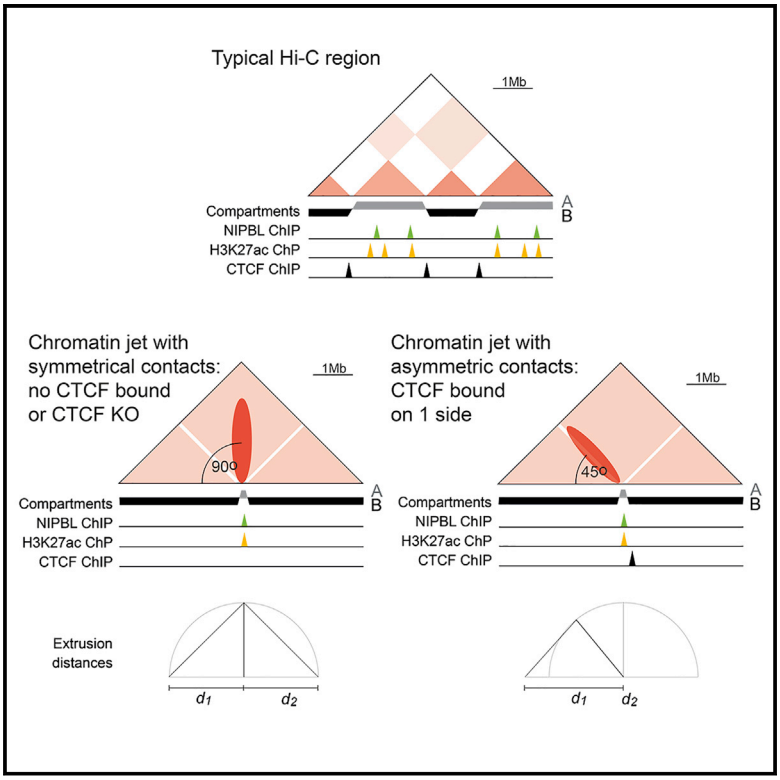
Molecular Cell

Article

Chromatin jets define the properties of cohesin-driven *in vivo* loop extrusion

Graphical abstract



Highlights

- Cohesin-powered chromatin jets emerge from focal areas of accessible chromatin
- Unless constrained by CTCF, jets can propagate symmetrically for 1–2 Mb
- CTCF can block jets or deflect the angle of jet propagation
- These data suggest that loop extrusion is controlled independently in both directions

Authors

Ya Guo, Ediem Al-Jibury, Rosalba Garcia-Millan, ..., Amanda G. Fisher, Gunnar Pruessner, Matthias Merkenschlager

Correspondence

g.pruessner@imperial.ac.uk (G.P.), matthias.merkenschlager@lms.mrc.ac.uk (M.M.)

In brief

Cohesin organizes the genome in 3D nuclear space. In this issue of *Molecular Cell*, Guo et al. describe and characterize "jets," an unusual form of cohesin-mediated chromatin interactions. Jets offer insights into the physiological behavior of cohesin-mediated loop extrusion and the principles that underlie genome organization.

Guo et al., 2022, Molecular Cell *82*, 3769–3780
October 20, 2022 © 2022 The Authors. Published by Elsevier Inc.
https://doi.org/10.1016/j.molcel.2022.09.003

CellPress

Molecular Cell

CellPress
OPEN ACCESS

Article

Chromatin jets define the properties of cohesin-driven *in vivo* loop extrusion

Ya Guo,[1,2,3,12] Ediem Al-Jibury,[1,4,12] Rosalba Garcia-Millan,[5,6,7,12] Konstantinos Ntagiantas,[5,12] James W.D. King,[1] Alex J. Nash,[1] Niels Galjart,[8] Boris Lenhard,[1,9] Daniel Rueckert,[4,10] Amanda G. Fisher,[1] Gunnar Pruessner,[5,11,*] and Matthias Merkenschlager[1,11,13,*]

[1]MRC London Institute of Medical Sciences, Institute of Clinical Sciences, Faculty of Medicine, Imperial College London, London W12 0NN, UK
[2]Sheng Yushou Center of Cell Biology and Immunology, Joint International Research Laboratory of Metabolic and Developmental Sciences, School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai 200240, China
[3]WLA Laboratories, Shanghai 201203, China
[4]Department of Computing, Imperial College London, London SW7 2RH, UK
[5]Department of Mathematics, Imperial College London, London SW7 2RH, UK
[6]Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge CB3 0WA, UK
[7]St John's College, University of Cambridge, Cambridge CB2 1TP, UK
[8]Department of Cell Biology, Erasmus University Medical Center, 3015 GD Rotterdam, the Netherlands
[9]Sars International Centre for Marine Molecular Biology, University of Bergen, 5008 Bergen, Norway
[10]Chair for AI in Healthcare and Medicine, Technical University of Munich, 81675 Munich, Germany
[11]Senior author
[12]These authors contributed equally
[13]Lead contact
*Correspondence: g.pruessner@imperial.ac.uk (G.P.), matthias.merkenschlager@lms.mrc.ac.uk (M.M.)
https://doi.org/10.1016/j.molcel.2022.09.003

SUMMARY

Complex genomes show intricate organization in three-dimensional (3D) nuclear space. Current models posit that cohesin extrudes loops to form self-interacting domains delimited by the DNA binding protein CTCF. Here, we describe and quantitatively characterize cohesin-propelled, jet-like chromatin contacts as landmarks of loop extrusion in quiescent mammalian lymphocytes. Experimental observations and polymer simulations indicate that narrow origins of loop extrusion favor jet formation. Unless constrained by CTCF, jets propagate symmetrically for 1–2 Mb, providing an estimate for the range of *in vivo* loop extrusion. Asymmetric CTCF binding deflects the angle of jet propagation as experimental evidence that cohesin-mediated loop extrusion can switch from bi- to unidirectional and is controlled independently in both directions. These data offer new insights into the physiological behavior of *in vivo* cohesin-mediated loop extrusion and further our understanding of the principles that underlie genome organization.

INTRODUCTION

The interphase genome is organized by the separation of chromatin states into active (A) and inactive (B) compartments and the formation of TADs, contact domains, and CTCF-based chromatin loops by cohesin and CTCF (Rao et al., 2014; Nora et al., 2017; Schwarzer et al., 2017; Dekker and Mirny, 2016; Merkenschlager and Nora, 2016; Mirny et al., 2019). Chromatin loops form preferentially between CTCF molecules bound to DNA motifs in convergent orientation, suggesting that loop formation is based on one-dimensional (1D) tracking rather than three-dimensional (3D) diffusion (Rao et al., 2014; Vietri-Rudan et al., 2015; de Wit et al., 2015) and consistent with models of loop extrusion mediated by the structural maintenance of chromosomes (SMCs) protein complexes (Riggs et al., 1990; Kimura et al., 1999; Nasmyth, 2001; Alipour and Marko, 2012; Sanborn et al., 2015; Fudenberg et al., 2016, 2017; Banigan et al., 2022). SMC complexes translocate on DNA and can extrude DNA loops at rates of approximately 1 kb s^{-1} (Gruber and Errington, 2009; Ganji et al., 2018; Golfier et al., 2020; Kong et al., 2020; Terakawa et al., 2017; Tran et al., 2017; Wang et al., 2017; Wang et al., 2018; Anchimiuk et al., 2021; Davidson et al., 2019; Kim et al., 2019). Structural studies provide evidence for a conserved cohesin interface with the N terminus of CTCF (Li et al., 2020; Nora et al., 2020; Pugacheva et al., 2020). It is thought that a process akin to loop extrusion underlies the formation of contact domains and CTCF-based loops: (1) loop domains reform rapidly after transient cohesin depletion, providing a minimal estimate of the speed of cohesin-mediated chromatin contact propagation (Rao et al., 2017). (2) Hi-C "stripes" connect CTCF domain

This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

anchors with multiple points within domains, highly suggestive of loop extrusion (Vian et al., 2018). (3) Loops become longer when the residence time of cohesin on chromatin is increased by the removal of the cohesin unloading factor WAPL (Haarhuis et al., 2017; Wutz et al., 2017; Liu et al., 2021a).

During DNA replication in bacteria, SMC complexes are loaded from predefined sites, providing evidence that individual SMC complexes can align entire chromosome arms comprising several million base pairs of DNA (Marbouty et al., 2015; Tran et al., 2017; Wang et al., 2017, 2018; reviewed by Banigan and Mirny, 2020). In contrast to bacteria, loading sites for SMC complexes are not well defined in mammalian cells. Active regulatory elements are enriched for cohesin and the cohesin component and loading factor NIPBL. ATAC-seq accessibility and the acetylation of histone H3 on lysine 27 (H3K27ac) are part of the chromatin signature of active regulatory elements and therefore considered as chromatin features of potential cohesin loading sites (Vian et al., 2018). Biochemical, functional, and genetic interactions of NIPBL/cohesin with H3K27ac-associated BET proteins (Olley et al., 2018; Luna-Peláez et al., 2019; Linares-Saldana et al., 2021), the enhancer-associated MLL3/4 complex (Yan et al., 2018), and the components of the transcriptional machinery (Kagey et al., 2010; Liu et al., 2021b; van den Berg et al., 2017) implicate enhancers and super-enhancers as candidate loading sites for cohesin within chromatin domains and TADs (Vian et al., 2018; Kagey et al., 2010). However, in the absence of direct assays for cohesin loading in living cells, the identity of *in vivo* cohesin loading sites remains uncertain. Moreover, chromatin regions that give rise to domains, loops, and stripes in mammalian genomes typically contain multiple potential cohesin loading sites. Therefore, it is unclear whether cohesin-dependent contacts are established by the continuous action of individual cohesin complexes or by a succession of multiple cohesin complexes.

Here, we describe cohesin-dependent chromatin contacts that originate from small, isolated sites of accessible chromatin and traverse adjacent B compartment regions in a jet-like fashion in primary mammalian cells. Jet origins feature local chromatin accessibility, NIPBL, cohesin, and H3K27ac. Our experimental and computational modeling suggest that lone accessible sites favor focal cohesin loading and chromatin jet formation, whereas broad or multiple accessible sites drive diffuse cohesin loading and favor the formation of contact domains (Fudenberg et al., 2016; Fudenberg et al., 2017; Banigan et al., 2022). Once initiated, jets can propagate symmetrically for 1–2 Mb. Unilateral CTCF encounters can convert bi- to unidirectional extrusion, indicating that both directions of extrusion are controlled independently. These data provide new insights into the physiological behavior of *in vivo* cohesin-mediated loop extrusion in unperturbed mammalian cells.

RESULTS

To elucidate the properties of cohesin-mediated chromatin contact propagation *in vivo*, we performed *in situ* Hi-C on wild-type, $Ctcf^{-/-}$, and $Rad21^{-/-}$ $Ctcf^{-/-}$ primary DP thymocytes (Figure S1; Table S1). The efficiency of CTCF depletion was 80.0% ± 10.1% (n = 7), and the efficiency of RAD21 depletion was 83.3% ± 4.2% (n = 7; Figure S1C). We chose small DP thymocytes because they are quiescent and therefore require cohesin and CTCF for 3D genome organization in interphase but not for cell cycle-related functions such as DNA replication or chromosome segregation (Seitan et al., 2011; Nasmyth and Haering, 2009). We examined the resulting Hi-C maps for features consistent with loop extrusion. In addition to the familiar Hi-C patterns of ubiquitous pleat-like compartments and diamond-shaped domains (see Figure 1A for a schematic and illustrative example, Figure S1 for quantification), visual inspection revealed a smaller number of jet-like projections from the Hi-C diagonal (see Figure 1B for a schematic and illustrative example, Table S2 for quantification), reminiscent of Hi-C patterns that are formed by aligned chromosome arms in bacteria (Tran et al., 2017; Wang et al., 2017, 2018; reviewed by Banigan and Mirny, 2020). The examination of published Hi-C data (Kieffer-Kwon et al., 2017) shows that similar jets are present in B cells (Figure S2).

The majority of jets originated from small regions denoted as open chromatin, i.e., A compartment intervals based on Hi-C eigenvector. Jets projected into surrounding closed chromatin regions (B compartment intervals, Figure 1C; Table S2). We identified a set of n = 38 jets and developed a protractor tool to quantify jet strength and jet angles in individual Hi-C replicates (see method details section; Figures 1D and 1E; Table S2). Protractor scanning confirmed increased chromatin contacts around small A compartment intervals that contained visually identified jets compared with small A compartment intervals that lacked jets, including regions that contained contact domains, or neither jets nor contact domains (Figure 1F; Table S2). The propagation of chromatin jets perpendicular to the Hi-C diagonal is consistent with bidirectional loop extrusion *in vivo* (Figure 1G).

Jets can be constrained by CTCF and released by CTCF removal

In comparison with wild-type thymocytes (Figure 1), additional jets arise in $Ctcf^{-/-}$ thymocytes (Figures 2A, 2B, and S3; Table S2). Aggregate subtraction plots confirmed that average jet strength increased in the absence of CTCF (Figures 2C and S4). Protractor scan quantification showed that a subset of jets became significantly stronger in CTCF-deficient ($Ctcf^{-/-}$) compared with wild-type DP thymocytes. Integration with CTCF chromatin immunoprecipitation sequencing (ChIP-seq) data showed that this increase in strength was confined to those jets that were targets of CTCF binding in wild-type DP thymocytes (Figure 2D). The loss of CTCF led to a significant decrease in the variance of jet length distributions (Figure 2E). This is consistent with the role of CTCF in blocking cohesin-mediated loop extrusion (Li et al., 2020) and indicates that CTCF is a major determinant of jet propagation. Polymer simulations in which CTCF stalls cohesin reproduced the experimentally observed impact of CTCF on the propagation of jets in the presence and the absence of CTCF (Figure 2F; see method details section and Table S3 for parameters). These data indicate that jet propagation can be constrained by CTCF and released by CTCF removal (Figure 2G).

Jets are cohesin dependent

To ask whether jet formation depends on cohesin, we compared Hi-C maps of wild-type, CTCF-deficient ($Ctcf^{-/-}$), and cohesin-deficient ($Rad21^{-/-}$ $Ctcf^{-/-}$) DP thymocytes.

CellPress
OPEN ACCESS

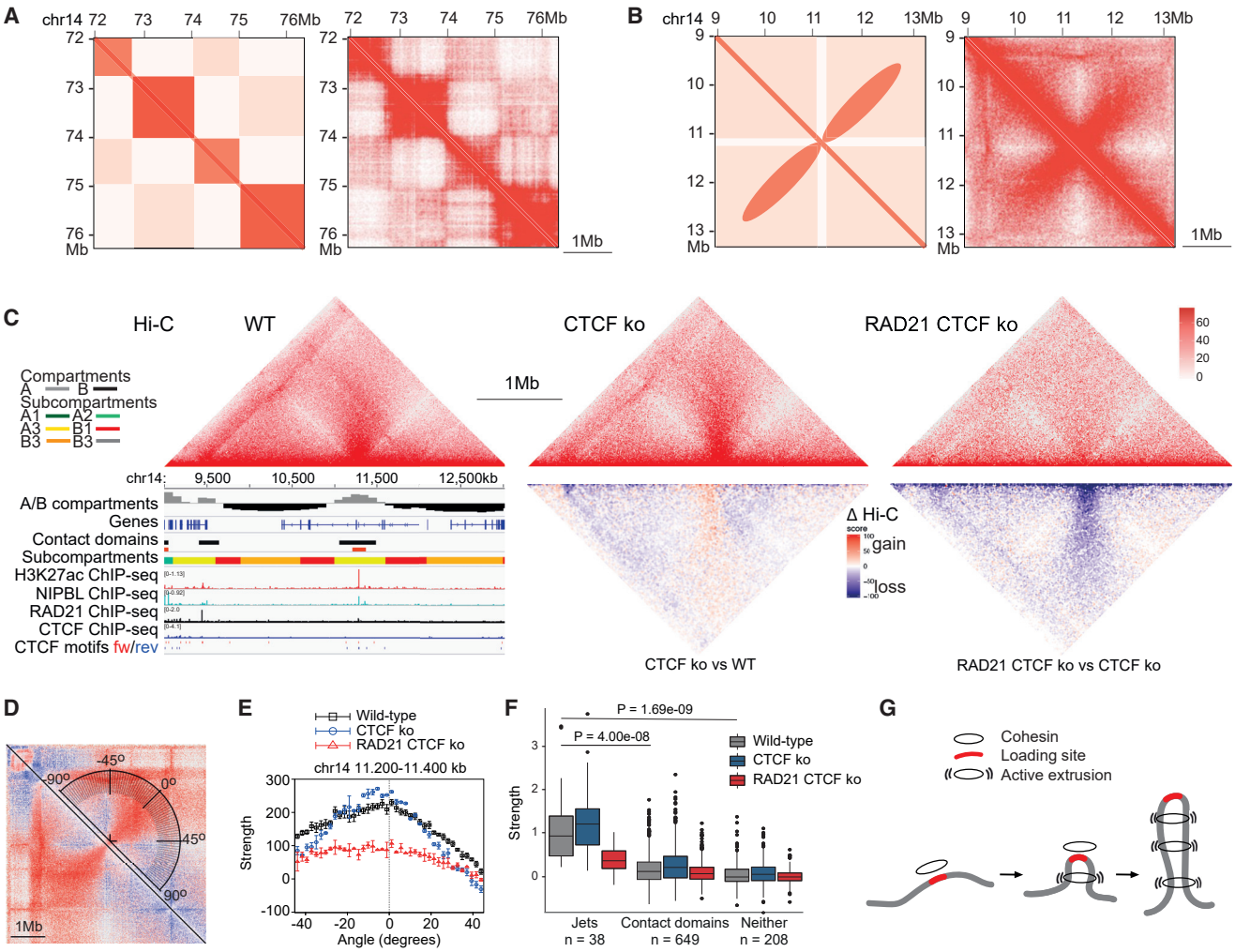


Figure 1. Identification of jets

(A) Schematic (left) and illustrative example (right) of a 4.25-Mb region on chr14 featuring chromatin domains (25-kb resolution).

(B) Schematic (left) and illustrative example (right) of a 4.25-Mb region on chr14 featuring a jet (25-kb resolution).

(C) Example of a jet in wild-type (top left) and *Ctcf*$^{-/-}$ DP thymocytes (top center). *Rad21*$^{-/-}$ *Ctcf*$^{-/-}$ DP double KO (DKO) thymocytes shown for comparison (top right). Genome browser view of chr14 11.25 Mb with A/B compartment eigenvector, subcompartments, contact domains (black 5 + 10 kb, red: 5 kb), ChIP-seq tracks, and directionality of CTCF motifs associated with CTCF ChIP-seq peaks.

(D) Protractor schematic. Observed over expected Hi-C map as background.

(E) Replicate-based protractor scan quantification of jet strength and angle.

(F) Replicate-based protractor scan quantification of contact strength around small A compartments with jets, with contact domains, or without jets and contact domains. p values: two-sided t test.

(G) Schematic representation of symmetric jet propagation.

We observed that jets were substantially weakened by the depletion of cohesin (Figures 1A and 2A; Table S2). The jet remnants visible in Hi-C maps of cohesin-deficient thymocytes are likely due to the imperfect depletion of the RAD21 protein (Figure S1C). The aggregate analysis and subtraction plots confirmed the weakening of jets in the absence of cohesin (Figure 3A). The protractor scan quantification of jet strength of wild-type, CTCF-deficient (*Ctcf*$^{-/-}$), and cohesin-deficient (*Rad21*$^{-/-}$ *Ctcf*$^{-/-}$) DP thymocytes showed that reduced jet strength after cohesin depletion was significant, whether or not CTCF was bound at jet origins (Figure 3B). The analysis of published dilution Hi-C data from wild-type and *Rad21*$^{-/-}$ DP thymocytes (Seitan et al., 2013) confirmed that depletion of cohesin alone was sufficient to weaken jet formation (Figure S5).

Jet origins have chromatin features consistent with cohesin loading sites

Compared with A compartment intervals genome wide, the subset of A compartment-like intervals that give rise to jets was

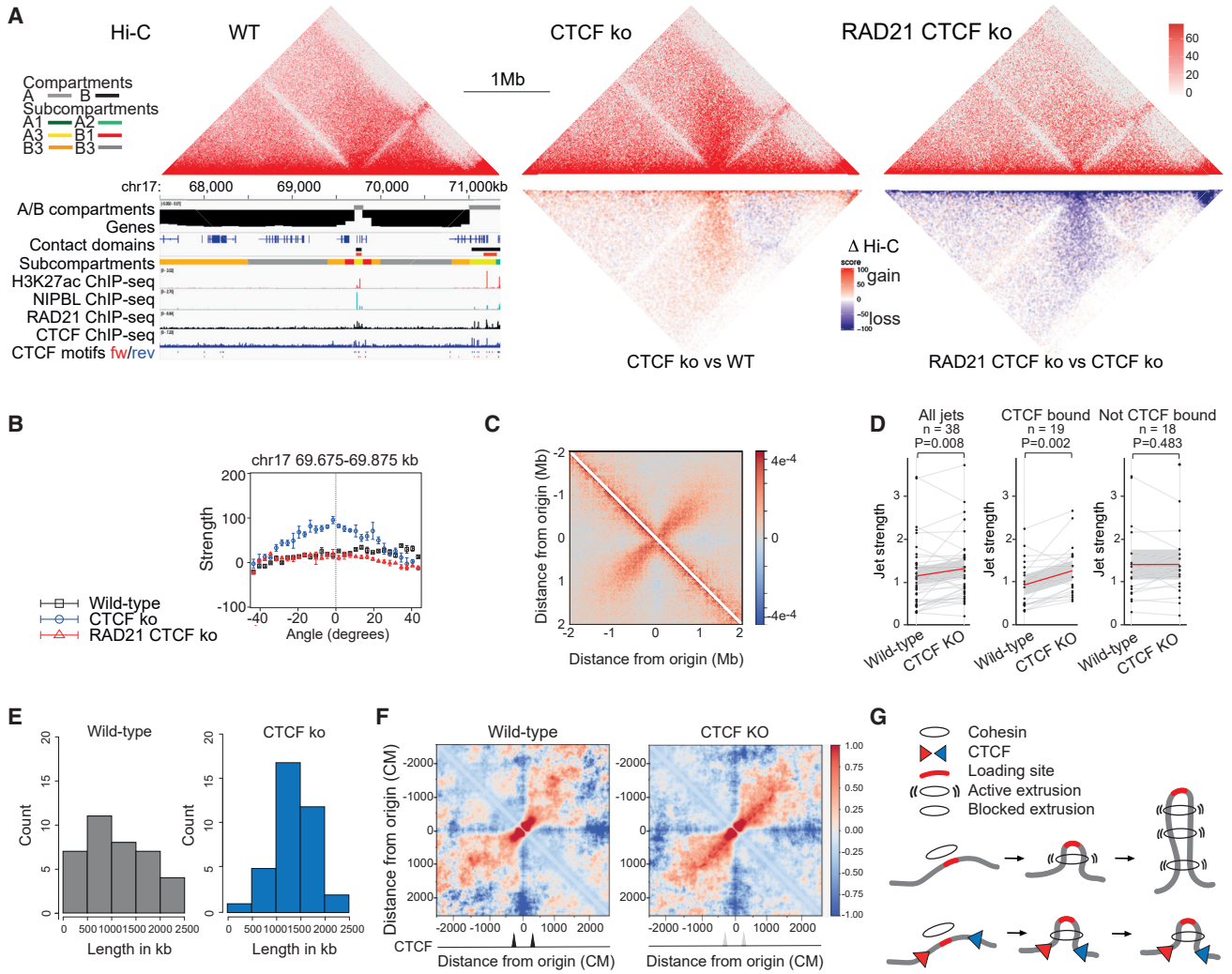


Figure 2. CTCF can constrain jets

(A) Example of a jet that gains in strength in *Ctcf*$^{-/-}$ DP thymocytes (center) compared with wild type (left). *Rad21*$^{-/-}$ *Ctcf*$^{-/-}$ DP thymocytes are shown for comparison (right). Genome browser view of chr17 69.8 Mb with A/B compartment eigenvector, subcompartments, contact domains (black 5 + 10 kb, red: 5 kb), ChIP-seq tracks, and directionality of CTCF motifs associated with CTCF ChIP-seq peaks.

(B) Replicate-based protractor scan quantification of the jet shown in (A).

(C) Mean Hi-C contact subtraction of n = 38 jets in CTCF KO minus wild-type thymocytes. Jet pileups for wild-type and CTCF KO are shown in Figure S4.

(D) Comparison of jet strength in wild-type and *Ctcf*$^{-/-}$ DP thymocytes. Jets overall increased in strength in *Ctcf*$^{-/-}$ compared with wild-type DP thymocytes (CTCF KO, p = 0.006, one-sided t test, n = 38 jets). Jets with at least one CTCF ChIP-seq peak within 200 kb of the jet origin increased in strength in *Ctcf*$^{-/-}$ compared with wild-type DP thymocytes (CTCF KO, p = 0.002, one-sided t test, n = 19 jets, 13/19 gained strength), whereas jets without a CTCF ChIP-seq peak within 200 kb of the jet origin did not (p = 0.483, one-sided t test, n = 18 jets, 6/18 gained strength). The mean (red) and 95% confidence interval are shown (gray).

(E) Comparison of jet reach in wild-type and *Ctcf*$^{-/-}$ DP thymocytes. The jet length shows a narrower distribution in *Ctcf*$^{-/-}$ than in wild-type thymocytes (variance of WT = 427,582.7, variance of CTCF KO = 224,744.8). p = 0.029 by one-sided F test to compare two variances (variance of CTCF KO < variance of wild type).

(F) Simulation of CTCF effects on jet propagation. Polymer simulation of the effect on jet propagation of CTCF binding positioned 200 chromatin monomers (CMs) up- and downstream of the center of the jet origin and loading was restricted to the origin (see method details section and Table S3 for model parameters).

(G) Cartoon representation. In the absence of CTCF, cohesin extrudes symmetrically. CTCF binding on both sides of the origin blocks jet propagation.

distinctly smaller in size (Figure 4A). Among A compartment intervals <1 Mb, jets preferentially originated from smaller A compartment intervals compared with contact domains (Figure 4B). To identify chromatin features of A compartment intervals that form jets, we integrated Hi-C with ATAC-seq, which identifies accessible chromatin, ChIP-seq for the cohesin subunit RAD21 and NIPBL, as well as H3K27ac as a mark for active regulatory elements. H3K27ac-associated BET proteins (Olley et al., 2018; Luna-Peláez et al., 2019; Linares-Saldana et al., 2021), the enhancer-associated MLL3/4 complex (Yan et al., 2018), and components of the transcriptional machinery (Kagey et al., 2010; Liu et al., 2021b) form biochemical, functional, and genetic interactions of NIPBL and cohesin (Olley et al., 2018; Luna-Peláez et al., 2019; Linares-Saldana et al., 2021; Yan

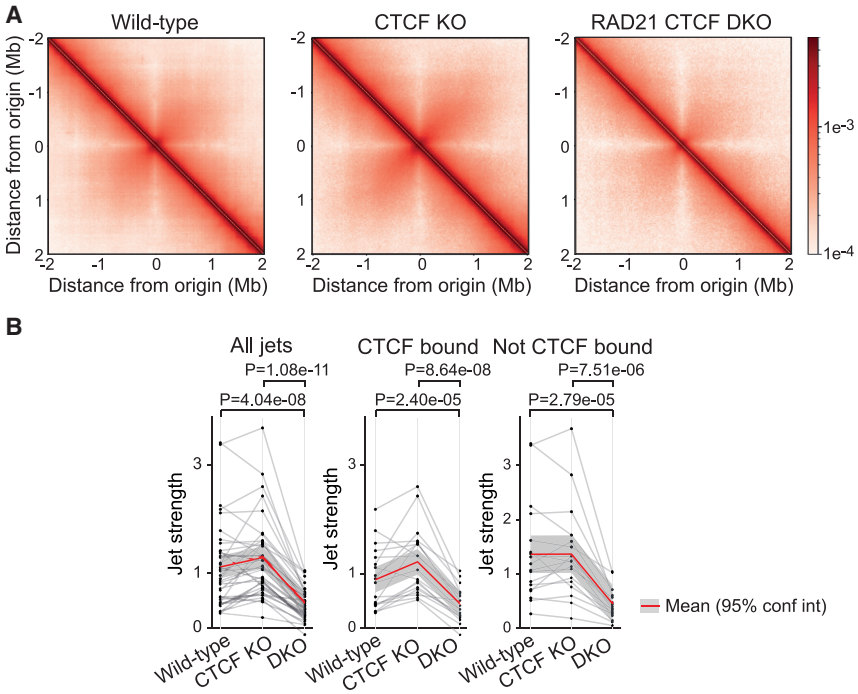


Figure 3. Jets are cohesin dependent

(A) Mean Hi-C contacts for n = 38 jets in wild-type and CTCF KO (as in Figure 2C) with *Rad21*$^{-/-}$ *Ctcf*$^{-/-}$ DP thymocytes (DKO, right). The difference in mean Hi-C signal CTCF KO minus wild type (right). x and y axes are in Mb.

(B) Numerical comparison of jet strength in wild-type and *Ctcf*$^{-/-}$ (as in Figure 2D) with *Rad21*$^{-/-}$ *Ctcf*$^{-/-}$ (DKO) DP thymocytes n = 38. Jets with at least one CTCF ChIP-seq peak within 200 kb of the jet origin (CTCF bound, n = 19) and jets without a CTCF ChIP-seq peak within 200 kb of the jet origin (not CTCF bound, n = 18) were affected by the loss of cohesin (one-sided t test). The mean (red) and 95% confidence interval are shown (gray).

et al., 2018; Kagey et al., 2010; Liu et al., 2021b). A compartment intervals <1 Mb that formed jets or contact domains showed greater ATAC-seq and H3K27ac signal, more RAD21 binding, and a tendency toward more NIPBL binding than A compartment intervals <1 Mb that formed neither jets nor contact domains (Figure 4C, quantification in Figure 4D). Cohesin loading has been associated with active chromatin regions and components of the transcriptional machinery (Kagey et al., 2010; van den Berg et al., 2017; Yan et al., 2018; Olley et al., 2018; Luna-Peláez et al., 2019; Linares-Saldana et al., 2021), and jet origins have chromatin features consistent with these studies.

Isolated focal areas of open chromatin favor jet formation

Jets showed a more focused distribution of ATAC-seq and H3K27ac ChIP-seq signals than contact domains (Figure 5A; quantification in Figure 5B), indicating that jets are more likely to arise from isolated focal areas of open chromatin. We used polymer simulations to model the impact of narrow versus broad cohesin loading areas on the resulting chromatin contacts. In these simulations, 100 monomers represented narrow cohesin loading 2,000 monomers represented broad areas of cohesin loading. All other parameters such as total polymer length, number of cohesins, drift, and unloading probability were kept constant (see simulations 1 and 2 in Table S3). Narrow cohesin loading resulted in jet-like features (Figure 5C, left), whereas broad loading gave rise to contact domains instead (Figure 5C, middle; Fudenberg et al., 2016, 2017). Interestingly, additional loading along the length of the jet disrupted jet formation (Figure 5C, right).

Taken together, the dependence of jets on cohesin and their origin from isolated areas of open chromatin marked by H3K27ac, RAD21, and NIPBL, the preference of jets for narrow origins, and the associated polymer simulations support the notion that jet origins have features of isolated cohesin loading sites. We propose that jets arise from isolated chromatin regions with focal accessibility and focal H3K27ac marks (Figure 5D).

Jets allow estimates of the range of *in vivo* loop extrusion

The observation that jets project from defined origins, are powered by cohesin, and are constrained by CTCF provides a unique opportunity to determine the properties of *in vivo* loop extrusion. We developed a stencil tool to quantify the range of jet propagation (see method details section, Figure 6A). We found that jets can propagate for ~1–2 Mb in wild-type and CTCF-deficient DP thymocytes (Figure 6B; Table S2) and for similar distances in resting B cells (Figure S2).

Jet angles are modulated by CTCF, providing experimental evidence for one-sided loop extrusion

We next analyzed the impact of CTCF on the angle of jet propagation. Jets that propagate perpendicular to the diagonal in a cohesin-dependent manner illustrate that *in vivo* loop extrusion by cohesin can progress bidirectionally (e.g., Figure 1A; Table S2). However, we found that a subset of jets deviates from the perpendicular (Figures 7A and S6; Table S2). The jet depicted in Figure 7A has a negative projection angle, indicating that this jet is deflected toward upstream sequences in wild-type cells. This jet is flanked by downstream CTCF sites. The jet depicted in Figure S6A has a positive projection angle in wild-type cells, indicating that it is deflected toward downstream sequences in wild-type cells. This jet is flanked by upstream CTCF sites.

In contrast to the jets shown in previous figures, these angled jets are not symmetrical in wild-type cells. Accordingly, the distance traveled by loop extrusion is greater in one direction (Figure 7C; Table S2). Interestingly, the projection angles of these deflected jets change in CTCF-deficient cells, and jets align more closely with the perpendicular (Figures 7A, 7B, S6A, and S6B; Table S2). Consequently, the range of jet angles is broader in wild-type than in CTCF-deficient cells (Figure 7D). We next counted the number of CTCF motifs with CTCF ChIP-

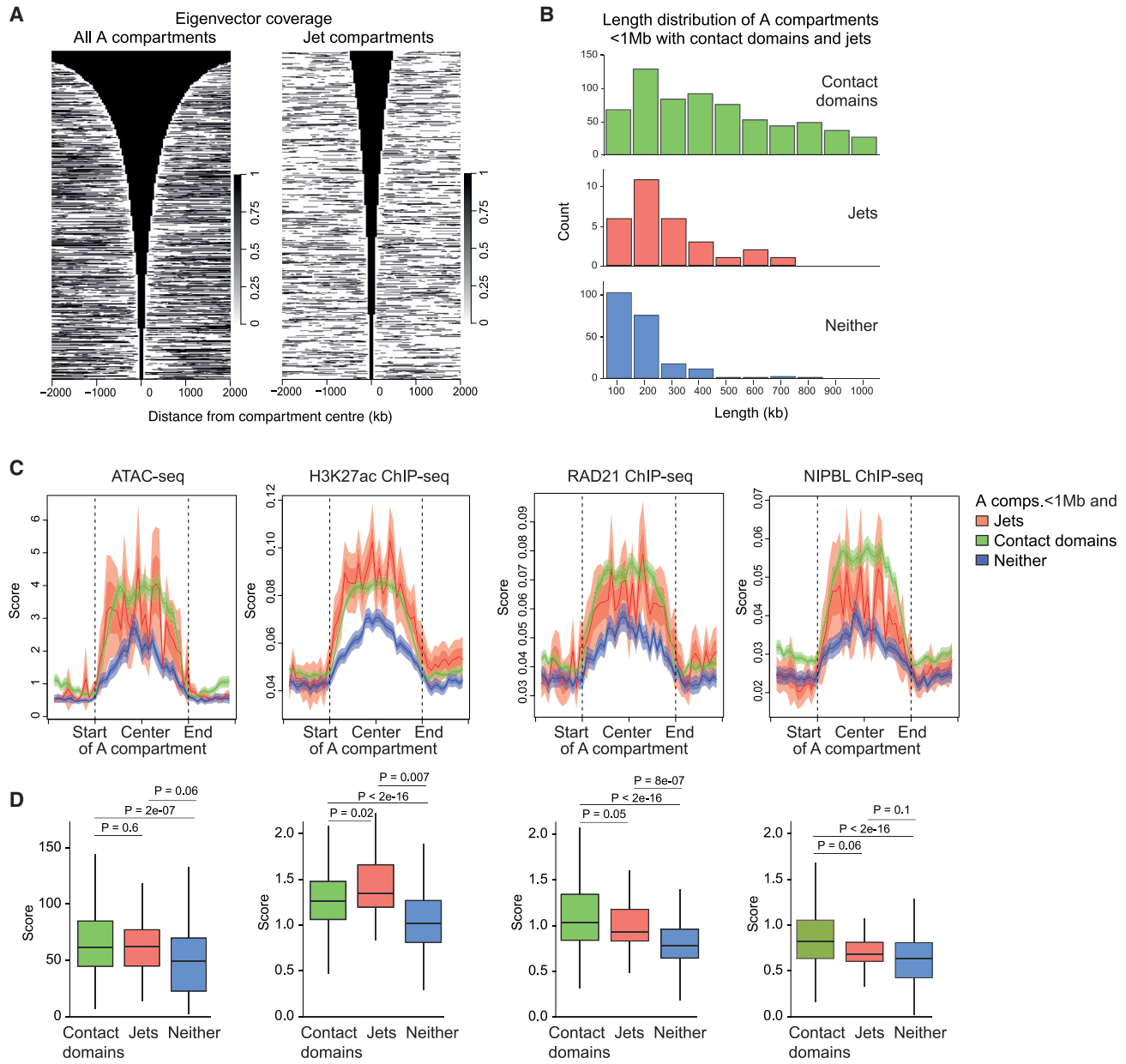


Figure 4. Jet origins have chromatin features consistent with cohesin loading sites

(A) Length distribution of all A compartment intervals called genome wide versus A compartment intervals that give rise to jets. Based on n = 38 jets and n = 1,267 A compartment intervals.

(B) Length distribution of A compartment intervals <1 Mb that contain contact domains, jets, or neither.

(C) Strength of ATAC-seq (Miyazaki et al., 2020) and ChIP-seq signal for A compartment intervals <1 Mb that contain contact domains, jets, or neither.

(D) Quantification of ATAC-seq and ChIP-seq signals shown in (C).

seq peaks in DP thymocytes as CTCF motif scores, where positive scores indicate CTCF downstream binding and negative motif scores with CTCF upstream binding. Positive CTCF motif scores were associated with jet deflection toward upstream sequences and negative CTCF motif scores were associated with jet deflection toward downstream sequences (Figure 7E).

The observation that CTCF can alter the projection angle of jets is informative with respect to the behavior of the loop extrusion machinery upon encounter with CTCF. The loop extrusion model posits that extrusion is blocked by CTCF bound to sites in convergent orientation (Fudenberg et al., 2016), and recent studies provide a mechanism for how the encounter of cohesin with the N terminus of CTCF arrests loop extrusion and stabilizes cohesin at CTCF sites (Li et al., 2020; Nora et al., 2020; Pugacheva et al., 2020). If loop extrusion encounters CTCF in only one direction, extrusion may either stop completely or continue

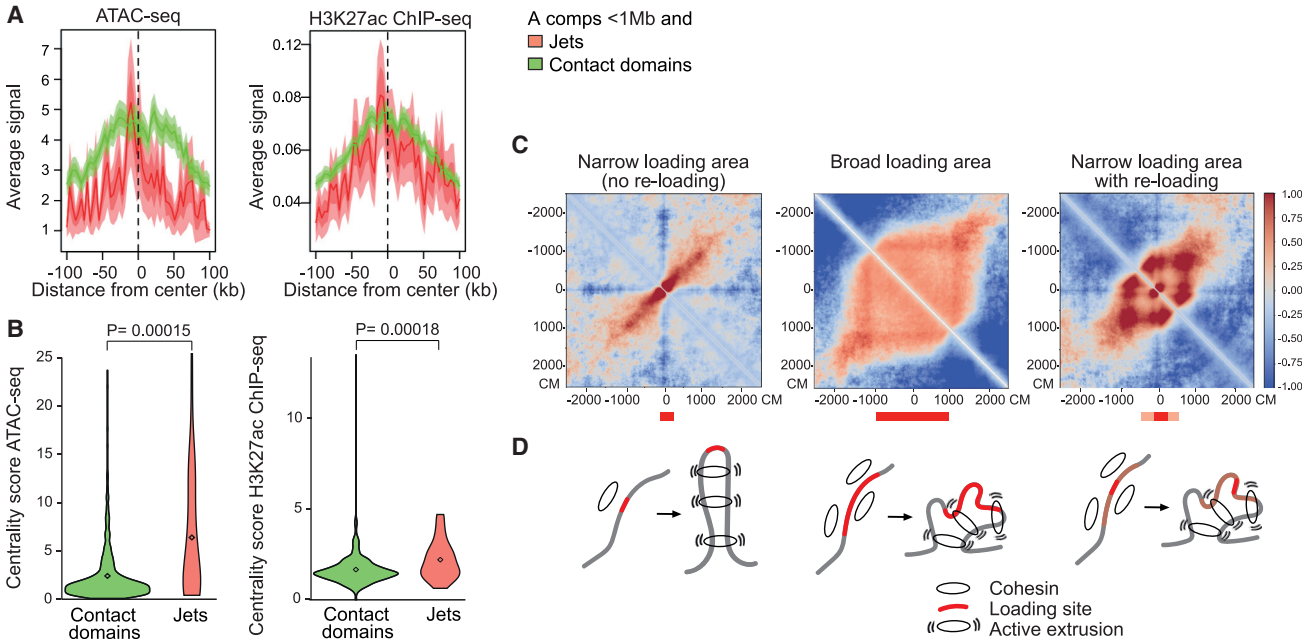


Figure 5. Isolated focal areas of open chromatin favor jet formation

(A) The distribution of ATAC-seq and H3K27ac ChIP-seq signal around the center of jets and contact domains within A compartments <1 Mb. The scale reflects the average size of contact domains within A compartments <1 M, which is 215 kb (n = 866).

(B) Centrality score as a measure of the extent to which ATAC-seq and H3K27ac ChIP-seq signals are focused around the center of the jets and contact domains depicted in (A).

(C) Polymer simulation of the effect of the size of cohesin loading area on the resulting chromatin contacts, jets versus chromatin domains. The size of the cohesin loading area is 100 chromatin monomers (CMs) for the narrow loading area (left) and 2,000 chromatin monomers for the broad loading area (middle). Right: focal loading from a narrow area of 100 chromatin monomers complemented by additional diffuse loading over an area of 2,000 chromatin monomers (see method details section and Table S3 for model parameters).

(D) Schematic. Focal loading can result in jet formation. Broad loading or focal loading with additional diffuse re-loading results in more complex contact patterns.

in the direction that is not blocked by CTCF (Fudenberg et al., 2017; Figure 7F). Our observation that jets can be deflected by an encounter with CTCF indicates that loop extrusion can switch from bidirectional to unidirectional in a CTCF-dependent manner. Using polymer simulations to model how unilateral CTCF encounter affects jet angles, we find that models where a unilateral encounter with CTCF imposes a bidirectional block on cohesin-mediated loop extrusion do not reproduce the experimental observation of altered jet angles (Figure 7G). We next modeled CTCF effects on jet angles under the assumption that unilateral encounter with CTCF imposes a unidirectional block on cohesin-mediated loop extrusion (Figure 7H). The results of this simulation reproduce the experimental observations that (1) jet angles change in response to unilateral CTCF encounter and that (2) both arms of the jets show different lengths in response to unilateral CTCF encounter (Figures 7C and 7F; Table S2). These findings provide experimental evidence for previous assumptions that when blocked in one direction, loop extrusion continues in the other direction (Fudenberg et al., 2017). Taken together, our experimental data and *in silico* simulations support a model where cohesin-driven loop extrusion *in vivo* is by default symmetrical and therefore bidirectional. However, CTCF can significantly deflect the projection angle of cohesin-driven chromatin contacts, indicating that extrusion can switch from bidirectional to unidirectional.

Taken together, these data provide a mechanistic dissection of *in vivo* loop extrusion. A model for jet formation is depicted in Figure S7, including symmetrical extrusion in the absence of CTCF (Figure S7A) the impact of bi- (Figure S7B) and unidirectional (Figure S7C) CTCF encounters and propagation from narrow versus broad origins (Figures S7A and S7D).

DISCUSSION

Here, we describe and quantitatively characterize cohesin-dependent chromatin jets and provide insights into the physiological behavior of *in vivo* cohesin-mediated loop extrusion.

Loading sites for SMC complexes are well defined in bacteria (Gruber and Errington, 2009; Ganji et al., 2018; Golfier et al., 2020; Kong et al., 2020; Terakawa et al., 2017; Tran et al., 2017; Wang et al., 2017; Wang et al., 2018; Anchimiuk et al., 2021). In mammalian cells, active regulatory elements are enriched for cohesin and NIPBL, ATAC-seq accessibility and H3K27ac, and considered as chromatin features of potential cohesin loading sites (Vian et al., 2018; Olley et al., 2018; Luna-Peláez et al., 2019; Linares-Saldana et al., 2021; Yan et al., 2018; Kagey et al., 2010; Liu et al., 2021b; van den Berg et al., 2017). However, the identity of *in vivo* cohesin loading sites has remained uncertain. Jets originate from isolated areas of accessible chromatin enriched for H3K27ac, NIPBL, and RAD21. Jets that project directly

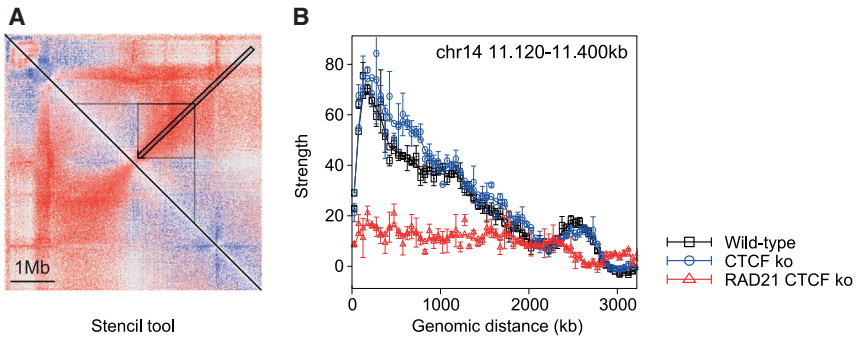


Figure 6. Jets can propagate for at least 1–2 Mb from their origin

(A) Description of the Stencil tool to measure jet length.

(B) Illustration of Hi-C replicate-based measurement of jet reach.

from their origins and propagate perpendicular to the Hi-C diagonal provide powerful evidence for the location of cohesin loading. Computational modeling supports the interpretation that cohesin is loaded at jet origins. In these simulations, narrow sites of cohesin loading give rise to jets. Previous simulations suggested that focal cohesin loading at promoters might generate jet-like features (Banigan et al., 2022), although these were not observed experimentally. In contrast to narrow loading sites, broad areas of cohesin loading lead to domain formation (Fudenberg et al., 2016, 2017). This interpretation is further strengthened by experimental observations that the likelihood of jet formation is inversely related to the size of the potential cohesin loading area. Jets echo the SMC-driven process that aligns bacterial chromosome arms, both in terms of Hi-C pattern and the requirement for defined loading sites (Gruber and Errington, 2009; Ganji et al., 2018; Marbouty et al., 2015; Golfier et al., 2020; Kong et al., 2020; Terakawa et al., 2017; Tran et al., 2017; Wang et al., 2017, 2018; Anchimiuk et al., 2021). Akin to the alignment of bacterial chromosome arms, jets may reflect the continuous, linear activity of stacked SMC complexes (Figure S7C). Indeed, our simulations indicate that the "re-loading" of additional cohesin complexes along the length of the jet would disrupt jet formation (Figure S7D).

Recent reports describe interesting chromatin features that are similar to jets in appearance and likely related in terms of mechanism. In yeast, pericentromeric jet-like features form when centromeres are engaged by mitotic spindles (Paldi et al., 2020), and the jet-like alignment of chromatin fragments are associated with double-strand break repair (Piazza et al., 2021; Arnould et al., 2021). These features are characterized by their short range of ~25 kb. Chromatin flares have been found to form in the specialized chromatin environment of zebrafish sperm (Wike et al., 2021). Flares show evidence of linear alignment of DNA sequences, but with an average length of ~175 kb, they are on a smaller scale than jets. Flare formation has not been linked to specific SMC complexes (Wike et al., 2021). A recent preprint observes the transient occurrence of cohesin-dependent jet-like chromatin features in response to the co-depletion of WAPL and CTCF in mouse ES cells, referred to as plumes (Liu et al., 2021b). As described here for jets, these structures originate from small accessible regions surrounded by B compartments (Liu et al., 2021b). However, because plumes form only after the depletion of the cohesin unloading factor WAPL, they cannot serve to establish the physiological range of cohesin-mediated loop extrusion. Although WAPL depletion was required, it was not sufficient for plume formation unless CTCF was also depleted (Liu et al., 2021b), which masks unidirectional *in vivo* loop extrusion upon a one-sided CTCF encounter.

Similar to flares and plumes, jets arise from focal origins. However, unlike flares (Wike et al., 2021), jets form in somatic cells, rather than in the specialized chromatin environment of sperm, and are unequivocally cohesin dependent. Unlike plumes (Liu et al., 2021b), jets form in unperturbed primary cells at steady state and therefore reflect *in vivo* loop extrusion under conditions of physiological cohesin residence time. Therefore, jets allow measurements of the genomic distances traversed by cohesin complexes loaded at specific sites, and jet length may reflect the range of individual cohesin complexes on chromatin *in vivo*. From their origins at isolated areas of accessible chromatin, jets propagate across neighboring closed (B compartment) chromatin to cover distances of approximately 1–2 Mb.

Our findings are consistent with the assumption that cohesin-mediated loop extrusion is symmetrical in the absence of CTCF and therefore bidirectional by default (Banigan et al., 2020). We find that unilateral CTCF encounter deflects the angle of jet propagation. This provides strong evidence that cohesin-mediated loop extrusion can switch from bi- to unidirectional *in vivo* (Fudenberg et al., 2017; Vian et al., 2018) and therefore is independently controlled in both directions. Jets provide new insights into the physiological behavior of *in vivo* cohesin-mediated loop extrusion and further our understanding of the principles that underlie genome organization.

Limitations of the study

Although our data show that jets are cohesin dependent and can be blocked or deflected by CTCF, we do not yet fully understand the rules that govern jet formation and, in particular, why not all focal sites of accessibility give rise to jets. We do not know the residence time of cohesin in quiescent primary lymphocytes and, as a result, can only speculate on cohesin extrusion speeds during jet propagation. Estimates of cohesin residence times in interphase range between 10 and 25 min in other cell types (Gerlich et al., 2006; Tedeschi et al., 2013; Hansen et al., 2017; Wutz et al., 2020). Assuming that cohesin residence times are similar in quiescent primary lymphocytes, our data suggest that cohesin can traverse "closed" B compartment chromatin at speeds that are comparable with—and possibly in excess of—SMC complex extrusion speed in bacteria and cohesin-mediated extrusion *in vitro* (Ganji et al., 2018; Golfier et al., 2020; Kong et al., 2020; Terakawa et al., 2017; Tran et al., 2017; Wang et al., 2017, 2018; Davidson et al., 2019; Kim et al., 2019). Our current model for the formation of jets is that individual cohesin complexes mediate continuous

CellPress
OPEN ACCESS

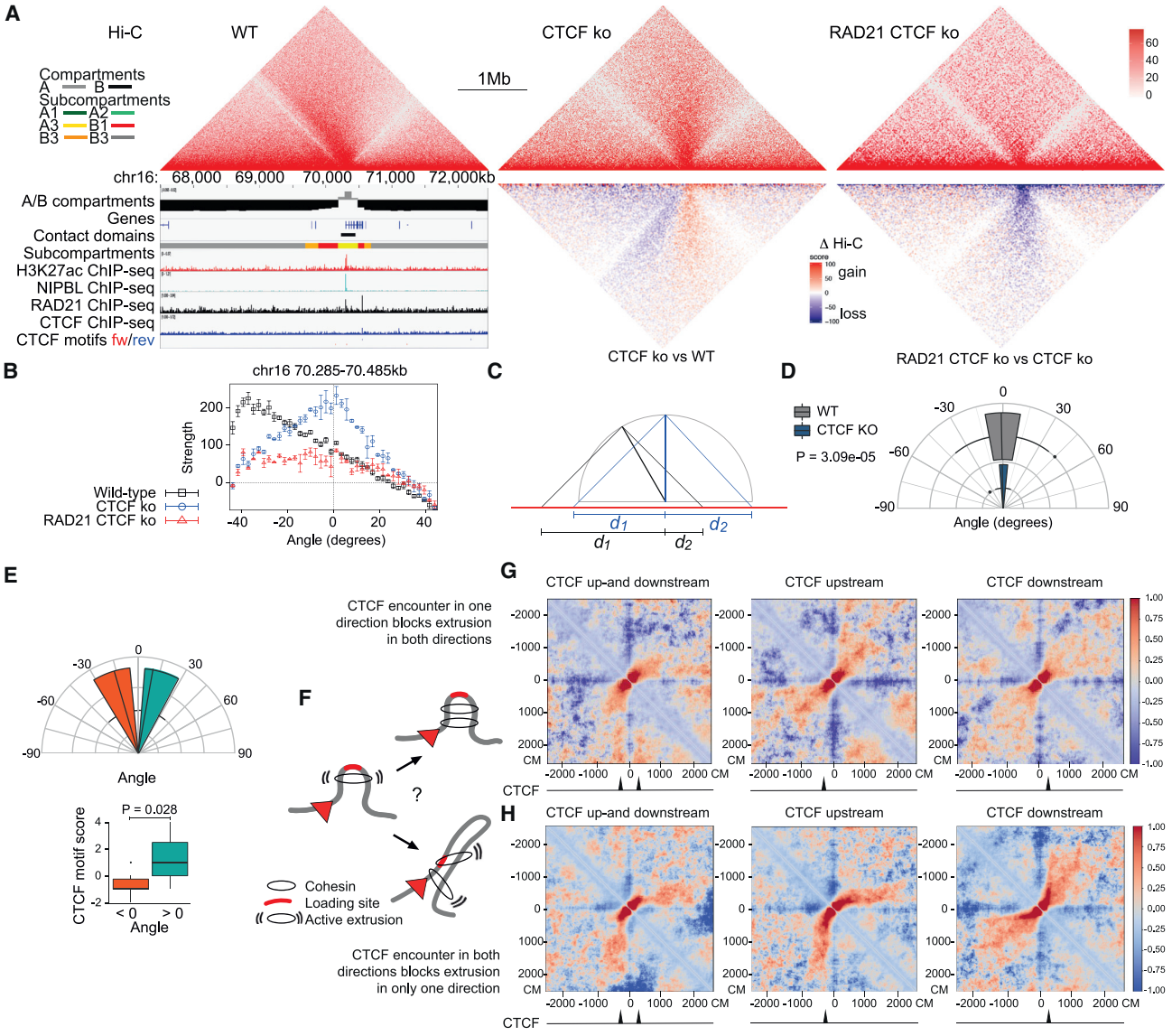


Figure 7. CTCF can modulate jet angles

(A) A jet that is deflected by CTCF (WT) and released by the removal of CTCF (CTCF KO).

(B) Replicate-based protractor scan quantification of the jet shown in (A).

(C) Cohesin travels the same distance in both directions for jets that are perpendicular to the diagonal (d1 = d2, blue). Cohesin travels further in one direction than in the other for jets that are not perpendicular to the diagonal (d1 ≠ d2, black).

(D) Summary of jet angles in the presence and absence of CTCF. 0° describes the perpendicular to the Hi-C diagonal. P indicates the significance of the change in variance (F test).

(E) Experimentally observed jet angles reflect CTCF binding. CTCF motif scores represent the number of CTCF ChIP-seq peaks with convergent motifs downstream within 150 kb of the jet origin minus the number of CTCF ChIP-seq peaks with convergent motifs upstream within 150 kb of the jet origin. Downstream > upstream, n = 8 jets; downstream < upstream, n = 10 jets. High CTCF motif scores indicate CTCF binding downstream and are associated with jet deflection toward upstream sequences and vice versa.

(F) Schematic. Unilateral CTCF encounter (encounter with CTCF in one direction) may inactivate cohesin-mediated loop extrusion completely (top) or convert bidirectional (two-sided) extrusion into unidirectional (one-sided) extrusion (bottom).

(G) Polymer simulation of how CTCF binding (200 monomers up- and/or downstream of the center of the loading area) affects jet angles under the assumption that unilateral encounter with CTCF imposes a bidirectional block on cohesin-mediated loop extrusion

(H) Polymer simulation of how CTCF binding (200 monomers up- and/or downstream of the center of the loading area) affects jet angles under the assumption that unilateral encounter with CTCF imposes a unidirectional block on cohesin-mediated loop extrusion. CMs, chromatin monomers (see method details section).

extrusion along the full length of the jet. It is possible that additional experiments and/or modeling approaches may provide alternative views of jet formation that may not involve continuous extrusion along the full length of the jet. Even in the absence of CTCF, not all jets are the same length, and further work will be required to understand the differences between jets with faster and slower decay rates. Similarly, it will be important to further investigate obstacles in the extrusion path that affect cohesin processivity to result in jets that remain narrowly focused versus jets that become more diffuse as they propagate. Finally, our study indicates that cohesin-mediated loop extrusion is independently controlled in both directions but does not address whether cohesin complexes work as dimers (Kim et al., 2019), monomers (Yatskevich et al., 2019; Davidson et al., 2019), or multimers (Xiang and Koshland, 2021).

STAR★METHODS

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

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Animals
- METHOD DETAILS
  - Mice and cells
  - Hi-C and analysis
  - ChIP-seq
  - Centrality score
  - *In silico* modelling
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.molcel.2022.09.003.

ACKNOWLEDGMENTS

We thank J. Elliott and B. Patel for cell sorting, L. Game for sequencing, and A. Thomas for computational support. This work was supported by the Medical Research Council UK, The Wellcome Trust (222845/Z/21/Z to K.N. and Investigator Award 099276/Z/12/Z to M.M.), EMBO (ALTF 620-2016 to Y.G.), a UK Government Industrial Strategy Rutherford Fund Fellowship (Y.G.), a UK National Productivity Investment Fund Ph.D. studentship in Data Science or Artificial Intelligence (E.A.-J.), and the Shanghai Science and Technology Commission (20PJ1405500/21DZ2210200 to Y.G.). R.G.-M. was supported in part by the European Research Council under the EU's Horizon 2020 Programme (grant number 740269) and acknowledges support from a St John's College Research Fellowship, University of Cambridge.

AUTHOR CONTRIBUTIONS

Y.G., E.A.-J., R.G.-M., K.N., B.L., D.R., A.G.F., G.P., and M.M. conceptualized the study. Y.G. generated the data. Y.G., E.A.-J., R.G.-M., K.N., J.W.D.K., A.J.N., G.P., and M.M. analyzed and visualized the data. K.N. did the computer simulation. N.G. contributed unique reagents and expertise. Y.G., E.A.-J., R.G.-M., K.N., B.L., D.R., A.G.F., G.P., and M.M. wrote the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

INCLUSION AND DIVERSITY

We support inclusive, diverse, and equitable conduct of research.

Received: March 29, 2022
Revised: July 26, 2022
Accepted: September 1, 2022
Published: September 30, 2022

REFERENCES

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

Anchimiuk, A., Lioy, V.S., Bock, F.P., Minnen, A., Boccard, F., and Gruber, S. (2021). A low SMC flux avoids collisions and facilitates chromosome organization in bacillus subtilis. eLife *10*, e65467.

Arnould, C., Rocher, V., Finoux, A.L., Clouaire, T., Li, K., Zhou, F., Caron, P., Mangeot, P.E., Ricci, E.P., Mourad, R., et al. (2021). Loop extrusion as a mechanism for formation of DNA damage repair foci. Nature *590*, 660–665.

Banigan, E.J., and Mirny, L.A. (2020). Loop extrusion: theory meets single-molecule experiments. Curr. Opin. Cell Biol. *64*, 124–138.

Banigan, E.J., Tang, W., van den Berg, A.A., Stocsits, R.R., Wutz, G., Brandão, H.B., Busslinger, G.A., Peters, J.-M., and Mirny, L.A. (2022). Transcription shapes 3D chromatin organization by interacting with loop-extruding cohesin complexes. Preprint at bioRxiv. https://doi.org/10.1101/2022.01.07.475367v1.

Banigan, E.J., van den Berg, A.A., Brandão, H.B., Marko, J.F., and Mirny, L.A. (2020). Chromosome organization by one-sided and two-sided loop extrusion. eLife *9*, e53558.

Davidson, I.F., Bauer, B., Goetz, D., Tang, W., Wutz, G., and Peters, J.M. (2019). DNA loop extrusion by human cohesin. Science *366*, 1338–1345.

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.

Dekker, J., and Mirny, L. (2016). The 3D genome as moderator of chromosomal communication. Cell *164*, 1110–1121.

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

Feng, J., Liu, T., Qin, B., Zhang, Y., and Liu, X.S. (2012). Identifying ChIP-seq enrichment using Macs. Nat. Protoc. *7*, 1728–1740.

Fudenberg, G., Abdennur, N., Imakaev, M., Goloborodko, A., and Mirny, L.A. (2017). Emerging evidence of chromosome folding by loop extrusion. Cold Spring Harb. Symp. Quant. Biol. *82*, 45–55.

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.

Ganji, M., Shaltiel, I.A., Bisht, S., Kim, E., Kalichava, A., Haering, C.H., and Dekker, C. (2018). Real-time imaging of DNA loop extrusion by condensin. Science *360*, 102–105.

Gerlich, D., Koch, B., Dupeux, F., Peters, J.-M., and Ellenberg, J. (2006). Live-cell imaging reveals a stable cohesin-chromatin interaction after but not before DNA replication. Curr. Biol. *16*, 1571–1578.

Golfier, S., Quail, T., Kimura, H., and Brugués, J. (2020). Cohesin and condensin extrude DNA loops in a cell cycle-dependent manner. eLife *9*, e53885.

Gruber, S., and Errington, J. (2009). Recruitment of condensin to replication origin regions by ParB/SpoOJ promotes chromosome segregation in B. subtilis. Cell *137*, 685–696.

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., van Steensel, B., et al. (2017). The cohesin release factor Wapl restricts chromatin loop extension. Cell *169*, 693–707.e14.

Hansen, A.S., Pustova, I., Cattoglio, C., Tjian, R., and Darzacq, X. (2017). CTCF and cohesin regulate chromatin loop stability with distinct dynamics. eLife *6*, e25776.

Heath, H., de Almeida, C.R., Sleutels, F., Dingjan, G., van de Nobelen, S., Jonkers, I., Ling, K.-W., Gribnau, J., Renkawitz, R., Grosveld, F., et al. (2008). CTCF regulates cell cycle progression of αβT cells in the thymus. EMBO J. *27*, 2839–2850.

Imakaev, M., Goloborodko, A., and Brandao, H.B. (2019). mirnylab/polychrom, v0.1.0, (p. v0.1.0). Zenodo. https://doi.org/10.5281/zenodo.3579473.

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.

Kieffer-Kwon, K., Nimura, K., Rao, S.S.P., Xu, J., Jung, S., Pekowska, A., Dose, M., Stevens, E., Mathe, E., Dong, P., et al. (2017). Myc regulates chromatin decompaction and nuclear architecture during B cell activation. Mol. Cell *67*, 566–578.e10.

Kim, Y., Shi, Z., Zhang, H., Finkelstein, I.J., and Yu, H. (2019). Human cohesin compacts DNA by loop extrusion. Science *366*, 1345–1349.

Kimura, K., Rybenkov, V.V., Crisona, N.J., Hirano, T., and Cozzarelli, N.R. (1999). 13s condensin actively reconfigures DNA by introducing global positive writhe: implications for chromosome condensation. Cell *98*, 239–248.

Kong, M., Cutts, E.E., Pan, D., Beuron, F., Kaliyappan, T., Xue, C., Morris, E.P., Musacchio, A., Vannini, A., and Greene, E.C. (2020). Human condensin I and II drive extensive ATP-dependent compaction of nucleosome-bound DNA. Mol. Cell *79*, 99–114.e9.

Langmead, B. (2010). Aligning short sequencing reads with Bowtie. Curr. Protoc. Bioinformatics *Chapter 11*. Unit 11.7. https://doi.org/10.1002/0471250953.bi1107s32.

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

Lee, P.P., Fitzpatrick, D.R., Beard, C., Jessup, H.K., Lehar, S., Makar, K.W., Pérez-Melgosa, M., Sweetser, M.T., Schlissel, M.S., Nguyen, S., et al. (2001). A critical role for dnmt1 and DNA methylation in T cell development, function, and survival. Immunity *15*, 763–774.

Li, Y., Haarhuis, J.H.I., Sedeño Cacciatore, Á.S., Oldenkamp, R., van Ruiten, M.S., Willems, L., Teunissen, H., Muir, K.W., de Wit, E., Rowland, B.D., et al. (2020). The structural basis for cohesin–CTCF-anchored loops. Nature *578*, 472–476.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; 1000 Genome Project Data Processing Subgroup (2009). The sequence alignment/map format and SAMtools. Bioinformatics *25*, 2078–2079.

Linares-Saldana, R., Kim, W., Bolar, N.A., Zhang, H., Koch-Bojalad, B.A., Yoon, S., Shah, P.P., Karnay, A., Park, D.S., Luppino, J.M., et al. (2021). Brd4 orchestrates genome folding to promote neural crest differentiation. Nat. Genet. *53*, 1480–1492.

Liu, N.Q., Maresca, M., van den Brand, T., Braccioli, L., Schijns, M.M.G.A., Teunissen, H., Bruneau, B.G., Nora, E.P., and de Wit, E. (2021a). Wapl maintains a cohesin loading cycle to preserve cell-type-specific distal gene regulation. Nat. Genet. *53*, 100–109.

Liu, N.Q., Magnitov, M., Schijns, M., van Schaik, T., van der Weide, R.H., Teunissen, H., van Steensel, B., and de Wit, E. (2021b). Rapid depletion of CTCF and cohesin proteins reveals dynamic features of chromosome architecture. Preprint at bioRxiv. https://doi.org/10.1101/2021.08.27.457977.

Luna-Peláez, N., March-Díaz, R., Ceballos-Chávez, M., Guerrero-Martínez, J.A., Grazioli, P., García-Gutiérrez, P., Vaccari, T., Massa, V., Reyes, J.C., and García-Domínguez, M. (2019). The Cornelia de Lange syndrome-associated factor NIPBL interacts with Brd4 ET domain for transcription control of a common set of genes. Cell Death Dis. *10*, 548.

Marbouty, M., Le Gall, A., Cattoni, D.I., Cournac, A., Koh, A., Fiche, J.B., Mozziconacci, J., Murray, H., Koszul, R., and Nollmann, M. (2015). Condensin- and replication-mediated bacterial chromosome folding and origin condensation revealed by Hi-C and super-resolution imaging. Mol. Cell *59*, 588–602.

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

Mirny, L.A., Imakaev, M., and Abdennur, N. (2019). Two major mechanisms of chromosome organization. Curr. Opin. Cell Biol. *58*, 142–152.

Miyazaki, K., Watanabe, H., Yoshikawa, G., Chen, K., Hidaka, R., Aitani, Y., Osawa, K., Takeda, R., Ochi, Y., Tani-ichi, S., et al. (2020). The transcription factor E2A activates multiple enhancers that drive *Rag* expression in developing T and B cells. Sci. Immunol. *5*, eabb1455.

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

Nasmyth, K., and Haering, C.H. (2009). Cohesin: its roles and mechanisms. Annu. Rev. Genet. *43*, 525–558.

Nora, E.P., Caccianini, L., Fudenberg, G., So, K., Kameswaran, V., Nagle, A., Uebersohn, A., Hajj, B., Saux, A.L., Coulon, A., et al. (2020). Molecular basis of CTCF binding polarity in genome folding. Nat. Commun. *11*, 5612.

Nora, E.P., Goloborodko, A., Valton, A.-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.

Olley, G., Ansari, M., Bengani, H., Grimes, G.R., Rhodes, J., von Kriegsheim, A., Blatnik, A., Stewart, F.J., Wakeling, E., Carroll, N., et al. (2018). Brd4 interacts with NIPBL and BRD4 is mutated in a Cornelia de Lange–like syndrome. Nat. Genet. *50*, 329–332.

Paldi, F., Alver, B., Robertson, D., Schalbetter, S.A., Kerr, A., Kelly, D.A., Baxter, J., Neale, M.J., and Marston, A.L. (2020). Convergent genes shape budding yeast pericentromeres. Nature *582*, 119–123.

Piazza, A., Bordelet, H., Dumont, A., Thierry, A., Savocco, J., Girard, F., and Koszul, R. (2021). Cohesin regulates homology search during recombinational DNA repair. Nat. Cell Biol. *23*, 1176–1186.

Pugacheva, E.M., Kubo, N., Loukinov, D., Tajmul, M., Kang, S., Kovalchuk, A.L., Strunnikov, A.V., Zentner, G.E., Ren, B., and Lobanenkov, V.V. (2020). CTCF mediates chromatin looping via N-terminal domain-dependent cohesin retention. Proc. Natl. Acad. Sci. USA *117*, 2020–2031.

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

Rao, S.S.P., Huang, S.C., Glenn St Hilaire, B.G.S., Engreitz, J.M., Perez, E.M., Kieffer-Kwon, K.R., Sanborn, A.L., Johnstone, S.E., Bascom, G.D., Bochkov, I.D., et al. (2017). Cohesin loss eliminates all loop domains. Cell *171*, 305–320.e24.

Riggs, A.D. (1990). DNA methylation and late replication probably aid cell memory, and type I DNA reeling could aid chromosome folding and enhancer function. Philos. Trans. R. Soc. Lond. B Biol. Sci. *326*, 285–297.

Robinson, J.T., Turner, D., Durand, N.C., Thorvaldsdóttir, H., Mesirov, J.P., and Aiden, E.L. (2018). Juicebox.js provides a cloud-based visualization system for Hi-C data. Cell Syst. *6*, 256–258.e1.

Sanborn, A.L., Rao, S.S.P., Huang, S.-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.H., Mirny, L., et al. (2017). Two independent modes of chromatin organization revealed by cohesin removal. Nature *551*, 51–56.

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 pre-existing architectural compartments. Genome Res. *23*, 2066–2077.

Seitan, V.C., Hao, B., Tachibana-Konwalski, K., Lavagnolli, T., Mira-Bontenbal, H., Brown, K.E., Teng, G., Carroll, T., Terry, A., Horan, K., et al. (2011). A role for cohesin in T-cell-receptor rearrangement and thymocyte differentiation. Nature *476*, 467–471.

Servant, N., Varoquaux, N., Lajoie, B.R., Viara, E., Chen, C.J., Vert, J.P., Heard, E., Dekker, J., and Barillot, E. (2015). HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. Genome Biol. *16*, 259.

Tachibana-Konwalski, K., Godwin, J., van der Weyden, L., Champion, L., Kudo, N.R., Adams, D.J., and Nasmyth, K. (2010). Rec8-containing cohesin maintains bivalents without turnover during the growing phase of mouse oocytes. Genes Dev. *24*, 2505–2516.

Tedeschi, A., Wutz, G., Huet, S., Jaritz, M., Wuensche, A., Schirghuber, E., Davidson, I.F., Tang, W., Cisneros, D.A., Bhaskara, V., et al. (2013). Wapl is an essential regulator of chromatin structure and chromosome segregation. Nature *501*, 564–568.

Terakawa, T., Bisht, S., Eeftens, J.M., Dekker, C., Haering, C.H., and Greene, E.C. (2017). The condensin complex is a mechanochemical motor that translocates along DNA. Science *358*, 672–676.

Tran, N.T., Laub, M.T., and Le, T.B.K. (2017). SMC progressively aligns chromosomal arms in Caulobacter crescentus but is antagonized by convergent transcription. Cell Rep. *20*, 2057–2071.

van den Berg, D.L.C., Azzarelli, R., Oishi, K., Martynoga, B., Urbán, N., Dekkers, D.H.W., Demmers, J.A., and Guillemot, F. (2017). Nipbl Interacts with Zfp609 and the Integrator complex to regulate cortical neuron migration. Neuron *93*, 348–361.

Vian, L., Pękowska, A., Rao, S.S.P., Kieffer-Kwon, K.R., Jung, S., Baranello, L., Huang, S.C., El Khattabi, L., Dose, M., Pruett, N., et al. (2018). The energetics and physiological impact of cohesin extrusion. Cell *173*, 1165–1178.e20.

Vietri Rudan, M., Barrington, C., Henderson, S., Ernst, C., Odom, D.T., Tanay, A., and Hadjur, S. (2015). Comparative Hi-C reveals that CTCF underlies evolution of chromosomal domain architecture. Cell Rep. *10*, 1297–1309.

Wang, X., Brandão, H.B., Le, T.B.K., 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.

Wang, X., Hughes, A.C., Brandão, H.B., Walker, B., Lierz, C., Cochran, J.C., Oakley, M.G., Kruse, A.C., and Rudner, D.Z. (2018). In vivo evidence for ATPase-dependent DNA translocation by the bacillus subtilis SMC condensin complex. Mol. Cell *71*, 841–847.e5.

Wike, C.L., Guo, Y., Tan, M., Nakamura, R., Shaw, D.K., Díaz, N., Whittaker-Tademy, A.F., Durand, N.C., Aiden, E.L., Vaquerizas, J.M., et al. (2021). Chromatin architecture transitions from zebrafish sperm through early embryogenesis. Genome Res. *31*, 981–994.

Wutz, G., Ladurner, R., St Hilaire, B.G.S., Stocsits, R.R., Nagasaka, K., Pignard, B., Sanborn, A., Tang, W., Várnai, C., Ivanov, M.P., et al. (2020). Esco1 and CTCF enable formation of long chromatin loops by protecting cohesin Stag1 from Wapl. eLife *9*, e52091.

Wutz, G., Várnai, C., Nagasaka, K., Cisneros, D.A., Stocsits, R.R., Tang, W., Schoenfelder, S., Jessberger, G., Muhar, M., and Hossain, M.J. (2017). Topologically associating domains and chromatin loops depend on cohesin are are regulated by CTCF, WAPL, and PDS5 proteins. EMBO J *36*, 3573–3599. https://doi.org/10.15252/embj.201798004.

Xiang, S., and Koshland, D. (2021). Cohesin architecture and clustering in vivo. eLife *10*, e62243.

Yan, J., Chen, S.-A.A., Local, A., Liu, T., Qiu, Y., Dorighi, K.M., Preissl, S., Rivera, C.M., Wang, C., Ye, Z., et al. (2018). Histone H3 lysine 4 monomethylation modulates long-range chromatin interactions at enhancers. Cell Res. *28*, 204–220.

Yatskevich, S., Rhodes, J., and Nasmyth, K. (2019). Organization of chromosomal DNA by SMC complexes. Annu. Rev. Genet. *53*, 445–482.

Zhang, N., Kuznetsov, S.G., Sharan, S.K., Li, K., Rao, P.H., and Pati, D. (2008). A handcuff model for the cohesin complex. J. Cell Biol. *183*, 1019–1031.

CellPress
OPEN ACCESS

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit polyclonal anti-CTCF	Abcam	Cat# ab70303; RRID:AB_1209546
Rabbit polyclonal anti-H3K27ac	Active Motif	Cat# 39133; RRID:AB_2561016
Rabbit polyclonal anti-Rad21	Abcam	Cat# ab992; RRID:AB_2176601
Rabbit polyclonal anti-SMC3	Abcam	Cat# ab9263; RRID:AB_307122
Rabbit polyclonal anti-Histone H3	Abcam	Cat# ab1791; RRID:AB_302613
Rat monoclonal anti-CD4	Biolegend	Cat# 100514; RRID:AB_312717
Rat monoclonal anti-CD8a	Thermo Fisher Scientific	Cat# 12-0081-85; RRID:AB_465532
Arm.Hamster monoclonal anti-CD69	BD Biosciences	Cat# 553236; RRID:AB_394725
Chemicals, peptides, and recombinant proteins		
16% Formaldehyde, Methanol-free	Pierce	Cat# 28908
Deposited data		
Hi-C, ChIP-seq data (CTCF, Rad21, H3K27ac)	This Study	GEO:GSE199059
ChIP-seq data (Nipbl)	Seitan et al., 2013	GEO:GSE48763
Experimental models: Organisms/strains		
Mouse: Rad21tm1.1Mmk	Our colony	MGI:5293824 https://doi.org/10.1038/nature10312
Mouse: Ctcf conditional allele	Our colony	https://doi.org/10.1038/emboj.2008.214
Mouse: Tg(Cd4-cre)1Cwi	Jax B6.Cg-Tg(Cd4-cre)1 Cwi/BfluJ	MGI:2386448 https://doi.org/10.1016/S1074-7613(01)00227-8
Critical commercial assays		
NEBNext ChIP-seq Library Prep Master Mix Set for Illumina	NEB	E6240
NEBNext Multiplex Oligos for Illumina	NEB	E7335, E7500
AMPure XP	Beckman Coulter	A63881
Software and algorithms		
Bowtie2 (v2.3.2)	Langmead and Salzberg, 2012	http://bowtie-bio.sourceforge.net/bowtie2/index.shtml
HiC-Pro (v2.7.8)	Servant et al., 2015	https://github.com/nservant/HiC-Pro
Juicer (v0.7.5)	Durand et al., 2016	https://github.com/aidenlab/juicer
Juicebox	Durand et al., 2016	https://github.com/aidenlab/Juicebox
R	The R Foundation	https://www.r-project.org/
Samtools (v1.4)	Li et al., 2009	http://samtools.sourceforge.net/
Stencil tool	This paper	https://github.com/garciamillan/DNA_jets
Protractor tool	This paper	https://github.com/garciamillan/DNA_jets
Polychrom	Imakaev, 2019	Version v0.1.0, 10.5281/zenodo.3579473
bowtie (v1.0.0)	Langmead, 2010	http://bowtie-bio.sourceforge.net/index.shtml
MACS2 (v2.1.0)	Zhang et al., 2008	https://github.com/macs3-project/MACS/wiki/Advanced%3A-Call-peaks-using-MACS2-subcommands

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to the lead contact, Matthias Merkenschlager (matthias.merkenschlager@lms.mrc.ac.uk).

Materials availability

This study did not generate new unique reagents.

Data and code availability

In situ Hi-C data for Hi-C CD69^{-} DP wild-type, CD4 Cre Ctcf$^{-/-}$ (CTCF KO), and CD4Cre Rad21$^{-/-}$ Ctcf$^{-/-}$ (DKO) thymocytes and ChIP-seq data for RAD21, CTCF, and H3K27ac in wild-type DP thymocytes generated in this study have been deposited at the NCBI Gene Expression Omnibus (GEO) under accession number GSE199059 and are publicly available as of the date of publication. ATAC-seq data in wild-type DP thymocytes are publicly available at GEO GSE141223 (Miyazaki et al., 2020). Dilution Hi-C for wild-type and CD4 Cre Rad21$^{-/-}$ DP thymocytes and ChIP-seq data for NIPBL in wild-type DP thymocytes are publicly available at GEO GSE48763 (Seitan et al., 2013). Mouse B cell Hi-C data are publicly available at GEO GSE82144 (Kieffer-Kwon et al., 2017).

All original code has been deposited at Zenodo and is publicly available as of the date of publication. The DOI for code for jet simulations is https://doi.org/10.5281/zenodo.7028262. The DOI for code for jet quantification is https://doi.org/10.5281/zenodo.7034657.

Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Animals

Mice harbouring a conditional *Ctcf* allele (*Ctcf*$^{lox/lox}$, Heath et al., 2008) were crossed with CD4Cre (Lee et al., 2001, RRID:IMSR_EM:01139). Mice harbouring a conditional *Rad21* allele (*Rad21*$^{lox/lox}$, Tachibana-Konwalski et al., 2010; Seitan et al., 2011, RRID:MGI:5293828) were crossed with CD4Cre (Lee et al., 2001, RRID:IMSR_EM:01139). Mice were maintained under specific pathogen free conditions, 12 hour light/dark cycle, food and water ad libitum, under the care of Imperial College London animal care staff and veterinary services. Males and females of the following genotypes were used between the ages of 6 and 12 weeks: *Ctcf*$^{wt/wt}$, *Rad21*$^{wt/wt}$ CD4Cre (controls), *Ctcf*$^{lox/lox}$, *Rad21*$^{wt/wt}$ CD4Cre (CTCF ko), *Ctcf*$^{lox/lox}$, *Rad21*$^{lox/lox}$ CD4Cre (double ko). Experimental groups were defined by genotype. Mouse work was performed according to the Animals (Scientific Procedures) Act Mouse work was done under a project licence issued by the UK Home Office, UK following review by the Imperial College London Animal Welfare and Ethical Review Body (AWERB).

METHOD DETAILS

Mice and cells

Mouse strains harbouring conditional alleles for *Ctcf* (*Ctcf*$^{lox/lox}$, Heath et al., 2008) and *Rad21* (*Rad21*$^{lox/lox}$, Tachibana-Konwalski et al., 2010; Seitan et al., 2011, RRID:MGI:5293828) were crossed with CD4Cre (Lee et al., 2001, RRID:IMSR_EM:01139) for lineage- and developmental stage-specific deletion at the transition to the CD4+ CD8+ (DP) stage of T cell development in the thymus (Lee et al., 2001) as described (Seitan et al., 2011). Quiescent CD4^{+} CD8^{+} CD69^{-} DP thymocytes were isolated by flow cytometry (Figure S1B).

Hi-C and analysis

In situ Hi-C was performed as described (Rao et al., 2014) with the following modifications. Cells were cross-linked in 1% formaldehyde for 10 min at room temperature. Chromatin was digested with 100 units of MboI overnight at 37 °C while shaking at 900 rpm. After overnight incubation, 100 units of MboI were added for the second-round digestion for 2-4 hours. The DNA polymerase I, large (Klenow) fragment, was used to fill in MboI-digested fragment overhangs in the presence of biotin-labelled dATP. Fragment ends were then ligated with 4000 units of T4 DNA ligase by incubating at room temperature for 4-6 h with rotation at 12 rpm. After ligation, cell nuclei were collected by centrifugation, and chromatin was reverse crosslinked overnight at 65 °C. RNase A was used to remove RNA after decrosslinking. Genomic DNA was isolated and precipitated by sodium acetate-ethanol precipitation, and then purified DNA was sheared for 8 min using a sonicator (Bioruptor) with high power setting, and 30 seconds on and 30 seconds off per minute. DNA fragments in the range of 300–500 bp were then selected using the AMPure XP beads (Beckmann Coulter). After biotin-labelled DNA fragments were captured on Dynabeads MyOne Streptavidin T1 (Thermo Fisher), fragment ends were repaired in a mixture of enzymes containing T4 polynucleotide kinase, T4 DNA polymerase I and DNA polymerase I, large (Klenow) fragment, and then NEBNext adaptors for Illumina sequencing were ligated to the dA-tailed fragment ends. After the USER enzyme digestion, NEBNext oligos for Illumina sequencing were used for PCR for library preparation. A PCR titration was carried out to determine the lowest number of PCR cycles (8-10 cycles in this work). Final Hi-C sequencing libraries were quantified and checked for size

distribution using Bioanalyzer (Agilent Technologies), and then sequenced on an HiSeq 2500 sequencer (Illumina). Hi-C raw sequencing data (2 × 100 bp paired-end reads) were mapped using the bowtie 2 (version 2.3.2) and then processed by the HiC-Pro (version 2.7.8) with default settings (Servant et al., 2015; Langmead and Salzberg, 2012). The reference genome used for mapping in this study was the UCSC assembly mm9, NCBI build 37. After converting using the script, HiC-Pro hicpro2juicebox.sh, valid chromatin contacts were normalized and.hic files were created by the Juicer Pre (Juicer tools version 0.7.5) (Durand et al., 2016). Compartments were called by eigenvector decomposition, performed using the cooltools eigdecomp module (https://github.com/open2c/cooltools) at 25kb resolution. A and B compartments were assigned by GC-content. Hi-C contact domains were called using Arrowhead with default parameters at 5kb and 10kb resolution as described (Rao et al., 2014). For cases of proximal 5kb and 10kb domains (defined as <10kb Euclidean distance) the 10kb domains were removed. To reduce redundancy in downstream analysis nested domains were then pruned such that only internal and non-nested domains were retained.

Subcompartments were determined at 100kb resolution based on methods described (Rao et al., 2014) with the following alterations: intra-chromosomal rather than inter-chromosomal contacts were used for classification, k-means clustering was used to assign subcompartments instead of a hidden Markov mode, a 5Mb mask was applied around the diagonal of each Hi-C map to remove the impact of local interactions, and Hi-C matrices were interpolated to reduce noise. Hi-C data were visualised in Juicebox (Robinson et al., 2018).

ChIP-seq

Cells for were cross-linked in 1% formaldehyde solution for 10 minutes at room temperature and then the reaction was quenched with 125 mM glycine at room temperature for 5 minutes. Cell nuclei were isolated using a RIPA lysis buffer (10mM Tris pH7.5, 1mM EDTA, 1% Triton X-100, 0.1% Sodium deoxycholate, 0.1% SDS, 0.15M NaCl, 1X Roche protease inhibitors). Following isolation of cell nuclei, genomic DNA was sonicated in a high-salt RIPA buffer (10mM Tris pH7.5, 1mM EDTA, 1% Triton X-100, 0.1% Sodium deoxycholate, 1% SDS, 0.8M NaCl, 1X Roche protease inhibitors) for 30 minutes on a sonicator (Bioruptor) with high power setting, and 30 seconds on and 30 seconds off per minute. Immunoprecipitation was performed in a RIPA buffer (10mM Tris pH7.5, 1mM EDTA, 1% Triton X-100, 0.1% Sodium deoxycholate, 0.1% SDS, 0.15M NaCl, 1X Roche protease inhibitors), and antibody-antigen complex was precipitated using Dynabead Protein G (Thermo Fisher). Protein was digested using Proteinase K, and then DNA-protein complex was reverse crosslinked overnight at 65°C. RNase A was added to remove RNA after decrosslinking. ChIP-seq library preparation were carried out using the NEBNext ChIP-Seq Library Prep Master Mix Set for Illumina (NEB) according to the manufacturer's protocol. Final sequencing libraries were sequenced on an HiSeq 2500 sequencer (Illumina). ChIP-seq raw reads (50 bp single-end) were mapped using the bowtie (version 1.0.0) (Langmead, 2010). After alignment and duplicates removal, peaks were identified using MACS2 (Feng et al., 2012).

Centrality score

The centrality score is the sum of ATAC-seq or ChIP-seq signal at the 50kb centred on the feature divided by the sum of the signal from the 50-75kb upstream and downstream of the feature.

In silico modelling

We use simulations to investigate the effect of cohesin loop extrusion and CTCF binding on Hi-C maps. The simulations work in two stages. One-dimensional movement of cohesin subunits across chromatin is followed by three-dimensional polymer folding simulations to produce simulated Hi-C maps. To model chromatin folding we use the publicly available software polychrom (Version v0.1.0, 10.5281/zenodo.3579473, Imakaev et al., 2019) and manually add pairs of cohesin cuffs as extra bonds that define the structure, as described in the following.

Cohesin is modelled as a pair of two random walkers that move along a straight line representing chromatin, modelled as a chain of monomers. Each monomer is a possible position for a walker. We initialise the simulation by placing with uniform probability both walkers of a fixed number of cohesin molecules within a certain loading area. We may further place CTCF molecules along the polymer. After this initial setup, the one dimensional simulation follows the process described in Algorithm 1 (Table S4). All of the actions in the algorithm have predefined probabilities. This is explicitly shown for the unloading action, and for the rest it is abbreviated.We model the movement of the two cuffs of each cohesin as a random walk with drift. The left cuff has a drift towards the left and the other towards the right. The cuffs cannot switch places or bypass each other and may therefore stall others. They can further be stopped and captured by CTCF molecules.

Two CTCF capture mechanisms and their impact on the formation of jets are investigated. In the first case, a suitably oriented CTCF captures only the cuff that it encounters while the other cuff of the same cohesin is free to move and further extrude chromatin. In the second case, when a CTCF captures one cuff of a cohesin, both cuffs are considered immobilized and cohesin stops extruding. The drift is the difference between the jump probabilities of the cuffs. Cohesin is unloaded with different probabilities. Cohesin captured by CTCF has a longer residence time on the polymer, which is reflected by the smaller unloading probability compared to free cohesin. We allow for 0, 1 or 2 CTCF sites, at the positions indicated. At each of these sites, we place 5 CTCFs equidistantly, every 50 monomers. Each CTCF may capture a cohesin cuff once it steps on the site with the probability stated and each may release them again with the much smaller release probability indicated.

We chose a 5000-monomers-long polymer for our simulations. The Hi-C maps are based on typically about twenty polymer stochastic systems with the same settings. Once the one dimensional simulation has run, the trajectories of all of the cohesin cuffs, are input as extra bonds in a three dimensional polychrom simulation, from which finally the average contacts are calculated, and the contact map is generated. The three dimensional simulations are run on Nvidia RTX6000 GPUs and use a variable Langevin integrator. Model parameters are listed in Table S3.

QUANTIFICATION AND STATISTICAL ANALYSIS

Quantification of jets: The Protractor tool quantifies jet strength and jet angles, and the Stencil tool quantifies jet reach. Both use observed over expected Hi-C data with KR normalisation. We denote observed data at $(x,y) \in [0,L]^2$ by $c_{x,y}$, and expected data by $e_{x,y}$. Jet candidates are defined by start and end coordinates (α_1, α_2) along a chromosome (equivalently (α_2, α_1)). To determine jet strength and orientation we define the protractor as the set of points on the Hi-C map enclosed by a semi-circle, typically of a 2Mb radius, whose origin is placed at (α_1, α_2). The straight side of the semi-circle is parallel to the Hi-C diagonal (Figure 1B). We divide the area of the protractor into n equal sectors (typically n = 80), labelled with a number $i \in \{0,\ldots, n-1\}$ and symmetric around an angle $\theta_i = 90° ((1 + 2_i)/n-1)$. Angles are within a range from $-90°$ to 90° where 0° is perpendicular to the Hi-C diagonal. We denote Ω_i as the set of data points that are included in sector i. The number of data points within a sector affected by missing data (white lines in the Hi-C map) is given by:

$$W_i = \sum_{(x,y) \in \Omega_i} (w(x) + w(y) - w(x)w(y)) \quad \text{(Equation 1)}$$

where $w(x)$ is an indicator function of a stripe of missing data at $x \in [0,L]$, defined as:

$$w(x) = \begin{cases} 1 \text{ if } c_{x,y} = 0 \text{ for all } y \in [0,L], \\ 0 \text{ otherwise,} \end{cases} \quad \text{(Equation 2)}$$

so that $(w(x) + w(y) - w(x)w(y)) = 1 - (1 - w(x))(1 - w(y))$ in Equation 1 is unity whenever $w(x) = 1$ or $w(y) = 1$. Since the Hi-C map satisfies the symmetry $c_{x,y} = c_{y,x}$, it is equivalent to identify stripes of missing data in the x coordinate or in the y coordinate, which is why $w(x)$ in Equation 2 is defined by probing for missing data along y only. We define f_i as the intensity of observed relative to expected interactions within each sector i,

$$f_i = \begin{cases} \mathcal{N} \sum_{(x,y) \in \Omega_i} (c_{x,y} - e_{x,y}) & \text{if } W_i < \tau, \\ \text{not defined} & \text{otherwise.} \end{cases} \quad \text{(Equation 3)}$$

where $\mathcal{N} = 10^8/T$ and T is the total number of counts, and the tolerance is $\tau = 0.1Q/n$, where Q is the total number of points in the protractor. We estimate the mean and standard deviation of f_i based on Hi-C replicates (Figure 1B). The values f_i may be subject to fluctuations inherent to the Hi-C data, the radius of the protractor, and the total number of sectors n. To obtain robust estimates, we apply a Gaussian kernel to f_i,

$$\overline{f_i} = \sum_{j=0}^{n-1} f_j \frac{1}{B\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{i-j}{B}\right)^2} \quad \text{(Equation 4)}$$

where we set the bandwidth $B = 1$. We determine the index M among those indeces i such that $\theta_i \in [-45°,45°]$, where $\overline{f_i}$ reaches its maximum,

$$\overline{f_M} = \max\left\{\overline{f_i} : i = \left\lceil \frac{n}{4} - \frac{1}{2} \right\rceil, \ldots, \left\lfloor \frac{3n}{4} - \frac{1}{2} \right\rfloor\right\} \quad \text{(Equation 5)}$$

which immediately gives the jet orientation θ_M. A nonzero orientation is indicative of asymmetric extrusion. We define the jet strength as

$$\mathcal{S} = 10^{-2}\overline{f_M}. \quad \text{(Equation 6)}$$

To measure the reach of jets that quantifies interactions between neighbouring regions of (α_1, α_2), we devised the stencil, which is a rectangle 3.5Mb in length and 0.1Mb in width. The stencil is placed on the Hi-C map so that the mid-point of one of the short sides is at (α_1, α_2) and its long sides are aligned with the jet orientation given by θ_M. We include in the stencil only contacts between locations outside the interval $[\alpha_1, \alpha_2]$. We divide the stencil along the direction θ_M into $m = 70$ equally sized pieces (Figure 6A). Similarly to Equations 1, 3, and 4, we define the functions

$$W_i = \sum_{(x,y) \in \Delta_i} (w(x) + w(y) - w(x)w(y)) \quad \text{(Equation 7)}$$

$$g_i = \begin{cases} \mathcal{N} \sum_{(x,y) \in \Delta_i} \left(c_{x,y} - e_{x,y}\right) & \text{if } W_i < \tau, \\ \text{not defined} & \text{otherwise,} \end{cases} \quad \text{(Equation 8)}$$

$$\overline{g_i} = \sum_{j=0}^{n-1} g_j \frac{1}{B\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{i-j}{B}\right)^2}, \quad \text{(Equation 9)}$$

where Δ_i is the set of points in the Hi-C map that are contained in each piece i along thestencil, $i \in \{0,\ldots,m-1\}$. We subtract the background given by Rad21$^{-/-}$ Ctcf$^{-/-}$ from the stencil g_i and find the piece i where the subtracted stencil falls below a threshold of 5. We then calculate the up- and downstream distances d_1 and d_2 (Figure 7C).

All statistical analysis and software used is listed in the Software and algorithms section of the STAR Methods table, and all statistical details of experiments can be found in the figure legends, figures, results, and method details sections, including the statistical tests used, the number of replicate experiments, definition of center, dispersion, and precision measures, and the definition of significance. Experimental and control groups were defined by genotypes, no randomization was performed, and no data were excluded from the analysis.