Targeted cohesin loading characterizes the entry and exit sites of loop extrusion trajectories

Ruiqi Han[1]†, Yike Huang[1]†, Iwan Vaandrager[1], Amin Allahyar[1], Mikhail Magnitov[2], Marjon J.A.M. Verstegen[1], Elzo de Wit[2], Peter H.L. Krijger[1], Wouter de Laat[1]*

Affiliations:

[1]Oncode Institute, Hubrecht Institute-KNAW and University Medical Center Utrecht, Utrecht, the Netherlands.

[2]Division of Gene Regulation, Oncode Institute, The Netherlands Cancer Institute, Amsterdam, the Netherlands

†These authors contributed equally to this work

*Corresponding author. Email: w.laat@hubrecht.eu

The cohesin complex (SMC1-SMC3-RAD21) shapes chromosomes by DNA loop extrusion, but individual extrusion trajectories were so far unappreciable in vivo. Here, we site-specifically induced dozens of extrusion trajectories anchored at artificial loading sites in living cells. Extruding cohesin transports loading proteins MAU2-NIPBL over megabase DNA distances to blocking CTCF sites that then loop back to the loading sequences, showing that CTCF-CTCF interactions are unnecessary for stabilized contacts between loop extrusion obstacles. When stalled, cohesin can block other extruding cohesin from either direction. Without RAD21, MAU2-NIPBL exclusively accumulate at loading sites, here genome-wide defined as enhancers. SMC1 now also selectively accumulates here, suggesting that cohesin may load modularly on chromatin. Genes inside high cohesin extrusion trajectories are collectively hindered in transcription. This work characterizes the impact, entry and exit sites of individual cohesin loop extrusion trajectories.

Main text

The evolutionary conserved cohesin complex is a tripartite ring-shaped structure consisting of RAD21, SMC1 and SMC3, associated with STAG1/2. The complex functions to hold sister chromatids together during mitosis and shape chromosomes in interphase cells by sub-dividing them into topologically associating domains (TADs) [1–3]. Cohesin establishes TADs presumably through a loop extrusion process [4–6], in which it is loaded on chromatin and subsequently reels in flanking sequences to build progressively larger DNA loops in an ATP-dependent manner [7–11]. Cohesin cycles between a chromatin-bound and -extruding state and an unbound state [12–14], requiring NIPBL and MAU2 for stable chromatin association [15,16] and stimulating loop extrusion [17–19]. When reaching convergently oriented CTCF proteins that demarcate domain boundaries, cohesin is protected against release by WAPL [12,20–22], explaining why more stabilized chromatin loops between opposite domain boundaries are observed. How loop extrusion impacts transcription remains unclear, but recent evidence suggests that continuous cohesin-mediated loop extrusion is required for the regulation of developmental genes by distal enhancers [13,23–26]. Studying the impact of cohesin loop extrusion activity remains challenging though, as individual loop extrusion trajectories cannot be discerned in living cells. Furthermore, *in vivo* studies of cohesin largely rely on (acute) cohesin protein depletion, which may lead to widespread changes in chromatin structure and functioning, cell cycle arrest and cell death [1,12,13].

To accurately monitor direct consequences of altered loop extrusion activity, we developed the TArgeted Cohesin Loader (TACL), a genetic platform for site-specific initiation and manipulation of individual loop extrusion trajectories *in vivo*. TACL employs the Tet-off system with TetR fused to the cohesin loading factor MAU2/SCC4 to conditionally recruit cohesin and initiate loop extrusion trajectories from chromosomally integrated Tet operator sequences (Fig.1A). We utilized the PiggyBac transposon system [27] to create a human eHAP1 cell line with twenty-seven randomly inserted 50xTetO platforms (hereafter "TetO") across the haploid genome (fig. S1A). These cells were transduced with lentivirus to stably express TetR fused to FLAG-MAU2 (TACL) or to FLAG-mCherry (control) (Fig. 1B). ChIP-qPCR confirmed that

both fusion proteins were recruited to TetO and that TetR-MAU2 dissociated from the platforms by the treatment of doxycycline (Dox) for 1 hour (Fig. 1C). TetR-MAU2 not only co-recruited its binding partner NIPBL to TetO, but it also attracted cohesin, as seen before in yeast [28]. These endogenous factors disassociated rapidly from TetO upon Dox treatment (Fig. 1C). To test whether TACL initiated loop extrusion events from TetO, we performed ChIP-seq for SMC1 and RAD21. Notably, we observed increased SMC1 and RAD21 deposition at many endogenous sites in the vicinity of the TetO (Fig. 1D). This specific and local accumulation of cohesin was lost upon Dox treatment (Fig.1, D and E). ChIP experiments of other cohesin subunits STAG1 and STAG2 suggested that mainly STAG2 was co-loaded at TetO and deposited at their surrounding sites (Fig. S2, A and B). These suggested that TACL induced loop extrusion in the vicinity of the defined loading sites.

Figure 1. TACL induces cohesin loading and loop extrusion

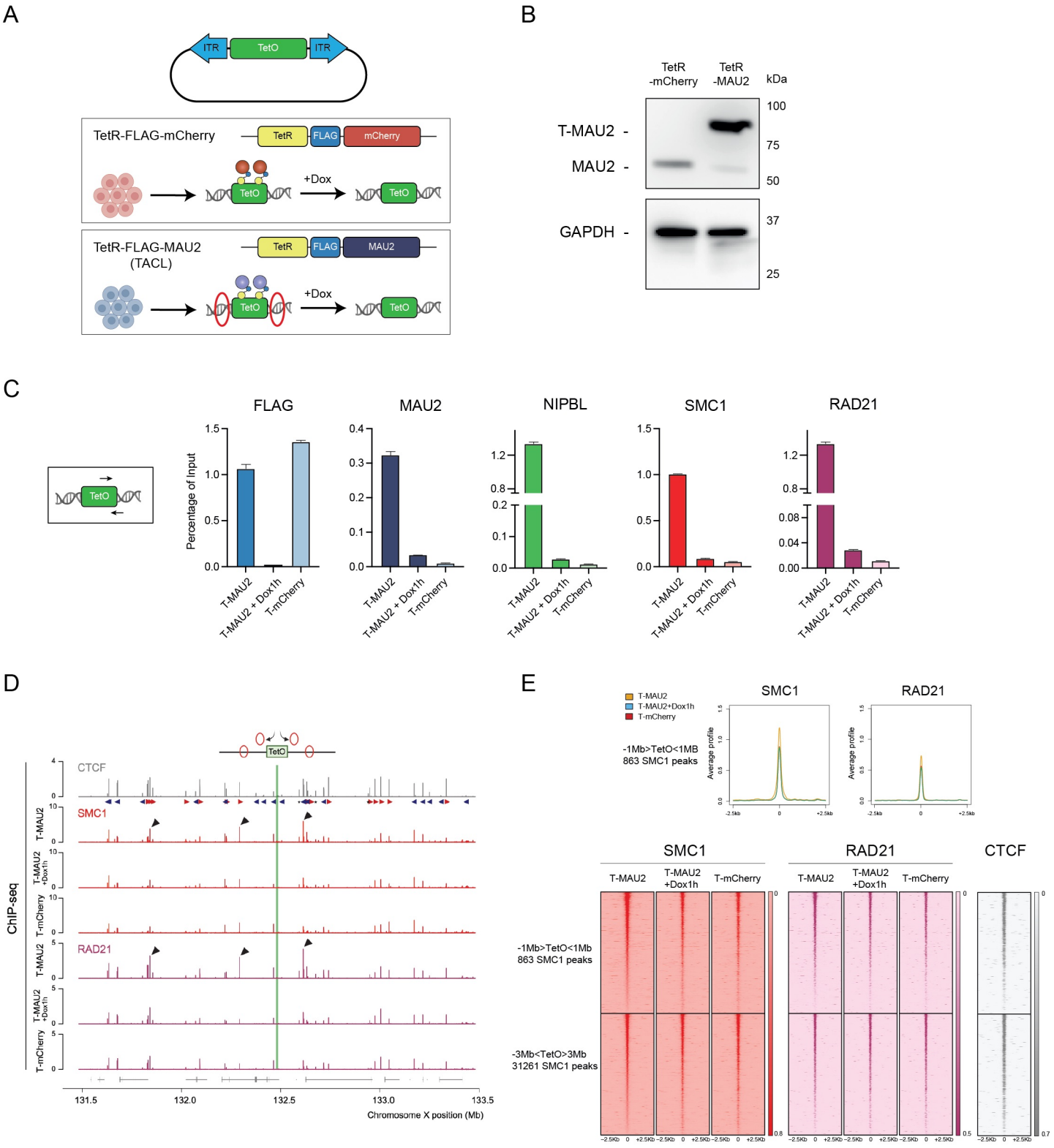
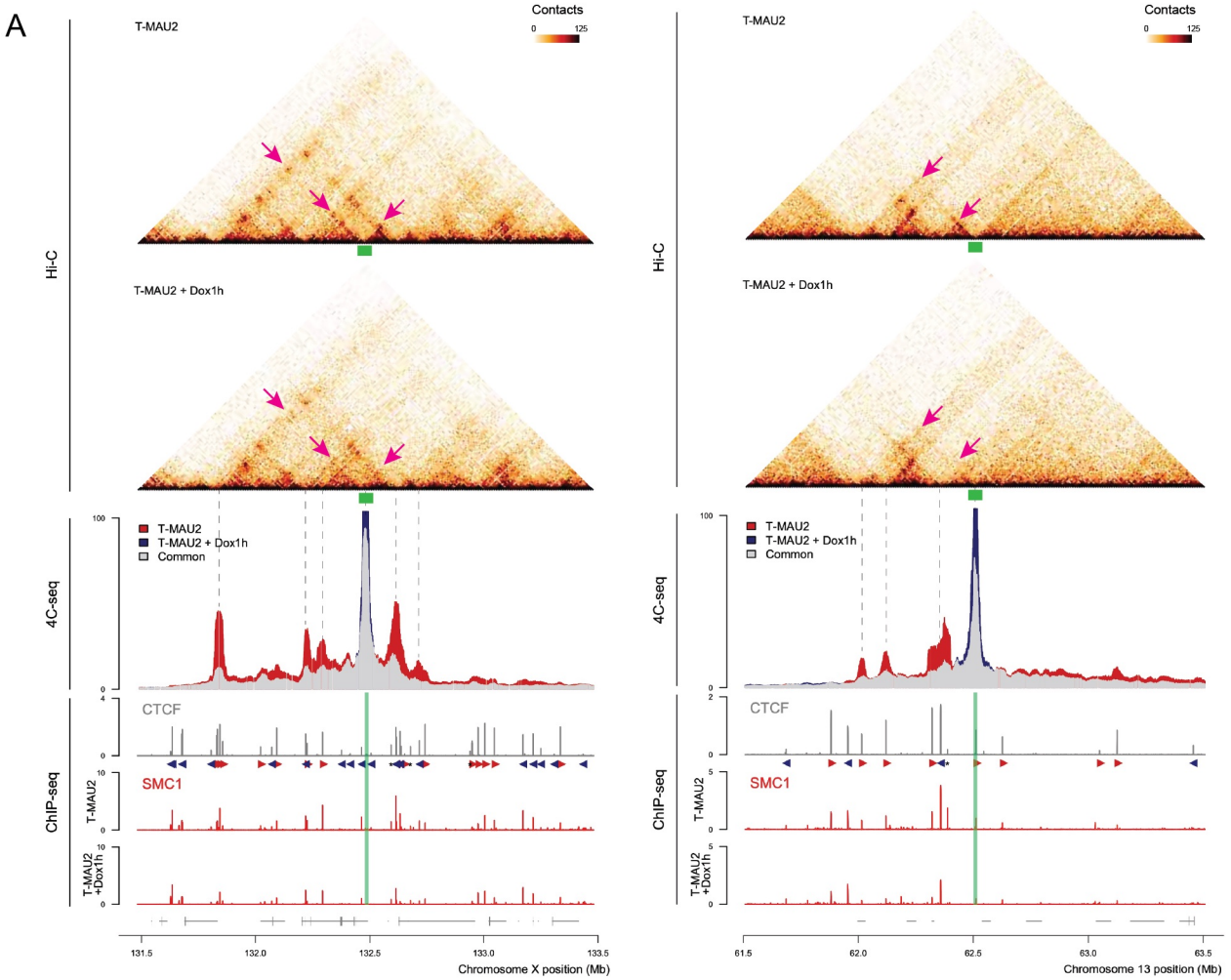


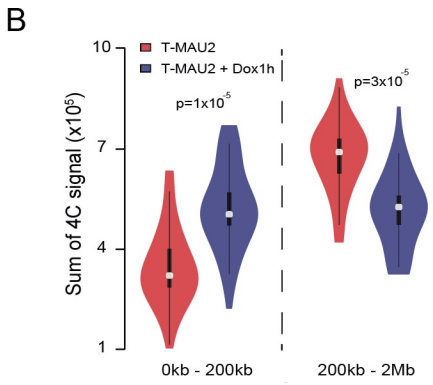
Fig. 1. TACL induces cohesin loading and loop extrusion. (A) Schematic illustration of the TACL system. **(B)** Western blot images with antibodies against MAU2 in T-mCherry and T-MAU2 cells. GAPDH serves as a loading control. **(C)** ChIP-qPCR analysis in the shown conditions with the corresponding antibodies. Values are shown in percentage of input (mean ± SD). At least three independent experiments were performed. **(D)** ChIP-seq profiles of CTCF, SMC1, and RAD21 ±1Mb from a TetO site, illustrating enrichment of cohesin proteins at CTCF sites (marked with arrows next to peaks). Green bar indicates the TetO location. **(E)** Average signal profiles and heatmaps of ChIP-seq signals centered at SMC1 peaks. **The upper panel** represents the average signal profiles of SMC1 and RAD21 enrichment at SMC1 peaks 2Mb surrounding all TetO sites for different conditions. **The lower panel** shows the heatmaps of SMC1, RAD21, and CTCF signals at SMC1 peaks within 2Mb or further than 3Mb away from all TetO sites. Profiles are shown ± 2.5kb around the peaks.

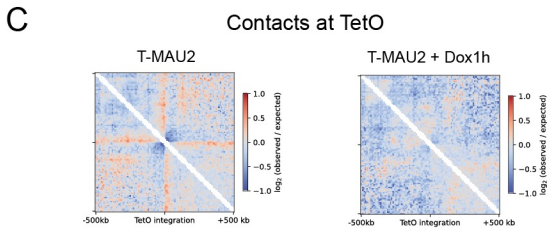
To test whether targeted cohesin recruitment via TACL could induce formation of local chromatin loops, we performed 4C-seq using TetO sequences as the viewpoint (VP), which allowed us to simultaneously assess contact profiles of all TetO locations. We observed that cohesin recruitment stimulated TetO to engage in long-range contacts (>200kb) at the expense of shorter-range contacts (Fig. 2, A and B, and fig. S3, A and B). Additionally, TACL activated many of the TetO platforms to form strong specific interactions with surrounding CTCF sites (Fig. 2A), suggesting that single entry sites enabled multiple consecutive convergent CTCF sites to engage in looping. The acquired topological contacts dismantled within 1 hour of Dox treatment, confirming that they were indeed induced by MAU2 recruitment to TetO (Fig. 2A and fig. S4). The TACL-stimulated, local, Dox-reversible, conformational changes to the chromatin surrounding TetO were further confirmed by Hi-C analysis. Stripes, not chromatin jets, were seen emerging in a bi-directional manner from multiple TetOs (Fig. 2, C and D and fig. S5A). Chromatin jets describe a recently identified Hi-C signature observed at some locally dominant cohesin loading sites that reflect bouquets of unanchored, bi-directionally extruding cohesin molecules that all initiated extrusion at the same

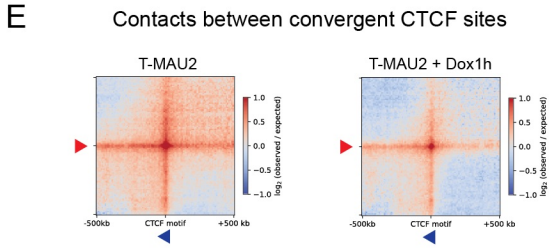
site [29]. In contrast, stripes are normally observed at strong CTCF boundaries [12,30] and are believed to reflect differently sized chromatin loops formed by uni-directional extruding cohesin molecules anchored at the same site. The high levels of cohesin associated with TetO may anchor an extruding cohesin molecule and only allow its one-sided extrusion, reeling in either upstream or downstream sequences. HiC meta-analysis showed that TetO indeed acquired boundary capacity, insulating up- from downstream sequences (Fig. 2D and fig. S5A). In some instances, the recruitment of extruding cohesin not only stimulated chromatin loops between TetO and endogenous CTCF sites, but also increased the contacts between pairs of endogenous CTCF sites (Fig. 2, A and E and fig. S5B). This might suggest that not all recruited cohesin complexes stayed anchored to TetO, but could also diffuse or extrude away to anchor and extrude loops between neighboring CTCF sites.

Figure 2 | Local stimulation of cohesin loop extrusion activity induces topological changes









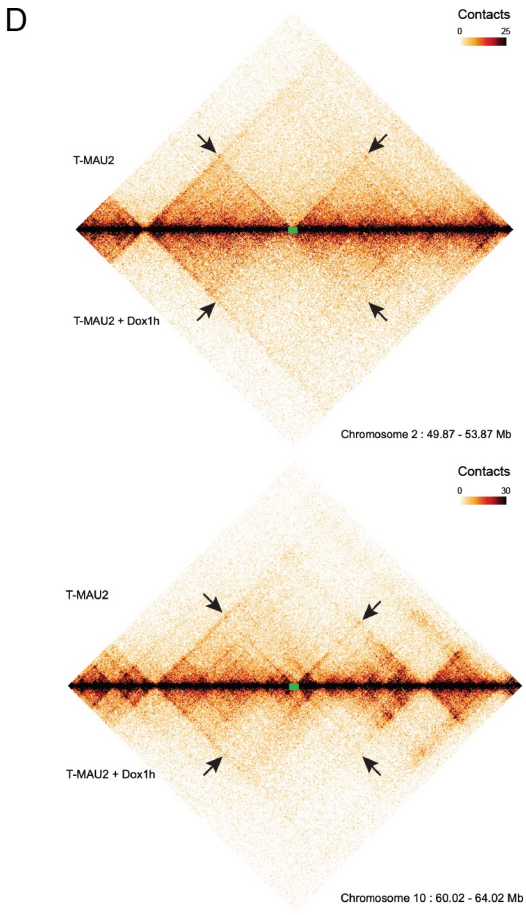


Fig. 2. Local stimulation of cohesin loop extrusion activity induces topological changes. (A) Hi-C, 4C-seq, and ChIP-seq overlays showing increased contacts induced by TACL. Increased contacts are indicated with magenta arrows on Hi-C maps, dashed line on 4C-seq profiles. Green bars at the center of the profile depict the TetO locations. **(B)** Sum of 4C signal within 0kb-200kb and 200kb-2Mb in T-MAU2 and T-MAU2+Dox1h conditions. P-values are obtained from two-sided t-test. **(C)** Aggregate contact analysis centered at TetO showing stripes emerging from TetO. **(D)** Hi-C profiles of two examples of stripe formation from TetO. Green bars mark TetO locations and black arrows indicate stripes in T-MAU2 cells. **(E)** Aggregate contact analysis centered at convergent CTCF sites in the vicinity of TetO.

Recent *in vitro* experiments suggested that cohesin needs to associate with NIPBL-MAU2 to act as an active loop-extruding holo-enzyme [17,18]. Here we searched for evidence that NIPBL and MAU2 co-traveled with cohesin from loading sites to stalling sites. NIPBL and MAU2 ChIP-seq showed that they strongly and specifically accumulated at sites surrounding TetO (Fig. 3, A and B). TetR-MAU2 carried the FLAG epitope to help distinguishing TACL-induced from endogenously induced MAU2 binding events. We observed that only in cells expressing TetR-FLAG-MAU2, but not TetR-FLAG-mCherry, FLAG was found deposited at these same sites surrounding TetO (Fig. 3, A and B, and fig. S6). Local deposition of TetR-FLAG-MAU2 was Dox-responsive, showing that NIPBL and MAU2 required prior loading onto TetO to accumulate at flanking sites (Fig. 3, A and B). Flanking sites collecting TACL-dependent MAU2 and NIPBL were nearly always pre-existing CTCF sites that also naturally halted cohesin in wildtype cells. In addition, some of these sites marked the anchors of TACL-induced chromatin loops with TetO, as detected by 4C-seq (fig. S4). Detailed inspection of the binding events revealed that induced deposition of MAU2, NIPBL and cohesin specifically occurred at the TetO-facing sides of convergently oriented CTCF sites, as expected from a direct interaction between cohesin and CTCF [22]. At the strongest binding sites, the ChIP-seq binding patterns, particularly of NIPBL, revealed what seemed to be the queuing of multiple extruding cohesin holoenzymes in front of CTCF sites (Fig. 3C and fig. S7, upper panel) reminiscent of cohesin

'traffic jams' [31]. TACL-induced extruding cohesin and auxiliary factors were also seen to accumulate at the 'illegal' side (C-terminus) of divergently oriented CTCF molecules loaded on their 'legal' side (N-terminus) with cohesin that presumably came from more distal sequences (Fig. 3C and fig. S7, lower panel). This strongly suggested that *in vivo*, extruding cohesin complexes are often stalled when encountering another stalled cohesin complex. Plotting the chromosomal distances of FLAG peaks to their nearest TetO revealed that cohesin loop extrusion trajectories can span up to two megabases of chromatin (Fig. 3D). Taken together, we propose that NIPBL and MAU2 load cohesin on chromatin, whereby cohesin associates with and transports these auxiliary factors while extruding loops across up to two megabases of DNA. Loop extrusion can be blocked when cohesin encounters chromatin-bound convergently oriented CTCF molecules, but also when cohesin encounters another stalled cohesin complex.

Figure 3. Extruding cohesin transports MAU2 and NIPBL along the chromatin fiber

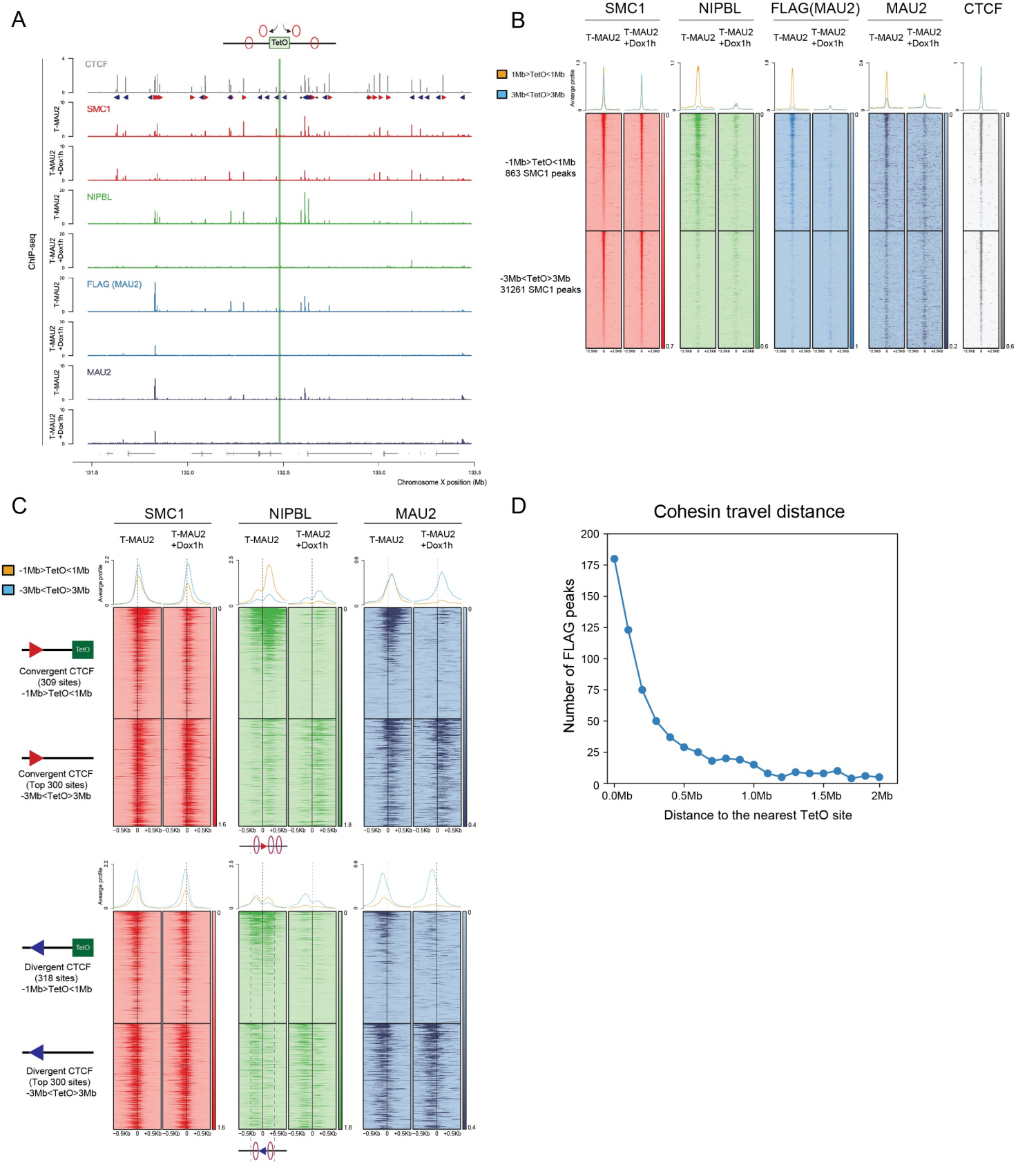


Fig. 3. Extruding cohesin transports MAU2 and NIPBL along the chromatin fiber. (A) ChIP-seq profiles of CTCF, SMC1, NIPBL, FLAG(MAU2), and MAU2 in a ±1Mb region from a TetO site, illustrating co-occupied sites by SMC1, NIPBL, FLAG(MAU2), and MAU2. Green bar indicates the TetO location. **(B)** Average signal profiles and heatmaps of SMC1, NIPBL, FLAG(MAU2), MAU2, and CTCF signals at SMC1 peaks ±1Mb or further than 3Mb away from all TetO sites. Profiles are shown ± 2.5kb around the peaks. **(C)** Average signal profiles and heatmaps of SMC1, NIPBL, and MAU2 at SMC1 peaks within ±1Mb or further than 3Mb away from all TetO sites. Profiles are shown ± 0.5kb around the peaks. CTCF sites are separated based on their relative orientation to TetO: convergent CTCF are all CTCF sites facing TetO; divergent CTCF are all CTCF sites facing the outside of TetO. Different orientations (relative to the genome) within the same category are flipped to the same direction. Signals are ranked by FLAG peak strength in T-MAU2 cells. **(D)** The number of FLAG peaks within a 2Mb window surrounding TetO, representing the travel distance of TACL-loaded cohesin.

To further investigate whether NIPBL-MAU2 deposition to flanking CTCF sites was the consequence of their "hitchhiking" with extruding cohesin complexes, we knocked-in an auxin-inducible degron (AID2) to deplete endogenous RAD21 in the TACL and control cells (Fig.4A). Treating the cells for two hours with 5-Ph-IAA (IAA) completely removed RAD21 protein in the cells (fig. S8). ChIP-qPCR confirmed that both cohesin subunits SMC1 and SMC3 disassociated from TetO (Fig. 4B). Consequently, TACL-induced topological contacts dismantled, and TetR-MAU2 and NIPBL no longer accumulated at TetO-flanking sequences (Fig. 4, C and D, and fig. S9 and S10). Yet, MAU2 and NIPBL remained bound to the TetO (Fig.4, B and D, and fig. S9), suggesting that these auxiliary factors can stably associate with loading sites independent of cohesin but required cohesin and loop extrusion to localize at CTCF sites. To further investigate such cohesin loading sites, we compared the genome-wide binding patterns of NIPBL-MAU2 in the presence and absence of RAD21. The great majority of NIPBL and MAU2 binding sites disappeared upon IAA treatment, but a small percentage (<17%) of the sites retained or even more strongly accumulated

both proteins in the absence of RAD21. Neither RAD21 nor SMC3 was present at these retained sites (fig. S11 and S12, category III, IV, and V), but when analyzing the genome-wide binding patterns of SMC1 in RAD21-depleted cells, we observed that SMC1 remained associated with the same sites that also recruited NIPBL-MAU2 in a RAD21-dependent manner (Fig. 4E and fig. S11, categories III, IV and V). The overexpression of TetR-MAU2, we noticed, stimulated additional loading of both NIPBL-MAU2 onto chromatin, particularly at these same sites that retained the factors in the absence of RAD21 sites (Fig. 4E and fig. S11, category III, IV, and V). The same sites were also identified as RAD21-independent NIPBL and SMC1 binding sites in the TetR-mCherry control cells, though having a lower ChIP-signal. (fig. S12).

Upon deeper inspection of the RAD21-dependent and -independent binding sites for MAU2, NIPBL and SMC1, we observed that the dependent sites typically were CTCF binding sites lacking active enhancer (H3K27Ac) and active promoter (H3K4me3) marks (category I, 16025 sites). In contrast, the independent sites typically lacked CTCF (with exception of category II, 173 sites). Twenty-six percent of the independent sites (785 sites: category III) lacked any of these marks, but a large percentage (69%, 2116/3074) could be classified as active enhancer (1021, category IV) or active promoters (1095, category V) (Fig. 4E and fig. S11). This, we believe, is direct demonstration that active enhancers and active promoters serve as entry, and CTCF sites as exit or pause sites of extruding cohesin complexes. WAPL is known to release cohesin from CTCF bound genomic locations. Indeed, the RAD21-dependent association of SMC1 to CTCF binding sites was previously shown to be WAPL-independent, while oppositely, its RAD21-independent recruitment to enhancers was previously found to be WAPL-dependent[13]. Our results also suggested that cohesin can build up modularly at its entry sites: SMC1 association to chromosomes can occur independent of RAD21 and SMC3, but the tripartite cohesin complex is required for loop extrusion and the migration of auxiliary factors MAU2 and NIPBL along chromosomes to flanking CTCF sites.

Figure 4. Exploring the cohesin loading dynamics with TAC!

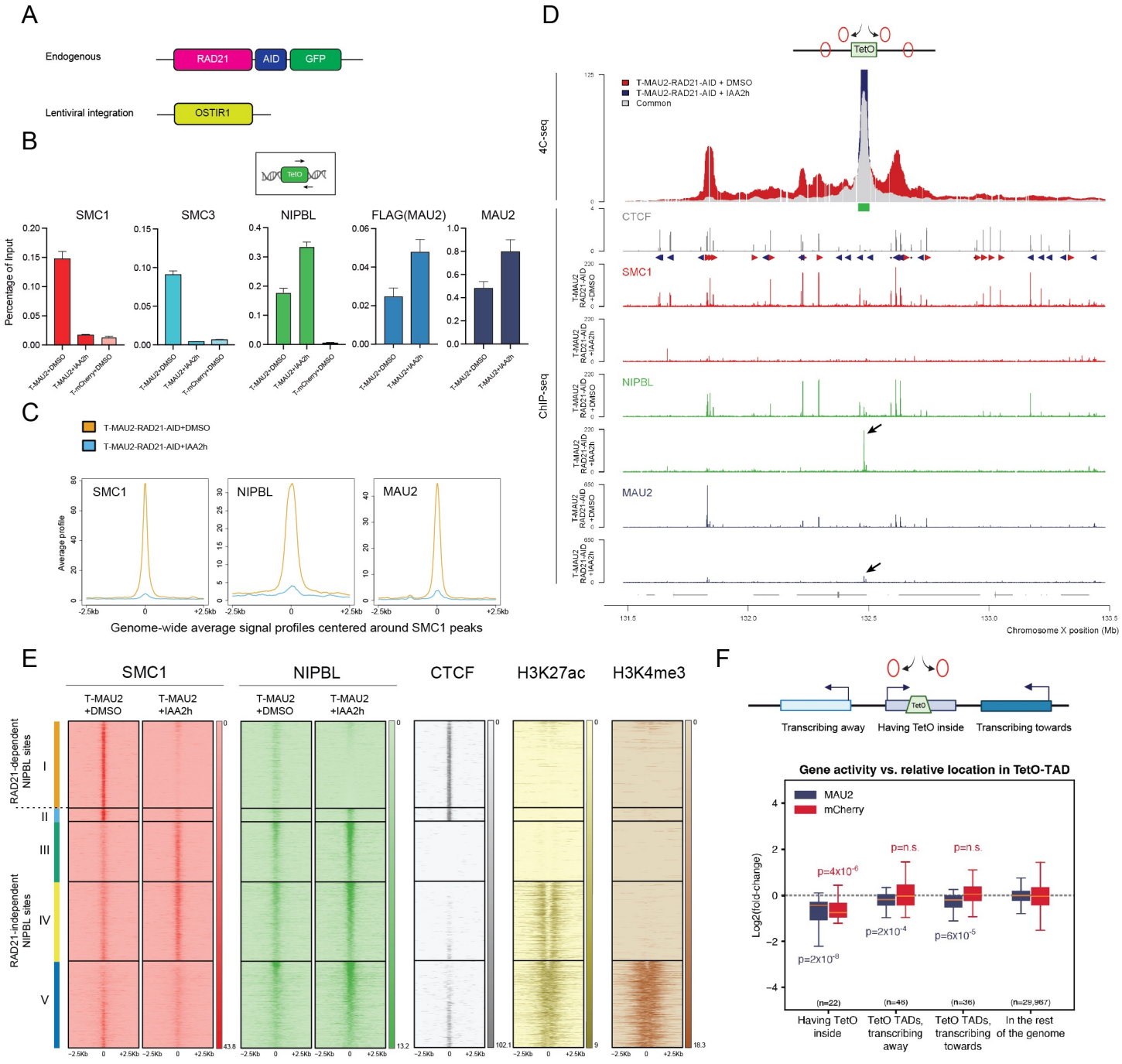


Fig. 4. Exploring the cohesin loading dynamics with TACL. (A) An illustration of the construction of RAD21-AID cells. **(B)** ChIP-qPCR analysis in the shown conditions with the corresponding antibodies. Values are shown in percentage of input (mean ± SD). At least two independent experiments were performed. **(C)** Genome-wide average signal profiles of SMC1, NIPBL, and MAU2 ChIP-seq signals ±IAA centered around SMC1 peaks in a ±2.5kb window. **(D)** 4C-seq and ChIP-seq profiles of T-MAU2 cells depleted of RAD21at a TetO location. The green bar indicates the TetO location. Enrichment of NIPBL and MAU2 at TetO after RAD21 depletion are depicted with black arrows. **(E)** Heatmaps of SMC1, NIPBL, CTCF, H3K27ac, and H3K4me3 ChIP-seq signals centered at SMC1/NIPBL peaks > 3Mb from TetO. Sites are divided into different categories by the presence of NIPBL, H3K27ac, and H3K4me3 after RAD21 depletion. **(F)** Gene activities within the TAD of TetO locations. The upper panel illustrates the three different categories of genes used in this study. The lower panel shows the transcription activities of the genes in the three groups compared to other genes in the rest of the genome (control). Y-axis represents the log2 fold change of gene activities between T-MAU2/T-mCherry cells and the cells treated with Dox for 1h. n represents the number of genes within each category and p-values are calculated between the indicated group of genes with its corresponding control group (genes in the rest of the genome).

To investigate the impact of directional encounters between extruding cohesin and the transcription machinery, we measured nascent transcription before and after Dox treatment for one hour. As expected from PiggyBac insertions [32], twenty-two of the twenty-seven TetO integrations were located inside active genes. The 27 hosting TADs further had 36 intact active genes transcribing towards, and 46 transcribing away from the TetO. Genes carrying TetO inside their gene bodies showed reduced transcription activity, irrespective of whether TetR-mCherry or TetR-MAU2 was accumulated at the TetO (Fig. 4F). Flanking genes elsewhere in the same TADs did not respond to active TetR-mCherry recruitment. In contrast, they collectively showed a small but significant decrease in transcription activity in response to TetR-MAU2-stimulated cohesin loop extrusion. Reduced transcription activities were independent of transcriptional

direction and thus independent of whether RNAPII tracked along DNA in the same or opposite direction of extruding cohesin machinery (Fig. 4F, fig. S13). These suggested that transcription by RNAPII is hindered, but not blocked, upon bypassing of extruding cohesin complexes.

In summary, by creating a system for targeted recruitment of extruding cohesin complexes in living cells, we were able to define and characterize the starts and ends of cohesin loop extrusion trajectories. Chromatin entry of cohesin predominantly takes place at enhancers and active promoters that associate with NIPBL and MAU2 to recruit SMC1, which has been shown previously to directly interact through its hinge domain with NIPBL [33]. SMC1 can associate without SMC3 and RAD21, suggesting that the cohesin complex can be built up on chromatin from individual components. Formation of the tripartite cohesin holocomplex is needed to initiate chromatin loop extrusion and extruding cohesin is responsible for longitudinal transport of NIPBL and MAU2 from their binding sites to other chromosomal locations. Chromatin-bound convergently oriented CTCF, but also, as previously suggested [34,35], cohesin-loaded divergently oriented CTCF molecules, can block extrusion, resulting in localized deposition of co-migrating NIPBL and MAU2. Thus, previously observed binding of NIPBL to CTCF sites [12,36] generally marks the end of cohesin extrusion trajectories, not their start. We showed that activated loop extrusion across genes negatively impacts, but not blocks, their transcriptional activities, no matter whether transcription proceeds towards or away from the loop extrusion initiation site.

References:

1. Rao, S. S. P. *et al.* Cohesin Loss Eliminates All Loop Domains. *Cell* (2017) doi:10.1016/j.cell.2017.09.026.
2. Schwarzer, W. *et al.* Two independent modes of chromatin organization revealed by cohesin removal. *Nature* **551**, (2017).
3. Wutz, G. *et al.* Topologically associating domains and chromatin loops depend on cohesin and are regulated by CTCF, WAPL, and PDS5 proteins. *EMBO J* **36**, (2017).
4. Sanborn, A. L. *et al.* Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. *Proc Natl Acad Sci U S A* **112**, (2015).
5. Fudenberg, G. *et al.* Formation of Chromosomal Domains by Loop Extrusion. *Cell Rep* (2016) doi:10.1016/j.celrep.2016.04.085.
6. Ganji, M. *et al.* Real-time imaging of DNA loop extrusion by condensin. *Science (1979)* (2018) doi:10.1126/science.aar7831.
7. Marko, J. F., de Los Rios, P., Barducci, A. & Gruber, S. DNA-segment-capture model for loop extrusion by structural maintenance of chromosome (SMC) protein complexes. *Nucleic Acids Res* **47**, (2019).
8. Ryu, J. K. *et al.* The condensin holocomplex cycles dynamically between open and collapsed states. *Nat Struct Mol Biol* **27**, (2020).
9. Shaltiel, I. A. *et al.* A hold-and-feed mechanism drives directional DNA loop extrusion by condensin. *Science (1979)* **376**, (2022).
10. Fudenberg, G., Abdennur, N., Imakaev, M., Goloborodko, A. & Mirny, L. A. Emerging Evidence of Chromosome Folding by Loop Extrusion. *Cold Spring Harb Symp Quant Biol* **82**, (2017).
11. Davidson, I. F. & Peters, J. M. Genome folding through loop extrusion by SMC complexes. *Nature Reviews Molecular Cell Biology* vol. 22 Preprint at https://doi.org/10.1038/s41580-021-00349-7 (2021).
12. Vian, L. *et al.* The Energetics and Physiological Impact of Cohesin Extrusion. *Cell* **173**, (2018).
13. Liu, N. Q. *et al.* WAPL maintains a cohesin loading cycle to preserve cell-type-specific distal gene regulation. *Nat Genet* **53**, (2021).
14. Mirny, L. & Dekker, J. Mechanisms of Chromosome Folding and Nuclear Organization: Their Interplay and Open Questions. *Cold Spring Harb Perspect Biol* **14**, (2022).
15. Ciosk, R. *et al.* Cohesin's binding to chromosomes depends on a separate complex consisting of Scc2 and Scc4 proteins. *Mol Cell* **5**, (2000).
16. Petela, N. J. *et al.* Folding of cohesin's coiled coil is important for scc2/4-induced association with chromosomes. *Elife* **10**, (2021).
17. Davidson, I. F. *et al.* DNA loop extrusion by human cohesin. *Science* (2019) doi:10.1126/science.aaz3418.
18. Kim, Y., Shi, Z., Zhang, H., Finkelstein, I. J. & Yu, H. Human cohesin compacts DNA by loop extrusion. *Science (1979)* **366**, (2019).
19. Bauer, B. W. *et al.* Cohesin mediates DNA loop extrusion by a "swing and clamp" mechanism. *Cell* **184**, (2021).
20. Rao, S. S. P. *et al.* A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping. *Cell* **159**, (2014).
21. de Wit, E. *et al.* CTCF Binding Polarity Determines Chromatin Looping. *Mol Cell* (2015) doi:10.1016/j.molcel.2015.09.023.
22. Li, Y. *et al.* The structural basis for cohesin–CTCF-anchored loops. *Nature* **578**, (2020).
23. Thiecke, M. J. *et al.* Cohesin-Dependent and -Independent Mechanisms Mediate Chromosomal Contacts between Promoters and Enhancers. *Cell Rep* **32**, (2020).
24. Calderon, L. *et al.* Cohesin-dependence of neuronal gene expression relates to chromatin loop length. *Elife* **11**, e76539 (2022).
25. Rinzema, N. J. *et al.* Building regulatory landscapes reveals that an enhancer can recruit cohesin to create contact domains, engage CTCF sites and activate distant genes. *Nat Struct Mol Biol* **29**, 563–574 (2022).
26. Kane, L. *et al.* Cohesin is required for long-range enhancer action at the Shh locus. *Nat Struct Mol Biol* **29**, 891–897 (2022).
27. Redolfi, J. *et al.* DamC reveals principles of chromatin folding in vivo without crosslinking and ligation. *Nat Struct Mol Biol* (2019) doi:10.1038/s41594-019-0231-0.
28. Petela, N. J. *et al.* Scc2 Is a Potent Activator of Cohesin's ATPase that Promotes Loading by Binding Scc1 without Pds5. *Mol Cell* **70**, (2018).
29. Guo, Y. *et al.* Chromatin jets define the properties of cohesin-driven in vivo loop extrusion. *Mol Cell* **82**, 3769-3780.e5 (2022).

30. Vietri Rudan, M. *et al.* Comparative Hi-C Reveals that CTCF Underlies Evolution of Chromosomal Domain Architecture. *Cell Rep* **10**, (2015).
31. Allahyar, A. *et al.* Enhancer hubs and loop collisions identified from single-allele topologies. *Nat Genet* (2018) doi:10.1038/s41588-018-0161-5.
32. Galvan, D. L. *et al.* Genome-wide mapping of piggybac transposon integrations in primary human T cells. *Journal of Immunotherapy* **32**, (2009).
33. Hu, B. *et al.* ATP hydrolysis is required for relocating cohesin from sites occupied by its Scc2/4 loading complex. *Current Biology* **21**, (2011).
34. Liu, N. Q. *et al.* Rapid depletion of CTCF and cohesin proteins reveals dynamic features of chromosome architecture. *bioRxiv* (2021).
35. Anania, C. *et al.* In vivo dissection of a clustered-CTCF domain boundary reveals developmental principles of regulatory insulation. *Nat Genet* **54**, (2022).
36. Banigan, E. J. *et al.* Transcription shapes 3D chromatin organization by interacting with loop-extruding cohesin complexes. *bioRxiv* 2022.01.07.475367 (2022).
37. van de Werken, H. J. G. *et al.* Robust 4C-seq data analysis to screen for regulatory DNA interactions. *Nat Methods* **9**, (2012).
38. Krijger, P. H. L., Geeven, G., Bianchi, V., Hilvering, C. R. E. & de Laat, W. 4C-seq from beginning to end: A detailed protocol for sample preparation and data analysis. *Methods* **170**, (2020).
39. Vermeulen, C. *et al.* Multi-contact 4C: long-molecule sequencing of complex proximity ligation products to uncover local cooperative and competitive chromatin topologies. *Nat Protoc* **15**, (2020).
40. Roberts, T. C. *et al.* Quantification of nascent transcription by bromouridine immunocapture nuclear run-on RT-qPCR. *Nat Protoc* **10**, (2015).
41. Yesbolatova, A. *et al.* The auxin-inducible degron 2 technology provides sharp degradation control in yeast, mammalian cells, and mice. *Nat Commun* **11**, (2020).
42. Lawrence, M., Gentleman, R. & Carey, V. rtracklayer: An R package for interfacing with genome browsers. *Bioinformatics* **25**, (2009).
43. Lawrence, M. *et al.* Software for Computing and Annotating Genomic Ranges. *PLoS Comput Biol* **9**, (2013).
44. Grant, C. E., Bailey, T. L. & Noble, W. S. FIMO: Scanning for occurrences of a given motif. *Bioinformatics* **27**, 1017–1018 (2011).
45. Khan, A. *et al.* JASPAR 2018: update of the open-access database of transcription factor binding profiles and its web framework. *Nucleic Acids Res* (2017) doi:10.1093/nar/gkx1126.
46. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, (2014).
47. Abdennur, N. & Mirny, L. A. Cooler: Scalable storage for Hi-C data and other genomically labeled arrays. *Bioinformatics* **36**, (2020).
48. Abdennur, N. *et al.* Cooltools: enabling high-resolution Hi-C analysis in Python. *bioRxiv* **2022.10.31.514564**, (2022).
49. Abdennur, N. *et al.* Bioframe: Operations on Genomic Intervals in Pandas Dataframes. *bioRxiv* (2022).
50. Durand, N. C. *et al.* Juicebox Provides a Visualization System for Hi-C Contact Maps with Unlimited Zoom. *Cell Syst* **3**, (2016).
51. Gilleland, E. Two-dimensional kernel smoothing: Using the R package smoothie. *NCAR Technical Note* (2013).

Methods

Cell culture

eHAP1 cells were cultured in Iscove's Modified Dulbecco's Medium (IMDM) medium supplemented with Glutamax (Thermofisher), 25mM Hepes, 10% FBS, and 1% Penicillin-Streptomycin following standard procedures. Cells were routinely checked and sorted for haploidy. 293TX cells were cultured in Dulbecco's Modified Eagle Medium (DMEM) medium supplemented with 10% FBS and 1% Penicillin-Streptomycin.

Antibodies

Anti-SMC1 (A300-055A, Bethyl), anti-SMC3 (A300-060A, Bethyl), anti-RAD21 (05-908, Merck), anti-NIPBL (A301-779A, Bethyl), anti-FLAG (F1804, Merck), anti-SCC4/MAU2 (ab183033, Abcam), anti-STAG1 (A302-579A, Bethyl), anti-STAG2 (A300-159A, Bethyl).

Plasmid construction

The plasmids expressing TetR-FLAG-MAU2 and TetR-FLAG-mCherry cassette was cloned into a Lentivirus backbone under the control of EF1 promoter. TetR, FLAG, and MAU2/mCherry sequences were PCR amplified with 20bp overhang for In-Fusion cloning. The final expression cassette is composed of EF1-TetR-FLAG-MAU2/mCherry-P2A-Puromycin. To insert the auxin inducible degron (AID2) tag into the endogenous RAD21 gene, a sgRNA (CCAAGGTTCCATATTATATA) targeting the C-terminus of RAD21 was cloned into a vector containing SpCas9-T2A-BFP. To construct the donor template for AID2 tag insertion, a cassette containing AID2-GFP was cloned between two homology arms of about 1kb surrounding the sgRNA cut site.

Generation of cell lines containing the TetO platforms

The plasmids bearing the TetO platforms and the piggybac transposase were originally obtained from Luca Giorgetti [27]. Briefly, eHap1 cells were trypsinized and resuspended in serum-free IMDM medium. A vector containing the piggybac transposase (pBroad3_hyPBase_IRES_tagRFPt) were mixed with a piggybac donor vector bearing 50x TetO binding sites and Polyethylenimine (PEI, polysciences) in serum-free IMDM. The DNA mix was incubated at room temperature for 10min, after which the cells and the DNA mix were incubated together for another 10min. The cells were then plated in a 6-well plate. After 24h, the medium was refreshed. 48-72h after the transfection, the cells were sorted for RFP signal, expressing the transposase. Sorted cells were plated in 15cm dish and cultured for at least 14 days. Colonies were picked and sub-cultured in 96-well plates. To genotype the clones with a sufficient number of integration sites, cells were lysed in DirectPCR lysis reagent (Viagen). Lyastes were subsequently assessed by running qPCR with primers annealing to the transposon sequences. A primer targeting a part of human FSIP2 gene was used as the reference among different clones. An estimation of the number of integration sites was calculated as: 2^- (Ct_TetO primer – Ct_reference). The exact number of integrations sites was validated by 4C-seq.

Lentivirus production and transduction

$4x10^6$ 293TX cells were plated in 10cm dish 24h prior virus production. Lentiviral vectors were co-transfected with pVSV-G, pMDL RRE, and pRSV-REV in serum-free DMEM with Polyethylenimine (PEI, polysciences). The medium was refreshed 18h after transfection. The medium containing the virus particles were harvested 48h after transfection by passing through a 0.45µm filter. For transduction, eHap1 cells were plated in a 6-well plate 24h before transduction. The transduction was performed by adding the virus particle directly onto the cells supplemented with 6µg/mL polybrene (Merck). The cells were refreshed 24h after transduction and the antibiotics (puromycin/blasticidin) were added 48h after transduction. Cells were selected with antibiotics until the cells in the control plate (without transduction) were completely dead.

Western blot

Cells were washed in PBS and lysed in RIPA buffer with protease inhibitor (Roche) on ice for 15min. The cell lysate was further disrupted by sonication with Bioruptor Pico (Diagnode). The cell lysate was cleared by spinning at 1000xg for 5min. The supernatant was incubated with Laemmli buffer and boiled for 10min. The sample was then loaded on a 4–15% Mini-PROTEAN® TGX™ Precast Protein Gel (Biorad) and ran at 100V for 90min. Proteins were transferred onto a nitrocellulose or PVDF membrane and incubated with the primary antibody overnight at 4°C. The membrane was then washed in PBS- 0.25% Tween and incubated with the secondary antibody at room temperature for 1h. Finally, the membrane was incubated with SuperSignal™ West Pico PLUS Chemiluminescent Substrate (Thermofisher) for 1min before visualized on ImageQuant 800 imager (Amersham).

Chromatin immunoprecipitation (ChIP)

100 million cells were crosslinked with 1% formaldehyde for 10min. Cells were subsequently quenched with 125mM glycine for 10min and washed twice with cold PBS. Cells were scraped from culture dishes and cell pellets were subsequently lysed in LB1 buffer (50mM Hepes, 140mM NaCl, 1mM EDTA, 10% glycerol, 0.5% NP40, 0.25% Triton X-100), washed in LB2 buffer (10mM Tris, 200mM NaCl, 1mM EDTA, 0.5mM EGTA), and resuspended in LB3 buffer (10mM Tris, 100mM NaCl, 1mM EDTA, 0.5mM EGTA, 0.1% sodium deoxycholate, 0.5% N-lauroylsarcosine) prior sonication. Chromatin was sonicated using Bioruptor Pico (Diagnode) with a setting of 30s on, 30s off for 8 cycles. Fragmented chromatin was then incubated with 6ug of antibodies pre-coupled to Dynabeads protein G beads (Thermofisher) overnight at 4°C. Beads-bound chromatin was then washed 10x with RIPA buffer (50mM Hepes, 500mM LiCl, 1mM EDTA, 1% NP40, 0.7% sodium deoxycholate), once with TBS buffer, and decrosslinked in elution buffer (50mM Tris, 10mM EDTA, 1% SDS) at 65°C for 18h. Eluted DNA was then treated with protease K and RNAse A, and subsequently purified with phenol/chloroform/isoamyl alcohol 25:24:1. Purified DNA was either assessed with qPCR or continued with ChIP-seq NGS sequencing library preparation. Sequencing libraries was constructed using NEBnext Ultra II DNA library prep kit (NEB) following the manufacture's protocol. Briefly, DNA was end-repaired and poly-A tailed, ligated to NEBnext adapters, and digested with USER enzyme. Annealed libraries were then purified with AMPure XP beads (Beckman Coulter) and PCR amplified with indexing primers for 4-12 cycles. Sequencing

libraries were checked with Bioanalyzer HS DNA chip (Agilent) and sequenced on the Illumina Nextseq 500 (single end reads, 75bp) and Nextseq 2000 platforms (pair end reads, 50bp).

4C-seq

4C templates preparation was performed as described in van de *Werken et al.* [37] with modifications as described by Krijger et al [38]. In brief, ten million cells per sample were crosslinked with 2% formaldehyde, followed by quenching by glycine at final concentration of 0.125 M. Four-cutter restriction enzyme MboI (New England BioLabs) was used for in situ digestion (300U/10 million cells). Digested DNA fragments were ligated, reverse-crosslinked and subsequently purified through isopropanol and magnetic beads (Macherey-Nagel NucleoMag PCR Beads). Four-cutter restriction enzyme Csp6I (CviQI, ThermoFisher ER0211, 50U/sample) was used for template trimming. Re-ligated and purified 4C templates were further proceed through *in vitro* Cas9 digestion as described below.

In vitro Cas9 digestion of 4C templates

To prevent PCR amplification and sequencing of TetO repeats due to tandem ligation of two or more TetO DpnII fragments in a given 4C circle, an *in vitro* digestion of 4C templates was performed as described in [39] with the following modifications: two sgRNA were used to target Cas9 into the TetO repeats between viewpoint primers; pre-incubation of Cas9 protein and sgRNA template at room temperature. In brief, two sgRNA templates were obtained using the Megashortscript T7 transcription kit (Invitrogen) followed by 4x AMPure XP (Agencourt) purification. Purified Cas9 protein (generated by Hubrecht protein facility) was pre-incubated with the sgRNAs for 30min at room temperature. 4C templates were subsequently added to the pre-incubated Cas9/sgRNA complexed for overnight digestion at 37°C. Cas9 protein was inactivated by incubating at 70 °C for 5 min. Resulting products were purified with 1x AMPure XP and used as PCR template for TetO dedicated 4C.

Nascent RNA sequencing (BrU-seq)

BrU-seq was performed based on the protocol from *Roberts et al* [40]. Cultured cells were incubated with 2mM Bromouridine (BrU, Merck) for 10min and subsequently lysed in TRIzol reagent (Thermofisher). RNA was isolated following the manufacture's protocol. Briefly, lysed cells were mixed with chloroform and centrifuged for 15min. The aqueous phase was transferred to a new tube and mixed with isopropanol. After centrifugation, RNA pellet was washed once with 70% ethanol and dissolved in DEPC water. To capture BrU-labelled nascent RNA, 6ug anti-BrdU antibodies (BD biosciences) pre-coupled with Dynabeads protein G beads (Thermofisher) were incubated with the total RNA for 1h at room temperature. Beads were then washed 3x with PBS/0.1% Tween-20/RNaseOUT. To purify the beads-bound RNA, TRIzol reagent was directly added to the beads and RNA was purified as described above. NGS sequencing libraries were generated using NEBnext Ultra II directional RNA library prep kit (NEB) following manufacture's protocol. Briefly, RNA was fragmented to about 200bp in size. First strand and second strand cDNA were synthesized. Double strand cDNA was then end repaired, poly-A tailed, ligated to NEBnext adapters, and digested with USER enzyme. Annealed libraries were then purified with AMPure XP beads (Beckman Coulter) and PCR amplified with

indexing primers for 7 cycles. Sequencing libraries were checked with Bioanalyzer HS DNA chip (Agilent) and sequenced on the Illumina Nextseq 2000 platforms (pair end reads, 50bp).

Hi-C

Hi-C template preparation was performed as described in *Rao et al* [20]. In brief, ten million cells per sample were crosslinked with 2% formaldehyde followed by quenching by glycine at final concentration of 0.2 M. Four-cutter restriction enzyme DpnII (New England BioLabs) was used for in situ digestion (400U/ 10 million cells). Digested DNA were repaired with biotin-14–dATP (Life Technologies) in a Klenow end-filling reaction. End-repaired, ligated, and reverse-crosslinked DNA was subsequently purified using isopropanol and magnetic beads (Macherey-Nagel NucleoMag PCR Beads). Purified DNA was sheared to 300-500 bp with Covaris and subsequently size-selected by AMPure XP (Agencourt). Appropriately-sized ligation fragments marked by biotin were pulled down with MyOne Streptavidin C1 DynaBeads (Invitrogen) and prepped for Illumina sequencing.

Generation of auxin-inducible degron cells

To deplete RAD21 in cells, we utilized the AID2 system [41] generate eHap1 cells stably expressing OsTIR1 (F74G), we transduced the cells with lentivirus containing an expression cassette of OSTIR1-P2A-hygromycin. After antibiotic selection with hygromycin, cells were co-transfected with a vector expressing a sgRNA against RAD21 and SpCas9-T2A-BFP, and the donor template containing AID-GFP flanked by homology arms. GFP-positive cells were analyzed and sorted with flow cytometry. Single cell clones were expanded and used for downstream analysis. To test the responsiveness of RAD21 depletion upon auxin (5-Ph-IAA, BioAcademia) treatment, we treated the cells with auxin for 2h and analyzed the expression of GFP with flow cytometry.

Data analysis

4C-seq

4C-seq reads were mapped to the hg38 reference genome and processed using pipe4C [38] (https://github.com/deLaatLab/pipe4C) with the following parameters: local (2Mb genomic region centered tetO) normalization to 1 million reads in cis, non-blind fragments only, window size 41, top 2 read counts removed. Profile overlays were produced using R (r-project.org).

Meta-TAD analysis

Per TetO integration site, locally (2Mb) normalized 4C reads were separated into 200kb centered the TetO repeats and the rest of region. Sum of 4C signal per region were calculated per tetO. Change in 4C signals of all TetO repeats were plotted for aggregative chromosomal interactions from TetO repeats,

ChIP-seq

ChIP-seq reads were mapped to the hg38 reference genome and processed using the 4DN ChIP-seq pipeline (https://github.com/4dn-dcic/chip-seq-pipeline2). P-val signal bigwigs were used for all heatmaps and example plots. For WT, T-MAU2, T-MAU2 treated with doxycycline or T-mCherry cells, the p-val signals were normalized based on the average p-val signal for all peaks of the same factor (except STAG1 and STAG2 for which the SMC1 peaks were used) in WT (CTCF, H3K27ac, H3K4me3) or T-MAU2 cells (all other factors) located further than 3MB from the TetO integration sites. Briefly, ChIPseq peaks were filtered for a signalValue which represented clear peaks by visual inspection (CTCF: 35, H3K27ac: 4, H3K4me3: 30, FLAG-MAU2: 35, SMC1: 35, RAD21: 40, MAU2: 0, NIPBL: 30) and for overlapping peaks, such that for overlapping peaks the peak with the highest signalValue was kept. Filtered peaks were resized to 10 bp and the signal was calculated using the GenomicRanges and rtracklayer R/Bioconductor [42,43]. The average signal was used as the scaling factor. For RAD21-AID cells p-val signals were normalized based on the average signal of the regions flanking the filtered T-MAU2 peaks. Briefly, peaks were filtered for signalValue and for overlapping peaks as described above. Next, peaks were resized to 5kb and the signals of the outer 1kb regions (2.5kb-1.5kb upstream and downstream of the peak center) were calculated. The average signal of the outer 1kb regions was used as the scaling factor. For heatmaps, the signal coverage was calculated per 10 bp bin as described above and normalized using the previously determined scaling factor. For the meta plot the average signal for each 10 bp bin was calculated.

CTCF motif analysis

The presence and orientation of CTCF motifs below each CTCF peak were identified using FIMO v5.3.0 [44] with motif MA0139.1 [45] and max-stored-scores 50000000. CTCF peaks for which all identified motifs were located on the plus strand were classified as forward CTCF peaks, while peaks for which all identified motifs were located on the minus strand were classified as reverse CTCF peaks. Forward CTCF motifs located upstream of TetO sites within 1 Mb and reverse CTCF motifs located downstream of CTCF within 1 Mb were classified as convergent CTCF binding sites. Reverse CTCF motifs located upstream of TetO sites within 1 Mb and forward CTCF motifs located downstream of CTCF within 1 Mb were classified as divergent CTCF binding sites.

Cohesin travel distance calculation

We used differentially-identified FLAG peaks (MAU2 vs. MAU2-Dox) as a proxy to ascertain how far the cohesin molecules are able to traverse along the chromatin after being loaded on our TetO platforms. Specifically, we partitioned the 2mb downstream and upstream of each TetO platform into 100kb bins. Then for each partition, we counted the number of overlapping FLAG peaks that are exclusively strong when the TetO platforms are active (i.e., log2(peak_coverage (MAU2)) > 3 and log2(peak_coverage(MAU2) / peak_coverage(MAU2-Dox)) > 1). These counts are then aggregated across the downstream and upstream of all (n=27) TetO platforms.

Bru-seq

After mapping to the reference genome (hg38, using BWA-MEM, default settings), we counted the number of reads that map over each gene body and used these counts as a measure of genes activities. We then followed DESeq2's median of ratios approach [46] to normalize the counts against sequencing depth. Next, inactive genes were excluded from the analysis by removing 40% of genes with smallest average activity (averaged across MAU2 and MAU2-Dox

experiments). Finally, genes were grouped into four categories depending on their relative position to the nearest TetO platform: 1. Genes having the TetO platform inside their body. 2. (and 3.) Genes located in the same TAD as the nearest TetO platform and transcribing towards (or away from) the TetO platform. 4. The rest of the genes. As a control, we repeated the same procedure for a case where mCherry is loaded on our TetO platform (i.e., mCherry, mCherry-Dox experiments).

Hi-C analysis

Hi-C data was processed using the distiller pipeline from Open2C (https://github.com/open2c/distiller-nf). The reads were mapped to the human reference genome assembly hg38 with bwa mem v0.7.17-r1188 with "-SP" flags. The alignments were parsed and filtered for duplicates using the pairtools v0.3.0 (https://github.com/open2c/pairtools). The complex walks in long reads were masked with "--walks-policy mask", the maximal allowed mismatch for reads to be considered as duplicates "max_mismatch_bp" was set to 1, and the mapping quality threshold was set to 30. Filtered reads pairs were aggregated into genomic bins of different sizes using the cooler v0.8.11 (https://github.com/open2c/cooler) [47]. The resulting Hi-C matrices were normalised using the iterative correction procedure.

Aggregate stripes analysis

The aggregate stripes analysis of the TetO integrations was performed using cooltools v0.5.1 (https://github.com/open2c/cooltools)[48] and bioframe v0.3.0 (https://github.com/open2c/bioframe)[49] for 10 kb observed-over-expected Hi-C contact matrices. Expected contact matrices were obtained using the cooltools.expected_cis function for each chromosomal arm. The pile-ups were created using the cooltools.pileup function with 500 kb regions around the integration coordinates as flanks.

Aggregate CTCF-CTCF motifs analysis

The aggregate analysis of the CTCF-CTCF motifs interactions was performed using cooltools v0.5.1 (https://github.com/open2c/cooltools)[48] and bioframe v0.3.0 (https://github.com/open2c/bioframe) [49] for 10 kb observed-over-expected Hi-C contact matrices. For this analysis, we focused on CTCF motifs that are located ±1Mb around the TetO integrations and overlap with FLAG peaks. These CTCF motifs were clustered by 10 kb distance with bioframe.cluster function and the strongest motif within each cluster was retained. We then created a list of pairwise CTCF-CTCF motifs of distance more than 50 kb in convergent orientation. This resulted in N=337 convergent CTCF-CTCF motif pairs. The pile-ups were created using the cooltools.pileup function with 500 kb regions around the motifs intersection as flanks.

HiC plots

Hi-C example plots were generated using Juicebox input files [50]. KR-normalized counts were extracted using Straw [50], normalized to 1 million reads, two-dimensional boxcar kernel smoothened using the kernel2dsmooth function from the smoothie package [51] and plotted using R.

Acknowledgments: We thank L. Giorgetti for providing the initial TetO constructs; M. Tanenbaum for providing the TetR construct; Q. Liu for assistance with cloning at the early stage of the project; C. Valdes and R. Neijts for valuable discussions; R. van der Weide for help with ChIP-seq analysis; and all members of the de Laat lab for support and advice.

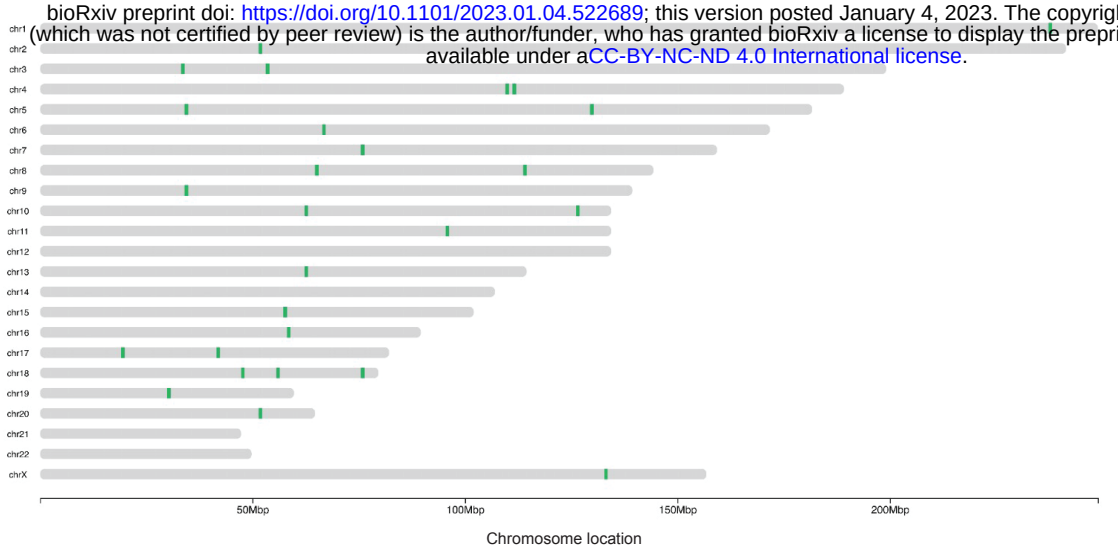
Author contributions: R.H. and W.d.L. conceived and initiated the project. R.H. cloned the construct. R.H. and Y.H. generated the cell lines. R.H. performed the ChIP-seq and BrU-seq experiments. Y.H. performed the 4C-seq experiments. M.V. performed western blot and Hi-C experiments. Y.H. and P.H.L.K. performed 4C-seq analysis. I.V. and P.H.L.K. performed integration site mapping and ChIP-seq analysis. A.A. performed cohesin travel distance and BrU-seq analysis. M.M. performed Hi-C analysis with input from E.d.W. R.H. and W.d.L. drafted the manuscript with input from other authors. R.H. and Y.H. edited the figure panels.

Competing interests: Authors declare that they have no competing interests.

Funding: This work is part of the Oncode Institute, and was funded by the NWO Groot grant (2019.012).

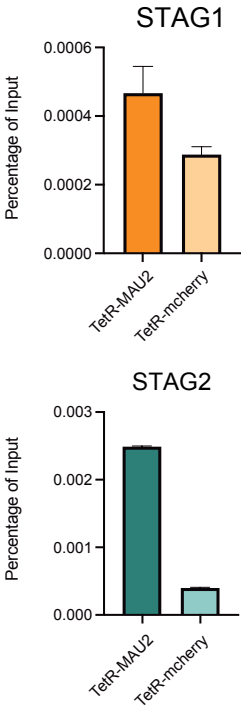
Data and materials availability: Cell lines, plasmids, and other materials are available upon request. Sequencing data are available at GSE218803.

Supplemental figure 1. Integration sites of TetO platforms

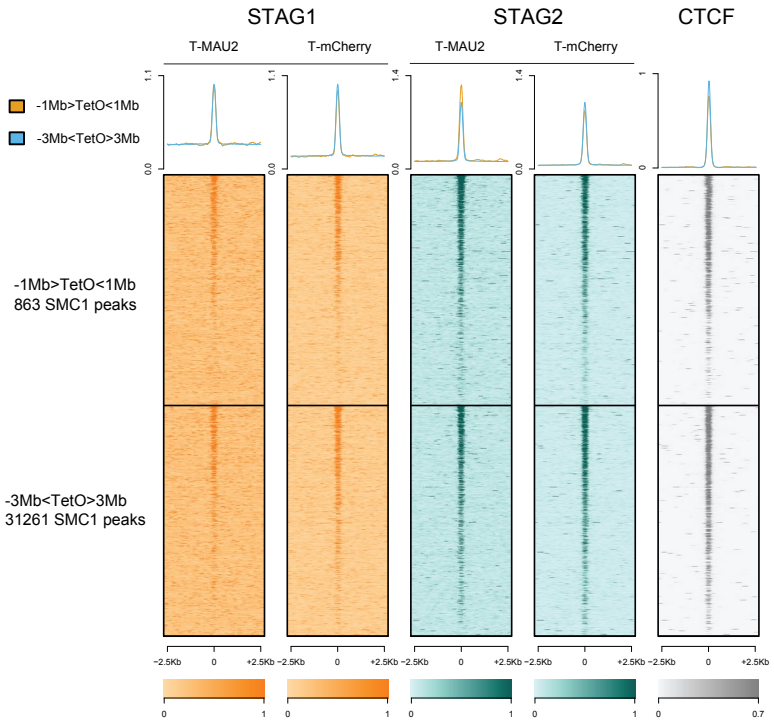


Supplemental figure 2. STAG1 and STAG2 ChIP-seq experiments

A

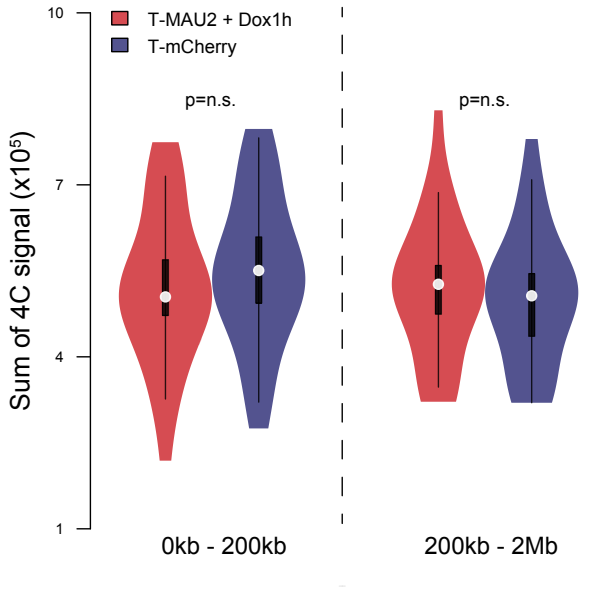


B

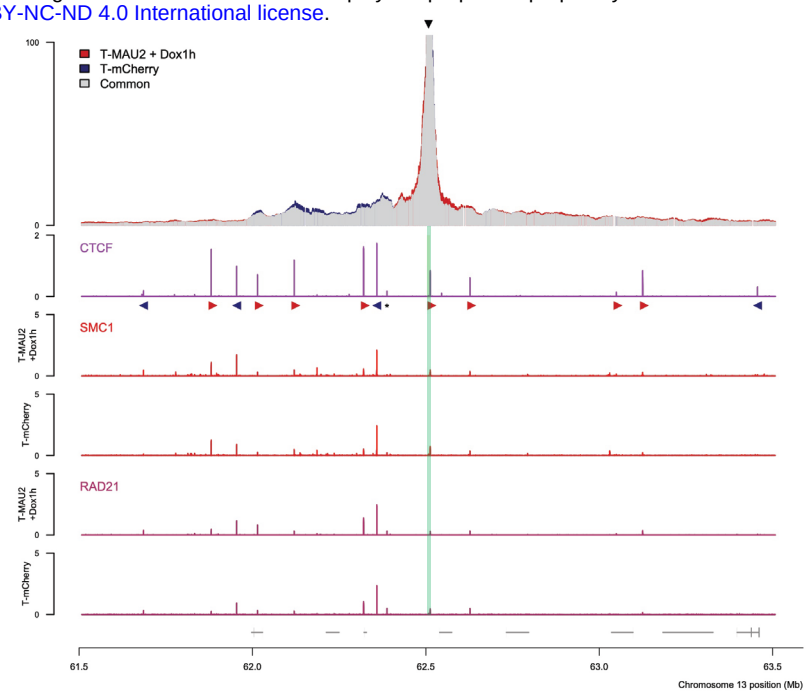


Supplemental figure 3. Meta-TAD analysis and an example of 4C-seq (T-MAU2 + Dox1h vs T-mCherry)

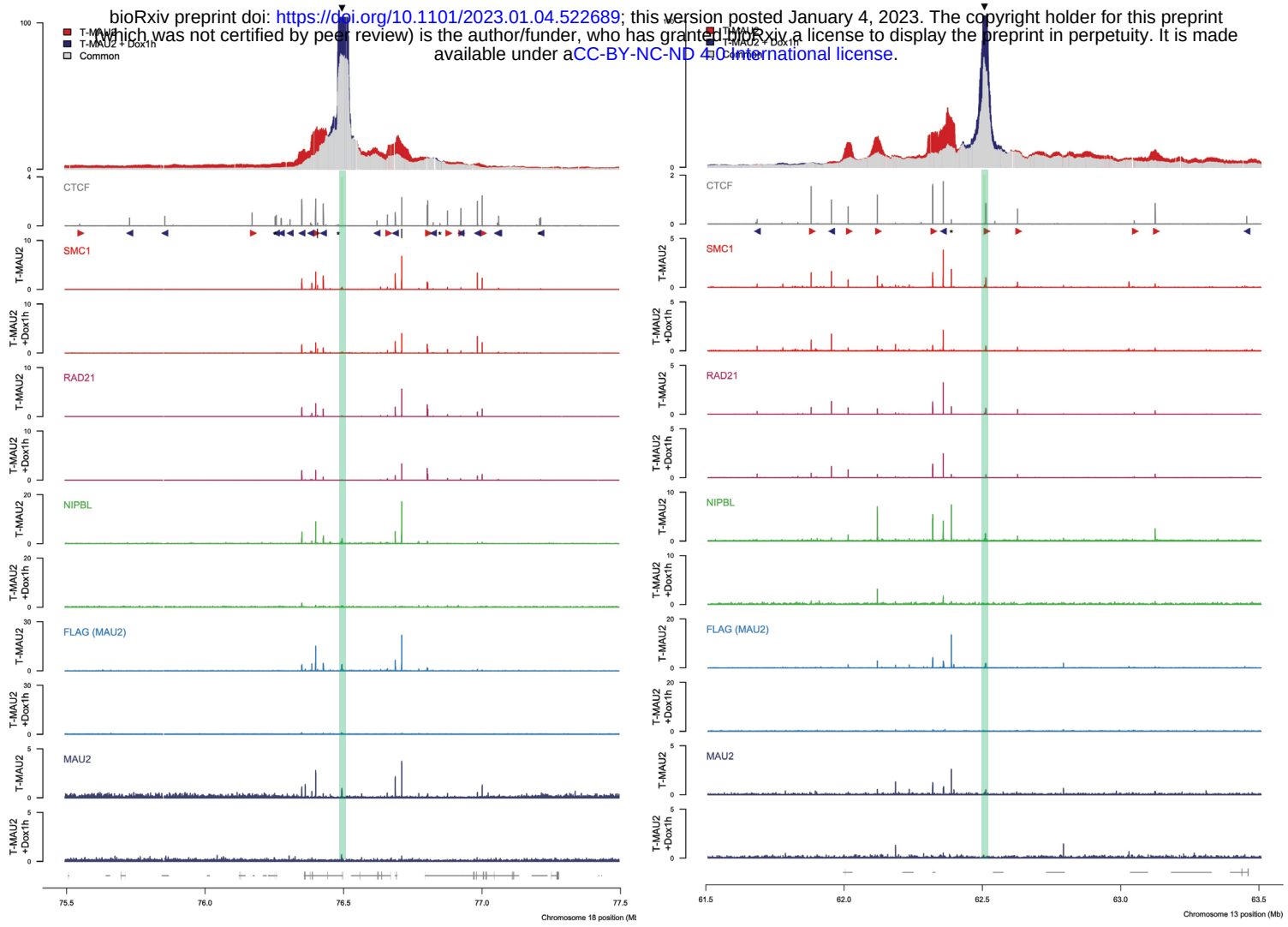
A



B

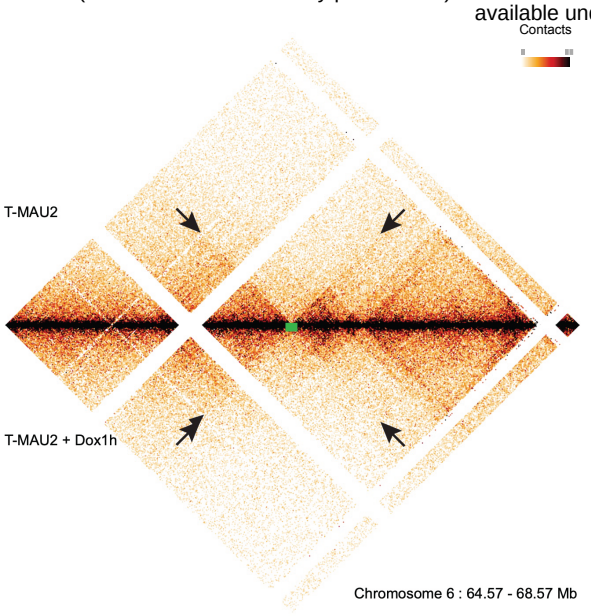


Supplemental figure 4. 4C-seq and cohesin ChIP-seq tracks near TetO sites (T-MAU2 vs T-MAU2 + Dox1h)

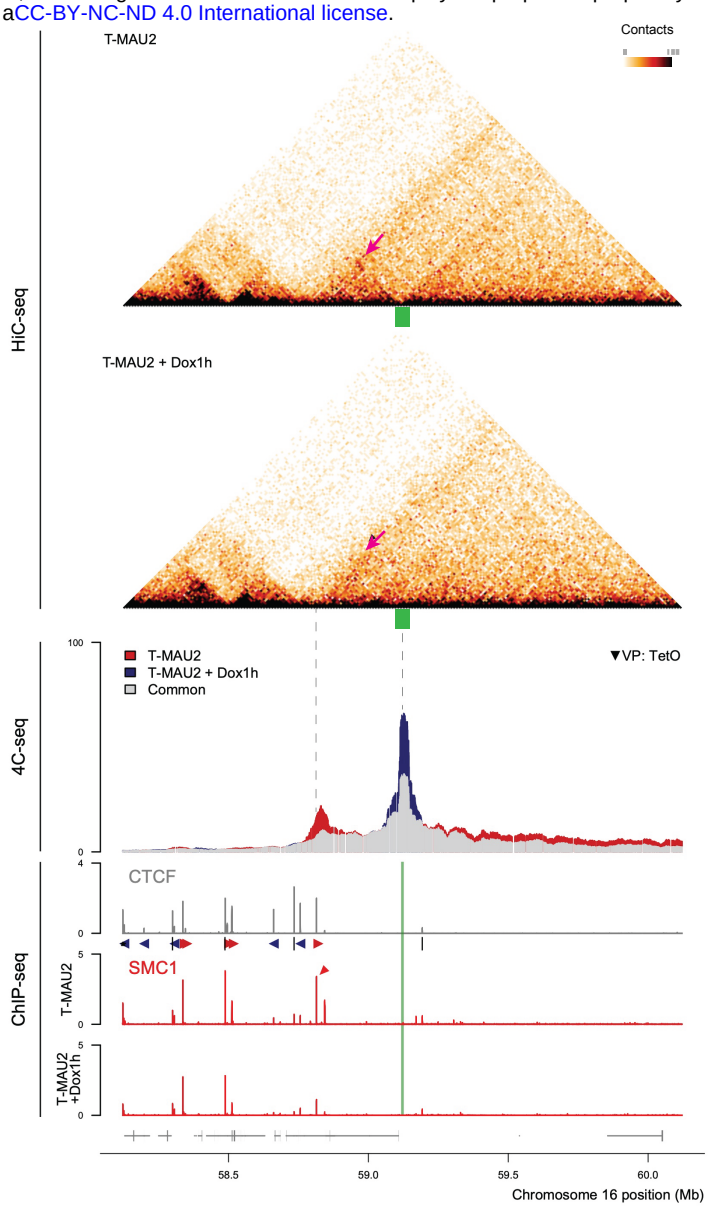


Supplemental figure 5. Hi-C profiles of TetO sites

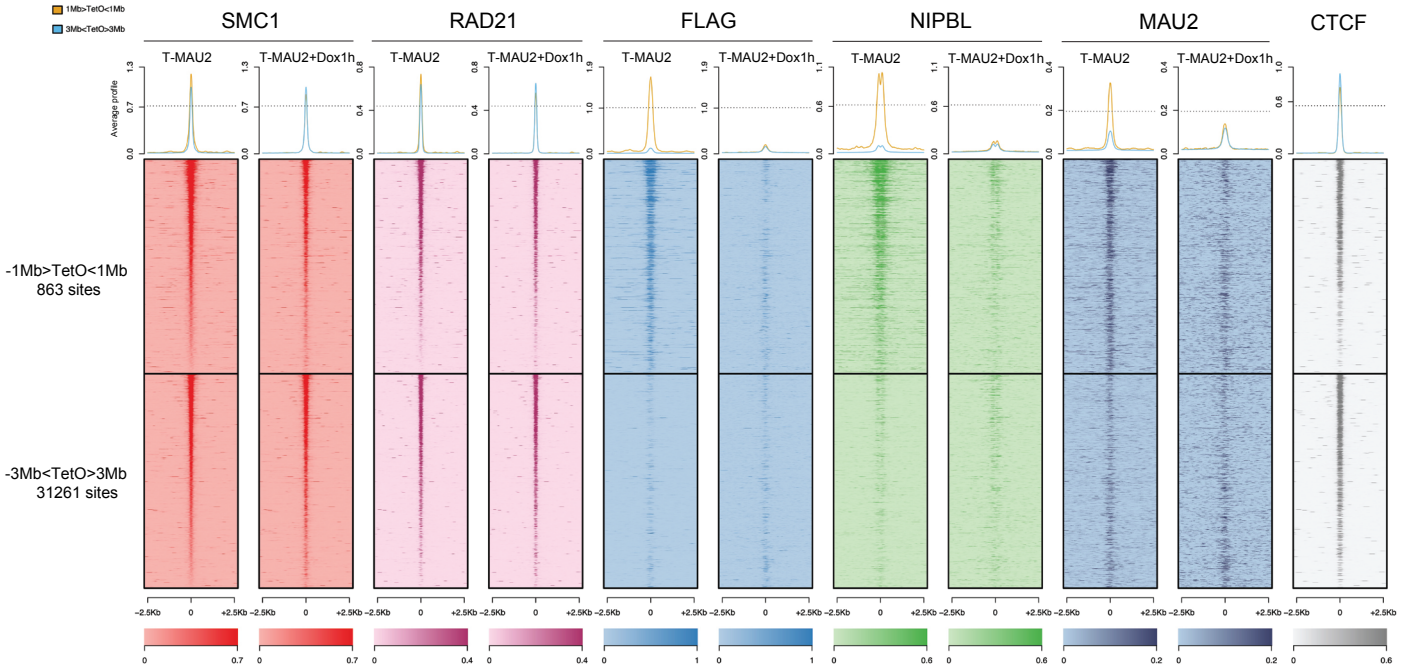
A



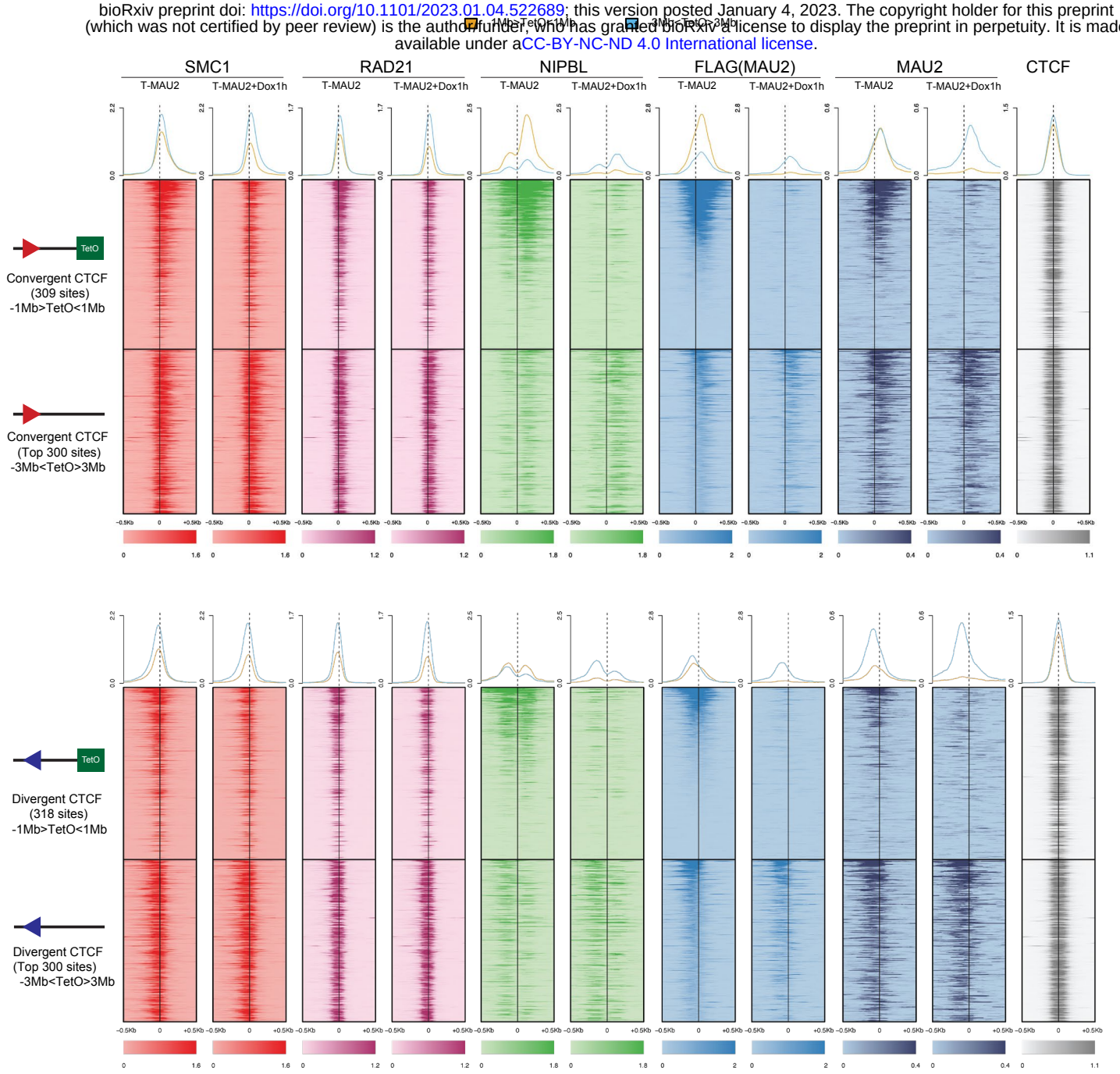
B



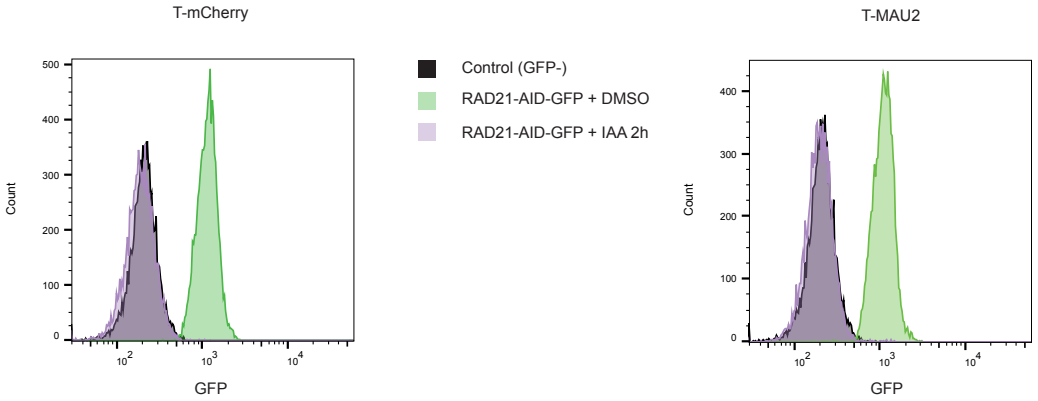
Supplemental figure 6. Average signal profiles and heatmaps of cohesin factor ChIP-seq



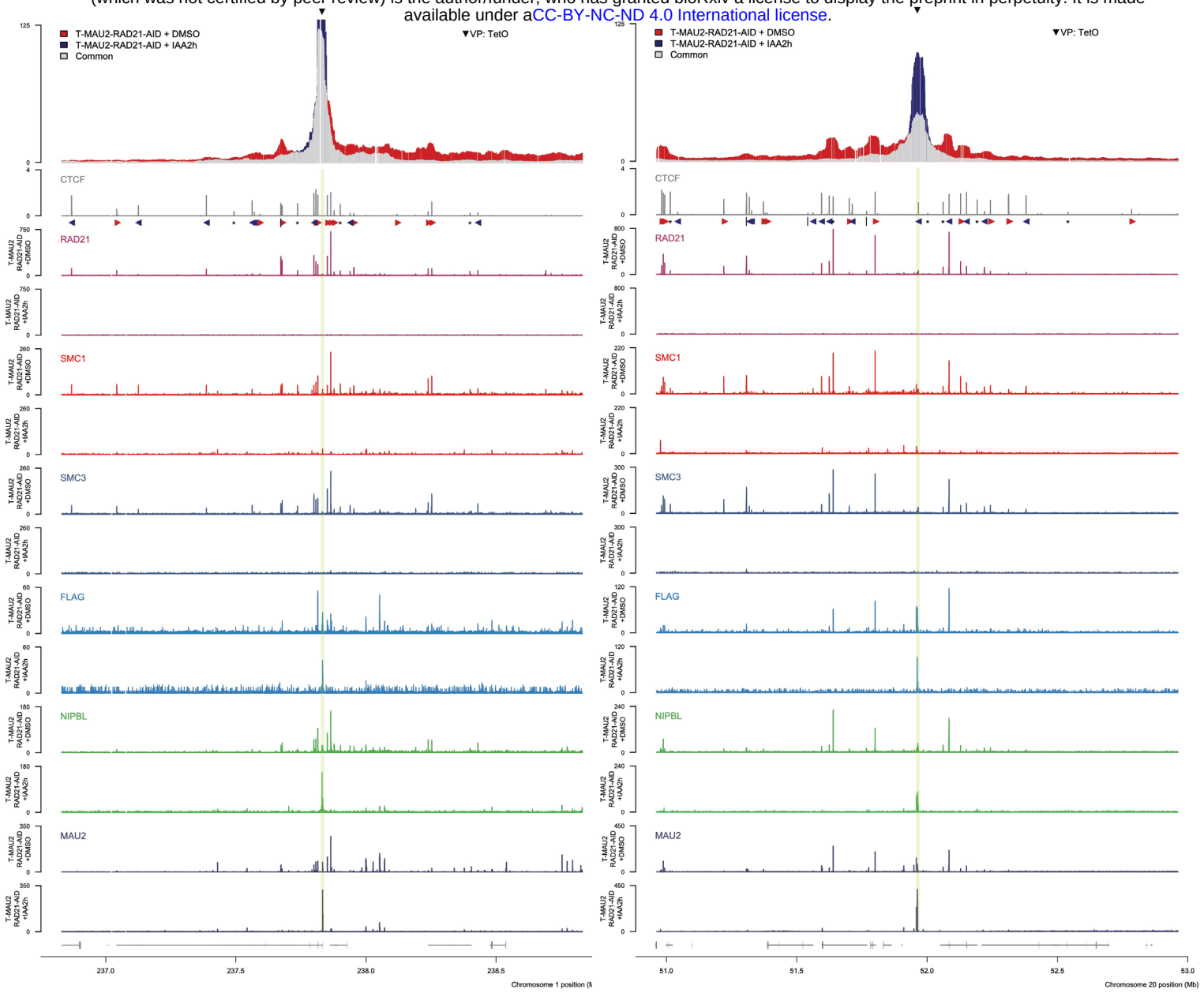
Supplemental figure 7. Average signal profiles and heatmaps of cohesin ChIP-seq separated by CTCF orientations



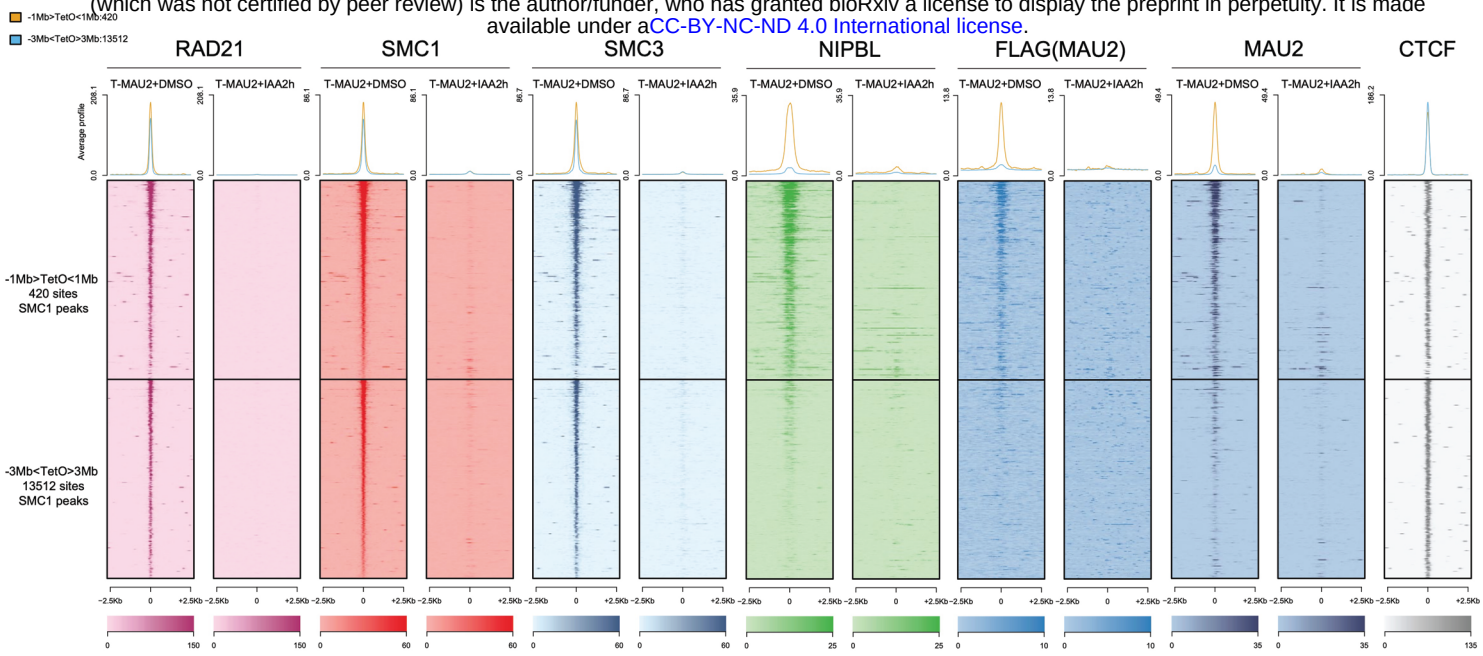
Supplemental figure 8. RAD21 depletion in RAD21-AID cells



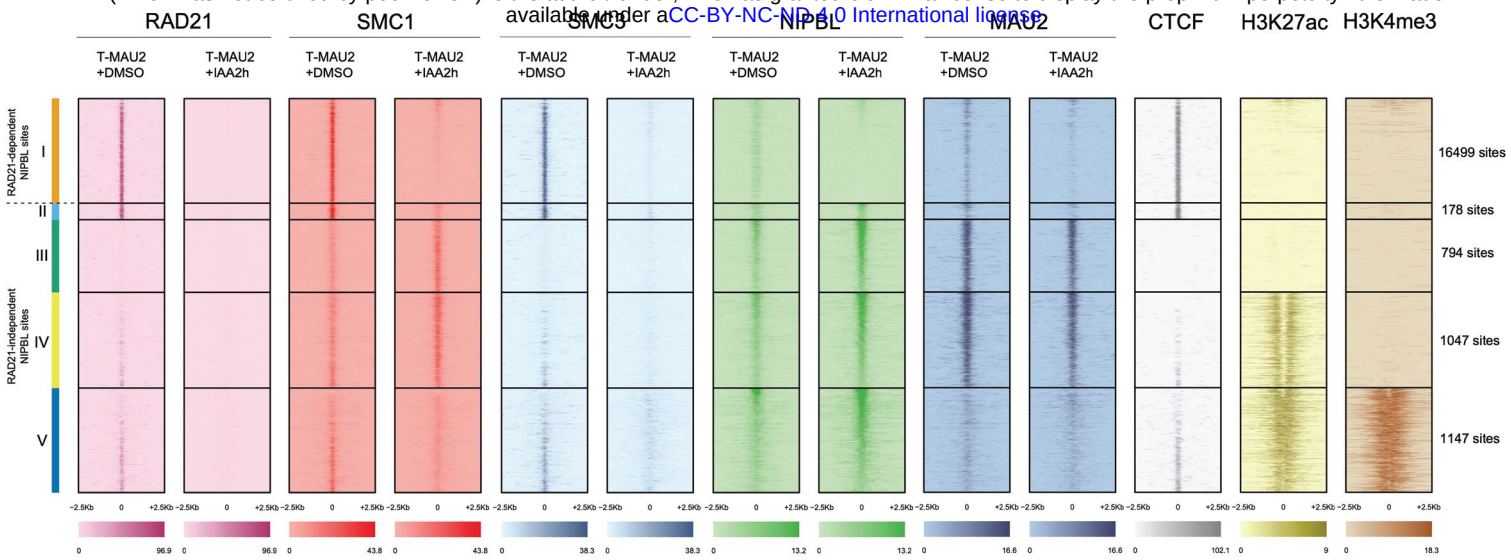
Supplemental figure 9. 4C-seq and ChIP-seq tracks upon RAD21 depletion in RAD21-AID cells



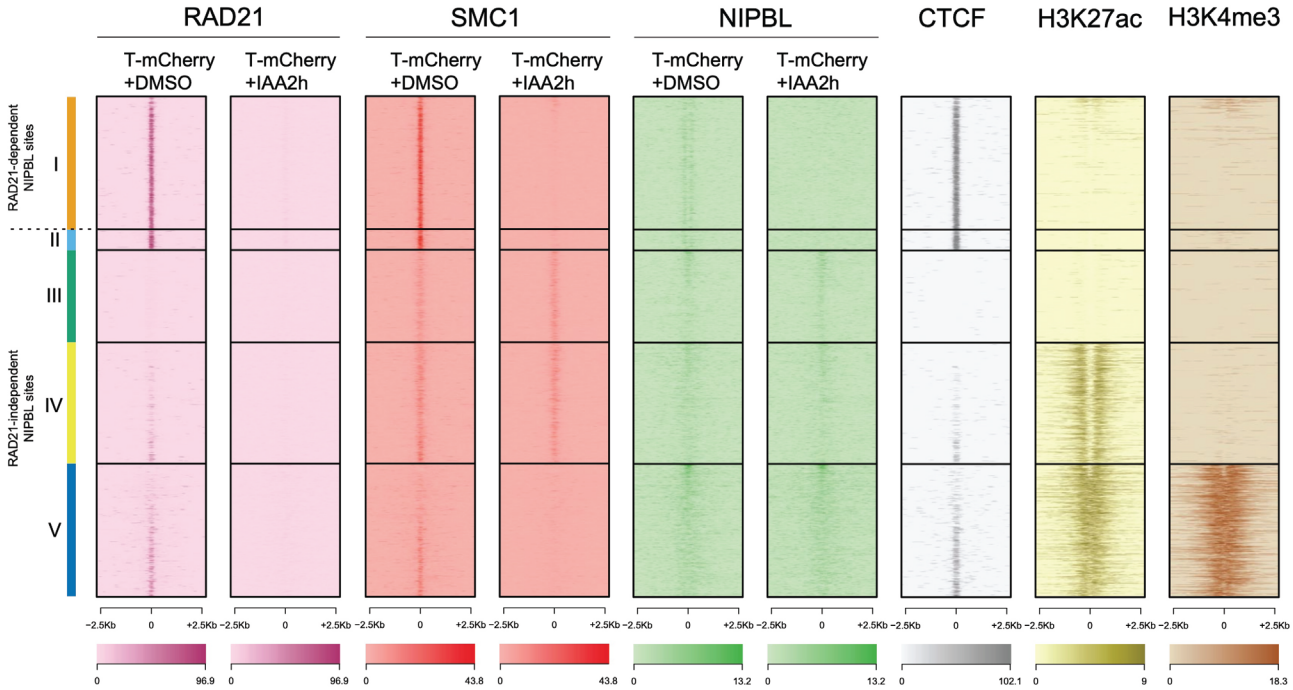
Supplemental figure 10. ChIP-seq upon RAD21 depletion in RAD21-AID cells



Supplemental figure 11. Heatmaps of ChIP-seq signals separated by dependency on RAD21 depletion (T-MAU2 cells)



Supplemental figure 12. Heatmaps of ChIP-seq signals separated by dependency on RAD21 depletion (T-mCherry cells)



Supplemental figure 13. Transcription activities of genes within the TetO-TADs

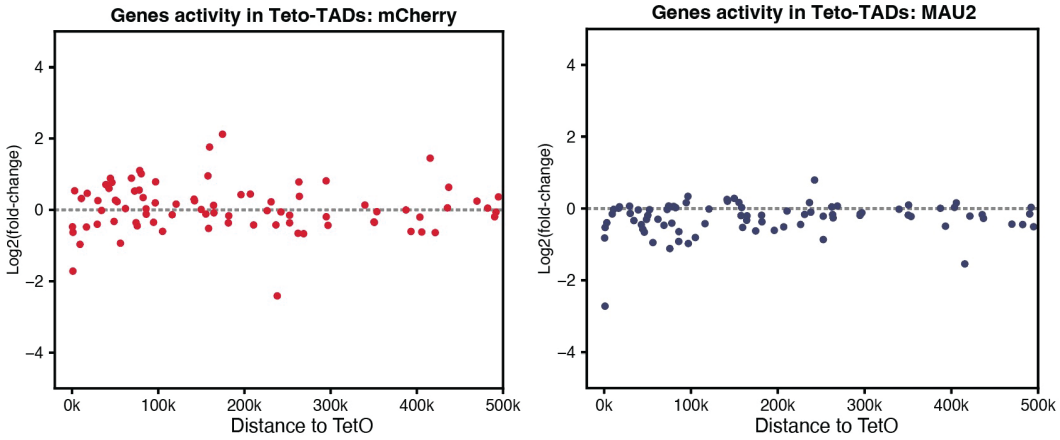


Fig. S1. Integration sites of TetO platforms. Green bars mark the location of TetO integration sites. In total, 27 sites are mapped.

Fig. S2. STAG1 and STAG2 ChIP-seq experiments. (A) ChIP-qPCR analysis of ChIP enrichment at TetO. Values are shown in percentage of input. **(B)** Average signal profiles and heatmaps of STAG1 and STAG2 ChIP-seq. Peaks are centered around SMC1 and signals are ranked by SMC1 peak strength. ChIP-seq signals are separated based on the distance to TetO into ±1Mb around TetO or more than 3Mb away.

Fig. S3. Meta-TAD analysis and an example of 4C-seq (T-MAU2 + Dox1h vs T-mCherry). (A) Sum of 4C signals 0kb-200kb and 200kb-2Mb in T-MAU2+Dox1h and T-mCherry conditions. P-values are obtained from two-sided t-test. (B) An example 4C-seq and ChIP-seq tracks from a TetO. Green bar indicates the TetO location.

Fig. S4. 4C-seq and cohesin ChIP-seq tracks near TetO (T-MAU2 vs T-MAU2+Dox1h). 4C-seq are plotted as overlays between T-MAU2 and T-MAU2+Dox1h conditions. Green bars indicate the location of TetO.

Fig. S5. Hi-C profiles of TetO sites. (A) Stripe formation from a TetO site. Green bar indicates the TetO location and black arrows depict the strips. **(B)** Hi-C, 4C-seq, and ChIP-seq overlays showing increased contacts induced by TACL. Increased contacts are indicated with magenta arrows on Hi-C maps, dashed line on 4C-seq profiles. Green bars at the center of the profile depict the TetO locations

Fig. S6. Average signal profiles and heatmaps of cohesin factor ChIP-seq. Profiles are centered around SMC1 peaks and signals are ranked by SMC1 peak strength. ChIP-seq signals are separated based on the distance to TetO into ±1Mb around TetO or more than 3Mb away.

Fig. S7. Average signal profiles and heatmaps of cohesin ChIP-seq separated by CTCF orientations. ChIP-seq peaks are centered at SMC1 peaks ±1Mb or > 3Mb away from all TetO sites. Profiles are shown ± 0.5kb around the peaks. CTCF sites are separated based on their relative orientation to TetO: convergent CTCF are all CTCF sites facing TetO (upper); divergent CTCF are all CTCF sites facing the outside of TetO (lower). Different orientations (relative to the genome) within the same category are flipped to the same direction. Signals are ranked by FLAG peak strength in T-MAU2 cells.

Fig. S8. RAD21 depletion in RAD21-AID cells. Flow cytometry analysis of T-mCherry and T-MAU2 cells with AID-GFP inserted at the endogenous RAD21 gene. Treatment of IAA for 2h completely removes RAD21-AID-GFP from the cell population.

Fig. S9. 4C-seq and cohesin ChIP-seq tracks upon RAD21 depletion in RAD21-AID cells. 4C-seq are plotted as overlays between T-MAU2-RAD21-AID and T-MAU2-RAD21-AID + IAA2h conditions. Green bars indicate the location of TetO.

Fig. S10. ChIP-seq upon RAD21 depletion in RAD21-AID cells. Average signal profiles and heatmaps of ChIP-seq signals centered at SMC1 peaks ± 1Mb or >3Mb away from all TetO sites. Profiles are shown ± 2.5kb around the peaks.

Fig. S11. Heatmaps of ChIP-seq signals separated by dependency on RAD21 depletion (T-MAU2 cells). Heatmaps of RAD21, SMC1, SMC3, NIPBL, MAU2, CTCF, H3K27ac, and H3K4me3 ChIP-seq signals centered at SMC1 and NIPBL peaks. Sites are divided into different categories by the presence of SMC1 and NIPBL after RAD21 depletion.

Fig. S12. Heatmaps of ChIP-seq signals separated by dependency on RAD21 depletion (T-mCherry cells). Heatmaps of RAD21, SMC1, NIPBL, CTCF, H3K27ac, and H3K4me3 ChIP-seq signals centered at SMC1 and NIPBL peaks. Sites are divided into different categories by the presence of SMC1 and NIPBL after RAD21 depletion.

Fig. S13. Transcription activities of genes within the TetO-TAD. Each dot represents a gene within the TetO-TAD. Y-axis represents the log2 fold change of gene activities measured by BrU-seq between T-MAU2/T-mCherry cells and the cells treated with Dox for 1h. X-axis represents the relative distance of the indicated gene to the nearest TetO.