

Published in final edited form as:

*Curr Biol.* 2011 October 11; 21(19): 1624–1634. doi:10.1016/j.cub.2011.08.036.

## Cohesin Selectively Binds and Regulates Genes with Paused RNA Polymerase

Avery Fay<sup>1</sup>, Ziva Misulovin<sup>1</sup>, Jian Li<sup>2</sup>, Cheri A. Schaaf<sup>1</sup>, Maria Gause<sup>1</sup>, David S. Gilmour<sup>2</sup>, and Dale Dorsett<sup>1,\*</sup>

<sup>1</sup>Edward A. Doisy Department of Biochemistry and Molecular Biology, Saint Louis University School of Medicine, Saint Louis, Missouri, USA 63108

<sup>2</sup>Center for Eukaryotic Gene Regulation, Department of Biochemistry and Molecular Biology, Pennsylvania State University, University Park, Pennsylvania, USA 16802

### Summary

**Background**—The cohesin complex mediates sister chromatid cohesion and regulates gene transcription. Prior studies show that cohesin preferentially binds and regulates genes that control growth and differentiation, and that even mild disruption of cohesin function alters development. Here we investigate how cohesin specifically recognizes and regulates genes that control development in *Drosophila*.

**Results**—Genome-wide analyses show that cohesin selectively binds genes in which RNA polymerase II pauses just downstream of the transcription start site. These genes often have GAGA factor (GAF) binding sites 100 base pairs (bp) upstream of the start site, and GT dinucleotide repeats 50 to 800 bp downstream in the plus strand. They have low levels of histone H3 lysine 36 trimethylation (H3K36me3) associated with transcriptional elongation, even when highly transcribed. Cohesin depletion does not reduce polymerase pausing, in contrast to depletion of the NELF (Negative Elongation Factor) pausing complex. Cohesin, NELF and Spt5 pausing/elongation factor knockdown experiments indicate that cohesin does not inhibit binding of polymerase to promoters or physically block transcriptional elongation, but at genes that it strongly represses, hinders transition of paused polymerase to elongation at a step distinct from those controlled by Spt5 and NELF.

**Conclusions**—Our findings argue that cohesin and pausing factors are recruited independently to the same genes, perhaps by GAF and the GT repeats, and that their combined action determines the level of actively elongating RNA polymerase.

### Keywords

*bbg*; *big bang*; bivalent gene; DSIF; EcR; ecdysone; *engrailed*; *Enhancer of split*; E(spl)-C; insulator; *invected*; Nipped-B; Polycomb; Rad21; Set2; Smc1; *terribly reduced optic lobes*; *trol*

---

© 2011 Elsevier Inc. All rights reserved.

\*Correspondence: dorsettd@slu.edu .

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

## Introduction

The cohesin protein complex is essential for sister chromatid cohesion, participates in DNA repair, regulates gene expression, and mild disruption of its activity causes developmental deficits in *Drosophila*, zebrafish, mice, and humans [1-3]. The Smc1, Smc3 and Rad21 (Mcd1/Sccl) cohesin subunits form a ring-like structure, and Stromalin (SA, Sccl, Stag) interacts with Rad21. Cohesin binds to chromosomes, and likely participates directly in cohesion, repair, and transcription. How cohesin carries out these functions is not well understood.

To understand cohesin function, it is essential to know how, when, and where it binds to chromosomes. Cohesin binds chromosomes throughout interphase in multiple modes that differ in stability [4, 5]. These binding modes are regulated by multiple factors [6]. Key among these is Nipped-B (Sccl, Mis4, Nipbl), which complexes with the Mau-2 (Sccl, Ssl3) protein and loads cohesin onto chromosomes [7-11]. Nipped-B and cohesin both bind chromosomes in a stable mode with a half-life of several minutes, and Nipped-B dosage dictates the amount of stable cohesin [4]. Substantial evidence indicates that stable cohesin topologically encircles DNA [12, 13].

In *Drosophila*, Nipped-B co-localizes with cohesin along chromosome arms [14, 15]. Cohesin and Nipped-B bind some 20% of active genes, with the highest levels at transcription start sites and DNA replication origins [15, 16]. Cohesin is excluded from inactive genes, including genes silenced by Polycomb complexes [15]. The pattern in mammalian cells is similar to that in *Drosophila*, except that cohesin, but not Nipped-B (NIPBL), associates with sites that bind the CTCF protein, which regulates transcription by various mechanisms [17-20].

Cohesin also regulates genes via multiple mechanisms. It facilitates looping and communication between transcriptional enhancers and gene promoters, and between sites that bind CTCF [17, 21-26]. Cohesin represses many genes as well, acting in concert with the Polycomb group (PcG) repressor proteins at a few genes in *Drosophila* cells, and many in mouse embryonic stem cells [1, 17, 27]. The genes bound and regulated by cohesin are enriched for those that control growth and development [1, 17, 27].

It is unknown how Nipped-B and cohesin selectively bind growth and development genes. To tackle this question we looked for DNA sequence and other features that are enriched in cohesin binding genes in *Drosophila* cells. We find that Nipped-B and cohesin bind genes in which RNA polymerase II is paused just downstream of the transcription start site. Our results indicate that at genes repressed by cohesin, cohesin hinders an early step in transcription distinct from, but closely coupled to the steps regulated by the DSIF and NELF pausing factors.

## Results

### Cohesin Binding Genes Are Enriched for GAGAG Motifs and GT Repeats

To find DNA motifs that might target Nipped-B and cohesin to specific genes, we compared sequences around the transcription start sites of genes that bind cohesin, and active genes that do not. Using genome-wide chromatin immunoprecipitation (ChIP-chip) data, we defined active genes that bind cohesin as those that bind RNA polymerase II (PolII), and cohesin (Nipped-B and Smc1) at the transcription start sites (TSS) in both the Sg4 (S2 subline) cells of embryonic origin, and ML-DmBG3 (BG3) cells derived from larval central nervous system [15]. Genes that do not bind cohesin were defined as those that bind only

PolII at the TSS in both cell lines. This identified 506 cohesin-binding and 1040 non-binding genes (Supplemental Files S1 and S2).

We counted the occurrences of all possible five nucleotide sequences in the plus strand from 1 kb upstream to 1 kb downstream of the TSS in both gene sets. This identified GAGAG and its reverse complement CTCTC as enriched in the cohesin-binding over non-binding genes. It also identified GTGTG, but not the reverse complement, as enriched in cohesin-binding genes. GAGAG and CTCTC occur 50 to 150 bp upstream of the TSS, and enrichment increases with the stringency of the p value used to call cohesin binding (Figure 1A,B).

GTGTG is enriched 50 to 800 bp downstream of the TSS, is plus strand-specific, occurs only in non-coding sequences (5' UTR and intron), and thus likely functions in nascent RNA (Figures 1C-E). The GTGTG motif is a subset of extended GT or TG repeats, as revealed by sorting the cohesin binding and non-binding genes according to the frequency with which a GT or TG dinucleotide follows another (Figure 1F). The TBPH protein binds UG repeats [28], but RNAi depletion of TBPH in BG3 cells had minimal effects on expression of several cohesin-regulated genes (data not shown).

### **Cohesin Associates with Genes that Bind GAGA Factor (GAF) and Negative Elongation Factor (NELF)**

The GAGAG sequence binds GAGA factor (GAF) encoded by *Trithorax-like (Trl)*. GAF has been mapped genome-wide in BG3 and S2 cells by ChIP-chip [29, 30]. GAF binds near the TSS of many cohesin-binding genes, indicating that the GAGAG sites identified by sequence analysis (Figure 1A) are functional. Accordingly, there is a 0.42 genome-wide correlation between Nipped-B binding in Sg4 cells and GAF in S2 cells, and a 0.50 correlation between Nipped-B and GAF in BG3 cells (Figure S1E,F).

GAF predicts genes that bind the NELF complex, which causes promoter-proximal polymerase pausing [30-32]. The genome-wide correlation between Nipped-B in Sg4 cells and NELF in S2 cells (0.45, Figure S1B) is similar to the GAF-NELF correlation (0.40, Figure S1D), indicating that Nipped-B also predicts NELF binding. As positive controls, the correlation between the B and E subunits of NELF in S2 cells is 0.77, and between Nipped-B and the Smc1 cohesin subunit in BG3 cells is 0.90 (Figure S1A,C). Thus Nipped-B and cohesin associate preferentially with genes that bind GAF and NELF. Nearly identical correlations between cohesin (Rad21) and GAF (0.39), and between cohesin and NELF (0.45) were found in quantitative co-immunostaining experiments with salivary polytene chromosomes (data not shown).

Both cohesin and the NELF pausing factor selectively associate with a subset of genes that lack the histone H3 lysine 36 trimethylation (H3K36me3) mark made during transcriptional elongation, providing further evidence that they bind the same genes. Comparing cohesin ChIP-chip data in BG3 cells to genome-wide data [29] for H3K36me3 and H3K36me1 (monomethylation) revealed that cohesin-binding genes have H3K36me1 ( $r = 0.53$ , Figure S1I) but not H3K36me3 ( $r = -0.01$ , Figure S1H). Like cohesin, NELF-binding genes also do not have H3K36me3 ( $r = 0.08$ , Figure S1G). Cohesin-NELF binding genes are expressed, and thus we conclude that elongation does not require H3K36me3.

Paused RNA polymerase produces short transcripts with 3' ends just downstream of the TSS, which have been mapped genome-wide in S2 cells [33]. This data reveals that cohesin-binding genes are highly enriched for short promoter transcripts compared to non-binding genes, confirming that cohesin selectively binds genes with paused polymerase (Figure 2A).

## Cohesin Depletion Alters Gene Expression Without Reducing Pausing

We tested if cohesin regulates polymerase pausing by  $\text{KMnO}_4$  footprinting after cohesin depletion in BG3 cells.  $\text{KMnO}_4$  modifies T residues in single-stranded DNA created by paused polymerase [34]. We examined five genes that bind NELF and cohesin, two that are activated by cohesin (*path*, *dm*), and three that are repressed (*HLHm3*, *invected*, *rho*) [27]. As a positive control, NELF depletion decreased  $\text{KMnO}_4$  reactivity at all five genes (Figure 2C). In contrast, Rad21 or Nipped-B depletion had no discernable effects. GAF knockdown decreased pausing at *path*, *dm* and *HLHm3*, had little effect at *rho*, and slightly increased pausing at *invected*. Western blots showed that NELF subunits were reduced by at least 90%, and that GAF was reduced by some 70% (Figure 2B). Nipped-B and Rad21 were reduced by ~80% (not shown) as described previously [27]. RT-PCR experiments confirmed that Nipped-B and Rad21 knockdown increased *rho*, *HLHm3*, and *invected* transcripts, and decreased *path* and *dm* transcripts (not shown).

## Cohesin and NELF Do Not Regulate Transcriptional Induction or Elongation at the *EcR* Gene in BG3 Cells

*Ecdysone receptor (EcR)* gene expression decreases in the mushroom body  $\gamma$  neuron and salivary glands when cohesin is absent, suggesting that cohesin activates *EcR* transcription in vivo [35, 36]. Cohesin binds throughout the 80 kb *EcR* transcribed region in BG3 and Sg4 cells (Figure 3A) [15] and NELF binds the active promoters in S2 cells [30]. GAF binds all three promoters in BG3 and S2 cells [29, 30]. Although expressed at a high basal level, *EcR* lacks H3K36me3 (Figure 3A).

We knocked down cohesin and NELF to compare how they regulate basal and ecdysone-induced *EcR* transcription. All three promoters (p1, p2, p3) are active in BG3 cells prior to ecdysone hormone treatment, and depletion of Rad21, Nipped-B, NELF-B, or GAF slightly increased transcripts from p1 (Figure 3B). Ecdysone increased transcripts from the p3 promoter and total transcripts 4-fold within an hour, but knockdown of Nipped-B, Rad21, NELF-B, or GAF did not alter the induced levels. Thus cohesin does not activate *EcR* in BG3 cells.

Cohesin-dependent activation of *EcR* in vivo may require tissue-specific enhancers that are inactive in BG3 cells, but we also considered the possibility that the cohesin binding along *EcR* interferes with elongation, so that cohesin knockdown could simultaneously decrease activation and increase elongation, resulting in little change in transcript levels. We thus used ecdysone-induction time course experiments to see if depletion of Rad21, Nipped-B, NELF or GAF altered induction or elongation kinetics, following the induced wave of RNA synthesis along the gene with probes shown in Figure 3A. In the control, the first increases in p3 and intron A RNA occurred 10 min after induction, the first increases in intron B and C RNA at 20 min, and the first 3' exon RNA increase at 30 min. We infer from this wave of RNA synthesis that elongating polymerase moves from p3 to the intron C site at between 2 to 2.5 kb per min. An increase in terminal exon RNA was delayed, suggesting that splicing slows movement from intron C to the end of the gene. Within the resolution of the experiments, knockdown of Rad21, Nipped-B, NELF-B or GAF did not alter the induction or elongation kinetics, indicating that cohesin does not alter polymerase recruitment or elongation (Figure 3C, Figure S2).

Depletion of target proteins was confirmed by westerns (not shown), and we detected effects on other genes. For the experiments in Figure 3 and Figure S2, Rad21 and Nipped-B knockdown increased *HLHm3* transcripts up to 30-fold, GAF knockdown increased *HLHm3* 5-fold and *HLHm3* transcripts 2-fold, and NELF-B knockdown increased *HLHm3* expression 2-fold.

## Cohesin Hinders Transcription of Repressed Genes After PolII Recruitment

Cohesin has more direct negative than positive effects on gene expression in BG3 cells [27]. Little is known about how cohesin represses transcription, and the above data show that cohesin depletion can increase transcripts without reducing polymerase pausing, and that cohesin does not physically hinder transcriptional activation or elongation at the *EcR* gene. We thus considered the possibility that cohesin may repress transcription at a step that occurs just after pausing.

*big bang* (*bbg*, Figure 4A), *terribly reduced optic lobes* (*trol*, Figure 5A), the *Enhancer of split* complex [E(spl)-C], which contains *HLHm $\delta$* , *HLHm3*, and *HLHm7* and other bHLH genes, and the *invected-engrailed* (*inv-en*) complex (Figure 6A), are 11 of the 32 cohesin-binding genes that increase four-fold or more in expression with cohesin or Nipped-B knockdown in BG3 cells, and thus typify cohesin-repressed genes [27]. All play roles in nervous system and imaginal disc development. They bind Nipped-B, cohesin and GAF in BG3 cells (Figures 4-6), and NELF in S2 cells (not shown). None have H3K36me3 in the transcribed region with the exception of the 3' end of *trol*, where cohesin binding is negligible (Figure 5A). All have H3K36me1 in the transcribed region except the E(spl)-C, although H3K27me3 made by the E(z) PcG protein shows that histone H3 is present [15, 27]. The E(spl)-C and *inv-en* complex are targeted by PcG silencing proteins in BG3 cells, but *bbg* and *trol*, like most cohesin-binding genes, are not [15, 27].

RNAi knockdown of Rad21 or Nipped-B in BG3 cells reduced cohesin and Nipped-B binding to *bbg*, *trol*, *invected*, *engrailed*, and E(spl)-C genes (Figure S3 and not shown). Rad21 and Nipped-B knockdown over 3 to 6 days also increased transcripts from all these genes, as previously observed [27]. Total *bbg* transcripts, as measured by a 3' exon probe, increased 5- to 20-fold (Figure 4C, Figure S4) and total *trol* transcripts increased 4- to 15-fold, with increases from all active promoters (Figure 5B, Figure S5). E(spl)-C transcripts increased 5- to 20-fold (Figure S6), and *invected* and *engrailed* transcripts increased 5- to 30-fold (Figure 6B, Figure S7). Rad21 and Nipped-B depletion had nearly identical effects, except that Nipped-B RNAi often required longer treatment. These data agree with the finding that Nipped-B and Rad21 depletion have virtually identical effects on gene expression genome-wide in BG3 cells [27]. Rad21 and Nipped-B proteins were reduced ~80% by day 4 (Figure 4B and not shown). With this reduction cell division is slightly delayed, but there is no discernable change in sister chromatid cohesion or chromosome segregation [27].

Despite the transcript increases, cohesin depletion did not increase polymerase binding at the promoters as measured by ChIP. Rad21 knockdown did not alter Ser2P PolII presence at the *bbg* or *trol* promoters, but using anti-Rpb3, decreased total PolII at the proximal p3 promoter of *bbg* by 40% (Figure 4D, Figure 5C), although transcripts from this promoter increased. Rad21 knockdown increased Ser2P elongating RNA polymerase on the 3' terminal exons of both *bbg* and *trol* by some 2-fold (Figure 4D, Figure 5C). Cohesin depletion increased total and Ser2P PolII 1.5-fold at the *invected* promoter, and 2-fold or more 700 bp downstream (Figure 6C). Low transcription precluded quantitative assessment of elongating polymerase for *engrailed* and the E(spl)-C. We conclude that cohesin does not repress these genes by inhibiting polymerase binding, and posit that it diminishes how much paused polymerase transitions to elongation.

## NELF and Spt5 Pausing Factors Have Promoter-Specific Effects on Cohesin-Repressed Genes

NELF-B depletion had smaller and more promoter-specific effects than cohesin depletion. At *bbg*, NELF knockdown had no significant effect on transcripts from the p2 promoter, but



increased p3 transcripts 3- to 5-fold (Figure 4C, Figure S4). NELF RNAi had no significant effects on *trol*, E(spl)-C or *engrailed* transcripts, but increased *invected* transcripts 3- to 5-fold (Figure 5B, Figure 6B, Figures S5-S7).

Spt5 depletion (>90%, Figure 4B and not shown) also increased transcripts of the cohesin-repressed genes to a lesser extent and in a more promoter-specific manner than cohesin depletion. *bbg* p3 promoter and total transcripts increased 2- to 3-fold (Figure 4C, Figure S4). Similar results were obtained using a different Spt5 RNAi target, and Spt4 RNAi (not shown). Spt5 depletion increased *trol* transcripts 2- to 3-fold, primarily from the p2 promoter, E(spl)-C transcripts 2- to 15-fold, and *invected* and *engrailed* transcripts by 5- to 10-fold (Figure 5B, Figure 6B, Figures S5-S7). Simultaneous knockdown of NELF and Spt5 had virtually the same effect as Spt5 depletion alone at all genes, consistent with the idea that Spt5 cooperates with NELF to induce pausing (Figures 4-6, Figures S4-S7).

### Spt5 Influences PolII Binding and Is Not Required for Cohesin Binding

Spt5 depletion did not detectably increase Ser2P or total PolII at the terminal exons of *bbg* and *trol*, despite the transcript increases (Figure 4D, Figure 5C). Unexpectedly, Spt5 knockdown decreased both Ser2P and total PolII at the distal promoters of *bbg* and *trol*, although transcripts from these promoters did not change. Spt5 depletion had minor effects on Ser2P and total PolII at the *bbg* p3 promoter (Figure 4D), although transcripts from this promoter increased as much as 5-fold, and increased total and Ser2P PolII at the *trol* p2 promoter 2- to 3-fold (Figure 5C), which correlates with the increase in p2 transcripts. Thus Spt5 depletion alters PolII binding at *bbg* and *trol* promoters, but except for the *trol* p2 promoter, the changes do not correlate with transcript levels.

Spt5 depletion did not alter Nipped-B or cohesin binding to *bbg*, *invected*, or *engrailed*, but increased their binding to the *trol* p2 promoter, where PolII also increased (Figure S3). Thus Spt5 is not required for cohesin binding, and we posit that an increase in *trol* p2 transcription caused by Spt5 depletion increases cohesin binding at this promoter.

### NELF Depletion Modifies Effects of Cohesin Knockdown in a Gene-Specific Manner

We performed simultaneous knockdowns of cohesin and pausing factors to test for epistatic and functional interactions. At *bbg*, combined NELF and cohesin or Nipped-B knockdown caused a 50-fold synergistic increase in total transcripts (Figure 4C, Figure S4). In contrast, at *trol*, combining NELF-B with Nipped-B or Rad21 knockdown had effects similar to Rad21 or Nipped-B knockdown alone (Figure 5B, Figure S5). In sharper contrast, combining NELF-B with Nipped-B or Rad21 knockdown decreased the level of E(spl)-C transcripts obtained with Nipped-B or Rad21 depletion alone (Figure S6). Combining NELF-B RNAi with Rad21 or Nipped-B depletion increased the effect of cohesin knockdown on *invected*, but decreased the effect on *engrailed* (Figure 6B, Figure S7). Because combined cohesin-NELF depletion rarely had the same effect as NELF or cohesin knockdown alone, we conclude that NELF and cohesin are not epistatic to each other. The robust interactions in some cases, however, indicate that the steps they regulate are closely linked. Because cohesin depletion did not increase PolII binding to promoters or reduce pausing, these findings argue that cohesin regulates a step downstream of pausing.

### Spt5 Is Limiting for Transcription After Cohesin Knockdown

In contrast to NELF depletion, Spt5 depletion is epistatic to Nipped-B and cohesin knockdown. Combining Spt5 RNAi with Rad21 or Nipped-B knockdown diminished the transcript increase caused by cohesin or Nipped-B knockdown for most genes, including *bbg*, the E(spl)-C genes, *invected*, and *engrailed* (Figure 4C, Figure 6B, Figure S4, Figure S6, Figure S7). Spt5 knockdown decreased the effect of cohesin depletion on *trol* p1

transcripts, but had little effect, or with extended knockdown, increased the effect of cohesin RNAi on p2 and total *trol* transcripts (Figure 5B, Figure S5). Because Spt5 is also required for elongation in addition to pausing, the finding that Spt5 depletion is epistatic to cohesin knockdown argues that cohesin regulates a step between pausing and elongation.

## Discussion

Genome-wide studies show that Nipped-B (NIPBL) and cohesin preferentially bind a subset of active genes, and have direct negative and positive effects on gene expression in [15, 17, 18, 27, 35]. Many positive effects likely reflect a role for cohesin in facilitating enhancer-promoter looping [17, 21, 22], similar to its role in looping between CTCF sites (23-26). Our studies address the less-understood mechanisms by which cohesin represses genes. They indicate that cohesin selectively binds genes with paused RNA polymerase, and at repressed genes, interferes with transition of paused polymerase to elongation at a step distinct from those regulated by the DSIF and NELF (Figure 7).

### Cohesin Binds Genes with Paused Polymerase But Is Not Required for Pausing

Genome-wide correlation between cohesin and NELF binding, and high levels of short transcripts at cohesin-binding promoters provide compelling evidence that cohesin selectively binds genes with paused polymerase. This agrees with the finding that cohesin preferentially regulates genes that function in growth and development [17, 27], because these ontologies are also enriched among genes with paused polymerase [37, 38].

If cohesin enhanced pausing by physically impeding elongation, it would provide a simple explanation for the association of cohesin with genes with paused polymerase. We find, however, by permanganate footprinting, that cohesin depletion altered gene expression without reducing pausing, and by time course experiments, that cohesin does not impede elongation along the *EcR* gene. Thus we posit that cohesin regulates a specific step in transcription in a context-dependent manner.

Based on an ecdysone-induced wave of RNA synthesis we infer that PolII moves at just over 2 kb per minute along *EcR*. Cohesin, which has a chromosomal binding half-life of 2 to 9 minutes depending on the cell cycle stage [4], binds to the entire length of *EcR* [15]. Thus the cohesin half-life appears incompatible with the elongation rate unless polymerase passes through cohesin rings, or unless cohesin opens and recloses without release from the chromosome. The interior diameter of cohesin is approximately 35 by 50 nm [39], and the elongating holopolymerase is less than 20 nm in diameter [40], and thus the idea that polymerase passes through cohesin is tenable.

### Cohesin Regulates an Early Step in Transcription Distinct from Those Controlled by DSIF and NELF

The step regulated by cohesin at repressed genes is not polymerase recruitment. Cohesin depletion did not increase polymerase binding to *bbg* and *trol* promoters, and also did not decrease pausing at the genes tested, although it altered transcript levels. Thus at genes it represses, cohesin likely hinders a step downstream of polymerase pausing, and prior to elongation (Figure 7). Consistent with this idea, depletion of Spt5, which is required for both pausing and elongation [41-43], diminished the increase in transcripts caused by cohesin knockdown, and cohesin depletion increased the presence of Ser2P PolIII, the elongating form, further downstream in the genes examined.

What is the step regulated by cohesin? The Set2 histone methyltransferase is required for H3K36me3, binds to the phosphorylated CTD of elongating PolII, and in yeast lacking Set2, transcription is sensitive to 6-azauracil, suggesting impaired elongation [44-47]. Cohesin-

binding genes lack H3K36me3, and thus we posit that cohesin might inhibit binding of Set2 to the phosphorylated PolII CTD, or Set2 activity (Figure 7B). At genes repressed by cohesin, Set2 may be limiting for transition from pausing to elongation.

Cohesin does not repress all genes, and thus the cohesin-regulated step is not always limiting for transcription. One possibility is that the net effect of cohesin is a balance between facilitating long-range activation and promoter repression, and another is that the cohesin-regulated step is limiting when other proteins also repress a gene. This latter idea is consistent with the finding that the *E(spl)-C* and *invected-engrailed* genes are partially repressed by PcG proteins in BG3 cells [27]. *bbg* and *trol* are not PcG targets in BG3 cells, but may be targeted by unknown repressors.

Cohesin usually does not bind PcG-targeted regions [15], but in the rare overlaps, cohesin repression is stronger than average [27]. These genes show expression levels that range from low to high, indicating that co-repression by cohesin and PcG proteins restrains, but doesn't silence transcription. This restrained state is frequent in mouse embryonic stem cells, and many genes that increase in expression with cohesin depletion in these cells have both the H3K4me3 histone modification characteristic of active genes, and the H3K27me3 PcG silencing modification [1, 17, 48]. Cohesin and PcG proteins may target closely linked steps in the transition from pausing to elongation, because PcG proteins also bind preferentially to genes with paused polymerase [49], and cohesin interacts with PcG complexes in nuclear extracts [50].

### **Implications of Cohesin and Paused Promoters for Gene Regulation by Chromatin Structure**

Association of cohesin with paused polymerase genes has implications for factors that control gene expression through chromatin structure, including domain boundaries, insulators, and dosage compensation complexes. Paused promoters have insulator activity that blocks long-range enhancer promoter interactions [51], and cohesin may facilitate this activity, as it does for CTCF in mammalian cells [19, 20]. H3K36me3 facilitates binding of the MSL dosage compensation complex that upregulates X-linked gene expression of in males [52], and thus lack of H3K36me3 may decrease compensation of cohesin-binding genes. MSL increases elongating polymerase levels [53], suggesting that cohesin and MSL may directly oppose each other at the transcriptional level.

### **What Recruits Cohesin to Genes with Paused Polymerase?**

Our data argue that the presence of cohesin and the NELF and DSIF pausing factors at the same genes are not interdependent. We postulate that the GAF factor binding sites, and/or the downstream TG repeats at these genes facilitate recruitment of both pausing factors and cohesin. GAF may recruit NELF at some genes [30], but GAF depletion does not mimic cohesin knockdown, indicating that GAF is not required for cohesin binding. The strand-specific enrichment of TG repeats in the transcribed region is particularly striking. We speculate, therefore, that UG repeats in the initial nascent transcripts may bind proteins that recruit Nipped-B, which then loads cohesin. The same proteins might also recruit NELF, which is transferred back to promoter-bound polymerase to induce pausing.

## **Experimental Procedures**

### **Computational Analysis**

Published microarray chromatin immunoprecipitation (ChIP-chip) data was processed using the MAT program [54]. The MAT score is a measure of enrichment that accounts for differences in oligonucleotide sequences, and is used to estimate probability of binding at



defined p values. Nipped-B, Smc1 and PolII ChIP-chip data in BG3 and Sg4 cells was from reference 15 (GEO GSE16152), and GAF, NELF-B and NELF-E data from S2 cells was from reference 30 (ArrayExpress E-MEXP-1547). Histone H3K36 methylation and GAF ChIP-chip data for BG3 cells were from the modENCODE project [submit.modencode.org/submit/public/list, 29]. Correlations between MAT scores were calculated using R [R Foundation for Statistical Computing, Vienna, 2007; ISBN 3-900051-07-0; www.R-project.org].

We wrote computer programs to identify DNA sequences that are enriched in cohesin-binding genes. Using annotation type “gene” in *Drosophila* genome release 5.28, we identified genes that bind PolII, Nipped-B and Smc1 with p value  $\leq 10^{-3}$  at the start site in both BG3 and Sg4 cells (Supplemental File S1), and genes that bind PolII but not Nipped-B or Smc1 in both cell types (File S2). We determined the number of occurrences of all possible five nucleotide sequences at each position from  $-1000$  to  $+1000$  relative to the TSS in both gene sets. The frequency of a sequence in a gene group was defined as the total number of occurrences in the group divided by the total number of five nucleotide DNA sequences (1,009,976 for cohesin-binding genes, 2,075,840 for non-binding). Differences in sequence frequencies between the two groups were normalized to identify large differences:

$$|(frequency\_in\_binding)-(frequency\_in\_non-binding)| \div (frequency\_in\_binding+frequency\_in\_non-binding).$$

To determine if an enriched sequence occurred in a specific position, we measured the frequency of that sequence at each nucleotide position relative to the TSS, where the frequency was defined as the number of occurrences at each position divided by the number of genes in each group. Frequency plots were smoothed by averaging the frequencies at the surrounding 25 nucleotides for each position.

To compare sequence frequencies in non-coding (5' UTR or intron) and coding sequences, we plotted the frequency at each nucleotide position relative to the TSS, but genes in which that nucleotide position was either coding or non-coding were removed from the gene group.

### RNAi Knockdown and Transcript Quantification

BG3 cells were cultured and proteins depleted by dsRNA treatment as described [27]. PCR primers used to make templates for dsRNA synthesis are in Table S1, except for Rad21 and Nipped-B, which are described elsewhere [27]. Templates were designed to avoid off-target matches of  $\geq 19$  nucleotides using *Drosophila* RNAi Screening Center (<http://www.flyrnai.org/>) online tools. The *EcR* gene was induced by addition of ecdysone (20-hydroxyecdysone, Sigma) to the culture media at 2 micromoles per liter.

To quantify protein depletion, total cell protein extracts were subjected to SDS-PAGE western blots as described [27]. The GAF, NELF-B, NELF-E, Rad21, Nipped-B, and Spt5 antibodies are described elsewhere [14, 30, 55, 56].

Total RNA was isolated using columns according to the manufacturer's protocol (Zymo Research), and quantified by real-time quantitative PCR [27]. Primers for quantitative real-time reverse transcription-PCR are in Table S1 and elsewhere [27].

### Permanganate Footprinting

KMnO<sub>4</sub> footprinting was performed as described [34], using primers in Table S1.

## Chromatin Immunoprecipitation

Chromatin immunoprecipitation was performed as described [15, 55]. Karen Adelman (NIEHS) provided Rpb3 antibodies. Precipitated DNA was quantified using RT-PCR. Enrichment was calculated relative to a genomic region upstream of *engrailed* that doesn't bind PolII or cohesin (Chr 2R nucleotides 7436002-7436080).

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

The authors thank Karen Adelman and John Lis for generously providing antibodies, and Tomasz Heyduk for helpful comments on the manuscript. This work was supported by grants from the NIH to DD (GM055683) and DG (GM047477).

## References

1. Dorsett D. Cohesin: genomic insights into controlling gene transcription and development. *Curr. Opin. Genet. Dev.* 2011; 21:199–206. [PubMed: 21324671]
2. Dorsett D, Krantz ID. On the molecular etiology of Cornelia de Lange syndrome. *Ann. N. Y. Acad. Sci.* 2009; 1151:22–37. [PubMed: 19154515]
3. Nasmyth K, Haering CH. Cohesin: its roles and mechanisms. *Annu. Rev. Genet.* 2009; 43:525–558. [PubMed: 19886810]
4. Gause M, Misulovin Z, Bilyeu A, Dorsett D. Dosage-sensitive regulation of cohesin chromosome binding and dynamics by Nipped-B, Pds5, and Wapl. *Mol. Cell. Biol.* 2010; 30:4940–4951. [PubMed: 20696838]
5. Gerlich D, Koch B, Dupeux F, Peters JM, Ellenberg J. Live-cell imaging reveals a stable cohesin-chromatin interaction after but not before DNA replication. *Curr. Biol.* 2006; 16:1571–1578. [PubMed: 16890534]
6. Xiong B, Gerton JL. Regulators of the cohesin network. *Annu. Rev. Biochem.* 2010; 79:131–153. [PubMed: 20331362]
7. Bernard P, Drogat J, Maure JF, Dheur S, Vaur S, Genier S, Javerzat JP. A screen for cohesin mutants uncovers Ssl3, the fission yeast counterpart of the cohesin loading factor Scc4. *Curr. Biol.* 2006; 16:875–881. [PubMed: 16682348]
8. Ciosk R, Shirayama M, Shevchenko A, Tanaka T, Toth A, Shevchenko A, Nasmyth K. Cohesin's binding to chromosomes depends on a separate complex consisting of Scc2 and Scc4 proteins. *Mol. Cell.* 2000; 5:243–254. [PubMed: 10882066]
9. Seitan VC, Banks P, Laval S, Majid NA, Dorsett D, Rana A, Smith J, Bateman A, Krpic S, Hostert A, et al. Metazoan Scc4 homologs link sister chromatid cohesion to cell and axon migration guidance. *PLoS Biol.* 2006; 4:e242. [PubMed: 16802858]
10. Takahashi TS, Basu A, Bermudez V, Hurwitz J, Walter JC. Cdc7-Drf1 kinase links chromosome cohesion to the initiation of DNA replication in *Xenopus* egg extracts. *Genes Dev.* 2008; 22:1894–1905. [PubMed: 18628396]
11. Tomonaga T, Nagao K, Kawasaki Y, Furuya K, Murakami A, Morishita J, Yuasa T, Sutani T, Kearsey SE, Uhlmann F, et al. Characterization of fission yeast cohesin: essential anaphase proteolysis of Rad21 phosphorylated in the S phase. *Genes Dev.* 2000; 14:2757–2770. [PubMed: 11069892]
12. Ivanov D, Nasmyth K. A topological interaction between cohesin rings and a circular minichromosome. *Cell.* 2005; 122:849–860. [PubMed: 16179255]
13. Ivanov D, Nasmyth K. A physical assay for sister chromatid cohesion in vitro. *Mol. Cell.* 2007; 27:300–310. [PubMed: 17643378]

14. Gause M, Webber HA, Misulovin Z, Haller G, Rollins RA, Eissenberg JC, Bickel SE, Dorsett D. Functional links between *Drosophila* Nipped-B and cohesin in somatic and meiotic cells. *Chromosoma*. 2008; 117:51–66. [PubMed: 17909832]
15. Misulovin Z, Schwartz YB, Li XY, Kahn TG, Gause M, MacArthur S, Fay JC, Eisen MB, Pirrotta V, Biggin MD, et al. Association of cohesin and Nipped-B with transcriptionally active regions of the *Drosophila melanogaster* genome. *Chromosoma*. 2008; 117:89–102. [PubMed: 17965872]
16. MacAlpine HK, Gordân R, Powell SK, Hartemink AJ, MacAlpine DM. *Drosophila* ORC localizes to open chromatin and marks sites of cohesin complex loading. *Genome Res*. 2010; 20:201–211. [PubMed: 19996087]
17. Kagey MH, Newman JJ, Bilodeau S, Zhan Y, Orlando DA, van Berkum NL, Ebmeier CC, Goossens J, Rahl PB, Levine SS, et al. Mediator and cohesin connect gene expression and chromatin architecture. *Nature*. 2011; 467:430–435. [PubMed: 20720539]
18. Liu J, Zhang Z, Bando M, Itoh T, Deardorff MA, Clark D, Kaur M, Tandy S, Kondoh T, Rappaport E, et al. Transcriptional dysregulation in *NIPBL* and cohesin mutant human cells. *PLoS Biol*. 2009; 7:e1000119. [PubMed: 19468298]
19. Parelho V, Hadjur S, Spivakov M, Leleu M, Sauer S, Gregson HC, Jarmuz A, Canzonetta C, Webster Z, Nesterova T, et al. Cohesins functionally associate with CTCF on mammalian chromosome arms. *Cell*. 2008; 132:422–433. [PubMed: 18237772]
20. Wendt KS, Yoshida K, Itoh T, Bando M, Koch B, Schirghuber E, Tsutsumi S, Nagae G, Ishihara K, Mishiro T, et al. Cohesin mediates transcriptional insulation by CCCTC-binding factor. *Nature*. 2008; 451:796–801. [PubMed: 18235444]
21. Rollins RA, Morcillo P, Dorsett D. Nipped-B, a *Drosophila* homologue of chromosomal adherins, participates in activation by remote enhancers in the *cut* and *Ultrabithorax* genes. *Genetics*. 1999; 152:577–593. [PubMed: 10353901]
22. Chien R, Zeng W, Kawauchi S, Bender MA, Santos R, Gregson HC, Schmiesing JA, Newkirk D, Kong X, Ball AR Jr. et al. Cohesin mediates chromatin interactions that regulate mammalian  $\beta$ -globin expression. *J. Biol. Chem*. 2011; 286:17870–17878. [PubMed: 21454523]
23. Hadjur S, Williams LM, Ryan NK, Cobb BS, Sexton T, Fraser P, Fisher AG, Merckenschlager M. Cohesins form chromosomal cis-interactions at the developmentally regulated *IFNG* locus. *Nature*. 2009; 460:410–413. [PubMed: 19458616]
24. Hou C, Dale R, Dean A. Cell type specificity of chromatin organization mediated by CTCF and cohesin. *Proc. Natl. Acad. Sci. U. S. A.* 2010; 107:3651–3656. [PubMed: 20133600]
25. Mishiro T, Ishihara K, Hino S, Tsutsumi S, Aburatani H, Shirahige K, Kinoshita Y, Nakao M. Architectural roles of multiple chromatin insulators at the human *apolipoprotein* gene cluster. *EMBO J*. 2009; 28:1234–1245. [PubMed: 19322193]
26. Nativio R, Wendt KS, Ito Y, Huddleston JE, Uribe-Lewis S, Woodfine K, Krueger C, Reik W, Peters JM, Murrell A. Cohesin is required for higher-order chromatin conformation at the imprinted *IGF2-H19* locus. *PLoS Genet*. 2009; 5:e1000739. [PubMed: 19956766]
27. Schaaf CA, Misulovin Z, Sahota G, Siddiqui AM, Schwartz YB, Kahn TG, Pirrotta V, Gause M, Dorsett D. Regulation of the *Drosophila* *Enhancer of split* and *invected-engrailed* gene complexes by sister chromatid cohesion proteins. *PLoS One*. 2009; 4:e6202. [PubMed: 19587787]
28. Ayala YM, Pantano S, D'Ambrogio A, Buratti E, Brindisi A, Marchetti C, Romano M, Baralle FE. Human, *Drosophila*, and *C. elegans* TDP43: nucleic acid binding properties and splicing regulatory function. *J. Mol. Biol*. 2005; 348:575–588. [PubMed: 15826655]
29. Celniker SE, Dillon LA, Gerstein MB, Gunsalus KC, Henikoff S, Karpen GH, Kellis M, Lai EC, Lieb JD, MacAlpine DM, et al. Unlocking the secrets of the genome. *Nature*. 2009; 459:927–930. [PubMed: 19536255]
30. Lee C, Li X, Hechmer A, Eisen M, Biggin MD, Venters BJ, Jiang C, Li J, Pugh BF, Gilmour DS. NELF and GAGA factor are linked to promoter-proximal pausing at many genes in *Drosophila*. *Mol. Cell. Biol*. 2008; 28:3290–3300. [PubMed: 18332113]
31. Gilchrist DA, Nechaev S, Lee C, Ghosh SK, Collins JB, Li L, Gilmour DS, Adelman K. NELF-mediated stalling of Pol II can enhance gene expression by blocking promoter-proximal nucleosome assembly. *Genes Dev*. 2008; 22:1921–1933. [PubMed: 18628398]

32. Yamaguchi Y, Takagi T, Wada T, Yano K, Furuya A, Sugimoto S, Hasegawa J, Handa H. NELF, a multisubunit complex containing RD, cooperates with DSIF to repress RNA polymerase II elongation. *Cell*. 1999; 97:41–51. [PubMed: 10199401]
33. Nechaev S, Fargo DC, dos Santos G, Liu L, Gao Y, Adelman K. Global analysis of short RNAs reveals widespread promoter-proximal stalling and arrest of Pol II in *Drosophila*. *Science*. 2010; 327:335–338. [PubMed: 20007866]
34. Gilmour DS, Fan R. Detecting transcriptionally engaged RNA polymerase in eukaryotic cells with permanganate genomic footprinting. *Methods*. 2009; 48:368–374. [PubMed: 19272453]
35. Pauli A, van Bemmel JG, Oliveira RA, Itoh T, Shirahige K, van Steensel B, Nasmyth K. A direct role for cohesin in gene regulation and ecdysone response in *Drosophila* salivary glands. *Curr. Biol*. 2010; 20:1787–1798. [PubMed: 20933422]
36. Schuldiner O, Berdnik D, Levy JM, Wu JS, Luginbuhl D, Gontang AC, Luo L. piggyBac-based mosaic screen identifies a postmitotic function for cohesin in regulating developmental axon pruning. *Dev. Cell*. 2008; 14:227–238. [PubMed: 18267091]
37. Muse GW, Gilchrist DA, Nechaev S, Shah R, Parker JS, Grissom SF, Zeitlinger J, Adelman K. RNA polymerase is poised for activation across the genome. *Nat. Genet*. 2007; 39:1507–1511. [PubMed: 17994021]
38. Zeitlinger J, Stark A, Kellis M, Hong JW, Nechaev S, Adelman K, Levine M, Young RA. RNA polymerase stalling at developmental control genes in the *Drosophila melanogaster* embryo. *Nat. Genet*. 2007; 39:1512–1516. [PubMed: 17994019]
39. Anderson DE, Losada A, Erickson HP, Hirano T. Condensin and cohesin display different arm conformations with characteristic hinge angles. *J. Cell. Biol*. 2002; 156:419–424. [PubMed: 11815634]
40. Gnatt AL, Cramer P, Fu J, Bushnell DA, Kornberg RD. Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 Å resolution. *Science*. 2001; 292:1876–1882. [PubMed: 11313499]
41. Hartzog GA, Wada T, Handa H, Winston F. Evidence that Spt4, Spt5, and Spt6 control transcription elongation by RNA polymerase II in *Saccharomyces cerevisiae*. *Genes Dev*. 1998; 12:357–369. [PubMed: 9450930]
42. Wada T, Takagi T, Yamaguchi Y, Ferdous A, Imai T, Hirose S, Sugimoto S, Yano K, Hartzog GA, Winston F, et al. DSIF, a novel transcription elongation factor that regulates RNA polymerase II processivity, is composed of human Spt4 and Spt5 homologs. *Genes Dev*. 1998; 12:343–356. [PubMed: 9450929]
43. Wu CH, Yamaguchi Y, Benjamin LR, Horvat-Gordon M, Washinsky J, Enerly E, Larsson J, Lambertsson A, Handa H, Gilmour D. NELF and DSIF cause promoter proximal pausing on the *hsp70* promoter in *Drosophila*. *Genes Dev*. 2003; 17:1402–1414. [PubMed: 12782658]
44. Krogan NJ, Kim M, Tong A, Golshani A, Cagney G, Canadien V, Richards DP, Beattie BK, Emili A, Boone C, et al. Methylation of histone H3 by Set2 in *Saccharomyces cerevisiae* is linked to transcriptional elongation by RNA polymerase II. *Mol. Cell. Biol*. 2003; 23:4207–4218. [PubMed: 12773564]
45. Li B, Howe L, Anderson S, Yates JR III, Workman JL. The Set2 histone methyltransferase functions through the phosphorylated carboxylterminal domain of RNA polymerase II. *J. Biol. Chem*. 2003; 278:8897–8903. [PubMed: 12511561]
46. Li J, Moazed D, Gygi SP. Association of the histone methyltransferase Set2 with RNA polymerase II plays a role in transcription elongation. *J. Biol. Chem*. 2002; 277:49383–49388. [PubMed: 12381723]
47. Xiao T, Hall H, Kizer KO, Shibata Y, Hall MC, Borchers CH, Strahl BD. Phosphorylation of RNA polymerase II CTD regulates H3 methylation in yeast. *Genes Dev*. 2003; 17:654–663. [PubMed: 12629047]
48. Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, Fry B, Meissner A, Wernig M, Plath K, et al. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell*. 2006; 125:315–326. [PubMed: 16630819]

49. Enderle D, Beisel C, Stadler MB, Gerstung M, Athri P, Paro R. Polycomb preferentially targets stalled promoters of coding and noncoding transcripts. *Genome Res.* 2011; 21:216–226. [PubMed: 21177970]
50. Strübbe G, Popp C, Schmidt A, Pauli A, Ringrose L, Beisel C, Paro R. Polycomb purification by in vivo biotinylation tagging reveals cohesin and Trithorax group proteins as interaction partners. *Proc. Natl. Acad. Sci. U. S. A.* 2011; 108:5572–5577. [PubMed: 21415365]
51. Chopra VS, Cande J, Hong JW, Levine M. Stalled Hox promoters as chromosomal boundaries. *Genes Dev.* 2009; 23:1505–1509. [PubMed: 19515973]
52. Larschan E, Alekseyenko AA, Gortchakov AA, Peng S, Li B, Yang P, Workman JL, Park PJ, Kuroda MI. MSL complex is attracted to genes marked by H3K36 trimethylation using a sequence-independent mechanism. *Mol. Cell.* 2007; 28:121–133. [PubMed: 17936709]
53. Larschan E, Bishop EP, Kharchenko PV, Core LJ, Lis JT, Park PJ, Kuroda MI. X chromosome dosage compensation via enhanced transcriptional elongation in *Drosophila*. *Nature.* 2011; 471:115–118. [PubMed: 21368835]
54. Johnson WE, Li W, Meyer CA, Gottardo R, Carroll JS, Brown M, Liu XS. Model-based analysis of tiling-arrays for ChIP-chip. *Proc. Natl. Acad. Sci. U. S. A.* 2006; 103:12457–12462. [PubMed: 16895995]
55. Dorsett D, Eissenberg JC, Misulovin Z, Martens A, Redding B, McKim K. Effects of sister chromatid cohesion proteins on *cut* gene expression during wing development in *Drosophila*. *Development.* 2005; 132:4743–4753. [PubMed: 16207752]
56. Andrulis ED, Guzmán E, Döring P, Werner J, Lis JT. High-resolution localization of *Drosophila* Spt5 and Spt6 at heat shock genes in vivo: roles in promoter proximal pausing and transcription elongation. *Genes Dev.* 2000; 14:2635–2649. [PubMed: 11040217]



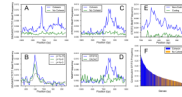
**Highlights**

Cohesin selectively binds genes that have paused RNA polymerase

Cohesin depletion alters gene expression without reducing polymerase pausing

Cohesin and pausing factors regulate closely linked steps at repressed genes

Cohesin inhibits transition of paused polymerase to elongation at repressed genes



**Figure 1.**

**Enrichment of GAF Binding Sites and GT Repeats in Cohesin-Binding Genes**

(A) Combined frequency of GAGAG and CTCTC, binding sequences for GAGA factor (GAF), at positions from  $-1$  kb to  $+1$  kb around the TSS of 506 cohesin-binding genes (blue) and 1040 genes that do not bind cohesin (green).

(B) GAGAG and CTCTC frequency increases with the stringency with which cohesin binding is called (blue,  $p \leq 10^{-15}$ ; green,  $p \leq 10^{-9}$ ; gray,  $p \leq 10^{-3}$ ).

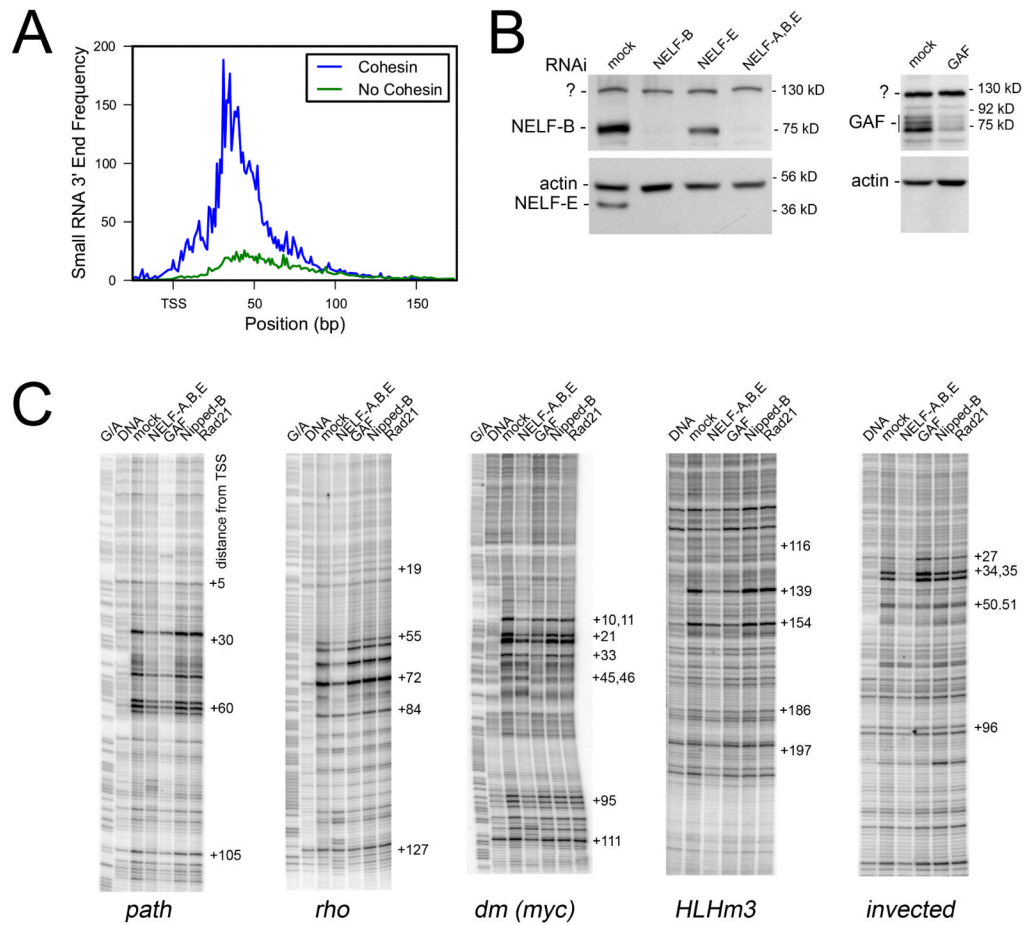
(C) Frequency of GTGTG from  $-1$  kb to  $+1$  kb in cohesin-binding (blue) and non-binding (green) genes.

(D) Frequency of GTGTG (blue) and CACAC (green) in cohesin-binding genes.

(E) Frequency of GTGTG in non-coding (blue) and coding (green) sequences of cohesin-binding genes.

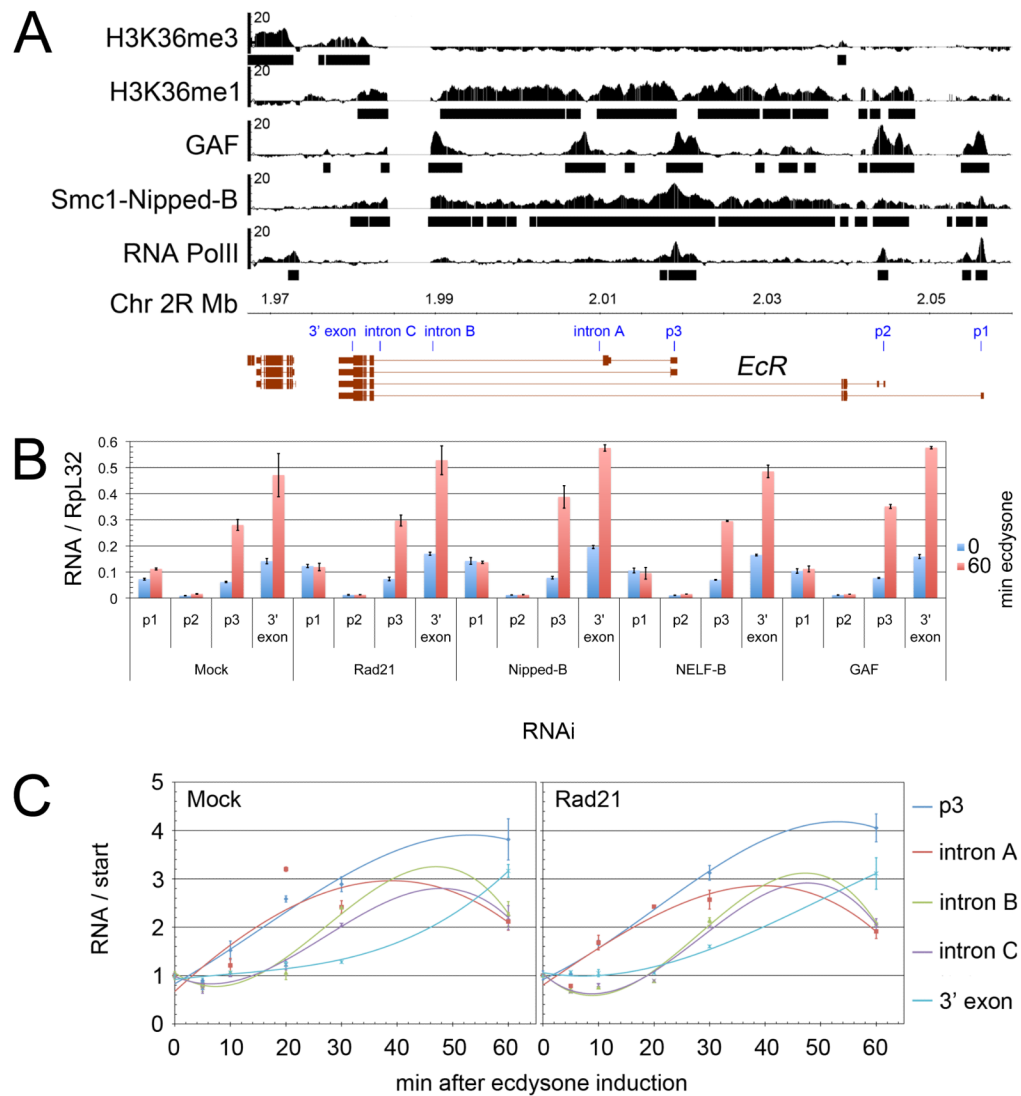
(F) Frequency of a GT or TG dinucleotide sequence following another in the 506 cohesin-binding (blue) and 1040 non-binding genes (yellow-orange) in the 1 kb region downstream of the TSS. The genes are in order of decreasing consecutive dinucleotide frequency.

Comparison of genome-wide binding of Nipped-B to GAGA factor (GAF) in Figure S1 indicates that the GAGAG sequences enriched in cohesin-binding genes are functional binding sites.

**Figure 2.**

Cohesin Selectively Binds Genes with Paused Polymerase but Cohesin Depletion Does Not Reduce Pausing

(A) Average abundance of 3' ends of short RNA transcripts, a marker of paused polymerase [33] in cohesin-binding (blue) and non-binding (green) genes at the transcription start sites. (B) Western blots showing RNAi depletion of NELF-B, NELF-E, and GAF in BG3 cells. (C) Effects of NELF, GAF, Nipped-B and Rad21 depletion on polymerase pausing at the *path*, *rho*, *dm*, *HLHm3* and *invected* genes as measured by  $\text{KMnO}_4$  footprinting. G/A lanes contain products from BG3 genomic DNA from the Maxam-Gilbert G/A reaction, and DNA lanes contain products from  $\text{KMnO}_4$ -treated purified genomic DNA. The remaining lanes contain products produced from genomic DNA isolated from  $\text{KMnO}_4$ -treated BG3 cells after the indicated RNAi treatments.

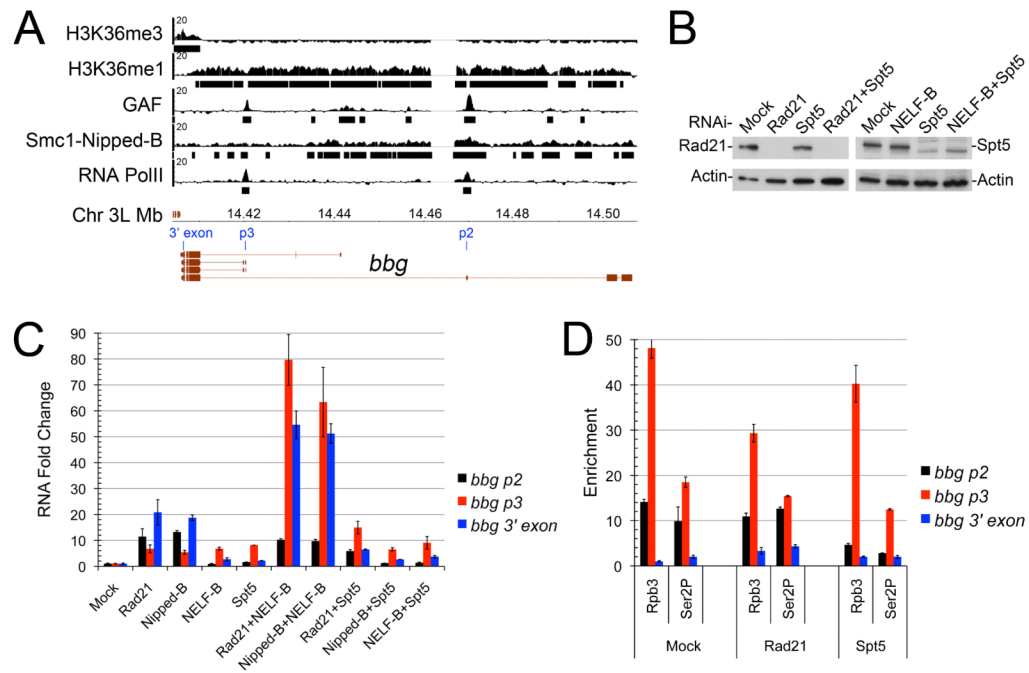


**Figure 3.**  
 Cohesin Does Not Interfere with Transcriptional Induction or Elongation at *EcR* Gene in BG3 Cells

(A) Map of *EcR* and association of PolII, combined Smc1-Nipped-B, GAF, and the H3K36me1 and H3K36me3 histone modifications determined by ChIP-chip in BG3 cells [15, 29, 30]. Bars underneath the ChIP profiles indicate where binding is called at  $p \leq 10^{-3}$ . Positions of RT-PCR probes are indicated in blue.

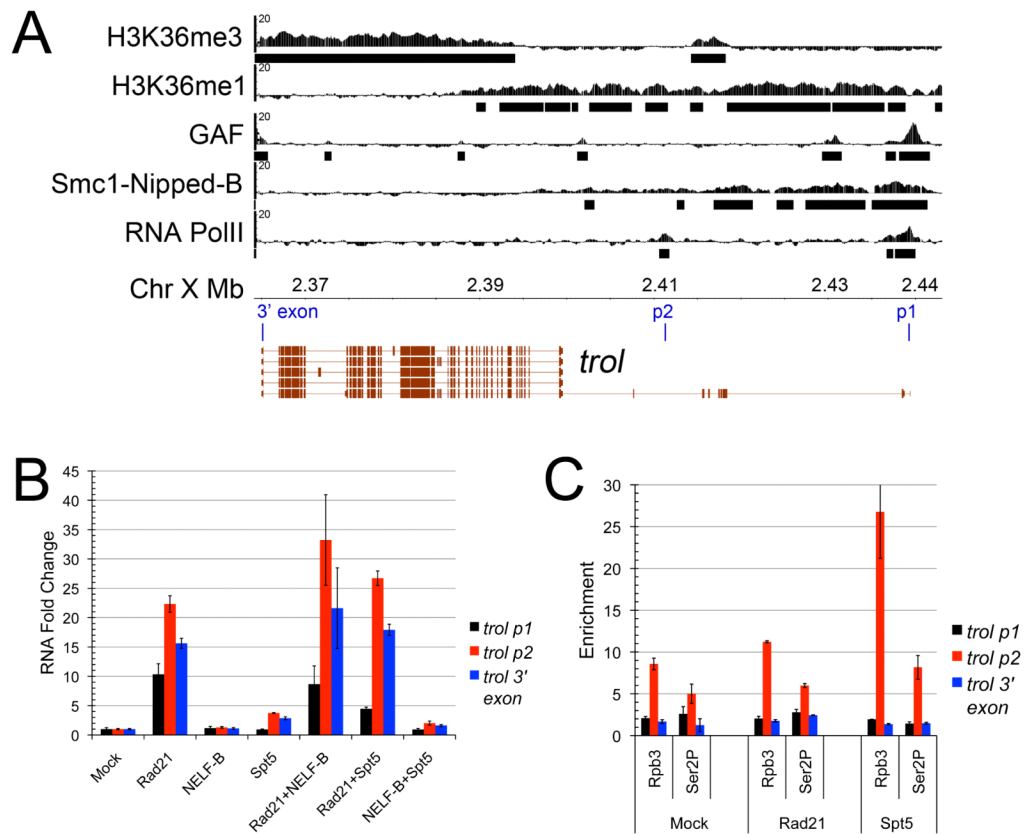
(B) Levels of transcripts containing p1, p2, p3 and 3' exon sequences before (blue) and after (red) ecdysone induction for 60 min. Cells were untreated (Mock) or pretreated with the indicated RNAi (Rad21, Nipped-B, NELF-B, GAF) for 4 to 6 days. Error bars are standard errors, calculated using all RT-PCR replicates in three independent experiments.

(C) Time courses of ecdysone induction showing the relative increases in transcripts containing the indicated probe sequences, with or without Rad21 depletion. Time courses after Nipped-B, GAF and NELF-B knockdown are in Figure S2. The curves shown are third order polynomial fits. Some replicates included 45 min time points (not shown) showing that intron B and C RNA reach peak levels around 45 min and then decrease. Error bars are standard errors, calculated using all RT-PCR replicates from three independent experiments.



**Figure 4.** Effects of Cohesin (Rad21), Nipped-B, Spt5 and NELF-B Depletion on *big bang* (*bbg*) Transcripts and PolII binding in BG3 Cells  
 (A) Map of *bbg* with ChIP-chip profiles. PolII ChIP identified p2 (unannotated) and p3 as the active promoters. RT-PCR probes are indicated in blue.  
 (B) Western blots showing typical RNAi depletion of Rad21 and Spt5.  
 (C) Changes in levels of transcripts containing the indicated probe sequences relative to control (Mock) after five days of the indicated RNAi treatments. Mock control p2 transcripts are ~0.6% the level of *RpL32* transcripts, p3 are ~0.15%, and 3' exon transcripts are ~0.22%. Error bars are standard errors, calculated using all RT-PCR replications from two independent experiments. Other times of RNAi treatment and other independent experiments are shown in Figure S4.  
 (D) Enrichment of the indicated probe sequences relative to an intergenic region (see Methods) after ChIP using anti-Rpb3 and anti-Ser2P PolII antibodies five days after RNAi treatment. Error bars are standard errors calculated using all RT-PCR replicates from ChIP of two to three independent chromatin preparations. ChIP for Rad21 and Nipped-B binding after Rad21 and Spt5 depletion are shown in Figure S3.



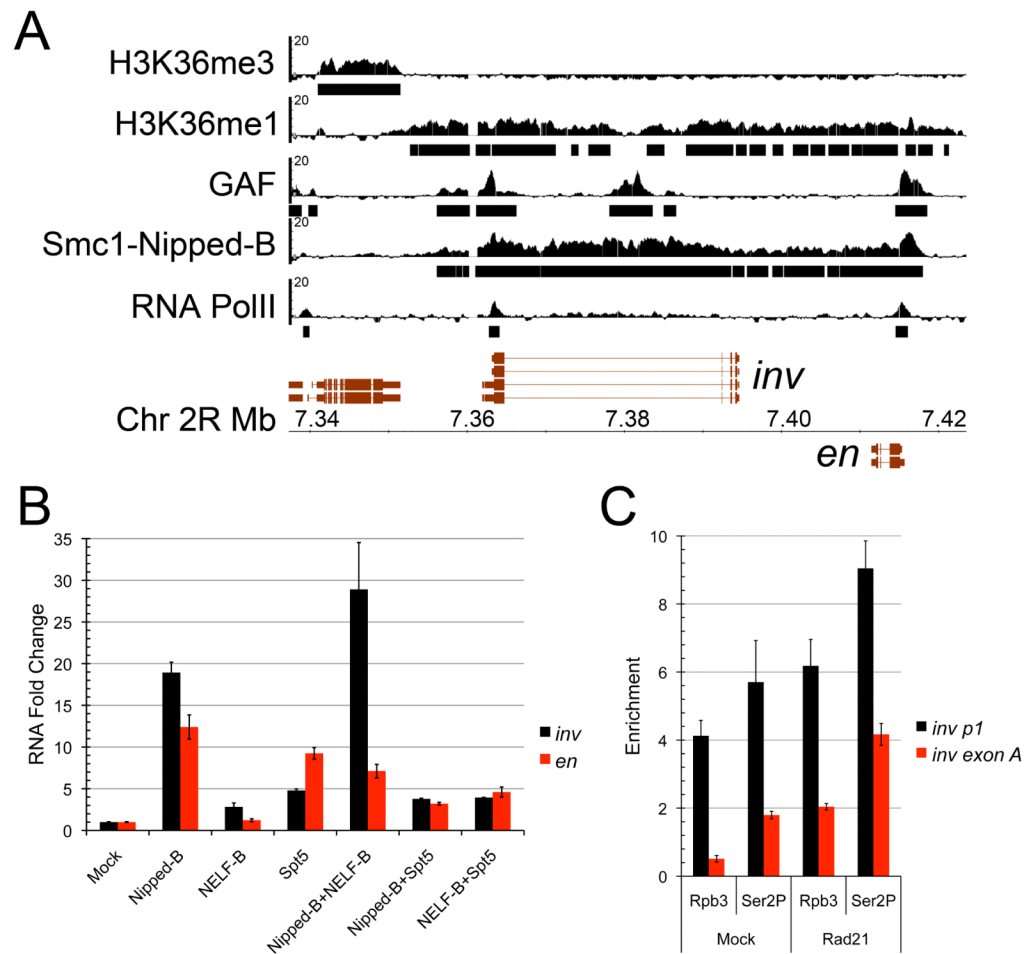


**Figure 5.** Effects of Rad21, NELF-B and Spt5 Depletion on Transcripts of *terribly reduced optic lobes* (*trol*) and PolII Binding in BG3 Cells

(A) Map of *trol*, with ChIP-chip data and locations of RT-PCR probes. PolII ChIP identified p1 and p2 (unannotated) as active promoters.

(B) Increase in transcripts containing the indicated probe sequences relative to control (Mock) after RNAi treatment for five days. Mock control p1 transcripts are ~0.5% the level of *RpL32* transcripts, p2 are ~0.6% and 3' exon transcripts are ~1.7%. Error bars are standard errors, calculated using all RT-PCR replications from two independent experiments. Other times of RNAi treatment, other RNAi treatments (Nipped-B, Nipped-B combined with NELF and Spt5) and RNAi treatment times are shown in Figure S5.

(C) ChIP using anti-Rpb3 and anti-Ser2P PolII antibodies five days after RNAi treatment. Error bars are standard errors calculated using all RT-PCR replicates from ChIP of two to three independent chromatin preparations. ChIP for Rad21 and Nipped-B binding after Rad21 and Spt5 depletion are shown in Figure S3.

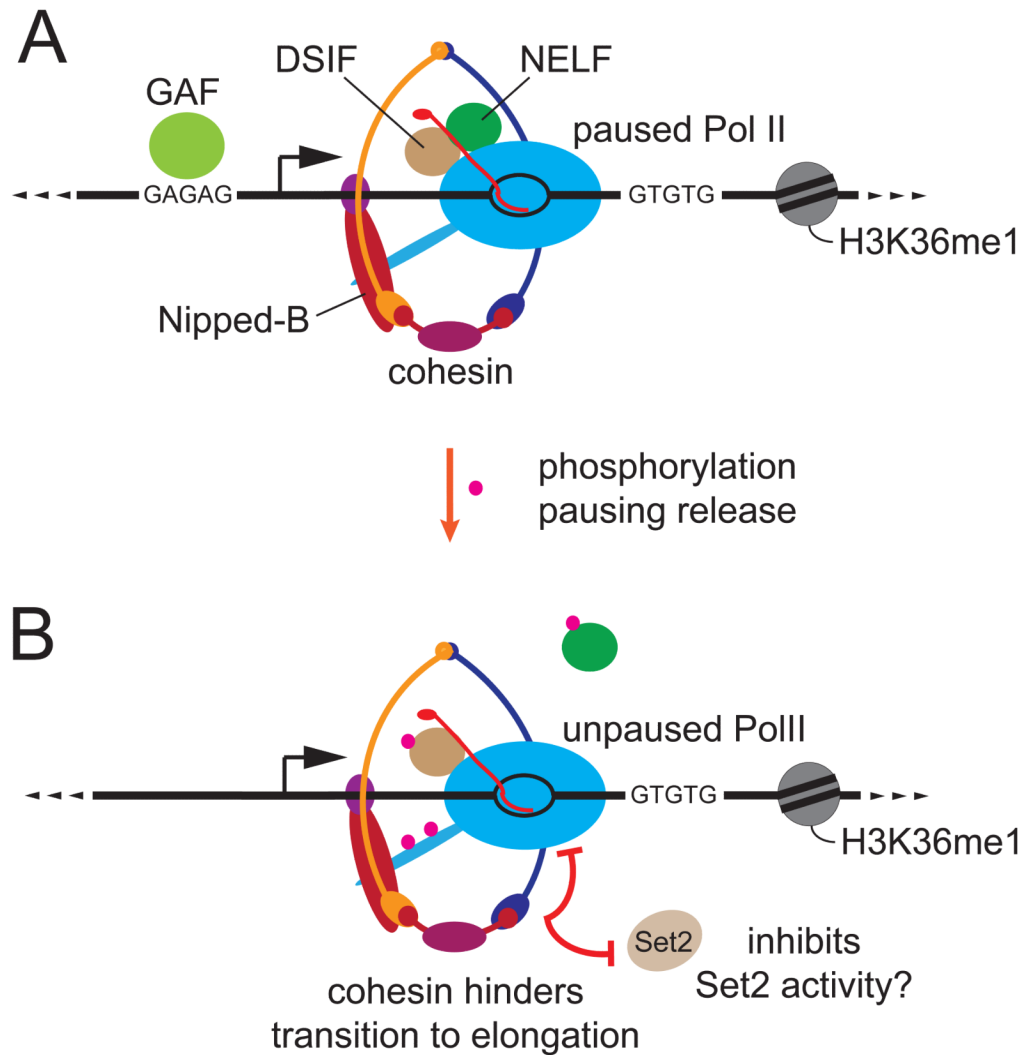
**Figure 6.**

Effects of Nipped-B, NELF-B and Spt5 Depletion on *invected* and *engrailed* Transcripts in BG3 Cells

(A) Map of the *invected-engrailed* complex with ChIP profiles. The RT-PCR probes for both genes are in the 5' exon (Table S1).

(B) Changes in transcript levels five days after the indicated RNAi treatments. Mock control *invected* transcripts are ~0.1% the level of *RpL32* transcripts, and *engrailed* transcripts are ~0.01%. Error bars are standard errors, calculated using all RT-PCR replications from two independent experiments. Additional days of treatment and Rad21 RNAi treatments are shown in Figure S7. Effects of Nipped-B, Rad21, NELF, and Spt5 depletion on E(spl)-C genes, which like *invected* and *engrailed*, are also PcG targets in BG3 cells [27], are shown in Figure S6.

(C) PolII ChIP for *invected* after five days of Rad21 RNAi treatment. Error bars are standard errors calculated using all RT-PCR replicates from two to three independent chromatin preparations. The *inv p1* probe covers nt 23-109 downstream from the TSS and *inv exon A* probe covers nt 691-802. ChIP for Rad21 and Nipped-B binding after Rad21 and Spt5 depletion are shown in Figure S3.



**Figure 7.**

Model for Effects of Cohesin on Repressed Genes with Paused RNA Polymerase

(A) Typical paused polymerase promoter with encircling cohesin, Nipped-B, upstream GAF, DSIF and NELF pausing factors, short nascent transcript (red), downstream GT repeats, and nucleosomes with H3K36me1. We theorize that GAF and the GU repeats in the initial nascent transcripts may independently recruit the pausing factors and Nipped-B, which then loads cohesin.

(B) After phosphorylation of NELF, DSIF (Spt5), and PolII by activator-recruited pTEFb, NELF is released, unpausing PolII. At repressed genes, we posit that cohesin hinders transition of newly unpaused PolII to elongation, possibly by blocking association of Set2 methyltransferase with PolII, or inhibiting Set2 activity.