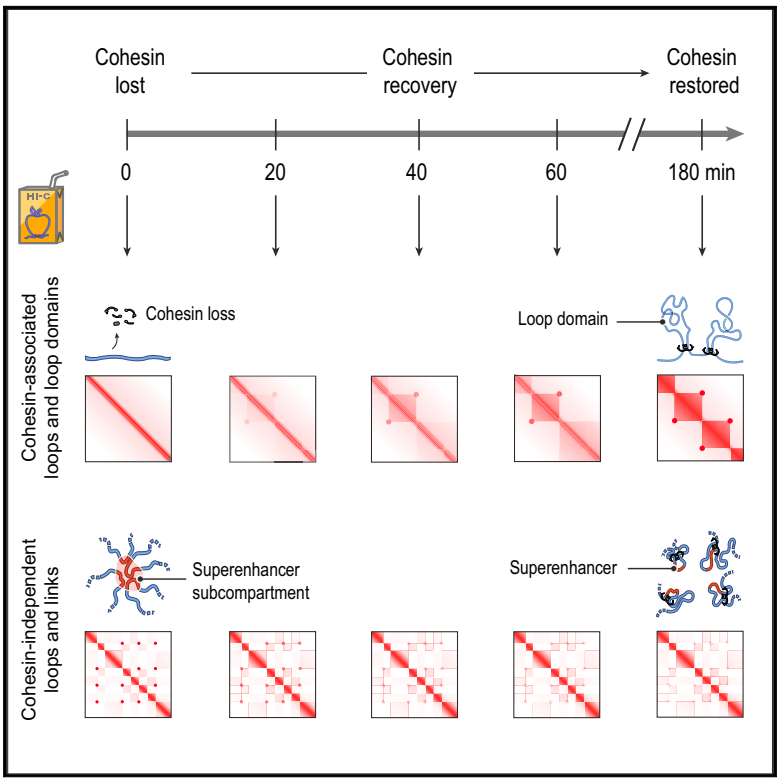
Article

Cell

Cohesin Loss Eliminates All Loop Domains

Graphical Abstract



Highlights

- We track the 4D Nucleome during cohesin loss and recovery, with 10 kb/20 min resolution
- After cohesin loss, loop domains disappear; effects on transcription are modest
- During cohesin recovery, loop domains form in minutes, consistent with fast extrusion
- Superenhancers form loops, interchromosomal links, and higher-order hubs

Authors

Suhas S.P. Rao, Su-Chen Huang, Brian Glenn St Hilaire, ..., Rafael Casellas, Eric S. Lander, Erez Lieberman Aiden

Correspondence

erez@erez.com

In Brief

Mapping the nucleome in 4D during cohesin loss and recovery reveals that cohesin degradation eliminates loop domains but has only modest transcriptional consequences.

Rao et al., 2017, Cell *171*, 305–320
October 5, 2017
https://doi.org/10.1016/j.cell.2017.09.026

CellPress

Cell

Cohesin Loss Eliminates All Loop Domains

Suhas S.P. Rao,[1,2,3] Su-Chen Huang,[1,2] Brian Glenn St Hilaire,[1,2,4] Jesse M. Engreitz,[5] Elizabeth M. Perez,[5] Kyong-Rim Kieffer-Kwon,[6] Adrian L. Sanborn,[1,4,7] Sarah E. Johnstone,[5,8] Gavin D. Bascom,[9] Ivan D. Bochkov,[1,2] Xingfan Huang,[1,10] Muhammad S. Shamim,[1,2,10,11] Jaeweon Shin,[1,10] Douglass Turner,[1,12] Ziyi Ye,[1,10] Arina D. Omer,[1,2] James T. Robinson,[1,5,12] Tamar Schlick,[9,13,14] Bradley E. Bernstein,[5,8] Rafael Casellas,[6,15] Eric S. Lander,[5,16,17] and Erez Lieberman Aiden[1,2,4,5,10,18,*]

[1]The Center for Genome Architecture, Baylor College of Medicine, Houston, TX 77030, USA
[2]Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX 77030, USA
[3]Department of Structural Biology, Stanford University School of Medicine, Stanford, CA 94305, USA
[4]Center for Theoretical Biological Physics, Rice University, Houston, TX 77030, USA
[5]Broad Institute of MIT and Harvard, Cambridge, MA 02139, USA
[6]Lymphocyte Nuclear Biology, NIAMS, NIH, Bethesda, MD 20892, USA
[7]Department of Computer Science, Stanford University, Stanford, CA 94305, USA
[8]Department of Pathology and Center for Cancer Research, Massachusetts General Hospital and Harvard Medical School, Boston, MA 02114, USA
[9]Department of Chemistry, New York University, New York, NY 10003, USA
[10]Departments of Computer Science and Computational and Applied Mathematics, Rice University, Houston, TX 77030, USA
[11]Medical Scientist Training Program, Baylor College of Medicine, Houston, TX 77030, USA
[12]Department of Medicine, University of California, San Diego, La Jolla, CA 92037, USA
[13]Courant Institute of Mathematical Sciences, New York University, New York, NY 10012, USA
[14]NYU-ECNU Center for Computational Chemistry, NYU Shanghai, Shanghai 200062, China
[15]Center of Cancer Research, NCI, NIH, Bethesda, MD 20892, USA
[16]Department of Biology, MIT, Cambridge, MA 02139, USA
[17]Department of Systems Biology, Harvard Medical School, Boston, MA 02115, USA
[18]Lead Contact
*Correspondence: erez@erez.com
https://doi.org/10.1016/j.cell.2017.09.026

SUMMARY

The human genome folds to create thousands of intervals, called "contact domains," that exhibit enhanced contact frequency within themselves. "Loop domains" form because of tethering between two loci—almost always bound by CTCF and cohesin—lying on the same chromosome. "Compartment domains" form when genomic intervals with similar histone marks co-segregate. Here, we explore the effects of degrading cohesin. All loop domains are eliminated, but neither compartment domains nor histone marks are affected. Loss of loop domains does not lead to widespread ectopic gene activation but does affect a significant minority of active genes. In particular, cohesin loss causes superenhancers to co-localize, forming hundreds of links within and across chromosomes and affecting the regulation of nearby genes. We then restore cohesin and monitor the re-formation of each loop. Although re-formation rates vary greatly, many megabase-sized loops recovered in under an hour, consistent with a model where loop extrusion is rapid.

INTRODUCTION

Many studies have shown that the insulator protein CTCF and the ring-shaped cohesin complex colocalize on chromatin (Wendt et al., 2008) and lie at the anchors of loops (Rao et al., 2014; Splinter et al., 2006) and the boundaries of contact domains (also called "topologically constrained domains," "topologically associated domains," or "physical domains") (Dixon et al., 2012; Lieberman-Aiden et al., 2009; Nora et al., 2012; Rao et al., 2014). This suggests that these proteins help regulate genome folding (Merkenschlager and Nora, 2016). Consistent with this, deletion of CTCF sites interferes with loop and contact domain formation (Guo et al., 2015; Sanborn et al., 2015; de Wit et al., 2015). However, initial, low-resolution experiments examining genome-wide depletion of CTCF and cohesin observed only limited effects, reporting that compartments and contact domains still appear to be present (Seitan et al., 2013; Sofueva et al., 2013; Zuin et al., 2014). These results have made it difficult to ascertain the role of CTCF and cohesin in regulating genome architecture.

Here, we examine the effects of cohesin loss on nuclear architecture, epigenetic state, and transcription.

RESULTS

Rapid Degradation of RAD21 Using an Auxin-Inducible Degron System

We employed an auxin-inducible degron (AID) (Natsume et al., 2016) to destroy RAD21, a core component of the cohesin complex. In this system, constitutive expression of the auxin-activated ubiquitin ligase TIR1 leads, in the presence of auxin, to rapid degradation of proteins tagged with an AID domain. We used this system in HCT-116, a human colorectal

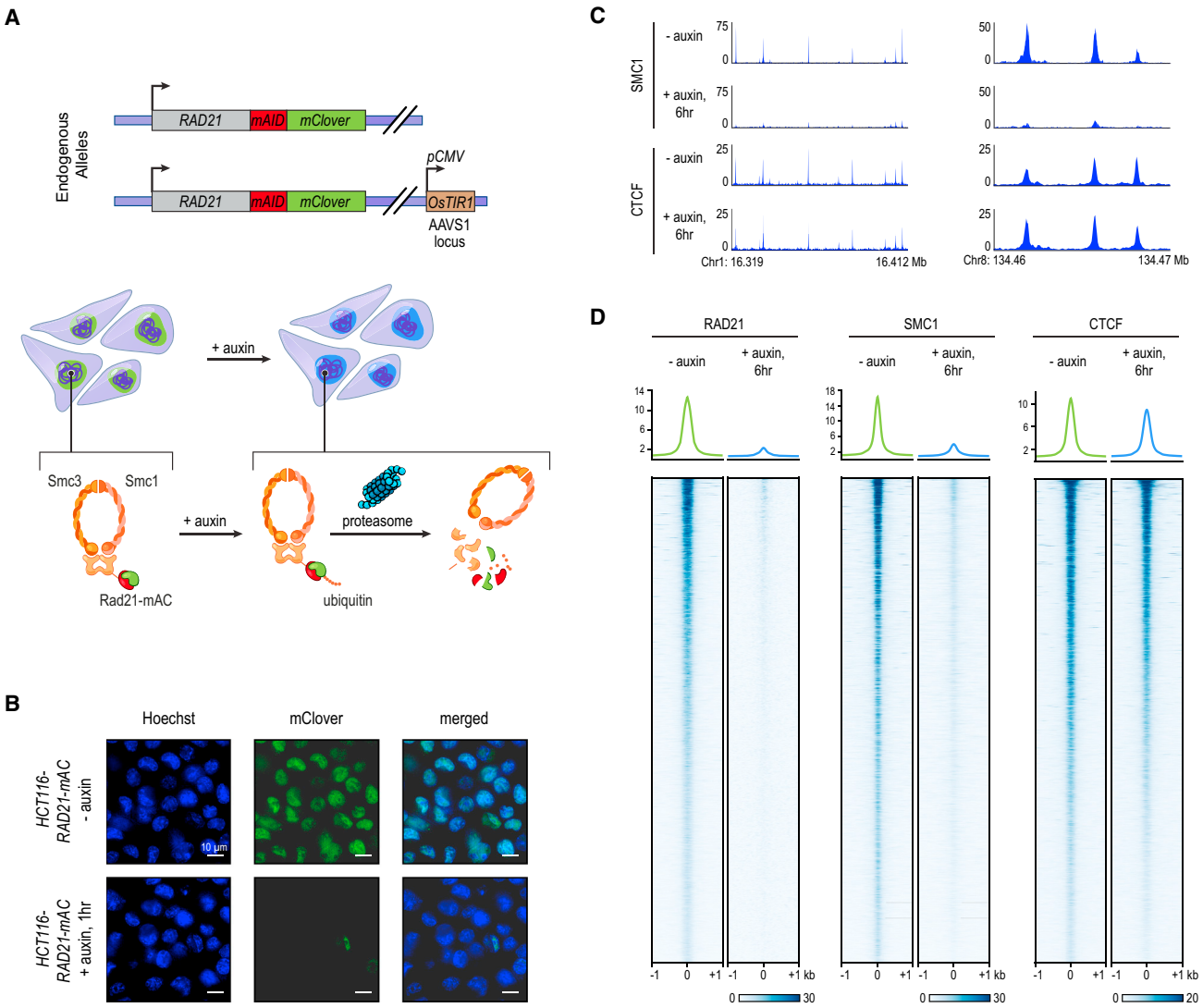


Figure 1. Tagging of Endogenous *RAD21* with an Auxin-Inducible Degron Allows for Rapid, Near Complete Cohesin Loss

(A) In HCT-116-RAD21-mAC cells, both *RAD21* alleles are tagged with auxin-inducible degrons and an mClover reporter, and the *OsTIR1* gene is integrated at the *AAVS1* locus. Auxin treatment leads to proteasomal degradation of RAD21.

(B) Live cell imaging after Hoechst 33342 staining to label nuclei. Nuclear mClover fluorescence corresponding to tagged RAD21 was lost after 1 hr of auxin treatment. See Figure S1.

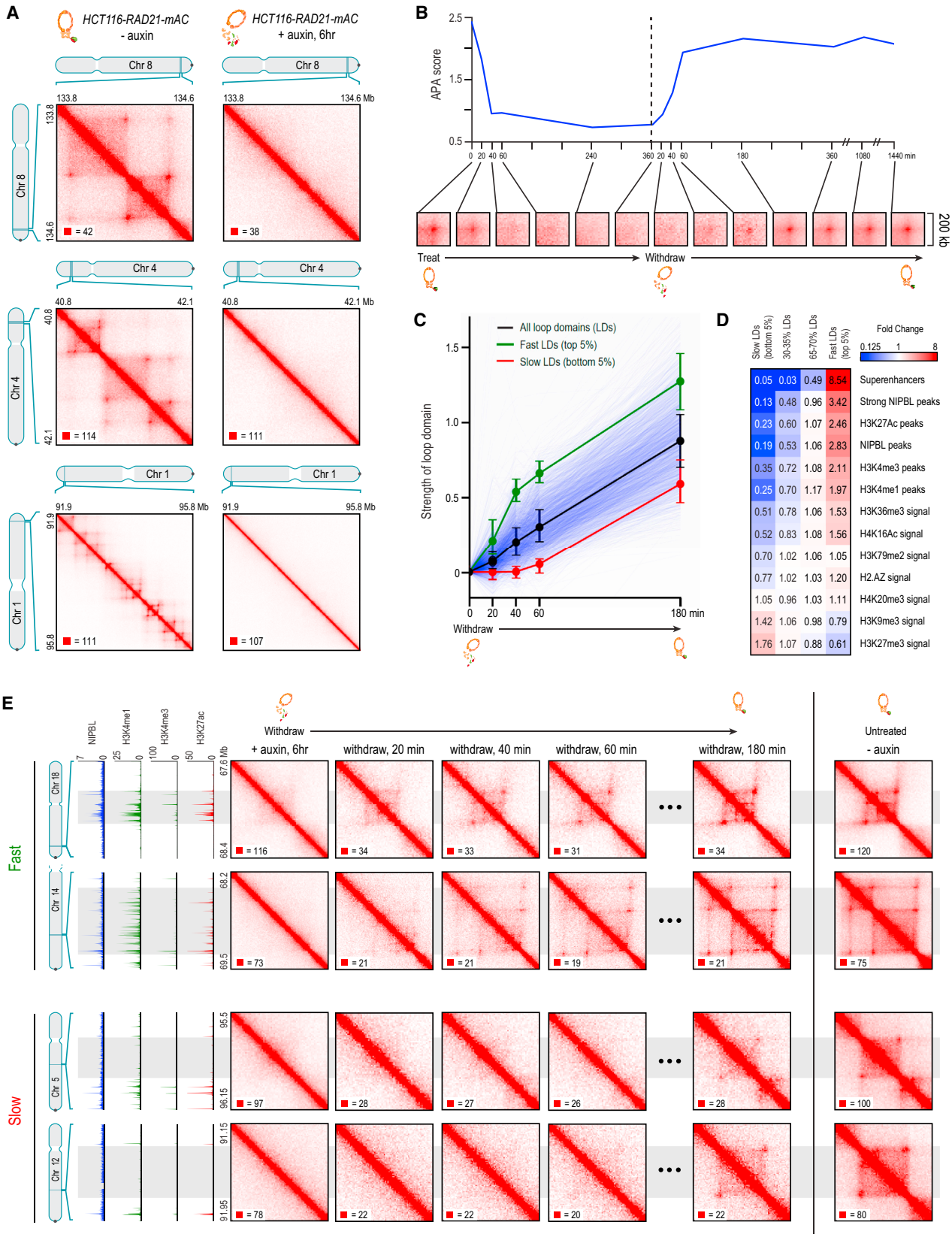
(C) SMC1 and CTCF ChIP-seq signal with and without auxin treatment.

(D) RAD21, SMC1, and CTCF ChIP-seq signal (left, middle, right) across all peaks called for each of the proteins in untreated RAD21-mAC cells. Top: Average enrichments for each protein. After RAD21 degradation, the cohesin complex no longer binds to chromatin. CTCF binding is unaffected.

carcinoma cell line. This cell line had been modified previously by Natsume et al. (2016) so that both *RAD21* alleles were tagged with an AID domain and a fluorescent mClover (RAD21-mAC) (Figure 1A). We confirmed that RAD21-mAC was efficiently degraded after 6 hr of auxin treatment using fluorescence microscopy and ChIP (chromatin immunoprecipitation) (Figures 1B and S1; STAR Methods) and that cohesin could no longer associate with DNA using ChIP-seq with antibodies for SMC1, a different cohesin subunit (Figures 1C and 1D).

Histone Modification Patterns Are Unaffected by Cohesin Loss

First, we first examined the effects of cohesin degradation on key epigenomic features associated with genome folding using ChIP-seq. We examined the distribution of CTCF (associated with loop anchors) and the histone modifications H3K27me3, H3K36me3, H3K27Ac, H3K4me1, H3K4me3, H3K9me3, H4K16Ac, H4K20me3, H3K79me2, and H2.AZ (associated with compartment intervals). Cohesin loss had little effect on these features (Figures 1C, 1D, and S1C–S1E; STAR Methods).



(legend on next page)

Loop Domains Are Rapidly Lost After Degradation of Cohesin

We then turned to study genome folding itself, beginning with loop domains. Loops arise when two loci on the same chromosome are tethered together. (For clarity, the loci will be referred to as "loop anchors," the tethered pair as a "link," and the interval between them as a "loop.") Loop anchors are typically a pair of DNA motifs in the convergent orientation (i.e., the motifs face each other) that bind CTCF and cohesin (Rao et al., 2014). Loops frequently form a contact domain—that is, an interval in which all pairs of loci exhibit higher contact frequency with one another than random pairs of loci at similar distances along the genome sequence. This structure is called a "loop domain" (Rao et al., 2014).

To examine loop domains, we used *in situ* Hi-C (Rao et al., 2014), which combines DNA-DNA proximity ligation and high-throughput sequencing to create maps showing the frequency of physical contact between all pairs of loci across the genome. Loop domains are manifest in Hi-C maps as a bright "peak" pixel (indicating the link between the two loop anchors) at the corner of a bright square (indicating the presence of a contact domain).

We generated roughly 6 billion Hi-C contacts from HCT-116 cells before (3.0 billion) and immediately after (2.9 billion) auxin treatment. In the untreated cells, our algorithms annotated 3,170 loops, of which 2,140 were loop domains. Strikingly, the loop domains disappeared upon cohesin loss. The result was evident by visual examination (Figures 2A and S2C; Data S1, I; [Durand et al., 2016a]). Moreover, the algorithms found only nine loop domains after auxin treatment. Upon close inspection, all were found to be false positives (STAR Methods). (We return below to examine loops *not* associated with contact domains.)

To see whether these changes were reversible, we performed a time-course analysis in which untreated cells were exposed to auxin for 6 hr, after which auxin was withdrawn (Figures 2B and S2B). Low-resolution Hi-C was performed immediately before treatment, as well as at a series of time points during treatment (at 20, 40, 60, 240, and 360 min) and after withdrawal (at 20, 40, 60, 180, 360, 1,080, and 1,440 min). To assess whether the anchors of the loop domains seen in the pre-treatment data continued to be linked (that is, co-located in space) subsequently, we used a method called Aggregate Peak Analysis (APA) (Durand et al., 2016b), which superimposes the signals from a set of peak pixels, thus allowing us to observe an aggregate signal even in sparse datasets where individual signals cannot be resolved (Figure 2B). The APA signal was initially strong, but it was gone by 40 min after initiation of treatment and remained absent throughout the rest of the treatment period (Figures 2B and S2B). The disappearance of the loop-domain links closely mirrored the depletion of cohesin levels in the samples during the treatment period, as ascertained by measuring mClover fluorescence (Figure S1). After auxin was withdrawn, the APA signal quickly increased, largely recovering by 1 hr (Figures 2B and S2B).

These results indicate that the formation of loop domains requires cohesin, that loop domains rapidly disappear after the cohesin tethering the link has been degraded, and that the restoration of cohesin rescues the loop domains.

Loop Domains Spanning NIPBL and Superenhancers Recover More Rapidly

To explore the formation process for individual loop domains, we generated 2.6 billion additional contacts from our Hi-C experiments 20, 40, 60, and 180 min after auxin withdrawal (678 million, 665 million, 618 million, and 675 million contacts, respectively). Using these improved maps, we were able to calculate individual recovery curves for 2,038 of the 2,140 loop domains observed in the untreated cells (Figure 2C; STAR Methods). (The time-resolved maps were not sufficiently deep to assess the 102 loop domains shorter than 100 kb.) Recovery rates for individual loop domains varied dramatically. Faster recovery was strongly associated with high levels of NIPBL binding between the loop anchors, as well as enrichment of promoter and enhancer elements, and of activating histone marks like H3K36me3 and H4K16Ac. By contrast, loop domains that recovered slowly were typically depleted for these features but enriched for repressive marks like H3K27me3 and H3K9me3 (Figures 2D, 2E, and S2D–S2F; Data S1, II). The most extreme difference we found was in the presence of superenhancers (also called "stretch enhancers," or SEs), which are regions of the mammalian genome containing a very high density of enhancer elements and which are marked by extremely high levels of H3K27 acetylation (Hnisz et al., 2013; Parker et al., 2013). Fast loop domains (recovery rate >95th percentile) were 159-fold more likely than slow loop domains (<5th percentile) to span an SE (2.94 SEs/Mb versus 0.02 SEs/Mb) and 26-fold more likely to contain strong NIPBL binding sites (4.23/Mb versus 0.17/Mb) (Figures S2E and S2F).

Taken together, these results indicate that the rate of loop domain re-formation varies greatly across the genome and is

Figure 2. Cohesin Degradation Eliminates Loop Domains

(A) Contact matrices show that loop domains in untreated RAD21-mAC cells (left) disappear after auxin treatment (right). Three representative loci are shown (at 5 [top] 10 kb [middle, bottom] resolution): chr8:133.8-134.6Mb (top), chr4:40.8-42.1Mb (middle) and chr1:91.9-95.8Mb (bottom).

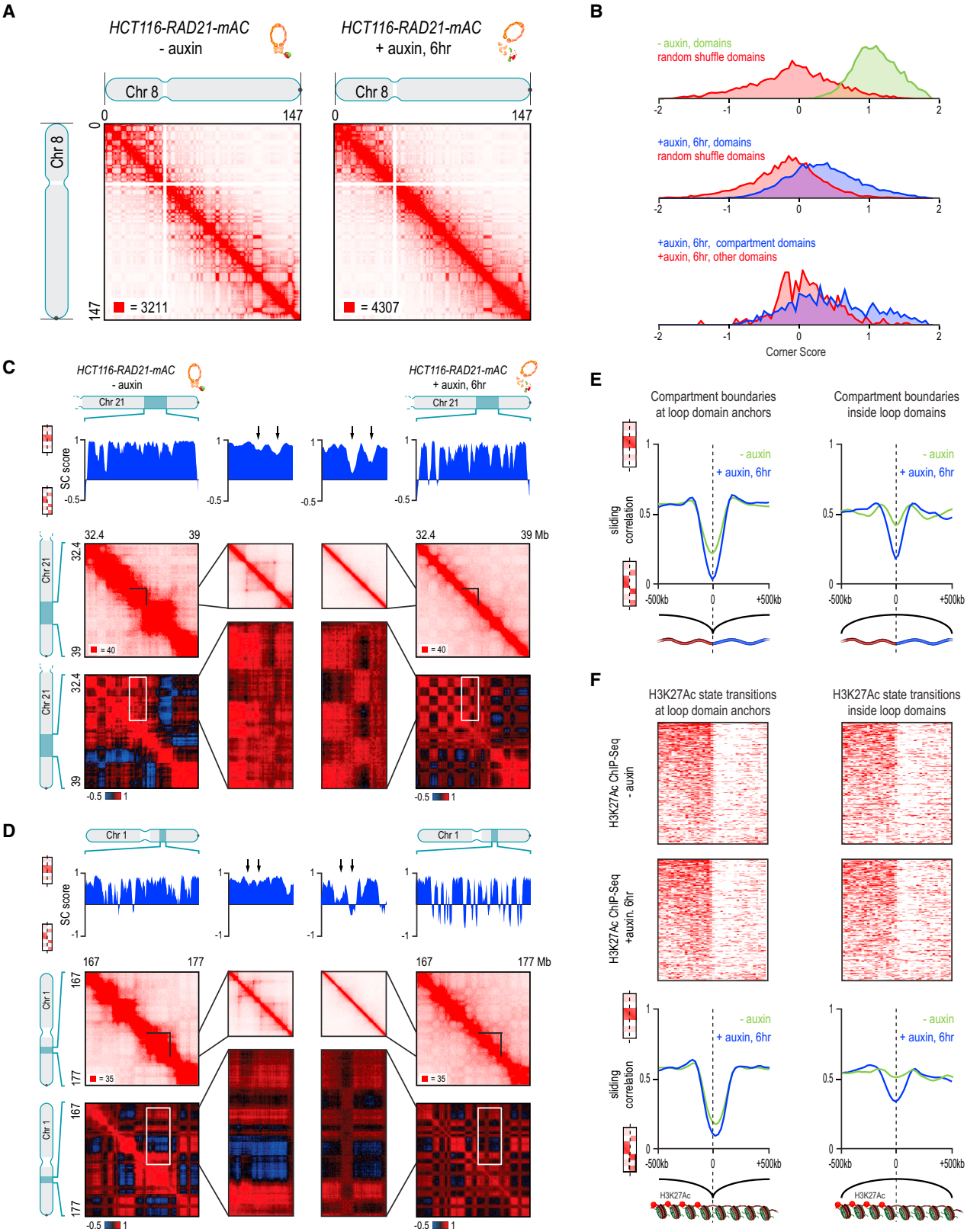
(B) Aggregate peak analysis (APA) was used to measure the aggregate strength of the links associated with all loop domains in low-resolution Hi-C maps generated across a time course of auxin treatment and withdrawal. Top: APA scores; values greater than 1 indicate the presence of loops. Bottom: APA plots; loop strength is indicated by the extent of focal enrichment at the center of the plot. See Figure S2B.

(C) Individual loop reformation curves for each of 1,988 loop domains (blue lines); the number of contacts in the untreated map corresponds to a value of 1, and the number of contacts in the auxin-treated map corresponds to 0. We highlight the median curve for all loops (black), loops at the 5th percentile or lower (red), and loops at the 95th percentile or higher (green) in terms of speed of recovery. See STAR Methods. Error bars indicate 25th and 75th percentile within each subset.

(D) Enrichment of epigenetic features within a loop domain versus speed of recovery. Enrichment is with respect to all intervals spanned by loop domains.

(E) Regions containing fast loop domains (1st row, chr18:67.6-68.4Mb; 2nd row, chr14:68.2-69.5Mb) and slow loop domains (3rd row, chr5:95.5-96.15Mb; 4th row, chr12:91.15-91.95Mb) are shown, along with ChIP-seq tracks (from auxin-treated cells) for NIPBL, H3K4me1, H3K4me3, and H3K27Ac. For fast loop domains, reformation is apparent by 20–40 min after auxin withdrawal, whereas for slow loop domains, reformation is not seen until 3 hr after auxin withdrawal.

An interactive version of this figure is available at http://www.cell.com/cell/9802-figure-2.



(legend on next page)

associated with factors ranging from cohesin loading to local regulatory activity.

Loss of Cohesin Is Associated with Stronger Genome Compartmentalization

Next, we examined the effects of cohesin loss on compartmentalization. Compartmentalization refers to the fact that the genome is partitioned into intervals (which can range from 14 kb to more than 5 Mb) belonging to a small number of types, such that intervals of the same type exhibit an enhanced contact frequency with one another (Lieberman-Aiden et al., 2009; Rao et al., 2014). Intervals are thereby assigned to two compartments (A or B, closely associated with open and closed chromatin, respectively) and, more finely, into six subcompartments (A1, A2, B1, B2, B3, and B4). The "compartment intervals" that lie in a particular compartment are associated with distinctive patterns of chromatin marks (Rao et al., 2014). Because loci within a compartment interval are of the same type, they exhibit an increased contact frequency with one another and frequently form contact domains. In this case, we call the contact domain a "compartment domain." The enhanced contact frequency between compartment intervals in the same subcompartment also gives rise to a plaid pattern in Hi-C maps (Lieberman-Aiden et al., 2009).

Whereas loop domains disappear entirely after cohesin loss, compartmentalization is preserved (Figure 3A). Following auxin treatment, there is no significant change in either the compartment domains, as defined by the presence of the corresponding squares along the diagonal in the Hi-C contact map (Figure 3B; STAR Methods), or in the plaid pattern, as defined by the eigenvectors of the Hi-C correlation map (Figure 3A; mean Pearson's $r = 0.968$ across all chromosomes). Our data are consistent with previous reports that genome compartmentalization is preserved after depletion of cohesin (Seitan et al., 2013) or CTCF (Nora et al., 2017).

We then examined the interaction between compartments and loop domains. Specifically, we examined the compartment boundaries (transition points between compartment intervals) that either (1) lay in the interior of a loop domain in untreated cells or (2) coincided with a loop-domain anchor in untreated cells (Figures 3C–3E). In the former case, the correlation in the genome-wide contact pattern on opposite sides of compartment boundaries showed a much greater decrease in treated versus untreated cells—that is, the plaid pattern across the genome became much stronger in the absence of cohesin (Figures 3C–3E; Data S1, III). The results were similar when we examined boundaries between intervals that were enriched versus depleted for H3K27Ac (which marks intervals in the A compartment [Rao et al., 2014]) or intervals that were enriched versus depleted for H3K27me3 (which marks intervals in the B1 subcompartment [Rao et al., 2014]) (Figures 3F, S3A–S3D; STAR Methods). These data indicate that the compartmentalization process that brings together loci with similar histone marks does not rely on cohesin. On the contrary, the strengthening of the plaid pattern after cohesin loss suggests that the formation of cohesin-dependent loop domains interferes with compartmentalization by promoting the co-localization of locus pairs with different histone modification patterns. Our data are consistent with the observation that genome compartmentalization is weakened in cells where the cohesin-antagonist *WAPL* is knocked out, leading to larger loop domains (Haarhuis et al., 2017).

Links between Superenhancers Are Strengthened after Loss of Cohesin

Next, we examined loops not associated with contact domains. Whereas 1,030 such loops were annotated in untreated cells, only 72 were annotated after cohesin loss. Upon close examination, 57 were false positives (STAR Methods). (The loop-detection algorithms have a higher false-discovery rate after cohesin loss, since true positives are so rare.) The remaining 15 loops were much larger than those seen in untreated cells (median: 1.75 Mb versus 0.275 Mb). Given their large size, we found that loops could be more reliably identified in treated cells by running our peak detection algorithm at coarser resolution (50–100 kb versus 5–10 kb) (STAR Methods). This identified an additional 46 loops that were confirmed by manual inspection (Figures 4A and S4A; STAR Methods). Afterward, the size

Figure 3. Genome Compartmentalization Is Strengthened after Cohesin Degradation

(A) Contact matrices of chromosome 8 at 500 kb resolution. The plaid pattern in the Hi-C map, indicating compartmentalization, is preserved after auxin treatment.

(B) Strength of contact domains called in untreated cells versus random intervals measured using the corner score (STAR Methods) in untreated (top) and treated cells (middle). Contact domain strength is reduced but not entirely lost. The remaining signal comes from compartment domains (bottom). The signal in treated maps from contact domains where both boundaries are contained completely inside a compartment interval ("other domains") is not enriched versus random pixels.

(C and D) Examples ([C], chr21:32.4-39Mb; [D], chr1:167-177Mb) showing that the loss of cohesin-associated loops after auxin treatment results in increased fine-scale compartmentalization. Top: Sliding correlation scores; valleys imply strong differences in long-range contact pattern observed at a locus as compared to neighboring loci, indicating a change in compartment (STAR Methods). Middle: Observed contact matrices. Bottom: Pearson's correlation maps for the local region shown (STAR Methods). Deeper valleys in the sliding correlation score and increased plaid patterning in the observed and Pearson's correlation maps indicate stronger fine-scale compartment interactions after auxin treatment. Blowouts: Loss of a loop domain results in strengthening of a compartment boundary spanned by the loop. Blown-out regions are indicated on zoomed out maps for both the observed (black upper triangle) and Pearson's correlation maps (white rectangle). Observed and Pearson's correlation maps are both shown at 25 kb resolution for the zoomed out matrices and 10 kb and 25 kb resolution respectively for the blown-out matrices.

(E) Sliding correlation scores before and after auxin treatment for compartment boundaries,which either coincide with loop domain anchors (left) or are located in the interior of a loop domain (right).

(F) Sliding correlation scores before and after auxin treatment for H3K27ac boundaries in untreated cells, which either coincide with loop domain anchors (left) or are located in the interior of a loop domain (right). Top and middle: H3K27Ac modification patterns are unchanged after auxin treatment.

Interactive figure: http://www.cell.com/cell/9802-figure-3.

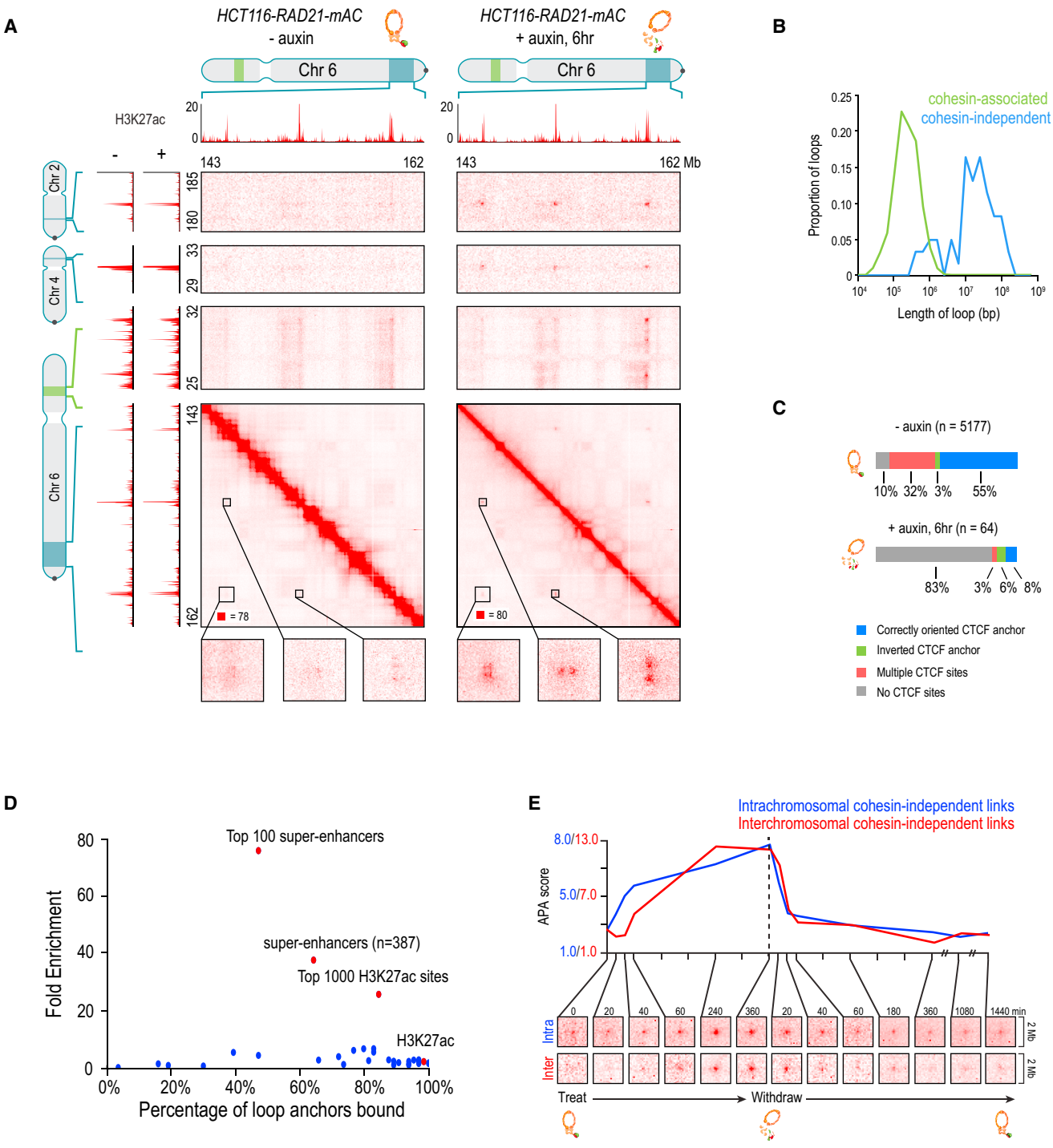


Figure 4. Cohesin Loss Causes Superenhancers to Co-localize, Forming Hundreds of Links within and across Chromosomes

(A) A network of intra- and interchromosomal cohesin-independent links between superenhancers on chr6, chr4, and chr2. H3K27 acetylation does not change with auxin treatment, but cohesin-independent links are significantly strengthened upon treatment. Intrachromosomal matrices are shown at 25 kb (on-diagonal) and 50 kb (off-diagonal) resolutions; interchromosomal matrices are shown at 100 kb resolution. Maximum color intensities are 28 reads for the off-diagonal intrachromosomal matrices and 20 reads for the interchromosomal matrices.

(B) Length distribution of cohesin-associated loops (green) versus cohesin-independent loops (blue).

(C) Top: CTCF binding patterns at cohesin-associated loop anchors. Bottom: CTCF binding patterns at cohesin-independent loop anchors.

(legend continued on next page)

difference between the 61 "cohesin-independent loops" and the cohesin-associated loops was even more dramatic (Figure 4B, median size: 23.15 Mb).

We sought to understand the basis of these cohesin-independent loops. We found that they do not demarcate the boundaries of contact domains (0 of 61 [0%] versus 2,140 of 3,170 [68%] for cohesin-associated loops). Remarkably, many cohesin-independent loop anchors form links with each another—manifest as focal peaks in the Hi-C heatmap—even when the anchors reside on different chromosomes (Figure 4A; Data S1, IV). In total, we identified 203 such interchromosomal links. Interchromosomal links are not seen between the anchors of cohesin-associated loops.

The anchors of cohesin-independent loops also exhibit very different patterns of protein binding. The proportion that binds CTCF is much lower (20% versus 90% for cohesin-associated loops; Figures 4C and S4A). Moreover, there is no tendency for the CTCF motifs at loop anchors to point into the loop (5 of 9 [56%] point inward versus 2,770 of 2,919 [95%] for cohesin-associated loops).

Notably, the cohesin-independent loop anchors are highly enriched for superenhancers. We found that 41 of the 64 cohesin-independent loop anchors overlapped with the 387 superenhancers in HCT-116 cells—a 37.5-fold enrichment ($p < 10^{-15}$). For the 100 strongest superenhancers, the enrichment was 76-fold (30 of 64, $p < 10^{-15}$; Figures 4D and S4B). Interestingly, cohesin-independent loops and the associated loops and links between superenhancers could be seen in the untreated cells as well but were much weaker (Figures 4A, 4E, 5A, and S4C–S4E; Data S1, IV; STAR Methods).

Strikingly, we observed large cliques forming between the anchors of the cohesin-independent loops (Figure 5A; Data S1, V). Large cliques are not seen for cohesin-associated loops (Figure S2A).

In many respects, the cohesin-independent loops resemble the superloops we previously observed on the inactive X chromosome (Darrow et al., 2016; Rao et al., 2014): they are very large (up to 77 Mb), the intervals they span do not form contact domains, and their anchors tend to form cliques and are marked by H3K27 acetylation (Figure S4F). We also found that the superloops tend to occur simultaneously, forming hubs containing three or more loci. We therefore wondered whether cohesin-independent loop anchors would exhibit the same behavior.

To probe this question, we examined concatemers—that is, Hi-C reads that bring together three or more loci and indicate that the loci in question were simultaneously co-located in a particular cell during the Hi-C experiment. We identified 57 million unique "triples," which bring together three loci; 32 million unique quadruples; and 130 thousand quintuples (Figure 5B; STAR Methods). Instead of a 2D heatmap, concatemers are naturally represented as an n-dimensional matrix, or tensor, showing the collision frequency (i.e., the frequency of simultaneous physical contact) between any set of n loci in the genome (Figure 5C). Hubs involving n loci manifest in the n-dimensional contact tensor as peaks in collision frequency with respect to the local n-dimensional neighborhood.

Because the number of entries in an n-dimensional contact tensor scales as the genome size to the n^{th} power, contact tensors can be exceedingly sparse. We therefore did not expect to see triples corresponding to any particular set of three cohesin-independent loop anchors. Instead, we developed a variant of APA for contact tensors, superimposing the signal from all possible sets of three cohesin-independent loop anchors that lie on a single chromosome (STAR Methods). This analysis revealed 11 collisions involving three cohesin-independent loop anchors in the auxin-treated data, as compared to an expected value of 0.41 collisions based on the density of collisions in the local neighborhood (Figures 5D–5F, S5A, and S5B). These findings indicate that, like superloops, cohesin-independent loop anchors tend to form hubs involving three or more loci. By contrast, no collisions were found in the untreated data. This is consistent with our finding that cohesin-independent loops are much weaker in the presence of cohesin.

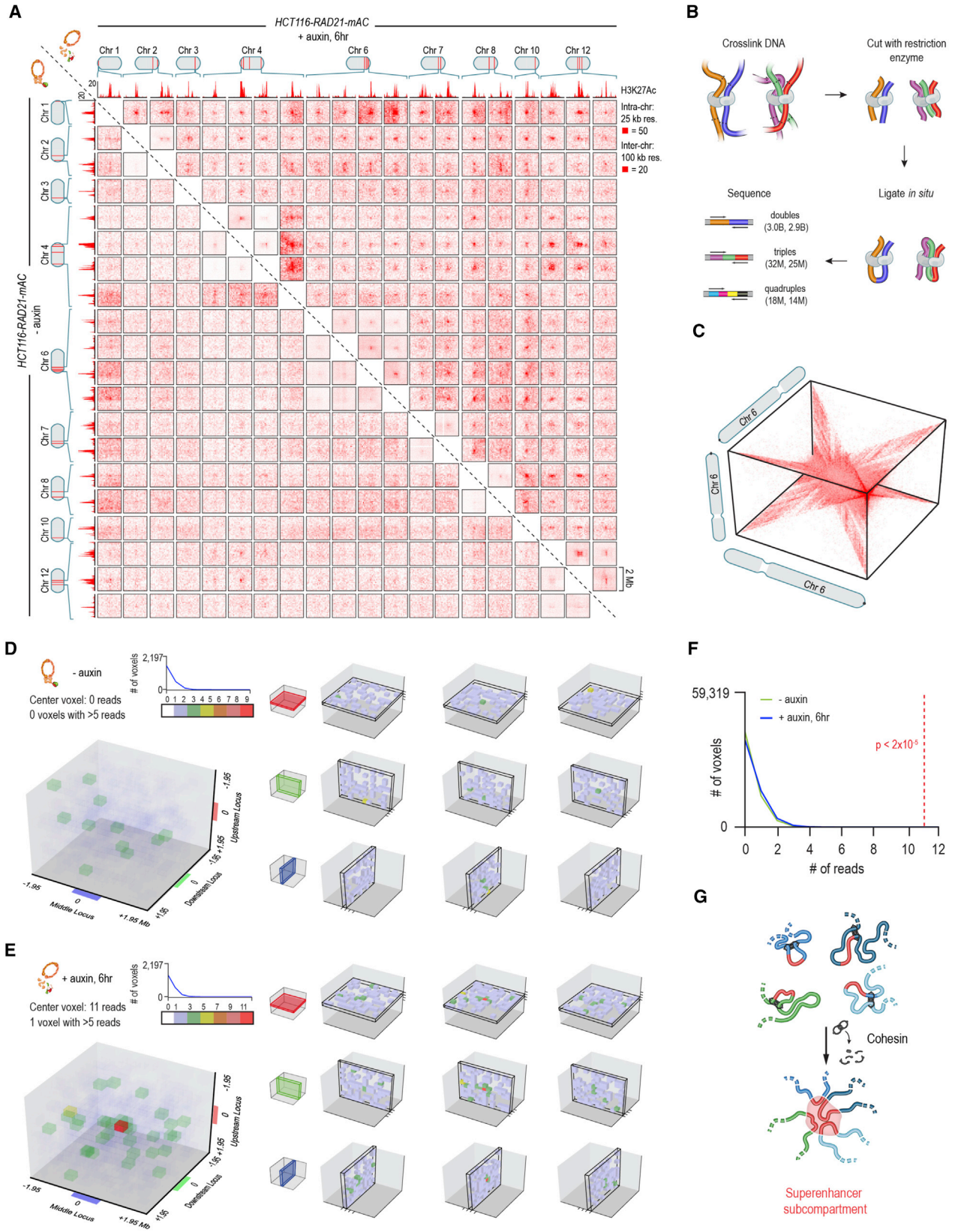
Molecular Dynamics Simulations Integrating Extrusion and Compartmentalization Can Recapitulate Hi-C Experimental Results

To test the hypothesis that the Hi-C contact maps we observed are consistent with the presence of two distinct folding mechanisms, we modeled a 2.1 Mb region on chromosome 3 (Sanborn et al., 2015). Our model treated the region as a block copolymer consisting of two types of chromatin, A or B, determined by classifying loci based on ChIP-seq binding data and containing CTCF binding sites, whose position and strength were derived from CTCF and SMC1 ChIP-seq tracks and whose orientation was determined by examining the human genome reference (Figure 6A; STAR Methods). We used molecular dynamics simulations to examine the behavior of this polymer in a solvent containing extrusion complexes (thus modeling loop extrusion [Fudenberg et al., 2016; Sanborn et al., 2015]) and in the presence of attractive forces between like monomers (thus modeling compartmentalization [Di Pierro et al., 2016]). The resulting ensemble was used to create an *in silico* contact map for the region.

We found that the resulting contact maps accurately recapitulated the experimental results in both untreated and treated cells (Figures 6A and 6B). The maps also illustrate the change in long-range contact pattern that is seen when a loop spans a compartment boundary (Data S1,VI). These findings suggest that the most prominent features observed in Hi-C data sets (loops, domains, and compartments) can be recapitulated by simulations that use only ChIP-seq data as input, in the context of a model combining extrusion and compartmentalization.

The above simulations assume that phase separation leads to the compartmentalization of chromatin intervals bearing similar patterns of histone marks. There are multiple models that could account for such a tendency. In one, phase separation is facilitated by protein chaperones that recognize histone tails.

(D) Percent of cohesin-independent loop anchors bound versus fold enrichment for 36 DNA-binding proteins and histone modifications.
(E) APA for intrachromosomal (blue) and interchromosomal (red) cohesin-independent links across a time course of auxin treatment and withdrawal. Top: APA scores; bottom: APA plots.
Interactive figure: http://www.cell.com/cell/9802-figure-4.



(legend on next page)

Alternatively, similar nucleosomes might directly attract one another through histone tail interactions. To see whether tail interactions can guide folding at the oligonucleosome scale given the mechanical, electrostatic, and entropic constraints on chromatin fibers, we simulated short fibers using a mesoscale approach (Bascom and Schlick, 2017; Bascom et al., 2017; Collepardo-Guevara et al., 2015; Grigoryev et al., 2016). These simulations have three components: linker DNA beads (each representing 9 bp), nucleosome core particles (rigid bodies with charged surfaces), and histone tail beads (each representing 5 aa) (Figures S6A and S6B). We found that histone tail interactions overcame constraints on the chromatin chain, leading either to focal contacts between short chromatin intervals or to global condensation, depending on the initial configuration and epigenetic state of the fiber (Figures S6C–S6E). Of course, these findings do not rule out a central role for protein chaperones *in vivo*.

Cohesin Loss Results in Strong Down-Regulation of Genes near Superenhancers but Does Not Bring About Widespread Ectopic Activation

Finally, we sought to investigate the role of cohesin in regulating gene expression. Cohesin has been proposed to facilitate interactions between enhancers and promoters (Kagey et al., 2010; Merkenschlager and Nora, 2016). Loop domains are thought to regulate this process by preventing enhancers from forming ectopic interactions with targets that lie in a different loop domain (Lupiáñez et al., 2015; Flavahan et al., 2016). We therefore characterized the effects of cohesin loss on nascent transcription by performing precision nuclear run-on sequencing (PRO-seq) in treated and untreated HCT-116 cells (Engreitz et al., 2016; Jonkers and Lis, 2015) (Figure 7A). We chose an early timepoint—6 hr after auxin treatment—with the aim of examining direct consequences rather than indirect effects due to changes in cell state.

To look for signs of ectopic activation, we examined the 14,853 genes that were not expressed (RPKM [reads per kb per million mapped reads] < 0.5) in untreated cells. Of these genes, 1% (216) were activated after treatment ($p < 0.05$, >30% change in RPKM, RPKM > 0.5 in treated cells). Thus, while cohesin plays a role in preventing ectopic activation, most genes remain inactivated even in the absence of cohesin and loop domains.

We next looked for changes in the 12,222 genes that were expressed (RPKM > 0.5) in untreated cells (Figure 5B). Here again, most genes (87%, 10,615) exhibited similar levels of transcription after cohesin degradation (RPKM changed by less than 30%). Strong effects were infrequent: 64 genes (0.5%) showed a 2-fold change, and 2 genes showed a 5-fold change (Figure 7B). While the quantitative impact may seem modest, such changes can have important biological impacts (Flavahan et al., 2016).

Of genes that exhibited a strong change in transcription, more were down-regulated than up-regulated (61% versus 39%), suggesting that cohesin-associated loops may both facilitate activation of promoters by their distal enhancer elements and block activation by inappropriate enhancers, with the former being somewhat more common.

We wondered how cohesin facilitated these promoter-enhancer contacts. We noticed that many of the genes that were down-regulated (by >1.75-fold) were located within 500 kb of superenhancers (23 of 49, 4.8-fold enrichment, Figures

Figure 5. In the Absence of Cohesin, a Clique Spanning More Than 20 Superenhancers Forms Pairwise Links and Higher-Order Hubs

(A) The interactions between 20 cohesin-independent loop anchors spread across nine chromosomes are shown before (lower triangle) and after (upper triangle) auxin treatment. Each matrix shows a 2 × 2 Mb region centered on the respective anchors. Intrachromosomal interactions are shown at 25 kb resolution; interchromosomal interactions are shown at 100 kb resolution. The anchors are strongly enriched for H3K27 acetylation both before and after auxin treatment. (ChIP-seq data is shown at 25 kb resolution.) Cohesin loss causes the anchors to form a clique, with focal interactions seen between nearly all pairs of loop anchors, regardless of whether they lie on the same chromosome.

(B) In addition to pairwise contacts, *in situ* Hi-C generates concatemers spanning three or more fragments. There are millions of triples (chimeric reads that align to three loci) and quadruples (chimeric reads that align to four loci) in both our untreated and auxin-treated *in situ* Hi-C data sets for RAD21-mAC cells. The numbers in parentheses indicate the number of n-mer contacts observed in the untreated (left) and auxin-treated (right) data.

(C) 3D tensor showing collisions between three loci on chromosome 6 at 1 Mb resolution. See STAR Methods.

(D) Left: 3D aggregate peak analysis (APA) using the untreated *in situ* Hi-C data for all 131 intrachromosomal trios of cohesin-independent loop anchors, chosen so that each anchor in a trio lies on the same chromosome as the other two anchors, but no two anchors in a trio lie within 10 Mb of one another. To create a 3D APA cube, we excise a 3.9 × 3.9 × 3.9 Mb subtensor centered on each trio and superimpose the results. The cube is shown at 300 kb resolution (i.e., each voxel corresponds to all collisions between three loci, each 300 kb in length). The subtensors are oriented such that the locus closest to the p terminus of a chromosome is always located on the z axis, the one closest to the q terminus is located on the y axis, and the locus in between is located on the x axis. The number of collisions in a voxel is indicated by its color; the histogram above the color scale shows the number of voxels of each color. No voxel contains more than five collisions, and the center voxel—reflecting all collisions between three cohesin-independent loop anchors—contains no collisions at all. Right top row: The central cross-section in z is shown, flanked by the two adjacent cross-sections. Middle row: The central cross-section in y, flanked by the adjacent cross sections. Bottom row: The central cross section in x, flanked by the adjacent cross sections. There is no enrichment at the center of the 3D APA cube.

(E) The preceding analysis is repeated using the auxin-treated data. Now, the center voxel contains 11 collisions, whereas no other voxel contains more than 5 collisions. These findings indicate that, in the absence of cohesin, cohesin-independent loop anchors tend to co-localize to form hubs containing three or more anchors.

(F) Histogram of number of voxels versus number of collisions for the two 3D-APA cubes shown in (D) and (E), as well as for 52 control 3D-APA cubes obtained by shifting one or more of the loci in each of the above trios by 3.9 Mb. With the exception of the central voxel in the auxin-treated 3D-APA cube, which contains 11 collisions, no voxel contains more than 8 collisions. This indicates that the observation of 11 collisions purely by chance is exceedingly unlikely.

(G) Under normal circumstances, loop extrusion facilitates short-range contacts between superenhancers and neighboring loci. Upon cohesin loss, superenhancers begin to co-localize, even when located on different chromosomes, and thereby form a subcompartment.

Interactive figure: http://www.cell.com/cell/9802-figure-5.

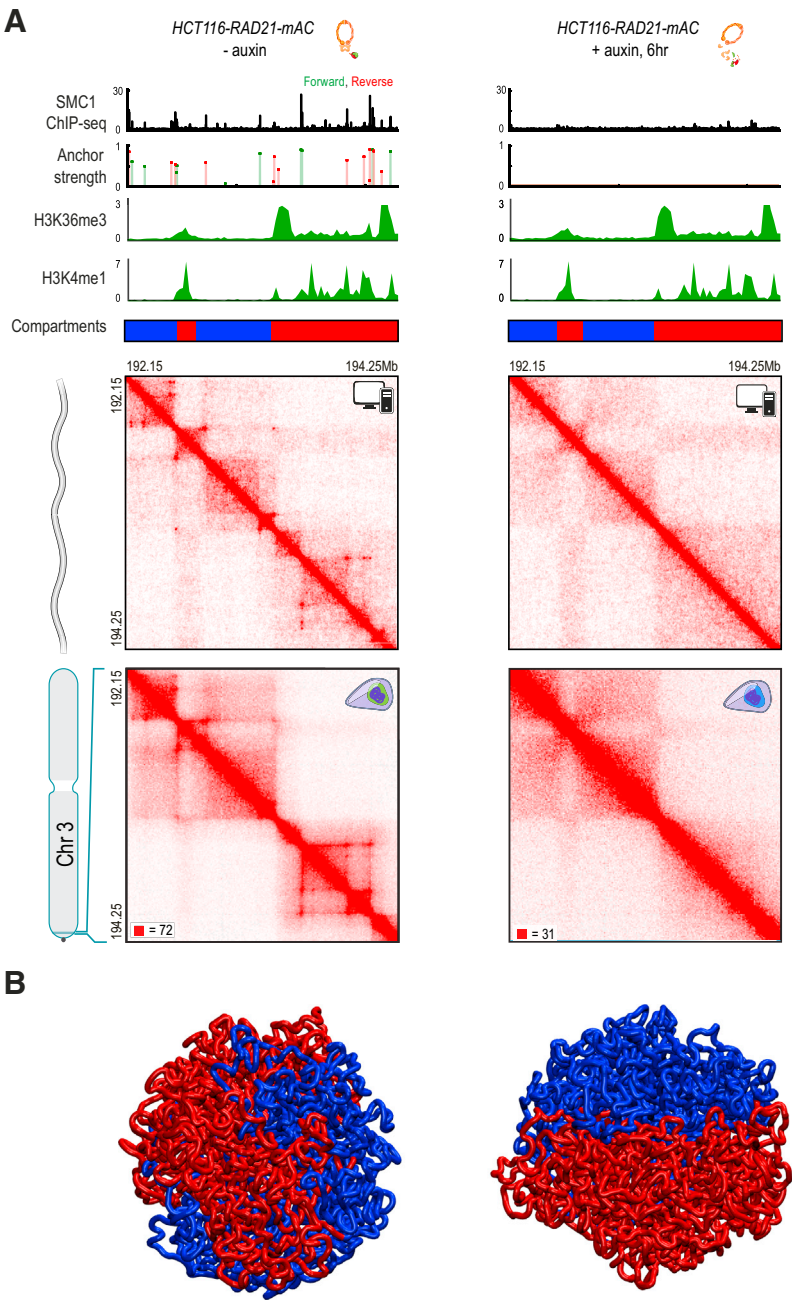


Figure 6. Molecular Dynamics Simulations Combining Extrusion and Compartmentalization Can Recapitulate Hi-C Experimental Results

(A) We use loop extrusion and compartmentalization to simulate a 2.1 Mb region on chromosome 3 in RAD21-mAC cells before (left) and after (right) auxin treatment. CTCF and SMC1 ChIP-seq signals are normalized and converted into binding probabilities for the simulated extrusion complex (first and second rows). Each peak is assigned a forward (green) or reverse (red) orientation based on the corresponding CTCF motif. ChIP-seq data for nine histone modifications were used to classify loci into two compartments (red and blue, fifth row). Histone modification data for H3K36me3 and H3K4me1 is shown, illustrating the correspondence between the classification tracks and the underlying ChIP-seq signals (third and fourth rows). The simulations yield an ensemble of polymer configurations. We show contact maps from the simulated ensemble (top) and from the corresponding Hi-C experiments (bottom).

(B) Examples of globules from simulations of compartmentalization with extrusion (left) and without (right). The globule without extrusion shows stronger segregation of compartment types.

Interactive figure: http://www.cell.com/cell/9802-figure-6.

7C, 7D, S7A, and S7B). Of these genes, 29% were located within 500 kb of one of the top 100 superenhancers (8.5-fold enrichment). Strikingly, these superenhancers were often located at the anchors of the cohesin-independent links seen in treated cells (8 of 19, a 13.7-fold enrichment).

The above results are interesting from the standpoint of transcriptional regulation. In the absence of cohesin, superenhancers associated with the down-regulated genes exhibit a strong tendency to form links with one another. By contrast, in the presence of cohesin, the majority of these superenhancers were located in the interior of a cohesin-associated loop (in 13 of 19 cases) and the long-distance cohesin-independent links were much weaker.

DISCUSSION

Here, we explore the 4D Nucleome of a human colon cancer cell line during cohesin loss and recovery, achieving Hi-C map resolutions of 5 kb with a time resolution of 20 min. We find that cohesin is required for the establishment and maintenance of loop domains. After cohesin loss, we also find that (1) histone marks are unchanged; (2) compartment structure is strengthened in the absence of cohesin, as loop domains spanning multiple compartment intervals lead to mixing among loci in different compartments; and (3) only a small subset of genes exhibit large changes in transcription level. As auxin is withdrawn, cohesin levels recover, and we are able to measure the rate of formation for nearly every loop domain, genome-wide. Loop domains that recover quickly are much more likely to span superenhancers and binding sites of the cohesin loading factor NIPBL. Finally, we identify a class of cohesin-independent loops, links, and hubs connecting superenhancer loci on many chromosomes.

Comparison with Other Studies

It is important to set our results in the context of other studies related to loop formation. While there is clear agreement that deletion of individual CTCF binding sites can result in a loss of cohesin binding and can abolish loops and contact domains (Guo et al., 2015; Sanborn et al., 2015; de Wit et al., 2015), there have been conflicting reports about the effects of depleting cohesin or CTCF genome-wide.

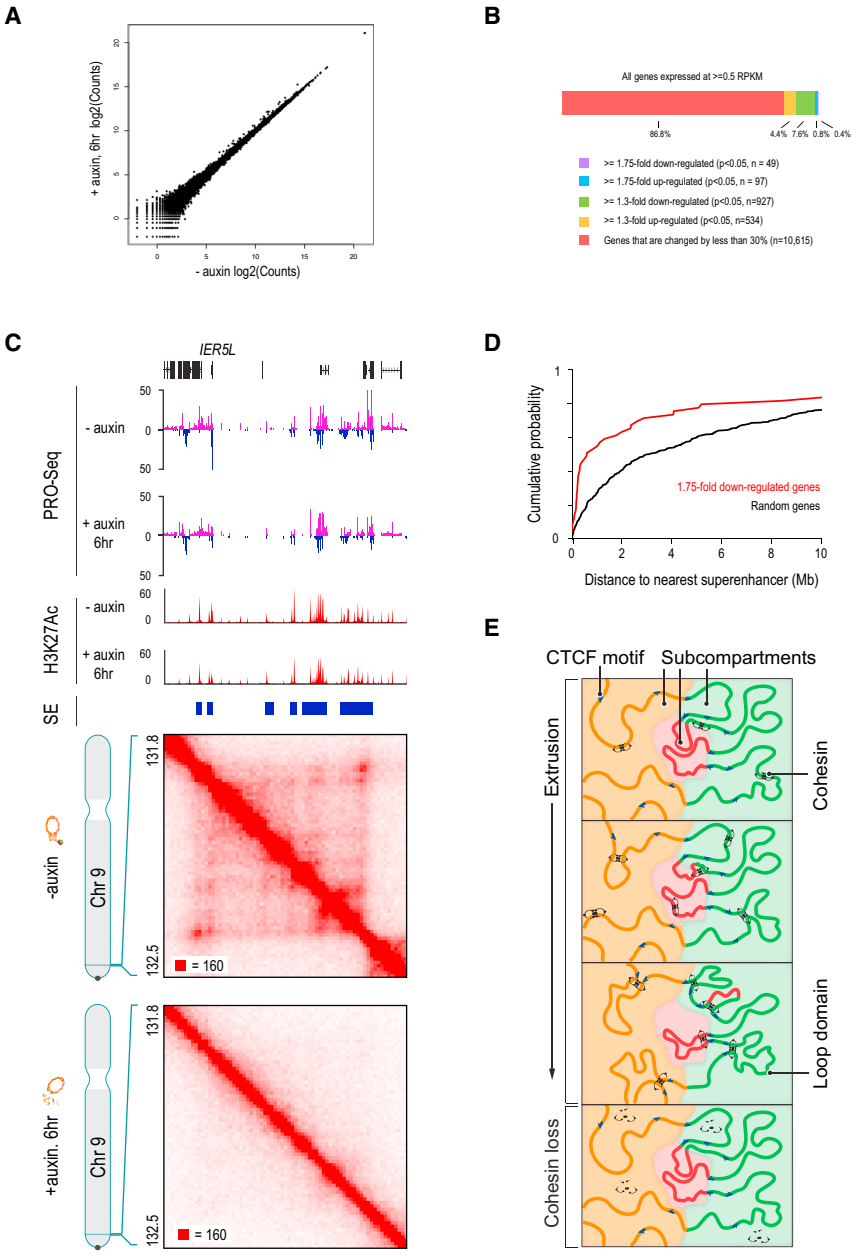


Figure 7. Cohesin Degradation Results in Strong Down-Regulation of Genes near Superenhancers but Does Not Result in Widespread Ectopic Gene Activation

(A) Scatter plot of gene-wide PRO-seq counts in RAD21-mAC cells before (x axis) and after (y axis) treatment.

(B) Genes that are expressed in untreated cells rarely undergo substantial changes in expression level after cohesin loss.

(C) An example of a strongly down-regulated gene near a superenhancer. In untreated cells, a series of cohesin-associated loops form between the *IER5L* promoter and nearby superenhancers. Upon auxin treatment, these loops are lost, and *IER5L* expression is 2.6-fold down-regulated.

(D) Cumulative probability distributions of distances to the nearest superenhancer for 1.75-fold down-regulated genes after auxin treatment (red) versus random genes (black).

(E) A model of how extrusion and compartmentalization combine to shape the spatial organization of the genome inside the nucleus. Intervals of chromatin with similar patterns of histone modification co-localize in nuclear subcompartments. Loop extrusion facilitates short-range contacts between nearby loci as the two subunits of the cohesin-based extrusion complex translocate in opposite directions on chromatin. The extrusion subunits halt at CTCF motifs facing inward, thus forming a loop domain between a pair of motifs in the convergent orientation. Loop domains represent dynamic structures that are maintained by cohesin; only a subset of them may be present at any given time. When the loop anchor motifs span multiple compartment intervals, the dynamics of loop extrusion interfere with compartmentalization by facilitating contacts between loci in different compartments. Loss of cohesin leads to the disappearance of loop domains and to a closer correspondence between genome compartmentalization patterns and histone modification patterns.

Interactive figure: http://www.cell.com/cell/9802-figure-7.

Early Hi-C studies of cohesin and CTCF depletion, using gene knockouts, knockdowns, and proteolytic cleavage, reported that contact domains remained (Seitan et al., 2013; Sofueva et al., 2013; Zuin et al., 2014). The discrepancy may be due to (1) the fact that low-resolution Hi-C analysis cannot distinguish between loop domains, which are sensitive to cohesin depletion, and compartment domains, which are not; and (2) the possibility that the cohesin depletion was incomplete (Figure S2C). More recently, CTCF depletion followed by higher-resolution Hi-C (Nora et al., 2017) revealed the disappearance of a subset of contact domains. Our data for cohesin depletion are consistent with this study.

Two recent studies have also sought to examine the effects of depletion of NIPBL, which encodes a cohesin loader protein. They report opposite conclusions. The Schwarzer et al. (2016) preprint reported, on the basis of new experiments, that the genome-wide aggregate signal from loop domains is absent following NIPBL deletion. By contrast, a recent publication reported the continued presence of individual loop domains after the near-complete depletion of NIPBL, although these loop domains were abnormally small (Haarhuis et al., 2017). In any case, it is unclear what effect NIBPL depletion would be expected to have on loop domain formation because, although NIPBL facilitates cohesin loading, it may not be essential for cohesin loading. Indeed, cohesin loading independent of NIPBL has been observed *in vitro,* albeit at low efficiency (Davidson et al., 2016; Stigler et al., 2016).

Finally, an exciting recent experiment demonstrated that deletion of *WAPL*, a cohesin antagonist that removes cohesin from chromatin, results in the formation of thousands of new loops

and loop domains, which are larger than those found when WAPL is intact (Haarhuis et al., 2017). Our results are consistent with these findings.

Two Mechanisms that Guide Genome Folding

Our results highlight two distinct mechanisms that guide genome folding.

The first is the cohesin-dependent formation of loop domains. The data presented above are consistent with several models of this process. We (Sanborn et al., 2015) and others (Alipour and Marko, 2012; Nichols and Corces, 2015; Fudenberg et al., 2016; Nasmyth, 2001) have hypothesized that the underlying physical process is the formation of loops by extrusion. In this model, loop domains form when a hypothetical cohesin-based extrusion complex (Xcom), which comprises two physically tethered subunits, binds chromatin at a particular location; subsequently, the subunits slide in opposite directions until they arrive at a bound CTCF protein. Thus, the disappearance of cohesin can eliminate all loop domains without influencing CTCF binding. Other models include the initial formation of loops via 3D diffusion of anchor loci, followed by cohesin-mediated stabilization. See Sanborn et al. (2015) for a fuller discussion.

The second mechanism is the cohesin-independent compartmentalization of chromatin intervals with similar histone marks (Lieberman-Aiden et al., 2009; Rao et al., 2014). This observation is also compatible with several models: histone modifications might drive the formation of compartments (i.e., "phase separation" [Hnisz et al., 2017; Jost et al., 2014; Di Pierro et al., 2016]), compartmentalization might lead to histone remodeling, or both processes might be caused by a third mechanism.

The data presented above—specifically, the fact that cohesin loss does not affect histone modifications but does cause long-range contact patterns to better match patterns of histone marks—is more consistent with histone patterns governing genome compartmentalization, rather than the reverse. This phase separation process could involve histone-tail interactions or the binding of reader proteins that target modified histones to specific locations in the nucleus (Wijchers et al., 2016). Interestingly, two recent studies have provided evidence that the protein HP1a, which binds the heterochromatic H3K9me mark, forms liquid droplets *in vivo* via phase separation, such that H3K9me heterochromatin is contained within the droplets (Larson et al., 2017; Strom et al., 2017). Of course, it remains possible that independent mechanisms may shape both histone mark patterns and genome compartmentalization.

The Speed of Loop Extrusion

We show that loop domains disappear shortly after auxin-induced cohesin loss and reappear shortly after auxin withdrawal, implying that (1) they require cohesin both for formation and maintenance and (2) they do not represent stable states of chromosome condensation.

In the loop extrusion model, the two physically tethered subunits of the Xcom bind chromatin at a single site and then slide in opposite directions along chromatin. Measurements of loop re-formation enable estimates of extrusion speed. For instance, in Figure 2E, we show a ~900 kb loop that is restored within 40 min of auxin withdrawal. Thus, this loop is extruded at an average rate of no less than 375 bp/s, with each Xcom subunit sliding at no less than 188 bp/s. Our estimates are lower bounds, as they ignore the time needed for auxin to disappear and for cohesin to re-form and be loaded on chromatin. (Given the correlation between loop re-formation speed and Nipbl, loading time may be a significant consideration.)

These estimates are similar to ones obtained studying the SMC complex in *B. subtilis* (500–1,000 bp/s) (Wang et al., 2017).

The rate estimates bear on the protein motors involved when Xcom subunits slide. For instance, it is possible that cohesin itself serves as a motor during the extrusion process. However, single-molecule studies of cohesin translocation *in vitro* have yielded an estimated sliding rate of only 1–2 bp/s on chromatin (Davidson et al., 2016; Stigler et al., 2016). These could indicate that cohesin alone—under the conditions probed in those experiments—is unlikely to be the principal source of translocase activity. Similarly, RNA polymerase II can push cohesin along DNA (Davidson et al., 2016), but elongation rate estimates (9–90 bp/s; [Jonkers and Lis, 2015]) are slower than what we observe. This suggests that other translocases may be involved in loop extrusion.

Two Classes of Loops

We observe a population of loops that are frequently anchored at superenhancers and do not depend on cohesin.

One explanation is that these loops, too, form by extrusion but using alternative protein complexes, such as condensin, instead of cohesin. However, the superenhancer anchors also form links with one another when they lie on different chromosomes, whereas loop extrusion cannot form interchromosomal links. Our data are therefore less consistent with a model where superenhancer loops form by extrusion and more consistent with the presence of an alternative mechanism, perhaps based on some form of facilitated diffusion or phase separation (Sanborn et al., 2015).

In particular, superenhancer links may represent compartmental co-segregation of small, H3K27-acetylated intervals (Figure 5G), which accounts for why these links can join loci on different chromosomes, why they are weaker in the presence of cohesin, and why their anchors form large cliques and higher-order hubs.

Notably, loops and links between superenhancers increase in strength rapidly following cohesin loss, reaching a plateau within hours. This implies that compartmentalization is capable of inducing intrachromosomal loops and interchromosomal links at rapid rates, comparable to those of loop-domain formation. (Our findings may be related to those of other studies, which have noted enhanced higher-order intrachromosomal interactions between domains containing superenhancers [Beagrie et al., 2017].)

The Interplay between Loop Extrusion and Compartmentalization

Using our high-resolution contact maps, we are also able to examine the ways in which loop domain formation and compartmentalization interact. It is commonly thought that compartment intervals are typically megabases in length and are subdivided into smaller domains in a hierarchical fashion (Dixon et al.,

2012; Nora et al., 2012). Here, we demonstrate that compartment intervals can be as short as tens of kilobases and can overlap loop domains in complex ways. For instance, we observe numerous examples of loop domains spanning multiple compartment intervals.

In such cases, we find that loop extrusion, by facilitating contacts between all loci in the loop domain, can enhance the contact frequency of loci that would ordinarily lie in different subcompartments. Thus, the long-range contact pattern seen for each locus is a mixture of the pattern that would ordinarily be seen for loci in the corresponding subcompartment and the pattern seen for other loci in the loop. This mixing disappears upon cohesin depletion (Figure 7E). Similarly, deletion of WAPL appears to increase the processivity of the Xcom and thus increases the size of loops (Haarhuis et al., 2017). Consistent with our observations, these larger loops are associated with extensive mixing, which obscures long-range compartment patterns.

The Interplay between Cohesin and Gene Regulation

Many studies have proposed that cohesin facilitates interactions between enhancers and promoters, thereby upregulating the transcription of many genes (Kagey et al., 2010; Merkenschlager and Nora, 2016). Moreover, studies have also suggested that loop domains formed between CTCF and cohesin binding sites create insulated regulatory neighborhoods—partially protecting genes with a loop domain from the influence of enhancers outside the domains (Flavahan et al., 2016; Lupiáñez et al., 2015). Our study, combining rapid depletion of cohesin and measurement of nascent transcription using PRO-seq, allows us to more clearly dissect the direct effects of cohesin loss on transcription.

We find that a very small set of genes, often lying near superenhancers, becomes strongly down-regulated after cohesin loss. However, most genes are not strongly affected. This suggests that cohesin-dependent loop domains themselves play at most a modest role in facilitating or disrupting interactions between promoters and enhancers. Of course, we cannot dismiss modest effects on overall levels of transcription as unimportant. For example, modest increases in the expression of receptor tyrosine kinase genes can have meaningful effects on cell proliferation (Flavahan et al., 2016).

Nevertheless, it is particularly interesting to compare the above findings with earlier studies, using similar methods, that showed a strong correlation between the presence of a loop domain and many-fold increases in the expression of genes at the loop domain anchor (Kagey et al., 2010; Rao et al., 2014). Taken in isolation, these earlier results are consistent with a model where the formation of loop domains routinely causes many-fold changes in gene expression. By contrast, the data presented here are more consistent with a model where the formation of loop domains influences gene transcription, but rarely causes many-fold changes.

One possibility is that both processes might be independent consequences of upstream regulatory events. For example, changes in the accessibility of a gene promoter might facilitate both transcription factor binding, activating the gene, and cohesin arrest, activating the loop domain. Alternatively, large increases in gene expression may alter the accessibility of CTCF motifs at the promoter and cause loop domain formation.

Our study suggests a model where cohesin-associated looping, by increasing the frequency of contact between loci within loop domains and by disturbing patterns of compartmentalization, facilitates mixing between elements (such as genes and superenhancers) that would otherwise be segregated. Thus, compartmentalization and extrusion—through independent and complementary mechanisms—interact to shape transcription.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - HCT-116 cells
- METHOD DETAILS
  - Microscopy
  - *In situ* Hi-C
  - ChIP-Seq
  - PRO-Seq
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Hi-C Data Processing
  - ChIP-Seq Data Processing
  - PRO-Seq Data Processing
  - Random Shuffle Annotations
  - Analysis of CTCF and cohesin binding
  - Analysis of histone modification patterns
  - Evaluation of loops and loop domains
  - Analysis of previous cohesin-depletion Hi-C data sets
  - Evaluation of the dynamics of loop domain formation
  - Evaluation of genome compartmentalization
  - Annotation and analysis of cohesin-independent links
  - Analysis of higher order contacts
  - Simulations of extrusion and compartmentalization
  - Oligonucleosome Resolution Simulations of Chromatin Fibers
  - Assessment of changes in transcription after cohesin loss
- DATA AND SOFTWARE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information includes seven figures, one table, and one data file and can be found with this article online at https://doi.org/10.1016/j.cell.2017.09.026.

AUTHOR CONTRIBUTIONS

S.S.P.R. and E.L.A. conceived this project and designed all experiments. S.-C.H., B.G.S.H., and A.D.O. performed Hi-C experiments. J.M.E. aided with the design of the PRO-seq experiments, and J.M.E. and E.M.P. performed PRO-seq experiments. K.-R.K.-K., S.E.J., and S.-C.H. performed ChIP-seq experiments. A.L.S. performed extrusion and compartmentalization simulations. G.D.B. and T.S. designed and analyzed nucleosome simulations, which were performed by G.D.B. I.D.B. performed microscopy experiments.

D.T., Z.Y., and J.T.R. designed a system for creating interactive figures. S.S.P.R., J.M.E., X.H., M.S.S., J.S., B.E.B., R.C., E.S.L., and E.L.A. analyzed data. S.S.P.R., E.S.L., and E.L.A. prepared the manuscript with input from all authors.

ACKNOWLEDGMENTS

This work was supported by a Paul and Daisy Soros Fellowship, a Fannie and John Hertz Foundation Fellowship, a Cornelia de Lange Syndrome Foundation grant, and a Stanford Medical Scholars Fellowship to S.S.P.R.; NIGMS award R01GM055164 to T.S.; and an NIH New Innovator Award (1DP2OD008540-01), an NSF Physics Frontier Center Grant (PHY-1427654, Center for Theoretical Biological Physics), the NHGRI Center for Excellence for Genomic Sciences (HG006193), the Welch Foundation (Q-1866), an NVIDIA Research Center Award, an IBM University Challenge Award, a Google Research Award, a Cancer Prevention Research Institute of Texas Scholar Award (R1304), a McNair Medical Institute Scholar Award, an NIH 4D Nucleome Grant (U01HL130010), an NIH Encyclopedia of DNA Elements Mapping Center Award (UM1HG009375), and the President's Early Career Award in Science and Engineering (4DP2OD008540) to E.L.A. B.E.B. holds equity in Fulcrum Therapeutics. We thank Masato Kanemaki for sharing the HCT-116 RAD21-mAID-mClover cell line; Roger Kornberg, Olga Dudchenko, Peter Geiduschek, and Miriam Huntley for their thoughtful comments on the manuscript; Neva Durand for assistance with computation; Fabio Stossi and Hannah Johnson for assistance with microscopy; Sigrid Knemeyer, Nathaniel Musial, and Lynn Zhu for assistance with figures; and Novogene for assistance with sequencing. The experiments in this study were informed by discussions with E. Nora regarding his experiments on domain structure after CTCF degradation, and we thank him and his collaborators. All contact maps reported here can be explored interactively via Juicebox at http://www.aidenlab.org/juicebox/.

Received: May 15, 2017
Revised: August 2, 2017
Accepted: September 18, 2017
Published: October 5, 2017

REFERENCES

Alipour, E., and Marko, J.F. (2012). Self-organization of domain structures by DNA-loop-extruding enzymes. Nucleic Acids Res. *40*, 11202–11212.

Anderson, J.A., Lorenz, C.D., and Travesset, A. (2008). General purpose molecular dynamics simulations fully implemented on graphics processing units. J. Comput. Phys. *227*, 5342–5359.

Bascom, G., and Schlick, T. (2017). Linking chromatin fibers to gene folding by hierarchical looping. Biophys. J. *112*, 434–445.

Bascom, G.D., Kim, T., and Schlick, T. (2017). Kilobase pair chromatin fiber contacts promoted by living-system-like DNA linker length distributions and nucleosome depletion. J. Phys. Chem. B *121*, 3882–3894.

Beagrie, R.A., Scialdone, A., Schueler, M., Kraemer, D.C., Chotalia, M., Xie, S.Q., Barbieri, M., de Santiago, I., Lavitas, L.-M.M., Branco, M.R., et al. (2017). Complex multi-enhancer contacts captured by genome architecture mapping. Nature *543*, 519–524.

Collepardo-Guevara, R., Portella, G., Vendruscolo, M., Frenkel, D., Schlick, T., and Orozco, M. (2015). Chromatin unfolding by epigenetic modifications explained by dramatic impairment of internucleosome interactions: a multiscale computational study. J. Amer. Chem. Soc. *137*, 10205–10215.

Darrow, E.M., Huntley, M.H., Dudchenko, O., Stamenova, E.K., Durand, N.C., Sun, Z., Huang, S.-C.C., Sanborn, A.L., Machol, I., Shamim, M., et al. (2016). Deletion of DXZ4 on the human inactive X chromosome alters higher-order genome architecture. Proc. Natl. Acad. Sci. USA *113*, E4504–E4512.

Davidson, I.F., Goetz, D., Zaczek, M.P., Molodtsov, M.I., Huis In 't Veld, P.J., Weissmann, F., Litos, G., Cisneros, D.A., Ocampo-Hafalla, M., Ladurner, R., et al. (2016). Rapid movement and transcriptional re-localization of human cohesin on DNA. EMBO J. *35*, 2671–2685.

de Wit, E., Vos, E.S., Holwerda, S.J., Valdes-Quezada, C., Verstegen, M.J., Teunissen, H., Splinter, E., Wijchers, P.J., Krijger, P.H., and de Laat, W. (2015). CTCF Binding Polarity Determines Chromatin Looping. Mol. Cell *60*, 676–684.

Di Pierro, M., Zhang, B., Aiden, E.L., Wolynes, P.G., and Onuchic, J.N.N. (2016). Transferable model for chromosome architecture. Proc. Natl. Acad. Sci. USA *113*, 12168–12173.

Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature *485*, 376–380.

Durand, N.C., Robinson, J.T., Shamim, M.S., Machol, I., Mesirov, J.P., Lander, E.S., and Aiden, E.L. (2016a). Juicebox provides a visualization system for Hi-C contact maps with unlimited zoom. Cell Syst. *3*, 99–101.

Durand, N.C., Shamim, M.S., Machol, I., Rao, S.S., Huntley, M.H., Lander, E.S., and Aiden, E.L. (2016b). Juicer provides a one-click system for analyzing loop-resolution Hi-C experiments. Cell Syst. *3*, 95–98.

ENCODE Project Consortium (2012). An integrated encyclopedia of DNA elements in the human genome. Nature *489*, 57–74.

Engreitz, J.M., Haines, J.E., Perez, E.M., Munson, G., Chen, J., Kane, M., McDonel, P.E., Guttman, M., and Lander, E.S. (2016). Local regulation of gene expression by lncRNA promoters, transcription and splicing. Nature *539*, 452–455.

Flavahan, W.A., Drier, Y., Liau, B.B., Gillespie, S.M., Venteicher, A.S., Stemmer-Rachamimov, A.O., Suvà, M.L., and Bernstein, B.E. (2016). Insulator dysfunction and oncogene activation in IDH mutant gliomas. Nature *529*, 110–114.

Fudenberg, G., Imakaev, M., Lu, C., Goloborodko, A., Abdennur, N., and Mirny, L.A. (2016). Formation of chromosomal domains by loop extrusion. Cell Rep. *15*, 2038–2049.

Glaser, J., Nguyen, T.D., Anderson, J.A., and Lui, P. (2015). Strong scaling of general-purpose molecular dynamics simulations on GPUs. Comput. Phys. *192*, 97–107.

Grigoryev, S.A., Bascom, G., Buckwalter, J.M., Schubert, M.B., Woodcock, C.L., and Schlick, T. (2016). Hierarchical looping of zigzag nucleosome chains in metaphase chromosomes. Proc. Natl. Acad. Sci. USA *113*, 1238–1243.

Guo, Y., Xu, Q., Canzio, D., Shou, J., Li, J., Gorkin, D.U., Jung, I., Wu, H., Zhai, Y., Tang, Y., et al. (2015). CRISPR inversion of CTCF sites alters genome topology and enhancer/promoter function. Cell *162*, 900–910.

Haarhuis, J.H.I., van der Weide, R.H., Blomen, V.A., Yáñez-Cuna, J.O., Amendola, M., van Ruiten, M.S., Krijger, P.H.L., Teunissen, H., Medema, R.H.H., van Steensel, B., et al. (2017). The cohesin release factor WAPL restricts chromatin loop extension. Cell *169*, 693–707.e14.

Hnisz, D., Abraham, B.J., Lee, T.I., Lau, A., Saint-André, V., Sigova, A.A., Hoke, H.A., and Young, R.A. (2013). Super-enhancers in the control of cell identity and disease. Cell *155*, 934–947.

Hnisz, D., Shrinivas, K., Young, R.A., Chakraborty, A.K., and Sharp, P.A. (2017). A phase separation model for transcriptional control. Cell *169*, 13–23.

Jonkers, I., and Lis, J.T. (2015). Getting up to speed with transcription elongation by RNA polymerase II. Nat. Rev. Mol. Cell Biol. *16*, 167–177.

Jost, D., Carrivain, P., Cavalli, G., and Vaillant, C. (2014). Modeling epigenome folding: formation and dynamics of topologically associated chromatin domains. Nucleic Acids Res. *42*, 9553–9561.

Kagey, M.H., Newman, J.J., Bilodeau, S., Zhan, Y., Orlando, D.A., van Berkum, N.L., Ebmeier, C.C., Goossens, J., Rahl, P.B., Levine, S.S., et al. (2010). Mediator and cohesin connect gene expression and chromatin architecture. Nature *467*, 430–435.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359.

Larson, A.G., Elnatan, D., Keenen, M.M., Trnka, M.J., Johnston, J.B., Burlingame, A.L., Agard, D.A., Redding, S., and Narlikar, G.J. (2017). Liquid droplet

formation by HP1α suggests a role for phase separation in heterochromatin. Nature *547*, 236–240.

Li, H., and Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics *26*, 589–595.

Lieberman-Aiden, E., van Berkum, N.L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O., et al. (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science *326*, 289–293.

Liu, T. (2014). Use model-based Analysis of ChIP-Seq (MACS) to analyze short reads generated by sequencing protein-DNA interactions in embryonic stem cells. Methods Mol. Biol. *1150*, 81–95.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550.

Lupiáñez, D.G.G., Kraft, K., Heinrich, V., Krawitz, P., Brancati, F., Klopocki, E., Horn, D., Kayserili, H., Opitz, J.M., Laxova, R., et al. (2015). Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. Cell *161*, 1012–1025.

Merkenschlager, M., and Nora, E.P.P. (2016). CTCF and cohesin in genome folding and transcriptional gene regulation. Annu. Rev. Genomics Hum. Genet. *17*, 17–43.

Nasmyth, K. (2001). Disseminating the genome: joining, resolving, and separating sister chromatids during mitosis and meiosis. Annu. Rev. Genet. *35*, 673–745.

Natsume, T., Kiyomitsu, T., Saga, Y., and Kanemaki, M.T. (2016). Rapid protein depletion in human cells by auxin-inducible degron tagging with short homology donors. Cell Rep. *15*, 210–218.

Nichols, M.H., and Corces, V.G. (2015). A CTCF Code for 3D Genome Architecture. Cell *162*, 703–705.

Nora, E.P.P., Lajoie, B.R., Schulz, E.G., Giorgetti, L., Okamoto, I., Servant, N., Piolot, T., van Berkum, N.L., Meisig, J., Sedat, J., et al. (2012). Spatial partitioning of the regulatory landscape of the X-inactivation centre. Nature *485*, 381–385.

Nora, E.P.P., Goloborodko, A., Valton, A.-L.L., Gibcus, J.H., Uebersohn, A., Abdennur, N., Dekker, J., Mirny, L.A., and Bruneau, B.G. (2017). Targeted degradation of CTCF decouples local insulation of chromosome domains from genomic compartmentalization. Cell *169*, 930–944.e22.

Parker, S.C., Stitzel, M.L., Taylor, D.L., Orozco, J.M., Erdos, M.R., Akiyama, J.A., van Bueren, K.L., Chines, P.S., Narisu, N., Black, B.L., et al.; NISC Comparative Sequencing Program; National Institutes of Health Intramural Sequencing Center Comparative Sequencing Program Authors; NISC Comparative Sequencing Program Authors (2013). Chromatin stretch enhancer states drive cell-specific gene regulation and harbor human disease risk variants. Proc. Natl. Acad. Sci. USA *110*, 17921–17926.

Rao, S.S., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S., and Aiden, E.L. (2014). A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell *159*, 1665–1680.

Sanborn, A.L., Rao, S.S., Huang, S.-C.C., Durand, N.C., Huntley, M.H., Jewett, A.I., Bochkov, I.D., Chinnappan, D., Cutkosky, A., Li, J., et al. (2015). Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. Proc. Natl. Acad. Sci. USA *112*, E6456–E6465.

Schwarzer, W., Abdennur, N., Goloborodko, A., Pekowska, A., Fudenberg, G., Loe-Mie, Y., Fonseca, N.A., Huber, W., Haering, C., Mirny, L., et al. (2016). Two independent modes of chromosome organization are revealed by cohesin removal. bioRxiv. https://doi.org/10.1101/094185.

Seitan, V.C., Faure, A.J., Zhan, Y., McCord, R.P., Lajoie, B.R., Ing-Simmons, E., Lenhard, B., Giorgetti, L., Heard, E., Fisher, A.G., et al. (2013). Cohesin-based chromatin interactions enable regulated gene expression within preexisting architectural compartments. Genome Res. *23*, 2066–2077.

Sofueva, S., Yaffe, E., Chan, W.-C.C., Georgopoulou, D., Vietri Rudan, M., Mira-Bontenbal, H., Pollard, S.M., Schroth, G.P., Tanay, A., and Hadjur, S. (2013). Cohesin-mediated interactions organize chromosomal domain architecture. EMBO J. *32*, 3119–3129.

Song, Q., and Smith, A.D. (2011). Identifying dispersed epigenomic domains from ChIP-Seq data. Bioinformatics *27*, 870–871.

Splinter, E., Heath, H., Kooren, J., Palstra, R.-J., Klous, P., Grosveld, F., Galjart, N., and de Laat, W. (2006). CTCF mediates long-range chromatin looping and local histone modification in the beta-globin locus. Genes Dev. *20*, 2349–2354.

Stigler, J., Çamdere, G.Ö., Koshland, D.E., and Greene, E.C. (2016). Single-molecule imaging reveals a collapsed conformational state for DNA-bound cohesin. Cell Rep. *15*, 988–998.

Strom, A.R., Emelyanov, A.V., Mir, M., Fyodorov, D.V., Darzacq, X., and Karpen, G.H. (2017). Phase separation drives heterochromatin domain formation. Nature *547*, 241–245.

Wang, X., Brandão, H.B., Le, T.B., Laub, M.T., and Rudner, D.Z. (2017). Bacillus subtilis SMC complexes juxtapose chromosome arms as they travel from origin to terminus. Science *355*, 524–527.

Wendt, K.S., Yoshida, K., Itoh, T., Bando, M., Koch, B., Schirghuber, E., Tsutsumi, S., Nagae, G., Ishihara, K., Mishiro, T., et al. (2008). Cohesin mediates transcriptional insulation by CCCTC-binding factor. Nature *451*, 796–801.

Wijchers, P.J., Krijger, P.H.L., Geeven, G., Zhu, Y., Denker, A., Verstegen, M.J.A.M., Valdes-Quezada, C., Vermeulen, C., Janssen, M., Teunissen, H., et al. (2016). Cause and consequence of tethering a SubTAD to different nuclear compartments. Mol. Cell *61*, 461–473.

Zuin, J., Dixon, J.R., van der Reijden, M.I., Ye, Z., Kolovos, P., Brouwer, R.W.W., van de Corput, M.P.P., van de Werken, H.J., Knoch, T.A., van IJcken, W.F., et al. (2014). Cohesin and CTCF differentially affect chromatin architecture and gene expression in human cells. Proc. Natl. Acad. Sci. USA *111*, 996–1001.