Condensin, cohesin and the control of chromatin states

Luis Aragon\textsuperscript{1}, Enrique Martinez-Perez \textsuperscript{2}, and Matthias Merkenschlager\textsuperscript{3}

\textsuperscript{1}Cell Cycle Group,  
\textsuperscript{2}Meiosis Group,  
\textsuperscript{3}Lymphocyte Development Group,  
MRC Clinical Sciences Centre, Imperial College London,  
Du Cane Road, London W12 0NN, UK

\textsuperscript{1}luis.aragon@imperial.ac.uk  
\textsuperscript{2}enrique.martinez-perez@imperial.ac.uk  
\textsuperscript{3}matthias.merkenschlager@csc.mrc.ac.uk

COGEDE provisional deadline: 17th Sept 2012 01 Oct 2012

Overview

Cohesin and condensin complexes are essential for defining the topology of chromosomes through the cell cycle. Here we look at the emerging role of these complexes in regulating chromatin structure and gene expression and reflect on how these activities could be linked with chromosome topology.

SMC protein complexes

Structural Maintenance of Chromosomes (SMC) proteins are vital for chromosome structure and dynamics, gene regulation and DNA repair. SMC proteins consist of N- and C-terminal domains that fold back onto each other to create an ATPase ‘head’ domain, connected to a central ‘hinge’ domain via long coiled-coils. Eukaryotes have six different SMC proteins that selectively heterodimerise to form three SMC complexes, cohesin (Smc1-Smc3), condensin (Smc2-Smc4), and Smc5-Smc6 [1].

In addition to the Smc1-Smc3 dimer, the cohesin complex contains Scc1 and Scc3 subunits [2] (Figure 1). Two different Scc3 proteins are found in vertebrate cells, SA1 and SA2. The N- and C-terminal regions of Scc1 link the head domains of Smc3 and Smc1, forming a tripartite ring. Scc3 interacts with Scc1 to strengthen the ring structure of the complex. Cohesin is thought to topologically embrace DNA [3] and has the important role to maintain sister chromatids paired (cohesed) from the time they are replicated until they segregate at anaphase [3].

Condensin complexes contain three subunits in addition to the Smc2-Smc4 dimer. Vertebrates have two distinct condensin complexes, condensin I and condensin II [4], which differ in their non-Smc subunits. Condensin I contains CAP-D2, CAP-H, and CAP-G while condensin II contains CAP-D3, CAP-H2, and CAP-G2 (Figure 1). Condensins are best known for regulating chromosome shape and condensation during mitosis [4].
*C. elegans* has a specialized condensin complex that mediates dosage compensation and differs from condensin I by containing DPY-27 instead on Smc4 [5]. (Figure 1).

**Molecular basis for cohesin and condensin functions on DNA**

Despite their similar structure, Cohesin and Condensin have different effects on DNA.

Like Cohesin, replicative helicases and the DNA replication processivity factor PCNA form rings with a central opening large enough to entrap dsDNA [6]. However, in contrast to these proteins, cohesin’s exceptional size is predicted to topologically encircle not only one DNA helix, but two. The ATPase activity of cohesin’s SMC subunits is thought to enable the stable association of the complex with chromosomes [7], which is a prerequisite for the establishment of cohesion between sister chromatids. In this way, cohesion acts as a structural framework, or glue.

While cohesin may not actively change the shape of chromosomes, it should be noted that holding chromosomes in place could have dramatic effects when combined with factors that act on the topology of DNA or the structure of chromatin. Examples include RNA polymerases that mediate transcription, DNA polymerases that mediate replication, chromatin remodelers that move nucleosomes, topoisomerases that change the topology of DNA, and protein complexes that function as readers and writers of post-translational histone modifications. The binding of cohesin to specific sites may have an important role in containing the resulting changes in topology and chromatin structure within specific chromosomal domains.

In contrast to cohesin, condensin actively contributes to the progressive winding and folding of chromatin fibres that occurs in preparation for mitosis. Condensin mediates positive supercoiling of DNA in the presence of type I topoisomerases [8,9], an activity that requires condensin holocomplex and involves ATP hydrolysis by the SMC subunits [8,9]. In addition, the Condensin core subunits Smc2 and Smc4 are able to anneal complementary ssDNAs into dsDNA [10]. These activities are thought to enable condensin to organise interphase chromosomes into multiple supercoils that form ordered solenoids [9] and thereby shape chromatin fibers into structures that resemble mitotic chromosomes.

A recent study demonstrated that condensin is required for introducing positive supercoils into catenated sister DNA, and that this activity facilitates the decatenation of sister chromatids by topoisomerase II [11]. Furthermore, Topoisomerase II, and possibly other topoisomerases, were shown to counteract the action of condensin. These observations suggest that the shape of mitotic chromosomes could be determined, at least in part, by the balance between the condensin-dependent supercoiling (overwinding) and topoisomerase-mediated relaxation [12]. In this model, chromosome compaction would be achieved by increasing the supercoiling activity of condensin in the presence of constant relaxation activity by topoisomerases (Figure 2a). Interestingly, replicated Xenopus chromosomes exhibit different shapes depending on the ratio of condensin I versus condensin II complexes [13]. Long and slim chromosomes are formed when Condensin I quantities exceed those of Condensin II, while shorter and thicker chromosomes are formed when equal amounts of the two complexes are present. Hence, the relative presence and activities of factors involved in the formation of mitotic chromosomes can alter the compaction and appearance of chromosomes (Figure 2b).
The same study found that cohesin complexes are also important determinants of chromosome shape [13], since they affect the loading of condensin II and promote the juxtaposition of sister chromatids in metaphase chromosomes.

In addition to the essential cell cycle-related roles of cohesin and condensin described above, there is growing evidence that cohesin and condensin contribute to the regulation of chromatin structure and gene expression in interphase. Exactly how cohesin and condensin mediate these additional functions is unknown, but it appears likely that the very properties and enzymatic activities discussed above are involved, perhaps subject to context-dependent regulation (Figure 2c).

**Cohesin's role in regulating chromatin structure and gene expression in interphase.**

Early evidence that cohesin may regulate gene expression came from genetic studies in Drosophila, where the developmentally regulated expression of specific homeobox genes was dependent of the dosage of the cohesin loading factor Nipped-B [14]. Mutations in the human homolog of Nipped-B, NIPBL, were later found to cause the developmental disorder Cornelia de Lange Syndrome [15]. Heterozygous mutations in cohesin subunits [15], cohesin co-factors [16], and cohesin-modifying enzymes [17,18] also result in developmental abnormalities in humans and in model organisms. Interestingly, cultured cells derived from NIPBL heterozygous individuals did not show obvious defects in chromosome segregation, leading to suggestions that developmental defects in Cornelia de Lange Syndrome might instead be due to cohesin functions in gene regulation [15]. To support this idea, a clear distinction between cohesin's cell division-related and cell division-independent functions was required. This distinction was first achieved by depleting cohesin from post-mitotic neurons in Drosophila [19,20] and subsequently in non-cycling mouse thymocytes [21]. Cohesin-deficient Drosophila neurons showed defective function (axon pruning) as a result of de-regulated expression of the ecdyson receptor [19,20]. Genetic deletion of the cohesin subunit Rad21 in mouse thymocytes led to a defective chromatin architecture at the T cell receptor alpha locus, where cohesin binding sites flank key promoter and enhancer elements. Cohesin was required for long-range promoter-enhancer interactions, as well as for the developmentally regulated transcription and rearrangement of the T cell receptor alpha locus, which is required for thymocyte differentiation [21].

How does cohesin impact on interphase chromatin structure?

Cohesin determines the topology of chromosomes not just by holding sister chromatids together in trans, but also by forming long-range interactions between its binding sites [21-25]. In contrast to sister chromatid cohesion, these interactions appear to occur in cis, i.e. within individual chromosomes. This is suggested by studies in G1 cells [22] and in non-cycling cells [21], which do not have sister chromatids. Cohesin binding sites on mammalian chromosome arms are defined either by the sequence-specific DNA binding protein CTCF [26,27] or by the cohesin loading factor Nipbl, components of the Mediator complex, and tissue-specific transcription factors, which are found at active gene regulatory elements [24,28]. There are examples of long-range interactions between both CTCF- and Mediator-associated cohesin sites, but the global impact of cohesin on the three-dimensional organisation of interphase chromatin is yet to be established. In addition, it remains unknown how the topology adopted by cohesin to form long-range interactions in cis compares to the topological embrace thought to mediate sister
chromatid cohesion in trans [29].

How does cohesin impact on gene expression?

According to current models, cohesin affects gene expression by forming long-range interactions. It is easy to imagine how cohesin's impact on gene regulation would depend on the types of gene regulatory elements it connects. Interactions between enhancers and promoters may drive gene expression [21,24], while CTCF-based interactions may block enhancer-promoter interactions or demarcate chromatin domains [30]. Published examples show correlations between gene expression, long-range interactions and cohesin binding, but the causal relationships remain to be worked out. In addition, cohesin may facilitate the binding of transcription factors to sub-optimal sequence motifs [31] and impact on RNA polymerase elongation [31].

However, it is important to remember that just because cohesin can regulate the expression of certain genes independently of its role in the cell cycle [19-21], not all effects of cohesin depletion on gene expression are necessarily direct. In particular, the loss of cohesin from cycling cells can trigger damage responses that will radically alter the pattern of gene expression. In rapidly dividing cells such as embryonic stem (ES) cells, for example, cohesin depletion will result in the activation of p53 in response to DNA damage. This in turn antagonises the expression of pluripotency factors [32] (Figure 3).

Condensin's role in regulating chromatin structure and gene expression in interphase

The classical example for the control of gene expression by condensin is X chromosome dosage compensation in C. elegans. Here, a specialised condensin complex (condensin I$^{DC}$) associates with the X chromosomes of XX hermaphrodites to induce a chromosome-wide reduction of gene expression that equalizes levels of X-derived transcripts to those of males carrying a single X chromosome [33]. Condensin I$^{DC}$ appears to control gene expression by regulating transcription, since mutants that fail to recruit condensin I$^{DC}$ display a dramatic increase in the binding of RNA Pol II to promoters and coding regions of genes on the X chromosome [34]. Interestingly, binding of condensin I$^{DC}$ to the promoter or gene body does not predict the dosage compensation status of a gene, and direct binding of condensin I$^{DC}$ is neither required nor sufficient to elicit repression [35]. Consistent with models where condensin I$^{DC}$ regulates gene expression by controlling higher order chromosome structure, worms with mutations in the closely related condensin I complex (Figure 1) display elongated chromosomal axes during meiotic prophase and this results in higher meiotic recombination frequencies [36]. Similar to dosage compensation, meiotic recombination is regulated in a chromosome-wide fashion. Thus, both condensin I and condensin I$^{DC}$ may act through higher order chromosome structure.

Two recent studies point to an explanation for the impact of condensin I$^{DC}$ on interphase chromosome structure [37,38]. The histone modification H4K20me1 is enriched on dosage-compensated X chromosomes. H4K20me1 enrichment depends on condensin I$^{DC}$, suggesting that H4K20me1 has a role in dosage compensation. Indeed, depletion of SET-1 (PR-Set7/SET8), the acetyl transferase required for H4K20me1, elicits defects in dosage compensation without affecting recruitment of condensin I$^{DC}$ [37]. However, condensin I$^{DC}$ does not appear to promote SET-1 activity, but rather to inhibit the activity...
of SET-4 (Suv4-20), which catalyzes the di- and tri-methylation of H4K20me1, since condensin lDc mutants have elevated levels of H4K20me3 on the X chromosomes. H4K20me1 is also enriched on inactive mammalian X chromosomes [39], where it contributes to chromosome condensation both in mitosis and interphase [40]. Thus, condensin lDc may control gene expression by creating a compacted chromatin structure that reduces access by RNA polymerase, through increased H4K20me1 levels on the X chromosomes.

In contrast to C. elegans, dosage compensation in Drosophila is initiated by the MSL (male-specific lethal) complex. MSL contains the histone acetyl transferase MOF, which acetylates histone H4 at lysine 16 (H4K16ac) and promotes the recruitment of RNA polymerase II [41] and transcription from the single X chromosome in males. Interestingly, association of MSL also changes the shape of the male X chromosome [42]. Hence, dosage compensation in flies is achieved by mechanisms that affect histone acetylation primarily and has a secondary impact on chromosome shape, while in worms shape appears to influence function. In mammals, dosage compensation is achieved by the transcriptional inactivation of one X chromosome in female XX cells [43]. The primary event is the association of the long non-coding RNA XIST with one X chromosome. This affects chromatin structure, represses transcription, and ultimately re-shapes the inactive X chromosome, which becomes highly condensed and visible as the Barr body in XX human cells. These comparisons suggests that distinct primary events can trigger a chain of events that changes chromosome structure, chromosome shape and transcriptional activity, which appear to be intricately linked.

The Drosophila condensin II protein CAP-D3 not only contributes to uniform chromatin condensation during mitotic prophase, but also co-localises with the retinoblastoma protein RBF1 on non-dividing polytene chromatin to affect the expression of numerous developmentally regulated genes, many of which are located in gene clusters [44]. In fat body cells, CAP-D3 contributes to the transcription of antimicrobial peptide genes, which are important for innate immunity and confer the ability to clear bacterial infections [44].

A connection between condensin II and the immune system also exists in mammals, where ENU mutagenesis screens for genes that affect mouse T cell development uncovered Ncaph2, which encodes the kleisin beta subunit of condensin II [45]. A follow-up study found that Ncaph2-deficient thymocytes do not undergo the same condensation of nuclear chromatin as normal T cells [46]. Ncaph2-deficient T cells failed to reach the quiescent state characteristic of resting T lymphocytes as judged by chromatin structure and accessibility to STAT transcription factors [46]. When interpreting these results it may be relevant that the numbers of lymphocytes are reduced in Ncaph2-deficient mice. This so-called lymphopenia means that the few lymphocytes that remain are subject to cytokine signals that cause them to proliferate [47]. Because quiescent and proliferating lymphocytes differ dramatically in their chromatin structure, it remains uncertain whether the impact of condensin II on interphase chromatin structure is direct or indirect.

Similar to T cell quiescence, the differentiation of mammalian eythroid cells is associated with gradual chromatin condensation. The mouse homolog of the condensin II subunit CAP-G2 called MTB, for More Than Blood, participates in erythropoiesis. Overexpression of MTB promotes the terminal differentiation of mouse erythroleukemia
cells. It is not known whether this is due to a direct effect of CAP-G2 on chromatin condensation or to interactions between CAP-G2 and the basic helix-loop-helix stem cell leukemia protein SCL, which is important for the differentiation of erythroid cells [48].

Condensin and pluripotency.

An RNA interference screen interrogating regulators of pluripotent mouse embryonic stem (ES) cell chromatin structure identified the condensin subunits Smc2 and Smc4 among a set of genes required for the viability of pluripotent ES cells [49], but not of immortalised fibroblasts, which were chosen as an example of differentiated cells [50]. In ES cells, knockdown of Smc2 and -4 in ES cells not only delayed the formation of mitotic chromosomes, but also resulted in interphase nuclei that were enlargement due to defective chromatin compaction. Fluorescence in situ hybridisation (FISH) studies to assess the organization of individual loci showed altered higher-order chromatin folding. Altered chromatin compaction coincided with alterations in post-translational histone modifications and in decreased DNA methylation [50]. Mechanistically, knockdown of Smc2 and -4 in ES cells resulted in the activation of p53 in ES cells but not in differentiated cells, but p53 was required only for the induction of apoptosis, not for the disruption of chromatin structure. These data suggest that condensin contributes to maintaining the interphase chromatin structure of ES cells.

**Conclusions and future directions**

Cohesin and condensin complexes play essential roles in defining the topology of chromosomes in cycling as well as in resting cells. It will be exciting to investigate the mechanisms by which ATP binding and hydrolysis by the core SMC subunits of both complexes is coupled to the topology of DNA in the context of chromatin, and how this in turn affects the regulation of gene expression.
References


   • Together with Pauli 2008, 2010 this study demonstrates a role for cohesin in the function and gene expression of post-mitotic neurons in Drosophila.

   • Together with Schuldiner 2008 and Pauli 2010 this study demonstrates a role for cohesin in the function and gene expression of post-mitotic neurons in Drosophila.

   • 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.


This study maps cohesin loading factors to sites of active transcription in mouse ES cells.


- This study and [38]• show that X chromosomes display increased H4K20me1 and that this enrichment is dependent on the dosage compensation complex, stabilising a link between dosage compensation and chromatin modification.


- See [37]•


- This study links condensin 2 with the structure and the accessibility of interphase chromatin in T lymphocytes.


- This study links condensin 2 with the structure of interphase chromatin in embryonic stem cells.

Added in proof:

Figure legends

Figure 1. Eukaryotic SMC complexes.

Each Smc complex is composed of a specific Smc dimer and several non-Smc subunits. (a) cohesin complex, (b) condensin I complex, condensin II complex and condensin-like dosage compensation complex IDC in C. elegans, which differs from condensin I by a single subunit. (c) The Smc5/6 complex.

Figure 2. Chromosome shape/compaction depends on the relative activities of structural factors.

(a) The compaction of chromatin/chromosome fibers could be achieved by balancing of two activities; Condensin-dependent supercoiling (which would drive compaction) and topoisomerase-mediated relaxation (which would drive decompaction). (b) Mitotic chromosome shapes might be achieved by different contributions of condensin I and II complexes and topoisomerases. Cohesin also contributes to chromosome shape by providing cohesive forces between the chromatids. Condensin II has been shown to be important for axial compaction of mitotic chromosomes [13]. On the other hand, the activity of condensin I complex (which is has a low residency time on chromatin) is likely to be counteracted by topoisomerases (as described in a) contributes to the lateral compaction of chromatids. (c) Recently Condensin has been shown to contribute to the regulation of chromatin structure and gene expression in interphase. It is likely that these effects are achieved by context-dependent regulation of Condensin and
topoisomerase activities in the interphase nucleus. In this scenario, targeting of Condensin to certain chromosome domains might contribute to the compaction and transcriptional repression of specific regions.

Figure 3. Direct and indirect effects of cohesin on gene expression.

Cohesin depletion at various stages of the cell cycle can affect gene expression in different ways (see text for details).
Figure 1
Eukaryotic SMC complexes

a
Cohesin

b
Condensin

i
ii

SMC2 SMC4
CAP-G2 CAP-D3
CAP-G CAP-D2

i'
ii'

SMC2 SMC4
CAP-H CAP-D2
CAP-H CAP-G

SMC5/6

SMC6 SMC5
NSE3 NSE4
NSE3 NSE6

NSE6 NSE2

NSE3 NSE4

SMC5 SMC6

NSE2 NSE6

NSE3 NSE4
Figure 2
Chromosome shapes/compaction depends on relative activities of structural factors

a) Topoisomerases
  supercoil relaxation

b) Chromatin Compaction Gradient
  Condensins
  supercoil introduction

b) High axial compaction
  Condensin II (++)
  Cohesin
  Topoisomerases (+++)
  Condensin I (+)

b) Low axial compaction
  Condensin II (+)
  Cohesin
  Topoisomerases (+)
  Condensin I (+++)

c) Condensin (+)
  Heterochromatin/transcriptionally repressed
  Centromere
  Condensin (-)
  Decondensed/transcriptionally active
Figure 3
Direct and indirect effects of cohesin on gene expression