Published in final edited form as:
Curr Opin Cell Biol. 2013 June ; 25(3): 327–333. doi:10.1016/j.ceb.2013.02.003.

Cohesin at active genes: a unifying theme for cohesin and gene expression from model organisms to humans

Dale Dorsett[1,*] and **Matthias Merkenschlager**[2,*]

[1]Edward A Doisy Department of Biochemistry and Molecular Biology, Saint Louis University School of Medicine, Saint Louis, Missouri, 63104, USA

[2]Lymphocyte Development Group, MRC Clinical Sciences Centre, Faculty of Medicine, Imperial College London, Du Cane Road, London W12 0NN, UK

Summary

Cohesin is an evolutionarily ancient multi-subunit protein complex with a deeply conserved function: it provides cohesion between sister chromatids from the time of DNA replication in S-phase until mitosis. This cohesion facilitates repair of damage that occurs during DNA replication, and, crucially, enforces faithful segregation of chromosomes upon cell division. Cohesin also influences gene expression, and relative to sister chromatid cohesion, gene expression is exquisitely sensitive to moderate changes in cohesin activity. Early studies revealed differences in cohesin's roles in gene expression between various organisms. In all organisms examined, however, cohesin marks a subset of active genes. This review focuses on the roles of cohesin at active genes, and to what extent these roles are conserved between organisms.

Introduction

Cohesin holds sister chromatids together from DNA replication in S phase until cell division, allowing for post-replicative DNA repair based on homology between sister chromatids, and proper segregation of chromosomes from mother to daughter cells through mitosis and meiosis (1, 2). The cohesin subunits form a ring-like structure that encircles DNA (Figure 1). Topologically bound cohesin is actively loaded by a complex containing Nipped-B/NIPBL/Scc2 and Mau-2/Scc4, sometimes referred as "kollerin". The stability of cohesin association with chromatin is regulated by a number of co-factors, including sororin, Wapl, and Pds5 (the Wapl-Pds5 complex is sometimes referred to as "releasin") (Figure 1).

The roles of cohesin and cohesin-loading factors in gene regulation were originally revealed by genetic studies in Drosophila, where reducing the dosage of the Drosophila Nipped-B cohesin loading factor kollerin subunit alter expression of specific homeobox genes during development (3). Consistent with these and other data from model organisms, human genetics identified the *NIPBL* kollerin subunit gene as the most frequently mutated gene in Cornelia de Lange Syndrome, a human disorder that alters physical and intellectual development (4, 5).

*Corresponding authors: dorsettd@slu.edu and matthias.merkenschlager@csc.mrc.ac.uk.

In yeast and flies, substantial reductions in cohesin dosage of more than 85% are required to disrupt cohesion and chromosome segregation, while small to moderate reductions can affect gene expression and development (6, 7). Genes that bind cohesin are more likely to change in expression, and changes in cohesin activity also alter gene expression in non-dividing cells as demonstrated for post-mitotic Drosophila neurons and non-dividing mouse thymocytes (7, 8, 9, 10, 11*). Taken together, these findings argue that cohesin can regulate gene expression directly, and not just indirectly as a consequence of cohesin's roles in cell division.

Although cohesin and its role in sister chromatid cohesion are deeply conserved, differences between organisms as to where cohesin binds along chromosome arms originally gave the impression that cohesin's roles in gene regulation has diverged substantially between organisms. While differences exist, further investigation has revealed similarities such as the association of cohesin with active genes, indicating that there are likely conserved core mechanisms.

Cohesin sites in mammalian cells

Cohesin associates with active genes

Cohesin is located primarily between convergently transcribed genes along budding yeast chromosome arms, and many of these sites lack the cohesin loading complex (12, 13, 14). In yeast, the strongest kollerin binding sites are promoters that bind the TFIIIC transcription factor, such as tRNA genes, and other strong promoters, such as at ribosomal protein genes (15). This gave rise to the notion that cohesin slides along chromosomes after loading, and may be pushed to some of its major binding sites by RNA polymerase (12, 14, 16). It was a surprise, therefore, when cohesin and kollerin were mapped genome-wide in Drosophila, and it was found that cohesin and its loading factor Nipped-B co-localize virtually completely. Drosophila cohesin and Nipped-B are found almost exclusively at a subset of transcriptionally active genes and DNA replication origins, and are excluded from silenced genes (17, 18*). At active genes, cohesin and Nipped-B binding generally peaks at the promoter, but often spreads into the transcribed region. A similar pattern of cohesin and Nipbl binding has been described at the promoters and enhancers of active genes in mammalian cells (19*, 20*).

Cohesin associates with CTCF in mammalian cells

In addition, mammalian cells have a large number of cohesin binding sites that also bind the sequence-specific DNA binding protein CTCF (21–24). There are two forms of the stromalin cohesin subunit in mammalian cells, SA1 and SA2, and like the cohesin at gene promoters, the cohesin at CTCF sites largely contains SA1 (25). CTCF binding often marks transitions between domains of chromatin with distinct properties, including histone modifications (26), domains that contact the nuclear lamina (27), and long-range interaction data inferred from chromosome conformation capture data (28, 29). Despite this enrichment, the vast majority of CTCF sites do not appear to mark domain boundaries, and numerous domain transitions lack CTCF binding (26–29). CTCF can block functional cooperation between enhancers and promoters (30), and this 'insulator' activity may require cohesin (22).

Many, but not all, CTCF-cohesin sites are shared between different cell types and most are not associated with promoters. Mammalian CTCF has been shown to interact with the Scc3/SA2 subunit of the cohesin complex (31). Although Drosophila has a CTCF homolog, it does not significantly co-localize with cohesin. Most CTCF-cohesin sites lack Nipbl (19*) and it may be that cohesin translocates from its loading sites to CTCF binding sites. Alternatively, CTCF may recruit cohesin directly (23, 31).

What brings cohesin to active genes?

Nipbl is part of the cohesin loading complex, and thus data that map Nipbl to regulatory elements of active genes suggest that cohesin is loaded there. For completeness, however, we note that loading at these sites has not been shown directly, and that the published global Nipbl map in mammalian cells relied on an antibody that is difficult to use. The fraction of ChIP reads that maps to peaks is unusually small (19*; GSE8395). Recent mapping of the Mau-2 kollerin subunit in mammalian cells, however, gives similar results, with frequent binding at enhancers and promoters, and infrequent localization at CTCF sites (Katsuhiko Shirahige, personal communication - Email forwarded by DD Nov 21). While experimental evidence for site-specific loading is difficult to obtain, data from Smc hydrolysis mutant cohesin complexes in yeast are suggestive: these cohesin complexes are unstable and associate with chromatin transiently at the fixed sites of Scc2 binding, as if they can still interact with the kollerin complexes, but cannot be loaded onto the chromosome (32).

In Drosophila, in vivo fluorescence recovery after photobleaching (FRAP) experiments show that reduction of the Nipped-B kollerin subunit specifically decreases the fraction of chromosomally-bound cohesin that has the longest residence time, which is presumed to be topologically bound, and also reduces promoter-bound cohesin detected by ChIP, which argues that ChIP detects primarily loaded cohesin, and not cohesin that transiently interacts with chromosomes (33, 34*). Cohesin ChIP signals at promoters are also reduced in human and mouse cells heterozygous mutant for *Nipbl* (35, 36*). A significant fraction of Nipped-B, however, also has the same long residence time as stable cohesin in FRAP studies, and Nipped-B co-localizes almost completely with cohesin, suggesting that in higher organisms, loaded cohesin may also recruit kollerin complexes. Consistent with this idea, cohesin depletion reduces Nipped-B binding to various extents at the promoters assayed (34*). Thus it remains unresolved in higher organisms where most cohesin loading occurs, although it currently seems reasonable to assume that it is likely where kollerin levels are the highest. With these reservations in mind, what are the mechanisms that bring cohesin to regulatory elements of active genes?

Mediator

Cohesin co-purifies and co-localizes with the Mediator complex, which binds many active promoters and transcriptional enhancers (19*, 37). Mediator is implicated in several steps in transcription, including facilitating interactions between enhancer-bound activators and the basal transcription machinery, transcription initiation, promoter escape and elongation (38). Roughly half the Mediator-binding sites bind Nipbl, indicating that Mediator alone is insufficient to recruit Nipbl and cohesin (19*).

Transcription factors

In yeast, reduction of TFIIIC decreases the amount of Scc2 on chromosomes, suggesting that TFIIIC may recruit or stabilize kollerin binding (15). Similarly, cohesin co-localizes with several tissue-specific transcription factors, including the estrogen receptor in breast cancer cell lines, liver-specific factors in liver cells, and pluripotency factors in embryonic stem cells (20*, 39*, 40). In mouse liver cells, cohesin binding sites that lack CTCF, a third of which are extragenic, generally bind multiple liver-specific transcription factors, and often lack optimal binding site sequences for some of them, suggesting that cohesin may facilitate transcription factor binding (39*).

Tissue-specific transcription factors associate with many active genes in their respective cell types, so it is difficult to determine whether cohesin and the transcription factors directly influence each other’s binding, or if their overlap simply reflects cohesin’s preference for

active genes. In the case of the Nanog pluripotency factor, there is evidence for direct interaction with cohesin (40). To explain cohesin loading, however, it would have to be postulated that cohesin recruited by Nanog in turn recruits Nipbl, as discussed above.

Roughly a third of cohesin-binding genes in Drosophila bind the GAGA factor (GAF) zinc finger and BTB/POZ domain protein 100 bp upstream of the transcription start site (34*). GAF, encoded by the *Trithorax-like* (*Trl*) gene, plays roles in both gene activation and silencing, and also binds many Polycomb Response Elements (PREs) that are critical for silencing by Polycomb group (PcG) complexes (41). Vertebrate GAF homologs have only recently been described, and other than the fact that they bind within homeobox (HOX) gene complexes, little is known about their genomic DNA binding patterns (41). The potential roles of GAF in cohesin loading remain to be investigated.

Chromatin structure

In Drosophila, although cohesin-binding genes tend to be highly expressed, they are deficient for the histone H3 lysine 36 trimethylation (H3K36me3) mark made by the Set2 methyltransferase that travels with elongating polymerase at virtually all active genes that lack cohesin (34*). H3K36me3 is also not detected at most cohesin-binding genes in mouse liver cells, suggesting that this is conserved between flies and mammals (39*). Cohesin-binding patterns, however, are not substantially altered in developing Drosophila wings that lack Set2 and H3K36me3 (Amanda Koenig, Ziva Misulovin, Maria Gause, Cheri A Schaaf, DD, unpublished observations), indicating that H3K36me3 does not block cohesin binding at the genes that lack cohesin.

In Drosophila, cohesin, kollerin and RNA polymerase extensively co-localize with Mi-2, an ATP-dependent nucleosome remodeler, and increases in Mi-2 levels destabilize cohesin binding (42). Mi-2 is a conserved subunit of the NuRD repressor complex, which also has histone deacetylase and demethylase activity, but is often found at active genes, including the pluripotency genes that bind cohesin in mammalian embryonic stem cells (42, 43). Co-localization of cohesin and Mi-2 is likely conserved in mammals, and may involve direct interaction, because cohesin and NuRD components have been co-purified from human cells (44). One mechanism by which NuRD may attenuate and balance gene expression is by down-regulating cohesin binding. Supporting this idea, the dominant small wing phenotype of Drosophila *Nipped-B* kollerin mutants is fully suppressed by a heterozygous *Mi-2* null mutation (42).

The cohesin subunit RAD21 was reported to directly interact with the NuRD ATPase subunit SNF2h, and cohesin association with Alu repeat sequences, which frequently bind CTCF, required SNF2h activity (44). Consistent with this are as yet unconfirmed findings that the SNF2-like helicase CHD8 is required for CTCF insulator activity (45). Another potential link to CTCF is the observation that the DNA methyltransferase inhibitor 5Aza-C facilitates cohesin association with Alu elements (44), because CTCF preferentially binds unmethylated DNA. Thus it may be currently theorized that NuRD, or at least NuRD components, destabilize cohesin at many non-CTCF sites, but facilitate cohesin association with CTCF sites.

Other mechanisms

Cohesin-binding genes in Drosophila are highly enriched for TG repeats in the non-coding transcribed region downstream of the promoter (34*). These repeats occur specifically in the plus strand, and thus the nascent transcripts of cohesin-binding genes will contain UG repeats, which specifically bind TDP-43 and CELF family proteins (46–48). Although the possible enrichment of UG repeats in mammalian cohesin-binding genes has not yet been

investigated, these RNA-binding proteins are highly conserved between Drosophila and mammals, and thus there is a possibility that they could influence cohesin binding in Drosophila and mammals.

Roles of cohesin in transcription

Cohesin and long-range interactions between regulatory elements

Drosophila *Nipped-B* mutations were originally isolated in a genetic screen for factors needed for activation of the *cut* and *Ultrabithorax* homeotic genes by transcriptional enhancers located 50 to 80 kb from the promoters (3). This gave rise to the idea that cohesin may control gene expression by facilitating or interfering with long-range interactions between gene regulatory elements in 3-dimensional nuclear space. However, it was not until the advent of high-resolution chromatin conformation capture (3C) studies in mammalian cells that cohesin could be shown to help form or stabilize long-range interactions between its binding sites (11*, 19*, 36*, 49–51). In mouse embryonic stem (ES) cells, cohesin and Mediator co-localize at the extragenic enhancers that drive expression of key pluripotency genes, and 3C experiments showed decreased enhancer-promoter contacts at these genes when cohesin was depleted using RNAi. Because ES cells are rapidly dividing, and cohesin is required for stem cell maintenance, the impact of cohesin knockdown in this system may be a composite of cohesin's cell cycle and gene regulatory functions. Genetic deletion of cohesin in non-dividing mouse thymocytes reduced interactions between the *T-cell receptor α* (*Tcra*) gene enhancer and promoter, thereby reducing *Tcra* transcription and rearrangement, and thus thymocyte differentiation (11*). Inducing mouse erythroleukemia cells to differentiate into mature erythroid cells activates the adult *β-globin* gene, and increases cohesin and Nipbl binding at the adult *β-globin* promoter, and the enhancer sequences in the *β-globin* locus control region (LCR), but not at the fetal genes that remain inactive (36*). Binding of cohesin and adult *β-globin* transcription was accompanied by increased interaction between enhancer elements and the promoter as measured by 3C and ChIP-loop experiments. These interactions are reduced in *Nipbl* (+/−) mutant cells (36*).

The mechanisms by which cohesin facilitates long-range interactions between its binding sites remains unknown. The obvious idea is that cohesin holds these elements together through topological entrapment in cis, similar to the way it holds sister chromatids together in trans.

The association of cohesin with a large fraction of cis-regulatory modules (CRMs) defined by binding of multiple tissue-specific transcription factors suggests that cohesin likely supports enhancer-promoter interactions at thousands of genes (39*). Similarly, in cultured Drosophila cells derived from central nervous system, kollerin and cohesin bind to more than 95% of CRMs predicted on the basis of DNaseI hypersensitive sites and the H3K27ac and H3K4me1 histone modifications (Schaaf, Dorsett et al., unpublished). Pol II is detected at nearly half of all predicted extragenic CRMs by ChIP, even though most of these are not transcribed as determined by global run-on sequencing (GRO-seq). This suggests that the Pol II detected at the extragenic CRMs may be transcriptionally engaged at a promoter, and interacts with the CRM through cohesin-facilitated looping. Consistent with this idea, the Pol II ChIP signal decreases at a quarter of the predicted CRMs upon cohesin depletion, which is higher than the frequency of Pol II decreases at active promoters.

Transcriptional elongation

In Drosophila and primary mouse liver cells, cohesin preferentially binds genes with promoter-proximal paused RNA polymerase, in which transcriptionally-engaged polymerase transcribes several dozen nucleotides but is blocked from entering into further elongation by the NELF (negative elongation factor) and DSIF (DRB-sensitivity inducing factor) pausing

complexes (34*, 39*). At least in Drosophila, cohesin is not required for pausing, and depletion of pausing factors does not substantially alter cohesin binding or vice versa. This suggests that although cohesin and pausing factor binding strongly overlap, they are recruited independently of each other (34*).

It has been proposed that promoter-associated cohesin may physically impede Pol II movement to facilitate pausing at TNFα-inducible genes in mammalian cells (52). However, cohesin depletion in Drosophila cells does not increase the rate at which a wave of nascent RNA synthesis moves along the induced *Ecdysone receptor* (*EcR*) gene, indicating that cohesin does not measurably hinder Pol II movement in this setting (34*).

An unusual transcriptional pause site is located at a cohesin-CTCF site in the first intron of a multi-cistronic latency transcription unit of Kaposi's sarcoma-associated herpesvirus (KHSV) (53). Pausing at this site is evident by an accumulation of Pol II, NELF and DSIF but is lost upon treatment with glycyrrhizic acid. This natural product from licorice inhibits viral growth and leads to a corresponding decrease in viral transcription (53). Mechanistically, glycyrrhizic acid interacts with the cohesin subunit Smc3, reduces its acetylation and association with Rad21. This causes global defects in sister chromatid cohesion, but surprisingly does not appear to cause a loss of cohesin binding at the pause site. Similar effects on pausing and transcription were also seen at a downstream cohesin-binding pause site in the endogenous *MYC* gene (P2) that lacks CTCF. These findings suggest that in contrast to promoter-proximal pausing, cohesin may directly control pausing at some intragenic sites.

Pol II pausing is a key point of transcriptional regulation, in which Pol II, NELF and DSIF must be phosphorylated by the P-TEFb complex, which can be recruited by activator proteins, to enter into productive elongation in the gene body (54). In Drosophila cells, cohesin can both inhibit and facilitate the transition of paused Pol II to elongation. At genes that are repressed by cohesin, pausing factor and cohesin co-depletion experiments indicate that cohesin inhibits transition of paused Pol II to elongation at a step different from those controlled by NELF and DSIF (34*). Genome-wide mapping of Pol II activity by ChIP and GRO-seq revealed that cohesin depletion both increases and decreases the pausing index, which is the ratio of the amount of transcriptionally-engaged Pol II at the promoter to the amount in the gene body, at hundreds of genes (Schaaf, Dorsett et al., unpublished). Increases in pausing occur much more frequently at cohesin-binding genes than at genes that lack cohesin, indicating that cohesin facilitates the transition of paused Pol II to elongation at many of the genes that it binds. It is currently unclear why cohesin has opposite effects on pausing at different genes, although many of the genes in which cohesin inhibits the transition to elongation are rare cases in which active genes also targeted by the PRC2 PcG complex, and thus have the H3K27me3 histone modification that is normally associated with gene silencing. Thus the effect of cohesin on pausing may be determined by the specific activator and repressor proteins that are also present at a gene.

Similar studies on the global effects of cohesin on Pol II pausing and elongation have yet to be conducted in mammalian cells. However, as described below, cohesin facilitates enhancer-promoter interactions in both Drosophila and vertebrates. Given that enhancers can increase gene transcription by facilitating the release of paused Pol II (55), it seems likely that the relationship between cohesin and RNA polymerase elongation is mediated in part by cohesin's role in supporting long-range enhancer-promoter interactions.

Transcriptional termination

A role for cohesin in the termination of transcription in the fission yeast *S. pombe* (56*). Here, bi-directional transcripts activate RNA interference mechanisms to establish

repressive chromatin marks. These marks are read by heterochromatin-binding proteins, which recruit cohesin, and cohesin in turn terminates transcription (56*). It will be interesting to determine whether similar mechanisms of transcriptional control by cohesin operate in other organisms.

Acknowledgments

Work in the Dorsett laboratory is supported by grants from the US National Institutes of Health (HD052860, GM055683). Work in the Merkenschlager laboratory is supported by the Medical Research Council, UK and the Wellcome Trust. The authors thank Katsu Shirahige for helpful discussion and sharing unpublished information.

References

1. Nasmyth K, Haering CH. Cohesin: its roles and mechanisms. Annu Rev Genet. 2009; 43:525–558. [PubMed: 19886810]
2. Dorsett D, Ström L. The ancient and evolving roles of cohesin in gene expression and DNA repair. Curr Biol. 2012; 22:R240–R250. [PubMed: 22497943]
3. 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]
4. Krantz ID, McCallum J, DeScipio C, Kaur M, Gillis LA, Yaeger D, Jukofsky L, Wasserman N, Bottani A, Morris CA, Nowaczyk MJ, Toriello H, Bamshad MJ, Carey JC, Rappaport E, Kawauchi S, Lander AD, Calof AL, Li HH, Devoto M, Jackson LG. Cornelia de Lange syndrome is caused by mutations in NIPBL, the human homolog of *Drosophila melanogaster* Nipped-B. Nat Genet. 2004; 36:631–635. [PubMed: 15146186]
5. Tonkin ET, Wang TJ, Lisgo S, Bamshad MJ, Strachan T. NIPBL, encoding a homolog of fungal Scc2-type sister chromatid cohesion proteins and fly Nipped-B, is mutated in Cornelia de Lange syndrome. Nat Genet. 2004; 36:636–641. [PubMed: 15146185]
6. Heidinger-Pauli JM, Mert O, Davenport C, Guacci V, Koshland D. Systematic reduction of cohesin differentially affects chromosome segregation, condensation, and DNA repair. Curr Biol. 2010; 20:957–963. [PubMed: 20451387]
7. 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]
8. Pauli A, Althoff F, Oliveira RA, Heidmann S, Schuldiner O, Lehner CF, Dickson BJ, Nasmyth K. Cell-type-specific TEV protease cleavage reveals cohesin functions in Drosophila neurons. Dev Cell. 2008; 14:239–251. [PubMed: 18267092]
9. 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]
10. 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]
11. Seitan VC, Hao B, Tachibana-Konwalski K, Lavagnolli T, Mira-Bontenbal H, Brown KE, Teng G, Carroll T, Terry A, Horan K, et al. A role for cohesin in T-cell-receptor rearrangement and thymocyte differentiation. Nature. 2011; 476:467–471. [PubMed: 21832993] . * This study provides a clear dissociation between cell cycle-related and interphase functions of cohesin in mammalian cells and shows how the differentiation of non-cycling mammalian cells can depend on cohesin-mediated gene regulation.
12. Glynn EF, Megee PC, Yu HG, Mistrot C, Unal E, Koshland DE, DeRisi JL, Gerton JL. Genome-wide mapping of the cohesin complex in the yeast *Saccharomyces cerevisiae*. PLoS Biol. 2004; 2:E259. [PubMed: 15309048]
13. Kogut I, Wang J, Guacci V, Mistry RK, Megee PC. The Scc2/Scc4 cohesin loader determines the distribution of cohesin on budding yeast chromosomes. Genes Dev. 2009; 23:2345–2357. [PubMed: 19797771]

14. Lengronne A, Katou Y, Mori S, Yokobayashi S, Kelly GP, Itoh T, Watanabe Y, Shirahige K, Uhlmann F. Cohesin relocation from sites of chromosomal loading to places of convergent transcription. Nature. 2004; 430:573–578. [PubMed: 15229615]
15. D'Ambrosio C, Schmidt CK, Katou Y, Kelly G, Itoh T, Shirahige K, Uhlmann F. Identification of cis-acting sites for condensin loading onto budding yeast chromosomes. Genes Dev. 2008; 22:2215–2227. [PubMed: 18708580]
16. Ocampo-Hafalla MT, Uhlmann F. Cohesin loading and sliding. J Cell Sci. 2011 Mar 1; 124(Pt 5): 685–691. 2011. [PubMed: 21321326]
17. 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]
18. Misulovin Z, Schwartz YB, Li XY, Kahn TG, Gause M, MacArthur S, Fay JC, Eisen MB, Pirrotta V, Biggin MD, Dorsett D. Association of cohesin and Nipped-B with transcriptionally active regions of the *Drosophila melanogaster* genome. Chromosoma. 2008 Feb; 117(1):89–102. 2008. [PubMed: 17965872] . * This genome-wide study maps Drosophila cohesin complexes and cohesin loading factors to active genes.
19. 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. 2010; 467:430–435. [PubMed: 20720539] . * This study maps cohesin loading factors to sites of active transcription in mouse ES cells.
20. Schmidt D, Schwalie PC, Ross-Innes CS, Hurtado A, Brown GD, Carroll JS, Flicek P, Odom DT. A CTCF-independent role for cohesin in tissue-specific transcription. Genome Res. 2010; 20:578–588. [PubMed: 20219941] . * This study identifies cell type-specific cohesin binding to sites that lack detectable CTCF protein and CTCF consensus sequences. Instead, cohesin co-localised with estrogen receptor in MCF7 breast cancer cells and with liver-specific transcription factors in HepG2 hepatocellular carcinoma cells, suggesting a link with cell type-specific gene expression.
21. 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]
22. 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]
23. Rubio ED, Reiss DJ, Welcsh PL, Disteche CM, Filippova GN, Baliga NS, Aebersold R, Ranish JA, Krumm A. CTCF physically links cohesin to chromatin. Proc Natl Acad Sci U S A. 2008; 105:8309–8314. [PubMed: 18550811]
24. Stedman W, Kang H, Lin S, Kissil JL, Bartolomei MS, Lieberman PM. Cohesins localize with CTC Fat the KSHV latency control region and at cellular c-myc and H19/Igf2 insulators. EMBO J. 2008; 27:654–666. [PubMed: 18219272]
25. Remeseiro S, Cuadrado A, Gómez-López G, Pisano DG, Losada A. A unique role of cohesin-SA1 in gene regulation and development. EMBO J. 2012; 31:2090–2102. [PubMed: 22415368]
26. Cuddapah S, Jothi R, Schones DE, Roh TY, Cui K, Zhao K. Global analysis of the insulator binding protein CTCF in chromatin barrier regions reveals demarcation of active and repressive domains. Genome Res. 2009; 19:24–32. [PubMed: 19056695]
27. van Steensel B. Chromatin: constructing the big picture. EMBO J. 2011; 30:1885–1895. [PubMed: 21527910]
28. Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, Hu M, Liu JS, Ren B. Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature. 2012; 485:376–380. [PubMed: 22495300]
29. Sanyal A, Lajoie BR, Jain G, Dekker J. The long-range interaction landscape of gene promoters. Nature. 2012; 489:109–113. [PubMed: 22955621]
30. Wallace JA, Felsenfeld G. We gather together: insulators and genome organization. Curr Opin Genet Dev. 2007; 17:400–407. [PubMed: 17913488]
31. Xiao T, Wallace J, Felsenfeld G. Specific sites in the C terminus of CTCF interact with the SA2 subunit of the cohesin complex and are required for cohesin-dependent insulation activity. Mol

Cell Biol. 2011; 31:2174–2183. [PubMed: 21444719] . * This paper details the sites of interaction between cohesin and CTCF.

32. Hu B, Itoh T, Mishra A, Katoh Y, Chan KL, Upcher W, Godlee C, Roig MB, Shirahige K, Nasmyth K. ATP hydrolysis is required for relocating cohesin from sites occupied by its Scc2/4 loading complex. Curr Biol. 2011; 21:12–24. [PubMed: 21185190] . * This paper suggests a positive association between RNA polymerases and cohesin loading factors in budding yeast.
33. 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]
34. Fay A, Misulovin Z, Li J, Schaaf CA, Gause M, Gilmour DS, Dorsett D. Cohesin selectively binds and regulates genes with paused RNA polymerase. Curr Biol. 2011; 21:1624–1634. [PubMed: 21962715] . * This paper distinguishes for the first time that one of the key characteristics of the subset of active genes that bind cohesin and kollerin is that they have high levels of promoter-proximal paused RNA polymerase, and demonstrates that cohesin can control the transition of paused polymerase into active elongation, establishing a new principle by which cohesin controls transcription of active genes.
35. Liu J, Zhang Z, Bando M, Itoh T, Deardorff MA, Clark D, Kaur M, Tandy S, Kondoh T, Rappaport E, Spinner NB, Vega H, Jackson LG, Shirahige K, Krantz ID. Transcriptional dysregulation in *NIPBL* and cohesin mutant human cells. PLoS Biol. 2009; 7:e1000119. [PubMed: 19468298]
36. Chien R, Zeng W, Kawauchi S, Bender MA, Santos R, Gregson HC, Schmiesing JA, Newkirk DA, Kong X, Ball AR Jr, Calof AL, Lander AD, Groudine MT, Yokomori K. Cohesin mediates chromatin interactions that regulate mammalian β-globin expression. J Biol Chem. 2011; 286:17870–17878. [PubMed: 21454523] . * This paper shows (a) a gain of binding of cohesin and Nipbl to both a remote enhancer and its cognate promoter upon gene activation, (b) a role for cohesin in enhancer-promoter communication, and (c) a decrease in cohesin binding to an active enhancer and its cognate promoter in vivo in cells with only 30% decrease in Nipbl. Together these findings help establish and confirm cohesin function at active genes.
37. Ebmeier CC, Taatjes DJ. Activator-Mediator binding regulates Mediator-cofactor interactions. Proc. Natl Acad. Sci. USA. 2010; 107:11283–11288. [PubMed: 20534441]
38. Conaway RC, Conaway JW. The Mediator complex and transcription elongation. Biochim Biophys Acta. 2012 S1874-9399(12)00161-7.
39. Faure AJ, Schmidt D, Watt S, Schwalie PC, Wilson MD, Xu H, Ramsay RG, Odom DT, Flicek P. Cohesin regulates tissue-specific expression by stabilising highly occupied cis-regulatory modules. Genome Res. 2012; 22:2163–2175. [PubMed: 22780989] . * This study suggests that cohesin can stabilise the binding of transcription factors to gene regulatory elements.
40. Nitzsche A, Paszkowski-Rogacz M, Matarese F, Janssen-Megens EM, Hubner NC, Schulz H, de Vries I, Ding L, Huebner N, Mann M, Stunnenberg HG, Buchholz F. RAD21 cooperates with pluripotency transcription factors in the maintenance of embryonic stem cell identity. PLoS One. 2011; 6:e19470. [PubMed: 21589869]
41. Matharu NK, Hussain T, Sankaranarayanan R, Mishra RK. Vertebrate homologue of Drosophila GAGA factor. J Mol Biol. 2010; 400:434–447. [PubMed: 20471984]
42. Fasulo B, Deuring R, Murawska M, Gause M, Dorighi KM, Schaaf CA, Dorsett D, Brehm A, Tamkun JW. The Drosophila MI-2 chromatin-remodeling factor regulates higher-order chromatin structure and cohesin dynamics in vivo. PLoS Genet. 2012; 8:e1002878. [PubMed: 22912596]
43. Hu G, Wade PA. NuRD and pluripotency: a complex balancing act. Cell Stem Cell. 2012; 10:497–503. [PubMed: 22560073]
44. Hakimi MA, Bochar DA, Schmiesing JA, Dong Y, Barak OG, Speicher DW, Yokomori K, Shiekhattar R. A chromatin remodelling complex that loads cohesin onto human chromosomes. Nature. 2002; 418:994–998. [PubMed: 12198550]
45. Ishihara K, Oshimura M, Nakao M. CTCF-dependent chromatin insulator is linked to epigenetic remodeling. Mol Cell. 2006; 23:733–742. [PubMed: 16949368]
46. Buratti E, Baralle FE. TDP-43: gumming up neurons through protein-protein and protein-RNA interactions. Trends Biochem Sci. 2012; 37:237–247. [PubMed: 22534659]

47. Dasgupta T, Ladd AN. The importance of CELF control: molecular and biological roles of the CUG-BP, Elav-like family of RNA-binding proteins. Wiley Interdiscip Rev RNA. 2012; 3:104–121. [PubMed: 22180311]
48. Vlasova-St Louis I, Bohjanen PR. Coordinate regulation of mRNA decay networks by GU-rich elements and CELF1. Curr Opin Genet Dev. 2011; 21:444–451. [PubMed: 21497082]
49. Hadjur S, Williams LM, Ryan NK, Cobb BS, Sexton T, Fraser P, Fisher AG, Merkenschlager M. Cohesins form chromosomal cis interactions at the developmentally regulated IFNG locus. Nature. 2009; 460:410–413. [PubMed: 19458616]
50. 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]
51. Degner SC, Verma-Gaur J, Wong TP, Bossen C, Iverson GM, Torkamani A, Vettermann C, Lin YC, Ju Z, Schulz D, et al. CCCTC-binding factor (CTCF) and cohesin influence the genomic architecture of the Igh locus and antisense transcription in pro-B cells. Proc Natl Acad Sci U S A. 2011; 108:9566–9571. [PubMed: 21606361]
52. Wada Y, Ohta Y, Xu M, Tsutsumi S, Minami T, Inoue K, Komura D, Kitakami J, Oshida N, Papantonis A, Izumi A, Kobayashi M, Meguro H, Kanki Y, Mimura I, Yamamoto K, Mataki C, Hamakubo T, Shirahige K, Aburatani H, Kimura H, Kodama T, Cook PR, Ihara S. A wave of nascent transcription on activated human genes. Proc Natl Acad Sci U S A. 2009; 106:18357–18361. [PubMed: 19826084]
53. Kang H, Lieberman PM. Mechanism of glycyrrhizic acid inhibition of Kaposi's sarcoma-associated herpesvirus: disruption of CTCF-cohesin-mediated RNA polymerase II pausing and sister chromatid cohesion. J Virol. 2011; 85:11159–11169. [PubMed: 21880767]
54. Peterlin BM, Price DH. Controlling the elongation phase of transcription with P-TEFb. Mol Cell. 2006; 23:297–305. [PubMed: 16885020]
55. Maston GA, Landt SG, Snyder M, Green MR. Characterization of enhancer function from genome-wide analyses. Annu Rev Genomics Hum Genet. 2012; 13:29–57. [PubMed: 22703170]
56. Gullerova M, Proudfoot NJ. Cohesin complex promotes transcriptional termination between convergent genes in S. pombe. Cell. 2008; 132:983–995. [PubMed: 18358811] . * This paper demonstrates an involvement of the RNAi machinery in recruiting cohesin to sites of bi-directional transcription in fission yeast and suggests a role for cohesin in defining transcription termination sites.
57. Nasmyth K. Cohesin: a catenase with separate entry and exit gates? Nat Cell Biol. 2011; 3:1170–1177. [PubMed: 21968990]
58. Nishiyama T, Ladurner R, Schmitz J, Kreidl E, Schleiffer A, Bhaskara V, Bando M, Shirahige K, Hyman AA, Mechtler K, Peters JM. Sororin mediates sister chromatid cohesion by antagonizing Wapl. Cell. 2010; 143:737–749. [PubMed: 21111234]

Perspective

The suggestion of cohesin loading at active genes provides a unifying theme for the relationship between cohesin and gene expression from model organisms to humans. We do not yet fully understand how cohesin loading and removal at specific sites is controlled, or to what extent the mechanisms are conserved between organisms. As summarized above, there are several tantalizing leads and it seems likely that at least some of the mechanisms will be revealed in the near future.

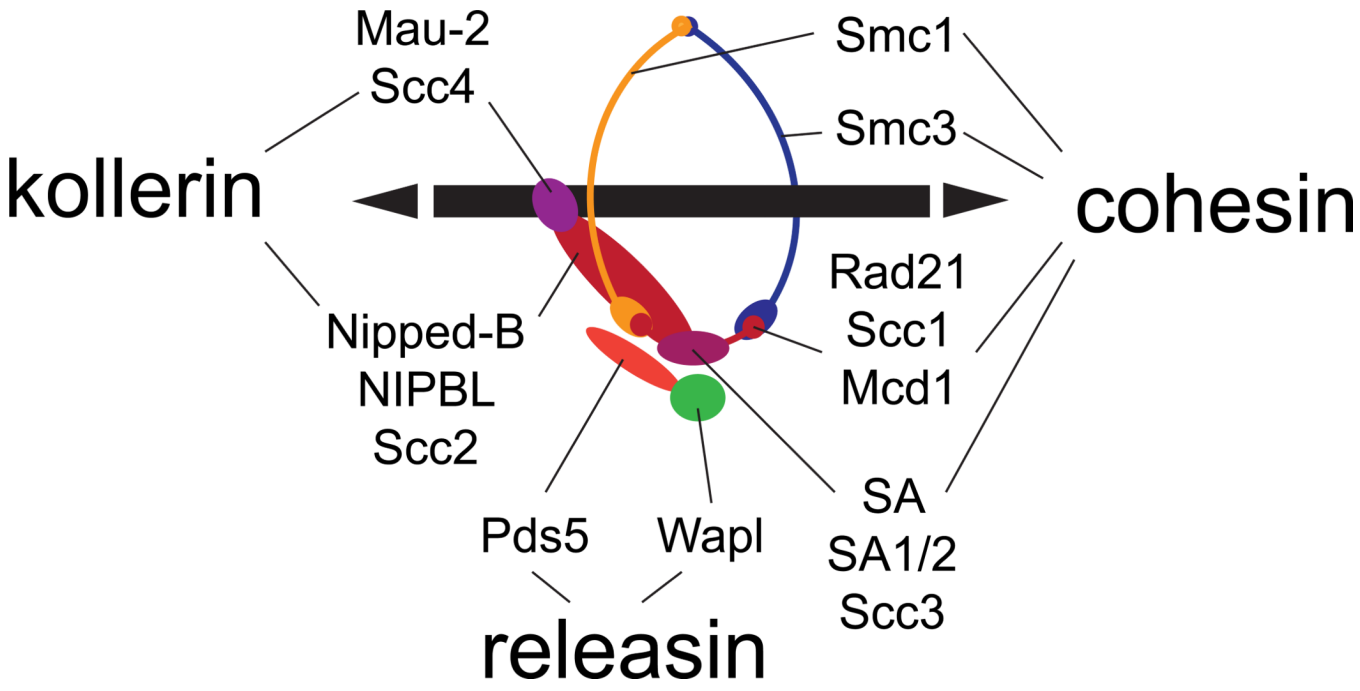

Figure 1.
Cohesin and the kollerin and releasin complexes that control cohesin chromosome binding dynamics during interphase. The various organism-specific names for each protein subunit of the complexes in Drosophila, mammals and yeast are given. Cohesin topologically encircles DNA, and during interphase, is loaded onto chromosomes by kollerin, which requires ATP hydrolysis by the cohesin Smc subunits, and is removed by the releasin complex (57). Releasin interacts with the cohesin SA subunit, and Sororin (not shown; Dalmatian in Drosophila) interacts with Pds5 to counteract Wapl and the cohesin removal activity (58).

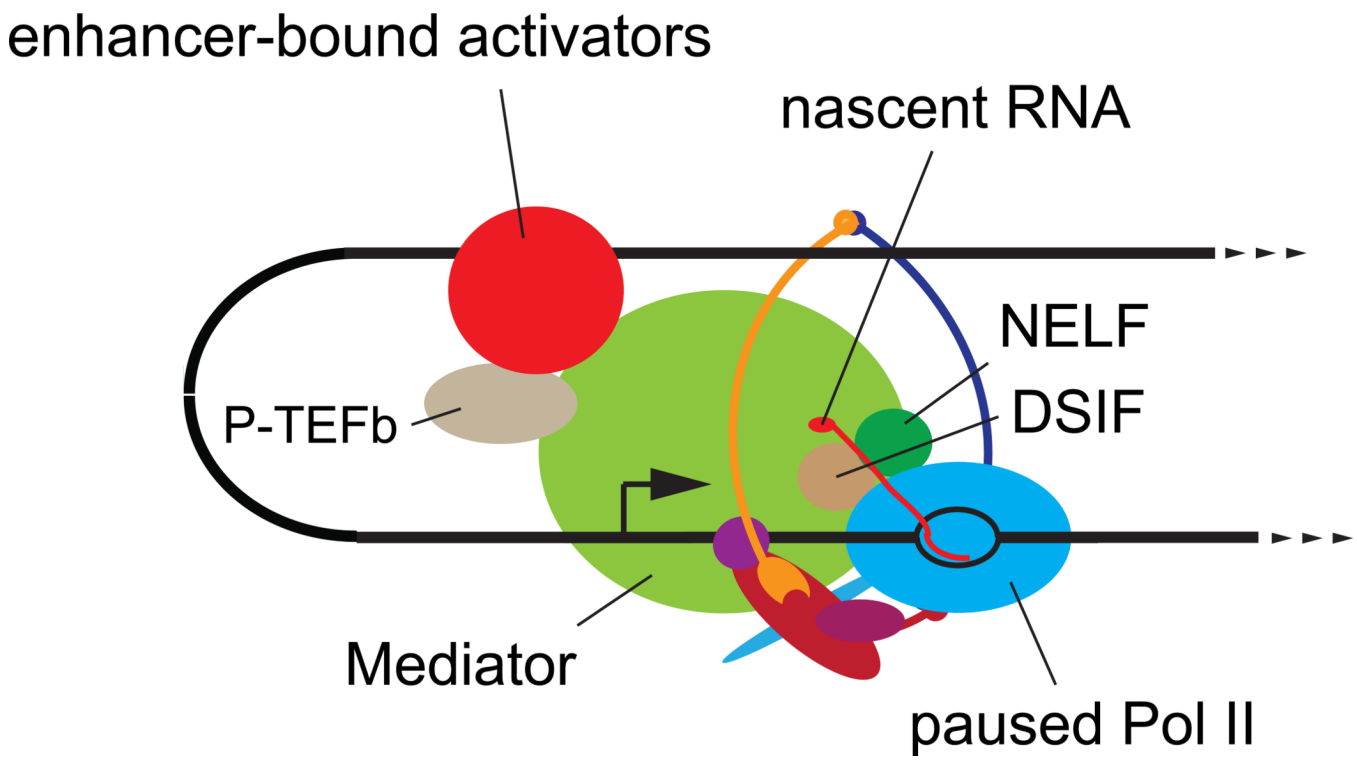

Figure 2.
Cohesin and it's proposed functions at active gene promoters in metazoan organisms. As described in the text, outside of CTCF sites in mammalian cells, kollerin and cohesin levels are highest at transcriptional enhancers and active gene promoters in which RNA polymerase (Pol II) pauses after transcribing several nucleotides. Pausing requires the NELF (negative elongation factor) and DSIF (DRB-sensitivity inducing factor), but they are not required for cohesin binding or *vice versa*. The Mediator complex interacts with cohesin and is present at virtually all cohesin-bound enhancers and promoters. Current evidence indicates that cohesin stabilizes DNA loops that bring the enhancer and promoter into contact, and thereby stimulate phosphorylation of Pol II, NELF, and DSIF by the P-TEFb complex (CycT, Cdk9) recruited by transcriptional activators and release the paused Pol II into active elongation.