Are SMC Complexes Loop Extruding Factors? Linking Theory With Fact

Jonathan Baxter,* Antony W. Oliver, and Stephanie A. Schalbetter

1. Introduction

In order to encode the genetic information required for cellular function, DNA polymers are extremely long. Every time a cell divides, one copy of the entire DNA molecule making up a chromosome needs to be faithfully segregated into each daughter cell. This process requires the intra-chromosomal compaction of each of the segregating chromosomes in a manner that ensures that different chromosomes are separated from each other, that sister-chromatids are fully resolved, and that the chromosome’s length during anaphase is shortened sufficiently to be fully segregated into the daughter cells. Although intra-chromosomal compaction is most visible during mitosis, it also actively occurs during interphase. Analysis of contact frequencies across interphase chromosomes has shown that they are partially compacted into distinct topologically associated domains (TADs).[1–3] The different types of compaction across chromosomes in mitosis and within TADs in interphase are linked by a common organizational principle – intra-chromosomal looping of DNA that appears to act sequentially along the length of the chromosome. Classical electron microscopy experiments and recent contact analysis have both shown that metaphase chromosomes are composed of an array of DNA loops that are linked to a chromosome axis.[4,5] Furthermore, during interphase, looping between the insulators that define TAD boundaries facilitates the correct, long-range cis-looping of DNA between enhancers and gene promoters.[6] Theories for how mitotic and interphase chromosomal interactions could be restricted to act in cis have led to a revival of the concept of DNA loop extruding factors acting along chromosomes.

2. The Loop Extrusion Model

Loop extruding factors (LEFs) are proposed to interact with DNA, pulling the DNA fiber into a loop, through specific contacts with its protein scaffold.[7,8] (Figure 1). Molecular dynamic simulations have been used to generate in silico models of how LEFs could generate TADs in interphase, and compact and resolve chromosomes in mitosis.[8–12] In these models, a LEF binds to DNA and starts extruding a loop, until the LEF is displaced from DNA or is blocked by a structure that inhibits further movement in one direction (Figure 1). The average size of a loop is determined by the number of LEFs stochastically loaded onto the chromosomes and the “processivity” of the loaded LEFs[11] – with processivity defined as the product of the LEF’s velocity and active residence time on DNA.[9] In such a system, high processivity and high numbers of LEFs would lead to dense loop arrays, whereas low numbers of LEFs with relatively low processivity would lead to sparse looping regimes. In these models “factors that decrease the speed of loop extrusion or reduce LEF residence time will decrease the processivity and thus decrease the average loop size and the degree of chromatid compaction.”[11] In mitosis, high numbers of LEFs with high processivity, acting in the absence of chromosomal boundaries, are sufficient to generate loop arrays across the entire length of chromosomes, driving intra-chromosomal compaction and inter-chromosomal resolution.[11,12] In interphase, the activity

DOI: 10.1002/bies.201800182
of structures that arrest or impede LEF activity (e.g., CTCF sites) complicates the outcomes of LEF activity. However, the action of fewer LEFs with lower processivity, acting within the boundaries generated by LEF-arresting structures, can fully account for the presence of distinct TADs along chromosomes.[9,10] The success of in silico LEF models in replicating the experimentally observed levels of chromosome compaction/organization leads directly to the question of which protein complexes could be capable of acting as LEFs in cells. One set of proteins which have been consistently linked to DNA-looping activities are the SMC (structural maintenance of chromosomes) family of protein complexes.[7,9,10,13]

3. Architecture of SMC Complexes

SMC complexes are found across all kingdoms of life and share a common structural architecture, which is composed of a dimer of two SMC proteins (either homo- or hetero-dimeric). Each SMC protein folds back on itself via a “hinge” domain, which, through long coiled-coil connections, also serves to bring two halves of a “head domain” together to form an ABC-type ATPase. Dimerization of SMC complexes is achieved through association of the hinge domains of the two folded-back proteins. The SMC dimer is always associated with a “kleisin”-type protein,[14] the C-terminus of which binds to the head domain of the \( \kappa \) SMC[15] and the N-terminus to the coiled-coil region proximal to the head domain of the \( \nu \) SMC protein.[16,17] Other non-SMC proteins are also recruited to the SMC complex, predominantly through interactions with distinct parts of the central section of the kleisin subunit. These proteins can be structurally sub-grouped into so-called HAWK (HEAT proteins associated with kleisins) and KITE (kleisin interacting winged-helix tandem elements) proteins.[18] KITEs are found to be associated with archaeal and bacterial SMCs and the eukaryotic SMC5/6 complex, whereas HAWK proteins interact with eukaryotic condensin and cohesin (Table 1).

In eukaryotes, the HAWK-containing SMC complexes, condensin and cohesin appear to be primary regulators of global chromosome structure. The condensin complex was originally identified as being essential for mitotic chromosome condensation in both Xenopus egg extracts and in yeasts,[19–21] while also being an abundant component of the chromosome axis of metazoan mitotic cells, where DNA loops are anchored.[22] In higher eukaryotes, there are two condensin sub-complexes, condensin I and condensin II, which have both distinct and overlapping in vivo functions.[23] Condensin II is constitutively nuclear and can be found stably bound to chromosomes throughout mitosis, while condensin I is excluded from the nucleus until nuclear envelope breakdown, when it dynamically binds to chromosomes.[23,24] Despite the distinct cell biology there appears to be considerable functional redundancy between the two complexes. Only depletion of both condensin I and II leads to the cessation of chromosome condensation.[25] However, conversion of chromosomes into a “mitotic conformation” can be achieved in vitro by just condensin I, with the addition of chromatin re-modellers, histones and topoisomerase II.[26] In contrast, the cohesin complex was identified as being essential for the formation and maintenance of sister chromatid cohesion in eukaryotic cells.[27] Sister chromatid cohesion depends on cohesin being “loaded” in an ATP hydrolysis-dependent manner, such that the SMC complex encircles either one or two DNA duplexes (topological entrapment). During DNA replication and following DNA

---

**Figure 1.** Putative enzymatic cycle of a loop extrusion factor (LEF). The LEF is posited to (a) load onto DNA and then (b) extrude DNA into a loop by translocating either in both or only one direction away from its loading site at a speed determined by its processivity. c) If the processivity of the LEF on DNA is low only a small loop will be generated before the LEF is displaced from DNA and the DNA loop destabilized. d) If the processivity of the LEF on DNA is high a larger loop of DNA will be extruded potentially being blocked by a protein-DNA complex capable of blocking the extruder. At some point the complex will become unstable leading to (c) eventual displacement of the LEF and loop destabilization.
damage, cohesin complexes are converted into a form that
entrap both duplexes of the newly replicated sister chroma-
tids. Following loading, cohesin is actively removed from
chromosomes by the action of one of two pathways: Rad61/Wapl
and Pds5 proteins catalyse the removal of cohesin from chromatin in a manner antagonized by post-translational
modification by the acetyl-transferase Eco1/Eso1/2. The
acetylated cohesin supporting sister chromatid cohesion is
completely ablated in metaphase through proteolytic cleavage of
the cohesin kleisin subunit by the caspase-like enzyme separase.
Cleavage of cohesin triggers the metaphase to anaphase transition, and segregation of the newly replicated chromosomes.

In addition to cohesin’s canonical and essential role in maintaining sister chromatid cohesion in proliferating cells, it has also been found to be involved in distinct roles that are linked to chromosome looping. Cohesin activity supports mitotic rDNA compaction in the budding yeast S. cerevisiae and normal gene expression in metazoan cells, accumulating at known insulator sites, and supporting functional chromosome interactions at
genes. Therefore, cohesin and condensin are closely linked to cis loop formation in interphase and mitosis, suggesting that they could be essential constituents of LEF machines.

4. Loop Extrusion and SMC Complex Dynamics

In silico simulations predict that if SMC complexes are acting as part of LEF machines, then they should display distinct biochemical and cell biological characteristics. First, ablation of SMC activity should lead to a cessation of sequential cis-looping along chromosomes, but not inhibit chromosome loops acting across other looped domains or in trans. Second, generation of loops by LEFs should be associated with the translocation of SMCs along DNA. Third, the coverage and size of looping should be a function of the number and processivity of SMCs on chromatids. These predictions have now been tested in several different systems and contexts, with the data collected to date consistent with SMC complexes acting as loop extruding factors.

4.1. SMC Complexes Are Required for Exclusively in cis DNA Looping

In multiple systems, SMC activity is required for cis-looping of chromosomes, in a manner fully consistent with LEF activity. In mitotic chicken cells, all detectable mitotic chromosomal looping is abolished upon acute degradation of SMC2. Loss of cohesin function causes a complete loss of the cis-loops that generate TADs in fission yeast, mouse liver cells, and human cell lines, and loss of mitotic looping along chromosome arms in budding yeast. As predicted by the loop extrusion model, cohesin is not required for the long-range interactions that constitute compartments, which are thought to be formed through stabilization of diffusive point to point interactions. In Babillus subtilis, the SMC complex (RfSMC) is also required for DNA looping. Therefore, in multiple contexts, both bacterial and eukaryotic SMC complexes are required for cis-loop formation in both interphase and mitosis.

4.2. SMCs Translocate Along DNA to Generate Loops

The in silico LEF model predicts that formation of cis-loops should be accompanied by translocation of the LEF machine along DNA away from its loading site. Translocation of SMC complexes along DNA has been observed in vitro and in vivo, and has also been closely linked with the formation of loops. In vitro the yeast condensin complex loads and translocates along DNA in an ATP-dependent manner and this activity directly facilitates loop formation along the DNA fiber. In contrast, yeast, Xenopus, and human cohesin complexes loaded onto DNA in vitro passively diffuse along DNA in a manner that is influenced by transcription and impeded by protein-DNA blocks indicating that the in vitro assembled complexes, as currently reconstituted, do not have an intrinsic translocase activity. The possibility that cohesin translocation is powered by an extrinsic factor is supported by the in vivo observations that local transcription promotes cohesin translocation away from its initial loading sites. However, elongating transcription complexes are not required for cohesin-dependent looping in interphase mammalian cells. These findings do not rule out...
the possibility that aborted transcription may still drive translocation, for example, through local diffusion of DNA supercoiling[53] or that translocation of cohesin along chromosomes is driven by osmotic pressure.[54,55] The translocation of BsSMC in vivo, away from its loading site, is spatially and temporally linked with DNA loop formation[56] in an ATP hydrolysis dependent manner.[57] In summary, evidence from multiple systems indicate that SMC complexes can translocate along DNA fibers, but it is not yet clear if complexes other than condensin are capable of intrinsically driving translocation and loop formation along the DNA fiber.

4.3. LEF Activity Is Linked to the Processivity of SMC Complexes on Chromatin

The final prediction of the LEF theory is that the processivity of SMC complexes should be directly related to the size of loops generated on the chromosome. Processivity is a function of both the velocity and residence time of the extruding complex. In vitro the velocity of loop extrusion by the yeast condensin complex is, on average, 0.6 kbp s⁻¹ for one-ended loop extrusion.[49] In vivo velocities of potential SMC loop extrusion machines have been inferred from the rates of formation of SMC-dependent loops; 0.9 kbp s⁻¹ for the BsSMC complex[56] and an average rate of no less than 0.375 kbp s⁻¹ for cohesin in mammalian cells.[41] However, at present we do not have comparative data for the loop formation rates of different SMC complexes within the same system. In terms of the residence time of the active complex, we can assume that chromatin stability of the complex (derived from either Fluorescence Recovery After Photobleaching experiments or by the complexes liability to salt extraction from chromatin) is directly relatable to its residence time as an actively extruding complex. This assumption is somewhat complicated for cohesin-type SMC complexes by the likelihood that different populations of chromatin-bound cohesin complexes enforce sister chromatid cohesion or loop extrusion.[58] Notwithstanding this issue, the predicted correlation between increased chromatin residency and loop size has now been examined and verified for cis-loops that are dependent on cohesin or condensin activity. Vertebrate condensin complexes I and II have different residence times of chromatin binding,[59,60] presumably enforced by the different paralogues of kleisin and HAWK proteins that bind to the core SM2/4 heterodimer in the complex (CAP-H, CAP-D2, and CAP-G for condensin I and CAP-H2, CAP-D3, and CAP-G2 for condensin II). Condensin II is stably bound to mitotic chromatin whereas the binding of condensin I is far more dynamic. Varying the ratio of condensin I and condensin II in Xenopus extracts significantly affects the type of compaction generated on chromosomes – a high ratio of condensin I relative to condensin II leads to long thin chromosomes, whereas the opposite ratio leads to relatively short and thick chromosomes.[61] The different chromosome compaction effects of condensin I versus condensin II have been rationalized by modeling the changes in Hi-C contact frequencies of DT40 mitotic chromosomes following acute degradation of either condensin I or condensin II.[38] The change in chromosome structure following acute loss of the more stable and processive condensin II complex can be fully accounted for by the loss of large cis-loops all along the chromosomes. In contrast, loss of the more dynamically binding and therefore relatively non-processive condensin I complex correlates with loss of smaller loops along the chromosomes.[38] The relative number of condensin I and II complexes binding to the chromosome is also consistent with this model.[60] Similarly, the different HAWK proteins found in cohesin complexes also appear to potentiate different types of chromatin loop formation. SA1, a mammalian orthologue of the yeast HAWK protein ScC3, mediates stable chromatin binding of cohesin and long-range looping, whereas a more distinct variant SA2 generates dynamically binding cohesin complexes and enforces short-range looping.[62] Finally, increasing the stability of cohesin on chromatin by depletion of factors involved in the active unloading of cohesin, Wapl, or Pds5A/B, results in larger loops forming along chromosomes.[63,64] In summary, the type of looping driven by SMC complex appears to be directly linked to the stability of their association with chromosomes, consistent with, and as predicted by, the LEF model.

5. How Could Different SMC Complexes Function as LEFs With Different Loop Forming Properties?

All SMC complexes are presumed to utilize the energetics of ATP binding and hydrolysis to effect conformational changes on interacting DNA segments. In the LEF model of SMC action, changes in processivity of the complex are presumably determined by factors that regulate ATP turnover and/or affect the stability of SMC–DNA interactions; multiple lines of evidence have indicated that kleisin/HAWK/KITE factors can influence both properties. Condensin and cohesin SMC heterodimers will only hydrolyse ATP in vitro and bind to DNA when associated with the kleisin/HAWK subcomplex.[65–67] HAWKs also act as both positive and negative regulators of ATP hydrolysis. The binding of budding yeast HAWK proteins ScC2 and Pds5 to cohesin is mutually exclusive, where ScC2 promotes hydrolysis, and Pds5 inhibits hydrolysis.[68] Different HAWKs are also likely to directly influence the translocation of SMC complexes through provision of direct DNA contacts. The variation in stability of such contacts could either vary the rates of loading/unloading onto DNA, or facilitate progress of the complex through its ATP-driven enzymatic cycle. In condensin, the interaction between the C-terminal region of the kleisin (Brn1/CAP-H) and its interacting HAWK (Ycg1/CAP-G) generates a stable but sequence-independent DNA binding site (“seat belt model”).[68] Other HAWK proteins also have DNA binding properties. The cohesin HAWK protein ScC3 binds DNA in a similar manner to Ycg1,[69] and ScC2 appears to have structural similarity to other DNA-binding HAWK proteins,[67] which could account for its direct DNA binding activity.[70] KITE proteins are also capable of direct DNA binding. The SCM5/S6 KITE Nse3 directly binds double-stranded DNA without obvious structural preferences.[71] At least for the condensin complex, disruption of DNA binding by the HAWK protein leads to impaired activity of the complex. Mutation of the Ycg1 DNA binding groove impedes loop extrusion and DNA compaction by condensin in vitro,[45,72] It is not known if disrupting the DNA
biological and biochemical evidence that SMC complexes could act as LEF machines on DNA, the mechanistic nature of how the ATP-dependent action of SMC complexes on DNA results in loop extrusion remains speculative. Here we outline several models that have been proposed for SMC-dependent loop extrusion. The first model (Figure 3a) builds on the paradigm established for how SMCs topologically entrap DNA. This model posits that SMC complexes that have utilized ATP hydrolysis to topologically...
Figure 2. Configuration of SMC complexes at different stages of the ATP binding and hydrolysis cycle. The cartoon of SMC structures is based on the electron microscopy of isolated *Xenopus* SMC complexes and structural analysis of the *Bs*SMC complex and the SMC-like Rad50 complex. a) In the absence of ATP binding the coiled coils are relatively closely juxtaposed and the head domains are not aligned. b) ATP binding to the head domains stabilizes head-head engagement of the nu (ν) and kappa (κ) SMC proteins and drives opening of the coiled coils. c) Isolated SMC complexes, particularly cohesin, are observed to have an open coiled coil arrangement with the head domains potentially separated. This is postulated to represent the configuration after ATP hydrolysis but before ADP has been displaced from the ATP binding pockets. d) Cartoon of an SMC complex in the ATP-bound form where the different proposed interaction spaces of DNA with the SMC complex discussed in Sections 6 and 7 are proposed to occur.
entrap two DNA duplexes can act as a slip link on the DNA if the two regions are in cis, thus able to form DNA loops through passive diffusion. In silico modeling of diffusive sliding of DNA-bound SMC complexes predicts that this action is capable of forming TADs, even without the action of extrinsic motors, such as transcription. Such action is relevant to either one SMC complex entrapping two strands (as shown in Figure 3a) or two linked SMC complexes each entrapping a single duplex.

Figure 3. Models of SMC action as a LEF. a) The “sliplink/diffusion” model of SMC loop extrusion. Following topological entrapment of two DNA duplexes connected in cis, the SMC complex could generate a loop through either passive translocation down the fibers, or could be “pushed” down the fiber by an accumulation of either DNA topological stress or transcription within the looping region. b) The “walking” model of SMC loop extrusion. As currently presented, this model predicts that ATP action at the heads promotes “walking” of the two head domains along a stretch of DNA in a manner analogous to kinesin translocation along microtubules. If this occurred while a second region of DNA was attached to the hinge domain, then one-ended loop extrusion would occur. c) In the “rock-climbing” model based on MukBEF dynamics in E. coli, sequential topological binding and extrusion of adjacent regions of DNA by two attached SMC complexes would generate loop extrusion activity. d) In the “loop capture ratchet/scrunching” model, the opening of the coiled coils by ATP binding to the head domains enables capture of a length of DNA by the DNA binding sites at the hinge and head domains. Following hydrolysis this loop of DNA is then forced by coiled coil closure into a lower compartment where it is captured between the closed coiled coils of the SMC and the kleisin-HAWK/KITE proteins. Sequential action of this nature will build up the size of the extruded loop.
At present it is not clear if passive diffusion of an SMC complex could account for the rapid rates of SMC-dependent loop formation that occur in vivo.[38,41,56] Also, this model argues that the intrinsic ATP hydrolysis-dependent movement of the yeast condensin complex is not a universal feature of SMC complex action in vivo.

If the in vitro DNA translocation and loop extrusion activities of yeast condensin are representative of all SMC looping action, then these experiments indicate that the ATP hydrolysis-dependent movement of the SMC complex takes place in distinct steps along the DNA, estimated to be around 50 nm in length. A stepwise action for SMC translocation argues against the continuous type of translocation seen in DNA helicases or polymerases. The combined characteristics of translocation along DNA have suggested several different models for how SMC complexes could act as LEF machines, including "walking," "rock climbing," or "loop capture ratchet" / "scrunching" mechanisms[44,89,99] (Figure 3b–d). The "walking" model requires three distinct DNA binding sites within the complex, at the hinge and proximal to each of the SMC head domains, to coordinate translocation along one section of DNA while maintaining a linkage with a second (Figure 3b). The model shown in Figure 3b assumes that the hinge generates one stable DNA binding site while each of the head domains and its associated proteins is capable of more transient, dynamic DNA binding, presumably through the direct DNA binding capability of associated HAWKs. However, the current biochemical data for SMC complexes indicates that a dsDNA binding site at the hinge is only accessible following ATP binding at the heads[93,95] and is not constitutive.

Alternatively, the co-ordinated action of at least two SMC complexes might facilitate a "walking"-type process. Coordinated action of two linked SMC complexes has been proposed for the action of the Escherichia coli SMC complex MukBEF.[100] MukBEF acts as a dimer of SMC dimers when it is loaded onto DNA,[100] which are connected by an N-terminal mediated dimerization of the kleisin MukF.[101] Their proposed form of action is a "rock climbing" movement on two DNA strands to generate LEF activity (Figure 3c), where staggered topological entrapment, followed by extrusion of DNA through the kleisin interface of two SMC complexes in close proximity would generate loop extrusion.[100,102]

Finally, the "ratchet/scrunching" model[44,89,99] is proposed to work by one or two SMC complexes opening their coiled coils (following ATP binding) before "snapping back" of the coiled coils after ATP hydrolysis and ADP displacement (Figure 3d). It is proposed that the process of snapping back of the coiled coils forces DNA, interacting with the hinge and head following ATP binding, into a lower compartment, formed by the interface between the SMC dimer and the bridging kleisin/non-SMC subcomplex.[89,99,103] A biophysical model based on this process has recently described this hypothetical mechanism in detail.[99]

8. Conclusions

Here we have reviewed how the current body of available cell biology and biochemical data for the SMC complexes agrees with the hypothesis that SMC complexes act as LEF machines generating chromatin loops in cells. While the available experimental data sits comfortably with this hypothesis, a number of key questions remain to be answered. In particular, further in vitro and in vivo evidence is required to demonstrate that SMC complexes other than condensin can actively translocate and generate cis-loops in DNA; it is not clear how loop extrusion by SMC complexes can occur in chromatized substrates and if these substrates require remodeling for SMC translocation, and, most crucially, it is unclear how an enzymatic cycle that can generate DNA loops can also be controlled and regulated to generate the other known cellular functions of SMC complexes, including sister chromatid cohesion and facilitation of DNA repair pathways. Potentially, the differences in LEF activity provided by each SMC complex, could account for their requirement in different genomic contexts.[42] How far LEF activity can be used to explain the numerous and essential functions of SMC complexes throughout evolution has yet to be determined.

Abbreviations

ATP, adenosine tri-phosphate; HEAT, Huntingtonin, elongation factor 3 (EF3), protein phosphatase 2A (PP2A), and the yeast kinase TOR1; HAWK, HEAT proteins associated with kleisins; KITE, kleisin interacting winged-helix tandem elements; LEF, Loop extruding factor; SMC, Structural Maintenance of Chromosomes; TAD, topologically associated domain; .

Acknowledgements

This work was funded by the U.K. Biotechnology and Biological Sciences Research Council (BBSRC Grant number BB/S001425/1) (J.B.), the Medical Research Council (MRC Grant Number MR/P018955/1) (A.W. O.), and the ERC (Grant code 311336) (S.A.S.).

Conflict of Interest

The authors declare no conflict of interest.

Keywords

chromosome structure, HAWK, KITE, loop extrusion, SMC complexes

Received: September 14, 2018
Revised: November 5, 2018
Published online: December 3, 2018

[1] J. R. Dixon, S. Selvaraj, F. Yue, A. Kim, Y. Li, Y. Shen, M. Hu, J. S. Liu, B. Ren, Nature 2012, 485, 376.
[2] E. P. Nora, B. R. Lajoie, E. G. Schulz, L. Giorgetti, I. Okamoto, N. Servant, T. Piolot, N. L. van Berkum, J. Meisig, J. Sedat, J. Gribnau, E. Barillot, N. Blüthgen, J. Dekker, E. Heard, Nature 2012, 485, 381.
[3] T. Sexton, E. Yaffe, E. Kenigsberg, F. Bantignies, B. Leblanc, M. Hoichman, H. Parrinello, A. Tanay, G. Cavalli, Cell 2012, 148, 458.
