PMSF for 20 min on ice. Nuclei were pelleted and re-suspended in 1 ml Nuclear Lysis Buffer (50 mM Tris pH 8, 10 mM EDTA, 1% SDS, fresh supplemented with PI and PMSF) for 10 min on ice. 0.6 ml IP dilution buffer (20 mM Tris pH 8, 2 mM EDTA, 150 mM NaCl, 1% Triton X-100, 0.01% SDS, fresh supplemented with PI and PMSF) was added followed by sonication (Epishear, Active Motif, 100% amplitude, 30 s ON and 30 s OFF) for 45 min. Samples were pelleted at 15,000 rpm for 10 min at 4 °C to remove cell debris. The supernatant was supplemented with 3.4 ml IP dilution buffer fresh supplemented with PI and PMSF, 50 μg isotope-matched IgG, and 50 μl protein A/G agarose beads (A:G = 1:1, ThermoFisher 15918014 and ThermoFisher 15920010) and rotated at 4 °C for 8 h to preclear the chromatin. 200 μl, chromatin was set aside as input chromatin. Precleared chromatin was then incubated with 35 μl A/G agarose beads (A:G = 1:1) pre-bound with anti-RNA PolII antibody (5 μg/IP) at 4 °C for overnight. Beads were washed once with IP wash buffer I (20 mM Tris pH 8, 2 mM EDTA, 50mMNaCl, 1% Triton X-100, 0.1% SDS), twice with high salt buffer (20 mM Tris pH 8, 2 mM EDTA, 500 mM NaCl, 1% Triton X-100, 0.01% SDS), once with IP wash buffer II (10 mM Tris pH 8, 1 mM EDTA, 0.25 M LiCl, 1% NP-40/Igepal, 1% sodium deoxycholate) and twice with TE buffer (10 mM Tris pH 8, 1 mM EDTA pH 8). All washing steps were performed on ice. Beads were moved to RT and eluted in 200 μl fresh-made elution buffer (100 mM NaHCO₃, 1% SDS). 12 μl of 5 M NaCl, 2 μl RNaseA (10 mg/ml) was added to IP and input samples and incubated at 65 °C for 2 h, followed by addition of 3 μl protease K (20 mg/ml) and incubated at 65 °C overnight to reverse crosslinking. Finally, IP and input samples were supplemented with 10 μl of 3 M sodium acetate (pH 5.2), and DNA was purified with a QIAquick PCR purification kit (QIAGEN 28106). ChIP-seq libraries were constructed using Illumina's TruSeq ChIP sample preparation kit (Illumina, catalog no. IP-202-1012). Libraries were size-selected using the SPRIselect beads (Beckman Coulter, catalog no. B23318) before PCR amplification. Libraries were then quantified through real-time PCR with the KAPA Library Quant Kit for Illumina (KAPA Biosystems catalog no. KK4835). Finally, libraries were pooled and sequenced on an Illumina NextSeq 500 platform using Illumina-sequencing reagents. Sequencing data was collected through NextSeq Control Software 2.2.0.

Quantification and data analysis

Hi-C data pre-processing. For each biological replicate, paired end reads were aligned to the mouse reference genome mm9 using bowtie2 (global parameters:–very-sensitive -L 30–score-min L,−0.6,−0.2–end-to-end –reorder; local parameters:–very-sensitive -L 20–score-min L,−0.6,−0.2–end-to-end–reorder) through the Hi-C Pro (2.10.0) software[41]. PCR duplicates were removed and uniquely mapped reads were paired to generate a validPair file. The output validPair file was converted into ".hic" file using the hicpro2juicebox utility. For merged samples, similar steps were taken on reads merged from each biological replicate.

A/B compartment calling and processing. Compartments were called based on the ".hic" files through eigenvector decomposition on the Pearson's correlation matrix of the observed/expected value of 100 kb binned, Knight–Ruiz (KR) balanced cis-interaction maps (Eigenvector utility of juicer_tools_1.13.02)[31]. Positive and negative EV1 values of each 100 kb bin were assigned to A- (active) and B- (inactive) compartments, respectively, based on gene density. Compartments were called on both replicate-merged samples and individual biological replicates across all conditions. Chromosome 3 was excluded from compartment analysis due to chromosomal translocation.

Saddle plotting and global compartment strength calculation. To visualize compartment strength, we generated saddle plots. Briefly, 100 kb binned Knight–Ruiz (KR) balanced *cis* observed/expected contact matrix was extracted from each ".hic" file through the DUMP utility of juicer_tools (1.13.02)[31]. For untreated control samples, the contact matrices were transformed in the same way such that each row and column of bins were reordered based on the eigenvector 1 (EV1) values associated with the mid-G1 sample, so that they fall into ascending order from top to bottom and from left to right. A similar transformation was applied onto auxin-treated samples across all cell cycle stages, based on their mid-G1 samples. Also, a similar transformation was performed on triptolide-treated G1-phase samples either with or without auxin treatment. After the transformation, bins at the top-

left corner are associated with B–B compartment interactions. Bins at the bottom-right corner are associated with A–A compartment interactions. Bins at the top-right and bottom-left corners are associated with B–A and A–B compartment interactions, respectively. The transformed contact maps from each chromosome were divided into 50 equal sections and averaged to create the genome-wide saddle plots. The compartment strength of each individual chromosome was computed as following: compartment strength = (median (top20% AA) + median (top20% BB))/(median (top20% AB) + median (top20% BA)). The compartment strength from each individual chromosome was averaged and log2 transformed as genome-wide compartment strength. The compartment strength of individual replicates and merged samples were computed independently.

Compartmentalization expansion curve R(s). The analysis of the progressive compartmentalization spreading across cell cycle stages has been described previously[1]. $R(s)$ curve was established to indicate the distance-dependent level of compartmentalization. To compute the $R(s)$ curve, 100 kb binned KR balanced *cis* observed/expected matrix was extracted from ".hic" files. For each interaction bin–bin pair separated by a given genomic distance (s), we computed the product of two EV1 values corresponding to the 2 bins. We then calculated the Spearman correlation coefficient R between EV1 products and the observed/expected values of all bin–bin pairs that are separated by s. $R(s)$ was then set to demarcate the level of compartmentalization for genomic distance s. To generate the $R(s)$ curve across different genomic distances, we computed R when s equals 100 kb, 200 kb, 300 kb……125 Mb. $R(s)$ curves of each chromosome were then averaged to generate the genome-wide $R(s)$ curve. $R(s)$ curve of individual biological replicates and merged samples for both untreated control and auxin treated samples were computed independently across all cell cycle stages. For interactions close to the diagonal of contact maps, well-compartmentalized regions, i.e. interactions between bins from the same type of compartments (A–A or B–B) tend to display high observed/expected values and positive (>0) EV1 products, whereas interactions between bins from different types of compartments (A–B or B–A) tend to exhibit low observed/expected values and negative (<0) EV1 product. Thus, R tends to be high in well-compartmentalized regions. At weakly compartmentalized regions, interactions between bins tend to be low also when distant from the diagonal regardless of whether the two bins are from the same type of compartment or not. Thus, R tends to be low in weakly compartmentalized regions.

Interactions between local B–B or A–A compartments. To quantify the interactions between closely positioned compartments, we adopted a high resolution (50 kb binned) A/B compartment profile from the mid-G1 untreated control sample as reference. To measure interactions between local B–B compartments (tier1, with one A-type compartment in between), we extracted genomic coordinates of all A-type compartments from the reference A/B compartment file. For each A-type compartment, we computed the average observed/expected values between the 250 kb region upstream of the start site and the 250 kb region downstream of the end sites across all cell cycle stages in both untreated control and auxin treated samples as well as G1-phase samples treated with triptolide. The resulting observed/expected values were denoted as interaction strengths between each closely spaced B–B compartment pair. We also computed more distally separated tier2 (with two A-type compartments in between) B–B interactions. For each given tier2 B–B compartment pair, we computed the average observed/expected values between the 250 kb region upstream of the start site of the first A-type compartment and the 250 kb region downstream of the end site of the second A-type compartment. A similar approach was taken to calculate tier3 (with three A-type compartments in between) and tier4 (with four A-type compartments in between) B–B interaction. Interactions between tier1 to 4 A–A compartment pairs were calculated using the same approach.

For distance-matched random controls, we randomly selected 500 A- or B-type compartments and shuffled their genomic coordinates for each entry using the "shuffle" function of bedtools[42]. Randomly shuffled bed files were used as input to compute the interaction strengths between 250 kb up-stream and down-stream flanking regions for each entry.

Loop calling and post-processing. Chromatin loops were called using a previously described HICCUPS method with modifications[31]. The following steps were taken to generate unique non-redundant lists of loops in untreated controls and auxin-treated samples across all cell cycle stages. (1) We used HICCUPS to call preliminary loops on the untreated control samples at each cell cycle stage using 10 kb binned matrices with the Juicer_tool_1.13.02. The inner and outer diameters of the donut filter were set to be 4 bins and 16 bins, respectively, and an FDR of 0.2 was adopted. (2) We repeated the above step on 10 kb binned auxin-treated samples across all cell cycle stages with the exact same parameter set. (3) Loop calls from (1) and (2) were merged to generate a non-redundant loop list across all cell cycle stages and auxin treatment conditions for 10 kb matrices. (4) False-positive calls introduced by exceptionally high outlier pixels were usually present in all samples irrespective of cell cycle stage and auxin treatment condition. To eliminate these artifacts, we removed pixels that were called in more than 6 of the 8 total samples. (5) In certain scenarios, pixels identified at different cell cycle stages or different auxin treatment conditions tended to cluster together. These clusters of pixels could actually be considered as one loop instead of many. Therefore, we

implemented a method to merge these clustered pixels. To begin with, for a given loop in the non-redundant list from step (4), we recorded a value q_{min} which represents the lowest q value across all cell cycle stages in both untreated controls and auxin treated samples. We then ordered the loops in ascending order based on their q_{min}. In this way, pixels at the top of the list were the most confident calls. We then focused on the top pixel and scanned through the rest of the list to identify pixels that were within a 20 kb radius of the top pixel. If no additional pixel were found nearby, we then considered the top pixel as a loop by itself without pixel clustering. If pixels existed that fulfilled the above requirement, we then consider these pixels together with the top pixel as a loop cluster. Pixels within the loop cluster were removed from the non-redundant loop list. We then recalculated the centroid of the loop cluster and computed the distance s between the centroid and the far cluster edge. We then started the second round of pixel merging by scanning the rest of the list to identify pixels that are within a radius of 20 kb + s from the centroid. Pixels within the loop cluster after the second collapsing step were removed from the non-redundant loop list. Next, we focused on the top pixel of the remaining list and repeated the above clustering steps, until no further pixels remained in the pixel list. After pixel merging, we generated a list of loop clusters that contained 1 or more pixels. For each loop cluster, we defined a cluster summit which was represented by the pixel with the lowest q_{min}. If a loop cluster only contained pixels called in the untreated controls but not in the auxin treated samples, we defined it as a "lost" loop. Conversely, if a loop cluster did not contain pixels from the control samples, it was defined as a "gained" loop. The remaining loop clusters were categorized into "retained", indicating that these loops were detected in both "-auxin" and "+auxin" samples. (6) We next performed step (1) through (5) on 25 kb binned matrices with an inner donut filter diameter of 1 bin and outer donut filter diameter of 6 bins. FDR of 0.01 was adopted. (7) Loops called on 25 kb binned matrices were then merged with those called on 10 kb matrices. If a 25 kb loop cluster overlaps with a 10 kb loop cluster, the 25 kb loop was dropped.

We noticed that some visually solid loop-like pixels were dropped by the HICCUPS due to a lack of surrounding significant pixels. To recover these potentially false-negative calls, we took advantage of our biological replicates. Specifically, we continued to complement our non-redundant loop calling list with the following steps. (8) For untreated control samples, we extracted all raw significant pixels from HICCUPS before clustering across cell cycle stages and combined them. (9) We then computed the donut FDR of the above pixels in all biological replicates across cell cycle stages in the untreated control samples through juicer_tools_1.13.02. (10) To determine if a pixel represented a loop, we implemented the below filters: (1) For ana/telophase or early-G1 phase, we required that a pixel must display an FDR < 0.2 in both replicate-merged and individual biological replicates. (2) For ana/telophase or early-G1 phase, a pixel was required to show an observed/donut-expected value of over 1.5 in replicate-merged and individual biological replicates. (3) For ana/telophase or early-G1 phase, a pixel had to exhibit an observed value of >10 in replicate-merged and individual biological replicates. (4) For the mid-G1 phase, the above three criteria had to be satisfied in at least 3 of the following 4: replicate-merged, biological replicate 1, 2, and 3. A pixel had to fulfill all the above filters to be viewed as valid in a given cell cycle stage, and it had to be valid in at least one post-mitotic cell cycle stage to be considered a valid loop for the untreated controls. (11) We then repeated steps (8) through (10) to get a list of valid pixels in auxin-treated samples. (12) We combined the valid pixels from untreated controls and auxin-treated samples and filter out pixels with the highest 5% observed/donut-expected values in prometaphase in untreated control samples as well as pixels with a distance of over 2 Mb. (13) We further removed pixels that were overlapping or next to the loops identified in step (9). In this manner, we obtained valid loops that had been previously missed by HICCUPS. (14) Finally, we performed step (5) on the remaining pixels at (13) to merge valid pixels that were clustered together. In total, we ended up with 16,370 non-redundant loops across all samples.

Loop categorization based on CTCF/cohesin/CRE. We categorized loops into different classes based on whether ChIP-seq peaks of CTCF and cohesin and annotations of promoters or enhancers were present at their anchors. For a peak to intersect with a loop anchor, it had to have at least 1 bp overlap with a 30 kb region centered on the midpoint of the loop anchor summit. We employed the CTCF/cohesin co-occupied peak list and the peaks of H3K27ac (CRE) from our previous study[1]. Our analysis focused on the following possibilities: (1) Two loop anchors harbor CTCF/cohesin co-occupied sites with neither harboring CREs. (2) Both loop anchors harbor CTCF/cohesin co-occupied sites with one anchor also harboring a CRE. (3) Both loop anchors each harbor a CTCF/cohesin co-occupied site and a CRE. (4) None of the loop anchors harbor CTCF/cohesin co-occupied sites but both contain CREs. (5) One loop anchor a harbors CTCF/cohesin co-occupied site and two anchors harbor CREs. Groups (1) and (2) loops were defined as "structural loops", loops from group (3) as "dual-function loops", and loops from the groups (4) and (5) as "CRE loops".

K-means clustering of loops. To measure the change of post-mitotic loop formation as well as the impact of CTCF depletion on loop formation, we defined a metric to measure the strength of each loop. For a given loop, we considered its summit pixel as well as eight surrounding pixels and computed their observed/donut-expected values across cell cycle stages in both untreated controls and auxin-treated samples.

For a specific cell cycle stage and auxin treatment condition, the loop strength was recorded as the average of the observed/donut-expected values from the 9 pixels. To dissect the wildtype CRE loop reformation patterns after mitosis, we focused on the 3232 CRE loops that were detected in untreated control samples. We then computed the z-scores of loop strength across all cell cycle stages in untreated control samples and performed k-means clustering using the three post-mitotic time points. We were able to recover four-loop clusters with distinct reformation kinetics. To assess the effect of CTCF depletion on cluster1 transient CRE loops, we attempted to further sub-categorize them using the loop strength from both untreated controls and auxin-treated samples. We computed the z-scores of loop strength across all cell cycle stages in both untreated as well as auxin-treated samples. We then performed k-means clustering using the three post-mitotic time points from both untreated controls and auxin-treated samples.

Measuring the interplay between structural and CRE loops. To quantify the degree to which CRE loops are disrupted or supported by structural loops, we performed the following enrichment analysis. We first focused on structural loop interpolation. For a given cluster of CRE loops (e.g. cluster1-R), we defined two scenarios based on whether or not they are interpolated by structural loops. The two scenarios were: (1) neither of the two CRE anchors were covered by a structural loop (not interpolated) and (2) either one or both of the CRE anchors were covered by structural loops (interpolated). For each scenario, we constructed a 2 × 2 contingency table based on in which scenario a given cluster of CRE loops fell, and whether or not the rest of the CRE loops fell into that scenario. Odds ratios and P-values were computed with the Fisher's exact test in R. Lastly, we defined for each cluster of CRE loops whether or not they were supported by structural loops. A similar approach was taken to calculate the odds ratio and P-values for each CRE loop cluster at each scenario.

Domain calling. Domains were independently identified in untreated controls and auxin-treated samples across all cell cycle stages, using the rGMAP algorithm (1.4) (https://github.com/tanlabcode/rGMAP)[30]. To call domains in untreated control samples, we extracted 10 kb binned KR balanced *cis* contact matrices from replicate-merged ".hic" files of each cell cycle stage, using the DUMP utility of juicer_tool (1.13.02)[31]. The contact matrices were used as input to feed in rGMAP for domain calling. For each cell cycle stage, we started by performing the following domain sweep: (1) Maximally three levels of domains were allowed (dom_order = 3). (2) Maximal contact distances of 2 Mb were allowed (maxDistInBin = 200). This step generated a basic list of domains. To capture sub-domain-like structures, we performed an additional domain sweep, which allowed a maximal contact distance of 500 kb (maxDistInBin = 50). Additional sub-domains were then added to the basal list to create a preliminary list of domains for each cell cycle stage. A similar approach was carried out to generate the preliminary domain list in the auxin-treated samples. As a reference, we also performed rGMAP to call domains in late-G1 phase parental cell samples, using the same criteria as above.

Domain and boundary detection across cell cycle stages. For untreated control samples, domains called at each cell cycle stage were merged to create a total domain list. Domains from the late-G1 wildtype samples were added to the above list to serve as a reference. To ensure the validity of our domain calls, we established the following filters: (1) Domains called at prometaphase had to overlap with at least three domains identified in the subsequent four cell cycle stages (ana/telo, early-G1, mid-G1, and late-G1 from wildtype sample) to be considered valid. (2) Domains called at ana/telo, early-G1 or mid-G1 had to overlap with at least one domain identified in subsequent cell cycle stages to be considered valid. To claim that a domain detected in prometaphase is also present at a later time point, we require that at least one domain exists in the later time point, whose upstream and downstream boundaries are within ±8 bins of those of the original domain. We performed this step across all subsequent cell cycle stages to identify all potentially "overlapping" domains. If at least three subsequently identified domains overlap with our query domain, we then separately average the up- and down-stream boundaries of all "overlapping" domains to replace the boundaries of the original prometaphase domain. A similar approach was carried out in the ana/telo, early-G1, and mid-G1 phases. These steps produced a list of high confidence domains that were detected across different cell cycle stages in the untreated control samples.

Next, we implemented a merging step to adjust boundary locations so that domains across cell cycle stages with highly similar boundaries would share a single consistent boundary (It is noteworthy that possibilities still remain that two highly similar boundaries represent true biological differences instead of technical differences.). As a start, we generated an overall non-redundant boundary list from all domains. Boundaries were then sorted based on their genomic coordinates from 5′ to 3′. Starting from the first boundary, we swept throughout the rest of the boundaries on the same chromosome and removed boundaries that are less than 80 kb away from the first boundary. We merged these boundaries into one and applied the mean of their genomic coordinates as the genomic coordinate of the final merged boundary. These boundaries were then removed from the overall boundary list. We then performed this step iteratively on the remaining boundaries until all boundaries were processed. The final averaged boundary coordinates were then reassigned to corresponding domains. For a boundary shared by multiple

domains, the time point of emergence of this boundary is determined by the earliest associated domain.

The same approach was carried out to process domains and boundaries identified in the auxin-treated samples across all cell cycle stages.

Insulation score profiling. Insulation scores were computed as previously described[43]. Briefly, we implemented a 12 bin × 12 bin window, which slides along the diagonals of the 10 kb binned KR balanced contact matrices. The sliding window was set to be one bin away from the diagonal. Genomic regions with low read counts (<12 counts) were discarded from the analysis. Windows interrupted by the starts or ends of chromosomes were also discarded. For each 10 kb bin, the sum of reading counts of each window was then normalized to the chromosomal average and log_2 transformed. A pseudo-read count was added to the chromosomal mean as well as each window before log transformation.

We noticed that in a few cases domain boundaries were shifted by several bins from the local minima of insulation, and thus did not accurately reflect the "real" boundary position. To solve this issue, we fine-tuned boundary positions such that boundaries were adjusted to the local minima of insulation. This adjustment was performed on untreated control samples. For each given boundary, we defined a wiggle room by sectioning a −6 bin to +6 bin genomic region that centered around the boundary. We then recorded the mid-G1 phase insulation scores of each bin within the wiggle room. The bin with the lowest insulation score for the mid-G1 phase was defined as the final position of the boundary, representing local minima of insulation scores. The adjusted boundary locations were then re-assigned to their corresponding domains. After boundary adjustment, domains smaller than 100 kb were filtered out to eliminate spurious domains. In some outlier cases, boundaries after adjustment ended up being extremely close to each other (within 20 kb). We, therefore, implemented a final step to merge boundaries using the same approach as described above.

Integration of domains and boundaries from untreated control and auxin treated samples. After boundary position adjustment, we obtained an intermediate list of domains and boundaries that were detected at each cell cycle stage for the untreated control samples. We then added domains and boundaries from "+auxin" samples to this list to create a final complete domain list before quality check. We carried out the flowing merging steps: For a given domain in the auxin treated samples, if both of its boundaries were <80 kb away from the up- and downstream boundaries of a "−auxin" domain, we then considered these two domains as "overlapping" and recorded the boundary coordinates of the "−auxin" domain in the final list. If the upstream (but NOT downstream) boundary of the "+auxin" domain was <80 kb away from any boundaries in the "−auxin" list, we then considered this "+auxin" domain a new domain and recorded the "−auxin" boundary coordinate as the upstream boundary for this "+auxin" domain. Similarly, if the downstream (but NOT upstream) boundary of the "+auxin" domain was <80 kb away from any boundaries in the "−auxin" list, we considered this "+auxin" domain as a new domain and recorded the "−auxin" boundary coordinate as the downstream boundary for this "+auxin" domain. Finally, if both upstream and downstream boundaries of the "+auxin" domain were more than 80 kb away from boundaries in the "−auxin" list, we considered this "+auxin" domain as a new domain and recorded its own boundary coordinates in the final list of domains.

Domain quality check and aggregated domain analysis (ADA). We noticed that in some rare cases, domains spanning large low-mappable regions were also called by the algorithm. To filter out low confidence domains, we implemented an aggregated domain analysis which measures the ratio between interactions just inside the domain and interactions just outside the domain. We computed ADA scores on the 10 kb-binned KR balanced observed/expected contact matrices as previously reported with modifications[1,12]. For each domain, the start and end coordinates were recorded as $i × 10,000$ and $j × 10,000$, respectively. Therefore, we could use (i, j) to mark the position of the corner pixel. We then marked our four horizontal stripes and four vertical stripes that were just inside the domains. The positions of the horizontal inner stripes are: $[i + 1, j−8: j−4]$, $[i + 2, j−7:j−3]$, $[i + 3, j−6:j−2]$ and $[i + 4, j−5:j−1]$, respectively. The positions of vertical inner stripes are: $[i + 1:i + 5, j−4]$, $[i + 2:i + 6, j−3]$, $[i + 3:i + 7, j−2]$ and $[i + 4:i + 8, j−1]$, respectively. We then also marked out an additional four horizontal and four vertical stripes that were outside of the domain. The positions of horizontal outer stripes were $[i−8, j−17:j−13]$, $[i−7, j−16:j−12]$, $[i−6, j−15:j−11]$ and $[i−5, j−14:j −13]$, respectively. The positions of vertical outer stripes were $[i + 10:i + 14, j + 5]$, $[i + 11:i + 15, j + 6]$, $[i + 12:i + 16, j + 7]$ and $[i + 13:i + 17, j + 8]$, respectively. The inner stripes and outer stripes had the same genomic separations. We computed the sum of observed/expected values for pixels within inner stripes and then divided this value by the sum of observed/expected values for pixels within outer stripes. The final result was log_2 transformed and recorded as the ADA score of the domain. Note that domains smaller than 150 kb was filtered to minimize the possibility of outer stripes stretching into another domain.

To eliminate domains covering low-mappable regions, we placed the following filters: (1) For a given domain, we examined all pixels in the inner and outer stripe regions. If any pixel displayed an observed/expected value of over 30, we dropped this domain from further analysis. This step was to filter out high outlier pixels that

are usually associated with low-mappable regions. (2) For a given domain, if either or both of the two outer stripe regions contained less than 5 non-zero pixels, the domain was dropped from further analysis. (3) For a given domain, if either or both of the two inner stripe regions contain less than 10 non-zero pixels, the domain was dropped from further analysis. To further ensure the validity of our domain calls, we also implemented a dynamic filter. This filter was established based on the rationale that true domains would gradually become stronger after mitotic exit and thus their ADA scores would be higher in post-mitotic time points compared to prometaphase. Specifically, we require that for a domain to be valid, at least 1 of the 6 post-mitotic samples ("ana/telo −auxin", "ana/telo + auxin", "early-G1 −auxin", "early-G1 + auxin", "mid-G1 −auxin", "mid-G1 + auxin") had to show at least a 1.25-fold ADA score enrichment compared to both prometaphase samples ("prometa −auxin", "prometa + auxin").

Dynamic clustering of boundaries. We quantified boundary strength as follows: We selected a −120 to +120 kb region centered on a boundary of interest and searched for the highest insulation score within this 240 kb region. This maxi-IS value was then unlogged and subtracted by the insulation score (unlogged) at the boundary itself. The resulting ΔIS was then denoted as the strength of the target boundary.

To examine the reformation dynamics of boundaries and measure the effect of CTCF depletion on the dynamic boundary formation after mitosis, we performed *k*-means clustering on the ΔIS of all boundaries across all 8 samples (4 cell cycle stages in both untreated control and auxin-treated samples). Specifically, for each boundary, we computed the *z*-scores of their ΔIS across all 8 samples. We then performed *k*-means clustering on the *z*-scores of 6 post-mitotic samples (ana/telo, early-G1, and mid-G1 × 2 treatment conditions). We found that when we chose *k* = 5 clusters, we were able to recover the most biologically interpretable clusters. Note, cluster5 was mostly spurious boundaries and thus was excluded from the analysis.

PCA based interrogation of chromatin state transition at boundaries. To assess histone modification features associated with different clusters of boundaries, we adopted a PCA-based approach as previously described[44]. For each target boundary, we selected a −50 to +50 kb genomic region and sectioned it into 10 bins (10 kb/bin). To assess chromatin state transition, we adopted the two histone marks H3K36me3 and H3K27me3, the former and latter representing transcriptionally active and inactive chromatin, respectively. We then calculated the mean G1E-ER4 ChIP-seq signals of these two marks (from asynchronously growing cells) in each 10 kb bin, using the UCSC toolkit (BigWigAverageOverBed). The ChIP-seq intensities of these marks were organized into two matrices such that each row represents the 100 kb region around a boundary, and each of the 10 columns represents a 10 kb bin. The columns were ordered based on their genomic positions from upstream to downstream. Each column was then normalized to the column sum such that the ChIP-seq intensity values from each column add up to 1. After normalization, the two matrices of H3K27me3 and H3K36me3 were stitched horizontally yielding a final matrix with 20 columns. We then applied principal component analysis (PCA) on the final matrix using the R function prcomp (3.6.1). We noticed that PC1 was able to accurately describe the transition of chromatin states in a way that boundaries with either highest or lowest PC1 projection values were typically at chromatin transition points (5′ active → 3′ inactive or 5′ inactive → 3″ active), whereas boundaries with median level PC1 projection values were not.

PolII ChIP-seq processing and peak calling. Reads were aligned against the mm9 reference genome using Bowtie2 (v2.2.9) with default parameters and soft clipping allowed ("–local")[45]. Alignments with MAPQ score lower than 10 and PCR duplicates were removed using SAMtools (v0.1.19)[46]. Reads aligned to mitochondria, random contigs, and ENCODE blacklisted regions were also removed for downstream analysis using BEDTools (v2.27.1)[42]. Peaks were called using MACS2 (v2.1.0) with default parameters and a 0.01 *q*-value cutoff. Fragment pileup and local lambda track files in bedGraph format were created during MACS2 peak calling and normalized to one million reads per library ("callpeak–bdg–SPMR")[47]. The latter was the track was subtracted from the former using MACS2 ("bdgcmp -m subtract"), negative values were reassigned as zeros and converted to bigwig format for visualization using the UCSC Toolkit ("bedGraphToBigWig"). Finally, a non-overlapping union set of peaks was created by merging peaks in all replicates using BEDtools such that all peaks that overlap by at least 1 bp were merged. Tracks of PolII ChIP-seq was generated using the python package pygenometracks (3.1).

Identification of active genes. Active genes were called based on the overall PolII ChIP-seq peak list with the following filters: (1) The TSS of a gene had to overlap with at least 1 positive PolII ChIP-seq peak. (2) The length of a gene had to be over 1 kb to ensure that enough reads were obtained over the gene body (+500 bp from TSS to TES) and discernible from the reads at the TSS. (3) We further filtered out the genes with the lowest (10%) H3K27ac signal or ATAC signal at the promoter regions (−250 to 250 bp of TSS, data from asynchronous G1E-ER4 cells)[48,49]. A PolII ChIP-seq peak at the TSS does not necessarily mean that the corresponding gene is active. In certain cases, inactive genes positioned closely downstream of the 3′ UTR of active genes could also display positive PolII signals at their TSS,

potentially leading to false assignment as an active gene. Therefore, we filtered out genes with low H3K27ac or ATAC signals to ensure that the genes were within "open" chromatin and more likely to be active. (4) The PolII ChIP-seq signal (+500 bp from TSS to TES) of at least 1 of the six post-mitotic samples ("early-G1 −auxin", "early-G1 + auxin", "mid-G1 −auxin", "mid-G1 + auxin", "late-G1 −auxin", "late-G1 + auxin") had to be ≥1.5 fold that of the two prometaphase ("−auxin" and "+auxin") samples.

PCA based interrogation of the post-mitotic gene activation pattern. PCA was performed separately on PolII ChIP-seq signals from control and auxin-treated samples. For untreated samples, we computed the replicate-merged PolII ChIP-seq signals (+500 bp from TSS to TES) of each active gene across all cell cycle stages, using the UCSC toolkit (BigWigAverageOverBed). The PolII signals from each cell cycle stage were then normalized such that they sum up to 1. PCA was performed on the last three cell cycle stages using the R package (prcomp). As described above, the PC1 values of each gene describe the "spikiness" of their post-mitotic reactivation pattern. We set the direction of PC1 projection values such that genes with high (positive) PC1 values were the most "spiky" after mitosis, whereas genes with low (negative) PC1 values displayed a gradual increase of PolII ChIP-seq signal after mitosis. The same procedure was performed on auxin-treated samples. The PC1 values of each gene from control and auxin-treated samples were highly correlated, suggesting that the post-mitotic transcriptional spiking was maintained after CTCF depletion.

Differential gene expression analysis. Gene expression levels were assessed by the number of PolII ChIP-seq read counts over the gene body (+500 from TSS to TES). To measure differential gene expression after CTCF depletion during mitotic exit, we first extracted raw PolII read counts over gene bodies from the bam files of each individual biological replicate using the "multicov" function of bedtools. This step was performed on the three post-mitotic cell cycle stages in both "−auxin" and "+auxin" samples. DESeq2 (1.24.0) was adopted to perform differential expression analysis between "−auxin" and "+auxin" samples for each post-mitotic cell cycle stage independently. Raw read PolII ChIP-seq read counts were used as input for DESeq2 with default parameters[50]. *P.adj* cutoff of 0.05 and fold change cutoff of 1.25 were adopted to call differentially expressed genes for each post-mitotic cell cycle stage. A gene was considered differentially expressed if it was significantly different in at least one post-mitotic time point. In total, we identified 426 differentially expressed genes during mitotic exit after CTCF depletion. To determine whether these genes were up- or down-regulated over time, we performed *k*-means clustering using the log$_2$FC output of DESeq2 across the early-, mid- and late-G1 phase. Finally, we identified 223 genes that were down-regulated after CTCF depletion, and 203 genes that were up-regulated after CTCF depletion during the mitosis to G1-phase transition.

ABC model to predict functional enhancers of active genes. To predict enhancers of active genes and establish E–P and P–P connections, we adopted a recently proposed ABC model (https://github.com/broadinstitute/ABC-Enhancer-Gene-Prediction)[32]. To simulate enhancer activity, we used H3K27ac ChIP-seq and ATAC-seq signals from asynchronously growing G1E-ER4 cells. These datasets were used in combination with six replicate-merged Hi-C datasets in this study (ana/telophase "−auxin", ana/telophase "+auxin", early-G1 "−auxin", early-G1 "+auxin", mid-G1 "−auxin", mid-G1 "+auxin") to predict enhancers in each of the three post-mitotic cell cycle stages with or without auxin treatment independently. We called E–P and P–P connections in each sample when the ABC score threshold equals to 0.01, 0.02, 0.03, 0.04, and 0.05, respectively. Higher the ABC score thresholds resulted in fewer but higher confidence connections. Note that we identified on average ~1.86 (fewer than the recommended 3) enhancers per gene when the ABC score threshold was set to 0.04, suggesting that 0.04 is a relatively stringent threshold. For a given ABC threshold (e.g. 0.01), we combined the predicted connections in each sample to generate an overall non-redundant list of high confidence E–P and P–P pairs. Each E–P or P–P pair was then assigned to genes with different responses (up-reg, down-reg, and non-reg) to CTCF depletion.

Differential interaction analysis between E–P contacts. Differential interaction analysis was carried out on E–P contacts called by ABC modeling with ABC score threshold set to 0.04. We adapted the E–P interaction strength values from ABC modeling and used them as the input for LIMMA (3.40.6) R package. Since the trend of CRE contact changing was overwhelmingly consistent between early and mid-G1 phase samples, we treated the samples from these two-time points as equal biological replicates. Thus, we had five biological replicates (two from early-G1 and three from mid-G1) for control and auxin-treated samples. LIMMA was used to determine differentially interacting E–P contacts. *P* values were calculated with the Bayes function within LIMMA and adjusted with the Benjamini–Hochberg method. An FDR of 0.1 was used to call significantly strengthened or reduced E–P contacts.

Aggregated plots for loops, domains, and compartments. To generated aggregated plots, we first cool files from ".hic" files using the python package hic2cool (0.8.0). Aggregated plots were then generated through the python package Coolpup (0.9.2) and Plotpup (0.9.2) using cool files as input[51]. For unscaled aggregated peak

analysis (APA), loops smaller than 100 kb were removed from the plots to avoid influence from pixels close to the diagonal. For unscaled aggregated plots of compartment transition points, compartments smaller than 300 kb were removed from the plots, again to minimize the influence from pixels near the diagonal.

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

Data availability

References

1. Zhang, H. et al. Chromatin structure dynamics during the mitosis-to-G1 phase transition. *Nature* **576**, 158–162 (2019).
2. Abramo, K. et al. A chromosome folding intermediate at the condensin-to-cohesin transition during telophase. *Nat. Cell Biol.* **21**, 1393–1402 (2019).
3. Gibcus, J. H. et al. A pathway for mitotic chromosome formation. *Science* **359**, eaao6135 (2018).
4. Naumova, N. et al. Organization of the mitotic chromosome. *Science* **342**, 948–953 (2013).
5. Pelham-Webb, B. et al. H3K27ac bookmarking promotes rapid post-mitotic activation of the pluripotent stem cell program without impacting 3D chromatin reorganization. *Mol. Cell* **81**, 1732–1748 e8 (2021).
6. Hsiung, C. C. et al. A hyperactive transcriptional state marks genome reactivation at the mitosis-G1 transition. *Genes Dev.* **30**, 1423–1439 (2016).
7. Palozola, K. C. et al. Mitotic transcription and waves of gene reactivation during mitotic exit. *Science* **358**, 119–122 (2017).
8. Dixon, J. R. et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376–380 (2012).
9. Fudenberg, G. et al. Formation of chromosomal domains by loop extrusion. *Cell Rep.* **15**, 2038–2049 (2016).
10. Sanborn, A. L. et al. Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. *Proc. Natl Acad. Sci. USA* **112**, E6456–E6465 (2015).
11. Nora, E. P. et al. Targeted degradation of CTCF decouples local insulation of chromosome domains from genomic compartmentalization. *Cell* **169**, 930–944 e22 (2017).
12. Rao, S. S. P. et al. Cohesin loss eliminates all loop domains. *Cell* **171**, 305–320 e24 (2017).
13. 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).
14. Stik, G. et al. CTCF is dispensable for immune cell transdifferentiation but facilitates an acute inflammatory response. *Nat. Genet.* **52**, 655–661 (2020).
15. Hyle, J. et al. Acute depletion of CTCF directly affects MYC regulation through loss of enhancer-promoter looping. *Nucleic Acids Res.* **47**, 6699–6713 (2019).
16. Owens, N. et al. CTCF confers local nucleosome resiliency after DNA replication and during mitosis. *Elife* **8**, e47898 (2019).
17. Oomen, M. E., Hansen, A. S., Liu, Y., Darzacq, X. & Dekker, J. CTCF sites display cell cycle-dependent dynamics in factor binding and nucleosome positioning. *Genome Res.* **29**, 236–249 (2019).
18. Ginno, P. A., Burger, L., Seebacher, J., Iesmantavicius, V. & Schubeler, D. Cell cycle-resolved chromatin proteomics reveals the extent of mitotic preservation of the genomic regulatory landscape. *Nat. Commun.* **9**, 4048 (2018).
19. Cai, Y. et al. Experimental and computational framework for a dynamic protein atlas of human cell division. *Nature* **561**, 411–415 (2018).
20. Hug, C. B., Grimaldi, A. G., Kruse, K. & Vaquerizas, J. M. Chromatin architecture emerges during zygotic genome activation independent of transcription. *Cell* **169**, 216–228 e19 (2017).
21. Rowley, M. J. et al. Evolutionarily conserved principles predict 3D chromatin organization. *Mol. Cell* **67**, 837–852 e7 (2017).
22. Hsieh, T. S. et al. Resolving the 3D landscape of transcription-linked mammalian chromatin folding. *Mol. Cell* **78**, 539–553 e8 (2020).
23. 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).
24. Zhang, D. et al. Alteration of genome folding via contact domain boundary insertion. *Nat Genet* **52**, 1076–1087 (2020).
25. You, Q. et al. Direct DNA crosslinking with CAP-C uncovers transcription-dependent chromatin organization at high resolution. *Nat Biotechnol* **39**, 225–235 (2021).
26. Morawska, M. & Ulrich, H. D. An expanded tool kit for the auxin-inducible degron system in budding yeast. *Yeast* **30**, 341–351 (2013).
27. Campbell, A. E., Hsiung, C. C. & Blobel, G. A. Comparative analysis of mitosis-specific antibodies for bulk purification of mitotic populations by fluorescence-activated cell sorting. *Biotechniques* **56**, 90–94 (2014).
28. Schwarzer, W. et al. Two independent modes of chromatin organization revealed by cohesin removal. *Nature* **551**, 51–56 (2017).
29. Hansen, A. S., Pustova, I., Cattoglio, C., Tjian, R. & Darzacq, X. CTCF and cohesin regulate chromatin loop stability with distinct dynamics. *Elife* **6**, e25776 (2017).
30. Yu, W., He, B. & Tan, K. Identifying topologically associating domains and subdomains by Gaussian Mixture model and Proportion test. *Nat. Commun.* **8**, 535 (2017).
31. Durand, N. C. et al. Juicer provides a one-click system for analyzing loop-resolution Hi-C experiments. *Cell Syst.* **3**, 95–98 (2016).
32. Fulco, C. P. et al. Activity-by-contact model of enhancer-promoter regulation from thousands of CRISPR perturbations. *Nat. Genet.* **51**, 1664–1669 (2019).
33. Bensaude, O. Inhibiting eukaryotic transcription: Which compound to choose? How to evaluate its activity? *Transcription* **2**, 103–108 (2011).
34. Chen, F., Gao, X. & Shilatifard, A. Stably paused genes revealed through inhibition of transcription initiation by the TFIIH inhibitor triptolide. *Genes Dev.* **29**, 39–47 (2015).
35. Rhodes, J. D. P. et al. Cohesin disrupts polycomb-dependent chromosome interactions in embryonic stem cells. *Cell Rep.* **30**, 820–835 e10 (2020).
36. Kubo, N. et al. Promoter-proximal CTCF binding promotes distal enhancer-dependent gene activation. *Nat Struct Mol Biol* **28**, 152–161 (2021).
37. Lee, J., Krivega, I., Dale, R. K. & Dean, A. The LDB1 complex co-opts CTCF for erythroid lineage-specific long-range enhancer interactions. *Cell Rep.* **19**, 2490–2502 (2017).
38. Thiecke, M. J. et al. Cohesin-dependent and -independent mechanisms mediate chromosomal contacts between promoters and enhancers. *Cell Rep.* **32**, 107929 (2020).
39. Zhang, D. et al. Alteration of genome folding via contact domain boundary insertion. *Nat. Genet.* **52**, 1076–1087 (2020).
40. Weiss, M. J., Yu, C. & Orkin, S. H. Erythroid-cell-specific properties of transcription factor GATA-1 revealed by phenotypic rescue of a gene-targeted cell line. *Mol. Cell Biol.* **17**, 1642–1651 (1997).
41. Servant, N. et al. HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. *Genome Biol.* **16**, 259 (2015).
42. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
43. Crane, E. et al. Condensin-driven remodelling of X chromosome topology during dosage compensation. *Nature* **523**, 240–244 (2015).
44. Behera, V. et al. Interrogating histone acetylation and BRD4 as mitotic bookmarks of transcription. *Cell Rep.* **27**, 400–415 e5 (2019).
45. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
46. Li, H. et al. The sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
47. Zhang, Y. et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
48. Dogan, N. et al. Occupancy by key transcription factors is a more accurate predictor of enhancer activity than histone modifications or chromatin accessibility. *Epigenet. Chromatin* **8**, 16 (2015).
49. Xiang, G. et al. An integrative view of the regulatory and transcriptional landscapes in mouse hematopoiesis. *Genome Res.* **30**, 472–484 (2020).
50. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
51. Flyamer, I. M., Illingworth, R. S. & Bickmore, W. A. Coolpup.py: versatile pile-up analysis of Hi-C data. *Bioinformatics* **36**, 2980–2985 (2020).
52. Wu, W. et al. Dynamic shifts in occupancy by TAL1 are guided by GATA factors and drive large-scale reprogramming of gene expression during hematopoiesis. *Genome Res.* **24**, 1945–1962 (2014).
53. Luan, J. et al. Distinct properties and functions of CTCF revealed by a rapidly inducible degron system. *Cell Rep.* **34**, 108783 (2021).

Author contributions

H.Z. and G.A.B. conceived the study and designed experiments. H.Z. performed all experiments with help from J.L., D.Z., M.W.V., C.A.K., B.G. and R.C.H. H.Z. performed all data analysis with help from Y.L. H.Z. and G.A.B. wrote the paper with contribution from all listed 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/s41467-021-25418-5.

Correspondence and requests for materials should be addressed to H.Z. or G.A.B.

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

Reprints and permission 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.