TU Delft

Delft University of Technology

Exploring SMC motor proteins and how to stop them

Barth, R.

DOI
10.4233/uuid:39950834-acc0-465e-959c-600fe40f2b57

Publication date
2024

Document Version
Final published version

Citation (APA)
Barth, R. (2024). *Exploring SMC motor proteins and how to stop them*. [Dissertation (TU Delft), Delft University of Technology]. https://doi.org/10.4233/uuid:39950834-acc0-465e-959c-600fe40f2b57

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

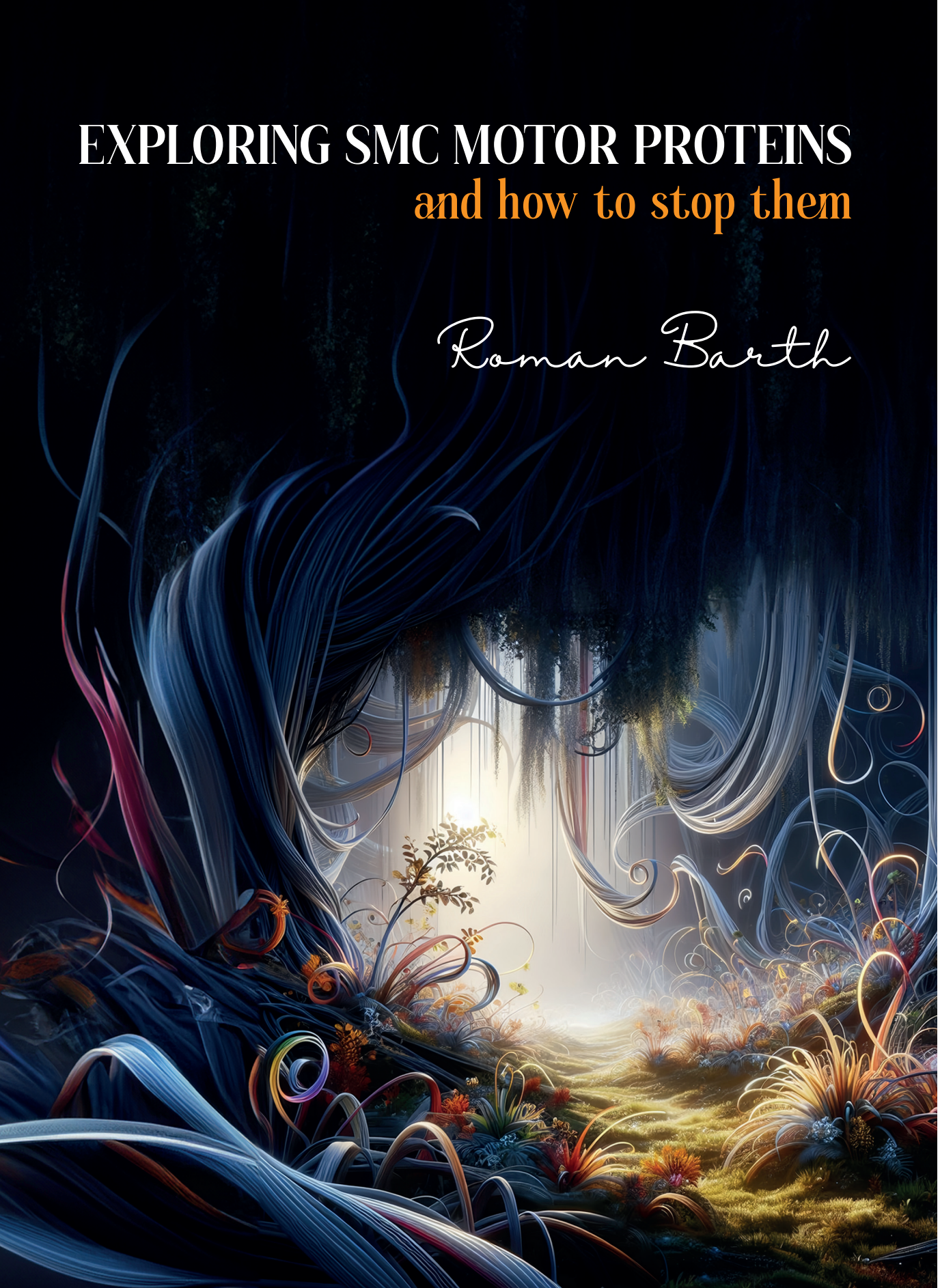
Copyright
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.
For technical reasons the number of authors shown on this cover page is limited to a maximum of 10.

EXPLORING SMC MOTOR PROTEINS
and how to stop them

Roman Barth



Exploring SMC motor proteins and how to stop them

Exploring SMC motor proteins and how to stop them

Dissertation

for the purpose of obtaining the degree of doctor
at Delft University of Technology
by the authority of the Rector Magnificus, Prof. dr. ir. T.H.J.J. van der Hagen,
chair of the Board for Doctorates,
to be defended publicly on
Friday 21 June 2024 at 12:30 o'clock

by

Roman BARTH

Master of Science in Applied Physics,
Technische Universiteit Delft, Netherlands,
born in Bruchsal, Germany.

This dissertation has been approved by the promotors.

Promotor:	Prof. dr. C. Dekker
Copromotor:	Prof. dr. A. M. Dogterom

Composition of the doctoral committee:

Rector Magnificus	Chairperson
Prof. dr. C. Dekker	Technische Universiteit Delft, promoter
Prof. dr. A. M. Dogterom	Technische Universiteit Delft, copromoter
Independent members:	
Prof. dr. B. Rowland	Technische Universiteit Delft
Prof. dr. L. A. Mirny	Massachusetts Institute of Technology
Prof. dr. W. L. de Laat	University Medical Center Utrecht
Dr. A. S. Hansen	Massachusetts Institute of Technology
Reserve member:	
Prof. dr. ir. S. J. J. Brouns	Technische Universiteit Delft
Other member:	
Prof. dr. J.-M. Peters	Medical University of Vienna

Keywords: SMC protein complexes, DNA loop extrusion, genome organization, single-molecule fluorescence, Magnetic Tweezers

Printed by: Proefschrift Specialist

Cover: Roman Barth using Microsoft Copilot GPT.

Style: TU Delft House Style, with modifications by Moritz Beller and Roman Barth

Copyright © 2024 by R. Barth

An electronic version of this dissertation is available at
`http://repository.tudelft.nl/`.

Contents

I General Introduction **3**

1 Looping the Genome with SMC complexes **5**
1.1 The hierarchical organization of the genome . . . 7
1.2 A short history of SMC proteins . . . 9
1.3 The versatile roles of SMCs in chromosome biology . . . 13
1.3.1 Activity of SMCs during cell division . . . 13
1.3.2 Activity of SMCs during interphase . . . 15
1.4 Biophysics of DNA loop extrusion by SMCs . . . 19
1.4.1 Single-Molecule visualization of loop extrusion by SMC complexes 19
1.4.2 Loop extrusion rate . . . 20
1.4.3 Stalling force . . . 20
1.4.4 Directionality of loop extrusion . . . 21
1.4.5 Topology of SMCs and DNA . . . 21
1.4.6 Conformational states of SMCs . . . 22
1.5 Models of DNA loop extrusion . . . 24
1.5.1 The Scrunching/Swing-and-Clamp Model . . . 24
1.5.2 The Brownian Ratchet Model . . . 24
1.5.3 The DNA Segment Capture/Hold-and-Feed Models . . . 25
1.5.4 The Reel & Seal model . . . 26
1.6 From naked DNA to chromatin . . . 28
1.7 Conclusion . . . 30

2 In this thesis **32**

II Towards a mechanistic understanding of DNA loop extrusion **37**

3 SMC complexes can traverse physical roadblocks bigger than their ring size **39**
3.1 Introduction . . . 41
3.2 Results . . . 42
3.2.1 Visualization of a condensin encountering an obstacle while extruding a DNA loop . . . 42
3.2.2 Loop-extruding condensin traverses DNA-bound proteins such as nucleosomes, RNA polymerase, and dCas9 . . . 44
3.2.3 Condensin and cohesin can pass roadblocks bigger than their ring size . . . 44
3.3 Discussion . . . 49

3.4 Limitations of the study . . . 50
3.5 Methods . . . 52
3.6 Supplementary Information . . . 60

4 The topology of DNA loop extrusion 67
4.1 Introduction . . . 69
4.2 Results . . . 71
4.2.1 Description of the pseudotopological models . . . 71
4.2.2 Predictions by the pseudotopological models . . . 72
4.2.3 Experimental tests of the predictions made above . . . 74
4.3 Conclusion . . . 81
4.4 Methods . . . 82
4.5 Supplementary Information . . . 84

5 SMC motor proteins extrude DNA asymmetrically and contain a direction switch 87
5.1 Introduction . . . 89
5.2 Results . . . 90
5.2.1 Phases of active DNA loop extrusion are interspersed by loop diffusion and loop slipping . . . 90
5.2.2 All eukaryotic SMC complexes extrude DNA asymmetrically . . . 91
5.2.3 NIPBL excess increases the occurrence of direction switches in cohesin-mediated loop extrusion . . . 95
5.2.4 In the absence of NIPBL, loops can diffuse and slip but cannot be extruded . . . 96
5.2.5 Extrusion direction switches coincide with an exchange of NIPBL on cohesin . . . 96
5.3 Discussion . . . 97
5.3.1 DNA loop extrusion is conserved, asymmetric, and can switch direction . . . 97
5.3.2 Mechanistic models that enable DNA strand exchange . . . 98
5.3.3 The physiological relevance of extrusion direction switching . . . 99
5.4 Limitations of the study . . . 101
5.5 Methods . . . 102
5.6 Supplementary Information . . . 109

6 All eukaryotic SMC proteins induce a twist of -0.6 at each DNA-loop-extrusion step 123
6.1 Introduction . . . 125
6.2 Results . . . 126
6.2.1 Assessing DNA twist that is induced at individual steps in DNA loop extrusion . . . 126
6.2.2 ATP-binding induces an equal amount of twist in each loop-extrusion step, irrespective of the step size . . . 128
6.2.3 All eukaryotic SMCs induce $\Delta L_k = -0.6$ per step into the extruded DNA loop . . . 130

6.2.4 A plasmid assay to measure cohesin-mediated DNA supercoiling . 130
6.2.5 Simulations based on single-molecule behavior recapitulate the bulk plasmid experiments . . . 133
6.2.6 A mutation of the DNA binding sites on cohesin's ATPase heads is predicted to lower cohesin's stalling torque . . . 134
6.2.7 SMC1 4E cohesin negatively supercoils the loop during single loop extrusion steps . . . 135
6.3 Discussion . . . 137
6.4 Methods . . . 141
6.5 Supplementary Information . . . 148

III Halting DNA loop extruders 159

7 CTCF is a DNA-tension-dependent barrier to cohesin-mediated loop extrusion 161
7.1 Main . . . 163
7.1.1 CTCF characterisation *in vitro* . . . 163
7.1.2 CTCF is a polar barrier to cohesin . . . 167
7.1.3 CTCF is a polar barrier to DNA looping . . . 167
7.1.4 CTCF is a DNA-tension-dependent barrier . . . 168
7.1.5 Transient loop anchoring by CTCF . . . 169
7.1.6 CTCF can switch the direction of looping . . . 171
7.2 Discussion and conclusions . . . 171
7.3 Methods . . . 174
7.4 Supplementary Information . . . 190

8 Two CTCF motifs impede cohesin-mediated DNA loop extrusion 207
8.1 CTCF's NTR impedes DNA loop extrusion activity of cohesin . . . 210
8.2 The CTCF NTR transforms cohesin into a unidirectional extruder . . . 213
8.3 The CTCF NTR impacts cohesin-driven DNA loop extrusion beyond the YDF motif . . . 214
8.4 The YDF motif transforms cohesin into a unidirectional extruder . . . 217
8.5 The N-terminal KTYQR motif inhibits DNA loop extrusion by cohesin . . . 217
8.6 An AlphaFold screen identifies a binding pocket of the KTYQR motif on STAG1/2 . . . 221
8.7 Conclusions . . . 222
8.8 Methods . . . 225
8.9 Supplementary Information . . . 231

IV Outlook and summary 241

9 Open questions and future directions 243
9.1 Is DNA temporarily pseudotopologically embraced during loop extrusion? . 246
9.2 What determines whether a loop extruder can switch directions? . . . 247

9.3 Which conformational changes account for the 'step' in the loop extrusion cycle? . . . 248
9.4 How do dimers of SMCs work and what are the minimal structural requirements for loop extrusion? . . . 248
9.5 How are SMC complexes regulated *in vivo* by their interactome?. . . . 249
9.6 Why so complicated?. . . . 249

Summary **253**

Samenvatting **257**

Zusammenfassung **261**

References **265**

Acknowledgments **303**

Curriculum Vitæ **305**

List of Publications **307**

I

General Introduction

1

Looping the Genome with SMC complexes

1

Abstract

Genetic information encoded in the DNA sequences is maintained, replicated, and transcribed across the tree of life. Organisms evolved multiple hierarchical layers of chromosome organization which ensure that DNA can be contained but also processed within individual cells. These layers range from transcriptionally active and inactive regions on the scale of several Mbp within individual chromosomes to loops and DNA and individual DNA-binding proteins. This thesis focuses on SMC (Structural Maintenance of Chromosomes) protein complexes that are an evolutionarily conserved family of motor proteins that hold sister chromatids together and fold genomes throughout the cell cycle by DNA loop extrusion. These complexes play a key role in a variety of functions in the packaging and regulation of chromosomes, and they have been intensely studied in recent years. Despite their importance, the detailed molecular mechanism for DNA loop extrusion by SMC complexes remains unresolved. Here, we provide an introduction to chromosome biology with a focus on the roles of SMCs in chromosome biology, and we particularly review in vitro single-molecule studies that have recently advanced our understanding of SMC proteins. We describe the mechanistic biophysical aspects of loop extrusion that govern genome organization and its consequences.

1.1 The hierarchical organization of the genome

The length of genomes in various species varies considerably, from the largest known genome of ~150 billion base pairs (bp) of *Paeonia japonica* [2] to 3.4 billion bp for *Homo sapiens*, down to 0.11 million bp for the bacterium *Nasuia deltocephalinicola* [3]. All cells require their DNA to be compacted at least 1,000-fold to fit within the finite space of a micrometer-sized nucleus (or cell volume for bacteria). Additionally, information encoded in the genome continuously needs to be accessed and processed. Accordingly, the genome is hierarchically organized as well as constantly and actively reshaped to sustain all cellular functions.

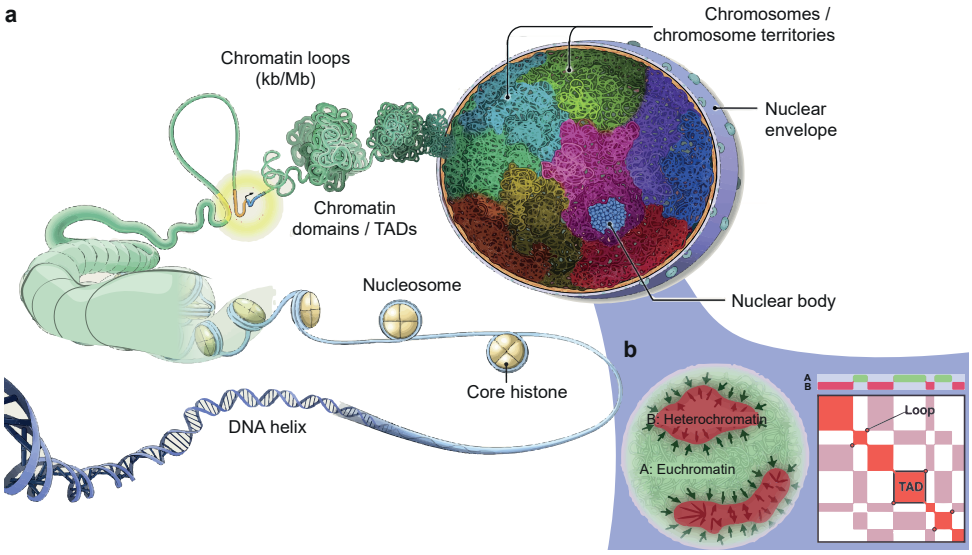


Figure 1.1 | The eukaryotic genome is hierarchically organized.
(**a**) The DNA double helix is wrapped around the histones to form nucleosomes, the basic packaging unit of chromatin. This chromatin fiber is extruded into loops by SMC complexes which results in chromatin domains and topologically associating domains (TADs). On a larger scale, individual chromosomes occupy their distinct region within the nucleus (chromosome territories). (**b**) (Left) Euchromatin and heterochromatin (or A/B compartments) spatially segregate on the scale of micrometers due to an attractive force between heterochromatin regions. (Right) Off-diagonal plaid patterns in Hi-C contact matrices are characteristic of A/B compartments while squares along the matrix diagonal indicate TADs, often demarcated by corner dots which represent cohesin-mediated DNA loops spanning the TAD. Adapted from [4], Hi-C matrix in (b) adapted from [5].

The quest to understand the organization of genomes across the tree of life has undergone significant advancements, providing a more and more detailed understanding of the 3D architecture from macroscopic to molecular scales. Traditionally elucidated through light microscopy (see Figure 1.2a for a light microscopy image from 1882), recent methodologies, including high-throughput chromosome conformation capture methods [6, 7], have refined how we think about the structure of chromatin [8].

In eukaryotes, the genome consists of multiple linear segments of DNA, the chromosomes. At the highest level of DNA organization, individual chromosomes occupy distinct regions within the nucleus, called chromosome territories [9, 10]. Each chromosome contains re-

gions which the cell needs to access and transcribe ('read out'). Which part of the genome is being transcribed is tightly regulated and depends on the cell cycle stage, stimuli the cell receives, as well as the developmental stage of a cell. The genome is thus broadly separated into regions which are transcriptionally more active (euchromatin, as identified by microscopy [11], or the A-compartments as identified by Hi-C [7]) is segregated from the transcriptionally more inactive heterochromatin/B-compartment (Figure 1.1b). This separation holds even at the level of individual chromosomes in individual cells [12] and relies on an attractive force between heterochromatic regions [13], likely mediated by (micro-)phase separation, for instance by DNA-bound proteins such as heterochromatin protein 1 (HP1) [14, 15]. Even though the compartmentalization into A/B compartments concerns domains spanning large stretches of DNA, up to several Mb, gene-regulatory elements may also separate into smaller hubs reminiscent of the A-compartment [16]. However, the spatial separation of these smaller regions is not necessarily due to exclusion from self-attracting heterochromatin but is potentially mediated by clustering of RNA polymerases and transcription factors [17–19]. The attraction of preferentially transcribed or silenced genomic regions can even overcome the spatial separation of individual chromosomes, causing chromosome territory borders to intermingle significantly [20, 21], depending on the transcriptional activity of the chromosome territory borders [20, 22].

At the molecular level, the fundamental unit of chromatin, the nucleosome, emerges as a key agent to enable the compaction of chromosomes [23–25]. Nucleosomes are wrapped around DNA, thereby compacting it, and separated by linker DNA from neighbouring nucleosomes, forming "beads on a string" architecture. Several nucleosomes in turn associate to form "clutches" or "blobs" of nucleosomes [26–30], allowing DNA to be more or less accessible for DNA-processing proteins.

Finally, these nucleosomal fibres are extruded into DNA loops by Structural Maintenance of Chromosomes (SMC) complexes. The genome of organisms harbouring DNA loop extruders is looped throughout the cell cycle. To enable cell division, the newly replicated sister chromatids have to be segregated into the respective daughter cells. For this process to be faithful, each sister chromatid is highly compacted. SMCs play a dual role at this stage. Cohesin ensures that the sister chromatids remain catenated before the spindle can pull each chromatid into its respective daughter cell while condensin complexes shrink the chromatids' volume by extruding the majority of its DNA into thousands of loops. However, the genome is also looped during interphase, which, among other influences, gives genomes across the tree of life their characteristic structure [31]. Furthermore, loop extrusion is utilized to fine-tune transcription regulation in response to stimuli or during development (see subsection 1.3.2).

While extruding, SMCs may encounter DNA-bound proteins, constituting a potential roadblock to loop extrusion. A specific molecular roadblock, CCCTC-binding factor (CTCF), is set along chromosomes by the genomic sequence. While most biological molecules do not appear to be a roadblock to loop extrusion, CTCF is a semi-permeable and directional roadblock to cohesin (see chapter 7 and [32–36]). Most CTCF binding sites are oriented convergently to each other [37, 38], delineating domains. Hi-C measurements, which average over a population of cells, display a square pattern emanating from the main diagonal of Hi-C matrices (Figure 1.1b) [39–41]. Such square patterns are called 'Topologically

Associating Domains' (TADs), are typically a few hundred kbp long, and originate from DNA loop extrusion and the interaction of extruders with boundary elements such as CTCF [42]. Note that the off-diagonal rectangular patterns reflect passive A/B compartment formation. Resulting from different mechanisms, compartmentalization and DNA loop extrusion are antagonistic but co-exist [43, 44]. Most frequently, TADs are demarcated by cohesin-mediated DNA loops between the TAD boundaries which colocalize with CTCF. Nested CTCF boundaries or weaker boundary elements to DNA loop extrusion (see section 1.6) give rise to sub-TADs. In organisms lacking CTCF, such as prokaryotes, domains of preferentially interacting chromatin can also be observed, albeit mediated largely by transcription rather than SMC-mediated looping (even though prokaryotes show another striking feature owing to looping of their genome, see below) [45–47].

As we show in chapter 6, SMC complexes also supercoil DNA during loop extrusion [48–51]. Negative and positive supercoiling is the under- or overwinding of the DNA double helix, respectively [52]. Nuclear processes like binding of the nucleosome to DNA [53] or transcription [54–57], as well as replication [58] generate supercoils. When accumulated, both negative and positive supercoiling cause the DNA to form plectonemes [59]. Further negative supercoiling causes the DNA double helix to separate ('DNA melting') [60], creating hotspots for DNA breakage. Excessive positive supercoiling may prevent proteins from binding to DNA [61]. DNA topoisomerases and gyrases thus resolve and/or introduce supercoils to maintain a viable degree of supercoiling *in vivo* [61].

1.2 A short history of SMC proteins

SMC complexes play vital roles at all stages of chromosome biology throughout the cell cycle, from forming the characteristic textbook X-shaped mitotic chromosomes to actively regulating gene expression, controlling replication timing, and conferring genome stability. To do so, SMCs have evolved to be regulators of DNA looping, as large-scale (0.1-1 Mbp) loops constitute the basic motif of the spatial structure of the genome in all domains of life. SMC complexes are multi-subunit motor proteins that use the energy of ATP hydrolysis to fold genomes into loops throughout the cellular life cycle. Eukaryotic cells have three classes of SMC complexes — cohesin, condensin, and SMC5/6 — that exhibit distinct roles. For example, condensin plays a central role in chromosome compaction and segregation, whereas cohesin is essential for sister chromatid cohesion and interphase genome organization, and SMC5/6 is implicated in DNA repair and checkpoint responses. Most archaea and bacteria, on the other hand, feature one major type of SMC proteins, for example, MukBEF in Escherichia coli and SMC-ScpAB in *B. subtilis*, whereas other SMC-like proteins (e.g. MksBEF) appear to coexist in certain bacteria [62]. These complexes facilitate the resolution of sister chromosomes before cell division, thus exhibiting functional similarity to eukaryotic condensin. DNA loop extrusion activity has been demonstrated on a single-molecule level for all eukaryotic SMCs, whereas some of the best *in vivo* evidence that SMC complexes extrude DNA loops comes from studies in bacteria.

Going back in history, the first documented observation of structural aspects of chromosomes was made in 1882 by Walther Flemming (Figure 1.2a) [65]. He observed emanating loops in axolotl oocyte lampbrush chromosomes, just a decade after DNA was identified

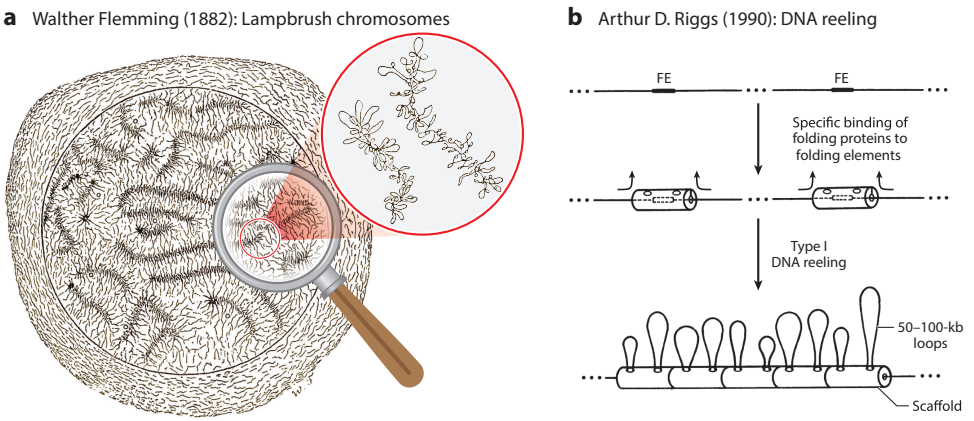
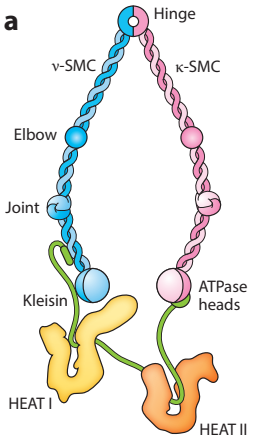


Figure 1.2 | Historic precursors to the DNA loop extrusion hypothesis.
(**a**) In 1882, Walther Flemming observed DNA loops in axolotl oocytes using light microscopy. These loops were due to the repulsion of highly expressed genes. Used with permission of The Royal Society (U.K.), from [63]; permission conveyed through Copyright Clearance Center, Inc. (**b**) Arthur D. Riggs first proposed a DNA reeling mechanism in *cis* to explain the formation of mitotic chromosomes, inspired by the bi-directional translocation of bacterial type I-restriction enzymes. Even though details changed, the principle of Riggs' mechanism remains valid until today. Adapted from [64]. Abbreviation: FE, folding element.

[66]. The notion of a proteinaceous scaffold to which loops of 30-90 kb are anchored prevailed for more than a century [67] after its initial description by Flemming. However, later research showed that lampbrush chromosomes actually result from high levels of transcription, which causes the transcript-decorated DNA to repel itself into a loop-like structure [68, 69]. DNA looping proved to be an attractive mechanism throughout history to explain a range of phenomena such as transcription initiation of prokaryotic operons [70], the action of enhancers on promoters in eukaryotes [71], and recombination [72]. Looping between genomic loci can occur over distances of hundreds of base pairs via dimerization of DNA-bound proteins or through supercoiling [73]. Such relatively small loops have to overcome an energy cost associated with the bending of DNA on the length scale of the 50 nm persistence length of DNA, for which DNA elasticity is a main determinant. As loops become longer, the DNA twisting and bending energies diminish, but this comes at the expense of an entropic cost, which is reflected in the decrease in the contact probability of any two genomic loci as their genomic separation increases [74]. To achieve DNA looping over long genomic distances, an active energy-consuming DNA reeling mechanism was proposed. Early evidence came from bacterial type I restriction enzymes, specifically EcoK, which were proposed to bind to recognition sites and translocate bidirectionally away until another EcoK is encountered [75]. In 1990, Arthur D. Riggs postulated that such a mechanism can be used to ensure specific looping in cis, i.e., along the same chromosome [64], an idea that was indeed confirmed with single-molecule experiments two decades later [76]. Quantitative estimations of the compaction of mitotic chromosomes by Riggs were sound: DNA reeling and encounters of 1,000 loops of 50-100 kb in length would result in an axial protein backbone of ~5 µm long, roughly the size of a mammalian mitotic chromosome (Figure 1.2b).

Even before Riggs' postulation of DNA reeling, a gene named stability of minichromosomes 1 (SMC1) was discovered in Saccharomyces cerevisiae [77, 78], while the discovery of more SMC genes in various organisms followed in the 1990s. Mutations of SMC1 inferred with faithful segregation of hybrid plasmids containing a centromeric locus [77]. The wildtype SMC1 gene was then discovered in 1993 by the Koshland lab and computational analysis of its sequence revealed early that the SMC1 protein contains two central coiled-coil regions flanked by a nucleotide-binding head [78]. A series of essential genes of various model organisms was discovered in the 1990s, which all turned out to be part of the family of SMC proteins: mukB in *E. coli* [79], which soon therafter was found to homodimerize [80], Smc2 in *S. cerevisiae* [81], XCAP-C and XCAP-E, members of the condensin complex in *Xenopus* egg extracts [82] and Smc1 and Smc3 forming the cohesin complex [83, 84] together with Rad21/Scc1 and STAG1/Scc3 [85–87] as well as Smc2 and Smc4 which form the condensin complex [82, 83] together with their accessory subunits CAP-D2, CAP-G, and CAP-H [88]. Finally, Smc5 and Smc6 [89, 90] comprise the SMC5/6 complex together with six non-SMC subunits Nse1-6 [91–95]. The annotation of the acronym SMC for this ring-shaped protein complex (Figure 1.3a,b) was later changed to **S**tructural **M**aintenance of **C**hromosomes. Although initially discovered for their deterministic impact on chromosome segregation, the relation of SMC proteins to Riggs' postulated DNA reeling activity was not straightforward. This model, one of many at the time, was largely ignored until the development of chromosome conformation capture technologies over a decade later. Two seminal papers, one focused on mammalian cells and the other on the bacterium *B. subtilis*, that were published within months of each other provided the first *in vivo* data that hinted toward the fact that SMC complexes could function as loop extruders [96, 97]. These studies led to further *in vivo* and *in vitro* experiments that culminated in single-molecule visualizations of SMC-driven DNA loop extrusion (Figure 1.3c) [98]. In this chapter, we discuss recent advances in the field of SMC proteins, focusing on the role of *in vitro* single-molecule studies, which continue to be pioneering in advancing our understanding of SMC proteins. After briefly reviewing the various roles of SMC proteins during mitosis and interphase, we dive into the various structures of SMC proteins found throughout the ATPase cycle. We discuss the currently unresolved but heavily debated mechanism of DNA loop extrusion and the effects of abundant roadblock DNA-binding proteins on chromatin that might interfere with loop extrusion *in vivo*.

1



b

		Bacteria		Eukaryotes		
		Bacillus subtilis	*Escherichia coli*	SMC5/6	Cohesin	Condensin
Core subunits	ν-SMC	SMC	MukB	Smc6	Smc3	Smc2
	κ-SMC	SMC	MukB	Smc5	Smc1	Smc4
	Kleisin	ScpA	MukF	Nse4	Scc1	Brn1
Accessory subunits	KITE	ScpB	MukE	Nse1 Nse3		
	HEAT I				Scc2 Pds5	Ycs4
	HEAT II				Scc3	Ycg1
				Nse2 Nse5 Nse6		

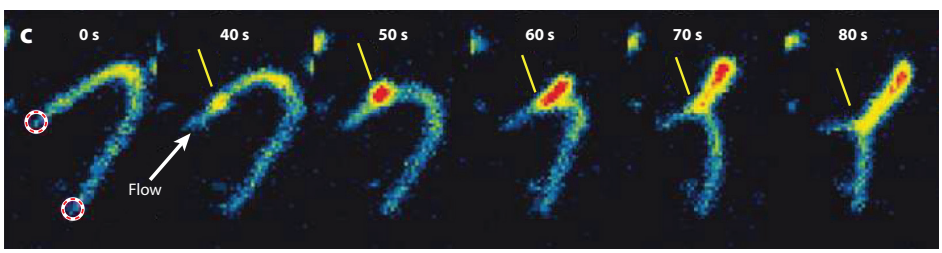


Figure 1.3 | SMC complexes extrude loops of DNA.
(a) General architecture of SMC complexes. The core of SMC complexes is characterized by a pair of ≈50 nm long antiparallel coiled-coil protein subunits named SMCs. These SMCs dimerize on one end (the hinge) and at their other end (the ATPase 'head' domains), mutually connected through a kleisin protein, resulting in a ring-like structure. The kleisin recruits additional regulatory subunits. One family of these regulators, the Kite dimers (Kleisin interacting winged-helix tandem elements) interact with SMC-kleisin rings from bacteria, archaea, and the eukaryotic SMC5/6 complex. Condensin and cohesin further evolved by replacing Kite domains with Hawk domains (HEAT-repeat proteins associated with kleisin) [99]. Condensin has two Hawks (Ycg1 and Ycs4 in budding yeast) and cohesin has three (Scc3, Scc2, Pds5 in budding yeast). While Ycg1 and Ycs4 are stably associated with condensin, as is Scc3 with cohesin, Scc2 and Pds5 transiently associate with cohesin and compete for cohesin binding [100, 101]. Both Kite and Hawk proteins play a crucial role in DNA binding and ATPase activity of the SMC complexes [102–104]. In the case of SMC5/6, three additional subunits exist, namely Nse2, Nse5 and Nse6. **(b)** Name of subunits comprising SMC complexes. **(c)** Single-molecule imaging demonstrated DNA loop extrusion. Single-molecule snapshots from Ganji et al. [98]. The flow is directed toward the top-right, leaders indicate the position of the extruded loop.

1.3 The versatile roles of SMCs in chromosome biology

While initially discovered and studied for their role in chromosome segregation, it became clear that SMCs are also involved in mitotic chromosome formation, interphase chromosome organization, transcriptional regulation, and DNA repair. Here we highlight the most prominent functions of SMC proteins through the stages of the cell cycle.

1.3.1 Activity of SMCs during cell division

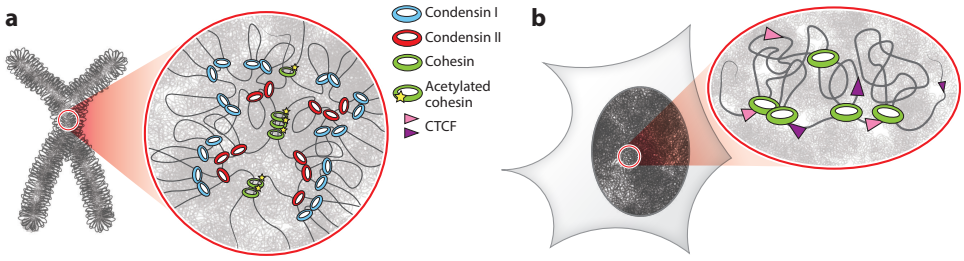


Figure 1.4 | SMC complexes are active throughout the cell cycle.
(**a**) Mitotic chromosomes in vertebrates are formed through looping the genome by condensin I and II, whereas cohesin is responsible for sister chromatid cohesion by topologically encircling the two sisters. Acetylation of the cohesin subunit Smc3 by the Eco1 acetyltransferase (denoted by a star) prevents loss of cohesin at this stage by preventing the unloader WAPL (Wings apart-like protein homolog) from acting on cohesin [105–108]. (**b**) In interphase, cohesin is the dominant SMC, as it extrudes DNA loops along the genome. Blocking of cohesin by convergently oriented CTCF (CCCTC-binding factor) sites (differently colored triangles denote CTCF in different orientations bound to DNA) results in TAD (Topologically Associated Domain) formation.

As cells replicate and divide, two of the most well-known SMCs work in symbiosis for faithful chromosome segregation: DNA compaction is driven by loop extrusion via condensin, while sister chromatid cohesion is established by cohesin. While the formation of DNA loops was suggested to underlie mitotic chromosome condensation early on [67, 109, 110], a novel spur of interest in the loop-extrusion hypothesis arose when it appeared that the elongated mitotic chromosome structure and segregation of the newly replicated chromosomes could be elegantly recapitulated with simulations [111–113]. Merely assuming a loop-extrusion mechanism, for which condensin was a strong candidate to be the driving protein, was sufficient to generate self-organized, compacted chromosomes. These simulations, combined with Hi-C chromosome-capture analyses, further explained the separate roles of two types of condensin present in vertebrates: Condensin II is nuclear and binds chromatin in prophase, whereas condensin I is cytoplasmic and thus can act on chromatin only after nuclear envelope breakdown [114, 115]. It was deduced that condensin II first extrudes large DNA loops ~400 kb long, whereas condensin I extrudes ~80-kb long nested loops within these 400-kb loops [116, 117] (Figure 1.4a). Notably, the estimates from the Hi-C and quantitative microscopy methods aligned remarkably well with very early estimates of loop sizes made by Riggs and Nasmyth [64, 118]. In contrast to eukaryotes, sister chromatid cohesion is absent in prokaryotes. In prokaryotes, the resolution and segregation of sister chromatids are coupled to replication. Bacterial SMCs load onto chromosomes at parS centromeric sequences adjacent to the replication origin, with the help of the

chromosome-partitioning protein ParB [119]. Hi-C analysis of *B. subtilis* suggested that SMCs translocate and perform loop extrusion of the entire megabase-sized chromosome until the terminus of replication is reached [120, 121]. This causes the two arms of the circular chromosome to become juxtaposed. SMC proteins are required for segregation of replication origins, and the chromosomes are segregated by the ParB-ParA system and entropic forces [122, 123]. In bacteria without the ParABS system, such as *E. coli*, the SMC-like MukBEF complex loads nonspecifically on the genome [124, 125]. By increasing chromosome occupancy of MukBEF and imaging the genome, it was observed that MukBEF complexes form a chromosomal axial core from which DNA loops emanate, reminiscent of mitotic chromosome condensation in eukaryotes [126]. Although it is difficult to come up with an alternative model for how SMC complexes organize bacterial chromosomes based on the *in vivo* evidence, a direct proof of loop extrusion by the bacterial complexes at the single-molecule level has been lacking to date, although multiple groups have tried to demonstrate this. The second important aspect of chromosome segregation in division, sister chromatid cohesion, depends critically on cohesin. Sister chromatid cohesion is established during S phase, concomitant with DNA replication [118]. Cohesin can be loaded onto the DNA via two independent pathways: Cohesin complexes can load de novo during or right after replication or existing chromatin-associated cohesins can be converted into cohesive ones [127–130]. Cohesin associates with both sister chromatids. The prevailing model is that a single ring-shaped cohesin complex [131] encircles one double-stranded (ds) DNA molecule from each sister chromatid [132] (Figure 1.4a) (for more focused reviews of cohesin's role in sister chromatid cohesion, see [133, 134]). While mechanistic details of the process are still lacking, *in vitro* studies showed that cohesin, when topologically bound to a dsDNA molecule, is capable of capturing and topologically embracing another ssDNA, but not dsDNA, molecule in the presence of Scc2 and ATP [135]. Such a scenario is reminiscent of replication, during which cohesin could topologically bind to the replicated leading strand and then capture the lagging strand before its conversion to dsDNA [136]. Cohesin and condensin can also stably bridge two dsDNA molecules in an ATP-dependent manner [137, 138]. Recently, a cohesin mutant was identified that is defective in sister chromatid cohesion but can still perform DNA loop extrusion [139]. This suggests that these functions can be genetically separated. The authors raise the intriguing possibility that CTCF may convert loop-extruding cohesin into cohesive cohesin, even though the mechanism remains as yet unknown [140]. CTCF and human minichromosome maintenance 3 (MCM3) (a part of the replisome) contain the same peptide sequence that is predicted to interact with a conserved binding site of cohesin, formed by its kleisin and STAG1/2 subunit [141]. It is therefore conceivable that such a proposed conversion can also occur when extruding cohesin complexes encounter the replisome. The final conversion from a topologically bound cohesin to a cohesive cohesin can then occur by catching the newly replicated lagging strand as proposed by Murayama et al. [135], by interacting with other replisome components [127], or at converging replication sites [142]. Cohesive cohesin must be protected from DNA release to prevent premature sister chromatid segregation. Acetylation of SMC3 by the acetyltransferase Eco1 in yeast (Esco1/Esco2 in humans [143–145]) protects cohesin from release by WAPL and PDS5 [106–108] (Figure 1.4a). SMC3 acetylation is likely coupled to cohesion establishment during replication, since Eco1 associates with the replication machinery [105]. It is thus conceivable that the topological encircling of one

or two dsDNA molecules occurs *in vivo* but is readily resolved by WAPL when cohesin is not acetylated. Also, co-entrapment of two dsDNA molecules has been observed by condensin [138, 146]. A bridging activity of two DNA molecules has also been suggested by experiments demonstrating that a loop extrusion-independent mechanism also contributes to the shaping of chromosomes [147].

1.3.2 Activity of SMCs during interphase

While condensin and cohesin join forces in forming mitotic chromosomes during cell division, their roles change during interphase; in this phase, chromosome organization is mainly governed by DNA loop extrusion by cohesin, while condensin plays only a modest role. First, we address the formation of topologically associated domains (TADs) and their implications for transcriptional regulation. Studies of the interphase organization of chromosomes were significantly accelerated when chromosome conformation capture (3C) methods advanced to a resolution of <100 kb. Self-interaction maps of chromosomes revealed regions, called TADs, of higher contact probabilities compared to their surroundings [39, 148]. These TADs span several hundred kilobase pairs and are often accompanied by strong chromosomal interactions between the loci demarcating TAD boundaries [149]. A hallmark of most TADs is that their boundaries are bound by CTCF and cohesin [39, 41, 150, 151], and the large majority of TADs disappear when either of these two components is deleted [33, 34, 152–154]. The loop-extrusion hypothesis, originally formulated to explain the condensation of mitotic chromosomes [111–113], proved to be successful in explaining TAD formation as well [42, 97, 155, 156]: Cohesin was suggested to bind and extrude DNA loops until it encounters a CTCF molecule, which acts as a boundary element (Figure 1.4b). This hypothesis could also explain two major observations in cells that are depleted of WAPL, a factor that releases cohesin from chromatin. Effectively increasing the residence time of cohesin on chromatin upon WAPL depletion led to longer loops in Hi-C experiments [34, 152] and dramatic compaction of chromatin and redistribution of cohesin into axial chromosome regions, termed vermicelli [157]. In chapter 7, we demonstrate on the single-molecule level *in vitro* that CTCF is indeed sufficient to stall loop-extrudin g cohesin, albeit in a DNA tension-dependent manner [35]. DNA loop extrusion by cohesin during interphase, and the associated formation of TADs, likely plays a fundamental role since roughly 2/3 of all TAD boundaries are invariant among cell types and between species [39, 158]. This high degree of evolutionary conservation is poorly understood, and recent single-cell analyses reveal the highly dynamic and cell-to-cell variant nature of loops and TADs [40, 159–162]. Yet, TAD boundaries may constitute a structural framework that enables or disfavors interactions between cis-regulatory sequences, which is corroborated by the fact that most enhancer-promoter interactions occur within a TAD [163, 164]. Even though the contact probability within TADs is only two- to fivefold higher than between TADs [165], computational studies found an up to twofold insulation and three- to fivefold facilitation of enhancer-promoter contacts, depending on the relative positioning of regulatory elements (Figure 1.5a) [165–167].

Rearrangements of TADs and the accompanying deregulation of genes may thus exert an evolutionary pressure on keeping TAD boundaries intact [168], providing an intriguing hypothesis to explain the strong evolutionary conservation of TADs. This hypothesis predicts that the structural integrity of the genome should have a broad effect on the

expression levels of genes. However, this is not ubiquitously the case, since the effect of cohesin and CTCF loss on transcription genome-wide is modest [153, 154] and occurs mainly after a few days of depletion [33], with some genes not being affected at all by structural alterations (for examples, see [169–172]). These contradictory views may be reconciled by the notion that TAD formation by DNA loop extrusion is solely a means to provide a rough structural framework in which specificity and maintenance is governed by transcription factors and the transcription machinery, which, in turn, may interact with DNA loop extrusion [173–175] (see 1.6). Active loop extrusion and blocking thereof constitutes a mechanism that can be regulated by the cell, e.g. by CTCF (un)binding or the establishment or dissolution of loops with the help of NIPBL (the human analogue of the cohesin Scc2 subunit) or WAPL. The specificity and selection of interaction partners from a pool of regulatory elements within a TAD may be tuned by the binding of transcription factors and the formation of transcriptional condensates, which may also ensure maintenance of interactions after establishment by loop extrusion [176–178] (Figure 1.5a). However, it is important to note that DNA loop extrusion is likely not a strict requisite for the regulation of a variety of elements such as stable housekeeping genes, promoters without partnering enhancers, promoter-enhancer pairs with exceptionally high affinity (such as Sox2 and its distal enhancers [179]), and promoters that can be readily looped over short distances without the additional help of the active looping machinery [180]. Instead, promoter-enhancer pairs separated by long genomic distances and genes with quick expression changes, e.g., in response to external stimuli or during development, might be more affected by the presence of cohesin and CTCF [181–186]. How does TAD formation relate to replication timing? Mammalian genomes replicate from multiple origins that fire in a distinct temporal order, called the replication-timing program. Domains spanning 400-800 kb can change their replication timing during development and are called replication domains (RDs) [187]. Remarkably, the boundaries of RDs have a near one-to-one correspondence to TAD boundaries (Figure 1.5b), and replication timing changes happen in genomic regions spanning one or multiple TADs [188]. Do RDs then depend on both cohesin and CTCF, just like TADs? Early-firing replication origins colocalize with loop anchors at TAD boundaries, demarcated by cohesin and CTCF binding [189] but not by cohesin alone [190]. Ablation of cohesin causes replication origins to become more diffuse, while ablation of WAPL, and thus a gain of longer loops compared to the wild type, localizes replication origins to TAD boundaries (Figure 1.5b). A molecular mechanism linking TADs and RDs is not yet resolved. Nevertheless, a prominent candidate involved in both processes is the minichromosome maintenance protein complex MCM2-7, a key enzyme of the replication machinery. Two recent studies have shown that the replication complex, in particular MCM2-7, is a semipermeable barrier to cohesin loop extrusion [191, 192]. This suggests that cohesin can be positioned by replication factors such as MCM2-7 and/or can be recruited de novo in order to perform both DNA loop extrusion and sister chromatid cohesion [130]. If responsible for the localization of replication origins, MCM2-7 should be preferentially loaded at or relocated to TAD boundaries in a fashion that depends on cohesin and CTCF [190] (Figure 1.5b). However, no correlation between MCM localization and TAD boundaries is known to date, and a molecular mechanism for this remains to be shown. Alternatively, other factors important for replication initiation might be localized by active loop extrusion to specifically activate MCM complexes at TAD bound-

aries (Figure 1.5c). In conclusion, while there is a striking correlation between replication timing and TAD boundaries, as well as documented interactions between cohesin and the prereplication complex, the molecular mechanism that couples DNA loop extrusion and replication timing await further investigation. Loop extrusion has also been implicated in DNA repair. DNA double-strand breaks (DSBs) threaten the integrity of the genome; they are highly deleterious and, if unrepaired, can cause genome instability. SMC proteins are essential components of the DNA repair pathways, yet details on their contribution remain largely unresolved. The high-fidelity homologous recombination pathway depends on cohesin's ability to mediate cohesion between sister chromatids [193]. Both SMC5/6 [194] and cohesin are recruited to DSBs, possibly even together [195]. In the case of cohesin, the recruitment depends on the presence of Scc2, indicating de novo loading of cohesin rather than conversion of previously bound complexes [196–198]. Both cohesin and SMC5/6 have been shown to bind ssDNA [199, 200], which may be essential for sensing ssDNA overhangs of DSBs in the repair through homologous recombination (Figure 1.5d). In a loose analogy to the concerted action of DNA loop extrusion and sister chromatid cohesion during mitosis, these two mechanisms may also be concomitantly at work during DNA repair [201]. One of the first signaling events following DSB is the phosphorylation of H2AX (in mammals) to form γH2AX [202]. Arnould et al. [196] observed that both cohesin and the ATM kinase remain bound at the DSB, yet the γH2AX mark was enriched up to 1 Mbp away from the DSB within a CTCF-bound TAD boundary. Given the observation that components of the DNA repair machinery block cohesin progression *in vivo* [203], the authors proposed that cohesin is stalled at DSBs and extrudes loops only on one side, which also generates the stripe pattern that has been observed around CTCF sites (Figure 1.5d). Based on these observations, the authors argued that cohesin's one-sided loop-extruding activity is necessary to reel DNA close to the DSB, while the ATM kinase deposits the γH2AX mark on the extruded DNA. Mechanistic details and whether the ATM kinase can directly interact with cohesin (rather than spreading from nucleosome to nucleosome) remain unknown. It is, however, also conceivable that, at least in part, SMC5/6 takes over the loop-extruding function at DSBs. The finding that SMC5/6 can act as a symmetric loop extruder with two independent motors *via* SMC5/6 dimerization [204] indicates that it may potentially bridge the DSB, consistent with an equal spreading of γH2AX on both sides of the DSB. However, for now, the roles of DNA loop extrusion and cohesion by cohesin and SMC5/6 in DNA repair remain barely understood. Another important biological process where loop extrusion plays a crucial role is the assembly of antigen receptor genes by V(D)J recombination in B and T cells of the vertebrate immune system (for a more detailed review on this topic, see [205]) and the isoform diversity of clustered Protocadherin (Pcdh) proteins in neurons by tuning the processivity of cohesin *via* WAPL across the Pcdhα gene cluster [182].

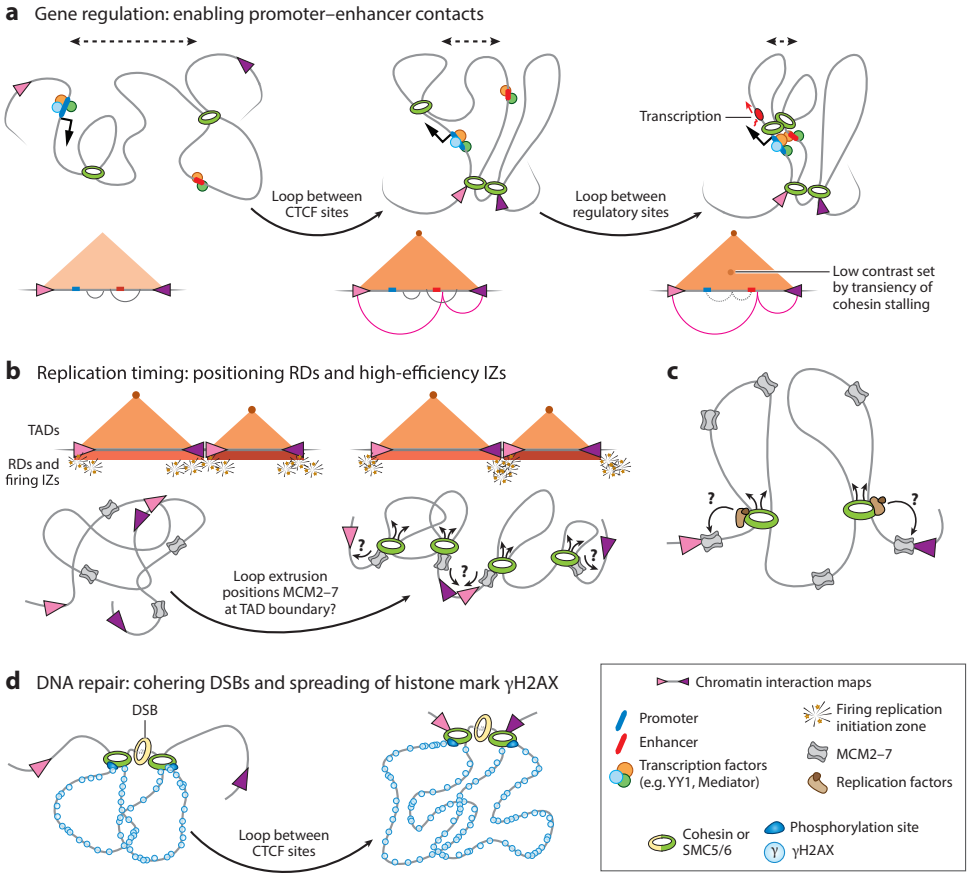


Figure 1.5 | DNA loop extrusion in gene regulation, replication, and DNA repair.
(**a**) An enhancer and promoter on an unconstrained genome segment may have a low contact probability due to the large intervening distance between the sites. DNA loop extrusion between CTCF (CCCTC-binding factor) sites located at the encompassing TAD (Topologically Associated Domain) can constrain the polymer and enhance contacts between regulatory sites. When additional loops are enriched between regulatory sites, even for a short time by transiently stalling loop extrusion, for example (cf. Figure 1.8), their contact probability is boosted. Due to the transiency and low occurrence, these interactions may remain rather hidden in chromatin interaction maps (triangle-shaped cartoons). (**b**) Replication domains (RDs) overlap with TADs. Without loop extrusion, initiation zones (IZs) appear diffuse, while ongoing loop extrusion localizes IZs to TAD boundaries. (**c**) Enrichment of cohesin at CTCF sites due to loop extrusion may bring other replication factors to MCM2-7 (Minichromosome maintenance complex 2-7) complexes residing at CTCF sites. (**d**) DNA double-strand breaks (DSBs) may be sensed and cohesed by cohesin and/or SMC5/6 to enable homologous recombination. Loop extrusion by cohesin or SMC5/6 may enable the spreading of the DSB response marker γH2AX (a phosphorylated variant of the H2A protein family) by reeling DNA through phosphorylation sites located at the DSB.

1.4 Biophysics of DNA loop extrusion by SMCs

1.4.1 Single-Molecule visualization of loop extrusion by SMC complexes

Despite extensive evidence for DNA loop extrusion by condensin and cohesin, the high complexity of the cellular environment, the existence of a plethora of interacting partners, and their dissimilarity to any other characterized DNA translocases posed considerable barriers to identifying the molecular functions of these complexes. Biochemical and biophysical *in vitro* analyses of the isolated complexes were therefore performed to elucidate the protein activities. In early years, inspired by the discovery that cohesin holds sister chromatids in mitotic chromosomes, as well as by the fact that the ATPase hydrolysis rate is very low, SMC complexes were often thought of as static chromatid linker molecules rather than active loop extruders. The first ATP-dependent activity of purified condensin to be observed in *in vitro* biochemistry analysis was that 13S condensin in Xenopus laevis egg extracts introduced positive supercoils into a circular DNA in the presence of topoisomerase I [48, 206] (see also chapter 6). The possibility that condensin could actively compact DNA in an ATP-dependent fashion first emerged from single-molecule magnetic tweezers experiments with condensin I from Xenopus laevis eggs [207] and later from budding yeast [208]. Evidence that condensin is a motor protein was obtained using single-molecule imaging, which showed that individual fluorescently labeled budding yeast condensin complexes translocated unidirectionally along stretched and surface-anchored fluorescently labeled DNA in an ATP-dependent manner [209]. Soon thereafter, the first direct evidence that condensin is a DNA-loop-extruding motor was provided using another single-molecule visualization assay [98]. In this breakthrough experiment, a DNA molecule was tethered at its two ends in a loose configuration such that the tension within the molecule was low. Upon stretching the DNA by applying a sideways liquid flow, the DNA was visualized, and a fluorescently labeled condensin could be observed to rapidly reel DNA into a loop upon ATP hydrolysis [98]. These real-time visualization experiments unambiguously proved that a single condensin complex can extrude a large (tens of kilobase pairs) loop of DNA. Loop extrusion was found to occur in an asymmetric manner in which DNA was reeled into the loop from only one side, while the other side of the DNA was anchored at the safety belt of the HEAT II Ycg1-kleisin subunits of condensin [98, 210]. Condensin was found to be a very fast motor, with DNA loop extrusion occurring at speeds up to 1.5 kbp/s (mean 0.6 kbp/s), but also a weak motor, as a tension of less than 1 pN on the DNA that was reeled in was sufficient to stall the motor action. Unlike condensin, early *in vitro* single-molecule studies of cohesin focused on the ATP-independent diffusion of human [211] and Saccharomyces pombe cohesin [102] along DNA. These experiments showed that cohesin can topologically load onto DNA and diffuse along the strand, even passing over some DNA-bound proteins (EcoRI, dCas9, TetR), but not over transcribing RNA polymerase (RNAp) or CTCF [35]. Direct evidence that cohesin can drive DNA loop extrusion was provided using human cohesin with a similar assay previously employed for visualizing condensin-mediated loop extrusion [173, 212]. These experiments revealed that human cohesin requires NIPBL-Mau2, the HEAT I subunit of cohesin that acts as a cohesin loader onto chromatin, not only for initiating but also for maintaining ongoing loop extrusion [173]. Based on the structural similarities among SMC complexes, it seems plausible that

loop extrusion activity might not be a function limited to condensin and cohesin but might be a universal feature shared by all SMC complexes. Indeed, *in vitro* magnetic tweezers experiments showed that SMC5/6 compacts DNA in an ATP-dependent manner under low tension [213], resembling ATP-dependent compaction by condensin [207, 208]. Recently, DNA loop extrusion induced by the budding yeast SMC5/6 complex was observed [204]. The study showed that Nse5/6, the subunits specific only to the SMC5/6 complex, negatively regulates DNA loop extrusion by reducing loop initiation, findings that were corroborated in Figure S5.2 of chapter 5 of this thesis. Real-time imaging of the DNA loop extrusion process by purified SMC complexes has proven to be a powerful approach as it not only allows the direct visualization of the extrusion process but also provides mechanistic insight and enables quantitative estimations of the properties of the loop extrusion at the single-molecule level. In 1.4.2-1.4.6, we discuss several important characteristics of loop extrusion that were revealed by *in vitro* single-molecule experiments.

1.4.2 Loop extrusion rate

Ganji et al. [98] observed that yeast condensin reels DNA into a loop at speeds of up to 1.5 kbp/s, with an average of 0.6 kbp/s. Notably, this is a very fast rate that exceeds the speed of other DNA translocases (such as helicase or polymerases) by orders of magnitude, indicating that these SMC complexes constitute a unique new class of motor proteins. Similar values were also found for different SMC complexes, e.g., for human cohesin (0.5-1 kbp/s [173, 212]), for cohesin from interphase Xenopus egg extract (2 kbp/s [214]), for fission yeast cohesin (≈1 kbp/s [215]), for human condensin I and II (≈1 kbp/s and 0.5 kbp/s, respectively [216]), for fission condensin (≈0.9 kbp/s [138]), and for yeast SMC5/6 (1 kbp/s [204], see also Figure S5.2 of chapter 5). Given the fact that all of these purified SMC complexes exhibit similarly low ATP hydrolysis rates, i.e., ≈ 2 ATP molecules per second, the length of DNA that the SMC complexes reel into loop per ATP hydrolysis cycle (i.e., the step size) is likely to be conserved across SMC complexes. In the case of budding yeast condensin [217], human cohesin (see chapter 7 [35]), and budding yeast SMC5/6 (chapter 6, [50]), single step sizes were recently measured to be ~20-45 nm, that is, up to 200 bp per step. Furthermore, it has not escaped our notice that the step size roughly equals the size of the SMC complex, which immediately suggests that a complex-wide conformational change is involved in the stepping. The rate of extrusion was also experimentally determined *in vivo* using Hi-C analyses. The rate of chromosome arm juxtaposition by bacterial SMC complexes in *B. subtilis* [120] and *C. crescentus* [218] was found to be 0.3-0.8 kbp/s, corroborating the range observed in single-molecule experiments. In eukaryotes, however, from matching Hi-C experiments to simulations, the DNA extrusion rates of cohesin and condensin were estimated to be lower, on the order of 0.1 kb/s [219], which may suggest that other factors, such as cellular crowding, roadblocks (e.g., DNA-bound CTCF and other proteins, transcription machinery, etc.) play a role in slowing down these motor proteins.

1.4.3 Stalling force

Another important property shared by all loop-extruding SMCs is that they stall if the tension on the DNA that is reeled in becomes too large. The growth of a loop thus saturates and reaches a plateau when extrusion is performed on double-tethered DNA substrates as more and more DNA is transferred to the loop and tension within the non-extruded DNA

inevitably increases [98]. The stalling force of SMC complexes is low, sub-piconewton, which is much lower than that of most other DNA-processing enzymes [220]. The estimated stalling forces for different SMC complexes exhibit similar values: 0.4-1 pN for condensin (21, 134), 0.1-0.8 pN for cohesin [35, 212, 214] (see also chapter 7), and 0.5 pN for SMC5/6 [204].

1.4.4 Directionality of loop extrusion

SMC complexes can reel DNA into the loop from one side or from both sides. Ganji et al. [98] first made the surprising observation that DNA loop extrusion by yeast condensin proceeds in a one-sided manner, i.e., asymmetrically. This posed a major theoretical challenge: Could one-sided extrusion perform the large variety of tasks that were thus far credited to two-sided extrusion [112, 116]? Experiments with a condensin mutant showed that robust one-sided extrusion depends on Ycg1-Brn1, which firmly anchors the condensin onto DNA [98, 210]. Recent work showed that deletion of Ycg1 in *Chaetomium thermophilum* turned condensin from a one-sided to a two-sided motor [221]. The physiological reason behind the strong asymmetry of condensin-driven loop extrusion in yeast cells is so far not well understood. Since human condensin I and II seem to be able to extrude loops in both a symmetric and asymmetric manner [216], asymmetric extrusion appears to be a weakly conserved feature among eukaryotic condensin complexes. Unlike condensin, cohesin and SMC5/6 were observed to extrude loops in a two-sided manner. chapter 5 of this thesis demonstrates that two-sided extrusion by cohesin, at the microscopic level, does not occur simultaneously but rather in an alternative fashion in which the direction of extrusion is switched over time (chapter 5 and [222]). In contrast, SMC5/6 can extrude loops also as dimers [204], and these dimers may extrude DNA symmetrically by the two motor complexes independently translocating along DNA, thus reeling in DNA simultaneously. Importantly, analysis of DNA loop formation in bacteria suggested that bacterial SMC complexes are also dimeric extruders [120].

1.4.5 Topology of SMCs and DNA

Since the geometry of SMC proteins is a ring, the specific topology of how SMC proteins interact with DNA has long been actively investigated. *In vivo* chemical cross-linking [132] and *in vitro* biochemical studies [132, 221, 223] suggested that SMC complexes can topologically load onto DNA, or in other words, a DNA molecule can be encircled by the trimeric SMC-kleisin ring. It was suggested that the entry of the DNA would be achieved by opening the hinge, while the exit gate would be the SMC-kleisin interface [224, 225]. A recent study from Shaltiel et al. [221], however, showed that DNA remained bound when incubated with a fusion complex made by internal covalent linkage of the trimeric ring of condensin, and the DNA would reside within the two chambers created by the interface of the kleisin and the two HEAT-repeat proteins associated with kleisins (HAWK) domains, thus remaining as a pseudotopological configuration within the trimeric ring. It thus appears possible that the traditional immunoprecipitation method, in combination with a high salt wash, does not discriminate between pseudotopological and topological loading of SMCs onto DNA, casting some doubt on early claims of a strictly topological entrapment of DNA by SMC rings. The link between topological loading of SMCs and their loop extrusion activity has long remained unclear. Recent work from Davidson et

al. [173] and Shaltiel et al. [221] provided clues in this direction. They used cohesin and condensin in which the trimeric core ring complexes were covalently linked and showed that these complexes were, remarkably, still able to extrude DNA loops, thereby excluding the possibility that topological loading is required for loop extrusion. Recent experiments from Pradhan et al. (see chapter 3 and [226]) furthermore showed that the covalently linked cohesin can also accommodate DNA-bound obstacles that are much larger than the ring size into the extruded loop, thus challenging topological and pseudotopological mechanisms of loop extrusion and instead indicating a nontopological mechanism.

1.4.6 Conformational states of SMCs

Since the loop extrusion function of SMC complexes was discovered, a major open question has been to understand how these proteins extrude DNA loops at the mechanistic level. To this end, it is critical to identify the conformational states of SMC complexes and how these conformations link to one another during loop extrusion. Recently, many high-resolution cryo-electron microscopy (cryo-EM) structures of SMCs became available in different states of the ATP-hydrolysis cycle, and the dynamics of the conformational transitions of SMCs were captured with high-speed atomic force microscopy (AFM) and fluorescence resonance energy transfer (FRET), all providing clues for resolving the mechanism of loop extrusion by SMCs (for reviews on SMC structures and conformations, see [227, 228]). Recent structural analyses of SMCs captured in the ATP-free (or apo) and ATP/DNA-bound state reveal striking structural similarities across different SMC complexes. In the apo state, SMCs are in a folded conformation in which the juxtaposed SMC coiled-coil arms are folded toward the ATPase heads [138, 200, 228–230], with which the HEAT I [200, 231] and Kite [104, 232] domains are predominantly associated. In the case of condensin, another apo conformation, a so-called apo-bridged state, was reported in which two SMC ATPase heads are separated by a large distance and bridged by the HEAT I domain [233]. Upon ATP binding, the two SMC ATPase heads do engage by sandwiching two ATP molecules, which appears to induce a conformational change that allows stable clamping of DNA within the complex [200, 221, 234–236]. The SMC head engagement leads to a partial opening of the SMC arms, giving it a V-shape in which a DNA strand is bound above the engaged heads and below the HEAT I or Kite domains. In this clamped state, the DNA is sandwiched by the kleisin from above and the engaged heads from below. Although a juxtaposed folded conformation of SMCs in the apo state was commonly reported in cryo-EM data, conformational changes of condensin [237] and cohesin [238] observed by AFM and FRET reveal that SMC arms can be well separated and exhibit dynamic transitions between open and closed states. For condensin, dynamic and flexible SMC arms were reported [208, 238, 239], and the SMC complex appeared to toggle between an open state and a state in which the hinge collapsed onto the globular domain. Similarly, for cohesin, an alignment of coiled-coil arms and their bending toward the Smc3 head was reported, similar to the cryo-EM data [200, 238]. Upon ATP binding, condensin and cohesin seem to undergo different conformational transitions. While binding of ATP led predominantly to a folded or collapsed conformation of condensin [237], the ATP-binding-induced head engagement was reported to change cohesin from a folded conformation to an open conformation [238].

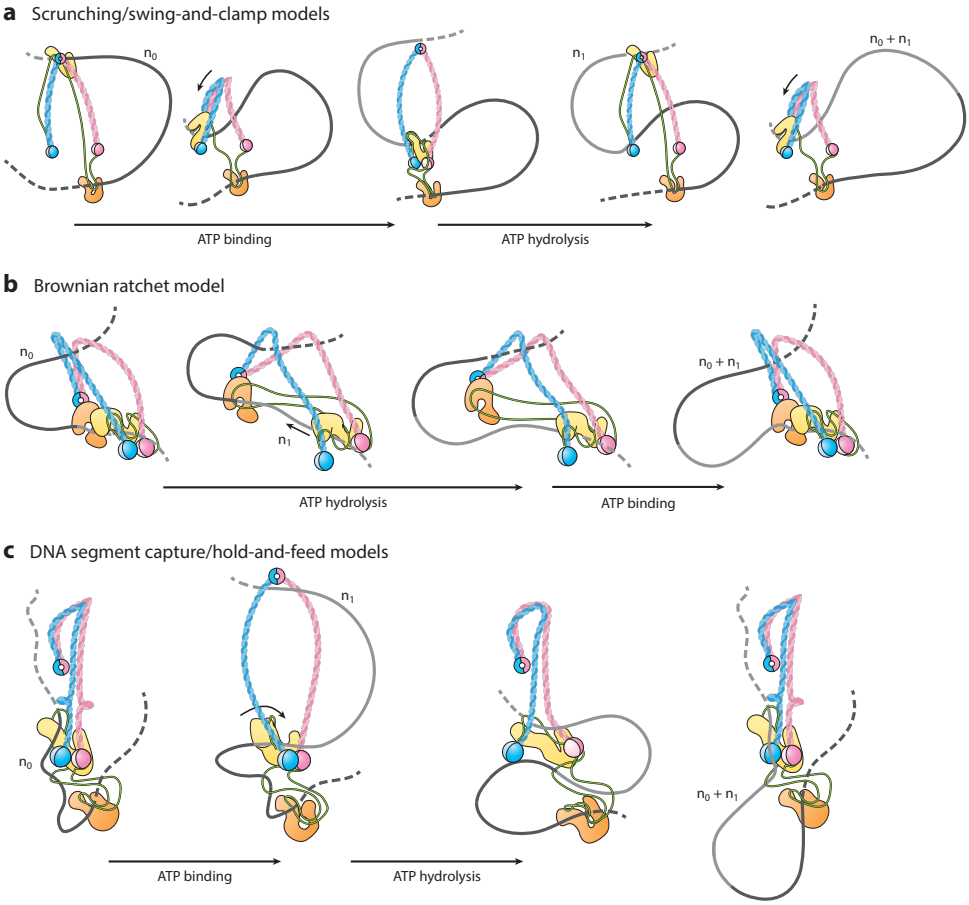


Figure 1.6 | Loop extrusion models.
(**a**) Scrunching and swing-and-clamp models. The SMC complex binds to DNA at the HEAT I (Huntingtin, elongation factor 3 (EF3), protein phosphatase 2A (PP2A), and the yeast kinase TOR1; located near the hinge) and HEAT II subunits. Folding of the SMCs and a subsequent ATP-head-engaged open conformation allow the transfer of a DNA segment between the hinge and the heads, thereby enlarging a loop. (**b**) Brownian ratchet model. Upon ATP binding, SMC complexes adopt a gripping state in which DNA is pseudotopologically loaded. A swinging motion of the HEAT I, driven by Brownian motion, turns the DNA bend into a loop. (**c**) Segment capture and hold-and-feed models. A DNA loop is initially captured by the interfaces of kleisin and the HEAT domains. ATP binding results in a power-stroke tilting motion of the HEAT I/kleisin chamber, allowing the complex to capture new DNA into the SMC ring. Upon ATP hydrolysis the opening of ATPase heads and the folding/zipping up action of the SMC arms allow the transfer of the new DNA into the previously established loop. The DNA trajectory is adopted from Oldenkamp & Rowland [227], which is compatible with all crosslinks tested by Shaltiel et al. [221]

1

1.5 Models of DNA loop extrusion

While the field has not converged on a final understanding of the molecular gymnastics that underlie loop extrusion by SMC proteins, progress has been made toward establishing a model for DNA loop extrusion at the molecular level. A variety of different models have been proposed that all aim to account for the available data, including the experimental finding that steps in the DNA extrusion process involve very large conformational changes—of the order of the size of the SMC ring itself [35, 50, 217]. Here we discuss recent models, which we categorize into three classes according to their shared characteristics (Figure 1.6).

1.5.1 The Scrunching/Swing-and-Clamp Model

Based on AFM observations of the large conformational transitions of condensin complexes between an extended open O shape to a collapsed butterfly B shape, triggered by the binding of ATP [237], a scrunching model was suggested for condensin-mediated loop extrusion. In this model, DNA binding occurs at the safety belt of HEAT II Ycg1-kleisin [210, 221], as well as at a second binding site at the hinge. Upon ATP binding, the hinge-bound DNA is transferred to the ATPase head domains by the conformational change from the O to the B shape and bound there at a third DNA-binding site. Indeed, recent magnetic tweezers data confirmed that it is the ATP binding, rather than hydrolysis, that is associated with the actual step of drawing DNA into the extruded loop [217]. Upon ATP hydrolysis, condensin returns to the O shape and stochastically captures a new DNA segment for the next step in the cycle. Recent work from Bauer et al. [238] proposed a more detailed model of loop extrusion driven by human cohesin, the swing-and-clamp model, which shares similarities with the scrunching model. In this model, the DNA transfer is controlled by the DNA-loader HEAT I (NIPBL), which in the apo state binds at the DNA-bound hinge and upon ATP binding dissociates from the hinge and associates with the heads while carrying DNA. The transfer of DNA-bound NIPBL from hinge to heads is achieved by the hinge swinging toward the SMC3 head in the apo state, whereupon the ATPase heads dimerize and DNA is clamped in between NIPBL and the engaged heads. The subsequent ATP hydrolysis leads to disengagement of the heads, disassembly of the clamp, and unfolding of the SMC elbows, thus bringing the hinge back to the original position and getting the motor ready for the next cycle. While the swing-and-clamp model may explain DNA translocation, it is not fully clear thus far how it leads to loop extrusion, which may involve a coordinated action with STAG1/2. Notably, the swing-and-clamp mechanism does not agree with data from yeast condensin, in which ATP binding was found to trigger the conformational change from the open to the collapsed shape [237]. For both the scrunching model and the swing-and-clamp model, it remains unclear how a mechanism that involves random grabbing of nearby DNA by the hinge provides a temporarily sustained directionality to the loop extrusion. However, computational analyses suggest that directional loop extrusion might also be a consequence of a geometric constraint governing DNA binding at the SMC hinge [240].

1.5.2 The Brownian Ratchet Model

This model is based on the recently resolved DNA-gripping state from fission yeast cohesin [215, 235]. This model suggests that a Brownian ratchet is built from the Scc3-hinge and Scc2-head modules, which are juxtaposed during the ATP-bound state but allow

unidirectional DNA diffusion upon ATP hydrolysis. According to this Brownian ratchet model, DNA arrives in the gripping state without passing the kleisin N-terminal gate, which prevents DNA passage upon the subsequent ATP-hydrolysis-induced ATPase head disengagement. The stochastic Brownian swinging motion of the Scc3-hinge module turns the DNA into a loop and is responsible for the subsequent loop growth. The low DNA affinity of the Scc3-hinge module results in DNA release, and the Scc3-hinge module subsequently returns to form a new DNA-gripping state upon ATP binding. This model, however, seems to be incompatible with some experimental observations, e.g., Scc2's association with the hinge in the apo state [200, 238] and the deletion of Ycg1 (analogous to Scc2) in *C. thermophilum* condensin, which did not impair the loop extrusion activity of condensin [221]. Furthermore, the Brownian ratchet model can explain only the passage of roadblocks that are smaller than the SMC lumen, while the passing of large roadblocks can be realized only by the opening of the SMC ring [215]. However, it was recently shown that a covalently closed SMC ring is able to bypass roadblocks 200 nm in diameter (see chapter 3 of this thesis and [226]).

1.5.3 The DNA Segment Capture/Hold-and-Feed Models

The DNA segment capture model [241, 242] was inspired by features of prokaryotic SMC complexes in which the conformation of SMC-ScpAB complexes transitions between the ATP head-disengaged juxtaposed and head-engaged opened states [233]. The model assumes that, in the apo state, a small loop of DNA is captured in the lower compartment created in between the juxtaposed SMC arms and the kleisin. Upon ATP binding, the opening of the SMC coiled-coil arms enables another loop to be captured in the upper compartment created within the head-dimerized SMC arms. Subsequent ATP hydrolysis and the associated opening of the ATPase heads merges the two compartments, followed by the zipping-up action of the SMC arms, which allows the transfer of the DNA loop within the upper compartment to the lower compartment. When iterated, this reaction extrudes a loop. A recent modification of this model [242] combines the loop capture in the upper compartment with an ATP-induced power stroke by the folding kleisin that feeds a DNA loop into the SMC ring. This is inspired by recent work by Shaltiel et al. [221], who proposed a hold-and-feed mechanism, which exhibits large similarities to the DNA segment capture model. This model specifies that a small loop is captured in the apo state by the kleisin-Ycs4 and kleisin-Ycg1 chambers. ATP binding-induced head engagement enables the power-stroke tilting motion of the Ycs4-kleisin chamber and swings the DNA within the chambers to allow the capture of new DNA into the SMC ring. Subsequent ATP hydrolysis and head disengagement changes the complex to the apo-bridged conformation [243], allowing the merging of the new loop into the lower compartment. At the start of the next cycle, the Ycs4-kleisin chamber is tilted back to the apo conformation together with a new segment of DNA to enlarge the loop. Although many aspects of this model are consistent with experimental observations and simulations [242], some predictions of this model regarding obstacle bypass remain inconsistent with experimental observations [226]. It appears that each of the models so far can explain a subset of the experimental data but not all of them (Figure 1.6c).

1

1.5.4 The Reel & Seal model

Oldenkamp and Rowland [227] recently provided a nontopological variant of the hold-and-feed model that is consistent with all cross-linking data by Shaltiel et al. [221], as well as obstacle bypass experiments (chapter 3 and chapter 4 of this thesis as well as [226, 244]). It implies that DNA is held temporarily in a pseudotopological manner by folding its kleisin subunit around DNA, using the HEAT-A and HEAT-B subunits to demarcate temporary pseudotopological 'chambers'. The extruded loop is not topologically inserted into the Smc-kleisin ring (Figure 1.7). DNA is initially held in a chamber made by kleisin, HEAT-A and binding of HEAT-A to the ATPase heads (Figure 1.7a, step 1) [221]. ATP binding engages the ATPase heads which forms an ATP-stimulated DNA binding site [245]. Presumably assisted by a conformational change of HEAT-A [246], DNA gets clamped onto the ATPase heads, triggering an opening of the Smc arms (Figure 1.7a, step 2) [238]. By moving the hinge away from the ATPase heads [238] and due to the strong DNA binding onto the ATPase heads, a small loop is captured inside the Smc lumen and the hinge contacts a downstream segment of DNA (Figure 1.7a, step 3). Note that the clamping of the DNA onto the heads provides a directionality for DNA loop extrusion. Subsequent ATP hydrolysis resolves the DNA-clamped state [238] and forces HEAT-A to unbind from the ATPase heads. The Smc arms bend over again [138, 200, 238] and force the loop through the Smc lumen and the now disengaged Smc ATPase heads which causes the extruded loop to grow (Figure 1.7a, step b). By rebinding of HEAT-A, a new DNA segment is captured into the pseudotopological chamber made by HEAT-A, kleisin and an Smc subunit, positioned for clamping at the next ATP binding and head engagement event (Figure 1.7a, step 5). By opening and re-sealing a latch formed by HEAT-A and kleisin (Figure 1.7a, step 3-5), a DNA-bound object of any size may pass the SMC nontopologically (Figure 1.7b). Because DNA acts as a rather stiff polymer with persistence length of 50 nm on the length scales of the SMC (~ 35 nm, see chapter 3, Figure S3.2e,f,h), a roadblock which cannot pass through the Smc lumen will be blocked either close to the hinge- or the head-proximal DNA binding site in order to minimize the DNA's bending energy (Figure 1.7b, step 3). One cycle of ATP hydrolysis and subsequent DNA binding relocates the roadblock proximal to the ATPase heads (Figure 1.7b, step4-5) while the next ATP hydrolysis step feeds the roadblock into the extruded loop (Figure 1.7b, step 6).

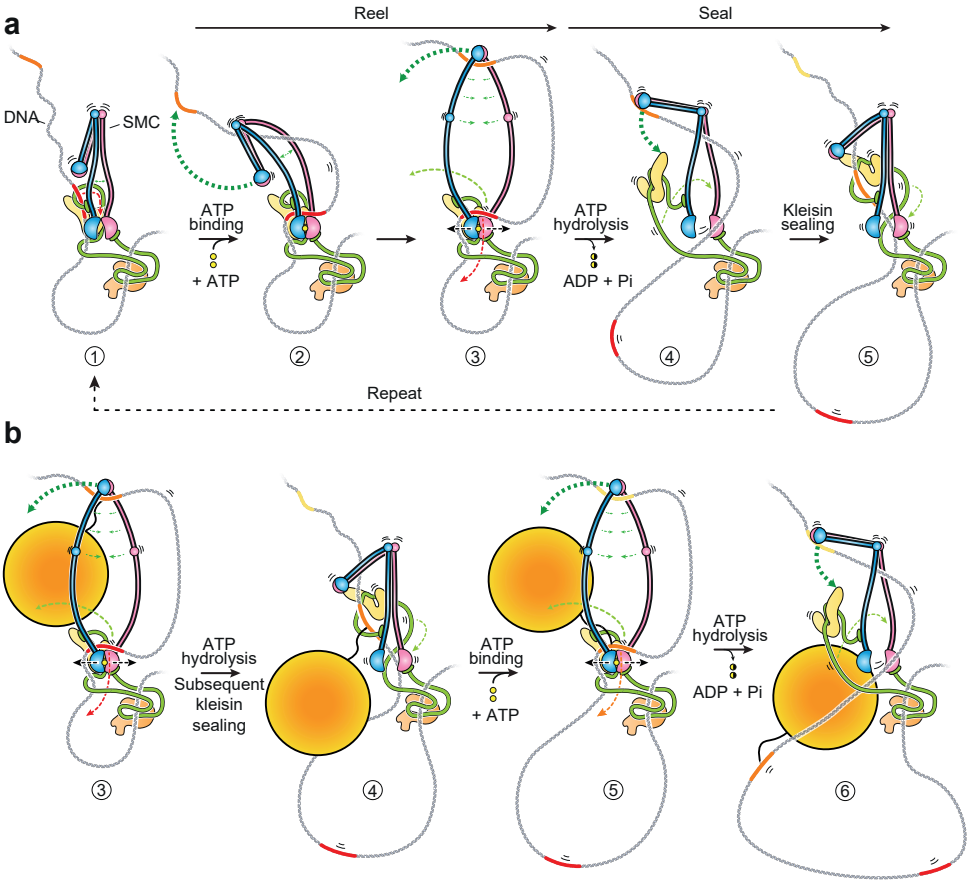


Figure 1.7 | Real & Seal Model.
(**a**) A potential pathway to DNA loop extrusion by the Real & Seal model. DNA is held within a chamber formed by HEAT-A, kleisin and one Smc subunit (step 1). ATP binding clamps DNA onto the ATPase heads, forcing a small DNA loop through the Smc lumen (step 2). Stretching and unfolding of the Smc arms allows the hinge-proximal DNA binding site to contact a downstream segment of DNA (step 3). Subsequent ATP hydrolysis disengages the ATPase heads, releasing HEAT-A and feeding DNA into the extruded loop (step 4). Re-sealing of the kleisin-HEAT-A latch poises the complex for the next loop extrusion cycle. (**b**) A roadblock will be located at the hinge-proximal DNA binding site of the SMC in order to minimize the bending energy acting on the DNA (step 3; step 1-2 are as shown in panel a). ATP hydrolysis relocates the roadblock close to the ATPase heads (step 4) while the subsequent ATP binding (step 5) and hydrolysis (step 6) feed the roadblock into the nontopologically extruded loop. Adapted from [247].

1

1.6 From naked DNA to chromatin

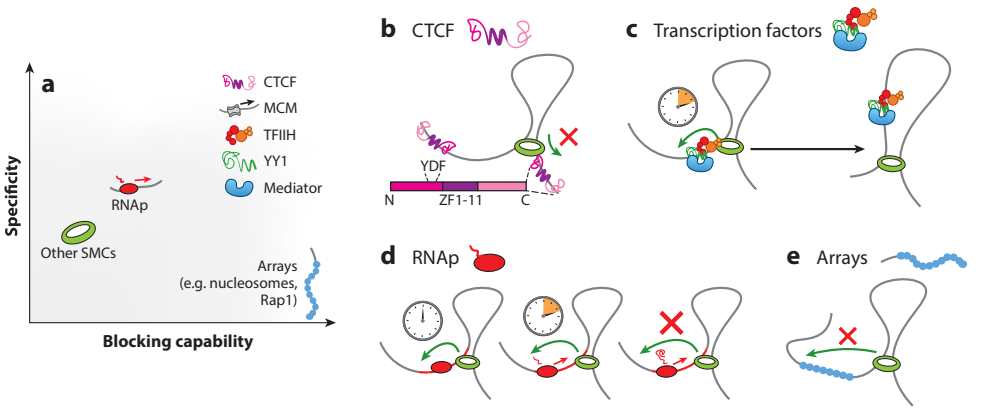


Figure 1.8 | DNA loop extrusion on chromatin.
(**a**) DNA-binding 'roadblocks' are characterized according to their specificity of interaction with SMCs and the expected capability to block loop extrusion. (**b**) CTCF (CCCTC-binding factor) interacts with cohesin primarily (though not exclusively) with its YxF motif located on the N-terminus. (**c**) Transcription factors such as TFIIH (Transcription factor II Human), YY1 (Ying-Yang 1 protein), and Mediator may temporarily stall loop extrusion, biasing interactions between regulatory sites. (**d**) While promoter-bound RNAp (RNA polymerase) does not constitute a strong roadblock to DNA loop extrusion, it may become progressively more effective in blocking loop extrusion upon transcription, as SMCs may interact with the RNA transcript. (**e**) A dense array of closely packed DNA-binding proteins constitutes a roadblock to loop extrusion.

The *in vitro* studies demonstrating DNA loop extrusion by condensin, cohesin, and SMC5/6 were performed on bare DNA [98, 138, 173, 204, 212, 214, 215]. While of crucial importance, a direct translation of the DNA loop extrusion mechanism to *in vivo* systems is not straightforward, as chromatin is decorated by a plethora of DNA-interacting proteins. While Hi-C data provide strong evidence for loop extrusion in vivo, it is not immediately clear how the various loop extruders interact with the DNA-binding proteins that putatively act as roadblocks for loop extrusion. In the remainder of this section, we discuss some frequently occurring DNA-binding proteins and sort them into two categories (Figure 1.8a): (i) roadblocks that specifically interact with the SMC through biochemical interactions [e.g. CTCF, MCM3, TFIIH, Yin Yang 1 (YY1), Mediator] and (ii) physical roadblocks (e.g. RNA and RNAp, nucleosomes, arrays of DNA binding proteins such as Rap1 [Repressor/activator protein 1], other SMCs). The best-studied and most prominent interaction partner of cohesin is CTCF [151, 248]. Loop extrusion by cohesin that is stalled by convergently oriented CTCF molecules [148, 155, 158, 249] gives rise to the appearance of TADs [39, 148]. An evolutionary conserved surface formed by STAG1/2 and kleisin specifically interacts with three N-terminally located amino acids, YxF, on CTCF [141] (Figure 1.8b) and on Shugoshin 1 [250]. *In vivo* deletion experiments suggest, however, that more than these three amino acids are responsible for the interaction of cohesin and CTCF [251] and the genomic insulation at CTCF sites [252]. Recent estimates from live-cell tracking of loci adjacent to TAD boundaries suggest that wildtype CTCF is an imperfect roadblock to loop extrusion [160, 253]. These results are supported by recent *in vitro* experiments showing that CTCF's ability to block cohesin's loop extrusion depends strongly on DNA tension and

that cohesin-CTCF encounters can result in loop extrusion direction reversal and stably bound loops, as well as loop slippage, all of which are dependent on DNA tension (see chapter 7, [35]). For now, it remains incompletely understood how the binding of CTCF's YxF motif to cohesin is modulated and how that binding interferes with its loop-extruding activity. Some other proteins also contain the YxF motif, e.g., human MCM3 (which is part of the replication machinery), TFIIH (part of the transcription machinery), and Shugoshin 1 [250]. While to date no known functional association between cohesin and TFIIH has been reported, MCM3 was indirectly shown to block DNA loop extrusion by noting that CTCF-anchored loops gained intensity upon depletion of MCM3, similar but independent from the WAPL depletion phenotype [191]. Single-molecule assays showed that diffusion of DNA-bound cohesin is indeed severely impeded by MCM3 and the YxF motif. However, loop extrusion stalling by MCM could not be measured *in vitro*. Beyond CTCF, other factors also modulate SMCs. Intriguingly, work in *B. subtilis* has identified the site-specific recombinase XerD as an SMC unloader [254], while MatP may have this role in *E. coli* [125]. Like CTCF, XerD bound to specific DNA elements limits the extent of loop extrusion in vivo, although in this case, it appears SMC is actively unloaded from the chromosome. Biophysical characterization of this novel unloader requires the development of an *in vitro* assay for the bacterial SMC complex. Furthermore, Mediator has been reported to interact with both cohesin [255, 256] and NIPBL-Mau2 [255, 257], potentially as a means of forming specific DNA loops between Mediator-bound promoters and enhancers. Similarly, Yin Yang 1 (YY1) preferentially occupies enhancers and promoters, and enhancer-promoter interactions are lost upon YY1 depletion [36, 258]. However, simple DNA binding of factors such as Mediator or YY1 is not likely to be sufficient to establish contacts between elements located up to several hundred kilobase pairs apart (Figure 1.4a). Cohesin could loop the intervening space and, if stalled at regulatory elements, facilitate the contact, which could thereafter be maintained by the dimerization of transcription factors (Figure 1.4a). An interaction between YY1 and cohesin, and also condensin, has been reported [259], supporting this view. If enhancer-promoter contacts need to be established, this might occur only infrequently, e.g., only upon external stimuli, and once established, they could be maintained by various factors without cohesin (Figure 1.8c). These factors might cause such interactions to be nearly invisible in population-averaging mapping methods, whereas single-molecule *in vitro* experiments have the ability to show whether and to what extent such an association exists. Next, we describe some proteins that do not feature a biochemical interaction with SMCs but may act as physical roadblocks for loop extrusion. For example, elongating RNAps have been suggested to interact with cohesin. In particular, the notion that cohesin is found in islands between two convergently oriented genes led to the idea that active transcription can push intragenic cohesin toward transcription termination sites [32, 211]. This pushing seems to impact not only passively bound cohesin but also loop-extruding cohesins, as evidenced by the fact that relocation of cohesin by transcription is accompanied by the formation of new *cis* contacts [175, 260]. Computational modelling of Hi-C maps from bacteria [174] and humans [175] suggests that active transcription affects the DNA-looping pattern around genes, as loop extruders are temporarily stalled at elongating RNAp, where the stalling time depends on the direction of loop extrusion with respect to the direction of transcription. This may present a mechanism by which contacts between regulatory elements and RNAp can be maintained during elongation

and termination [261, 262]. The first single-molecule experiments showed that transcribing RNAp can block diffusing cohesin [211]. However, loop-extruding yeast condensin overcomes promoter-bound RNAp with ease [226]. This result calls into question the efficiency of RNAp as an interacting partner to SMC proteins. But *in vitro* studies have so far investigated only loop extrusion encounters with inactive promoter-bound RNAp, and actively transcribing RNAp might resemble the *in vivo* scenario closer. In fact, it might be the RNA transcript rather than the RNAp that interacts with cohesin and thereby stalls loop extrusion, and significant blocking may be observed only once the RNA transcripts grow to a sizeable length (Figure 1.8d). This hypothesis is supported by the fact that the cohesin subunits STAG1/2 [263] and NIPBL [264] are RNA-binding proteins. Furthermore, SMC proteins may encounter each other on the same DNA during loop extrusion. For example, the bacterium *B. subtilis* carries eight parS sites within ~800 kbp around its origin of replication, onto which BsSMC loads with the help of ParB [265, 266], and SMC-SMC encounters are likely frequent. Single-molecule experiments involving two loop-extruding condensin complexes showed that condensins can pass each other and form a double loop, a new loop motif that was called a Z-loop [267]. Loop-extruding SMCs loaded at these clustered parS sites are thus able to traverse one another upon encounter [265, 266]. Furthermore, simulations have shown that even though yeast condensin is a one-sided loop extruder [98], the traversal of condensins can rescue the formation of mitotic chromosomes by allowing further compaction than one-sided extruders without traversal would allow [268]. The bypassing of nucleosomes by DNA loop extruders is also crucial for support of the loop-extrusion hypothesis *in vivo*. As expected, condensin from humans and yeast indeed incorporates nucleosomes into compacted extruded DNA loops [216, 226]. While DNA loop extrusion is not considerably affected by single nucleosomes, it is conceivable that dense arrays of nucleosomes may prohibit cohesin binding [269], and stiffening of nucleosome arrays of nucleosomes by histone tail acetylation [29, 270] may stall loop extrusion (Figure 1.8e). An analogous phenomenon may be exploited in *S. cerevisiae* to resolve accidental fusions of telomeres. Stalling of loop extrusion due to an array of closely packed Rap1-binding sites at the fusion may localize it to the abscission site upon cell division [271]. Even though condensin can take steps of up to several hundred base pairs, it is conceivable that a stiff segment of DNA may halt loop extrusion if the Rap1-bound array is longer than the step size in physical space (i.e., ~40 nm) [217] or when Rap1 binding causes stiffening (190) such that the bending energy of the Rap1-bound DNA segment cannot be overcome [221, 241, 272]. Unexpectedly, it appears quite difficult to rigorously stop SMCs since there is, to date, no known single impermeable roadblock to DNA loop extrusion. Even the biochemical interactions between cohesin and CTCF's YxF motif have a dissociation constant in the micromolar range [141], implying that CTCF is also an imperfect roadblock. Only closely spaced arrays of DNA-bound proteins appear to be the sole roadblock with the capability to completely block loop extrusion to date.

1.7 Conclusion

The looping of the genome by SMC complexes through DNA loop extrusion is an old idea that was rejuvenated by *in vivo* Hi-C experiments. Single-molecule demonstrations of DNA loop extrusion [98, 173, 204, 212] provided the key *in vitro* evidence and have enabled molecular insights into the underlying mechanism. These studies have demonstrated the

ability to quantify the loop extrusion, as well as the interaction kinetics between SMC complexes and potential roadblocks. This also informs simulations that use the observed kinetics of *in vitro* experiments to resemble *in vivo* data more closely. While initial single-molecule experiments used model substrates of naked DNA, recent *in vitro* work has been designed to approach more complicated molecular constructs, e.g., by incorporation of potential roadblocks [191, 216, 226] or deletion experiments [221]. This will allow some of the many open questions in the field to be addressed. One example concerns the heavily debated regulation of genes: the persistent question of whether enhancer-promoter contacts are necessary to induce transcription, and whether such contacts may be made by modulating loop extrusion around regulatory sites. Another outstanding challenge is to reconstitute loop extrusion by bacterial and archaeal SMC complexes to investigate the similarities and differences in the underlying mechanism of loop formation. Questions regarding the roles of SMC proteins in replication and DNA repair are starting to emerge only now. Unfolding the molecular mechanisms governing DNA loop extrusion will not only contribute to our understanding of its role throughout the cell cycle but eventually also shed light on the causes of associated diseases such as cohesinopathies [273] and offer potential diagnosis and treatment targets.

2

In this thesis

As illustrated by the extensive listing of potential DNA loop extrusion models and an ongoing debate concerning the topic [221, 228, 242, 247, 274, 275], it is clear that the molecular mechanism leading to DNA loop extrusion is much sought after. This thesis delves deeply into some of the aforementioned aspects of SMC biology with the aim to understand how DNA loop extrusion occurs on roadblock-decorated chromatin, and how the DNA loop extrusion mechanism enables this.

In chapter 3, we start interrogating how SMC proteins deal with obstacles on the DNA that they encounter as they reel in DNA. We first examine biological DNA-binding proteins such as nucleosomes, RNA polymerases, and dCas9. These roadblocks virtually constituted no barrier to DNA loop extrusion by yeast condensin. Motivated by the fact that SMC complexes (cohesin) can embrace up to two DNA strands topologically, it was thought that loop extrusion may also embrace the loop topologically or pseudotopologically. This led us to explore the encounter of SMCs with nanoparticles up to 200 nm in diameter that were coupled to dCas9 which was in turn bound to DNA. While such very large nanoparticles did pass into the loop less frequently, they were still extruded with a non-zero probability. Being even larger than the circular stretched-out peptide chain of Smc2-Smc4-Brn1 itself, these results meant that either any interface of the Smc lumen had to open for roadblock passage or that loop extrusion could occur entirely nontopologically, i.e. roadblocks never pass through the Smc ring, not even pseudotopologically. To test these hypotheses, we used an engineered version of human cohesin whose Smc-kleisin and Smc-Smc interfaces cannot be opened [173]. Strikingly, even this engineered version passed 200 nm large roadblocks, leading us to the conclusion that DNA loop extrusion can occur entirely nontopologically.

During the revision of our manuscript with these data, new results were published on the bioRxiv preprint server which argued for a pseudotopological model of DNA loop extrusion [221]. While the authors could explain particle bypass, their proposed model made a number of predictions. In chapter 4, we identify and explain these predictions and then test them against available experimental data. We conclude that the pseudotopological

model stands in contrast with the experimental data and that instead a nontopological model best describes all available data.

DNA loop extrusion was first unequivocally demonstrated for yeast condensin [98], one year later followed by the demonstration of loop extrusion by human cohesin [173, 212] and later others. A striking difference between yeast condensin and human cohesin was the directionality of DNA loop extrusion. Yeast condensin extruded DNA strictly asymmetrically from one side only while human cohesin reeled in DNA from both sides into the DNA loop. This variability despite the large structural similarities between the eukaryotic SMC complexes undermined whether the loop extrusion mechanism is shared among all SMC complexes. In chapter 5, we closely examine the dynamics and directionality of DNA loop extrusion of all eukaryotic SMC complexes. We find that all examined SMC complexes extrude DNA strictly asymmetrically. However, human cohesin and yeast SMC5/6 can switch the direction of loop extrusion frequently, evoking the impression of symmetric loop extrusion when averaged over time and in sideflow experiments. Focusing on human cohesin, we identified the subunit NIPBL as a molecular gearbox. NIPBL is not a stable part of the cohesin complex, but dynamically associates with and dissociates from the complex. By visualizing its exchange on cohesin directly using a pool of differentially fluorophore-labeled NIPBL complexes, we find that an exchange of NIPBL strongly correlates with the switch of the extrusion direction. This leads us to propose that DNA binding to SMCs (e.g. through NIPBL on cohesin) may be temporarily weakened (e.g. by dissociation of NIPBL from cohesin) in order to switch the downstream and upstream DNA strands for loop extrusion. Furthermore, these results suggest that asymmetric DNA loop extrusion is the common mechanism shared by all eukaryotic SMC complexes.

An early biochemical reconstitutions of a *Xenopus* condensin in 1997 [48] demonstrated that plasmids were supercoiled by condensin in an ATP-dependent manner. In the meantime, sister chromatid cohesion by single cohesin complexes [132] and DNA loop extrusion was demonstrated [98, 173, 212] and DNA supercoiling by SMCs remained unaddressed. A wave of excitement sparked new interest in SMC-mediated DNA supercoiling but it remained unclear whether or not supercoiling is a separate function from loop extrusion, like cohesion. We answer this question in chapter 6, where we use a magnetic tweezers setup that has previously proven powerful in resolving individual steps of DNA loop extrusion [217]. Using torsionally constrained DNA, we find that the extruded loop is negatively twisted at every step upon ATP binding. Strikingly, the quantity, on average -0.6 twists per step, appears to be conserved for condensin, cohesin, and SMC5/6. Previous experiments on DNA supercoiling by condensin used a bulk assay in which a torsionally constrained plasmid is incubated with an excess of SMC complexes, ATP, and a topoisomerase which fixes the SMC-induced linking number changes [48, 49] (some experiments using this bulk assay for human cohesin are reported in chapter 6). After deproteinization of the reaction products, the linking number of the plasmids is resolved *via* gel electrophoresis. This assay showed that the handedness of the reaction product depends on the excess of SMCs over plasmid, leading the authors to question whether SMCs really induce supercoils with only negative handedness into the loop. Equipped with the quantities obtained from single-molecule magnetic tweezers measurements, we computationally modelled the accumulation and distribution of twist within the plasmid and the twist fixation by topoisomerases. These

simulations recapitulated the experimental results of the plasmid bulk assay with the single-molecule data, requiring solely that SMCs induce -0.6 twists/step into the extruded loop.

DNA loop extrusion appears to be beneficial to the cell in two very different phases of the cell cycle. During metaphase, maximal compaction of chromatids is reached when SMCs are permitted to loop DNA undisturbed, i.e. without any obstacles to stop them [112, 113, 268]. During interphase, however, loop extrusion plays a more subtle, regulatory role in finetuning gene expression and contributing to genome maintenance [181, 196, 201, 276]. The characteristic organization of the genome at the level of TADs vanishes in the absence of CTCF, a key regulator of chromosome structure [32–34]. In chapter 7, we reconstituted DNA-bound CTCF on the single-molecule level and tested whether this system was able to halt cohesin-mediated loop extrusion in a directional manner. We found that CTCF alone is sufficient to stall both diffusing as well as extruding cohesin. Unexpectedly, however, we found that the efficiency of CTCF to stall extruding cohesin was strongly dependent on the tension acting on DNA at the moment of encounter, a biophysical parameter which was not considered so far. These results suggest that DNA mechanics, in particular the tension acting on DNA in the sub-pN range, modulates the permeability of CTCF boundaries and thus the length of extruded chromatin loops.

In chapter 8, we delve into different modes through which CTCF interferes with DNA loop extrusion by cohesin. Employing magnetic tweezers and single-molecule fluorescence visualization, we examine the impact of various peptides spanning the CTCF N terminus on cohesin's ability to extrude DNA loops. When exposed to the established YxF motif, cohesin frequently fails to take consecutive DNA loop extrusion steps. Furthermore, cohesin complexes typically exhibit unidirectional extrusion when exposed to the YxF motif. Additionally, we identify a second motif near the N terminus that produces a distinct effect. Peptides containing a KTYQR motif hinder the overall stepping activity of cohesin but do not alter the directionality of loop extrusion. This concluding chapter provides a summary of the progress achieved so far.

Finally, I summarize the findings of this thesis in chapter 9 and provide my perspective on developments and key experiments to advance the SMC field in the future.

II

Towards a mechanistic understanding of DNA loop extrusion



3

SMC complexes can traverse physical roadblocks bigger than their ring size

This chapter has been published as: Biswajit Pradhan, Roman Barth, Eugene Kim, Iain F. Davidson, Benedikt Bauer, Theo van Laar, Wayne Yang, Je-Kyung Ryu, Jaco van der Torre, Jan-Michael Peters, Cees Dekker: SMC complexes can traverse physical roadblocks bigger than their ring size, Cell Reports 2022 Oct 18;41(3):111491 [226].

Abstract

RING-SHAPED structural maintenance of chromosomes (SMC) complexes like condensin and cohesin extrude loops of DNA. It remains, however, unclear how they can extrude DNA loops in chromatin that is bound with proteins. Here, we use in vitro single-molecule visualization to show that nucleosomes, RNA polymerase, and dCas9 pose virtually no barrier to loop extrusion by yeast condensin. We find that even DNA-bound nanoparticles as large as 200 nm, much bigger than the SMC ring size, also translocate into DNA loops during extrusion by condensin and cohesin. This even occurs for a single-chain version of cohesin in which the ring-forming subunits are covalently linked and cannot open to entrap DNA. The data show that SMC-driven loop extrusion has surprisingly little difficulty in accommodating large roadblocks into the loop. The findings also show that the extruded DNA does not pass through the SMC ring (pseudo)topologically, hence pointing to a nontopological mechanism for DNA loop extrusion.

3.1 Introduction

Chromosome organization in eukaryotic cells is vital for genome segregation, gene regulation, and recombination [118, 277]. Structural maintenance of chromosome (SMC) complexes such as condensin and cohesin play a key role in the three-dimensional organization of the chromosome [278–280]. These SMC complexes are DNA-binding ATPases, in which a SMC heterodimer and kleisin subunit form a ring-like structure [243, 281]. SMC complexes organize the genome by extruding DNA loops in a processive and ATP-dependent manner [112, 282–284]. Both cohesin and condensin have been shown to extrude loops *in vitro* on bare DNA molecules that were tethered onto a surface [98, 173, 212, 214, 216].

It remains unclear, however, how SMC complexes deal with the abundant proteins that are part of chromatin in cells. *in vivo*, SMC complexes will encounter lots of DNA-binding proteins such as nucleosomes as well as large obstacles such as the DNA replisome, with a size of ~20 nm [285], or a transcribing RNA polymerase that, with its RNA transcript and spliceosome, can have a globular diameter of >70 nm [286], i.e., even bigger than ~35 nm ring size of SMCs. Some data indicate that SMCs can bypass DNA-bound objects, e.g. Kim et al. [212] and Kong et al. [216] found that nucleosomal DNA was compacted by cohesin and condensin, and yeast condensin was shown to be able to bypass another condensin to form higher-order DNA loops [267]. Yet, other evidence has indicated a slowing down or blocking of DNA loop extrusion by specific proteins. For example, the DNA-binding protein CCCTC-binding factor (CTCF) regulates topologically associating domains and long-range interactions by blocking loop extrusion through specific binding to cohesin [141]. Furthermore, *in vivo* data suggested that transcribing RNA polymerases are able to 'push' cohesin to the 3'-end of highly transcribed genes in eukaryotes [32, 260] and Brandão et al. theoretically predicted that transcription slows SMC complexes down but does not block SMC translocation in B. subtilis [174]. Under conditions in which cohesin diffuses along DNA, SMCs were reported to be blocked by objects larger than 20 nm [102, 211]. Overall, it therefore remains unresolved whether loop-extruding SMC complexes will block, stall, bypass, dissociate, or push obstacles that are present on the DNA upon encounter.

One might expect that this relates to the topology of the SMC complex during DNA loop extrusion. SMCs can embrace DNA in their ring structure, as cohesin was found to mediate sister chromatid cohesion through topological entrapment [287] while condensin was reported to topologically link chromatid arms for their structural rigidity in mitosis [146]. Whether topological entrapment is necessary for loop extrusion is under debate. While most models [132, 146, 229, 233, 235, 241, 288–290] assume some form of DNA entrapment, recent data on human cohesin with a covalently closed-ring structure indicated its ability to extrude loops [173], suggesting a possible pseudo- or non-topological loading of the SMC complex during DNA loop extrusion, where, respectively, the DNA gets inserted into the SMC ring without opening the ring, or where the DNA loop does not get embraced by the SMC ring at all but DNA binding occurs externally.

Here, we systematically study the effect of DNA-binding proteins on DNA-loop extrusion

for both yeast condensin and (covalently closed) human cohesin. Strikingly, we find that particles as large as 200 nm, i.e., much bigger than the ~35 nm condensin ring size, can be translocated into DNA loops during extrusion. This demonstrates that SMCs can handle protein-loaded chromatin substrates for loop extrusion without any problems, and provides evidence for a nontopological mechanism of SMC-driven DNA loop extrusion.

3.2 Results

3.2.1 Visualization of a condensin encountering an obstacle while extruding a DNA loop

Loop extrusion was visualized *in vitro* by monitoring an Alexa647 fluorescently labelled 'roadblock' protein on fluorescently labelled DNA (sytox orange (SxO); Figure 3.1a). Roadblock proteins were bound to 48.5kb λ-DNA, which was anchored onto a streptavidin-coated passivated glass surface at both its ends through biotin linkers (Figure S3.1) and imaged with homebuilt HILO sheet microscopy. Upon addition of condensin (0.5 nM) and ATP (5 mM), fluorescence spots locally formed on DNA, whose intensity grew until the slack was removed between the two DNA ends. In-plane buffer flow perpendicular to the DNA confirmed that these were extruded DNA loops. Consistent with previous findings [98], condensin extruded a DNA loop asymmetrically, and consequently, the DNA-bound roadblock was reeled towards the loop or remained fixed with respect to the loop position (Figure 3.1c, Figure S3.1i-l). Events, where the roadblock was moving towards the loop to subsequently co-localize with the loop, were identified as an encounter and considered for further analysis.

Upon encounter, loop extrusion by condensin either blocked, or the roadblock traversed the stem of the condensin-mediated DNA loop to subsequently translocate into the extruded DNA loop (Figure 3.1b). Figure 3.1c shows an example of such a traversal, where the DNA formed a loop that grew in size until it encountered the roadblock (here dCas9) whereupon it continued to grow such that the loop encompassed the roadblock protein. When the loop spontaneously disrupted after 9 minutes, the protein was still bound to the DNA and, as expected, at its initial position. Alternatively, a roadblock can get blocked at the stem of the loop upon encounter. We never observed pushing of the roadblock along the DNA or disruption of the loop upon the encounter. To avoid imposing additional tension within the DNA by the side flow, which can slow down or stall loop extrusion [98, 217], most roadblock experiments were performed without side flow (Figure 3.1d). Kymographs were used to visualize loop formation and roadblock positions over time, see Figure 3.1e,f. Upon encounter, a continued movement of the now co-localized loop and roadblock signaled the passing of the roadblock into the loop, since the SMC still drove the loop expansion. Quantification of the loop size from the intensities (Figure 3.1g) provided further details, see Figure 3.1i,j,k for examples of a passage, stalling and passing, and blocking event, respectively.

We found that the mean square displacement (MSD) of the roadblock along the DNA serves as a good marker for discriminating whether the roadblock got blocked or passed into the

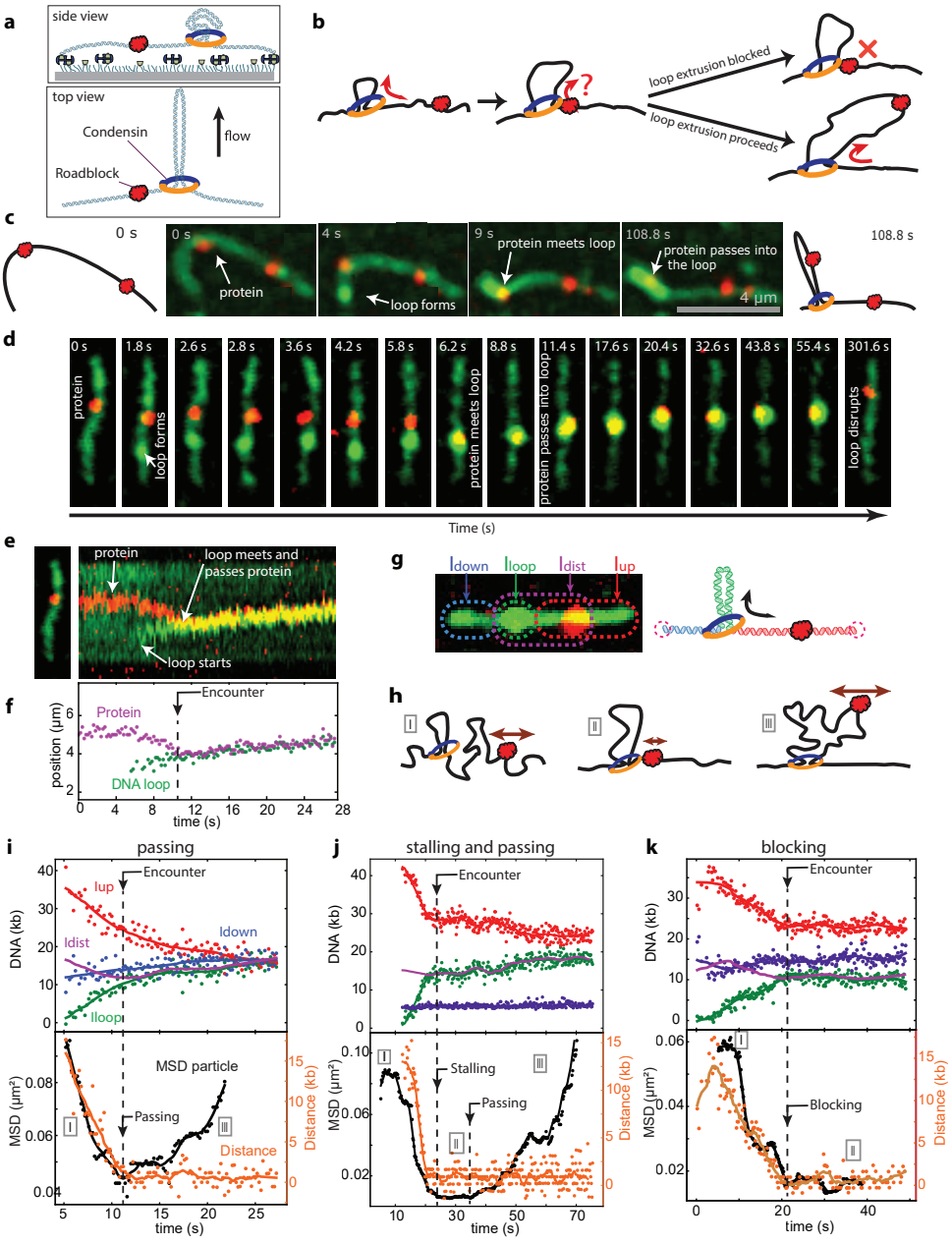


Figure 3.1 | Visualization of a condensin encountering an obstacle while extruding a DNA loop.
(**a**) Loop extrusion assay. (**b**) Two scenarios for condensin encountering an obstacle during loop extrusion: The obstacle can either pass into the loop or get blocked. (**c**) Snapshots of condensin bypassing a DNA-bound protein (dCas9) and including it in the loop during DNA loop extrusion with side flow. (**d**) Snapshots of DNA loop extrusion where the loop encounter accommodates a protein. (**e**) Kymograph for the event in (**d**). (**f**) Localizations of the protein (magenta) and DNA puncta (green) from the kymograph. (**g**) Color scheme denoting the various fluorescence intensities of different regions on DNA. (**h**) Cartoon showing the Brownian fluctuations of the roadblock at different stages of the encounter. (**i**) Kinetics of the DNA loop size formation for a passing event. Intensities are denoted in colors outlined in panel g. The bottom curve show MSD (moving average; black) and SMC-roadblock distance (orange). (**j**) Same as i, but for a stalling and subsequent passing event. (**k**) Same as i, but for a blocking event.

extruded loop. For a passage event, the MSD (Figure 3.1i, bottom) could be seen to rapidly decrease until the condensin-roadblock encounter, and rapidly increase again after the encounter. This is explained from, respectively, the tightening up of the DNA between its two ends in the early phase of loop extrusion (i.e., going from I to II in Figure 3.1h), and the increased distance from the roadblock (inside the loop) to the loop stem which is associated with an increased Brownian motion (III in Figure 3.1h) since the DNA within the loop is not under tension. The MSD (Figure 3.1i, bottom) changes much more distinctly over time than the loop size (Figure 3.1i, top), making it a discriminating measure for inferring whether the roadblock did or did not get incorporated into the extruded DNA loop.

3.2.2 Loop-extruding condensin traverses DNA-bound proteins such as nucleosomes, RNA polymerase, and dCas9

In probing the effects of roadblocks on condensin-driven DNA loop extrusion, we first examined nucleosomes, the basic building block of chromatin with a diameter of ~11 nm (Figure 3.2a). Nucleosomes were reconstituted from yeast histone octamers with Alexa647 on H3A (see Methods section 3.5), and assembled as single nucleosomes on DNA (Figure 3.2b; Figure S3.2a-d), yielding 2-5 isolated nucleosomes per λ-DNA in our optical assay. Both passing and blocking events were observed (Figure 3.2c-g), but passing events were much more frequent (88 ± 4%, Figure 3.2i). A small fraction of the encounters showed stalling and passing with a pausing time of ~10 s (Figure S3.2p). Notably, the blocking events almost always occurred at the end of a loop extrusion where the forces within the DNA were approximately similar to the stalling force for condensin (Figure S3.2m). Other proteins with similar diameters, specifically RNA polymerase (RNAP) from E. Coli, and dCas9 (Figure 3.2h), showed similar passing fractions as the nucleosome, namely 92 ± 5 % and 87 ± 8 %, respectively. This indicates that ~10 nm-sized DNA-binding proteins can be readily accommodated into the extruded loop.

3.2.3 Condensin and cohesin can pass roadblocks bigger than their ring size

Next, we asked whether yet larger obstacles on DNA would block DNA-extruding condensin. Gold nanoparticles were chosen to controlably vary the size of the roadblocks, as they can be obtained with a narrow size distribution with a median diameter of 10-125 nm. Gold particles were functionalized with SH-PEG-NH_2/SH-PEG-COOH (see Methods section 3.5) which minimized interactions with DNA, condensin, and the glass surface. The amine (-NH_2) groups facilitated binding to Alexa647 and to a dCas9-Snaptag that allowed sequence-specific anchoring of the nanoparticle to the DNA *via* dCas9. TEM imaging (Figure 3.3b-d), which visualized the gold core but not the outer organic layer, showed a narrow size distribution. The hydrodynamic diameter was estimated from fluorescence correlation curves (Figure S3.2j) and comparison of the TEM and hydrodynamic diameters yielded a thickness of 7.9 ± 1.2 nm for the passivation layer. For the largest particle in our study, we used polystyrene nanoparticles with the same functionalization that added 16 ± 2 nm to the 182 ± 8 nm diameter as measured by EM, leading to a 198 ± 10 nm total diameter of these '200 nm' polystyrene nanoparticles.

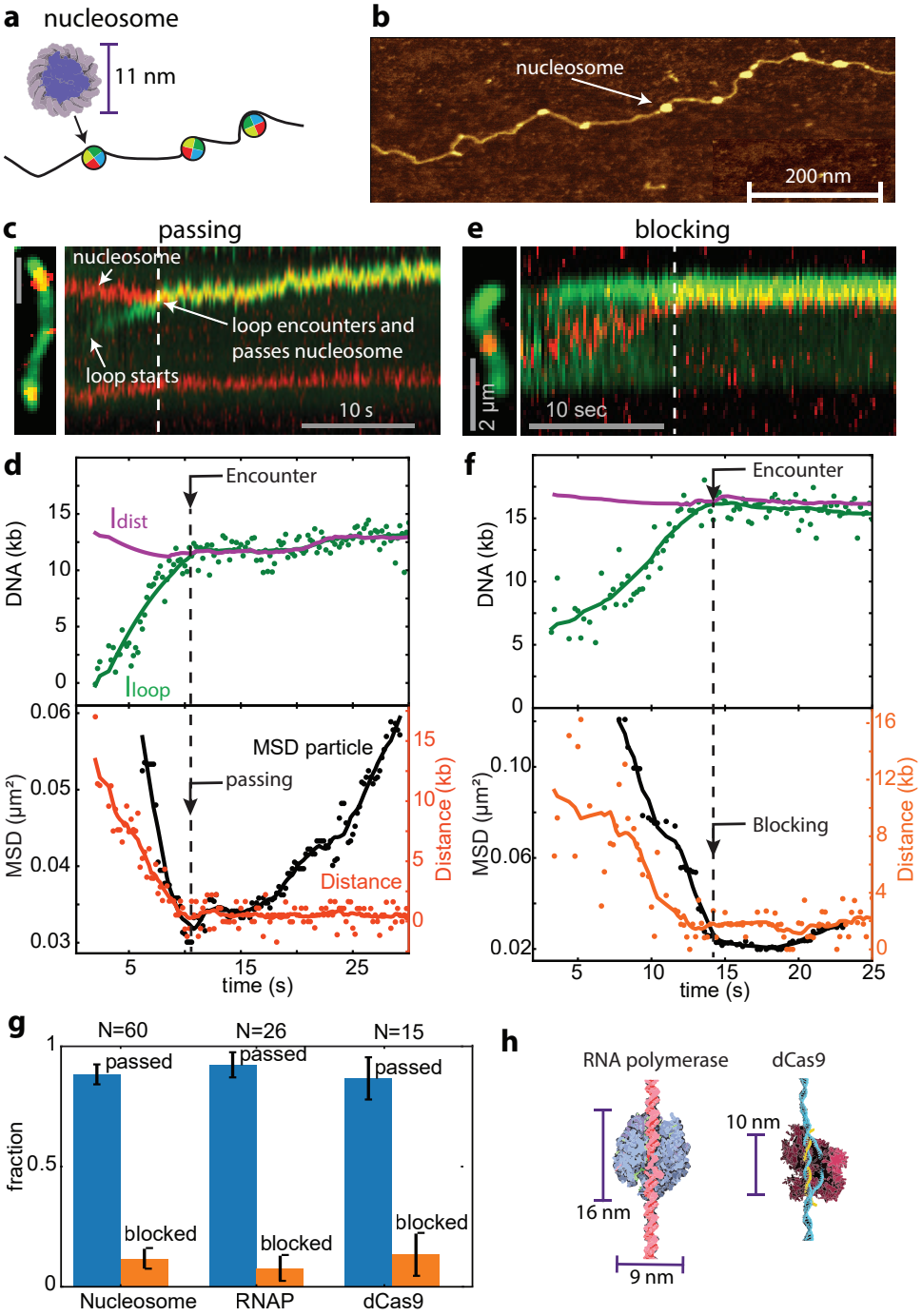


Figure 3.2 | Loop-extruding condensin traverses nucleosomes and other DNA-bound proteins.
(**a**) Schematic of nucleosomes on DNA. (**b**) AFM image of single nucleosomes on DNA. (**c**) Kymograph for a passage event. (**d**) Corresponding loop kinetics and MSD traces. (e, f) Same as c, d, but for a blocking event. (**g**) Fraction of the roadblocks passed (blue) and blocked (orange) for nucleosomes, RNAP and dCas9. (**h**) Schematics of RNAP and dCas9 (adapted from Wikimedia Commons).

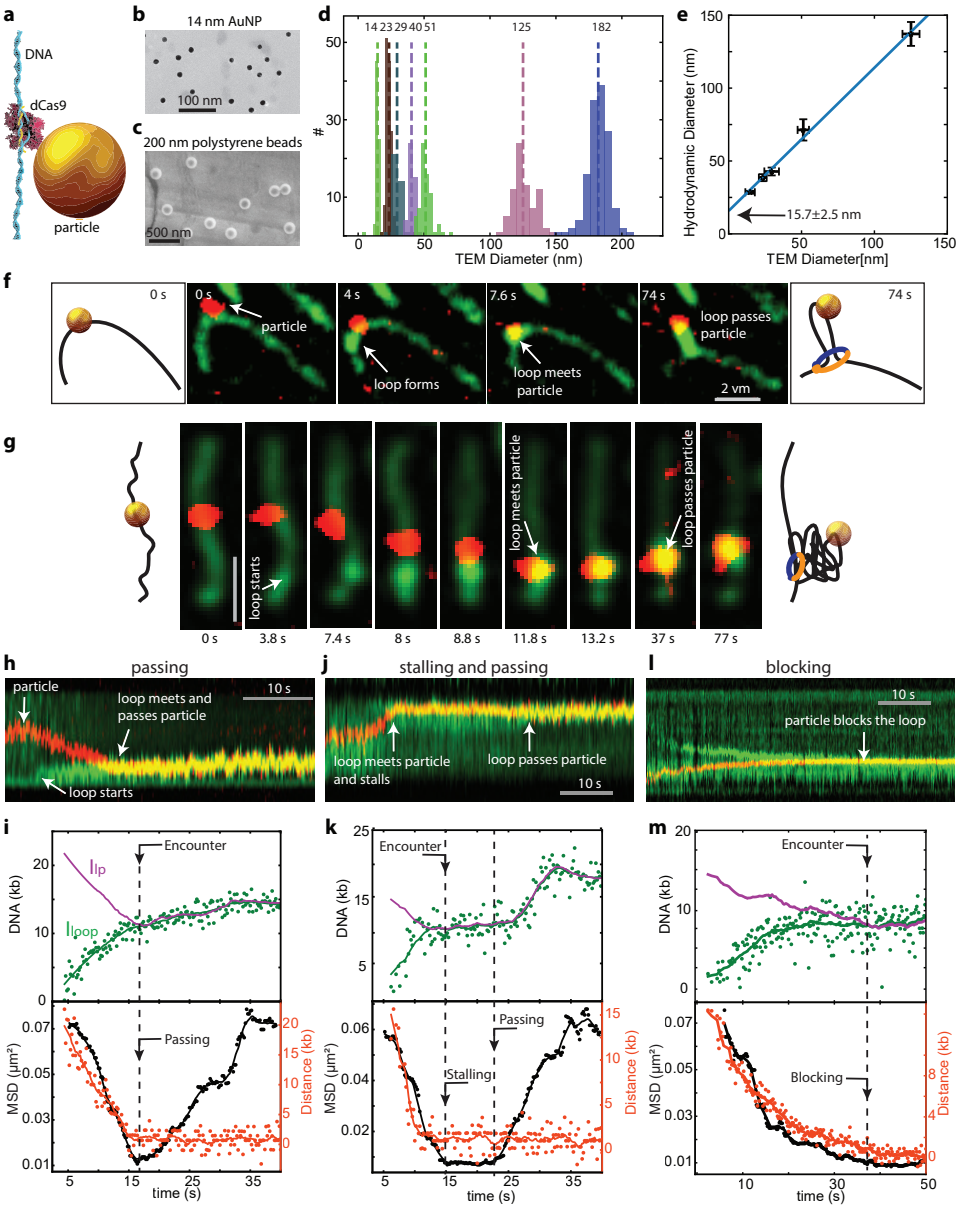


Figure 3.3 | Functionalized gold nanoparticles as controlled steric obstacles.
(**a**) Schematic showing the binding of a functionalized gold nanoparticle to the DNA through dCas9. (**b**) TEM image of functionalized gold nanoparticles of sizes 14 nm showing uniform sizes and absence of aggregates. (**c**) Idem for 186 nm polystyrene beads. (**d**) Size distribution of the functionalized gold nanoparticles from the TEM images. (**e**) Correlation between hydrodynamic diameter obtained from FCS and diameters obtained from TEM. A linear fit shows an additional thickness of 15.7 nm on the nanoparticles. (**f**) Snapshots of DNA loop extrusion with side flow for a 39 nm gold nanoparticle (red) on the DNA (green) that ends up within the loop. (**g**) Snapshots of loop extrusion without side flow with a 39 nm gold nanoparticle (red) on the DNA (green). (h, j, l) Kymograph of DNA (green) and nanoparticle (red) corresponding to passing (**h**), stalling before passing (**j**) and blocking (**l**) events. (i, k, m) Loop kinetics and MSD traces corresponding to passing (**i**), stalling before passing (**k**) and blocking (**m**) events.

Condensin-driven loop extrusion was examined for functionalized nanoparticles of different sizes that were bound onto DNA. Figure 3.3f-m provides examples of gold nanoparticles with a hydrodynamic diameter of 39 nm. We observed different outcomes upon encounter of the condensin with the roadblock, viz., passing (Figure 3.3f-i and Figure S3.4), stalling and passing (Figure 3.3j,k, Figure S3.2p), and blocking (Figure 3.3l,m). Application of a side flow unequivocally demonstrated that a 39 nm particle can be inserted into the loop (Figure 3.3f). From the statistics of passing versus blocking, 39 nm particles show similar high passing fractions (94 ± 6%; including 'stalling and passing' events) as the nucleosomes (88 ± 4%).

The most striking finding of our study is that DNA-bound particles with a diameter as large as 200 nm can pass into the loop during the process of DNA loop extrusion. Figure 3.4c shows representative side-flow images of DNA loop extrusion in the presence of a 200-nm particle where the particle can be seen moving towards the loop, then encountering the loop, to subsequently pass into the loop. Figure 3.4c-e shows another example of a passage event where a 200 nm particle gets incorporated into the extruded DNA loop, as evidenced by the increase in both the loop size and the clear minimum in the MSD at the encounter. Notably, this 200 nm size is much larger than the condensin ring (cf. Figure 3.4a for a comparison to scale), as the circular SMC complex has a diameter of about 35 nm [131], as is e.g. evident from AFM imaging (Figure S3.2e, f), whereas theoretically one could imagine a ~110 nm ring diameter (Figure S3.2i) when combining the ~50 nm long Smc arms and stretching the kleisin to its full maximal length, which however is non-physiological.

As displayed in Figure 3.4b, the fraction of DNA-bound nanoparticles that traversed the condensin to get incorporated into the DNA loop decreased from >90 % for particle sizes up to 40 nm, to 57 ± 9 % for the 200nm-size particles. The very large passing fractions for the smaller particles indicate that the vast majority of biologically relevant DNA-binding proteins will be readily accommodated into extruded DNA loops. While the passing fraction decreases with increasing particle size, it is remarkable that >50% of the 200 nm particles still get translocated into the extruded loop - implying that also very large protein complexes on DNA will traverse into DNA loops. About 10-20 % of all encounters exhibited stalling before subsequent passing, with a characteristic stalling time of about ~40 s (Figure S3.2p).

To explore the generality of these findings, we performed similar experiments with human cohesin, which was also shown to extrude DNA loops [173, 212]. For both 30 nm and 200 nm DNA-bound nanoparticles, we observed that these roadblocks were able to pass the cohesin complex during loop extrusion. Figure 3.4g shows a typical example of a passage event that shows that cohesin behaves very similarly to condensin, although the encounter/MSD analysis is slightly more involved due to the two-sided nature of loop extrusion for cohesin (see SI and Figure S3.3). Interestingly, both 30 nm and 200 nm particles traverse cohesin to pass into the loop with similar efficiencies, 47 ± 8 % and 45 ± 15 % respectively (Figure 3.4b, right).

Our observation that 200-nm particles can be translocated into DNA loops by extruding

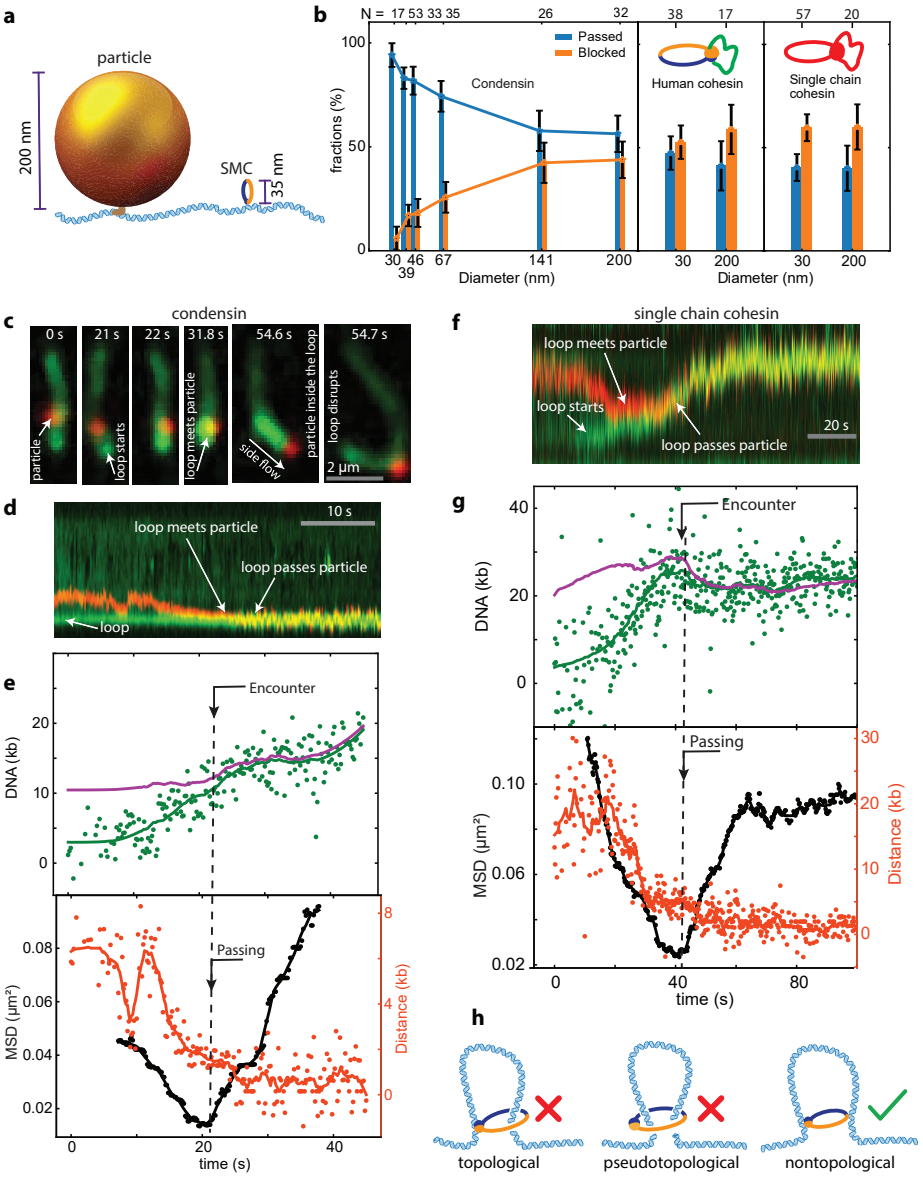


Figure 3.4 | Condensin and cohesin can pass roadblocks bigger than their ring size.
(**a**) Sketch of a Cas9-200nm nanoparticle and condensin on DNA, to scale. (**b**) Fraction of roadblocks passed into the loop or blocked by yeast condensin, human cohesin, and single-chain cohesin, for different particle sizes. (**c**) Snapshots of loop extrusion by condensin on a DNA that accommodates a 200 nm roadblock in a passing event. At 54.7 s, a side flow was applied and the particle can be seen inside the loop. (**d**) Kymograph of a passage event where condensin traverses a 200 nm roadblock. (**e**) Loop kinetics and MSD trace for the event in (**d**). (**f**) Kymograph of a passage event where single-chain human cohesin traverses a 200 nm roadblock. (**g**) Loop kinetics and MSD trace for event (f). (**h**) Models of topological, pseudotopological, and nontopological embracing of DNA underlying DNA loop extrusion.

SMC complexes suggests that the DNA that is being extruded is not encompassed by the ~35 nm SMC ring structure. Alternatively, large obstacles may get translocated across an SMC complex by transiently opening the SMC ring structure. To rigorously test this important question, we examined a single-chain variant of human cohesin in the roadblock assay. In this engineered version of cohesin (Figure S3.2g), the three ring-forming subunits SMC3, SCC1 and SMC1 are expressed as a single fused polypeptide chain. Similar to wildtype cohesin, this fusion protein folds into a ring-like structure in which SMC1 and SMC3 hetero-dimerize *via* the hinge domain to close the ring. This hinge-hinge interaction is furthermore crosslinked *via* cysteine residues, resulting in the formation of a covalently closed ring structure to which the other holocomplex subunits (STAG1 and NIPBL) can bind [173]. Remarkably, our data show that these single-chain cohesin complexes do pass the 30 nm and 200 nm particles during loop extrusion with a similar efficiency (40 ± 6 % and 44 ± 16 % respectively) to that of wildtype cohesin (Figure 3.4b,f,g). These data demonstrate that the roadblock passage during loop extrusion does not involve ring opening of the SMC complex.

3.3 Discussion

From systematically investigating the effects of DNA-binding roadblocks on DNA loop extrusion by SMC complexes, we thus conclude that an SMC complex can effectively traverse roadblocks, a finding that even holds for surprisingly large obstacles. Our results directly show that DNA-bound protein complexes such as nucleosomes and other ~10nm sized DNA-binding proteins pose no barrier to condensin-induced loop extrusion. Our finding that RNA polymerases do not block DNA loop extrusion is consistent with Hi-C experiments on bacterial condensin [174], and the ~14 s stalling times that we observed are similar to the estimated *in vivo* stalling time of ~10 s for highly transcribed operons [174].

Our data have important consequences for understanding the activity of SMCs in cells since the ability of SMC complexes to overcome very large protein complexes on DNA has obvious advantages in its processing of cellular chromatin. Vice versa, the finding that SMC complexes are largely unobstructed by physical barriers along the DNA indicates that SMC barriers such as CTCF must function according to other principles, e.g. must be governed by biochemical interactions, similar to what was recently proposed for how CTCF blocks cohesin by binding to the SA2-SCC1 subcomplex of cohesin [141].

The very large size of nanoparticles that can be accommodated in loop extrusion also has consequences for the mechanistic modeling of the SMC motor proteins, as it indicates that SMC complexes can readily bypass obstacles that are clearly larger than its ~35 nm ring size. Covalent linking of the interfaces of cohesin did not affect its passing ability, showing that a temporary opening of the ring structure is not needed for the loop extrusion process and roadblock traversals. This has implications for our understanding of the topology through which the SMC complex interacts with DNA during loop extrusion (cf. Figure 3.4h for a visual explanation of the possible topological modes). Early molecular models for loop extrusion suggested that the SMC ring topologically embraces the DNA

upon loading and during DNA loop extrusion (see e.g. [229, 241, 288]). More recently, also pseudotopological [173, 215, 282] and nontopological loading [173, 291] of the DNA were considered. Furthermore, after releasing the findings of this manuscript on bioRxiv, new models were proposed [221, 242] which considered a pseudotopological mechanism for loop extrusion. Prompted by our preprint data, these authors attempted to account for the passage of large obstacles within their model (SI Figure 25 in [221], and SI Figure 6 in [242]). However, this alternative interpretation of our data makes a number of predictions that are not supported by the experimental results [244] (see also chapter 4). Specifically, when the SMC encounters a large roadblock, these models predict the appearance of two loops, not one, since the roadblock - with its size that cannot pass the ring structure - obstructs the merging of the two preformed loops postulated in the models, and hence leads to the formation of a second loop. One could assume (as in [221, 242]) that the preformed loop might slip into the new loop, to again form one joint loop, but this is not the case, as can be seen from considering the force balance for the extruded DNA loop (which in our experiments are typically >10 kb). Summing up, the recently proposed pseudotopological models imply the appearance of two loops upon roadblock encounter as well as that the particle gets stuck at the SMC. Both predictions are however not confirmed by our experimental observations. Instead, the data from the current study provide evidence in favor of a nontopological mode for DNA loop extrusion, where DNA binds externally to the SMC complex.

The inclusion of large DNA-bound particles into extruded loop that are produced by SMCs (most strikingly, even by single-chain cohesin) discards the possibility of (pseudo) topological entrapment of the extruded DNA in the process of loop extrusion (in the strict definition of Figure 3.4h where the extruded DNA passes through the SMC ring). Notably, however, these experiments are agnostic about the question of whether or not a DNA strand is entrapped in other biological processes involving SMCs, where such a topological loading likely is very relevant. It will be interesting to expand the approach of the current roadblock study even further, e.g. by examining Smc5/6 SMC complexes [204] or cohesin for a variety of co-factor conditions, in order to examine the generality of our findings and to gain further understanding of the control of chromosome structure by SMC-driven DNA loop extrusion.

3.4 Limitations of the study

Our experiments provide a visually striking demonstration that SMC-driven loop extrusion has surprisingly little difficulty in accommodating large DNA-bound protein complexes into the extruded DNA loop. The experiments unambiguously show that these large objects do not block loop extrusion by SMC proteins. Like every technique, however, our methodology also has its technical challenges and limitations. For example, experiments on protein roadblocks (nucleosomes, RNAP, dCas9) can be hindered by the photobleaching and a low signal-to-noise ratio (SNR): The fluorescence of the roadblock protein has to survive long enough for the encounter with the SMC complex to be observed, and the SNR of protein roadblocks can be low as the signal comes from a single fluorophore on a single molecule that is fluctuating in position within the field of view. Such a lower SNR adds uncertainty to the localization precision of the protein roadblock, thus increasing the noise of the

calculated MSD, which can make it challenging to precisely localize the time point for the encounter between the roadblock and condensin. More generally, we present *in vitro* experiments, and inevitably, it is always a question to what extent *in vitro* data resemble the *in vivo* behavior of SMC complexes. While our experiments clearly demonstrate the intrinsic capabilities of SMCs to bypass huge roadblocks and accommodate them into the extruded DNA loops, it remains of interest to see how this plays out in the crowded environment of chromatin in the nucleus.

3.5 Methods

Experimental model and subject details

Yeast strains: *Saccharomyces cerevisiae* strains are derived of W303. Pentameric *S. cerevisiae* condensin complexes were purified as reported previously [98].

Insect Cells: *SF9* insect cells were used to express cohesin as described in [173]. After expression, cultures were harvested after 48 - 60 h at 27 °C, washed in PBS, frozen in liquid nitrogen and stored at -80°C.

Bacterial strains: *E. coli* strain BLR λDE3 was used for purification of RNA polymerase with SNAP-tag as described by Tetone et al. [292]. Histone proteins for nucleosome reconstitution were co-expressed in *Escherichia coli* strain BL21-codonplus-DE3-RIL (Agilent).

Method details

DNA and condensin preparation

Condensin holocomplex from *S. cerevisiae* was purified using our previously published protocol [98]. λ-DNA containing a biotin on both ends was made by hybridizing and ligating short oligonucleotides containing 5' phosphate and 3' Biotin on the single-stranded DNA ends of λ-DNA (NEB, N3011S). For this, we used oligonucleotides JT41 (P-GGGCGGCGACCT-Bio) and JT42 (P-AGGTCGCCGCCC-Bio) (IDT). Taq DNA ligase (NEB, M0208L) was used to ligate the oligonucleotide on Lambda, using 10 times molar access of oligonucleotide to λ-DNA. The mixture was incubated for 10' at 65°C and then for 1 hour at 50°C in Taq DNA ligase buffer. The biotin-λ-DNA was then cleaned up from free oligonucleotides and enzymes using an AKTA pure system (Cytiva), with a homemade gel filtration column containing approximately 46ml of Sephacryl S-1000 SF gel filtration media (Cytiva), run with TE supplemented with 150 mM NaCl buffer at 0.2 ml/min. The fractions containing the Biotin-λ-DNA were aliquoted and stored at -20°C.

Protein roadblocks - Nucleosome reconstitution and AFM imaging

Histone octamers were purified and assembled as established by Sanchez et al. [293]. pET-Duet.H3(D82C)-H4 and pCDFduet.H2A(K120C)-H2B were made, using site-directed mutagenesis with pfu-Turbo (Agilent) and Dpn1 (NEB), from pET-Duet.H3-H4 with primers TL-CH-13 (GCGCAGGATTTCAAAACCTGTCTGCGTTTTCAGAGCAGCG), TL-CH-14 (CGCTGCTCTGAAAACGCAGACAGGTTTTGAAATCCTGCGC) and pCDFduet.H2A-H2B with primers TL-CH-5 (CAAAACTTGTTGCCATGCAAGTCTGCCAAGACTGC), TL-CH-6 (GCAGTCTTGGCAGACTTGCATGGCAACAAGTTTTG). Histone proteins were co-expressed in *Escherichia coli* strain BL21-codonplus-DE3-RIL (Agilent) from pET-Duet.H3(D82C)-H4 and pCDFduet.H2A-H2B or pET-Duet.H3-H4 and pCDFduet.H2A(K120C)-H2B [294]. Cells were grown to ~OD 0.4 and expression was induced by adding 0.4 mM isopropyl 1-thio-ß-D-galactopyranoside (Santa Cruz Biotechnology Inc), and shaking at 180 rpm for 16 h at 18°C. Cells were subsequently sonicated in a Qsonica Q500 sonicator for 2 min with cycles of 5 s on and 5 s off and an amplitude of 40%, in 0.5 M NaCl, 20 mM Tris-HCl pH8, 0.1 mM EDTA, 1 mM DTT, 0.3 mM PMSF and protease complete inhibitor (Roche). The supernatant containing the histone complexes was purified on a 5 ml Hi-Trap Heparin

column (Cytiva) and eluted with a gradient of 0.5-2 M NaCl, 20 mM Tris-HCl pH8, 0.1 mM EDTA, 1 mM DTT. Fractions were analyzed using SDS-Page and further purified on a Superdex 200 increase (Cytiva) in 2 M NaCl, 20 mM Tris-HCl pH8, 0.1 mM EDTA and 1 mM DTT, analyzed on a 12% SDS-Page gel, concentrated and then labeled with C2-Maleimide-Alexa647 (ThermoFisher). For labeling, 1 mg/ml histone-octamers were dialyzed in 50 mM MOPS pH7, 2 M NaCl. TCEP (Sigma-Aldrich) was added to 1 mM and 0.2 mM C2-Maleimide-Alexa647 was added. This was incubated for 2 hours at 21°C, quenched with 1 mM DTT and then purified on a Superdex 200 increase column and checked with an SDS-page gel for intact labeled histone-octamers. Nucleosome assembly was done using salt dialysis [216]. For every assembly, different ratios of Alexa647-histone-octamer/DNA were made. 136 nM Alexa647-histone-octamer, ~0.15 nM Biotin-λ-DNA and 2.5-15 nM 10kb DNA (ThermoFisher, SM1751) were combined in 2 M NaCl, 10 mM Tris-HCL pH8, 1 mM EDTA and transferred to a slide-A-lyzer mini dialysis device, 3.5K MWCO (ThermoFisher, 88400). It was then dialyzed at 4°C in 100ml of 2 M NaCl, 10 mM Tris-HCl pH8, 1 mM EDTA, 5 mM β-mercaptoethanol. We added 10 mM Tris-HCl pH 8.0, 1 mM EDTA, pH 8, 5 mM β-mercaptoethanol at a rate of 0.5 ml/min using a peristaltic pump until the buffer reached ~0.4 M NaCl. We then transferred the slide-A-lyzer to a beaker glass with 100 ml 10 mM Tris-HCl pH 8.0, 1 mM EDTA, pH 8.0, 5 mM β-mercaptoethanol and incubated for another 2 hours. Subsequently, we analyzed the sample on an agarose gel and checked on a typhoon if labeled histones were comigrating with the DNA and checked proper nucleosome assembly on the AFM.

After treating a mica with 0.0001 % (wt/vol) poly-L-Ornithine for 1.5 min, we rinsed the mica using 3 ml MilliQ water and dried it using N_2 gas. Then, the sample of the nucleosome assembly was deposited onto the mica for 1 min and rinsed with 3 ml MilliQ water. After drying the sample using N_2 gas, we imaged it using a Bruker Multimode AFM, with a Nanoscope V controller and Nanoscope version 9.2 software [237] using Bruker ScanAsyst-Air-HR cantilevers (stiffness 0.4 N/m, tip radius 2 nm). PeakForce Tapping mode was used with an 8 kHz oscillation frequency. To reduce sample distortion induced by tip-sample interaction, we used low a peak force set-point value (< 100 pN). To obtain high resolution images, 2.5 × 2.5 μm^2 scan areas with 2,500 × 2500 pixels^2 were scanned at 0.5 Hz scanning speed. For imaging condensin and cohesin holocomplexes, we deposited 2 nM protein complexes onto a 0.00001 % (wt/vol) poly-L-Ornithine treated mica for 10 s, and the mica was washed using 3 ml MiliQ water and dried using N_2 gas. Then, the sample was imaged with the same AFM parameters described above. We performed image processing of the AFM images using Gwyddion version 2.53 by removing background subtraction and filtering transient noise [237]. Afterwards, we subtracted (planar and line by line) background polynomials after excluding the masked grains of DNA-protein images and applied plane background subtraction. Finally, we de-convoluted the AFM images using a blind tip estimation and surface reconstruction to minimize the tip convolution effect, the widening of images induced by non-zero AFM tip size. To measure the diameter of nucleosomes (Figure S3.2c,d), we masked each nucleosome complex bound to DNA on the AFM images and performed mean diameter measurement in the Gwyddion software.

Table 3.1 | Target sequences for binding dCas9 to λ-DNA. Rank and efficiency are given by ChopChop. Target sequences include the protospacer adjacent motif (PAM).

Rank	Target sequence	λ-DNA location	Efficiency (%)
16	AAGTGATGCGAAAAAAACAGCGG	seq:10643	75.57
5	TGTATGAAGATTCACAACCGGGG	seq:19879	79.29
3	GAAATCCACTGAAAGCACAGCGG	seq:31232	81.03
12	GCTTGGAACTGAGAAGACAGCGG	seq:38428	76.97

Protein roadblocks - RNA polymerase labeling and binding to DNA

Wild-type *E. coli* RNA polymerase core-enzyme ($\alpha 2\beta\beta'\omega$) with a SNAP-tag and transcription initiation factor σ_{70} were purified and expressed as previously described [295]. Purified RNAP was covalently attached to Alexa647-benzylguanne according to the protocol provided by the provider and further purified using a Sepharose 6 gel-filtration column (Cytiva). The labeling efficiency of the core-enzyme was 35 %. RNAP holo-enzyme was reconstituted by incubating labeled RNAP with σ_{70} in 1:10 ratio in a buffer containing 10 mM Tris-HCl, pH 8, 150 mM NaCl, 0.1 mM EDTA, 5 µm $ZnCl_2$, 1 mM DTT, 5 % glycerol. Prior to anchoring λ-DNA onto the surface, 1 nM λ-DNA was incubated with 10 nM RNAP holo-enzyme in imaging buffer (50 mM TrisHCl pH 7.5, 50 mM NaCl, 2.5 mM $MgCl_2$, 1 mM DTT, 5%(w/v) D-dextrose, 2 mM Trolox, 40 µg/mL glucose oxidase, 17 µg/mL catalase) for 5 min at room temperature.

Protein roadblocks - dCas9 binding to DNA

The protocol for dCas9 binding to DNA was adapted from the protocol described by [296]. gRNA was obtained by annealing a mixture of Alt-R Crispr-Cas9 tracrRNA and crRNA (IDT) at 90°C for 1 min. To prepare dCas9/gRNA construct, 120 nM of dCas9-Snaptag and 1.2 µm gRNA were incubated in NEBuffer3.1 for 30 min at 37°C. The dCas9/gRNA constructs were stored in aliquots at -80°C until further use. crRNA target sequences were chosen using the Chopchop webtool (`https://chopchop.cbu.uib.no/`). The dCas9/gRNA was attached to the target sequences on λ-DNA by incubating 120 pM biotinylated λ-DNA and 12 nM dCas9/gRNA in NEBuffer3.1 for 45 min at room temperature. We call this construct "λ-DNA/dCas9" in short. To add a label on dCas9 attached to λ-DNA, 100 nM SNAP-Surface Alexafluor-647 was incubated with the "λ-DNA/dCas9" for 45 min at room temperature.

Gold and polystyrene nanoparticle functionalization

For each size, 100 µl of gold nanoparticles (Nanopartz) stock was incubated on a shaker at 4 °C overnight with SH-PEG-COOH and SH-PEG-NH_2 with concentrations given in Table 3.2. Unbound PEG was removed by dilution (15x) with MilliQ water, centrifugation (speeds in the table below), and removal of the liquid supernatant after centrifugation. The dilution and centrifugation step was repeated at least five times to get rid of the unbound PEG. The PEG functionalized gold nanoparticles were incubated with 10 mM Alexa647 NHS and 0.1 mM Benzyl Guanine NHS (BG-NHS) in 50 mM borate buffer at 4 °C overnight. The unbound dye and BG-NHS were removed by dilution with MilliQ water, centrifugation and removal of the supernatant. The dilution and concentration step was repeated at least five

Table 3.2 | Concentration of functionalization groups for particle functionalization.

Size (nm)	SH-PEG-COOH (mM)	SH-PEG-NH_2 (mM)	Alexa647-NHS (mM)	BG-NHS (mM)	Centrifugation speed (g)
14	10	5	10	0.1	20000
21	10	2.5	10	0.1	10000
29	10	1	10	0.1	2500
50	10	0.4	10	0.1	1100
125	10	0.2	10	0.1	400
200 (polystyrene)	10	1	10	0.1	400

times to remove most of BG-NHS. We estimate around 1 to 10 BG and around 100 Alexa647 molecules per nanoparticle. The dye-functionalized gold nanoparticles were stored at 4 °C for further use. 200 nm amine-functionalized polystyrene particles (Nanocs) were incubated with 10 mM SMCC (succinimidyl4-[N-maleimidomethyl]cyclohexane-1-carboxylate) in borate buffer pH 8 for 30 mins in an ice bath which leaves maleimide groups on the surface for further reaction with thiol groups. The unreacted SMCC was removed by dialysis. The maleimide functionalized polystyrene particles were immediately functionalized with Alexa647 and BG-NHS as stated in Table 3.2. Functionalized gold nanoparticles were incubated with 0.5 mg/mL BSA for 5 min and then mixed with λ-DNA/dCas9 for 45 min at room temperature at a ratio of 10:1 (particles: DNA). λ-DNAs with nanoparticles bound to them were then used to anchor onto the functionalized glass substrate.

Nanoparticle size measurement from TEM images and FCS

Hydrodynamic diameters were estimated through FCS measurement. Fluorescence time traces and autocorrelations were recorded in a Picoquant MicroTime 200 confocal microscope. The functionalized gold nanoparticles with Alexa647 on their surface were excited with a 640 nm laser with powers giving to 10000 photons per second on average at the detector. The concentration of the particles was kept around 1 nM. The autocorrelation curves (Figure S3.2j) were fitted with

$$G(\tau) = \left[1 - \tau + \tau e^{-\frac{\tau}{\tau_T}}\right] \frac{1}{N(1-\tau)\left(1+\frac{\tau}{\tau_D}\right)} \sqrt{\frac{1}{1+\left(\frac{\tau}{\tau_D}\right)\left(\frac{\omega_0}{\omega_z}\right)^2}}, \tag{3.1}$$

where τ is the lag time, τ_T is the triplet blinking time of the Alexa647 fluorophore, N is the number of molecules in the detection volume, ω_0 and ω_z are the lateral and longitudinal width of the point spread function of the microscope, and τ_D is the diffusion time of the particle. ω_0 and ω_z were determined by calibrating the microscope with the known diffusion coefficient (3.3 $\mu m^2 s^{-1}$) of Alexa647. The hydrodynamic diameter ($2R_h$) was calculated from that as $R_h = \frac{k_B T}{6\pi\rho D}$, where $D = \frac{\omega_0^2}{4\tau_D}$ and T is the temperature, ρ is the viscosity and D is the diffusion coefficient. Functionalized gold nanoparticles were incubated with 0.5 mg/mL BSA for 5 min and then mixed with λ-DNA/dCas9 for 45 mins at room temperature at a ratio of 10:1 (particles:DNA). These λ-DNAs with nanoparticles bound on them were

then used to anchor onto the functionalized glass substrate. We performed TEM and SEM imaging to verify the sizes of the nanoparticles. The functionalized nanoparticles were diluted 1:20 times with MilliQ and then deposited on carbon grids coated copper TEM grids (Ted Pella) using a JEOL JEM-1400 transmission electron microscope with an acceleration voltage of 120 keV. For particles larger than 100 nm, the nanoparticles were deposited on a copper tape and imaged using a FEI Helios G4 CX in SEM mode at 10 KeV.

Single-molecule loop extrusion assay and data analysis

Flow cells were prepared with PEG/PEG-biotin passivated glass slides as described previously [98]. Microscope slides were cleaned with acid piranha (sulfuric acid (5 parts) and hydrogen peroxide (1 part)) and silanized with [3- (2-aminoethyl)aminopropyl] trimethoxysilane in methanol containing 5% glacial acetic acid which leaves free amine groups on the surface. The slides were then treated with 5 mg/ml methoxy-PEG-N-hydroxysuccinimide (MW 3500, Laysan Bio) and 0.05 mg/ml biotin-PEG-N-hydroxysuccinimide (MW3400, Laysan Bio) in 50 mM Borate buffer, pH 9. The pegylated slides were dried with a gentle flow of nitrogen, sealed, and stored at -20o C until further use. Flow cells were then assembled with the functionalized glass slides as described before [98]. The channels in the flow cell were incubated with 100 nM streptavidin in T20 buffer (40 mM Tris-HCl pH7.5, 20 mM NaCl, 0.2 mM EDTA) for 1 min. The unbound streptavidin was washed with T20 buffer. The surface was further passivated by incubating the channel with 0.25 mg/mL Bovine Serum Albumin (BSA) in T20 for 10 min. A volume of 30 µL containing 5 pM of λ-DNA with the DNA roadblock, along with 20 nM SxO in T20, was introduced into the channel at a flow rate of 4 µl/min. The free DNA in the channel were removed by flowing 100 µl T20. The buffer in the channel was replaced with imaging buffer containing 100 nM SxO. The imaging buffer scavenges oxygen and lengthens the observation time of the fluorophores by reducing photobleaching. We visualized the DNA and roadblocks in a home-built Highly Inclined and Laminated Optical sheet (HILO) microscope [297] with a 60x oil immersion, 1.49NA CFI APO TIRF (Nikon) objective. HILO allows the excitation of molecules in a thin sheet of light that is 1-10 µm above the glass substrate; hence reducing the background from free-floating dyes in the solution and increasing signal-to-noise ratio of the SxO on DNA and Alexa647 labeled roadblocks. The SxO on DNA was excited with 0.1 W/cm^2 561 nm laser light and single Alexa647 fluorophores on proteins were excited with 100 W/cm^2 637 nm in alternative excitation (ALEX) mode to simultaneously visualize the DNA and roadblock proteins. The gold nanoparticles with 10 to 100 Alexa647 molecules on the surface were excited with 10 to W/cm^2 637 nm laser light. The two lasers were illuminated alternatively with an exposure time of 100 ms allowing simultaneous observation of the DNA and the roadblocks.

Data analysis

Fluorescence images were analyzed in a custom-written software in python programming language. The images were denoised using a machine-learning-based method called "Noise2void" [298]. Fluorescence intensities along the DNA axis were averaged with 5 pixels on both sides of the axis to obtain a kymograph (e.g. Figure 3.1d). Kymographs for the roadblock were obtained with the same axis as the DNA. The intensities of the kymographs were corrected such that the non-DNA part of the kymograph had zero intensity. Each vertical line on the kymograph represents a snapshot of the DNA at one

time point. Peaks on each line were obtained using the "find_peaks" algorithm in scipy [299] which finds the local maxima by comparing the neighbouring values. The intense peaks from the peaks obtained were selected with >50% threshold of the peak prominence (relative peak intensities). The peak position represents the loop position and the loop intensity (Int_{loop}) was determined by summing intensities from 7 pixels around the peak position. The intensities of regions ahead (Int_{ahead}) and behind the loop (Int_{behind}), and loop-to-roadblock ($Int_{looptoroadblock}$) were determined. The intensities were converted to DNA bases as follows:

DNA size in the loop (bp): $I_{loop} = \frac{Int_{loop}}{Total\ DNA\ intensity}$

DNA size ahead of the loop (bp): $I_{up} = \frac{Int_{ahead}}{Total\ DNA\ intensity}$

DNA size below the loop (bp): $I_{down} = \frac{Int_{behind}}{Total\ DNA\ intensity}$

DNA size roadblock to and including loop (bp): $I_{dist} = \frac{(Int_{loop} + Int_{loop\ to\ roadblock})\cdot}{Total\ DNA\ intensity}$

The mean square displacement of the roadblock was obtained from the positions of the roadblock at different times as below:

$$MSD(t) = \frac{1}{50}\sum_{t}^{t+50} |x(t+1) - x(t)|^2, \tag{3.2}$$

where the number 50 denotes the window size of 50 frames. The loop kinetics were smoothed using the Savitzky-Golay method with a 2nd order polynomial with a moving window of 50-100 points. The result of the encounter between a roadblock and an SMC complex was inferred from both the kinetics and MSD traces. The event was considered an encounter when a roadblock moved towards a loop during loop extrusion and colocalized with the loop. The encounter was inferred from the kymograph. We categorized the results of the encounters into three options: i) the roadblock ended up inside the loop, corresponding to "passing", ii) it remained at the stem of the loop corresponding to "blocking", and iii) a transient pause occurred at the stem of the loop followed by passing into the loop, corresponding to "stalling and passing". An increase in both loop size and MSD after the encounter indicated passing. A constant loop size and MSD upon encounter indicated blocking. An increase in loop size and a constant MSD was also observed occasionally (particularly for cohesin) and was counted as blocking. A constant loop size and MSD, occurring for more than a second, followed by an increase in both corresponded to stalling and passing. Dissociation of roadblocks upon encounter with SMC complex was observed only very rarely.

Human cohesin and single-chain cohesin experiments

Recombinant human cohesin^{STAG1}, NIPBL-Mau2, STAG1 and single-chain cohesin were expressed and purified as described previously [173]. In brief, a cell pellet was thawed and lysed by Dounce homogenization. The soluble fraction after centrifugation was combined with Toyopearl AF-chelate-650M resin (Tosoh Bioscience) pre-charged with Ni^{2+} ions and incubated for 3 h. Beads were washed and supplemented with 15 mM imidazole pH 7.5, 0.01% Tween20. Bound protein was eluted in cohesin purification buffer (25 mM sodium

phosphate pH 7.5, 150 mM NaCl, 300 mM imidazole, 5 % glycerol, 0.01% Tween20). The eluate was combined with FLAG-M2 agarose resin (Sigma; A2220) and incubated for three hours. Elution was made with 25 ml of cohesin storage buffer (25 mM sodium phosphate pH 7.5, 150 mM NaCl, 50 mM imidazole, 5 % glycerol), concentrated, aliquoted, snap-frozen, and stored at -80°C. Recombinant cohesin was diluted to 10 nM and NIPBL-Mau2 was diluted to 50 nM in ice-cold storage buffer on ice and briefly mixed by pipetting. Cohesin and NIPBL were then immediately added to the imaging buffer (40 mM Tris pH 7.5, 50 mM NaCl, 2.5 mM MgCl2, 5% glucose, 0.25 mg/ml BSA, 1 mM DTT, 0.05% Tween20, 2 mM ATP) supplemented with the oxygen scavenging buffer as described above, equilibrated to 370C. Imaging buffer supplemented with protein was then immediately introduced in the flow cell at 10 - 30 pM final concentration with a 4-fold molar excess of NIPBL-Mau2 at a high flow rate (20-40 µl/min) for 30 sec and flow was turned off to observe loop extrusion events. Single-chain cohesin was crosslinked by incubation with 0.23 mM BMOE for 10 min on ice. The reaction was quenched by the addition of DTT to 10 mM. For crosslinking controls, BMOE was either omitted or BMOE was quenched with DTT before the addition of single-chain trimer (Figure S3.3a). Crosslinked single-chain cohesin was then incubated with 12-fold molar excess of recombinant STAG1 on ice for at least 30 min (up to 8 h). Single-chain cohesin was then added to imaging buffer at 10-30 pM with 4-fold molar excess of NIPBL-Mau2 as for recombinant human cohesin. To check the crosslinking efficiency, 500 fmol of non-crosslinked and crosslinked cohesin were run on a 6% Tris-Glycin protein gel and stained by Coomassie. All experiments with recombinant human cohesin and single-chain cohesin were performed at 37°C.

Analysis of cohesin-roadblock encounters

Loop extrusion by recombinant human cohesin and single-chain cohesin has been reported to be two-sided [173, 212, 222] (and chapter 5), in contrast to loop extrusion by yeast condensin that extrudes DNA asymmetrically [98]. This fact makes analysis of roadblock encounters with cohesin slightly more involved compared to encounters of roadblocks with condensin. For example, we found that cohesin has a tendency to start extruding loops close to the roadblock on the DNA ($\tilde{6}$0% of loops initiate at the position on the DNA that colocalizes with a roadblock). For loops which initiate sufficiently far from the roadblock, moving of the loop toward the roadblock makes it unequivocally possible to determine which side of the loop encounters the roadblock. However, this is not possible for loops initiating too close to the roadblock. Furthermore, two-sided extruders may keep extruding DNA into the loop from the non-blocked side, thus making loop growth not a sufficient proxy to classify encounters as passing or blocking. To carefully dissect roadblock encounters with two-sided loop extruders, we distinguish four scenarios, which are illustrated in Figure S3.3b. Blocking of the loop extrusion process (case 1) is demarcated by a colocalized DNA loop and roadblock puncta in the kymograph, by halting loop growth and by a decrease of roadblock MSD during initial loop extrusion, and a continuously low MSD as the roadblock is located at the loop base. This case is analogous to the blocking scenario for one-sided extruders. Alternatively, cohesin loop extrusion may be blocked on the side of the roadblock, but continue to extrude on the other side (case 2), in which case the DNA loop and roadblock appear to translocate together in the kymograph. The loop keeps growing, yet the MSD of the particle decreases steadily as the loop grows, which allows the assignment of the event as blocking, rather than passing as would be

judged from the growing loop intensity on its own. A similar argumentation holds for passing events. In the case where the loop continues to grow on the side of the roadblock (case 3), the loop and roadblock translocate together, the loop grows and the MSD shows a characteristic minimum at the moment of encounter. For situations in which the loop grows or slips from both sides (case 4), the loop and MSD translocate together in the kymograph, yet not necessarily unidirectionally, and the loop size can both grow and shrink. In these cases, a passing event is characterized by a minimum in the MSD as the loop size either remains constant (loop grows on the roadblock side and shrinks on the other side, effectively keeping the loop size constant; we refer to this as 'loop diffusion', see Figure S3.3f) or increases. A minimum in MSD might also be associated with a decrease in loop size (slipping of DNA out the loop), which releases tension in the DNA, in which case the event cannot be attributed as 'passing'. Thus, the combination of a minimum in MSD and a kymograph with a constant or rising loop size is used as a proxy to distinguish passing and blocking events.

Quantification and statistical analysis

The fitting of curves (e.g. Figure 3.3e) was done with Scipy package in Python (version 3.8). Smoothing of data (e.g. Figure 3.1j, k) was done with interpolation by Savitzky-Golay method in *scipy*. Error bars (Figure 3.4b) of binomial distributions were calculated using binomial proportion confidence intervals. Violin plots showing the kernel probability density were obtained using the *seaborn* package in python. Plots were created using *matplotlib* and *seaborn* packages.

3.6 SUPPLEMENTARY INFORMATION

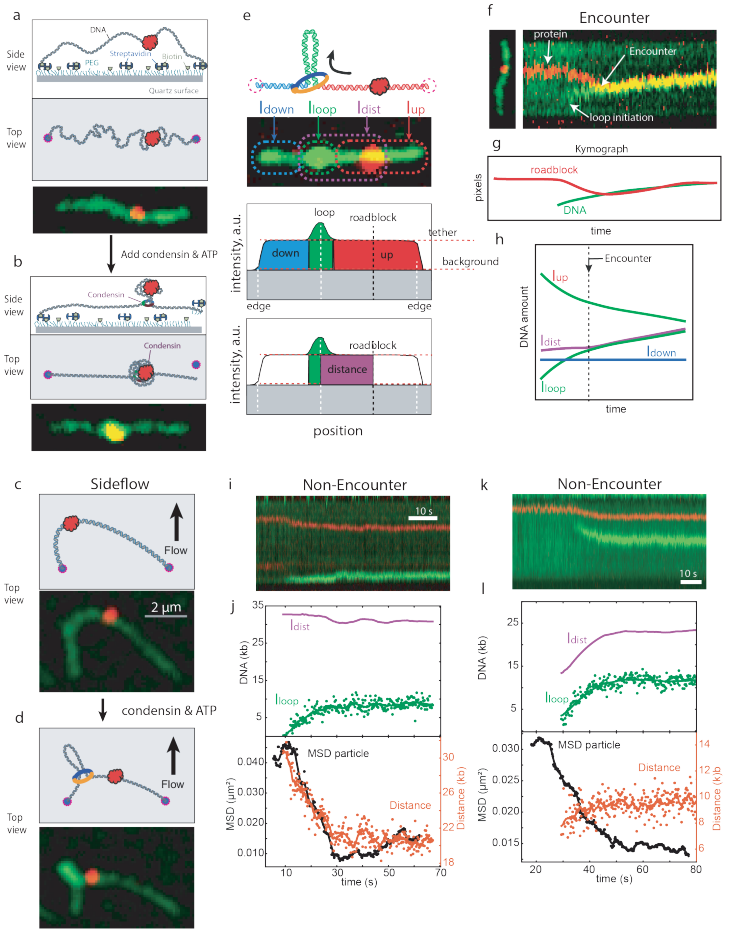


Figure S3.1 | Experimental assay and types of encounters.
(**a**) Side view and top view of lambda DNA double tethered on a streptavidin/PEG-coated surface with a fluorescence snapshot of the DNA (green) with the roadblock (red) which is strongly bound to the DNA through dCas9. (**b**) Side view and top view of an SMC-extruded DNA loop after addition of condensin and ATP. (**c**) Top view of the DNA loop extrusion with inplane buffer flow. (**d**) Addition of ATP and condensin resulted in a loop and the roadblock moving towards the stem of the loop. (**e**) Representation of intensities of different parts of a loop-extruding DNA molecule with a roadblock. (**f**) Kymograph of the loop position and roadblock position along the axis of the DNA, with x-axis as time and y-axis as pixels. (**g**) Kinetics of intensities of different parts of the DNA as represented in (**h**). The kymograph (**f**) and the schematic of its kinetics (**g, h**) represents an encounter of the roadblock with the loop. (**i-l**) Examples of condensin extruding DNA that has a dCas9 roadblock bound, but where the SMC did not meet the roadblock, and hence the event is not counted as an encounter. (**i, j**) Condensin extruded DNA from the side that contained the roadblock but never reached it due to the stalling force that emerged within the DNA. (**k, l**) An example of a roadblock that was on the anchor side of condensin and hence condensin never encountered it.

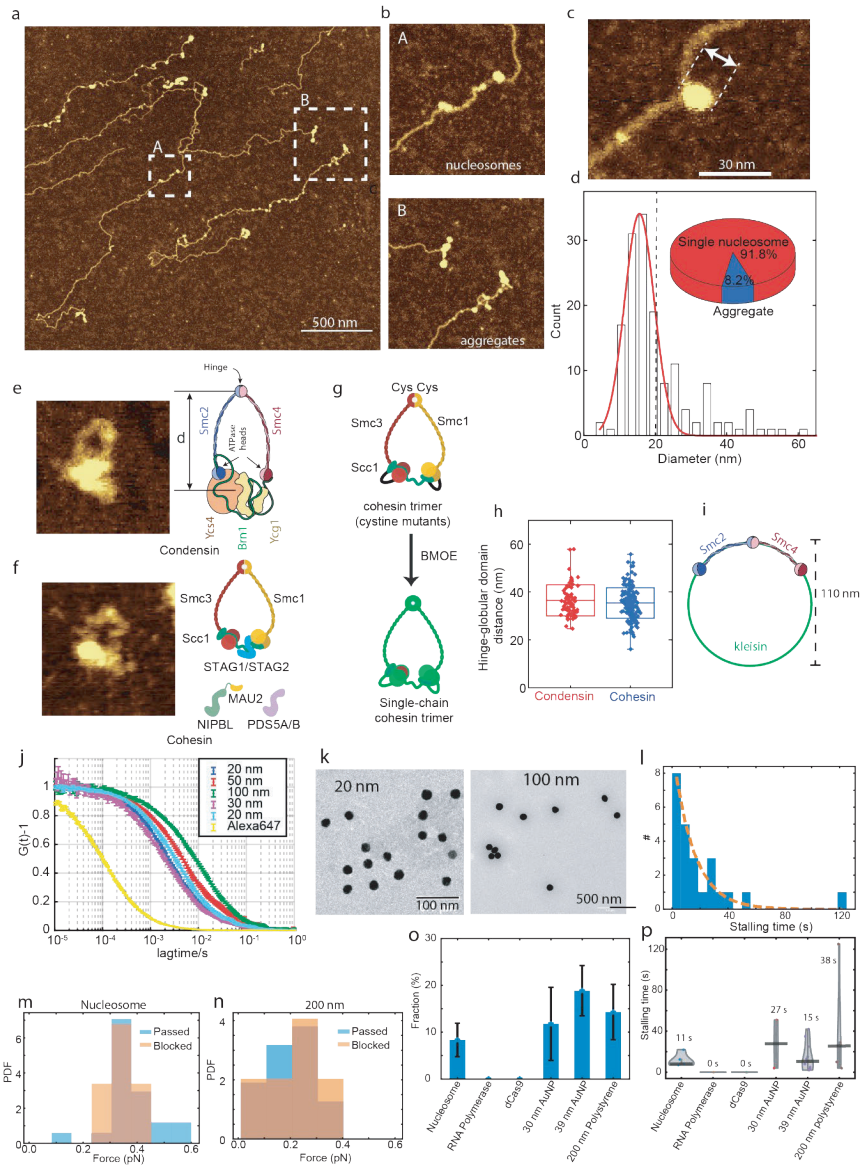


Figure S3.2 | Loop extrusion efficiencies as a function of the roadblock sizes and SMC complexes. (**a**) AFM image of nucleosomes on DNA. (**b**) Zooms of the rectangles shown in (a) for single nucleosomes (A) and some small aggregates (B). (**c**) Zoom in on a nucleosome with annotation to show the measured diameter. (**d**) Histogram of the measured diameter of nucleosomes. The solid red line is a single-Gaussian fit yielding a maximum at 15.6 nm. If conservatively, we assign all structures beyond the dashed line as aggregates, we observe that 92% of the nucleosomes are single nucleosomes, see inset, indicating that aggregation was minimal. AFM images of O-shaped condensin (**e**) and cohesin (**f**) with their schematic showing to the right. The AFM images are 100 x 100 nm. The diameters of both condensin and cohesin are distributed around a mean value of about 36 nm, as denoted in panel (h). *Caption continues on next page.*

Figure S3.2 | **(i)** Sketch indicating the theoretically maximal circumference that an SMC trimer would be able to attain if the kleisin were entirely stretched to its full length. The Smc2 and Smc4 arms are known to be about 50 nm in length each. The kleisin is an unstructured 700 amino acid chain that can, using 0.35 nm/amino acid, be estimated to have a 245 nm theoretical maximum length. This would suggest a maximum possible diameter of about 110 nm for the SMC ring including kleisin in this nonphysiological configuration. **(j)** Fluorescence correlation curves for functionalized gold nanoparticles exhibited a slower decay as the size of the nanoparticle increased. The hydrodynamic diameters that were obtained from these curves (see Methods section 3.5 for details) are denoted in the main text. **(k)** TEM images of functionalized gold nanoparticles of different sizes showed uniform distributions and the absence of aggregates. **(l)** Histogram of all stalling times for all the roadblocks shown in (q). The dashed line is an exponential fit with a time constant of 14 s. Force histograms for DNA loop extrusion with nucleosomes **(m)** and 200 nm particles **(n)** as roadblocks. Data in blue denote the stalling force for loop extrusion, and data in orange denote the force at which loop extrusion got blocked by the roadblocks. The very similar force values with and without roadblock suggest that blocking is mostly due to a lack of slack on the DNA. **(p)** Fraction of total encounters between condensin and roadblocks that show stalling and passing. **(q)** Stalling times for different roadblocks.

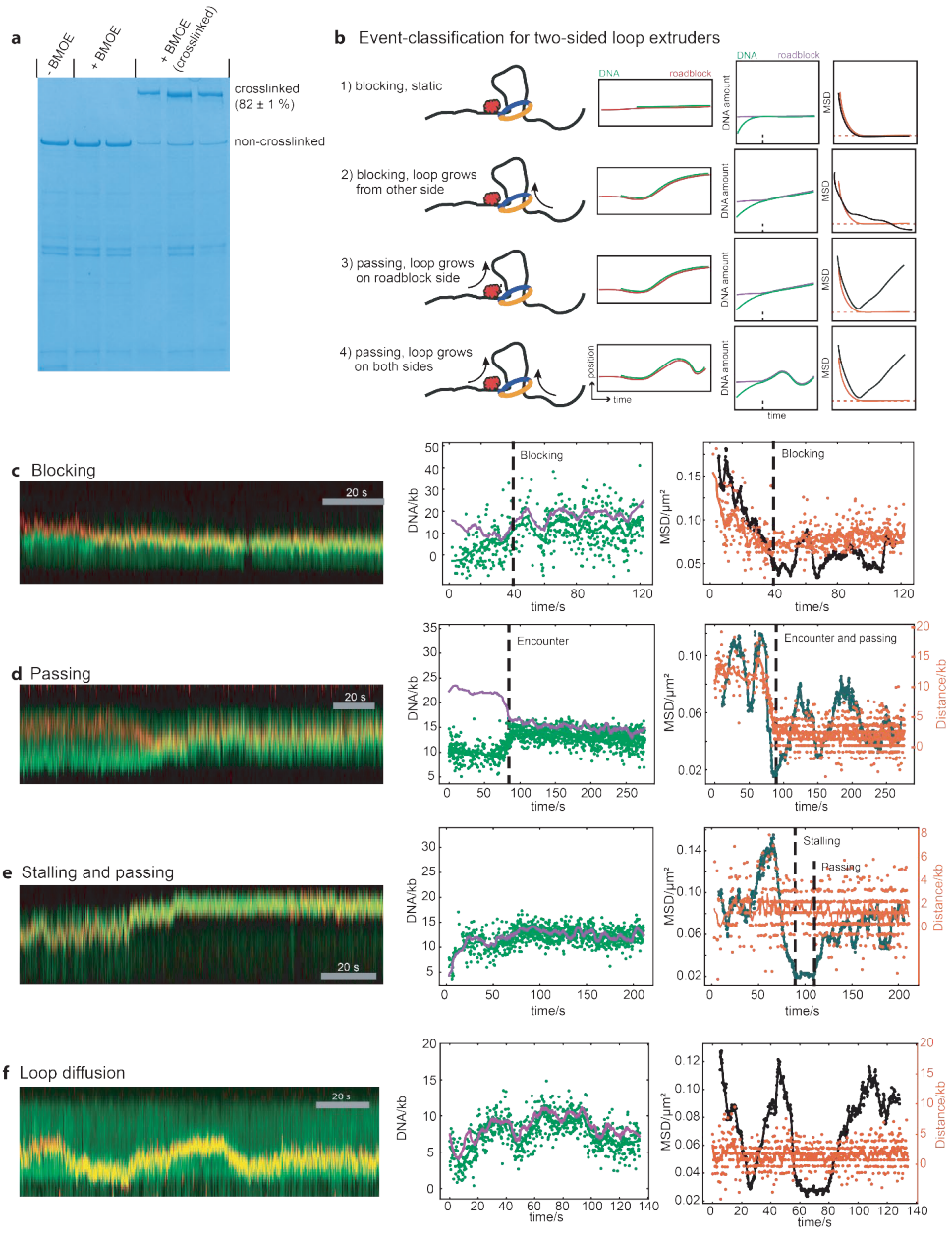


Figure S3.3 | Loop extrusion data for human cohesin on DNA with 30nm gold particles that act as roadblocks.

(**a**) Gel shift of crosslinked cohesin and wildtype cohesin showing an 82% crosslinking efficiency. (**b**) Illustrative classification of encounter events for two-sided loop extrudes, in particular for cases in which the loop initiates close to or at the roadblock: blocking without further loop growth (1), blocking but loop grows further from the non-blocked side of the DNA (2), passing while the loop only grows from the side of the roadblock (3), and passing while the loop grows and/or shrinks from both sides (4). See Methods section 3.5 (**c, d, e**) Examples of kymographs and position/MSD traces for loop-extruding cohesin that is getting blocked, passed, and stalled and passed when encountering a roadblock particle, as well as an example where the loop merely diffuses.

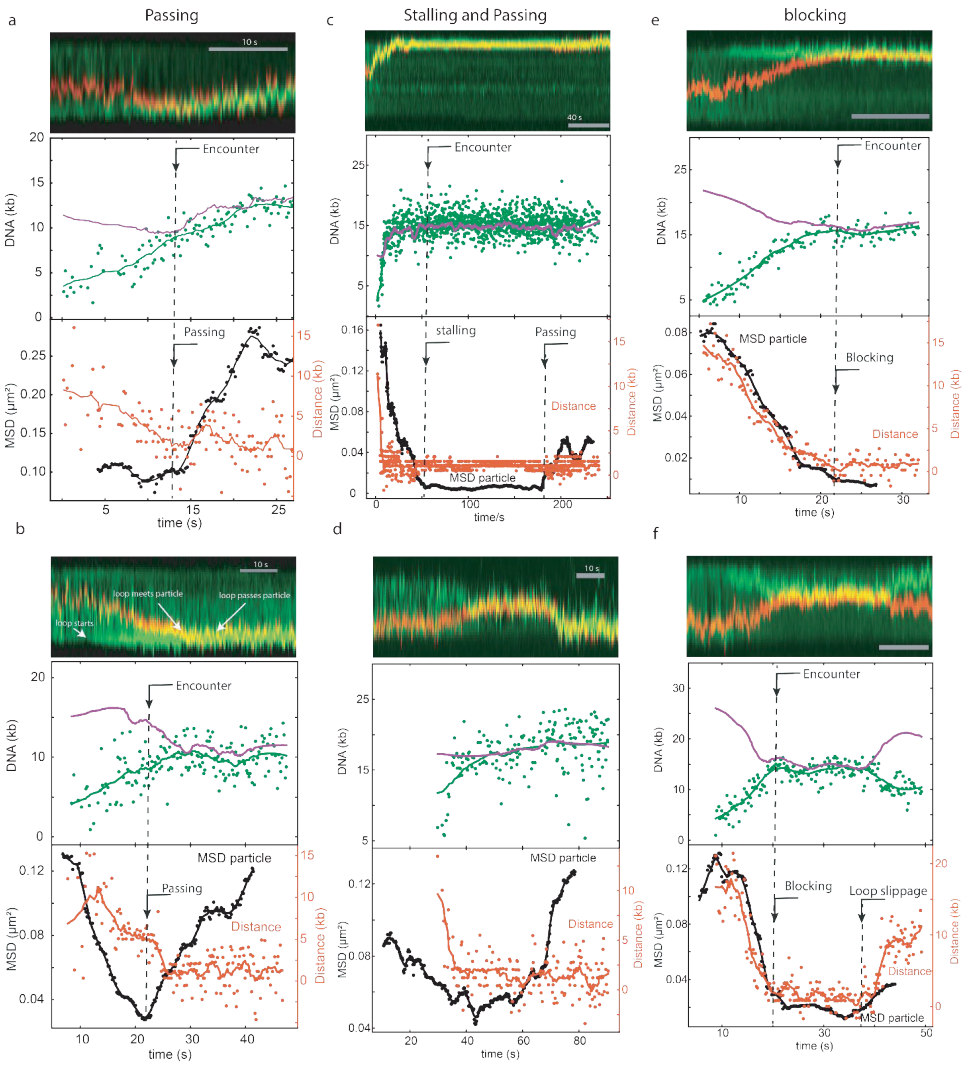
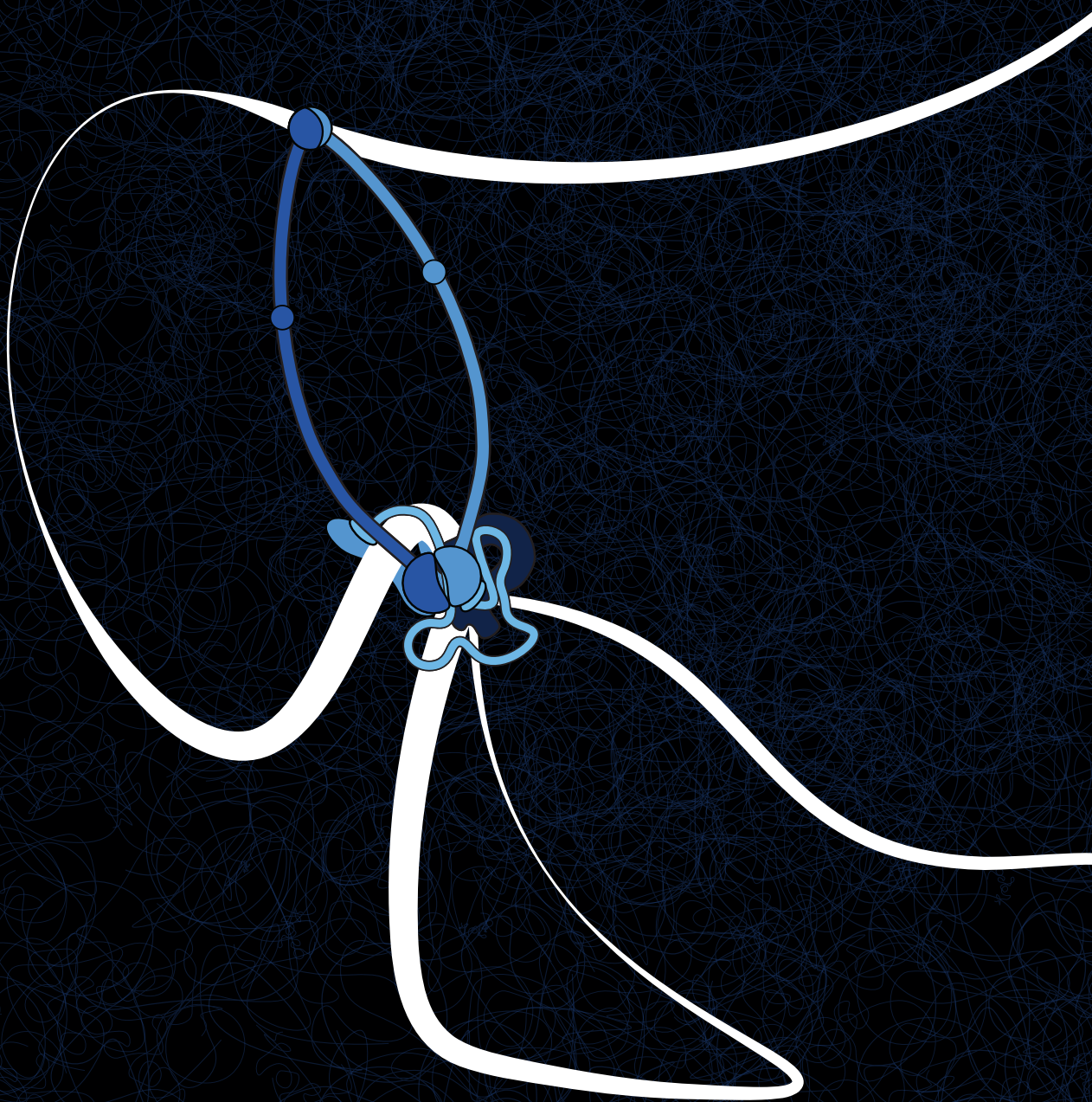


Figure S3.4 | More examples of 200 nm polystyrene beads that act as a roadblock for yeast condensin. (**a,b**) Passing event; (**c,d**) stalling and passing; (**e,f**) blocking.



4

The topology of DNA loop extrusion

Testing pseudotopological and nontopological models for SMC-driven DNA loop extrusion against roadblock-traversal experiments

This chapter has been published as: Roman Barth*, Biswajit Pradhan*, Eugene Kim, Iain F. Davidson, Jaco van der Torre, Jan-Michael Peters, Cees Dekker: Testing pseudotopological and nontopological models for SMC-driven DNA loop extrusion against roadblock-traversal experiments, Sci Rep. 2023 May 19;13(1):8100 [244].

Abstract

DNA loop extrusion by structural-maintenance-of-chromosome (SMC) complexes has emerged as a primary organizing principle for chromosomes. The mechanism by which SMC motor proteins extrude DNA loops is still unresolved and much debated. The ring-like structure of SMC complexes prompted multiple models where the extruded DNA is topologically or pseudotopologically entrapped within the ring during loop extrusion. However, recent experiments showed the passage of roadblocks much bigger than the SMC ring size, suggesting a nontopological mechanism. Recently, attempts were made to reconcile the observed passage of large roadblocks with a pseudotopological mechanism. Here we examine the predictions of these pseudotopological models and find that they are not consistent with new experimental data on SMC roadblock encounters. Particularly, these models predict the formation of two loops and that roadblocks will reside near the stem of the loop upon encounter - both in contrast to experimental observations. Overall, the experimental data reinforce the notion of a nontopological mechanism for extrusion of DNA.

4.1 Introduction

Higher order genome organization is largely dictated by structural maintenance of chromosome complexes (SMC) complexes [279, 282]. SMC complexes including condensin, cohesin, and SMC5/6 in eukaryotes, share a ring-like architecture consisting of multiple subunits where two coiled-coil SMC arms and an intrinsically disordered kleisin constitute a large (35 nm) ring. Additional subunits attach to the kleisin and ATPase heads of the SMC subunits, rendering different functions to the different SMC complexes. SMC proteins organize the chromosome by DNA loop extrusion, as demonstrated by single-molecule in vitro experiments [98, 173, 212]. However, the mechanistic understanding of the coordinated movement of SMC subunits that underlie loop extrusion is still in its infancy.

The topology of DNA involved during loop extrusion has been a major point of discussion in recent years [215, 227, 282, 300]. Three possibilities have been proposed: i) topological binding, where the SMC ring opens to encircle the DNA, ii) pseudotopological loading, where a small nascent loop enters the SMC ring without opening, and where loop extrusion continues pseudotopologically by further extruding this inserted loop through the ring, iii) nontopological loading where the DNA does not enter the ring to extrude loop but instead binds at the outer interfaces of the SMC ring.

Recently, we showed that condensin and cohesin can bypass very large (even 200 nm) DNA-bound proteins and particles while extruding DNA loops [226]. Upon encounter of such a DNA-bound particle by the loop-extruding SMC complex, the particle transfers into the extruded DNA loop in the majority of cases. We even observed such passage of large particles by single-chain cohesins where all interfaces of the SMC subunits and the kleisin were covalently linked [173]. The passage of roadblocks bigger than the ring size of cohesin, even when the ring cannot open, clearly indicated that the extruded DNA does not enter the ring topologically or pseudotopologically. Instead, these experiments pointed to a nontopological mechanism for loop extrusion.

After these data appeared on bioRxiv in July 2021 [226], two new pseudotopological models were proposed by Shaltiel et al. (“hold-and-feed mechanism” [221]) and by Nomidis et al. (“segment-capture model” [242]). In light of the roadblock experiments, the authors of these models pursued to explain the passage of the huge roadblocks within the context of their models.

Here we provide a detailed evaluation of these models with regard to the passage of large DNA-bound roadblocks upon encounter with a loop-extruding SMC complex. We spell out particular predictions that result from these models and evaluate whether or not these predictions are confirmed by experimental observations. We show that the predictions are not consistent with the experimental data for encounters of a loop-extruding SMC complex with DNA-bound roadblocks. The analysis points to a nontopological mechanism instead of a pseudotopological model for loop extrusion.

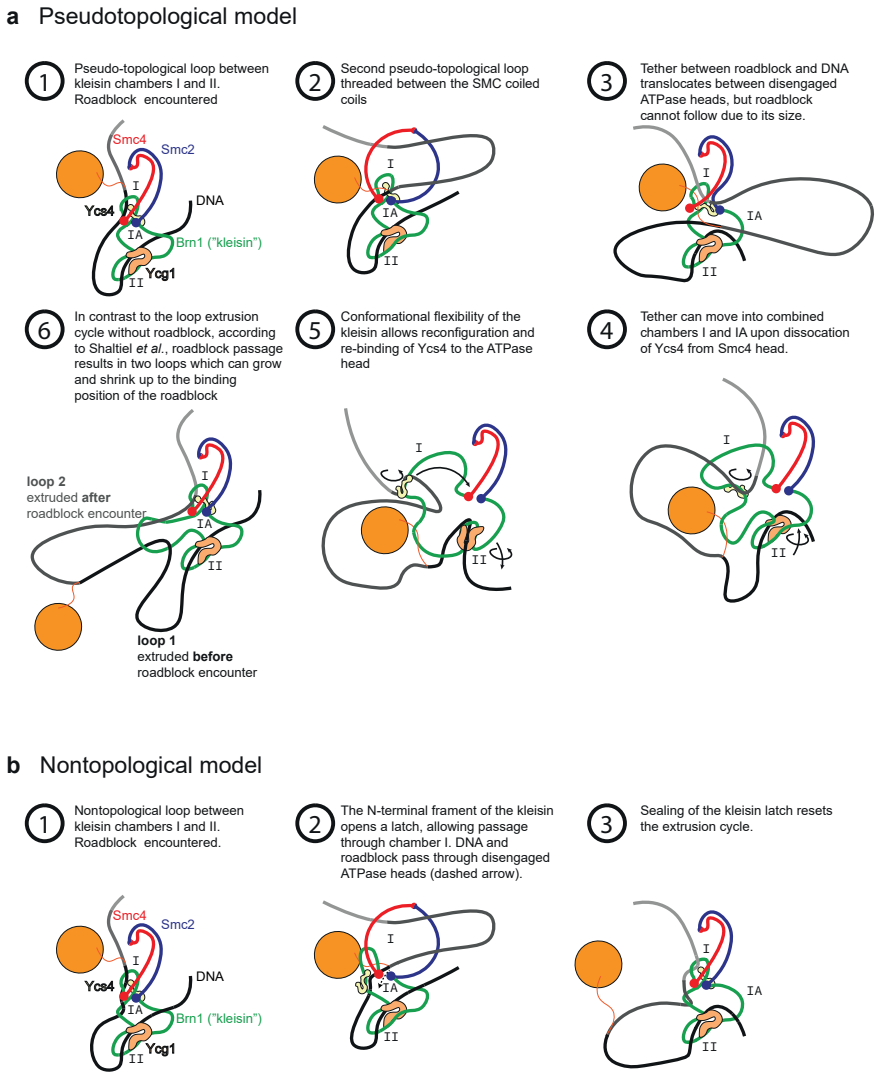


Figure 4.1 | Description of the mechanism postulated by Shaltiel et al. for roadblock passage into an extruded loop on the DNA and a potential nontopological model.
(**a**) The steps through the proposed DNA loop extrusion cycle are commented in more detail in steps 1-6 within the figure. Adapted from [221]. (**b**) Potential nontopological model which is closely analogous to the pseudotopological model, but with a slight variation in the DNA-SMC topology which allows particle bypass.

4.2 Results

4.2.1 Description of the pseudotopological models

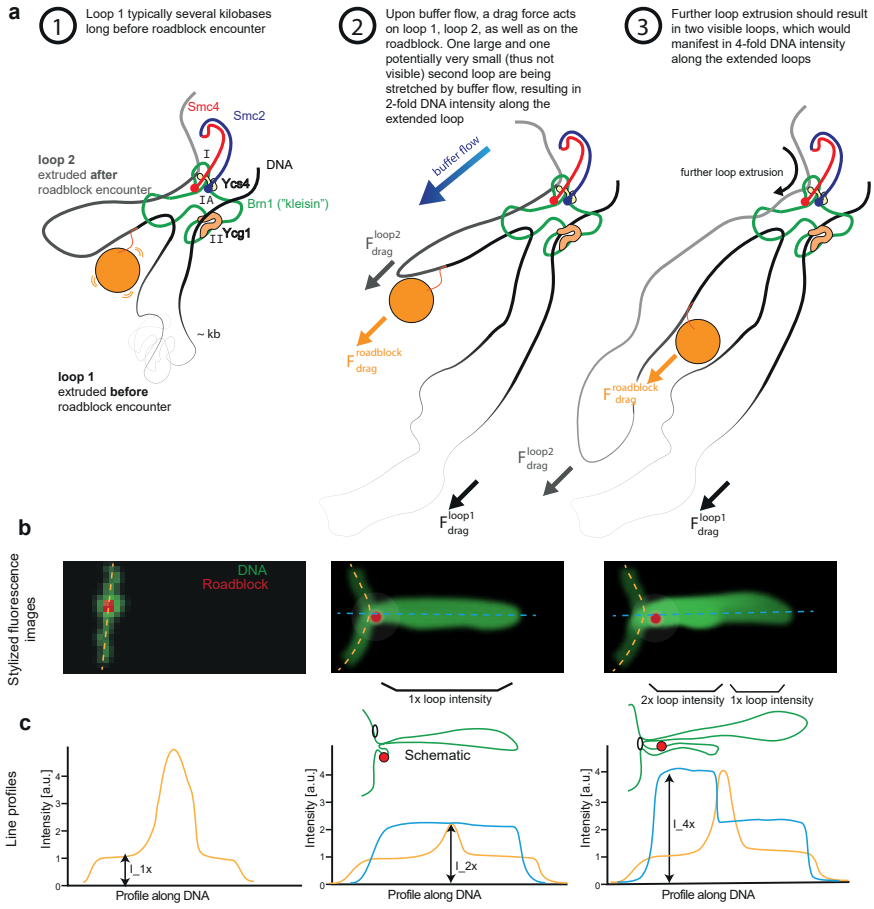


Figure 4.2 | Two (not one) loops are associated with the pseudotopological model.
(**a**) Presence of two loops during side-flow. (**b**) Predicted stylized intensity images (not real experimental data) of the double tethered DNA (green) with the roadblock (red). (**c**) Expected intensity profiles along the dashed lines in the images of panel b.

According to the pseudotopological models proposed by Shaltiel et al. [221] and Nomidis et al. [242], SMC complexes entrap DNA pseudo-topologically, i.e., the two DNA strands that form the base of a loop pass through the ring structure formed by the SMC and kleisin subunits of these complexes. In view of the clear observation of the passage of roadblocks bigger than the SMC ring size into extruded loops in our experiments, the authors of these papers attempted to explain our observations within the context of their models. In other words, they provided an alternative interpretation of our data reported in [226]. The authors of the papers argued that their model was not necessarily inconsistent with our observation that large roadblocks are able to enter extruded loops, even though

these roadblocks are larger than the ring size of SMC complexes. If they are correct, it would question our conclusion that the data point to a nontopological mechanism for DNA loop extrusion by SMCs. Below, we evaluate this alternative explanation in light of our experimental observations.

First, let us summarize the alternative model that Shaltiel et al. and Nomidis et al. proposed [221, 242]. We focus on the model of Shaltiel et al. who most explicitly pointed out the dynamics involved in the roadblock passage, according to their model, but the same considerations apply to the model by Nomidis et al. Without explicitly commenting on the detailed SMC conformational changes involved in these models (which also differ between these models), we describe how these authors pursue to explain the passage of roadblocks bigger than the ring size of SMCs. Figure 4.1a illustrates the model for roadblock passage by Shaltiel et al. Briefly, the model for loop extrusion involves the trapping of DNA in three chambers (I, IA, and II) in yeast condensin, where extrusion is proposed to proceed through a concerted action called hold-and-feed mechanism. DNA binding to Ycg1 and the associated kleisin region act as a "safety belt". DNA loading into chambers I and II initiates extrusion of DNA. Two pseudo-topologically entrapped DNA loops are thus formed; one between chambers I and II, and the second one threaded between the SMC coiled coils. Upon ATP hydrolysis, these two pseudo-topological loops merge into one loop (loop-1), and the SMC starts extrusion of the next segment of DNA. Note that the figure illustrates only tiny loops, whereas in our experiments the encounter of a large roadblock with a DNA-extruding SMC is observed when the loop has been already significantly extended. Specifically, this means that the loops sketched in most panels of the figures will in practice be a very large loop of many kbp of DNA in our experiments.

For comparison, Figure 4.1b depicts a potential nontopological model which is consistent with the crosslinking data by Shaltiel et al. while it also allows roadblock passage [227]. The basic mechanism is analogous to the pseudotopological model of Shaltiel et al. in the sense that a small DNA loop is temporarily reeled through the Smc lumen, triggered by binding of DNA on the engaged ATP heads. The N-terminal kleisin however can latch out which allows traversal of a DNA-bound protein through chamber I. Disengaging the ATP heads allows the roadblock to merge with the large, previously extruded, loop. Sealing of chamber I resets the loop extrusion cycle.

Rather than commenting on the plausibility of the various steps in these models, we here make predictions of what should be observed when a roadblock particle is encountered in the case that the pseudotopological model would hold. Subsequently, we turn to the experimental observations and evaluate whether or not these predictions are confirmed by the observations.

4.2.2 Predictions by the pseudotopological models

Specifically, the alternative models make the following predictions for the roadblock experiments:

Prediction 1: Appearance of two loops in side-flow visualization and

Prediction 2: Roadblocks will reside near the stem of the loop upon encounter

When condensin encounters a large roadblock, the model predicts the existence of two loops (Figure 4.2a), not one, since the roadblock - with its size that cannot pass the ring structure - obstructs the merging of the two preformed loops described above. Next to the already formed extruded loop, a new second loop is therefore formed, where the latter loop has the roadblock particle attached.

One could assume (as Shaltiel et al. appear to do without further consideration) that the preformed loop-1 might slip into loop-2, to again form 1 joint loop. This however is not the case, as can be seen from considering the force balance (Figure 4.3a): If side-flow is applied, several forces in action can be distinguished:

(i) A drag force on the particle $F_{drag}^{roadblock} = 6\pi\eta Rv$, where the viscosity is $\eta = 1\ mPa\ s$ for a water-based buffer, R is the particle radius, and v the buffer flow speed. We experimentally measured the flow speed in side-flow experiments from occasionally observed beads that were not attached to DNA and were flowing through the field of view. We measured a flow speed of 79 ± 26 µm/s (n=3 independent experiments). The observed forces for different particle sizes are shown in Figure 4.3b,c and the values are below 0.15 pN for the largest 200 nm particles, and smaller for smaller roadblocks, down to 0.02 pN for a 30 nm particle.

(ii) Drag force on loop-1 (F_{drag}^{loop-1}). This can be estimated by calculating the tension on a flow-stretched DNA, as described by Pedersen et al. [301] (see Methods section 4.4). Practically in our experiments, by the time that a roadblock encounters condensin, the size of loop-1 is observed to typically exceed 10 kilobases, which is associated with a drag force of value more than 0.27 pN (Figure 4.3c).

(iii) Drag force on loop-2 (F_{drag}^{loop-2}). As loop-2 is a very small nascent loop (~100 bp), the force acting on it is initially negligible, but it may become more significant as the loop continues to grow. The method to calculate the drag force acting on loop-2 is then analogous to loop-1.

It follows that upon encounter the drag force on loop-2 is nihil, while the drag force on loop-1 dominates and is larger than the drag force on the particle. As a result, the pre-encounter-formed loop-1 remains stable, while the particle remains localized close to condensin.

As loop extrusion proceeds, loop-2 grows and is slowly stretched alongside loop-1. Because of limitations in optical resolution, it is challenging to mutually distinguish the two loops spatially. However, the intensity of the two loops can be easily distinguished from that of a single loop by drawing intensity profiles across the loop during sideflow. Figure 4.2b and c illustrate what, according to the pseudotopological model, would be expected in stylized intensity images and intensity profiles along the DNA for the pre-encounter and post-encounter loops. In summary, the pseudotopological model predicts the observation of two loops that should be observed in a double intensity of that of the pre-encounter-formed loop-1 (prediction 1). The nontopological model instead predicts that the roadblock can be passed by SMC complexes, and accordingly only one loop forms (Figure 4.1b).

Furthermore, the pseudotopological model predicts that the roadblock, resides at the stem

of the loop after encounter (prediction 2), while the second loop keeps growing. It is of interest to note two features connected to this prediction: First, the effects of spontaneous Brownian motion can be estimated to be small. $F = -\Upsilon v + \delta F(t) \approx 1\,fN$; where $\Upsilon = 6\pi\eta R$, η is dynamic viscosity, R the radius of the particle, and the typical velocity of a 10 nm particle is conservatively estimated at its upper bound, which is ~1 µm/s for an active motor [302]. The fluctuating force is on average zero, $\langle \delta F(t) \rangle = 0$, but the autocorrelation function can be given as $\langle \delta F(t)\ \delta F(t+\tau) \rangle 2\Upsilon kT\delta(\tau)$, which leads to $\langle \delta F(t)^2 \rangle 2\Upsilon kT \approx (1\ fN)^2$. The Brownian forces to move the particle away from the SMC are thus very low, in the order of a few fN, so an appreciable spontaneous drifting away of the particle is very unlikely. Second, on long time scales, when the second loop would grow to a size comparable to the first loop (~10 kbp), the drag force on the particle plus that on loop-2 may exceed the drag force on loop-1, which may initiate a slipping process. However, this would occur only rarely in our assay due to the finite total length of the DNA which limits the overall growth of the loop. Indeed, the typical loop size of the loop at the moment of encounter (i.e. loop-1) was on average 12 kb in our experiments, while the added amount of DNA after that (i.e. loop-2 in the pseudotopological model) was 6 kb on average.

Summing up, according to the pseudotopological model, one would expect that the roadblock resides at the stem of the loop after the encounter (prediction 2). In contrast, only one loop is predicted to appear in the nontopological model (Figure 4.1b), and the roadblock is expected to move away from the stem of the loop as the loop continues to grow.

Prediction 3: Loop size will increase while the mean squared displacement may stay constant after a roadblock encounter

What happens when the loop extruding SMC encounters a roadblock on DNA? We note that the hold-and-feed model is not considering that loop extrusion will be blocked in this case, but rather that the roadblock ends up in a second loop. In the absence of side flow, upon encounter with the roadblock (as explained in the previous point), loop-2 with the roadblock particle grows through extrusion of DNA, but the particle stays close to the SMC due to the higher drag force on loop 1 than on the particle. This makes a specific prediction for the observation of simultaneously a continued growth of the loop size without, however, any significant increase in the mean squared displacement (MSD) of the particle, as illustrated in Figure 4.4a. The nontopological model instead predicts that the roadblock is in the single extruded loop which enlarges further upon extrusion after roadblock passage (see previous point). Hence, as the loop grows, the MSD of the particle is expected to grow.

4.2.3 Experimental tests of the predictions made above

Our previous study [226] presented an abundant set of experimental observations of encounters of a loop-extruding SMC complex and a DNA-bound roadblock. Here, we test whether or not the above predictions of the pseudotopological model are confirmed by the experimental observations. For the latter, we refer to the published data [226], but we here also present additional new data that were all acquired with the same methodology and under the same conditions as in ref. [226].

Test of Prediction 1. To test whether or not the extruded DNA forms one or two loops when a roadblock is encountered, we measured intensity profiles of DNA at different stages of loop extrusion after a roadblock encounter, see Figure 4.5. The intensity of a single loop (Figure 4.5a,b) was deduced from the intensity profile on the DNA loop *before* the encounter between roadblock and the loop. The intensity of a single loop (i.e. I_2x) and the predicted putative double loop (I_4x) is shown in the lines in Figure 4.5c. We observe that the maximum intensity of the DNA equals the value of the intensity I_2x of a *single* loop while it never comes close to the predicted doubled loop intensity I_4x at any stage during loop extrusion at and after the encounter. From these side-flow visualizations, we thus did not obtain any indication of a doubled loop intensity (e.g. Figure 4.5a-g) - in contrast to the prediction of the pseudotopological model. Our data are therefore not supporting the key predictions made by the pseudotopological model, but instead support the nontopological model.

Test of Prediction 2. The hold-and-feed model predicts that roadblocks will reside near the stem of the loop. This is, however, not observed. Instead, soon after the encounter, the roadblock is observed to be located within the extruded loop and separated from condensin or cohesin, see main figures from Pradhan et. al. [226] and Figure 4.5a,f,h,i.

We can illustrate this in more detail by an example given in Figure 4.6. Here, the loop contained 24 kb of DNA before encounter with the roadblock. Upon encounter, the loop-extrusion speed did not abruptly change but loop extrusion continued monotonously at an approximately constant rate until the loop disrupted when it reached its final size of 30 kb. For this example, Shaltiel et al. predict the appearance of a second loop (loop-2) of 6 kb. However, no increased intensity is observed close to the stem of the loop, see Figure 4.6b-e. The pseudotopological model further predicts that the roadblock initially resides at the stem of the loop until the cumulative drag force on loop-2 and roadblock exceeds the drag force on loop-1. Note that the drag force on the 24 kb-long loop-1 is very large, ~0.65 pN (Figure 4.3d), while the drag forces on the fully extruded loop-2 remain low at ~0.16 pN and the drag force on the particle is estimated at ~0.03 pN (Figure 4.3b-d). Hence, one does not expect any slippage of loop-1 into loop-2. The roadblock should therefore, according to the pseudotopological model, reside at the loop base while the loop continues to grow. However, the roadblock was experimentally observed to clearly travel continuously (and not suddenly as predicted for a potential slipping event) along one arm of the extruded loop (Figure 4.6a-h). Quantitatively, after the encounter, the extruded loop enlarged from 1.6 μm to 2.1 μm under sideflow conditions (due to its increase from 24 kb to 30 kb). Each arm of the extruded loop in the final state thus contains 15 kb (half of the entire loop). For continuous loop extrusion into one loop, one would thus expect the particle to have moved to a distance of $6\ kb/15\ kb \cdot 2.1\ \mu m$ = 0.84 μm from the loop stem, which indeed is in excellent agreement with measurements of the ~0.9 μm distance between roadblock and loop stem (Figure 4.6h). Yet another measure is to consider whether or not the particle moved into the loop with a speed comparable to the loop-extrusion speed. This would be expected in our scenario of continuous loop extrusion where the particle smoothly moves into the extruded loop, whereas the pseudotopological model would predict the particle to stall at the loop stem and then suddenly speed up in a slippage event. Experimentally, we observed that the particle moved over a distance of ~0.9 μm in 13 seconds, i.e., with a

speed of ~0.07 μm/s. This fits well with the loop-extrusion speed which for this example is ~0.08 μm/s (since it is ~0.5 kb/s (Figure 4.6i) and the 6 kb is extruded into 0.9 μm along one arm of the loop). We thus observe that the loop-extrusion speed is the same as the velocity with which the roadblock moves away from loop base.

The data thus are in full accordance with continuous extrusion of the DNA and the particle into one loop, as predicted by the nontopological model, while the data are not in accordance with the pseudotopological model.

Test of Prediction 3. How do the size of the extruded DNA loop and the MSD of the particle correlate? We previously reported ([226] and chapter 3) that we observed a strong correlation between an increase in loop size and an increase in MSD after the encounter with the roadblock. Here, we focus more in detail on stalling and passing events, where the MSD remained constant, and we searched for data that were compatible with prediction 3. In the model prediction, the loop size can grow while the MSD of the particle may stay constant (Figure 4.4a). Note also that here we discuss experiments without an applied side flow. Figure 4.4b shows an example of what we called a stalling event where the MSD (black) remained constant for more than 10 s. Clearly, the corresponding loop sizes during the period when the MSD was constant (e.g. between 30 s and 60 s in Figure 4.4b and above 30 s in Figure 4.4c) was also found to be constant. This is the typical behavior. We examined all our stalling events (N = 15), and we found that none of the events showed a concomitant increase in loop size (see Figure 4.4d for additional examples). The observation that the loop size never increased while the MSD remained constant is not consistent with one of the key predictions made by the pseudotopological while it meets the expectation of the nontopological model.

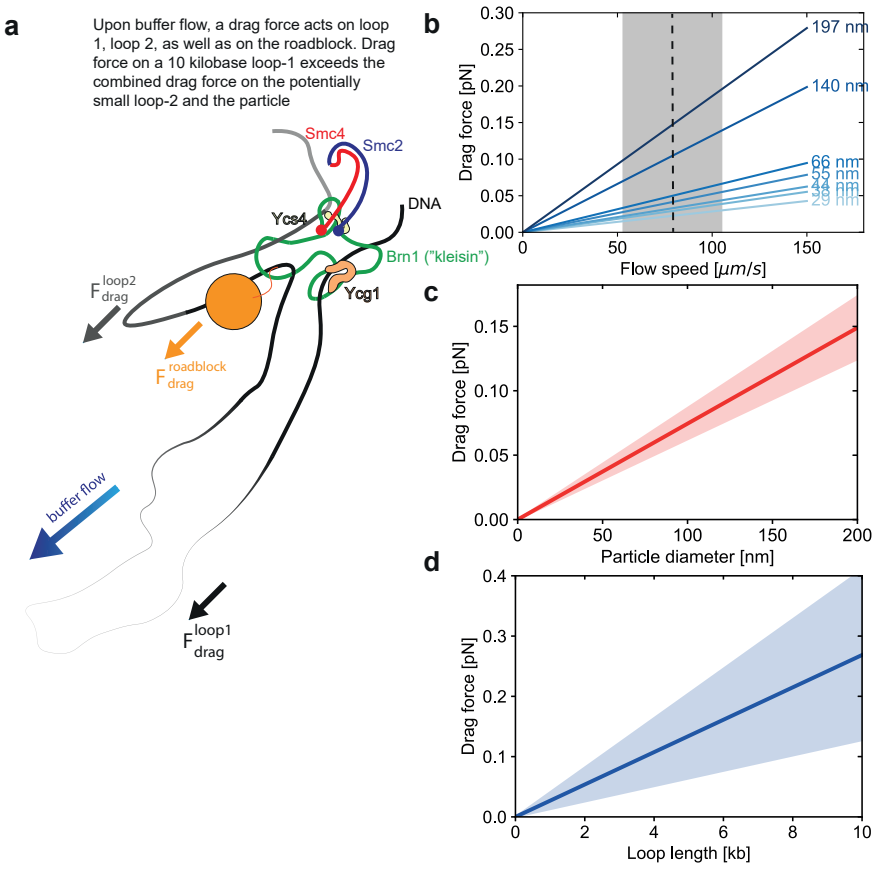


Figure 4.3 | Drag forces on loops and particle associated with the pseudotopological model.
(**a**) Schematic showing the blockage of the roadblock in presence of a large loop-1. (**b**) Estimated drag forces acting on particles of different sizes. The dashed line and grey area denote the experimentally measured flow speed of 79 ± 26 µm/s (N = 20). (**c**) The drag force on the roadblock (red, standard deviation based on uncertainty in buffer flow speed) versus particle size at the experimentally measured flow speed. (**d**) Drag force on DNA loop versus loop length (blue line and shaded area denote mean and standard deviation, respectively). Standard deviations are based on the uncertainty in buffer flow speed (panel b) and drag coefficient [301].

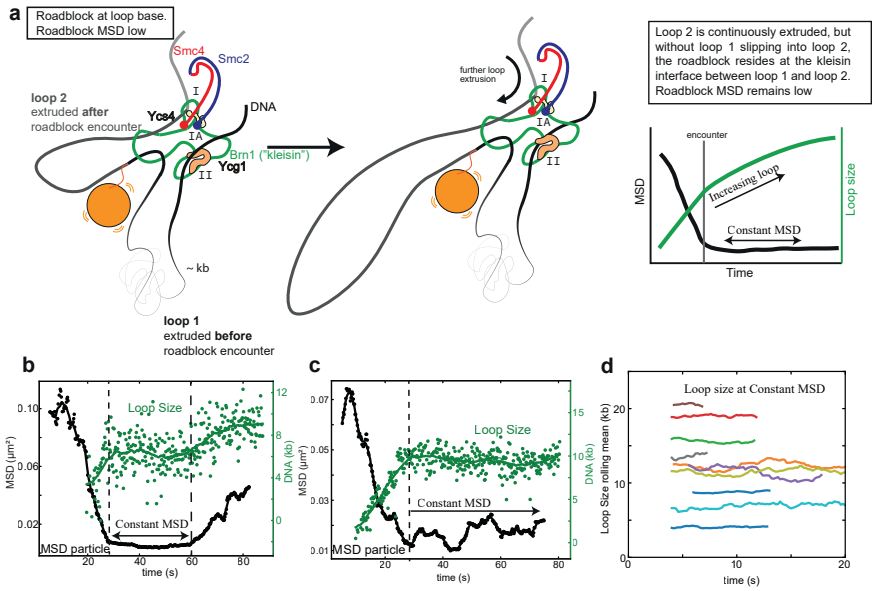


Figure 4.4 | Test of prediction 3 associated with the pseudotopological model.
(**a**) Correlation between DNA loop size and MSD of the particle after the encounter between condensin and roadblock, associated with the pseudotopological model. (**b-c**) Examples of kinetics of MSD (black) and loop sizes of extrusion (green) for events where the MSD remained low for a longer duration. (**d**) Rolling mean of loop size with a window size of 20 points as a function of time in events where the MSD remained constant. Different colors represent different DNA molecules (N = 10). λ-DNA and a gold nanoparticle of 39 nm was used in all panels.

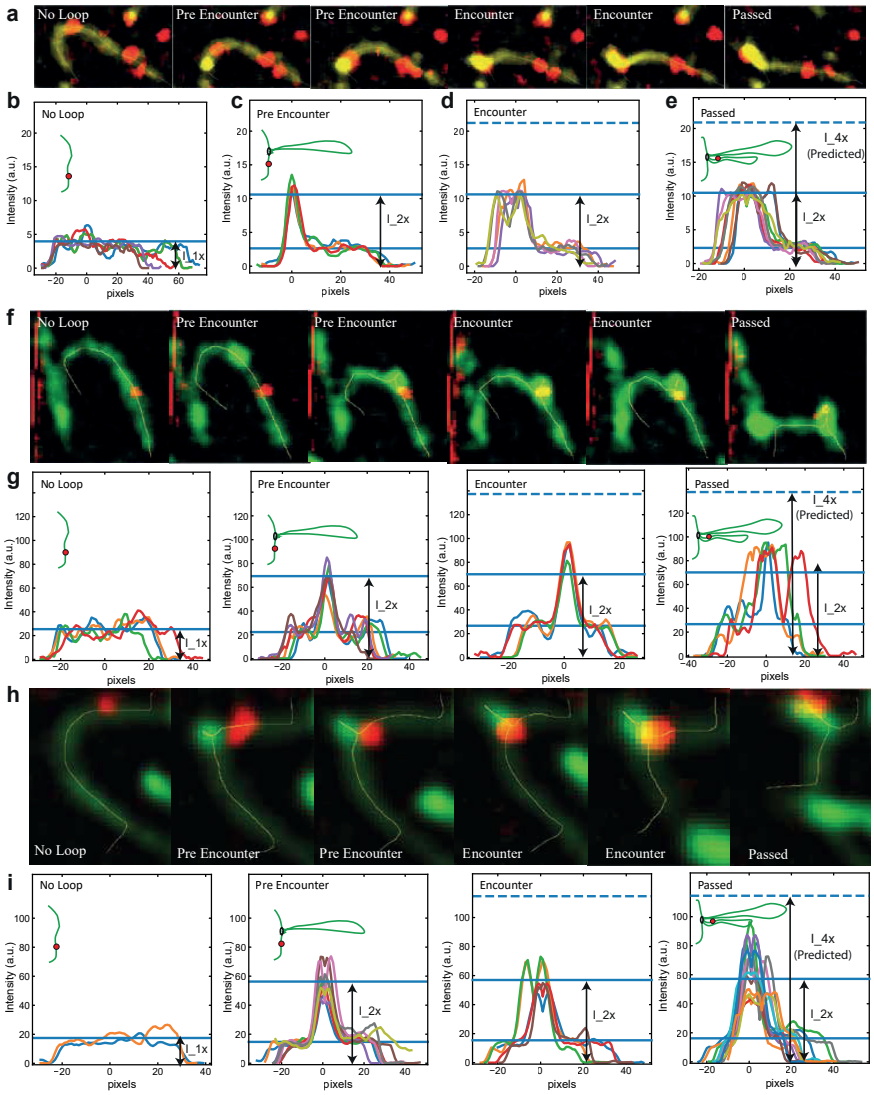


Figure 4.5 | Test of prediction 1 associated with the pseudotopological model.
(**a**) Snapshots of DNA and roadblock with sideflow representing different scenarios of intensity profiles in B-D namely No Loop, Pre Encounter, Encounter, and Passed the roadblock. Intensity profiles of DNA with a roadblock under sideflow before loop extrusion (b, No Loop), with loop but before encountering roadblock (**c**, Pre Encounter), at encounter when the roadblock is at the stem of the loop (**d**, Encounter), and after the encounter when the roadblock is inside the loop (**e**, Passed). Different colors represent different frames. (**f-i**) Two more examples of the intensity profiles on DNA at different stages of the loop extrusion. λ-DNA and a gold nanoparticle of 39 nm was used in all panels.

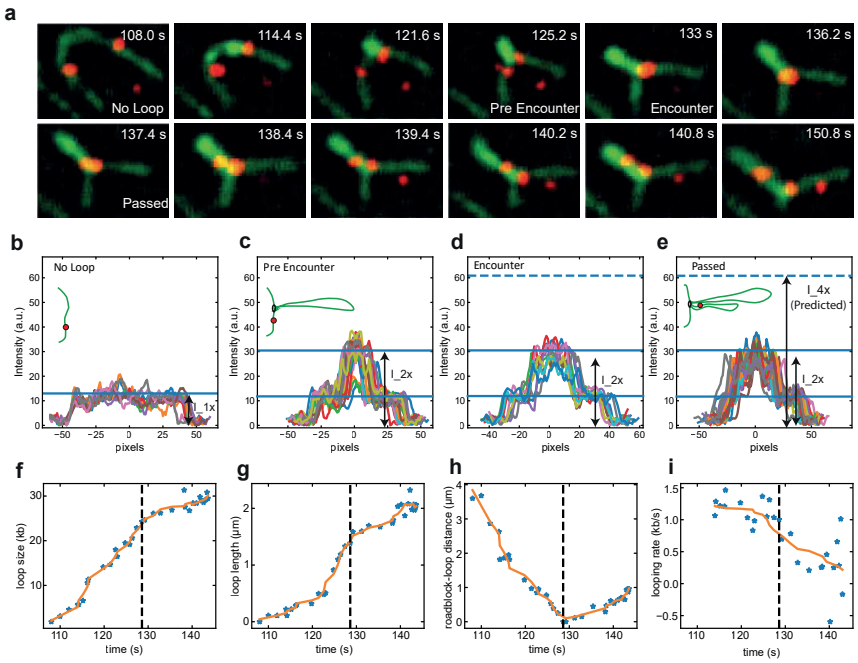


Figure 4.6 | Test of prediction 2 associated with the pseudotopological model.
(**a**) Snapshots of DNA (green) and a 39 nm-wide roadblock (red) with sideflow showing the continuous approach of the roadblock toward the loop stem before encounter (time points until 128 s) and the continuously increasing distance between roadblock and loop stem after encounter. (**b-e**) Loop intensities during the time series. (**f**) Loop size (kb) versus time. (**g**) Loop length (μm) versus time. (**h**) Distance (μm) between roadblock and loop stem versus time. (**i**) Looping rate versus time. Dashed lines in f-i show the time of encounter. λ-DNA and a gold nanoparticle of 39 nm was used in all panels.

4.3 Conclusion

We conclude that the predictions of the pseudotopological models by Shaltiel et al. and Nomidis et al. are not consistent with our experimental observations. Instead, our experimental data are fully consistent with a nontopological model of loop extrusion.

While the loop extrusion mechanism remains unresolved and is currently a matter of strong debate, a nontopological model for condensin and cohesin can be envisioned, as outlined by Oldenkamp and Rowland [227]. The high structural similarity between condensin and cohesin (similar length of the SMC arms and the kleisin, presence of both HEAT-A and HEAT-B) suggests that the underlying loop extrusion mechanism is highly analogous. It is unclear if SMC5/6, which has been observed to dimerize for loop extrusion *in vitro* [204] and whose kleisin is almost half shorter than for condensin (402 amino acids for yeast Nse4 versus 754 amino acids for yeast Brn1), also uses a similar mechanism. The available crosslink data by Shaltiel et al. for yeast condensin allow both a pseudotopological and a nontopological DNA entrapment. In the nontopological version, the HEAT-A-bound N-terminal fragment of the kleisin could flip open, alleviating particle bypass through the HEAT-A compartment. Upon ATP hydrolysis, the ATP heads disengage [238], allowing roadblocks to pass through the ATP heads, see Figure 4.1b (more details of this model in [247]).

Models such as these can be further tested in future experiments that probe the traversal of large roadblock proteins that symmetrically embrace DNA. Small (~10 nm diameter) roadblocks such as nucleosomes can be passed by SMC complexes directly, as distances of up to tens of nm between the ATPase heads have been observed in AFM images of condensin and cohesin [237, 238]. Roadblocks of any size which bind to DNA asymmetrically on one side (like the dCas9-nanoparticle constructs used in this study) can pass as long as their DNA-binding linker is long enough. However, this scenario could differ for a symmetric protein complex that fully wraps around DNA, like for example the MCM hexamer (a part of the eukaryotic replication machinery). Roadblock passage experiments with such DNA-enclosing roadblocks would be of interest to test this prediction.

4.4 Methods

Data collection

Data acquisition and analysis of loop extrusion and roadblock passages were performed as described in detail in [226].

Estimation of the flow velocity

Occasionally, bright autofluorescent particles and/or free fluorescently labeled nanoparticles (e.g. after breakage of dsDNA or rare dissociation of dCas9 from DNA) were observed flowing through a particular field of view under buffer flow. These particles were detected and tracked for as long as they were present in the field of view (Figure S4.1). The distance between successive localizations within one field of view were converted to the instantaneous flow speed using the imaging interval. Since the images were recorded using Highly Inclined and Laminated Optical sheet (HILO) microscopy [226], the recorded fluorescence of fly-by particles which were in focus are localized within a distance of about approximately a micrometer from the axial plane in which DNA is imaged such that the flow speed encountered by fly-by particles and roadblocks on DNA or the DNA itself was comparable.

Estimation of the drag force

The drag force on a spherical particle is computed according to Stokes' law $F_{drag} = 6\pi\eta R\, v$, where $\eta = 1\ mPa\ s$ for a water-based buffer, R is the particle radius, and v the buffer flow speed. The drag force on the extruded DNA loop is estimated based on [301]. Pedersen et al. compute the parallel and perpendicular drag coefficient per unit length on a cylinder of radius 1 nm and obtain values in the range $2-7\cdot10^{-3}\ pN\ s\ \mu m^{-2}$ numerically and from fits to experimental data. Since the majority of the DNA loop is aligned with the buffer flow in our experiments, we consider the lower bound of these estimates, i.e. $\gamma = 1\cdot10^{-3}\ pN\ s\ \mu m^{-2}$ per unit length. The drag force on the extruded DNA loop is computed as $F_{drag}^{loop} = \gamma v\ L_c$ where L_c is the contour length of a strand of DNA of number of basepairs N, i.e. $L_c = \alpha N\ \cdot 0.34\ nm$, where $0.34\ nm$ is the distance between basepairs, and α is a factor which accounts for a slight DNA stretching when using intercalating dyes [217] like Sytox Orange as in [226].

Measurement of DNA intensity profiles in sideflow images

Image frames were marked as "NoLoop" when the loop was not yet initiated; "Pre Encounter" when a loop was growing but the roadblock had not yet encountered (colocalized) with the stem of the loop; "Encounter" when the roadblock and stem of the loop colocalized; and "Passed" when "roadblock" was observed inside the loop. Contour lines were manually drawn on the DNA and intensity profiles were obtained using ImageJ. Line profiles from each group were plotted together in Python.

Quantification of loop length, size, distance to roadblock, and loop extrusion rate in side flow experiments

Intensities inside and outside the loop were extracted by manually drawing rectangular boxes with a width of 11 pixels. DNA loop size was calculated as 48.5 kb · loop-Intensity / total-Intensity. The roadblock-loop distance was calculated by manually drawing a line between the roadblock and the stem of the loop and converting the length of the line from

pixels to μm with a pixel size of 0.108 μm/pixel. Similarly, loop length was calculated by drawing a line from the stem of the loop to the end of the loop and image frames with loops in the optical focus were chosen for length calculations.

4.5 Supplementary Information

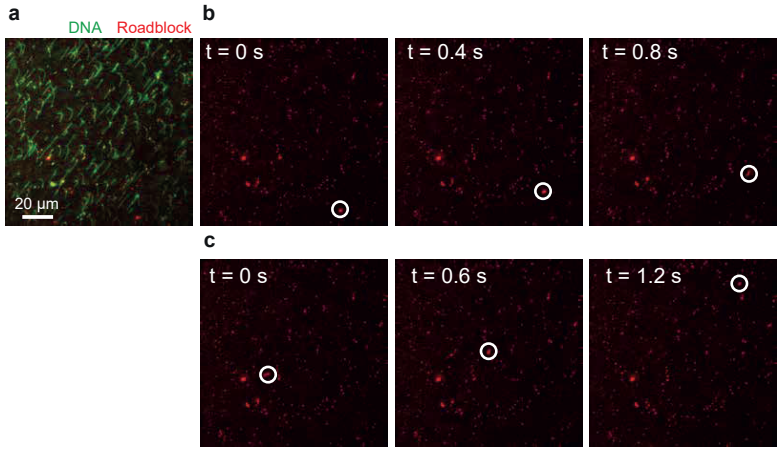
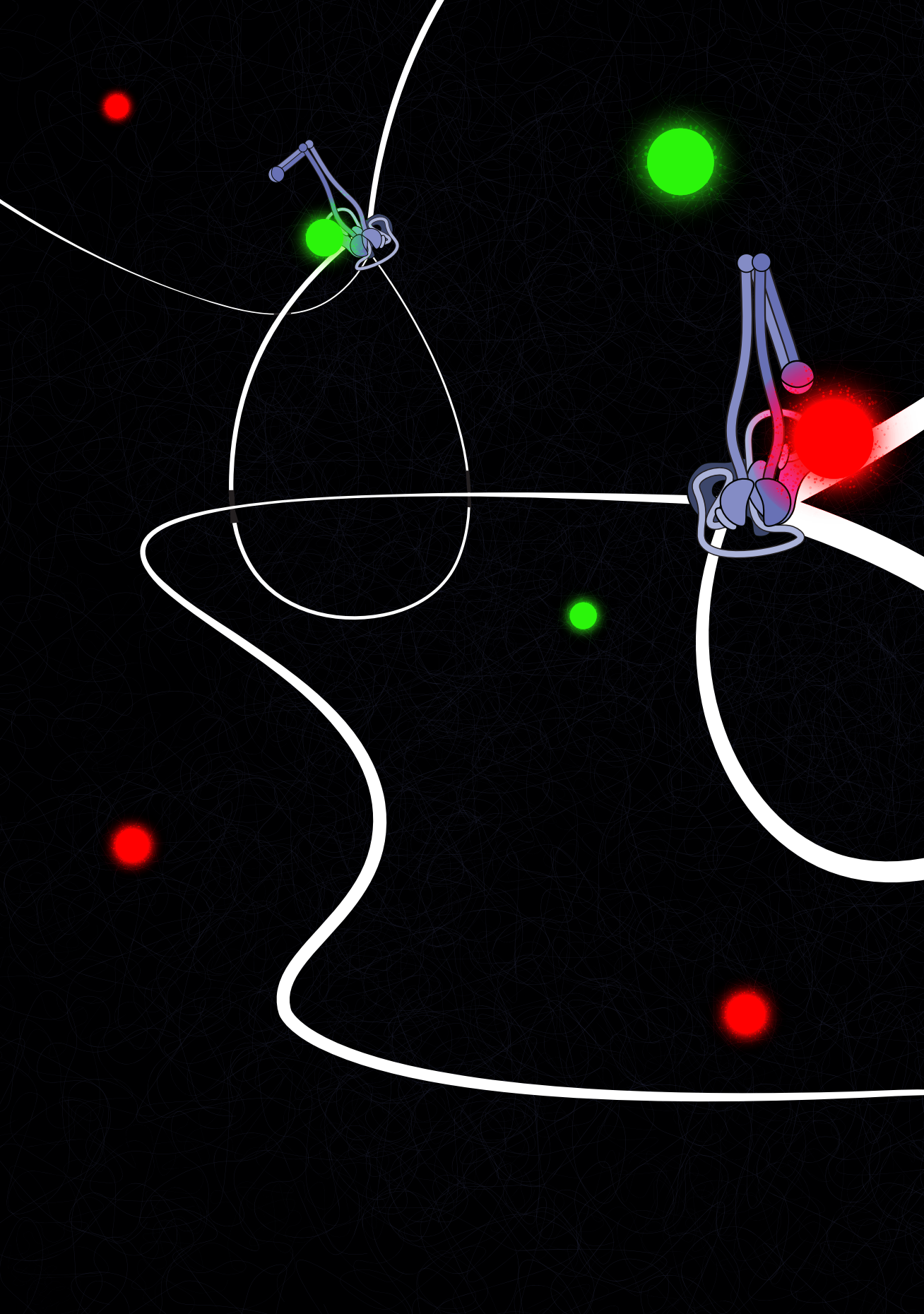
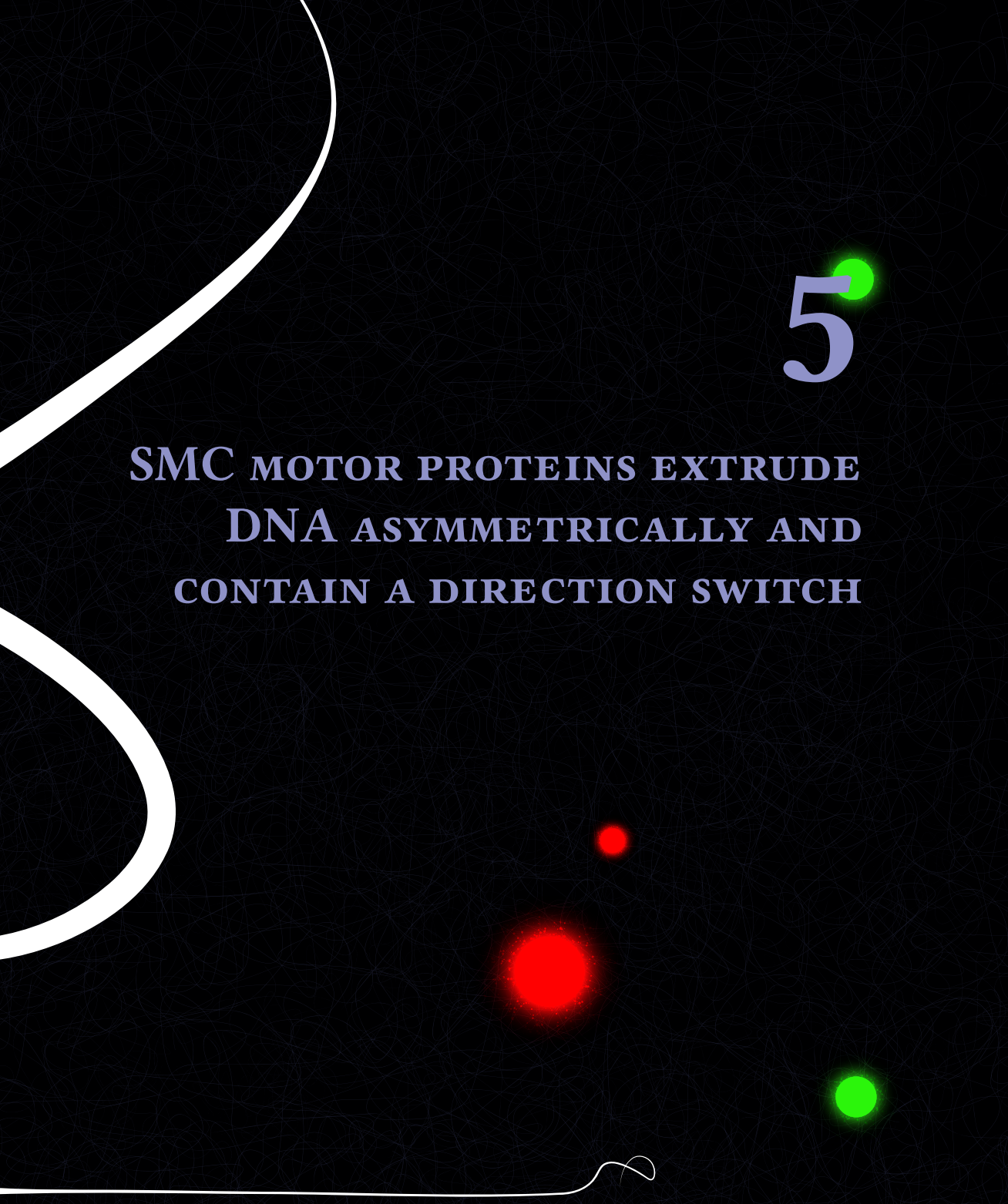


Figure S4.1 | Measurement of the flow speed using unattached fluorescent particles.
(**a**) A field of view is shown in which a fluorescently labelled roadblock (red) is attached to a Sytox Orange-stained DNA (green). The DNA is stretched by a buffer flow perpendicular to the DNA's end-to-end vector. (**b-c**) Two instances of free fluorescent particles (marked by a white circle) are shown that cross the field of view within 0.8 and 1.2 s, respectively. The flow speed is computed from the distance between localizations in successive frames.



5

SMC motor proteins extrude DNA asymmetrically and contain a direction switch



This chapter is available as preprint: Roman Barth, Iain F. Davidson, Jaco van der Torre, Michael Taschner, Stephan Gruber, Jan-Michael Peters, Cees Dekker: SMC motor proteins extrude DNA asymmetrically and contain a direction switch, bioRxiv 2023 Dec [222].

Abstract

Structural Maintenance of Chromosomes (SMC) complexes organize the genome via DNA loop extrusion. While some SMCs were reported to do so symmetrically, reeling DNA from both sides into the extruded DNA loop simultaneously, others perform loop extrusion asymmetrically toward one direction only. The mechanism underlying this variability remains unclear. Here, we examine the directionality of DNA loop extrusion by SMCs using in vitro single-molecule experiments. We find that cohesin and SMC5/6 do not reel in DNA from both sides, as reported before, but instead extrude DNA asymmetrically, while the direction can switch over time. Asymmetric DNA loop extrusion thus is the shared mechanism across all eukaryotic SMC complexes. For cohesin, direction switches strongly correlate with the turnover of the subunit NIPBL, during which DNA strand switching may occur. The findings reveal that SMCs, surprisingly, contain a direction switch subunit.

5.1 Introduction

Genomes across the tree of life are organized and constantly reshaped by active processes within the cell. Structural Maintenance of Chromosomes (SMC) protein complexes are important, owing to their major role in chromosome organization [1]. The three major members of the SMC family in eukaryotes are condensin, cohesin, and SMC5/6. All three are capable of manipulating DNA via an ATP-dependent process called DNA loop extrusion [50, 98, 138, 173, 204, 212, 214–216], where loops are created and enlarged stepwise by ATP-binding events to the SMC complex [50, 217] (and chapter 6). Condensin mainly extrudes DNA loops during mitosis to compact sister chromatids and facilitate chromosome segregation [112, 116, 116]. Cohesin organizes interphase chromosomes by looping DNA between convergently oriented CCCTC-binding factor (CTCF) proteins [35, 141, 155, 303], giving rise to topologically associated domains in population-average chromosome-conformation-capture [33, 41, 148, 153, 158] as well as imaging experiments [160, 161, 304], which contributes to transcriptional regulation [1, 181] as well as genome integrity [196, 305, 306]. SMC5/6 has poorly understood functions in chromosome segregation and genome maintenance [307, 308].

The composition and architecture of the three SMC complexes are highly similar (shown schematically for human cohesin in Figure 5.1a; reviewed previously [228, 230]). Two Smc subunits dimerize at their hinge domain and are connected *via* coiled-coil arms to their heads. The heads harbour ATP-binding domains related to those of ATP-binding cassette (ABC) transporters, which dimerize upon ATP binding to catalyse its hydrolysis. A flexible kleisin subunit bridges the Smc ATPase heads and provides binding sites for two HAWK (HEAT protein associated with a kleisin) proteins in cohesin and condensin, or KITE (kleisin interacting tandem winged-helix element) proteins in SMC5/6.

Despite the overall highly conserved architecture of eukaryotic SMCs and their common ability to extrude DNA loops [227, 228], they appear to significantly differ in their loop extrusion directionality. Here, we use the following terminology to describe the directionality: asymmetric loop extruders incorporate DNA only from one side into the loop, while symmetric extruders reel in DNA from both sides simultaneously into the loop. Unidirectional extruders undergo sequential phases of asymmetric extrusion (possibly interrupted by pauses) that always occur in the same direction, whereas bidirectional extruders exhibit phases of asymmetric extrusion but here the side from which DNA is incorporated into the loop switches direction over time. Yeast condensin is an asymmetric and unidirectional extruder, that is, it reels DNA into the loop strictly from one side [98, 221], while human cohesin was reported to reel DNA from both sides into the loop and was thus considered to be a symmetric DNA loop extruder [173, 212]. Dimeric yeast SMC5/6 similarly was reported to be a symmetric extruder [204]. Given the structural similarity of these complexes [138, 228, 230], the reason for these differences is unclear. This variability undermines whether the loop extrusion mechanism [247] is shared among all SMC complexes.

Here, we experimentally measured the directionality of DNA loop extrusion by human cohesin, yeast condensin, and yeast SMC5/6 at a single-molecule level. Close inspection of their loop extrusion dynamics revealed that cohesin and SMC5/6 undergo short phases of active extrusion, characterized by loop growth, which are interspersed by diffusion [173]

of the loop and SMC across the DNA, and loop slipping, the gradual loss of the previously extruded loop. Focusing on the extrusion phases, we find that all three monomeric SMC complexes extrude DNA strictly asymmetrically within these phases. While the condensin holocomplex is a strictly unidirectional extruder, cohesin and monomeric SMC5/6 undergo frequent direction switches, making them bidirectional loop extruders. Furthermore, we found that human cohesin undergoes direction switches upon exchange of its HEAT subunit NIPBL. These findings indicate that the DNA loop extrusion mechanism is inherently asymmetric and is likely common to all SMC complexes.

5.2 Results

5.2.1 Phases of active DNA loop extrusion are interspersed by loop diffusion and loop slipping

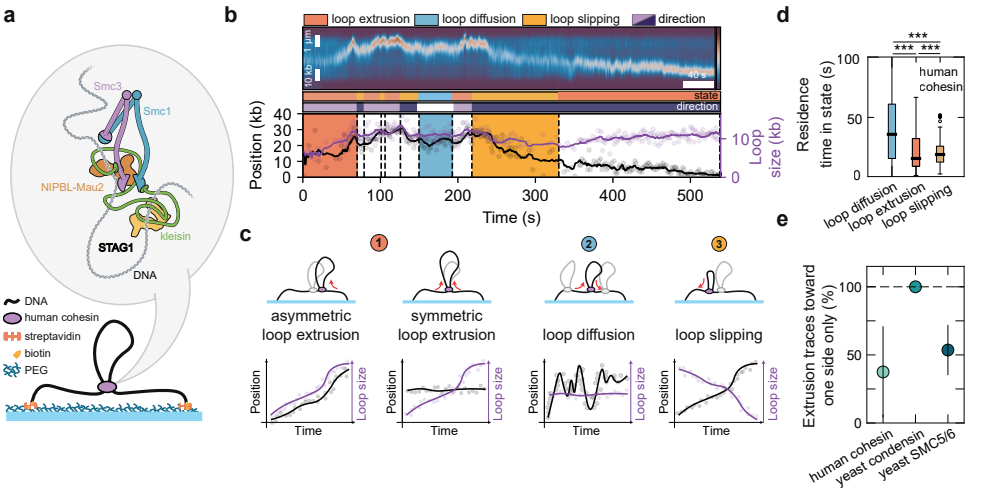


Figure 5.1 | SMC-mediated DNA loop extrusion is interspersed with loop-diffusion and loop-slipping events.
(**a**) Illustration of the DNA loop extrusion assay and the architecture of human cohesin containing Smc3, Smc1, Scc1 (kleisin), NIPBL-Mau2, and STAG1/2 (here only STAG1 is used). The corresponding subunit names for yeast condensin are Smc4, Smc2, Brn1, Ycs4, and Ycg1. The corresponding subunit names for yeast SMC5/6 are Smc5, Smc6, and Nse4 with additional Nse subunits Nse1, Nse2, and Nse3. (**b**) A typical kymograph of cohesin-mediated DNA loop extrusion. Loop position and size are quantified. Their time propagation allows us to discern phases of loop extrusion, diffusion and slipping. Extrusion and slipping have a direction given by the local sign of the derivative of the loop position with respect to time. (**c**) Illustration of loop position and size propagation in time for asymmetric and symmetric extrusion, loop diffusion, and slipping. (**d**) Duration of diffusion, extrusion, and slipping states (N = 93, 344, 257, respectively) for human cohesin. Black horizontal lines are median values, the box extends between the first and third quartile, and the whiskers extend to 1.5·IQR. Statistical significance was assessed by a Mann-Whitney test with Bonferroni correction. (***: p<0.001). (**e**) Fraction of SMC complexes with at least two extrusion events that displayed only unidirectional extrusion (N = 13, 49, 33 for cohesin, condensin, and SMC5/6, respectively). Error bars denote the binomial 95% confidence interval.

To characterize the dynamics of DNA loop extrusion by human cohesin in detail, we reconstituted DNA loop extrusion with purified proteins at the single-molecule level *in vitro*. To this end, λ-DNA, which served as a substrate for loop extrusion, was tethered at

both ends onto a polyethylene glycol-passivated glass surface at an extension of ~30% of its contour length. DNA was visualized using SYTOX Orange intercalating dye, and imaged using highly inclined and laminated optical sheet (HILO) microscopy (Methods section 5.5). Human cohesin and ATP were subsequently introduced into the flow channel and the flow was stopped to record cohesin-mediated loop extrusion in the absence of buffer flow (Figure 5.1a). Kymographs were constructed that served as the basis for quantifying the size of the loop and its position (Figure 5.1b, Methods section 5.5).

The time propagation of these two independent quantities, loop size and position, defines distinct phases into which loop extrusion traces were segmented (Figure 5.1c): (i) Phases of active DNA loop extrusion exhibited an increase in loop size (Figure 5.1c and red shading in Figure 5.1b), (ii) extended 'loop diffusion' periods in which the loop size remained approximately constant, while the loop position changed in a diffusive manner (Figure 5.1c, blue shading in Figure 5.1b, and Figure S5.1g), and (iii) gradual shrinking of loops (Figure 5.1b, yellow shading in Figure 5.1c, and Figure S5.1h), as noted previously as a possible outcome of cohesin-CTCF encounters [35]. Phases of loop stalling, e.g. by reaching the end of surface-tethered DNA molecules, were excluded (e.g. Figure 5.2). This gradual shrinkage is distinct from a sudden step-wise loop rupture, which resolves the loop within one or few frames (~100 ms) [98, 221]. Loop shrinkage typically led to only a partial (i.e. incomplete) loss (Figure S5.1a,b,g). Quantitatively, cohesin remained longest in the diffusive state (41 ± 44 s, median ± SD), while spending 16 ± 17 s (median ± SD) in active extrusion, which was comparable to the time spent in the loop slipping phases (20 ± 15 s, median ± SD; Figure 5.1d). We tested whether even shorter phases existed, which could have been missed by acquiring images at a frame rate of 5-10 frames per second (fps). Acquired images at 50 fps yielded no significantly shorter phases however (Methods section 5.5, Figure S5.1c-f).

To examine how the loop extrusion dynamics among the three eukaryotic SMC complexes differed, we repeated the experiments and analyses with budding yeast condensin and budding yeast SMC5/6 (Methods section 5.5, Figure S5.2). First, we tested whether the purified SMC5/6 used here is able to perform DNA loop extrusion, as reported previously [204] (Figure S5.2). Indeed, we confirmed that SMC5/6 exhibits loop extrusion of DNA with similar characteristics as reported (Figure S5.2b-e), although we found that yeast SMC5/6 - like human cohesin and yeast condensin - extrudes already as a monomer (Figure S5.2f-i), while Pradhan et al. [204] reported loop extrusion only for SMC5/6 dimers. This held true both for SMC5/6 expressed and purified from *E. coli* (ref. [309]) as well as from *S. cerevisiae* (ref. [204]; compare Figure S5.2g,i). While SMC5/6 exhibited dynamics comparable to cohesin, condensin exhibited only brief phases of diffusion (7 ± 4 s, median ± SD) while extrusion phases lasted 13 ± 9 s (median ± SD), and loop slipping was hardly ever observed (1/65 phases; Figure S5.1k).

5.2.2 All eukaryotic SMC complexes extrude DNA asymmetrically

The growth and concomitant change in the loop position is a proxy for the direction of DNA loop extrusion, as illustrated in Figure 5.1c. By monitoring the direction of travel of the loop extrusion phases during the experiment (from the start of the extrusion until loop dissolution or the end of acquisition), we measured the fraction of SMC complexes that

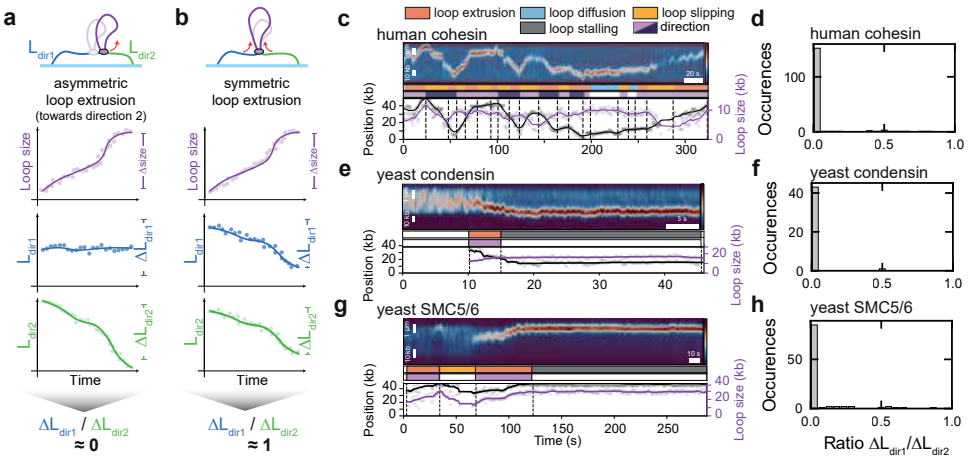


Figure 5.2 | All eukaryotic SMC complexes extrude DNA asymmetrically.
(**a**) Illustration of loop size and amount of DNA left and right of the loop in time for asymmetric loop extrusion. (**b**) Illustration of loop size and amount of DNA left and right of the loop in time for symmetric loop extrusion. The ratio of changes in the amount of DNA left and right of the loop is discriminative between asymmetric and symmetric extrusion. (**c**) Exemplary loop extrusion kymograph of cohesin with loop position and size quantification and segmentation into phases of extrusion, diffusion, and slipping. (**d**) Distribution of the ratio $\Delta L_{dir1}/\Delta L_{dir2}$ for extrusion events of human cohesin (N = 166 events). (**e**) Exemplary kymograph of unidirectional DNA loop extrusion for yeast condensin. (**f**) Distribution of the ratio $\Delta L_{dir1}/\Delta L_{dir2}$ distribution for yeast condensin (N = 44 events). (**g**) Exemplary kymograph of unidirectional DNA loop extrusion for yeast SMC5/6. (**h**) Distribution of the ratio $\Delta L_{dir1}/\Delta L_{dir2}$ distribution for yeast SMC5/6 (N = 100 events).

showed multiple extrusion phases in the same or opposite directions. While condensin is a strictly unidirectional extruder (all extrusion phases moved toward one side of the DNA), 40-60% of cohesin and SMC5/6 complexes that show at least two extrusion phases extruded toward both sides during the course of the experiment (Figure 5.1e).

Cohesin-mediated loop extrusion was mostly interspersed with either a diffusion or a slipping phase (or both) between subsequent extrusion phases (Figure S5.1i); switches between extrusion and slipping phases were most common (Figure S5.1j). Direction changes without intervening diffusion or slipping phases were less common, as only approximately 25% of direction changes occur during loop extrusion, i.e. without observable intermittent diffusion or slipping phase (Figure S5.1i). The sequence of extrusion, diffusion, and slipping phases can be described as a first-order Markov chain, that is, the transition probabilities between the states solely depend on the current state, indicating that the system is memory-less (Figure S5.3a).

Asymmetric and symmetric DNA loop extruders are characterized by their ability to reel DNA from one or both sides into the loop, respectively [98, 204]. For an asymmetric extruder, the stretch of DNA on one side of the SMC becomes progressively shorter, while the length of the other side remains constant [98] (Figure 5.2a). In contrast, both arms become simultaneously shorter as the loop is being enlarged by a symmetric extruder (Figure 5.2b). Thus, the ratio of the length changes on both sides of the loop during loop extrusion constitutes a quantity that we name a 'symmetry indicator' which is close to

0 for an asymmetric extruder (since $\Delta L_{dir1} \approx 0$ in Figure 5.2a) and ≈1 for a symmetric extruder. To determine whether cohesin extrudes DNA asymmetrically or symmetrically, we computed the symmetry indicator $\Delta L_{dir1}/\Delta L_{dir2}$ within the segmented loop extrusion phases. We found that 92 ± 4 % (mean ± 95% binomial confidence interval, N = 166) of all extrusion phases exhibited a value of $\Delta L_{dir1}/\Delta L_{dir2}$ below 0.1 (Figure 5.2d). Extrusion phases of yeast condensin, a known asymmetric extruder [98] (Figure 5.2e), used here as a positive control, exhibited a similar distribution with 98 ± 4 % (mean ± 95% binomial confidence interval, N = 44, Figure 5.2f) of all values of $\Delta L_{dir1}/\Delta L_{dir2}$ below 0.1. Monomeric yeast SMC5/6 (Figure 5.2g) also exhibited asymmetric extrusion phases (86 ± 7 % [mean ± 95% binomial confidence interval] of $\Delta L_{dir1}/\Delta L_{dir2}$ values below 0.1, N = 100, Figure 5.2h and Figure S5.1l). Notably, the slipping phases of cohesin and SMC5/6 were also asymmetric (Figure S5.1m,n), suggesting that loop slipping is the result of weak DNA binding at either the motor or anchor side of the SMC complexes, but not both.

We conclude that all eukaryotic SMC complexes extrude DNA asymmetrically, strictly toward one direction at a time; however, cohesin and SMC5/6 may switch the direction of extrusion between two successive extrusion phases.

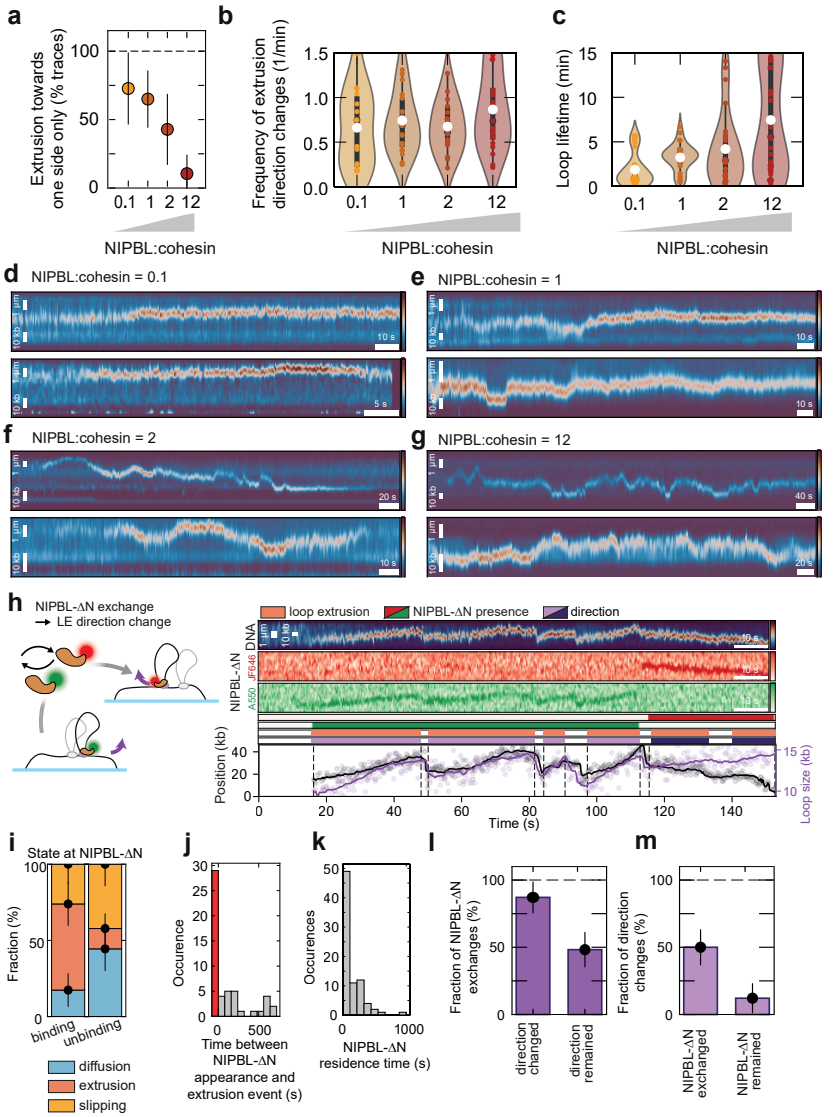


Figure 5.3 | DNA loop extrusion direction changes coincide with an exchange of NIPBL-ΔN.
(**a**) Fraction of cohesin complexes undergoing at least two extrusion events that displayed only unidirectional extrusion for varying ratios of NIPBL-Mau2:cohesin at 50 mM NaCl (N = 41, 30, 17, 26 from left to right). Error bars denote the binomial 95% confidence interval. (**b**) Frequency of extrusion direction changes for varying ratios of NIPBL-Mau2:cohesin at 50 mM NaCl (N = 41, 30, 17, 26). The white dot denotes the mean. The box shows the quartiles of the dataset and the whiskers extend to 1.5·IQR. (**c**) Loop lifetime between the first extrusion event and loop disruption for varying ratios of NIPBL-Mau2:cohesin at 50 mM NaCl (N = 41, 31, 25, 37). The white dot denotes the mean. The box shows the quartiles of the dataset and the whiskers extend to 1.5·IQR. (**d**) Exemplary kymographs of cohesin-mediated loop extrusion at NIPBL-Mau2:cohesin = 0.1 at 50 mM NaCl. (**e**) Exemplary kymographs of cohesin-mediated loop extrusion at NIPBL-Mau2:cohesin = 1 at 50 mM NaCl. (**f**) Exemplary kymographs of cohesin-mediated loop extrusion at NIPBL-Mau2:cohesin = 2 at 50 mM NaCl. (**g**) Exemplary kymographs of cohesin-mediated loop extrusion at NIPBL-Mau2:cohesin = 12 at 50 mM NaCl. *Caption continues on next page.*

Figure 5.3 | (**h**) Exemplary trace where the exchange of NIPBL-ΔN coincides with a change in the loop extrusion direction. Cohesin extrudes toward the lower end as long as the ATTO 550-labeled NIPBL-ΔN is bound to cohesin. After exchange to a JF646-labeled NIPBL-ΔN, cohesin extrudes toward the upper end. HeLa cohesin was used with NIPBL-ΔN-ATTO550 and NIPBL-ΔN-JF646. (**i**) Probability of states at which NIPBL-ΔN binds to cohesin (N = 46) and at which NIPBL-ΔN dissociates (N = 45). Binding and unbinding at extrusion phases was only observed concomitant with the start/end of extrusion phases, respectively. The bar height indicates the probability of a given state and the error bar denotes the binomial 95% confidence interval. (**j**) Distribution of time between NIPBL-ΔN appearance and the next extrusion event (N = 56). The red bar represents the occurrences when NIPBL-ΔN appearance coincided with the start of an extrusion event. (**k**) Distribution of NIPBL-ΔN residence times (N = 47). (**l**) Fraction of cohesin molecules for which NIPBL-ΔN exchanged in between direction changes (87 ± 12 %, N = 31) and the fraction of cohesin molecules on which NIPBL-ΔN exchanged without direction change between successive extrusion phases (48 ± 13 %, N = 56) at 50 mM NaCl. Error bars denote the binomial 95% confidence interval. (**m**) Fraction of direction changes between successive extrusion phases during which NIPBL-ΔN exchanged (50 ± 13 %, N = 54) and the fraction of direction changes between successive extrusion phases during which NIPBL-ΔN remained (12 ± 11 %, N = 33) at 50 mM NaCl. Error bars denote the binomial 95% confidence interval.

5.2.3 NIPBL excess increases the occurrence of direction switches in cohesin-mediated loop extrusion

Even though yeast condensin is a strictly unidirectional DNA loop extruder, the complex extrudes towards both directions alternatingly when Ycg1, condensin's 'anchor' (ref. [221]), is deleted, or when the strong DNA binding site within the safety belt [210], is mutated. These results suggest that weakening or dissolution of at least one DNA-binding site is required to permit direction switches during DNA loop extrusion. Human cohesin requires a reservoir of unbound NIPBL-Mau2 for ongoing DNA loop extrusion and loop maintenance [173] and *in vivo* FRAP experiments indicated that NIPBL-Mau2 hops from cohesin to cohesin in cells [310], suggesting that NIPBL-Mau2 is only transiently bound to cohesin as NIPBL-Mau2 heterodimers continuously exchange on cohesin. As NIPBL-Mau2 contains a strong DNA-binding site [238], we hypothesized that dissociation of NIPBL-Mau2 from the complex weakens the DNA-binding strength close to the Smc3 subunit, which could subsequently lead to extrusion direction switches.

To test this, we performed cohesin-mediated DNA loop extrusion experiments with varying concentrations of NIPBL-Mau2 to modulate the average occupancy of NIPBL-Mau2 on cohesin. In loop extrusion experiments where NIPBL-Mau2 was added at a sub-stoichiometric ratio to cohesin, mostly unidirectional extrusion was observed (Figure 5.3a, ratio of NIPBL-Mau2 to cohesin = 0.1, and Figure 5.3d). Strikingly, upon increasing the NIPBL-Mau2:cohesin ratio, more traces with extrusion phases toward both directions were observed (Figure 5.3a,e-g). In addition, the fraction of cohesin complexes that underwent diffusion and loop slipping increased with increasing NIPBL-Mau2:cohesin concentration ratios (Figure S5.3b,c). Bidirectional extrusion cannot be attributed to the dimerization of cohesin, as suggested previously [212], since in both uni- and bidirectional extrusions, mostly a single cohesin complex colocalized with the loop [173] (Figure S5.3d-f). In response to increasing the NIPBL-Mau2:cohesin ratio, cohesin could change the extrusion direction more frequently, resulting in an increased fraction of cohesin molecules which switch extrusion direction before the loop vanishes. However, we did not find an enhanced direction exchange frequency (Figure 5.3b). Instead, the loop lifetime, the time between

the beginning of the first extrusion phase until loop dissolution, increased as a larger excess of NIPBL-Mau2 over cohesin was supplied (Figure 5.3c), in line with previous loop maintenance assays [173], giving cohesin more time to change extrusion direction or enter diffusion and slipping phases before the loop vanishes. Subsequent extrusion phases were indistinguishable in terms of loop extrusion rate (Figure S5.3g). These results suggest that rebinding of NIPBL to cohesin stabilizes loops and that cohesin can undergo additional loop extrusion cycles in any direction upon rebinding of a NIPBL molecule.

5.2.4 In the absence of NIPBL, loops can diffuse and slip but cannot be extruded

Next, we directly visualized whether an exchange of NIPBL-Mau2 on cohesin occurs. For this, we expressed and purified an N-terminal truncation of NIPBL (NIPBL-ΔN, ref. [173, 238]) which was split into two batches that were labelled differentially with ATTO 550 (NIPBL-ΔN-A550) and Janelia Fluor 646 (NIPBL-ΔN-JF646) fluorophores (Methods section 5.5, Figure S5.4a). We then performed cohesin-mediated DNA loop extrusion experiments with an equimolar ratio of NIPBL-ΔN-A550 and NIPBL-ΔN-JF646 and monitored the presence and absence of NIPBL-ΔN in the two colors colocalizing with the loop (Figure S5.4b-e and Figure 5.3h). Cohesin did not show any extrusion phases in the absence of NIPBL-ΔN. NIPBL-ΔN binding to cohesin typically coincided with the start of a loop-extrusion phase, but NIPBL-ΔN also bound to cohesin during diffusion and slipping phases (Figure 5.3i, left bar). In the latter case, the next extrusion phase followed within a few minutes (142 ± 222 s [mean ± SD], Figure 5.3j). Dissociation of NIPBL-ΔN occurred most often during diffusion or slipping phases, on average 31 ± 53 s (mean ± SD) after the last extrusion phase (Figure S5.4f), and only rarely at the end of extrusion phases (Figure 5.3i, right bar) after NIPBL-ΔN spent on average 161 ± 186 s (mean ± SD) bound to cohesin (Figure 5.3k). Cohesin without bound NIPBL-ΔN was typically rebound by NIPBL-ΔN after ~1-2 min (Figure S5.4g) though cohesin without bound NIPBL-ΔN was observed for up to 5 min (at NIPBL-ΔN:cohesin ratio of 0.1-1 in these experiments), which is in agreement with the report that in the absence of NIPBL-Mau2 most loops dissociated within 8 min (ref. [173]). The diffusion constant of diffusing loops as well as the loop slipping rate were not significantly different in the presence or absence of NIPBL-ΔN (Figure S5.4h,i), suggesting that NIPBL does not contribute to DNA binding during the diffusion and slipping phases. Because NIPBL is known to contribute to DNA clamping onto the ATPase heads upon ATP binding [200, 234], this suggests that diffusion and slipping phases represent an ATP-unbound states of cohesin.

5.2.5 Extrusion direction switches coincide with an exchange of NIPBL on cohesin

Next, we analysed how an exchange of NIPBL-ΔN relates to the extrusion direction switching of cohesin. We considered an exchange of NIPBL-ΔN if the fluorescence intensity of one label was replaced with an intensity of the other label, but also when the fluorescence intensity of one label disappeared and re-appeared at least 2 frames later. Note that imposing this restriction allowed us to monitor exchanges between NIPBL-ΔN carrying the same label only as long as cohesin remained unbound by NIPBL-ΔN for at least 600 ms. Figure 5.3h shows an exemplary kymograph of cohesin-mediated loop extrusion, in which

first (15 - 112 s) four successive extrusion phases move toward the upper end of the DNA molecule with intermittent loop slipping phases. During all these unidirectional phases, a NIPBL-ΔN-A550 molecule colocalized with the DNA loop (presumably bound to cohesin [173]), which however dissociated from cohesin at t = 113 s. At t = 116 s, a NIPBL-ΔN-JF646 bound to the cohesin, which coincided with a direction switch as extrusion of the loop now proceeded toward the lower end of the DNA molecule. We monitored such an appearance and disappearance of NIPBL-ΔN in relation to direction switches for multiple molecules (additional examples in Figure S5.4b-e). In 86 ± 12 % of the cases in which a direction change was observed, there was also an exchange of NIPBL-ΔN between extrusion phases in opposite directions (e.g. Figure S5.4b). When subsequent extrusion phases occurred in the same direction, an exchange of NIPBL-ΔN was observed in 52 ± 12 % of the cases (Figure 5.3l, example in Figure S5.4d). Conversely, an exchange of NIPBL-ΔN coincided in 47 ± 12 % of the cases with a direction switch, while only 14 ± 11 % of the cases in which a direction switch was observed did not correlate with an exchange of NIPBL-ΔN (Figure 5.3m).

These data show that extrusion direction switches coincide an exchange of NIPBL, strongly suggesting that NIPBL exchange is a necessary step to permit extrusion direction switches. The opposite is however not true: an exchange of NIPBL does not necessarily yield a direction switch. It is possible that NIPBL molecules carrying the same fluorophore did exchange within a time period that could not be resolved (using the time resolution of 300 ms used in these experiments). Events in which direction changes were observed without concomitant NIPBL exchange might have been caused by such events (Figure 5.3l, first bar and Figure 5.3m, second bar).

5.3 Discussion

5.3.1 DNA loop extrusion is conserved, asymmetric, and can switch direction

DNA loop extrusion is an important molecular mechanism across the tree of life. While yeast condensin has been reported to extrude DNA from one side only [98], human cohesin [173, 212] and yeast SMC5/6 (ref. [204]) were found to reel in DNA from both sides, suggesting that cohesin and SMC5/6 are symmetric extruders. These differential loop extrusion dynamics questioned whether the DNA loop extrusion mechanism is conserved among eukaryotic SMC complexes. Using single-molecule *in vitro* reconstitution of DNA loop extrusion by all three Eukaryotic SMCs and quantification of their loop extrusion dynamics, we set out to clarify whether an asymmetric DNA loop extrusion mechanism may constitute the common *modus operandi* of the Eukaryotic SMCs. Indeed, we found that all monomeric Eukaryotic SMCs extrude DNA strictly asymmetrically (Figure 5.2). However, cohesin and SMC5/6 are able to switch the direction of extrusion (Figure 5.1e). These results strongly suggest that the DNA loop extrusion mechanism is shared among the eukaryotic SMCs (potentially including the prokaryotic SMCs as well), is inherently asymmetric, and that the DNA loop extrusion cycle permits direction switches.

Our results suggest that direction switches require an exchange of NIPBL on human cohesin (Figure 5.3). Strikingly, the propensity of cohesin and condensin to switch extrusion

directions can be tuned by modulating the ability to bind DNA at either the HEAT-A subunit (NIPBL/Ycs4, respectively), HEAT-B subunit (STAG1/Ycg1), or both. For cohesin, dissociation of NIPBL likely weakens a DNA binding site close to Smc3 (the motor side) because NIPBL forms a strong DNA-binding site on the cohesin holocomplex [238]. Note that the dissociation and rebinding of DNA-binding subunits is an extreme measure to temporarily weaken DNA binding at these subunits. In order to allow direction switches *via* DNA strand exchange (see below), DNA is only required to temporarily unbind from the motor (NIPBL/Ycg4 or equivalent) and anchor (STAG1/STAG2/Ycg1), while both HEAT subunits could remain bound to cohesin and condensin, respectively. The fact that direction switches are observed for ΔYcg1 condensin, even though Ycs4 was intact in these experiments [221], suggests that DNA can indeed temporarily unbind from Ycs4 (and potentially other HEAT and KITE subunits) without complete subunit dissociation from the complex. This might also explain why both uni- as well as bidirectional extrusion has been observed concomitantly for human condensin I and II (ref. [216]) and for *X. laevis* condensin and cohesin from egg extracts [214], where no subunit turnover is known. The loop extrusion activity of human cohesin, but not yeast condensin, is regulated *in vivo* by the replacement of NIPBL by Pds5/WAPL and by the balance of STAG1- and STAG2-bound cohesin [282]. This might explain why subunits (at least NIPBL) are dynamically exchanged on cohesin, while they are not for other SMC proteins.

5.3.2 Mechanistic models that enable DNA strand exchange

How can a DNA-loop-extrusion model account for direction switches? We envision two possible scenarios that extend the recently proposed reel-and-seal model [247], based on previous considerations in which DNA strands between the motor and anchor sides are exchanged ('DNA strand exchange') to enable extrusion direction switches [215, 221], see Figure 5.4 and Figure S5.5. The stem of the DNA loop is bound to at least two DNA binding sides: the motor subunit (NIPBL) and the anchor subunit (STAG1/2). Weak DNA binding at the STAG1-kleisin interface could hand over the previously bound DNA strand to a binding site on the Smc1 ATPase head [238] since STAG1 and Smc1 can be positioned in close spatial proximity. Indeed, Ycg1 was found in close proximity to both ATPase heads in an ATP-bound state (in the absence of DNA) [243] and a structure prediction by AlphaFold2 (ref. [311]) suggests that STAG1 may be bound to the Smc1 but not to the Smc3 ATPase head (Figure S5.5b,c). In the first variant, dissociation of NIPBL could similarly transfer DNA to the DNA-binding site on the Smc3 ATPase head [238] in the ATP-unbound state (Figure 5.4 upper arc and Figure S5.5d). The DNA strands can then exchange between the Smc subunits, potentially exploiting the spatial proximity induced by ATPase head engagement upon ATP binding. Rebinding of NIPBL to cohesin selects the Smc3-bound DNA strand as the strand which is subsequently extruded into the loop.

Instead of DNA strand exchange based on the DNA binding sites on both ATPase heads, a second scenario is that NIPBL could associate with the hinge upon rebinding [238]. In this stage, the NIPBL-hinge complex could bind either the Smc3- or Smc1-proximal DNA strand, poising each one of them to be extruded into the loop in the subsequent extrusion step (Figure S5.5e). Both possibilities, DNA strand exchange *via* the DNA binding sites on the ATPase heads or *via* NIPBL-hinge association may or may not lead to DNA strand exchange and require DNA to transiently dissociate from NIPBL and STAG1. This is in line

with the observation that NIPBL turnover can, but does not necessarily lead to direction switching. Strong DNA binding at either HEAT-A or HEAT-B would inhibit these DNA-strand-exchange pathways, explaining why yeast condensin does not undergo direction switches in the presence of the strong safety belt [98, 98, 210]. Note that both pathways do not necessarily require subunit exchange, as DNA unbinding from both subunits (and for model 2 unbinding of NIPBL from Smc3 and binding to the hinge) is sufficient. The models are also sufficiently general to be applicable for direction changes observed for human condensin I and II (ref. [216]) as well as for *X. laevis* condensin and cohesin from egg extracts [214]. It is unclear whether ATP binding or hydrolysis contribute to the conformational changes required to enable DNA strand exchange. Since direction changes occur much less frequently than loop extrusion steps, the potential transient states involved in DNA strand exchange will be difficult to resolve using structural methods like cryo-EM.

5.3.3 The physiological relevance of extrusion direction switching

In human and yeast cells, NIPBL is present at a slightly sub-stoichiometric ratio to Scc1 (ref. [312–315]). However, Pds5 is more abundant than NIPBL which could potentially result in an excess of NIPBL over Pds5-unbound cohesin *in vivo*. Based on this stoichiometry, we thus expect that NIPBL turnover and direction switching occur in vivo, in line with studies which found that Scc2 is not stably associated with chromosomal cohesin [316] and hops between cohesin complexes in HeLa cells [310]. We anticipate that the direction changes described here are essential for proper genome organization by SMC complexes [317]. Simulations have shown that unidirectional extruders are unable to recapitulate the experimentally observed interphase chromosome organization, in particular the formation of TADs, even if the anchor site is only diffusively bound to DNA [268], while asymmetric extrusion with Z-loop formation and direction switching could recapitulate mitotic chromosome formation [268, 318]. In contrast, asymmetric extrusion with switching adequately reproduces the interphase chromosome organization. These simulations suggested that cohesin has to undergo direction switches approximately once per minute, yielding in the order of 10 switches before cohesin dissociation from DNA. This estimate is in fair agreement with our *in vitro* results (Figure 5.3b). Further support that direction switching by cohesin is likely to occur *in vivo* comes from the observation of extrusion "fountains" (in zebrafish [319], in *C. elegans* [320], "plumes" in mouse cells [321], and "jets" in mouse thymocytes [322]), regions of enhanced contact frequency in Hi-C maps, that emanate from a narrow region at which cohesin is preferentially loaded and broadening as extrusion proceeds away. Polymer simulations suggest that extrusion direction switching can explain the formation of these fountains [319]. In the simulations, the authors supposed that direction switching resulted from collisions between SMC proteins, but the fact that SMCs can traverse each other [267] makes this unlikely. Instead, direction switching is likely inherent to cohesin-mediated DNA loop extrusion.

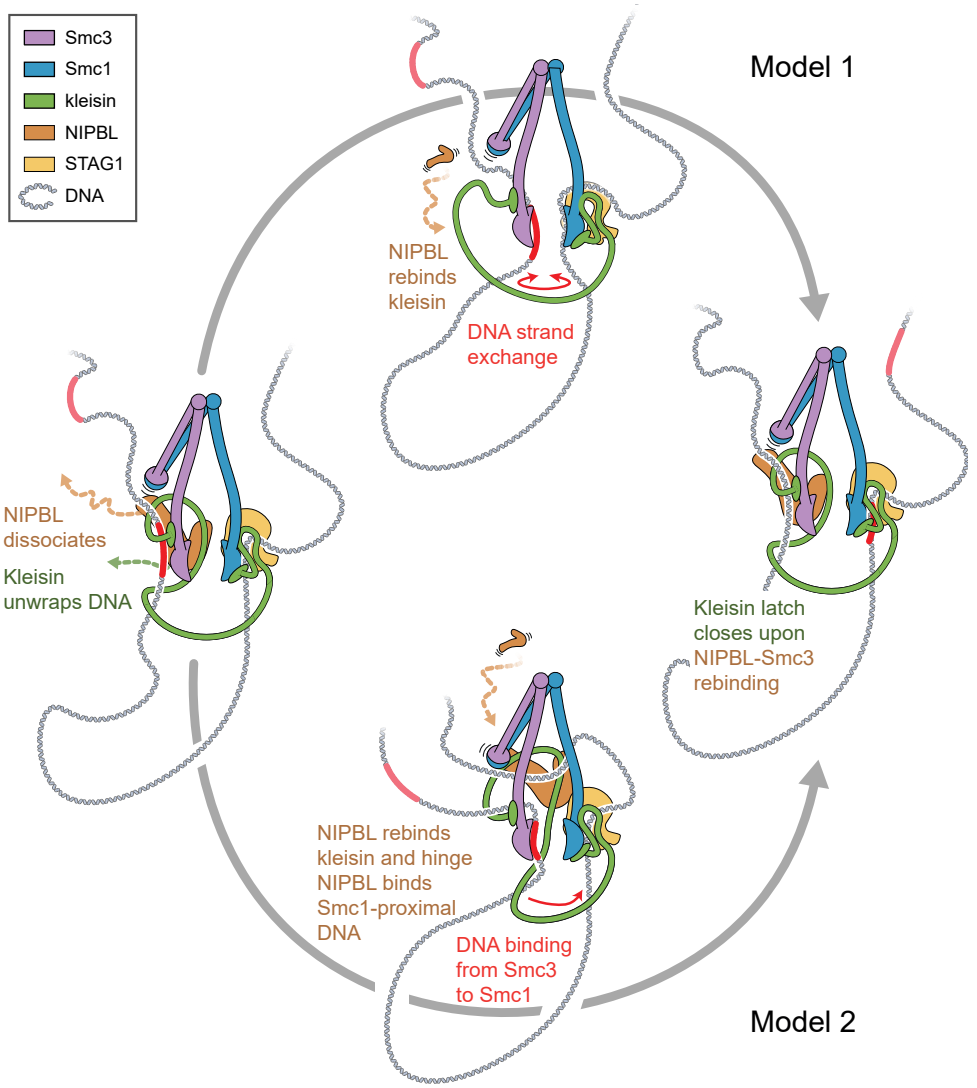


Figure 5.4 | Potential pathways of loop extrusion direction change *via* DNA strand exchange mediated by NIPBL exchange. Upon dissociation of NIPBL, the kleisin unwraps DNA (left panel). To prevent loss of the Smc3-proximal DNA (red), the Smc3 ATPase head binds DNA. DNA may unbind from STAG1-kleisin concomitantly, while the Smc1 ATPase head can bind the Smc1-proximal strand (faint DNA strand). DNA strand exchange may occur spontaneously between the ATPase heads (upper panel). Upon rebinding of NIPBL to kleisin (upper panel) and closing of the kleisin latch, the former Smc1-proximal strand is poised for the next loop extrusion cycle (right panel). Alternatively, NIPBL may re-bind the cohesin complex by binding to kleisin and the Smc hinge (lower panel). The NIPBL-kleisin-hinge complex may bind either DNA strand which is the one on which loop extrusion subsequently proceeds. If the former Smc1-proximal DNA strand is chosen, the Smc3-proximal DNA strand exchanges to the DNA binding site at the Smc1 ATPase head or at STAG1. The binding of NIPBL-kleisin to the Smc3 ATPase head prepares the complex for the next loop extrusion cycle.

5.4 Limitations of the study

In experiments which probe the turnover of NIPBL-ΔN, exchanges between NIPBL molecules carrying the same fluorophore faster than 600 ms (two time points) may appear like stably bound complexes. This might potentially account for the fact that not all direction switches were accompanied by a NIPBL exchange in these experiments (Figure 5.3l) and that direction switches were observed also within periods in which no NIPBL exchange was observed (Figure 5.3m). Early experiments demonstrating DNA loop extrusion by human cohesin were performed under sideflow [173]. In these experiments, both arms of the non-extruded DNA shrank simultaneously. The results obtained in the present manuscript, in the absence of continuous buffer flow, suggest that this symmetric extrusion is the result of rapid direction switches with concomitant NIPBL turnover. To check whether buffer flow might accelerate direction switching, we performed sideflow experiments as in Davidson et al. (ref. [173]) to probe whether short extrusion phases could be observed that alternatingly proceed along one arm or the other. We failed, within the imaging noise, to observe such periods at a time resolution of 100 ms. Furthermore, we attempted to visualize a potential (fast) NIPBL turnover on cohesin during sideflow experiments. However, due to a propensity of NIPBL to accumulate (which is to some extent alleviated by using NIPBL-ΔN), we observed only an increase of fluorescence owing to NIPBL at the loop stem and on DNA in general, preventing us from observing an exchange of single molecules.

Finally, we showed that monomeric SMC5/6 can extrude DNA loops asymmetrically. The fact that monomeric yeast SMC5/6 purified from *E. coli* as well as from yeast cells is able to extrude DNA loops is in contrast to a previous study which showed that only SMC5/6 dimers are able to extrude DNA [204]. We also observed that SMC5/6 is able to switch directions. However, our study focused on human cohesin, leaving the molecular mechanism underlying direction changes of SMC5/6 unresolved.

5.5 Methods

Experimental model and subject details

Protein expression and purification

Human cohesin, as well as NIPBL-Mau2, were expressed in and purified from *Sf9* insect cells and HeLa cells as described previously [173]. Recombinant cohesin from *Sf9* insect cells was fluorescently labelled with JF646 as described previously [173]. *S. cerevisiae* condensin was expressed in and purified from *S. cerevisiae* as described previously [98]. *S. cerevisiae* SMC5/6 was expressed, purified, and fluorescently labelled from *E. coli* as described previously [50] and chapter 6. *S. cerevisiae* SMC5/6 was expressed, purified, and fluorescently labelled from *S. cerevisiae* as described in Pradhan et al. (ref. [204]) with small modifications. The cells were lysed and bound to IgG Sepharose as described in Pradhan et al. (ref. [204]). Then the beads were washed in STO500 buffer (50 mM Tris-HCl pH 8.0, 500 mM NaCl, 2 mM $MgCl_2$, 0.5 mM tris(2-carboxyethyl)phosphine (TCEP), 10% glycerol, 0.1% IGEPAL CA-630) and eluted in 3 ml of STO500 containing TEV protease. NIPBL-ΔN was expressed and purified as described previously [173, 238] for NIPBL-MAU2 except that before elution from FLAG-M2 agarose, FLAG-M2 agarose beads were resuspended to one bead volume in 25 mM NaH_2PO_4/Na_2HPO_4 pH 7.5, 150 mM NaCl, 5 % glycerol, 50 mM imidazole pH 7.5 and split in half for fluorescent labelling with Atto550 and JF646 (see below).

Method details

Fluorescent labelling of SMC5/6 and estimation of the labelling efficiency

SMC5/6 hexamer purified from *E. coli* with a C-terminal HALO tag on Nse2 was incubated with a 2-fold molar excess of Janelia-Fluor646 HaloTag Ligand (Promega; 12 µM SMC5/6 hexamer + 24 µM label). After incubation for 1 h at 25˚C, the reaction was stopped and excess label was removed using a Zeba Spin desalting column (Thermo Fisher) equilibrated in SEC buffer (20 mM Tris pH 7.5, 250 mM NaCl, 1 mM TCEP). SMC5/6 purified from *E. coli* and JF646 concentrations were measured independently using absorption measurements at 280 nm for SMC5/6 and at 646 nm for JF646, yielding a fraction of 0.58 ± 0.11 (N = 8 independent measurements) JF646 molecules per SMC5/6 hexamer. Fluorescence Correlation Spectroscopy (FCS) was performed on samples of 10 nM SMC5/6 hexamers in a buffer containing 40 mM Tris-HCl pH 7.5, 50 mM NaCl, 2.5 mM $MgCl_2$, 1 mM TCEP at room temperature. The resulting autocorrelation functions were well described by a fit with a single component, indicating that no free fluorophores were present. SMC5/6 purified from yeast (roughly 1 µM concentration) was incubated with 20 µM of Alexa647-SNAP label overnight at 4˚C in the presence of 1 mM DTT. The complex was then purified using a Zeba Spin column equilibrated in STO500 buffer. The labelling efficiency of SMC5/6 purified from yeast was estimated similarly and resulted in a labelling efficiency of 0.22 ± 0.02 (N = 3 independent measurements).

Fluorescent labelling of NIPBL-ΔN

Janelia Fluor 646 SE (Tocris; 6148) was dissolved in dimethylformamide (DMF) at 20 mM. HaloTag Amine (O2) Ligand (Promega; P6711) was dissolved in dimethylformamide (DMF) at 50 mM. HaloTag Amine (O2) Ligand was added to JF646 SE in a 248.7 µl reaction in DMF

while stirring and was then supplemented with 1.3 µl diisopropylethylamine (Sigma) (final concentrations: 7 mM JF646 SE, 2.3 mM HaloTag ligand, 31 mM diisopropylethylamine). Labelling reactions were incubated at room temperature for ~ 16 hours protected from light. The reaction mixture was diluted 10-fold in solvent A (5 % acetonitrile, 0.1 % formic acid) and purified in five successive reverse phase HPLC runs using an Ultimate 3000 (ThermoFisher Scientific) equipped with a Kinetex 5µ XB- C18 100A, 250 x 4.6mm column using the following gradient of solvents: A (5 % acetonitrile, 0.1 % formic acid), B (acetonitrile, 0.1 % formic acid); 0 - 100 % B in 30 min at a flow rate of 0.8 ml/min. Peak fractions were pooled, lyophilized, resuspended in DMSO, frozen in liquid nitrogen and stored at -80°C. Fractions were analysed by incubating with purified HaloTag protein for 15 min at room temperature followed by SDS-PAGE and excitation using a Bio-Rad ChemiDoc MP. Atto550 NHS-ester (ATTO-TEC; AD 550-35) was dissolved in dimethylformamide (DMF) at 20 mM. HaloTag Amine (O2) Ligand was added to Atto550 NHS-ester in a 245 µl reaction in DMF while stirring and was then supplemented with 5 µl diisopropylethylamine (Sigma) (final concentrations: 25 mM Atto550 NHS-ester, 8.3 mM HaloTag ligand, 115 mM diisopropylethylamine). Labelling reactions were performed as for Janelia Fluor 646 SE, except that reaction mixture was diluted 10-fold in solvent A (40% acetonitrile, 0.1 % Trifluoroacetic acid) and reverse phase HPLC was performed using the following gradient of solvents: A (40% acetonitrile, 0.1 % Trifluoroacetic acid), B (acetonitrile, 0.1 % Trifluoroacetic acid); 0 - 100 % B in 30 min at a flow rate of 0.8 ml/min. To label recombinant NIPBL-ΔN with Atto550-HaloTag Ligand and JF646-HaloTag Ligand, the purification was split in half. Each aliquot was supplemented with excess Atto550-HaloTag Ligand or JF646-HaloTag Ligand and incubated for 15 min at room temperature protected from light. Beads were washed extensively with recombinant cohesin purification buffer 3 and bound protein was eluted and concentrated as described for NIPBL-MAU2 in [173].

Double-tethered DNA assay for single-molecule imaging of DNA loop extrusion

DNA loop extrusion experiments and imaging were performed essentially as described previously [35, 98, 226, 267, 323] with one round of pegylation with 5 mg/ml methoxy-PEG-N-hydroxysuccinimide (MW 3500, LaysanBio) and 0.05 mg/ml biotin-PEG-N-hydroxysuccinimide (MW 3400, LaysanBio) in 0.1 M NaHCO3, 0.55 M K_2SO_4 buffer [324]. Subsequently, three more rounds of pegylation with $MS(PEG)_4$ (MW 333.33, Thermo Fisher) in the same buffer were applied. Slides and coverslips were incubated overnight at 4°C in the dark and then washed with ultrapure water and dried by a N_2 stream. After assembly of flow chambers by sandwiching coverslip and drilled glass slide using double-sided tape and sealing the edges using epoxy glue, 0.5 mg/ml streptavidin was incubated for 1 min in loop extrusion buffer (see below) omitting glucose oxidase, intercalating dyes, ATP, and protein. The flow cell was then washed with 100 µl of the same buffer. End-biotinylated λ-DNA (48.5 kbp, ref. [98]) was then introduced at a flow speed to reach an end-to-end extension of about 6-7 µm (30% of the contour length). Unbound DNA was washed out with up to 200 µl buffer. The DNA was found to be nicked, preventing a buildup of plectonemic supercoils during DNA loop extrusion [48, 49, 323, 325]. DNA Loop extrusion reactions by human cohesin were carried out in a buffer containing 40 mM Tris-HCl pH 7.5, 50 mM NaCl (unless otherwise denoted), 2.5 mM $MgCl_2$, 2.5% D-glucose, 2 mM Trolox, 10 nM catalase, 18.75 nM glucose oxidase, 100 nM Sytox Orange, 0.5 mg/ml BSA, 1 mM DTT, 1 mM ATP with 10-100 pM

cohesin (from *Sf9* insect cells) and the indicated concentration of NIPBL-Mau2 at 37°C. For experiments in which cohesin-JF646 was photobleached, glucose oxidase was omitted from the loop extrusion buffer. For experiments with differentially labelled NIPBL-ΔN, the buffer was adjusted to 25 mM NaCl and 25 nM Sytox Green. Part of the experiments were performed with cohesin purified from HeLa cells in which case 1 nM HeLa cohesin was used and part was performed with cohesin purified from *Sf9* cells. The two cohesin purifications behaved similarly and the results of these experiments were pooled. DNA loop extrusion experiments with yeast condensin were carried out in a buffer containing 40 mM Tris-HCl pH 7.5, 50 mM NaCl, 2.5 mM $MgCl_2$, 2.5% D-glucose, 2 mM Trolox, 10 nM catalase, 18.75 nM glucose oxidase, 100 nM Sytox Orange, 0.5 mg/ml BSA, 1 mM DTT, 1 mM ATP with 1 nM cohesin at room temperature (22 °C). DNA loop extrusion experiments with yeast SMC5/6 from *E. coli* were carried out in a buffer containing 40 mM Tris-HCl pH 7.5, 50 mM NaCl, 2.5 mM $MgCl_2$, 2.5% D-glucose, 2 mM Trolox, 10 nM catalase, 18.75 nM glucose oxidase, 100 nM Sytox Orange, 0.5 mg/ml BSA, 1 mM TCEP, 1 mM ATP with 3 nM SMC5/6 hexamer at 30°C or in the same buffer but with 100 mM NaCl, 7.5 mM $MgCl_2$ as in chapter 6 [50]. DNA loop extrusion experiments with yeast SMC5/6 from yeast were performed in the same buffer with 100 mM NaCl, 7.5 mM $MgCl_2$ as in Pradhan et al. (ref. [204]). All data were acquired using an exposure time of 100 ms per channel unless otherwise stated. For DNA loop extrusion without monitoring differentially labelled NIPBL-ΔN, the 561 nm-channel was used to image Sytox Orange-stained DNA and the 640 nm-channel was used to alternatingly monitor the SMC complex. The effective time resolution in these experiments was thus 200 ms. For experiments with differentially labelled NIPBL-ΔN, the 480 nm channel was used to image Sytox Green-stained DNA, the 561 nm channel was used to image NIPBL-ΔN-A550, and the 640 nm channel was used to image NIPBL-ΔN-JF646. The time resolution in these experiments was thus 300 ms. Experiments to visualize exchanges of NIPBL-Mau2 were performed with the N-terminal truncation of NIPBL, NIPBL-ΔN-A550 and NIPBL-ΔN-JF646 because NIPBL-ΔN has a smaller propensity to accumulate/aggregate during DNA loop extrusion experiments and was thus more suitable to image single NIPBL molecules. To evaluate whether imaging at 200 ms time resolution potentially omits short phases of loop extrusion, diffusion, or slipping, experiments with human cohesin were also performed while imaging Sytox Orange-stained DNA at 20 ms exposure time, yielding thus a 10-fold higher sampling rate compared to the remainder of the experiments. Image series were acquired for 15 min after flush-in of SMC proteins.

Fluorescence Correlation Spectroscopy (FCS)

Fluorophore diffusion measurements were conducted using version 1 coverslips on a Picoquant Microtime 200 microscope, which was operated using Symphotime software at room temperature. To focus a 640 nm laser, a 60x Olympus UPSLAPO 60XW water immersion objective with a working distance of 280 µm and a numerical aperture of 1.2 was used. Prior to the experiment, the molecular brightness of an Alexa647 fluorophore solution was optimized by adjusting the correction collar of the objective. The emission light was subsequently directed through a 50 µm pinhole, split by a dichroic mirror, and filtered through a 600/75 optical band pass filter (Chroma, Bellow Falls). Fluorescence emission was collected by single-photon avalanche-diode detectors (PD5cTC and PD1cTC, Micro Photon Devices, Bolzano). For fitting the FCS curves, the size of the confocal volume was determined from measurements of the free Alexa647 by fitting a single-component

diffusion model with triplet state (a diffusion constant of 297 $\mu m^2/s$ was used [326]). The axial and lateral sizes of the confocal volume were fixed for further analysis. FCS amplitudes and diffusion coefficients were subsequently fitted for curves recorded on SMC5/6-JF646 containing samples.

Mass photometry experiments

Coverslips (Menzel, 24x60mm, #1.5) were washed three times alternatingly with isopropanol and milli-Q water, dried in a nitrogen stream and mounted onto a Refeyn One^{MP} device (Refeyn). Data was acquired using the AcquireMP software (Refeyn Ltd.) and analysed using DiscoveryMP (Refeyn Ltd.). Before the measurement, a mass calibration was recorded to allow conversion between contrast to mass using a protein standard (Invitrogen NativeMark^{TM} Unstained Protein Standard) with masses of 1236, 1048, 720, 480, 242, 146, 66, and 20 kDa, respectively. A buffer containing 40 mM Tris-HCl pH 7.5, 50 mM NaCl, 2.5 mM $MgCl_2$, and 1 mM DTT was filtered through a 0.22 µm syringe filter (Merck Millex®-GV). 15 µl buffer was added onto the mounted coverslip and the focus was found automatically and stabilized. Single-chain cohesin or STAG1 were added to the droplet to a final concentration of 10 nM, briefly mixed by pipetting and acquired immediately. Experiments were performed at room temperature (22°C). When single-chain cohesin was measured together with STAG1, single-chain cohesin and STAG1 were mixed in an equimolar ratio at their respective stock concentrations on ice without prolonged incubation and then immediately added to the droplet to yield a final concentration of 10 nM each. Immediately after dilution, a series of 6000 frames was acquired at 1 kHz. All samples were measured three times.

Protein complex structure prediction using AlphaFold2

AlphaFold2 predictions were generated using ColabFold, running on Google Colab [327] using the following parameters: msa_mode=mmseqs2_uniref_env; pair_mode=unpaired_paired; num_models=1 or 5 (predictions of large complexes quickly run out of GPU memory in which case only 1 model could be predicted); num_recycles=3; recycle_early_stop_tolerance=auto; max_msa=auto; num_seeds=1; use_dropout=True. The sequences used to generate the predictions are listed in Supplementary Information section 5.6. Protein structures were rendered using UCSF ChimeraX (Figure 5.4a-f, ref. [328]).

Quantification and statistical analysis

Quantification and segmentation of kymographs

Quantification of DNA loop size and position was performed as described previously [35]. In brief, the intensity of the loop is normalized to the intensity along the entire DNA molecule and multiplied by the known length of the DNA molecule (48.5 kbp). The loop position was determined as the relative position of the loop from one randomly selected end of the DNA for each molecule but kept constant for all analyses. Half of the loop size at every moment was added to this position. The reported loop position (in kbp) refers thus to the tip of the extruded loop. DNA loop size and position were plotted concomitantly versus time (Figure 5.1b) and in a two-dimensional plot (Figure S5.1a,b). Segmentation of the kymographs into phases of loop extrusion, loop diffusion, and loop slipping was performed manually. As illustrated in Figure 5.1c, the time course of the loop size was first

segmented into phases of loop growth (loop extrusion), stagnant loop size (loop diffusion), and loop size decrease (loop slipping). Phases in which loops colocalize with the DNA ends (5% of the DNA length at each side) were labelled as 'stalled' and excluded from further analysis. Loop extrusion and slipping phases were further divided into segments if the direction changed, as determined by the local slope of the loop position over time. After manual segmentation, the classification of segments into extrusion, diffusion, and slipping phases was done semi-automatically by applying a threshold used to distinguish loop diffusion from loop extrusion (loop size increase) and slipping (loop size decrease). The direction of loop extrusion and slipping phases were assigned based on the sign of the loop position derivative with respect to the time within each extrusion and slipping phase, respectively. To determine whether loop extrusion phases proceed asymmetrically or symmetrically, the amount of DNA towards both sides (L_{dir1} and L_{dir2} in Figure 5.2) were quantified by dividing the relative fluorescence of both DNA segments to the left and right of the loop by the length of the DNA molecule (48.5 kbp), as done previously [98, 204, 214]. DNA loop extrusion traces by human cohesin were usually acquired at a time resolution of 200 ms (100 ms exposure time for the DNA and cohesin-JF646 channel, respectively, in the absence of labelled NIPBL-ΔN). To check whether phases of loop extrusion, diffusion, or slipping exist which are too short to be segmented with a time resolution of 200 ms, cohesin-mediated DNA loop extrusion traces were also acquired at a time resolution of 20 ms (only the DNA channel was imaged). This data set was then segmented and classified using the time resolution of 20 ms. Additionally, the same data set was subsampled by averaging 5 subsequent frames of the DNA channel (thus effectively increasing the exposure time to 100 ms) and omitting the following 5 frames (thus mimicking the imaging of cohesin-JF646 for another 100 ms). The subsampled data set was segmented and classified independently. The segmentation results of both data sets were quantitatively compared to check for short phases in the 20 ms data set which were missed in the subsampled data set (Figure S5.1c-f). 76 ± 17 % (probability and 95% binomial confidence interval) of all traces were segmented into an equal number of phases (Figure S5.1c; Jaccard index of 0.80 [N_{union} = 184, $N_{intersection}$ = 148]) and the change points were set, on average, at the same time (mean time difference 0.14 s; Figure S5.1d) with low variation between individual change points (SD = 1 s; Figure S5.1d). The fraction of uni- vs. bidirectional traces was identical (Figure S5.1e) and there was no significant change between the duration of loop extrusion, diffusion, and slipping phases (Figure S5.1f). We conclude that a time resolution of 200 ms is sufficient to segment traces of DNA loop extrusion into phases of loop extrusion, diffusion, and slipping and that no significantly shorter phases exist.

5

Quantification of DNA loop extrusion, loop diffusion, and loop slipping kinetics

The following metrics were calculated which characterize DNA loop extrusion traces of all eukaryotic SMC complexes: loop extrusion rate, loop slipping rate, diffusion constant, loop lifetime, the fraction of unidirectional traces, frequency of direction change, and the fraction of diffusing and slipping loops. The loop extrusion (Figure S5.3g) and loop slipping rate (Figure S5.4i) were determined *via* a linear fit to the loop size across the loop extrusion or loop slipping phase, respectively. To quantify the diffusion constant of loops during diffusion phases (Figure S5.4h), the mean squared displacement (MSD) of the loop during

the diffusion phase was calculated. The diffusion constant was subsequently determined *via* linear regression of the form MSD(τ) = Dτ + o (diffusion constant D, time lag τ, and offset o due to finite localization uncertainty) to the first 10% of the available time lags for sufficient data points of the fit and to avoid flattening of the MSD curve when the loop reaches the DNA ends. The loop lifetime (Figure 5.3c) was determined as the interval from the beginning of the first loop extrusion phase until the complete dissociation of the loop. In cases in which acquisition was stopped before loop dissociation, the loop lifetime was computed from the beginning of loop extrusion until the end of acquisition. A trace was labelled as extruding toward one direction only (unidirectional; Figure 5.1e and Figure 5.3a) if a trace showed at least two extrusion phases in only one direction and no extrusion phase in the other direction. A trace was labelled as extruding bidirectionally if the trace exhibited at least one extrusion phase in each direction. Traces which exhibited only one extrusion phase were not considered because it is unclear whether the SMC complex could only undergo a single phase of extrusion despite being a bidirectional extruder or whether the SMC complex was indeed a unidirectional extruder. The average frequency of direction changes was computed for each loop by dividing the number of direction switches per trace by the loop lifetime (Figure 5.3b) for traces which showed at least one direction switch. The fraction of traces which showed loop diffusion (Figure S5.3c) and loop slipping (Figure S5.3d) was computed as the fraction of traces with at least one diffusion or slipping phase, respectively.

Bleaching and annotation of NIPBL-ΔN presence

The intensity of cohesin-JF646 (in bleaching experiments; Figure S5.3e-g) and NIPBL-ΔN-A550/JF646 was computed as the average intensity in the respective channel at the loop position over a window of 5 pixels around the loop centroid position. The number of bleaching steps of cohesin-JF646 was determined by plotting the cohesin-JF646 intensity at the loop over time and counting the number of step-wise decreases to the background level (Figure S5.3e). Only cohesin complexes associated with a DNA loop were considered, which allows an assignment of the number of cohesin-JF646 bleaching steps to traces showing uni- (Figure S5.3f) or bidirectional extrusion (Figure S5.3g). The appearance and disappearance of NIPBL-ΔN-A550/JF646 (Figure 5.3h-m and Figure S5.4b-i) was assessed analogously.

Testing the Markov property

Let us denote the state k of cohesin as being in an extrusion, diffusion, or slipping phase. A finite, first-order, discrete Markov chain is a stochastic process in which the probability to transition from state k to state l depends only on the current state (state k) but on states at previous points in time [329]: $P\{X(t+1) = l \,|X(0) = k_0, \ldots, X(t-1) = k_{t-1}, X(t) = k\} = P\{X(t+1) = l \mid X(t) = k\} = p_{kl}$. The Markov chain is thus determined by the transition matrix Π which summarizes all 3x3 possible transitions (here we consider three possible

states):

$$\Pi = \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix} \tag{5.1}$$

$$p_{kl} \geq 0 \tag{5.2}$$

$$p_{kk} = 0 \tag{5.3}$$

$$\sum_{l=1}^{3} p_{kl} = 1 \tag{5.4}$$

The transition probability p_{kl} is computed as the number of transitions from state k to state l, N, divided by the number of transitions from state k to any state, i.e. $p_{kl} = N_{kl} / \sum_l N_{kl}$.

Alternatively, the stochastic process might have memory, in which case the transition probability from state k (at time t) to state l (at time $t+1$) depends also on the previous state j (at time $t-1$), i.e. $p_{jkl} = P\{X(t+1) = l \mid X(t) = k,\ X(t-1) = j\}$. A three-state transition matrix is constructed analogously to the two-state transition matrix Π. The chi-square statistic is subsequently used to test whether the Markov chain is a first order (memory-less, the null hypothesis H_0) or second order (with memory, H_1) process [330]. The chi-square statistic for the null hypothesis of state k (in other words: the hypothesis that the transition from the current state k to the future state l is independent of the past state j) is:

$$Q_k = \sum_{jl} N_{jk} \left(p_{jkl} - p_{kl} \right)^2 / p_{kl} \tag{5.5}$$

Where $p_{kl} = \sum_{j=0}^{3} N_{jkl} / \sum_{j=0}^{3} \sum_{m=0}^{3} N_{jkm}$ and $N_{jk} = \sum_l N_{jkl}$. If H_0 is true, Q_k has a chi-square distribution with $(3-1)^2$ degrees of freedom. Similarly, the joint hypothesis that the transition from all states is memory-less, i.e. that $p_{jkl} = p_{kl}\ \forall\ j,\ k,\ l = 0,\ 1,\ 2$, can be computed *via* the sum $Q = \sum_k Q_k$ which has a chi-square distribution with $3(2-1)^2$ degrees of freedom. The null hypothesis H_0 is rejected if the test result $P(X > Q_k)$ is significant at the significance level of $\alpha = 0.05$. The p-values for the three states are all > 0.05 across all conditions (Figure S5.3a), thus cohesin switches from one state to another without memory of the past state.

5.6 Supplementary Information

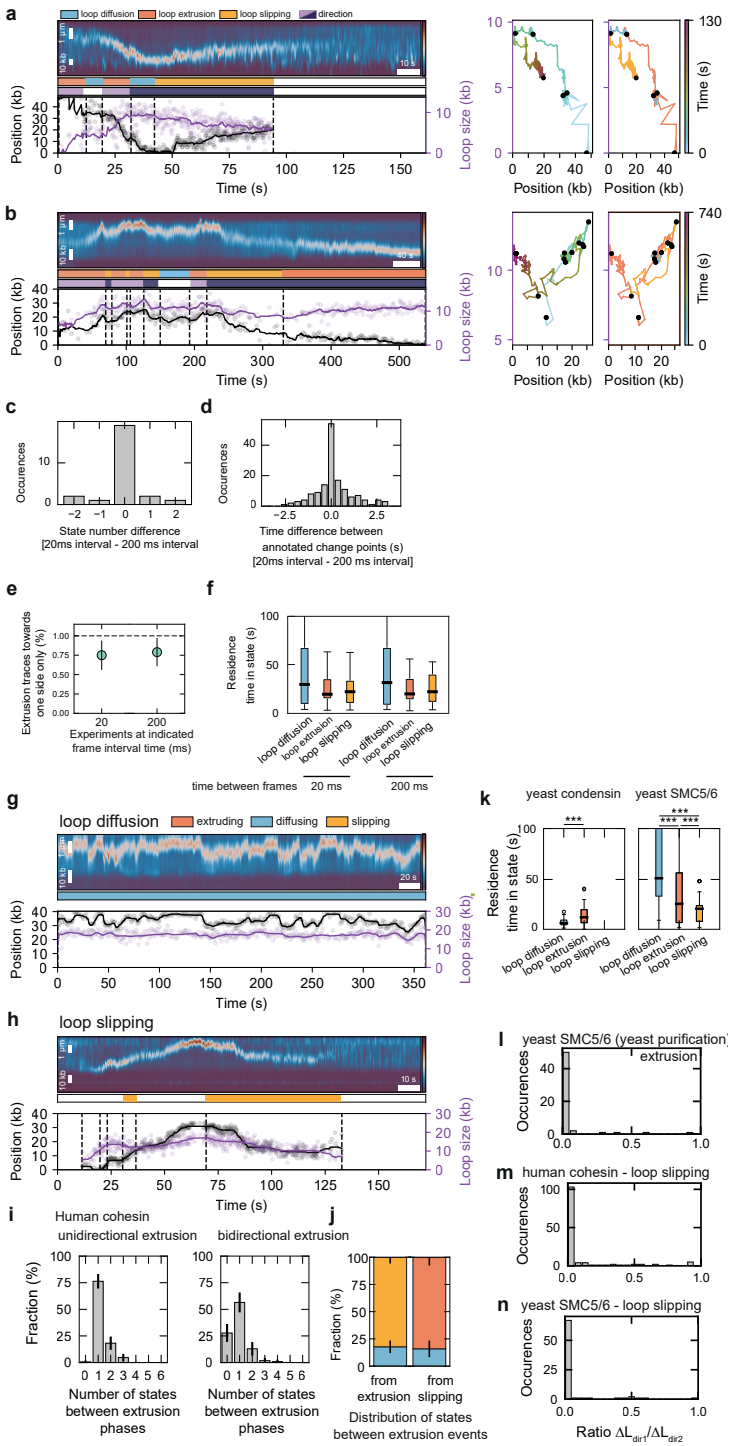


Figure S5.1 | Comparison of DNA loop extrusion trace segmentation between data acquisition with 20 ms and 200 ms time resolution. Loop diffusion and slipping examples, state residence times for yeast SMC5/6 and yeast condensin. Loop slipping is, like extrusion, asymmetric. *Caption continues on next page.*

Figure S5.1 | (**a**) Kymograph of cohesin-mediated DNA loop extrusion (left). Loop position and size are quantified and segmented into loop extrusion, diffusion, and slipping phases. The two panels on the right display a two-dimensional plot of the loop size versus loop position. Black dots demarcate boundaries between diffusion, extrusion, and slipping phases. Diffusion phases move on horizontal lines in this graph, and extrusion and slipping phases are demarcated by diagonal lines. (**b**) as (a) but for a loop extrusion trace in which extrusion phases in both directions are present as shown in Figure 5.1b. (**c**) Difference in the number of phases (segments) between the data set acquired at 20 ms frame interval time and the subsampled one (interval time 200 ms). Negative values indicate that fewer segments were segmented in the 20 ms interval data set than in the 200 ms data set. (**d**) Time between the annotated change points (time points between segments) in the data set with an interval time of 20 ms and 200 ms between frames. 74% of all change points coincide between the two annotations within 1 s (N = 147 segments). Negative values indicate that the change point in the 20 ms interval data set was annotated prior to the respective change point in the 200 ms data set. (**e**) Fraction of traces with at least two extrusion events that displayed only unidirectional extrusion (N = 25 traces each). Error bars denote the binomial 95% confidence interval. (**f**) Duration of diffusion, extrusion, and slipping phases in the annotated data sets acquired at 20 ms frame interval time and the subsampled one (interval time 200 ms). Black horizontal lines are median values, the box extends between the first and third quartile, and the whiskers extend to 1.5·IQR. No statistically significant difference was found between loop diffusion, extrusion, and slipping states between the two data sets. Statistical significance was assessed by a Mann-Whitney test with Bonferroni correction (p-values are 0.9514, 0.6504, and 0.4710 for loop diffusion, extrusion, and slipping states, respectively). (**g**) Exemplary kymograph of human cohesin showing loop diffusion. (**h**) Exemplary kymograph of human cohesin showing loop slipping. (**i**) Duration of diffusion, extrusion, and slipping states for yeast condensin (N = 16, 48, 1, respectively) and yeast SMC5/6 (N = 52, 124, 80, respectively). Black horizontal lines are median values, the box extends between first and third quartile, and whiskers extend to 1.5·IQR. Statistical significance was assessed by a Mann-Whitney test with Bonferroni correction. (***: p<0.001). (**j**) Number of intervening states between extrusion events for traces showing only unidirectional (N = 148) or bidirectional extrusion (N = 108). Bar height denotes the probability of finding the annotated number of steps between extrusion events. Error bars denote the binomial 95% confidence interval. (**k**) Distribution of states between direction exchange in extrusion (N = 90) and slipping (N = 88) events. Error bars denote the binomial 95% confidence interval. (**l**) Distribution of the ratio $\Delta L_{dir1}/\Delta L_{dir2}$ for extrusion events of yeast SMC5/6 purified from yeast cells (N = 56 events). Yeast SMC5/6 purified from yeast cells exhibits asymmetric extrusion phases (93 ± 7 % [mean ± 95% binomial confidence interval] of $\Delta L_{dir1}/\Delta L_{dir2}$ values below 0.1). (**m**) Distribution of the ratio $\Delta L_{dir1}/\Delta L_{dir2}$ for slipping events of human cohesin (N = 130 events). (**n**) Distribution of the ratio $\Delta L_{dir1}/\Delta L_{dir2}$ for slipping events of yeast SMC5/6, purified from *E. coli* (N = 77 events).

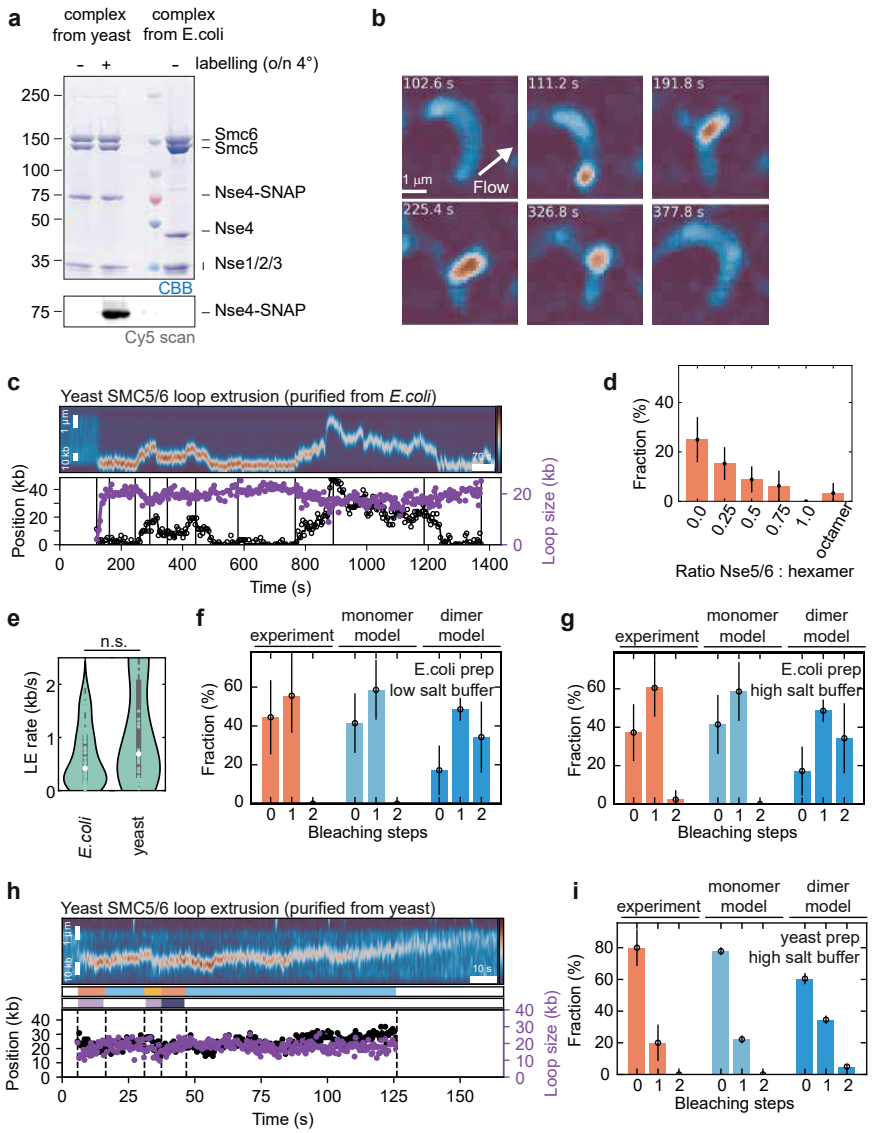


Figure S5.2 | DNA loop extrusion by single SMC5/6 hexamers.
(**a**) Detection of SMC5/6 hexamer subunits by SDS-PAGE and Coomassie staining for the unlabeled (1st lane), Nse4-Alexa647-labeled (2nd lane) SMC5/6 complex purified from yeast cells, as well as for the unlabeled SMC5/6 purified from *E. coli* as in [309]. (**b**) Fluorescence imaging snapshots showing DNA loop extrusion by an SMC5/6 hexamer purified from *E. coli* during side flow. (**c**) Exemplary kymograph of DNA loop extrusion by an SMC5/6 hexamer purified from *E. coli*. (**d**) Loop extrusion efficiency with varying ratio of Nse5/6 to the SMC5/6 hexamer purified from *E. coli*. No loop extrusion is observed at an equimolar ratio of Nse5/6 and the SMC5/6 hexamer. A co-purification of Nse5/6-bound SMC5/6 hexamer (octamer) shows a small fraction of loop extrusion, potentially due to dissociation of Nse5/6 from the SMC5/6 hexamer that allows loop extrusion. Bar heights denote the probability of finding a DNA with a loop and the error bar denotes the binomial 95% confidence interval (N = 92, 124, 124, 64, 74, 84 DNA molecules from left to right). (**e**) DNA loop extrusion rate of the SMC5/6 hexamer, purified from *E. coli* (ref. [309]) and from yeast. The white dot marks the median value (0.4 ± 0.5, median ± SD, N = 83 for *E. coli*; 0.7 ± 0.9, median ± SD, N = 30 for *E. coli*), the box extends to the quartiles, and the whiskers denote the interquartile range. Statistical significance was assessed by a two-sided Kolmogorov-Smirnoff test (p = 0.087). *Caption continues on next page.*

Figure S5.2 | (**f**) Bleaching step distribution of JF646-labeled SMC5/6 hexamers purified from *E. coli* in 50 mM NaCl, 2.5 mM $MgCl_2$ with a labeling efficiency of 58 ± 11% (N = 27 for the experiment; N = 1,000 simulated monomer/dimer complexes). The monomer model fits the experimental data while the dimer model does not. Error bars denote the 95% binomial confidence interval for experimental data and the range of the predicted monomer/dimer bleaching step distributions for a labeling efficiency of 47% (lower end) and 69% (upper end). (**g**) As for (f) but in a 100 mM NaCl, 7.5 mM $MgCl_2$ buffer, which is the buffer in which dimers of SMC5/6 hexamers were reported to extrude by Pradhan et al. [204]. (**h**) Exemplary kymograph of DNA loop extrusion by an SMC5/6 hexamer purified from yeast. (**i**) As for (g) but for SMC5/6 purified from yeast with a labeling efficiency of 22 ± 2% (N = 50 for the experiment; N = 1,000 simulated monomer/dimer complexes) in a 100 mM NaCl, 7.5 mM $MgCl_2$ buffer, in which dimers of SMC5/6 hexamers were reported to extrude by Pradhan et al. [204]. Error bars denote the 95% binomial confidence interval for experimental data and the range of the predicted monomer/dimer bleaching step distributions for a labeling efficiency of 20% (lower end) and 24% (upper end).

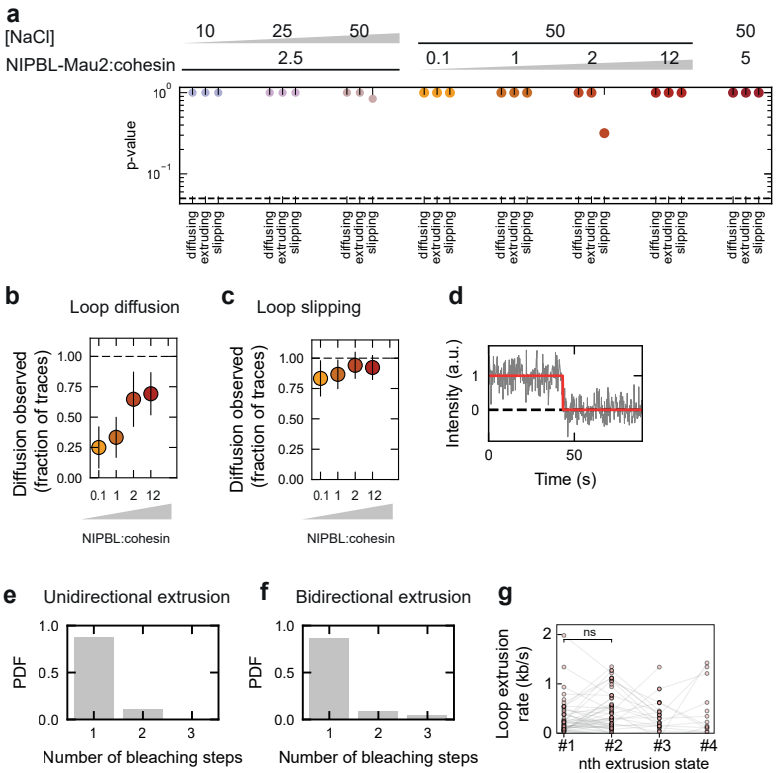


Figure S5.3 | The sequence of loop diffusion, extrusion, and slipping events is a Markov chain and the excess of NIPBL-Mau2 and STAG1 over cohesin modulate the occurrence of diffusion and slipping phases. Bidirectional extrusion is the result of monomeric cohesin complexes.
(**a**) p-value of the Markov property test for transitions starting from a loop diffusion, extrusion, and slipping state for all tested NaCl concentrations and NIPBL-Mau2:cohesin ratios. A value below 0.05 (dashed line) indicates the violation of the Markov property and would suggest a system with memory. (**b**) Fraction of diffusing cohesin complexes for varying NaCl concentrations (NIPBL-Mau2:cohesin = 2.5, N = 9, 20, 13 from left to right), for varying ratios of NIPBL-Mau2:cohesin (50 mM NaCl, N = 24, 30, 17, 26 from left to right). Dots denote the fraction, error bars denote the binomial 95% confidence interval. (**c**) As for (b) but for loop slipping. Number of data points as in (b). (**d**) Exemplary bleaching trace showing a single bleaching step. (**e**) Number of bleaching steps of human cohesin-JF646 for complexes which extrude unidirectionally (N = 128). (**f**) Number of bleaching steps of human cohesin-JF646 for complexes which extrude bidirectionally (N = 23). (**g**) Loop extrusion rate for the nth extrusion state within single kymographs.

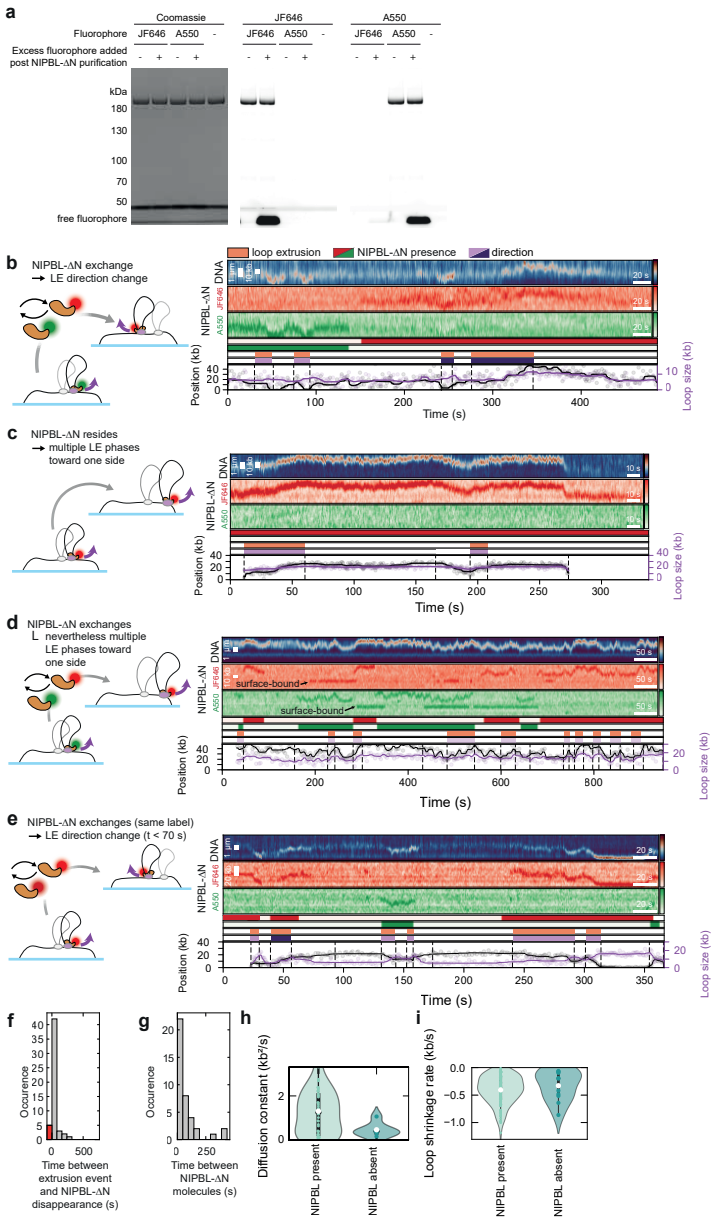


Figure S5.4 | Additional loop extrusion traces with differentially labeled NIPBL-ΔN.
(**a**) Purity and labeling of NIPBL-ΔN with JF646 and ATTO550, respectively. Molecular weights are indicated in kilo Daltons (kDa). An excess of dye was added to the second and fourth lane, respectively, to check whether further labeling can be achieved. (**b**) Exemplary trace where cohesin core complex associated with a A550-labeled NIPBL-ΔN first extrudes downward. Upon exchange to a JF646-labeled NIPBL-ΔN at ~150 s, the cohesin complex extrudes upwards. (**c**) Exemplary trace where a JF646-labeled NIPBL-ΔN resides at cohesin which extrudes solely unidirectionally (here towards the upper end). (**d**) Example of DNA loop extrusion with exchange of NIPBL-ΔN without direction change. (**e**) Example of DNA loop extrusion with exchange of NIPBL-ΔN and concomitant direction change between ~45 s and ~125 s. Another NIPBL-ΔN exchange occurs at ~240 s but the direction remains (downward here). *Caption continues on next page.*

Figure S5.4 | (**f**) Distribution of the time interval between NIPBL-ΔN disappearance and the end of the preceding extrusion phase (N = 45). The red bar represents the occurrences when NIPBL-ΔN disappearance coincided with the end of an extrusion event. (**g**) Distribution of time between disappearance of one and appearance of the next NIPBL-ΔN molecule (N = 34). (**h**) Diffusion constant of the DNA loop for diffusion phases in the presence (N = 45) and absence (N = 6) of NIPBL-ΔN. White dots denote median values, the box extends between first and third quartile, and whiskers extend to 1.5·IQR. Statistical significance was assessed by a Mann-Whitney test (not significant: $p > 0.05$). (**i**) Loop shrinkage rate in the presence (N = 109) and absence (N = 11) of NIPBL-ΔN. White dots denote median values, the box extends between first and third quartile, and whiskers extend to 1.5·IQR. Statistical significance was assessed by a Mann-Whitney test (not significant: $p > 0.05$).

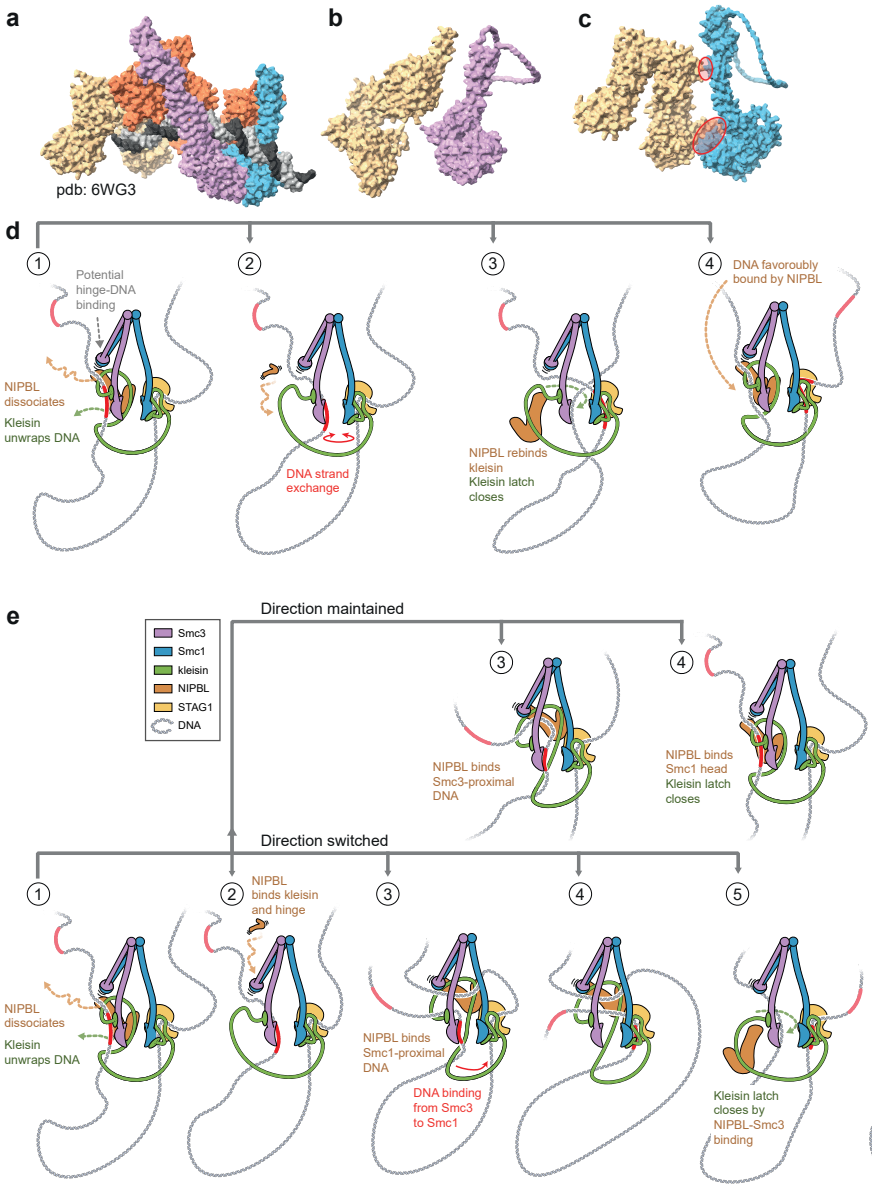


Figure S5.5 | Potential pathways of loop extrusion direction change *via* DNA strand exchange mediated by NIPBL exchange.
(**a**) Experimentally solved structure of the cohesin-NIPBL-DNA complex in the engaged state (pdb 6WG3 , ref. [200]). (**b**) AlphaFold2 prediction of the Smc3 head domain including a part of the coiled coils (connected by a 64x G linker) and STAG1. No interaction is apparent. Kleisin was included during the folding but is not shown for clarity. *Caption continues on next page.*

Figure S5.5 | (**c**) AlphaFold2 prediction of the Smc1 head domain including a part of the coiled coils (connected by a 64x G linker) and STAG1. Interactions between STAG1 and Smc1 are apparent close to DNA-binding patches 1 and 2 which were identified by Bauer et al. (ref. [238]). Kleisin was included during the folding but is not shown for clarity. (**d**) Upon dissociation of NIPBL, the kleisin unwraps DNA (step 1). To prevent loss of the Smc3-proximal DNA (red), the Smc3 ATPase head or the Smc hinge may bind DNA. DNA may unbind from STAG1-kleisin concomitantly (step 2). DNA strand exchange may occur spontaneously between the ATPase heads (step 2). Upon rebinding of NIPBL to kleisin (step 3) and closing of the kleisin latch, the former Smc1-proximal strand is poised for the next loop extrusion cycle (step 4). (**e**) Step 1 is as described in (d). Alternatively, to the pathway sketched in (a), NIPBL may re-bind the cohesin complex by binding to kleisin and the Smc hinge (step 2). The NIPBL-kleisin-hinge complex may bind either DNA strand which is the one on which loop extrusion subsequently proceeds. If the former Smc1-proximal DNA strand is chosen (step 3, lower panel), the Smc3-proximal DNA strand exchanges to the DNA binding site at the Smc1 ATPase head or at STAG1 (steps 3 and 4, lower panel). If the Smc3-proximal strand is chosen (step 3, upper panel), no direction change occurs. The binding of NIPBL-kleisin to the Smc3 ATPase head prepares the complex for the next loop extrusion cycle (step 5, lower panel; step 4, upper panel).

Sequences used to generate AlphaFold2 predictions

Smc3

GFLKLIEIENFKSYKGRQIIGPFQRFTAIIGPNGSGKSNLMDAISFVLGEKTSNLRVKTLRDLIHGAPVMGFLKLIEIENFKSYKGRQIIGPFQRFTAIIGPNGSGKSNLMDAISFVLGEKTSNLRVKTLRDLIHGAPVGKPAANRAFVSMVYSEEGAEDRTFARVIVGGSSEYKINNKVVQLHEYSEELEKLGILIKARNFLVFQGAVESIAMKNPKERTALFEEISRSGELAQEYDKRKKEMVKAEEDTQFNYHRKKNIAAERKEAKQEKEEADRYQRLKDEVVRAQVQLQLFKLYHNEVEIEKLNKELASKNKEIEKDKKRMDKVEDELKEKKKELGKMMREQQQIEKEIKEKDSELNQKRPQYIKAKENTSHKIKKLEAAKKSLQNAQKHYKKRKGDMDELEKEMLSVEKARQEFEERMEEESQSQGRDLTLEENQVKKYHRLKEEASKRAATLAQELEKFNRDQKADQDRLDLEERKKVETEAKIKQKLREIEENQKRIEKLEEYITTSKQSLEEQKKLEGELTEEVEMAKRRIDEINKELNQVMEQLGDARIDRQESSRQQRKAEIMESIKRLYPGSVYGRLIDLCQPTQKKYQIAVTKVLGKNMDAIIVDSEKTGRDCIQYIKEQRGEPETFLPLDYLEVKPTDEKLRELKGAKLVIDVIRYEPPHIKKALQYACGNALVCDNVEDARRIAFGGHQRHKTVALDGTLFQKSGVISGGASDLKAKARRWDEKAVDKLKEKKERLTEELKEQMKAKRKEAELRQVQSQAHGLQMRLKYSQSDLEQTKTRHLALNLQEKSKLESELANFGPRINDIKRIIQSRERERMKDLKEKMNQVEDEVFEEFCREIGVRNIREFEEEKVKRQNEIAKKRLEFENQKTRLGIQLDFEKNQLKEDQDKVHMWEQTVKKDENEIEKLKKEEQRHMKIIDETMAQLQDLKNQHLAKKSEVNDKNHEMEEIRKKLGGANKEMTHLQKEVTAIETKLEQKRSDRHNLLQACKMQDIKLPLSKGTMDDISQEEGSSQGEDSVSGSQRISSIYAREALIEIDYGDLCEDLKDAQAEEEEIKQEMNTLQQKLNEQQSVLQRIAAPNMKAMEKLESVRDKFQETSDEFEAARKRAKKAKQAFEQIKKERFDRFNACFESVATNIDEIYKALSRNSSAQAFLGPENPEEPYLDGINYNCVAPGKRFRPMDNLSGGEKTVAALALLFAIHSYKPAPFFVLDQIDAALDNTNIGKVANYIKEQSTCNFQAIVISLKEEFYTKAESLIGVYPEQGDCVISKVLTFDLTKYPDANPNPNEQ

Smc3 head complete (including 64G linker)

MGFLKLIEIENFKSYKGRQIIGPFQRFTAIIGPNGSGKSNLMDAISFVLGEKTSNLRVKTLRDLIHGAPVGKPAANRAFVSMVYSEEGAEDRTFARVIVGGSSEYKINNKVVQLHEYSEELEKLGILIKARNFLVFQGAVESIAMKNPKERTALFEEISRSGELAQEYDKRKKEMVKAEEDTQFNYHRKKNIAAERKEAKQGGG

NIPBL

MMDSSTFKRFTASIENILDNLEDMDFTAFGDDDEIPQELLLGKHQLNELGSESAKIKAMGIMDKLSTDKTVKVLNILEKNIQDGSKLSTLLNHNNDTEEEERLWRDLIMERVTKSADACL
TTINIMTSPNMPKAVYIEDVIERVIQYTKFHLQNTLYPQYDPVYRLDPHGGGLLSSKAKRAKCSTHKQRVIVMLYNKVCDIVSSLSELLEIQLLTDTTILQVSSMGITPFFVENVSELQLCAIK
LVTAVFSRYEKHRQLILEEIFTSLARLPTSKRSLRNFRLNSSDMDGEPMYIQMVTALVLQLIQCVVHLPSSEKDSNAEEDSNKKIDQDVVITNSYETAMRTAQNFLSIFLKKCGSKQGEEDY
RPLFENFVQDLLSTVNKPEWPAAELLLSLLGRLLVHQFSNKSTEMALRVASLDYLGTVAARLRKDAVTSKMDQGSIERILKQVSGGEDEIQQLQKALLDYLDENTETDPSLVFSRKFYIAQ
WFRDTTLTETKAMKSQKDEESSEGTHHAKEIETTGQIMHRAENRKKFLRSIIKTTPSQFSTLKMNSDTVDYDDACLIVRYLASMRPFAQSFDIYLTQILRVLGENAIAVRTKAMKCLSEVV
AVDPSILARLDMQRGVHGRLMDNSTSVREAAVELLGRFVLCRPQLAEQYYDMLIERILDTGISVRKRVIKILRDICIEQPTFPKITEMCVKMIRRVNDEEGIKKLVNETFQKLWFTPTPHND
KEAMTRKILNITDVVAACRDTGYDWFEQLLQNLLKSEEDSSYKPVKKACTQLVDNLVEHILKYEESLADSDNKGVNSGRLVACITTLFLFSKIRPQLMVKHAMTMQPYLTTKCSTQNDF
MVICNVAKILELVVPLMEHPSETFLATIEEDLMKLIIKYGMTVVQHCVSCLGAVVNKVTQNFKFVWACFNRYYGAISKLKSQHQEDPNNTSLLTNKPALLRSLFTVGALCRHFDFDLEDF
KGNSKVNIKDKVLELLMYFTKHSDEEVQTKAIIGLGFAFIQHPSLMFEQEVKNLYNNILSDKNSSVNLKIQVLKNLQTYLQEEDTRMQQADRDWKKVAKQEDLKEMGDVSSGMSSSIMQ
LYLKQVLEAFFHTQSSVRHFALNVIALTLNQGLIHPVQCVPYLIAMGTDPEPAMRNKADQQLVEIDKKYAGFIHMKAVAGMKMSYQVQQAINTCLKDPVRGFRQDESSSALCSHLYSM
IRGNRQHRRAFLISLLNLFDDTAKTDVTMLLYIADNLACFPYQTQEEPLFIMHHHIDITLSVSGSNLLQSFKESMVKDKRKERKSSPSKENESSDSEEEVSRPRKSRKRVDSDSDSDSDDDINSV
MKCLPENSAPLIEFANVSQGILLLLMLKQHLKNLCGFSDSKIQKYSPSESAKVYDKAINRKTGVHFHPKQTLDFLRSDMANSKITEEVKRSIVKQYLDFKLLMEHL

STAG1

TLFEVVKLGKSAMQSVVDDWIESYKQDRDIALLDLINFFIQCSGCRGTVRIEMFRNMQNAEIIRKMTEEFDEDSGDYPLTMPGPQWKKFRSNFCEFIGVLIRQCQYSIIYDEYMMDTVISLL
TGLSDSQVRAFRHTSTLAAMKLMTALVNVALNLSIHQDNTQRQYEAERNKMIGKRANERLELLLQKRKELQENQDEIENMMNSIFKGIFVHRYRDAIAEIRAICIEEIGVWMKMYSDAF
LNDSYLKYVGWTLHDRQGEVRLKCLKALQSLYTNRELFPKLELFTNRFKDRIVSMTLDKEYDVAVEAIRLVTLILHGSEEALSNEDCENVYHLVYSAHRPVAVAAGEFLHKKLFSRHDPQ
AEEALAKRRGRNSPNGNLIRMLVLFFLESELHEHAAYLVDSLWESSQELLKDWECMTELLLEEPVQGEEAMSDRQESALIELMVCTIRQAAEAHPPVGRGTGKRVLTAKERKTQIDDRN
KLTEHFIITLPMLLSKYSADAEKVANLLQIPQYFDLEIYSTGRMEKHLDALLKQIKFVVEKHVESDVLEACSKTYSILCSEEYTIQNRVDIARSQLIDEFVDRFNHSVEDLLQEGEEADDDDIY
NVLSTLKRLTSFHNAHDLTKWDLFGNCYRLLKTGIEHGAMPEQIVVQALQCSHYSILWQLVKITDGSPSKEDLLVLRKTVKSFLAVCQQCLSNVNTPVKEQAFMLLCDLLMIFSHQLMT
GGREGLQPLVFNPDTGLQSELLSFVMDHVFIDQDEENQSMEGDEEDEANKIEALHKRRNLLAAFSKLIIYDIVDMHAAADIFKHYMKYYNDYGDIIKETLSKTRQIDKIQCAKTLILSLQQ
LFNELVQEQGPNLDRTSAHVSGIKELARRFALTFGLDQIKTREAVATLHKDGIEFAFKYQNQKGQEYPPPNLAFLEVLSEFSSKLLRQDKKTVHSYLEKFLTEQMMERREDVWLPLISYRNN
SL

Smc3head:kleisin:SA1

MGFLKLIEIENFKSYKGRQIIGPFQRFTAIIGPNGSGKSNLMDAISFVLGEKTSNLRVKTLRDLIHGAPVGKPAANRAFVSMVYSEEGAEDRTFARVIVGGSSEYKINNKVVQLHEYSEELEK
LGILIKARNFLVFQGAVESIAMKNPKERTALFEEISRSGELAQEYDKRKKEMVKAEEDTQFNYHRKKNIAAERKEAKQGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG
GGGGGGGGGGGGGGGGGGGGGGGGGGLESVRDKFQETSDEFEAARKRAKKAKQAFEQIKKERFDRFNACFESVATNIDEIYKALSRNSSAQAFLGPENPEEPYLDGINYNCVAPGK
RFRPMDNLSGGEKTVAALALLFAIHSYKPAPFFVLDQIDAALDNTNIGKVANYIKEQSTCNFQAIVISLKEEFYTKAESLIGVYPEQGDCVISKVLTFDLTKYPDANPNPNEQ:MFYAHFVL
SKRGPLAKIWLAAHWDKKLTKAHVFECNLESSVESIISPKVKMALRTSGHLLLGVVRIYHRKAKYLLADCNEAFIKIKMAFRPGVVDLPEENREAAYNAITLPEEFHDFDQPLPDLDDIDV
AQQFSLNQSRVEEITMREEVGNISILQENDFGDFGMDDREIMREGSAFEDDDMLVSTTTSNLLLESEQSTSNLNEKINHLEYEDQYKDDNFGEGNDGGILDDKLISNNDGGIFDDPPALSE
AGVMLPEQPAHDDMDEDDNVSMGGPDSPDSVDPVEPMPTMTDQTTLVPNEEEAFALEPIDITVKETKAKRKRKLIVDSVKELDSKTIRAQLSDYSDIVTTLDLAPPTKKLMMWKETG
GVEKLFSLPAQPLWNNRLLKLFTRCLTPLVPEDLRKRRKGGEADNLDEFLKEFENPEVPREDQQQQHQQRDVIDEPIIEEPSRLQESVMEASRTNIDESAMPPPPPQGVKRKAGQIDPEPV
MPPQQVEQMEIPPVELPPEEPPNICQLIPELELLPEKEKEKEKEKEDDEEEEDEDASGGDQDQEERRWNKRTQQMLHGLQRALAKTGAESISLLELCRNTNRKQAAAKFYSFLVLKKQQ
AIELTQEEPYSDIIATPGPRFHII:TLFEVVKLGKSAMQSVVDDWIESYKQDRDIALLDLINFFIQCSGCRGTVRIEMFRNMQNAEIIRKMTEEFDEDSGDYPLTMPGPQWKKFRSNFCEFIGV
LIRQCQYSIIYDEYMMDTVISLLTGLSDSQVRAFRHTSTLAAMKLMTALVNVALNLSIHQDNTQRQYEAERNKMIGKRANERLELLLQKRKELQENQDEIENMMNSIFKGIFVHRYRDAI
AEIRAICIEEIGVWMKMYSDAFLNDSYLKYVGWTLHDRQGEVRLKCLKALQSLYTNRELFPKLELFTNRFKDRIVSMTLDKEYDVAVEAIRLVTLILHGSEEALSNEDCENVYHLVYSAHR
PVAVAAGEFLHKKLFSRHDPQAEEALAKRRGRNSPNGNLIRMLVLFFLESELHEHAAYLVDSLWESSQELLKDWECMTELLLEEPVQGEEAMSDRQESALIELMVCTIRQAAEAHPPVG
RGTGKRVLTAKERKTQIDDRNKLTEHFIITLPMLLSKYSADAEKVANLLQIPQYFDLEIYSTGRMEKHLDALLKQIKFVVEKHVESDVLEACSKTYSILCSEEYTIQNRVDIARSQLIDEFVD
RFNHSVEDLLQEGEEADDDDIYNVLSTLKRLTSFHNAHDLTKWDLFGNCYRLLKTGIEHGAMPEQIVVQALQCSHYSILWQLVKITDGSPSKEDLLVLRKTVKSFLAVCQQCLSNVNTP
VKEQAFMLLCDLLMIFSHQLMTGGREGLQPLVFNPDTGLQSELLSFVMDHVFIDQDEENQSMEGDEEDEANKIEALHKRRNLLAAFSKLIIYDIVDMHAAADIFKHYMKYYNDYGDIIK
ETLSKTRQIDKIQCAKTLILSLQQLFNELVQEQGPNLDRTSAHVSGIKELARRFALTFGLDQIKTREAVATLHKDGIEFAFKYQNQKGQEYPPPNLAFLEVLSEFSSKLLRQDKKTVHSYLEK
FLTEQMMERREDVWLPLISYRNSL

Smc1head:kleisin:SA1

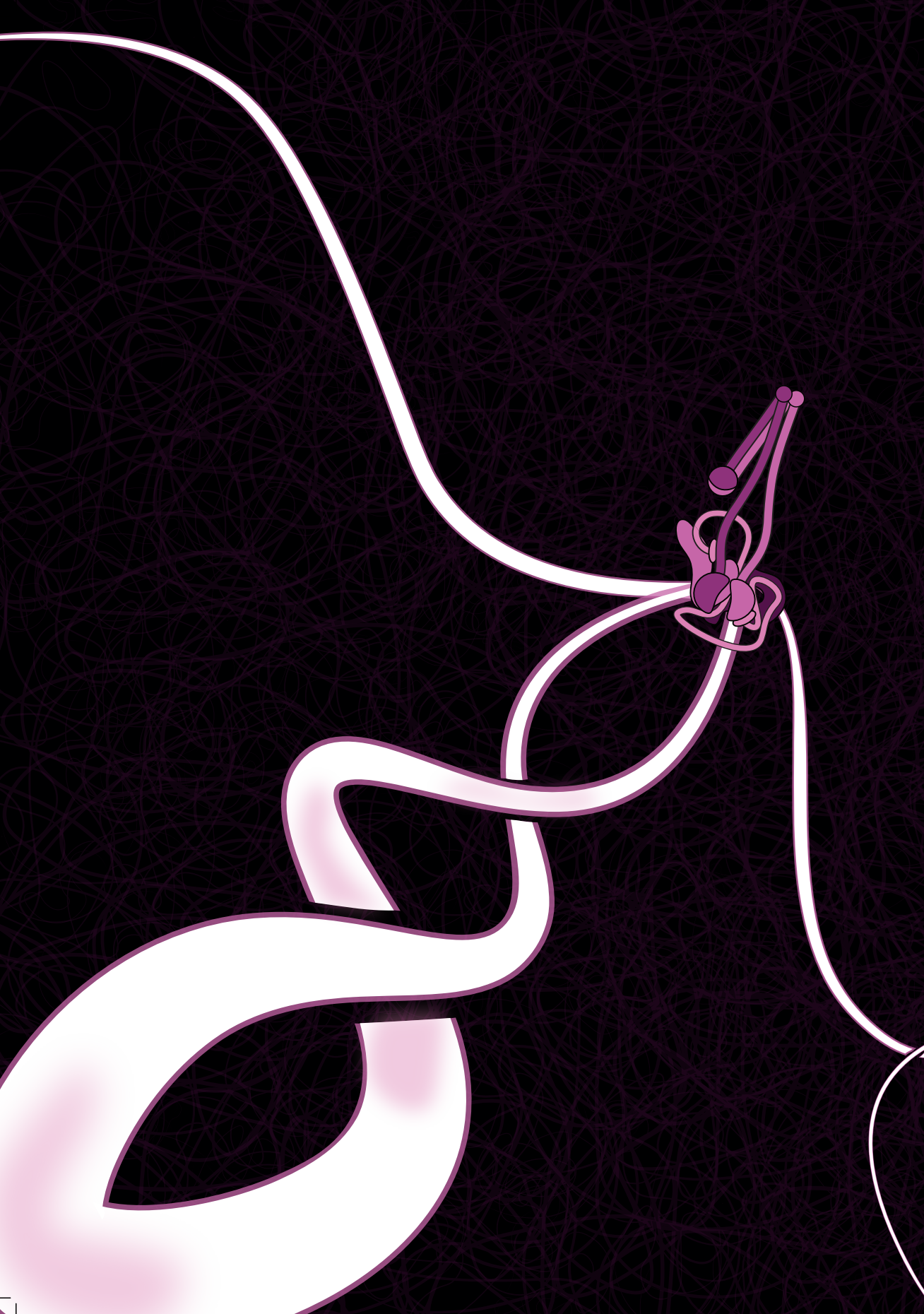
MYIKQVIIQGFRSYRDQTIVDPFSSKHNVIVGRNGSGKSNFFYAIQFVLSDEFSHLRPEQRLALLHEGTGPRVISAFVEIIFDNSDNRLPIDKEEVSLRRVIGAKKDQYFLDKKMVTKNDVMN
LLESAGFSRSNPYYIVKQGKINQMATAPDSQRLKLLREVAGTRVYDERKEESISLMKETEGKREKINELLKYIEERLHTLEEEKEELAQYQKWDKMRRALEYTIYNQELNETRAKLDEGGG
GGGGGGGGGGGGGGGGGGGG

Smc3head:kleisin:NIPBL:SA1

MGFLKLIEIENFKSYKGRQIIGPFQRFTAIIGPNGSGKSNLMDAISFVLGEKTSNLRVKTLRDLIHGAPVGKPAANRAFVSMVYSEEGAEDRTFARVIVGGSSEYKINNKVVQLHEYSEELEK
LGILIKARNFLVFQGAVESIAMKNPKERTALFEEISRSGELAQEYDKRKKEMVKAEEDTQFNYHRKKNIAAERKEAKQGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG
GGGGGGGGGGGGGGGGGGGGGGGGGGLESVRDKFQETSDEFEAARKRAKKAKQAFEQIKKERFDRFNACFESVATNIDEIYKALSRNSSAQAFLGPENPEEPYLDGINYNCVAPGK
RFRPMDNLSGGEKTVAALALLFAIHSYKPAPFFVLDQIDAALDNTNIGKVANYIKEQSTCNFQAIVISLKEEFYTKAESLIGVYPEQGDCVISKVLTFDLTKYPDANPNPNEQ:MFYAHFVL
SKRGPLAKIWLAAHWDKKLTKAHVFECNLESSVESIISPKVKMALRTSGHLLLGVVRIYHRKAKYLLADCNEAFIKIKMAFRPGVVDLPEENREAAYNAITLPEEFHDFDQPLPDLDDIDV
AQQFSLNQSRVEEITMREEVGNISILQENDFGDFGMDDREIMREGSAFEDDDMLVSTTTSNLLLESEQSTSNLNEKINHLEYEDQYKDDNFGEGNDGGILDDKLISNNDGGIFDDPPALSE
AGVMLPEQPAHDDMDEDDNVSMGGPDSPDSVDPVEPMPTMTDQTTLVPNEEEAFALEPIDITVKETKAKRKRKLIVDSVKELDSKTIRAQLSDYSDIVTTLDLAPPTKKLMMWKETG
GVEKLFSLPAQPLWNNRLLKLFTRCLTPLVPEDLRKRRKGGEADNLDEFLKEFENPEVPREDQQQQHQQRDVIDEPIIEEPSRLQESVMEASRTNIDESAMPPPPPQGVKRKAGQIDPEPV
MPPQQVEQMEIPPVELPPEEPPNICQLIPELELLPEKEKEKEKEKEDDEEEEDEDASGGDQDQEERRWNKRTQQMLHGLQRALAKTGAESISLLELCRNTNRKQAAAKFYSFLVLKKQQ
AIELTQEEPYSDIIATPGPRFHII:MMDSSTFKRFTASIENILDNLEDMDFTAFGDDDEIPQELLLGKHQLNELGSESAKIKAMGIMDKLSTDKTVKLNILEKNIQDGSKLSTLLNHNNDTEE
EERLWRDLIMERVTKSADACLTTINIMTSPNMPKAVYIEDVIERVIQYTKFHLQNTLYPQYDPVYRLDPHGGGLLSSKAKRAKCSTHKQRVIVMLYNKVCDIVSSLSELLEIQLLTDTTILQ
VSSMGITPFFVENVSELQLCAIKLVTAVFSRYEKHRQLILEEIFTSLARLPTSKRSLRNFRLNSSDMDGEPMYIQMVTALVLQLIQCVVHLPSSEKDSNAEEDSNKKIDQDVVITNSYETAMR
TAQNFLSIFLKKCGSKQGEEDYRPLFENFVQDLLSTVNKPEWPAAELLLSLLGRLLVHQFSNKSTEMALRVASLDYLGTVAARLRKDAVTSKMDQGSIERILKQVSGGEDEIQQLQKALL
DYLDENTETDPSLVFSRKFYIAQWFRDTTLETEKAMKSQKDEESSEGTHHAKEIETTGQIMHRAENRKKKFLRSIIKTTPSQFSTLKMNSDTVDYDDACLIVRYLASMRPFAQSFDIYLTQIL
RVLGENAIAVRTKAMKCLSEVVAVDPSILARLDMQRGVHGRLMDNSTSVREAAVELLGRFVLCRPQLAEQYYDMLIERILDTGISVRKRVIKILRDICIEQPTFPKITEMCVKMIRRVVNDEE
GIKKLVNETFQKLWFTPTPHNDKEAMTRKILNITDVVAACRDTGYDWFEQLLQNLLKSEEDSSYKPVKKACTQLVDNLVEHILKYEESLADSDNKGVNSGRLVACITTLFLFSKIRPQLM
VKHAMTMQPYLTTKCSTQNDFMVICNVAKILELVVPLMEHPSETFLATIEEDLMKLIIKYGMTVVQHCVSCLGAVVNKVTQNFKFVWACFNRYYGAISKLKSQHQEDPNNTSLLTNKP
ALLRSLFTVGALCRHFDFDLEDFKGNSKVNIKDKVLELLMYFTKHSDEEVQTKAIIGLGFAFIQHPSLMFEQEVKNLYNNILSDKNSSVNLKIQVLKNLQTYLQEEDTRMQQADRDWKKV
AKQEDLKEMGDVSSGMSSSIMQLYLKQVLEAFFHTQSSVRHFALNVIALTLNQGLIHPVQCVPYLIAMGTDPEPAMRNKADQQLVEIDKKYAGFIHMKAVAGMKMSYQVQQAINTCL
KDPVRGFRQDESSSALCSHLYSMIRGNRQHRRAFLISLLNLFDDTAKTDVTMLLYIADNLACFPYQTQEEPLFIMHHIDITLSVSGSNLLQSFKESMVKDKRKERKSSPSKENESSDSEEEVS
RPRKSRKRVDSDSDSDSEDDINSVMKCLPENSAPLIEFANVSQGILLLLMLKQHLKNLCGFSDSKIQKYSPSESAKVYDKAINRKTGVHFHPKQTLDFLRSDMANSKITEEVKRSIVKQYLD
FKLLMEHL:TLFEVVKLGKSAMQSVVDDWIESYKQDRDIALLDLINFFIQCSGCRGTVRIEMFRNMQNAEIIRKMTEEFDEDSGDYPLTMPGPQWKKFRSNFCEFIGVLIRQCQYSIIYDEY
MMDTVISLLTGLSDSQVRAFRHTSTLAAMKLMTALVNVALNLSIHQDNTQRQYEAERNKMIGKRANERLELLLQKRKELQENQDEIENMMNSIFKGIFVHRYRDAIAEIRAICIEEIGVW
MKMYSDAFLNDSYLKYVGWTLHDRQGEVRLKCLKALQSLYTNRELFPKLELFTNRFKDRIVSMTLDKEYDVAVEAIRLVTLILHGSEEALSNEDCENVYHLVYSAHRPVAVAAGEFLHK
KLFSRHDPQAEEALAKRRGRNSPNGNLIRMLVLFFLESELHEHAAYLVDSLWESSQELLKDWECMTELLLEEPVQGEEAMSDRQESALIELMVCTIRQAAEAHPPVGRGTGKRVLTAKE
RKTQIDDRNKLTEHFIITLPMLLSKYSADAEKVANLLQIPQYFDLEIYSTGRMEKHLDALLKQIKFVVEKHVESDVLEACSKTYSILCSEEYTIQNRVDIARSQLIDEFVDRFNHSVEDLLQEG
EEADDDDIYNVLSTLKRLTSFHNAHDLTKWDLFGNCYRLLKTGIEHGAMPEQIVVQALQCSHYSILWQLVKITDGSPSKEDLLVLRKTVKSFLAVCQQCLSNVNTPVKEQAFMLLCDLL
MIFSHQLMTGGREGLQPLVFNPDTGLQSELLSFVMDHVFIDQDEENQSMEGDEEDEANKIEALHKRRNLLAAFSKLIIYDIVDMHAAADIFKHYMKYYNDYGDIIKETL
SKTRQIDKIQCAKTLILSLQQLFNELVQEQGPNLDRTSAHVSGIKELARRFALTFGLDQIKTREAVATLHKDGIEFAFKYQNQKGQEYPPPNLAFLEVLSEFSSKLLRQDKKTVHSYLEKFLT
EQMMERREDVWLPLISYRNSL

Smc1head:kleisin:NIPBL:SA1

MYIKQVIIQGFRSYRDQTIVDPFSSKHNVIVGRNGSGKSNFFYAIQFVLSDEFSHLRPEQRLALLHEGTGPRVISAFVEIIFDNSDNRLPIDKEEVSLRRVIGAKKDQYFLDKKMVTKNDVMN
LLESAGFSRSNPYYIVKQGKINQMATAPDSQRLKLLREVAGTRVYDERKEEISLMKETEGKREKINELLKYIEERLHTLEEEKEELAQYQKWDKMRRALEYTIYNQELNETRAKLDEGGG
GGGGGGGGGGGGGGGGGGGG



6

All eukaryotic SMC proteins induce a twist of -0.6 at each DNA-loop-extrusion step

This chapter is available as two preprints:
Richard Janissen*, Roman Barth*, Iain F. Davidson, Michael Taschner, Stephan Gruber, Jan-Michael Peters, Cees Dekker: All eukaryotic SMC proteins induce a twist of -0.6 at each DNA-loop-extrusion step, bioRxiv, March 2024, [50]

Iain F. Davidson, Roman Barth, Sabrina Horn, Richard Janissen, Benedikt Bauer, Cees Dekker, Jan-Michael Peters: Cohesin supercoils DNA during loop extrusion, bioRxiv, March 2024, [51]

Abstract

Eukaryotes carry three types of Structural Maintenance of Chromosomes (SMC) protein complexes, condensin, cohesin, and SMC5/6, which are ATP-dependent motor proteins that remodel the genome via DNA loop extrusion. SMCs modulate DNA supercoiling, but it has remained incompletely understood how this is achieved. Here we present a single-molecule magnetic tweezers assay that directly measures how much twist is induced by an individual SMC in each loop-extrusion step. We demonstrate that all three SMC complexes induce the same large negative twist (i.e., a linking number change ΔL_k of -0.6 at each loop-extrusion step) into the extruded loop. This -0.6 twist is independent of step size. Using ATP-hydrolysis mutants and non-hydrolysable ATP analogues, we find that ATP binding is the twist-inducing event during the ATPase cycle, which coincides with the force-generating loop-extrusion step. The fact that all three eukaryotic SMC proteins induce the same amount of twist indicates a common DNA-loop-extrusion mechanism among these SMC complexes. Furthermore, we conducted bulk supercoiling assays using torsionally constrained plasmids and performed accompanying in silico analyses. This data suggest that DNA extrusion stops once cohesin reaches its stall torque during supercoiling.

6.1 Introduction

SMC (Structural Maintenance of Chromosomes) protein complexes are essential in all organisms to ensure proper chromosome organization. In eukaryotes, three SMC protein complexes exist: cohesin, condensin, and SMC5/6 (ref. [331]). These SMC complexes share a common architecture (Figure S6.1a) where two Smc subunits heterodimerize at their hinge and an intrinsically disordered kleisin subunit connects their ATPase domains, thus forming a ring structure. Cohesin and condensin associate with two HAWK (HEAT repeat-containing proteins Associated With Kleisins) subunits on the kleisin, while SMC5/6 carries two KITE (kleisin interacting winged-helix tandem element) on its kleisin and a third, Nse2, on the coiled coil of Smc5 (ref. [1]). The three SMCs participate in a large range of diverse biological processes *in vivo* [1]. From S-phase until metaphase, cohesin topologically embraces the two replicated sister chromatids [132] while condensin compacts the chromatids *via* DNA loop extrusion in mitosis [98, 116] for subsequent chromosome segregation. During interphase, cohesin also folds genomic DNA into loops [173, 252, 282], which has been implicated in the regulation of transcription [181], recombination [203, 305], and the local separation of sister chromatids [332]. SMC5/6 predominantly contributes to DNA repair [90], but its exact role remains elusive. Despite their varying roles, all three eukaryotic SMC complexes are able to extrude DNA into loops [98, 173, 204, 212, 214, 216] *via* an ATP-dependent DNA loop extrusion mechanism that remains incompletely understood [247].

SMCs have been reported to change the amount of DNA supercoiling. Early biochemical reconstitutions of condensin from *Xenopus* egg extracts demonstrated that condensin generates positive DNA supercoiling [48, 206, 333], i.e., an additional twist deposited in DNA in the same right-handed direction as the double helix. Supercoiling is ubiquitously present in cells and contributes to the packaging of chromosomes and the regulation of genetic processes. Recent data have shown that condensin [323] and SMC5/6 [325] complexes are able to loop already supercoiled DNA and that cohesin also actively induce DNA supercoiling using ATP (refs. [48, 49, 51, 323]). Condensin DC has been associated with local negative supercoils at its high-occupancy binding sites of the X chromosome in C. elegans [334].

Likely, DNA supercoiling and DNA loop extrusion stem from the same conformational changes during the ATPase cycle, suggesting that also SMC5/6 should exhibit a DNA supercoiling activity as well. Notably, DNA loop extrusion by condensin [217] and cohesin [35] occurs in large discrete steps of up to hundred base pairs each, which are generated by ATP binding to the SMC complex [217]. It remains unknown how SMC proteins induce DNA twist, whether or how this mechanism is related to the loop-extrusion mechanism, and if the DNA twist directionality and degree of twist are comparable across the SMC family.

Here, we develop a high-resolution magnetic tweezers assay to quantitatively examine, at the single-molecule level, the twist induced into DNA during loop extrusion by all three eukaryotic SMC complexes: human cohesin, yeast condensin, and yeast SMC5/6. We find that all three SMC complexes twist DNA negatively in each DNA loop-extrusion step, irrespective of the loop-extrusion step size. Notably, a large linking-number change

($\Delta L_k = -0.6$) was resolved in each step of DNA loop extrusion by individual SMCs on single DNA molecules. Using ATP-hydrolysis mutants and non-hydrolysable ATP analogues, we show that ATP binding is not only the DNA extrusion step but also the twist-generating event of the ATPase cycle. Strikingly, all three SMC complexes quantitatively induced the same amount of negative twist in the extruded DNA loop, indicating a conserved mechanism shared by all eukaryotic SMCs to generate supercoils that is an intrinsic property of the DNA loop-extrusion cycle.

6.2 Results

6.2.1 Assessing DNA twist that is induced at individual steps in DNA loop extrusion

With the objective to probe the degree of DNA twist induced by SMCs during single DNA loop extrusion steps, we turned to magnetic tweezers as this single-molecule technique has been proven to be able to resolve both single loop-extrusion steps of yeast condensin and human cohesin [35, 217], as well as minute changes in DNA twist [335–338]. Torsionally constrained 3.6 kbp-long dsDNA molecules were tethered between a superparamagnetic 1 μm bead and an anti-digoxygenin-functionalized glass surface *via* handles on the end of DNA molecules containing either multiple biotin or digoxygenin labels, respectively (Methods section 6.4). A pair of permanent magnets was mounted above the flow cell which allowed to exert a calibrated force on the surface-tethered DNA molecules [339] (Figure 6.1a). Here, we applied a force of 0.3 pN, at which individual SMC-mediated loop extrusion steps can be measured [35, 217]. Rotation of the magnet pair within the plane of the surface yielded a corresponding rotation of the magnetized bead. Since the attached DNA molecule is torsionally constrained, its linking number L_k changes proportionally with the rotation of the bead. The DNA extension is at its maximum when no rotation is applied, i.e., when $\Delta L_k = 0$. Applying a low number of positive (negative) rotations overwinds (underwinds) and twists the DNA. Beyond a buckling point [340], twist is converted into writhe which absorbs the applied turns into plectonemic supercoils, resulting in the shortening of the DNA end-to-end extension [335, 337, 338, 341]. Monitoring the DNA end-to-end extension versus positive and negative rotations thus generates a rotation curve which is symmetric around zero rotations ($\Delta L_k = 0$) at low forces ($\leq 0.6pN$, i.e., where DNA remains in its B-form [335, 337, 338, 341]) (Figure 6.1a). Rotation curves were measured for all DNA molecules in the field of view as a reference before the introduction of SMCs into the flow chamber.

This magnetic tweezers assay allows to resolve of both the steps and the twist induced by individual SMCs during their loop extrusion activity. In a first series of experiments, human cohesin and ATP were flushed into the flow chamber while a high force of 7 pN was applied to the tethers to prevent compaction of the DNA molecules during flush-in by SMC-mediated DNA loop extrusion. When the flow was stopped and the force was quickly lowered to 0.3 pN, a stepwise decrease in the DNA end-to-end extension was observed (Figure 6.1b). These can be attributed to discrete DNA loop extrusion steps as previously demonstrated for yeast condensin and human cohesin [35, 217]. Directly after monitoring the steps for 10 minutes, another rotation curve was acquired at the same force (Figure 6.1c). Interestingly, the peak of the rotation curve after loop extrusion (Figure 6.1c, red curve)

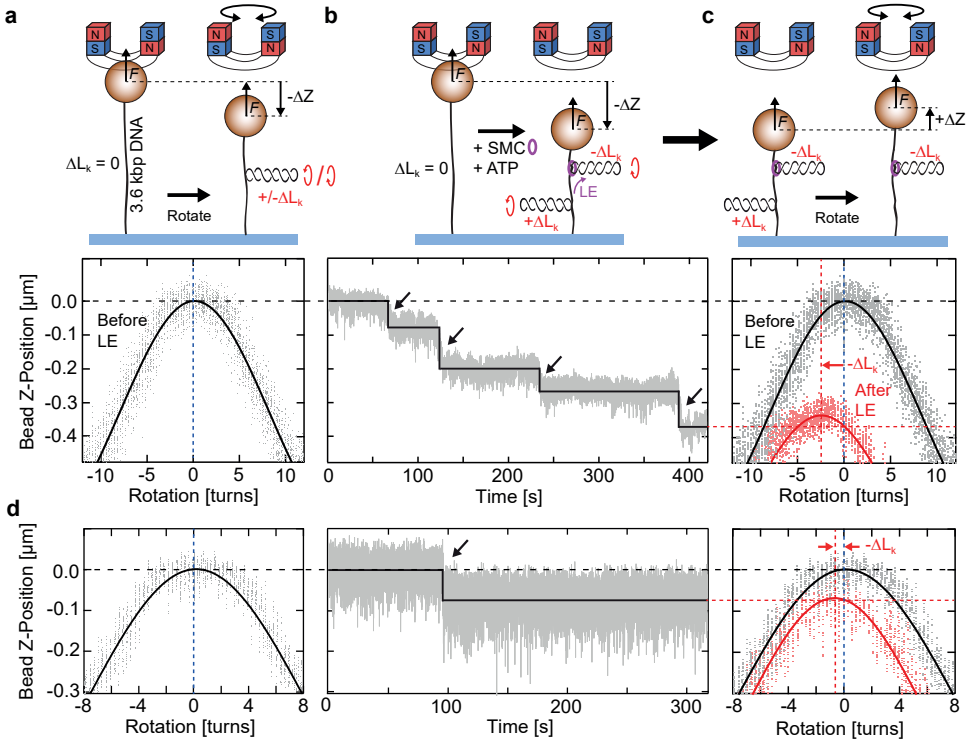


Figure 6.1 | Measuring twist induced by single DNA-loop-extrusion steps by use of magnetic tweezers. (**a**) Assay schematic (top) and corresponding experimental data (bottom) of the DNA end-to-end extension as a function of magnet rotation for a torsionally constrained 3.6 kbp DNA molecule. At a constant force of 0.3 pN, the DNA extension is maximal when no external rotations are applied. Upon applying positive/negative rotations, the over/underwinding of the DNA changes the linking number Lk and leads to the formation of plectonemic supercoils, which, at this low force, are symmetric for both positive and negative coiling. Solid line is a Gaussian fit to the rotation curve data. (**b**) Assay schematic (top) and experimental data (bottom) which depict a representative trajectory of human cohesin (100 pM human cohesin, 250 pM NIPBL-Mau2) showing the step-wise DNA loop extrusion in the presence of 1 mM ATP at 0.3 pN. Black line shows the fit result from a step-finding algorithm. During loop extrusion, human cohesin induces negative supercoils ($-\Delta L_k$) into the extruded DNA loop, while compensatory, positive supercoils ($+\Delta L_k$) form in the DNA molecule outside the loop. (**c**) Assay schematic (top) and experimental data (bottom) of DNA extension as a function of magnet rotation, similar to (a), conducted directly after the DNA loop extrusion experiment at 0.3 pN. The rotation curve after loop extrusion (red; line depicts Gaussian fit to the data) shows that the maximum DNA extension was shifted to negative magnet rotation compared to the initial rotation curve (black), which was caused by the uncoiling of the positive complementary supercoils ($+\Delta L_k$) that were formed in the DNA molecule outside the DNA loop. The degree of negative supercoils ($-\Delta L_k$) that are generated by human cohesin residing in the loop is thus directly measured from the shift of the rotation-curve maxima. Black arrows in the trajectories depict single steps. (**d**) Same as (a-c), for a trace showing only a single LE step (20 pM human cohesin with 50 pM NIPBL-Mau2 and 1 mM ATP). A lower amount of induced negative supercoils ($-\Delta L_k$) is observed, compared to (a-c).

was systematically shifted to a lower bead height and to a negative number of rotations, compared to the reference rotation curve acquired before the extrusion of DNA by SMCs (Figure 6.1a,c, black line). As the SMC reeled DNA into the loop, the DNA tether length decreased, explaining the decrease of the bead height [35, 217].

The shift of the rotation curve signals the twist that was inserted into the extruded loop by the SMC. To obtain the maximum DNA extension in the final rotation curve (Figure 6.1c, red curve), negative rotations have to be externally applied to the DNA tether (outside the extruded loop) in order to remove the writhe. In other words, the DNA tether outside the loop was positively twisted after loop extrusion. Since the overall linking number of the entire torsionally constrained DNA molecule was conserved, the positive linking number change ($\Delta L_k > 0$) that was measured in the non-looped DNA tether compensates the negative linking number change ($\Delta L_k < 0$) that was generated by the SMC inside the extruded loop (Figure 6.1c). In this way, our assay thus directly measures, at the single-molecule level, the induced twist that is generated into the extruded loop through one or more steps of the SMC. The magnitude of the induced twist ΔL_k is quantified as the difference of the peak positions of Gaussian fits to the rotation curves before and after loop extrusion (Methods section 6.4), and it can be correlated to the number of downward steps that was observed in the time trace. A small but still negative shift between the rotation curves before and after the induction of loop extrusion could even be resolved for DNA molecules that showed only a single downward step (Figure 6.1d). This, to our knowledge, constitutes the first direct demonstration that single SMC molecules negatively twist DNA, concomitant with individual loop extrusion steps.

6.2.2 ATP-binding induces an equal amount of twist in each loop-extrusion step, irrespective of the step size

Using magnetic tweezers, we previously determined that ATP binding (i.e. not ATP hydrolysis) is the step-generating event during the DNA loop-extrusion cycle of yeast condensin [217]. To test whether this is also the case for cohesin, as well as to assess whether the ATP-binding step is involved in DNA twisting, we used AMP-PnP, a non-hydrolysable ATP analogue, as well as an ATP-hydrolysis (EQ/EQ) deficient mutant during cohesin-mediated loop extrusion to obtain traces with a single loop-extrusion step. Mutations in cohesin's Walker B motif (E1157Q/E1144Q; EQ/EQ) of SMC1 and SMC3 block ATP binding, head dimerization and ATP hydrolysis. The DNA end-to-end extension exhibited only a single downward step (e.g. as in Figure 6.1d) or repeated downward-reverse step combinations (Figure S6.2a) for human cohesin in these conditions. No subsequent consecutive downward steps were observed, confirming that ATP binding is the step-generating event for human cohesin. Downward-reverse step combinations represent single loop extrusion steps which were followed by a reversal, possibly upon dissociation of DNA from one of the binding sites before the reeled in DNA was merged with the loop. Reversals were found to not contribute to the induced DNA twist (i.e. yielding $\Delta L_k = 0$). The change in DNA linking number ΔL_k for traces with a single downward step (Figure 6.2a), was -0.58 ± 0.16 (mean ± SD, N = 37).

By contrast, traces of wild type human cohesin in the presence of ATP often exhibited several consecutive downward steps (Figure S6.2a). The change in linking number after cohesin-mediated loop extrusion was found to scale linearly with the number of loop extrusion steps (Figure 6.2a). Division of the final linking number change by the number of loop extrusion steps in the trace showed that each step induced on average a DNA twist of $\Delta L_k = -0.59 \pm 0.16$ (mean ± SD, N = 78; Figure 6.2b). Strikingly, the linking number

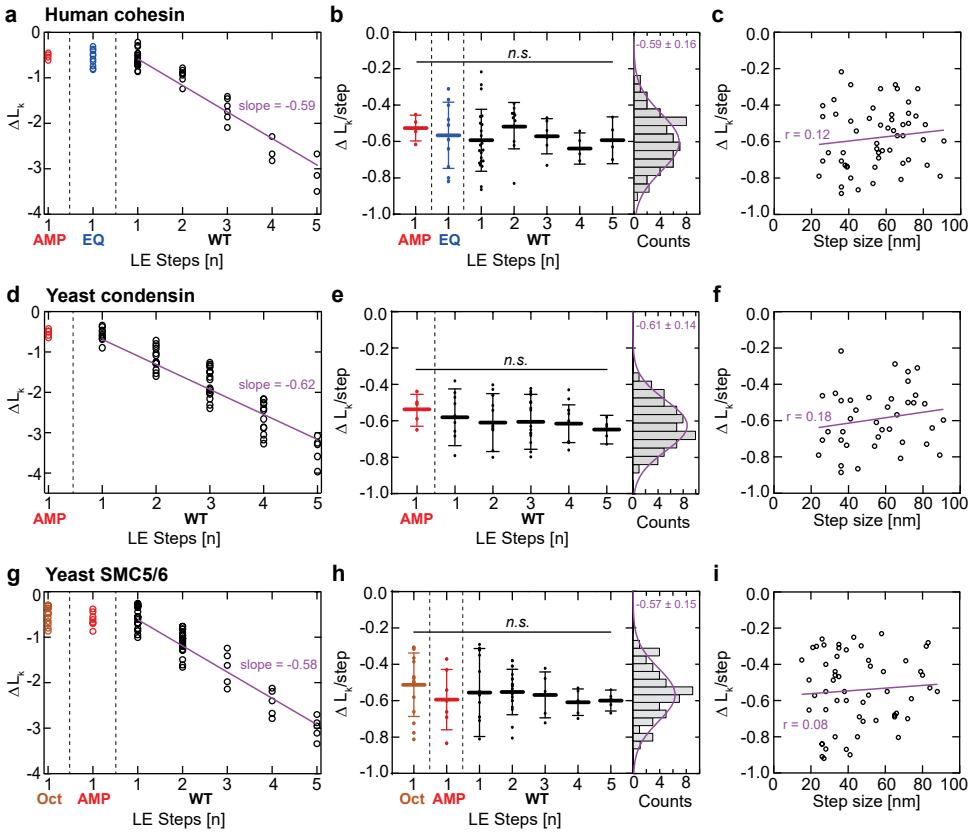


Figure 6.2 | Induced DNA twist is constant for each loop extrusion step and step size-independent for all eukaryotic SMC complexes.
(**a**) Linking number change ΔL_k versus the number of steps for wild type (WT) human cohesin in the presence of AMP-PnP (red, N = 5), for the EQ/EQ mutant of human cohesin in the presence of ATP (blue, N = 10), and for WT human cohesin in the presence of ATP (black, N = 63). Pink line is a linear fit without offset. (**b**) Same data as in panel (a), but displaying the linking number change per step (i.e. ΔL_k divided by the number of LE steps). The histogram on the right represents all data points (AMP-PnP, EQ, and WT) and is fitted by a Gaussian with -0.59 ± 0.16 ΔL_k/step (mean ± SD). Statistical significance was assessed using a one-way analysis of variance (ANOVA) with a significance level $\alpha = 0.05$ (95% confidence interval; n.s.: $p > 0.05$). (**c**) Linking number change versus the measured step size, for traces with only a single step (pooled from AMP-PnP, EQ, and WT experiments, N = 54). Linear fit and Pearson's correlation coefficient are shown in pink. (**d-f**) As (a-c) but for yeast condensin (N = 5 for AMP-PnP, N=64 for WT, N = 39 in the right panel). (**g-i**) As (a) but for the yeast SMC5/6 hexamer (N = 15 for the SMC5/6 octamer [SMC5/6 with Nse5-Nse6], N = 7 for AMP-PnP, N =56 for WT, N = 55 in the right panel).

change was independent of the size of the loop extrusion step (Figure 6.2c), as a similar -0.6 values were found for small (~20 nm) steps and for large (~100 nm) steps. This suggests that DNA twist is generated by local conformational changes within the SMC complex, independently of the amount of DNA that is fed into the loop.

6.2.3 All eukaryotic SMCs induce $\Delta L_k = -0.6$ per step into the extruded DNA loop

All three eukaryotic SMC complexes - cohesin, condensin, and SMC5/6 - share the ability to extrude DNA into loops and they do so with very similar characteristics, viz., at a comparable extrusion rate and up to a comparable sub-pN stalling force [35, 98, 173, 204, 212]. We therefore examined if their ability to twist DNA is also shared. To answer this question, we repeated magnetic tweezers experiments described above for human cohesin with budding yeast condensin (Figure 6.2d-f, Figure S6.1b, Figure S6.2b, Methods section 6.4) and budding yeast SMC5/6 (Figure 6.2g-i, Figure S6.1c, Figure S6.2c, Methods section 6.4, [222]). We then analysed whether individual loop extrusion steps can be observed using torsionally unconstrained DNA and measured the size of these steps (Methods section 6.4). At an applied force of 0.2 pN, we found that the median step size of SMC5/6 is ~30 nm, while cohesin and condensin take steps of median size ~40 nm (Figure S6.1d). The step size decreased as the DNA tension is increased, which is a phenomenon observed for all eukaryotic SMC complexes (Figure S6.1d and refs. [35, 217]).

Magnetic tweezers experiments on yeast condensin and SMC5/6 yielded very similar results as for human cohesin. In the presence of AMP-PnP, wild type yeast condensin and SMC5/6 showed only a single downward step and downward-reverse step combinations (Figure S6.2b,c), while in the presence of ATP, multiple consecutive downward steps were observed (Figure S6.2b,c). The SMC5/6 hexamer can be converted to an octameric version by the addition of the co-factor Nse5/6 (ref. [309, 342]). Octameric SMC5/6 does not loop DNA [204, 222] but promotes salt-stable loading of SMC5/6 onto DNA, which is indicative of topological loading [204, 342]. Interestingly, loop-extrusion traces of octameric SMC5/6 exhibited a single downward step or downward-reverse step combinations (but never multiple consecutive downward steps) (Figure S6.2c), similar to traces of wild type SMCs in the presence of a non-hydrolyzable AMP-PnP or an ATP-hydrolysis deficient mutant (Figure S6.2c and ref. [217], respectively). This suggests that Nse5/6 blocks ATP hydrolysis but not ATP binding to SMC5/6, consistent with an earlier report that ATP binding but not hydrolysis is required to topologically load SMC5/6 onto DNA [342].

As for cohesin, the linking number change induced by condensin and SMC5/6 hexamer was negative and scaled with the amount of loop extrusion steps, while the linking number change per step remained constant at a value of $\Delta L_k = -0.61 \pm 0.14$ and -0.57 ± 0.15, respectively (mean ± SD; Figure 6.2e,h). These values were also independent of the size of the loop extrusion steps (Figure 6.2f,i). Strikingly, all three SMC complexes thus induced a comparable amount of DNA twist per step of $\Delta L_k = -0.59 \pm 0.02$ (mean ± SD across the three SMCs, Figure 6.5a).

6.2.4 A plasmid assay to measure cohesin-mediated DNA supercoiling

Next to resolving the SMC-induced DNA twist during loop extrusion using magnetic tweezers, we used a plasmid supercoiling assay which was used previously to measure condensin-mediated supercoiling [48, 49]. In topologically constrained DNA molecules, any local SMC-mediated change in supercoiling, either *via* twist or writhe, is accompanied by a compensatory twist or writhe with the opposite sign elsewhere in the DNA molecule

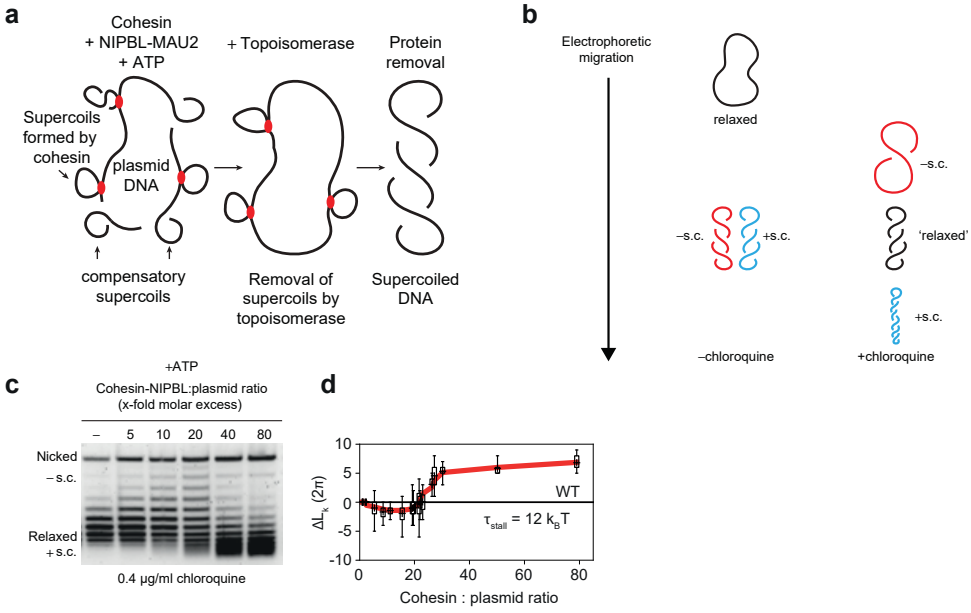


Figure 6.3 | Loop extrusion-induced torque simulations can recapitulate the transition from negative to positive supercoiling handedness at high cohesin:DNA ratios. **a** Illustration of the DNA supercoiling assay. Plasmid DNA is relaxed by incubating it with human topoisomerase I. Cohesin, NIPBL-Mau2 and ATP are then added in the continued presence of topoisomerase. Cohesin induces supercoils in the plasmid DNA, which is accompanied by compensatory supercoiling of the opposite handedness. Topoisomerase removes the compensatory supercoils but not those induced by cohesin, resulting in a change in the linking number of the plasmid DNA that persists following protein removal. Note that the activity of cohesin and topoisomerase are illustrated separately for clarity, although they occur concomitantly. **(b)** Chloroquine agarose gel electrophoresis allows discrimination between negatively and positively supercoiled plasmid DNA. **(c)** Negatively supercoiled plasmid DNA was relaxed by incubating with human topoisomerase I in the presence ATP for 45 minutes. Cohesin was added at the indicated molar ratio relative to plasmid DNA along with a two-fold excess of NIPBL-Mau2 relative to cohesin and incubations were continued for a further 90 minutes. Reactions were then terminated and purified DNA was separated by agarose gel electrophoresis in the presence of 0.4 µg/ml chloroquine. **(d)** Number and handedness of linking numbers induced by the combined action of SMCs and topoisomerases (ΔL_k) by human topoisomerase I at the indicated wild type cohesin:DNA ratios from simulations. Crosses denote the mean value, whiskers denote the span of the data. The red line is a guide to the eye and was constructed by applying a Savitzky-Golay filter of window length 5 and order 1. Data are from 50 independent simulations.

(Figure 6.3a). If either the actively induced or the compensatory twist or writhe is removed by a topoisomerase, this will result in a permanent change of the plasmid's linking number L_k. Following deproteinization of the plasmid, this change can be detected by an increase in the electrophoretic mobility of the plasmids. Condensin was reported to generate plasmids with a net positive linking number when incubated at high molar excess over plasmid DNA [48] but to generate net-negatively supercoiled plasmids when incubated at lower protein:plasmid ratios, i.e. under conditions that are more similar to those in which condensin extrudes DNA loops *in vitro* [49].

To investigate the handedness of the plasmids on which cohesin acted, we incubated cohesin and NIPBL-Mau2 with plasmid DNA at a range of different protein concentrations

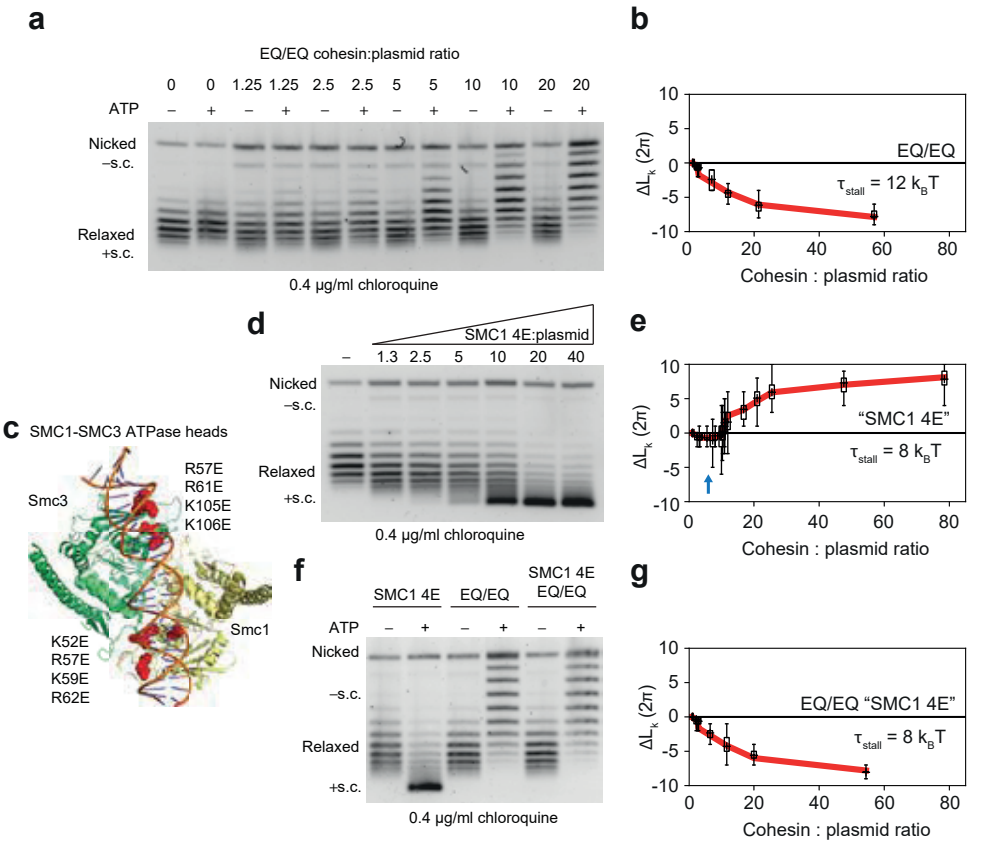


Figure 6.4 | Loop extrusion-induced torque simulations can recapitulate the transition from negative to positive supercoiling handedness at high cohesin:DNA ratios. **a** As Figure 6.3c except EQ/EQ cohesin was added at the indicated cohesin:plasmid ratio along with NIPBL-Mau2 at a two-fold NIPBL:cohesin ratio. Every second lane is a control in the absence of ATP. EQ/EQ cohesin generates negatively supercoiled products at every tested cohesin:plasmid ratio. **(b)** As Figure 6.3d but EQ/EQ cohesin was simulated by allowing only a single step per cohesin. **(c)** Structure of the Smc1-Smc3 ATPase heads (PDB: 6WG3). DNA binding site mutations are indicated as spheres [238]. **(d)** Negatively supercoiled plasmid DNA was relaxed by incubating with human topoisomerase I for 45 minutes. Smc1 4E cohesin was added at the indicated cohesin:DNA ratios of cohesin along with a two-fold excess of NIPBL-Mau2 relative to cohesin in the presence and absence of ATP and incubations were continued for a further 90 minutes. Reactions were then terminated and purified DNA was separated by agarose gel electrophoresis in the presence of 0.4 µg/ml chloroquine. **(e)** As Figure 6.3d except the torque at which loop extrusion stalls (τ_{stall}) was set to 8 k_BT. **(f)** The indicated forms of cohesin were added at a 20:1 cohesin:plasmid ratio in a buffer containing 25 mM NaCl. Smc1 4E-EQ/EQ phenocopies EQ/EQ cohesin. **(g)** As Figure 6.3c except the torque at which loop extrusion stalls (τ_{stall}) was set to 8 k_BT and only a single step per cohesin was allowed.

in the presence of human topoisomerase I, which is capable of relaxing both positively and negatively supercoiled DNA. We then deproteinized these samples and performed agarose gel electrophoresis in the presence of a defined concentration of the DNA intercalator chloroquine. Chloroquine induces positive supercoiling, leading to increased

electrophoretic mobility for DNA that was already positively supercoiled and reduced mobility for DNA that was negatively supercoiled Figure 6.3b). At low cohesin:plasmid ratios in the presence of ATP, cohesin induced a concentration-dependent increase in slower migrating DNA bands, indicative of negatively supercoiled plasmids after deproteinization (Figure 6.3c). In contrast, at cohesin:plasmid ratios above ~20:1, faster migrating bands became predominant, indicating the appearance of positively supercoiled reaction products. The appearance of both forms of supercoiling was reduced in the absence of ATP [51]. The appearance of negatively supercoiled plasmids at low cohesin:plasmid ratio was dependent on NIPBL-Mau2 and was slightly enhanced in the presence of STAG1 [51]. Summing up, like condensin, cohesin generates plasmids with net-negative supercoiling at low cohesin:plasmid ratios and net-positive supercoiling at high cohesin:plasmid ratios.

6.2.5 Simulations based on single-molecule behavior recapitulate the bulk plasmid experiments

It is a priori unclear why the plasmid assay reports a varying handedness of the deproteinized reaction depending on the molar excess of SMCs to plasmid. Based on our magnetic tweezers result that all SMC complexes induce an average of $\Delta L_k = -0.6$ into the extruded loop, we set out to model DNA supercoiling in the plasmid assay based on the following simple assumptions: (i) Each DNA loop extrusion step of every cohesin molecule acting on the plasmid adds a negative twist of -0.6 into the loop, and necessarily therefore also a compensatory +0.6 twist in the unlooped region of the plasmid. Furthermore, (ii) we take into account that twist that exists in the unlooped region is partially being transferred into the loop in each step according to the relative twist per length in the unlooped region and the DNA-loop-extrusion step size (see Methods section 6.4). (iii) Human topoisomerase I is assumed to bind with uniform probability across the plasmid and removes available excess linking number changes (for both positively and negatively supercoiled DNA) in the respective segment at rate k_{topo}. To compare to the experiments, we counted the number of cohesin complexes on the plasmid and the removed linking numbers throughout the simulation, and - like in the experiments - evaluated the final result after a set time.

This simple model recapitulates the previously unexplained transition from negatively to positively supercoiled products (Figure 6.3d). Note that topoisomerases do not reduce L_k for a specific handedness during the simulations, but both negative and positive supercoils are removed. Graphs as shown in Figure 6.3d report the net linking number of the plasmid at the end of the simulation (after 1800 steps, which roughly corresponds to 15 min; see Methods section 6.4. We further analyzed how the introduction of only negative twist into the loop (and consequently positive twist into the unlooped region of the plasmid) can result in net-negative at low and net-positive supercoiling at high cohesin:plasmid ratios. To this end, we computed the fraction of the plasmid which carried a negative or positive twist at every time point throughout the simulations and averaged the ratio of the positive and negative fraction over time and several independent simulations (Figure S6.4a). Wild-type cohesin generated a regime below a cohesin:plasmid ratio of ~30 at which most of the plasmid is positively supercoiled. In this regime, topoisomerases which bind uniformly across the plasmid have a higher chance of encountering a positively supercoiled

segment (a loop, Z-loop, or unlooped segment). This leads to a preferential relaxation of positive supercoils by topoisomerases, yielding a net-negative reaction product.

We next tested whether ATP hydrolysis is necessary to detect topoisomerase-fixed supercoils in the plasmid assay. To this end, we experimentally repeated bulk experiments with cohesin harboring ATP hydrolysis mutations (EQ/EQ, see Figure 6.2a,b and Figure S6.2a) for their ability to induce supercoiling at low cohesin:plasmid ratios in the presence of NIPBL-Mau2 (Figure 6.4a). Only negatively supercoiled plasmids were detected in the presence of EQ/EQ cohesin than in the presence of wild type cohesin (Figure 6.3c). In simulations, cohesin EQ/EQ is modelled as the wild type but the SMC is only allowed to make a single step whereupon it stalls indefinitely (up until SMC dissociation). The model reproduces the observation that products of reactions containing EQ/EQ cohesin are always negatively supercoiled (Figure 6.4b). This result can be understood from the preferential action of topoisomerase on the unlooped region: Because topoisomerase can only alter the linking number by 1 or multiples thereof, topoisomerase activity is excluded from loops which are only a single step long because those carry only -0.6 twist. The unlooped region, however, accumulates twist from several EQ/EQ mutants which leads to a twist accumulation which is relaxed by TopoI (Figure S6.4a).

6.2.6 A mutation of the DNA binding sites on cohesin's ATPase heads is predicted to lower cohesin's stalling torque

Cryogenic electron microscopy (cryo-EM) [200, 234, 235] and single-molecule Förster resonance energy transfer experiments [238] have revealed that EQ/EQ mutations stabilize the head-engaged state of cohesin in which DNA is clamped onto the ATPase heads by NIPBL, presumably because these mutants can bind but not hydrolyze ATP. Our finding that EQ/EQ cohesin and wild type cohesin in the presence of AMP-PnP - and thus ATP binding alone - is sufficient to induce the negative twist of -0.6 into the loop (Figure 6.2 and Figure 6.4a) suggests that cohesin induces negative supercoiling of the loop during DNA clamping. This hypothesis predicts that cohesin mutants that are defective in DNA clamping will be defective in their ability to supercoil the loop negatively. To test this prediction, we analyzed cohesin complexes in which the DNA binding sites on the ATPase heads of SMC1 were mutated (Figure 6.4c [238], more mutations were tested in [51]. Indeed, mutation of the DNA binding site on the SMC1 head (SMC1 4E) and on the SMC3 head (SMC3 4E [51]) both greatly reduced cohesin's ability to generate negatively supercoiled plasmids (Figure 6.4d). Unexpectedly, however, these experiments revealed that cohesin containing the SMC1 4E mutation was not only defective in generating negatively supercoiled plasmids but instead generated positively supercoiled plasmids in a strictly ATP-dependent manner (Figure 6.4d and [51]). This phenomenon was not specific to this particular charge-reversal mutant of the SMC1 ATPase head, since a five amino acid deletion in SMC1 that includes two of the residues mutated in SMC1 4E (SMC1 Δ58-62) also displayed increased positively supercoiled products [51]. Since wild type cohesin generates negative supercoiling at cohesin:plasmid ratios up to around 20:1 and positive supercoiling above this Figure 6.3c, it is possible that cohesin SMC1 4E generates positively supercoiled plasmids because it is more active in supercoiling than wild type cohesin, i.e. whether it would mimic a high

cohesin:plasmid ratio situation. If so, one might expect to observe negative supercoiling when SMC1 4E is incubated at cohesin:plasmid ratios below 20:1. However, only positively supercoiled plasmids could be detected at all cohesin:plasmid ratios tested, including the low ones (Figure 6.4d).

The results described above indicate that cohesin SMC1 4E and other ATPase head mutants generate positively supercoiled plasmids, presumably because they fail to clamp DNA. We therefore wondered whether wild type cohesin also switches from negative to positive supercoiling of the plasmids at high cohesin:plasmid ratios because cohesin fails to clamp DNA under these conditions. For example, it is conceivable that at high cohesin:plasmid ratios so many cohesin molecules initiate loop extrusion that the torque generated by these events makes it impossible for cohesin to further clamp and supercoil DNA. To test this possibility, we searched for other conditions in which the ability of wild type cohesin to clamp DNA might be reduced. Since DNA-protein binding events depend on electrostatic interactions we reasoned that simply increasing the ionic strength in supercoiling assays might have such an effect, i.e. might reduce cohesin's ability to clamp DNA. If so, then increasing the salt concentration in supercoiling assays containing wild type cohesin at low cohesin:plasmid ratios would be predicted to phenocopy the effect of the SMC1 4E mutation, i.e. generate positively supercoiled plasmids.

We therefore performed experimental DNA supercoiling reactions with wild type cohesin in the presence of varying concentrations of NaCl (Figure S6.4b). As expected, the supercoiling activities of cohesin decreased as the concentration of NaCl was increased. However, the handedness of supercoiling was also inverted: in the presence of 25 mM NaCl only negative supercoiling was detectable, at 50 mM NaCl both negative and positive supercoiling occurred, but at 75 mM NaCl positive supercoiling became predominant. This remarkable sensitivity of the supercoiling handedness to salt concentration suggests that wild type cohesin and SMC1 4E switch from generating negatively to positively supercoiled plasmids for similar reasons, namely in situations in which they are defective to clamp DNA. According to this hypothesis, the critical difference between wild type and SMC1 4E cohesin might be their stall torque τ_{stall}, i.e the torque in the DNA that is fed into the loop or the torque present within the loop that stalls their ability to transiently clamp DNA to extrude loops and thus to negatively supercoil extruded DNA.

6.2.7 SMC1 4E cohesin negatively supercoils the loop during single loop extrusion steps

To analyze the effect of single SMC1 4E cohesin complexes on individual stretched DNA molecules, we used the magnetic tweezers assay in which the degree of DNA twisting that occurs during individual loop extrusion steps can be measured (subsection 6.2.1). In this assay, SMC1 4E cohesin generated a negative twist of -0.63 ± 0.12 (mean ± SD) in each loop extrusion step (Figure S6.5a-d), which is a value that has the same sign and degree of twist as observed in the presence of wild type cohesin, EQ/EQ cohesin, as well as yeast condensin and yeast SMC5/6 (Figure 6.2 and Figure 6.5a). Thus, although the predominant phenotype of SMC1 4E in the plasmid assay is positively supercoiled products,

the complexes that bind to DNA and initiate DNA loop extrusion steps in the magnetic tweezers assay individually behave like wild type. Although SMC1 4E complexes are reduced in their ability to hydrolyse ATP and extrude DNA loops, they are not completely inactive [238]. It thus appears that both wild type and SMC1 4E cohesin can only generate negative supercoils, and that the switching to positive supercoiling is an indirect effect of the plasmid assay, in which the changes in ΔL_k that are detected by gel electrophoresis are generated by the interplay between topoisomerase I and cohesin.

Based on the hypothesis that the DNA clamping deficiency of SMC1 4E can be expressed as wild type cohesin with a reduced stall torque, τ_{stall}, we simulated the outcome of the plasmid assay when we reduced the stall torque from 12 k_BT for wild type cohesin (Methods section 6.4) to 8 k_BT. This led to a disappearance of the cohesin:plasmid regime in which negatively supercoiled plasmids are observed (Figure 6.4e), in striking agreement with the experiment (Figure 6.4d). Restricting the number of loop extrusion steps to 1 in order to simulate the combination of the EQ/EQ and SMC1 4E mutation, yielded only negatively supercoiled plasmids, indistinguishable from the reaction products of EQ/EQ cohesin and in agreement with the experiments (Figure 6.4f,g). SMC1 4E cohesin fails to generate negatively supercoiled plasmids because the plasmid is, on average, always predominantly negatively supercoiled, causing a bias of topoisomerases to remove negative supercoils which results in positively supercoiled reaction products at all tested cohesin:plasmid ratios (Figure S6.4a). With only a minimal number of assumptions, the model thus recapitulates all the experimentally observed trends.

We also tested whether the selective relaxation of negative supercoils by *E. coli* TopA can be recapitulated by simulations. When we restricted the topoisomerase to only act on negatively supercoiled DNA, the simulation accordingly produced only positively supercoiled products. While so far we obtained excellent agreement with all experiments, this finding stands in contrast to the experimental result which shows that a low cohesin:plasmid ratio in the presence of TopA did not result in any linking number changes (Figure S6.5e,f). However, we note that TopA selectively relaxes only negative supercoils by recognizing ssDNA that are only generated by very strong underwinding of DNA [343, 344] or at large forces acting on the DNA. In the absence of tension on the DNA, this occurs only at a supercoiling density $\sigma = (L_k - L_{k0})/(L_{k0}) \approx -0.065$ [345]. When incorporating the constraint that TopA can only act on segments of the plasmid with a supercoiling density of $\sigma \leq -0.065$, the simulations reproduced the experimental outcome, i.e. no linking number changes could be made by topoisomerases at a low wild type and EQ/EQ cohesin:plasmid ratio (Figure S6E-F). Furthermore, SMC1 4E cohesin is unable to generate plasmid segments with sufficiently low supercoiling densities since its stall torque is reduced compared to wild type cohesin. This results in the absence of any linking number changes in simulations (Figure S6.5g) with SMC1 4E cohesin and TopA, in contrast to the experiment (Figure S6.5f). This may hint toward an incompletely understood interaction of SMCs with topoisomerases which induce a bias towards the relaxation of supercoils in the looped versus the unlooped region of the plasmid.

We conclude that the simple modeling largely recapitulates all experimental data, where, notably, cohesin induces strictly negative supercoiling into the extruded DNA loop in an ATP-binding and NIPBL-dependent manner. Mutations that weaken the DNA binding affinity of DNA onto the ATPase heads and cohesin in high ionic strength conditions can be described as a phenotype of cohesin whose stall torque is reduced compared to wild type cohesin.

6.3 Discussion

We here presented a high-resolution magnetic tweezers assay on single torsionally constrained DNA molecules that directly measures how much twist is induced by an individual SMC in each loop-extrusion step. The data provide direct evidence that all eukaryotic SMC complexes negatively twist DNA by the same amount at every loop-extrusion step. The linking number change of $\Delta L_k = -0.6$ is generated upon ATP binding and is found to be a conserved value across the eukaryotic SMC family.

Previous deletion and mutation studies with bulk plasmid assays suggested that SMC proteins induce positive supercoiling of genomic DNA *in vivo* in eukaryotes [346] and in prokaryotes [347, 348]. Positive DNA supercoiling by *Xenopus* 13S condensin was also reconstituted *in vitro* [48]. Recently, however, it was discovered that two modes of supercoiling can be observed in plasmid assays for condensin [49] and cohesin (see subsection 6.2.4 [51]), depending on the molar ratio of SMC to plasmid as well as the absolute protein and DNA concentrations [49]. Using these assays, it appeared that condensin and cohesin twist DNA positively when it is in large excess to DNA, but negative twist is induced at low protein-DNA ratios and absolute concentrations (ref. [49, 51]). However, the magnetic tweezers experiments and simulations presented here (subsection 6.2.5) demonstrate that the conversion of negatively to positively supercoiled products can be recapitulated by solely requiring a negative twist of $\Delta L_k = -0.6$ which is induced into the loop at every loop extrusion step.

Kimura et al. [48] noted that ATP hydrolysis is necessary for the supercoiling reaction to occur. By contrast, for conditions where ATP hydrolysis was inhibited, we observed that SMCs generate at most one step with -0.6 turns (Figure 6.2a,d,g and Figure S6.2) - which shows that SMCs induce negative twist into DNA upon ATP binding alone. Possibly, the topoisomerases used in the experiments of ref. [48] was unable to catalyse sufficient changes in linking number to be resolved in these bulk assays. In contrast, Martínez-García et al. [49] found that incubation of the supercoiling reaction with ATP and subsequent incubation with an excess of AMP-PnP increases the amount of topo I-fixed supercoils 2-fold compared to an incubation with ATP alone. From the average levels of induced supercoiling, the authors estimated that a negative twist of $\Delta L_k = -0.8$ is induced upon ATP binding which is reduced to $\Delta L_k = -0.4$ upon ATP hydrolysis. While the induction of negative supercoils during each extrusion step aligns with our observation, the ΔL_k values differ from our direct single-molecule measurements, as we observed that ATP binding alone induces a twist of $\Delta L_k = -0.6$ which remains unchanged by subsequent ATP hydrolysis. Possibly, condensin is stabilized on DNA in the ATP-bound state in the plasmid assay, preventing the loss of condensin-mediated supercoils upon loop dissociation before topo I could have permanently induced a linking number change. The discrepancy

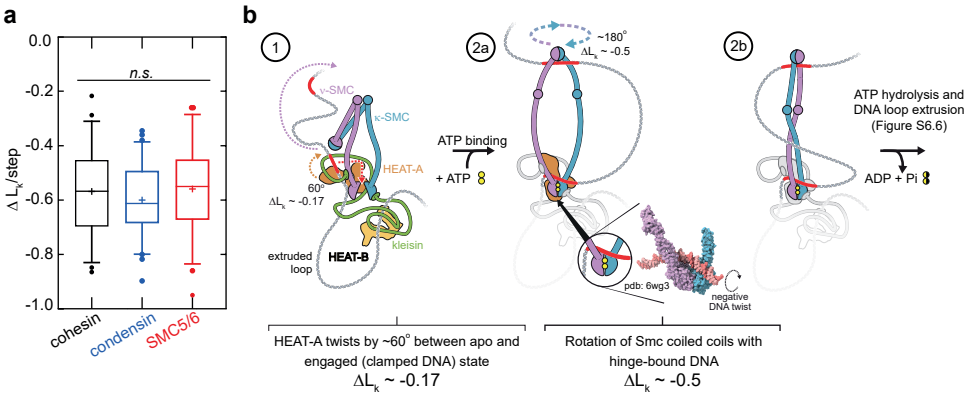


Figure 6.5 | Eukaryotic SMC complexes induce comparable negative DNA twist per loop extrusion step. (**a**) Induced twist per step for all three eukaryotic SMC complexes, where all values of the changes in the linking number ΔL_k per DNA-loop-extrusion step were pooled from Figure 6.2b,e,h (N = 59 for cohesin, N= 64 for yeast condensin, N= 56 for SMC5/6). Statistical significance was assessed using a one-way analysis of variance (ANOVA) with a significance level $\alpha = 0.05$ (95% confidence interval; n.s.: $p > 0.05$) (**b**) Potential mechanism yielding a negative DNA twist of $\Delta L_k \sim -0.6$ based on the Reel & Seal model (ref. [247]). DNA is nontopologically held by kleisin and HEAT-A (step 1). Upon ATP binding, clamping of DNA on top of the engaged ATPase heads temporarily forces a small loop of DNA pseudo-topologically into the Smc lumen (step 2a). Rotation of HEAT-A (if bound to DNA in the apo as well as the engaged state) induces a negative twist of about 60°, yielding $\Delta L_k \sim -0.17$ (Figure S6.6d,e), into the loop (see inset for twist direction [200]). Spontaneous rotation of the coiled coils with hinge-bound DNA by about 180° (ref. [238, 239, 349]) induces a further linking number change of $\Delta L_k \sim -0.5$ into the loop (step 2b).

between the estimates of ΔL_k per ATPase cycle by Martínez-García et al. [49] and our results may be explained by the assumptions that had to be made in the bulk assay to assess the supercoiling activity of condensin. Specifically, all condensin molecules were assumed to be active and topo I was assumed to be 100% efficient in removing the generated supercoils. In contrast, our single-molecule experiments directly measure the induced twist on the single-step level in loop extrusion by a single SMC complex, without any treatment by topoisomerases.

When we converted the SMC5/6 hexamer to a SMC5/6 octamer, we observed that the octamer can induce DNA twist in a single step (inducing a linking number change of $\Delta L_k = -0.51 \pm 0.17$, mean ± SD) despite its inability to hydrolyse ATP [342, 350] or extrude DNA loops (Figure 6.2g, Figure S5.2d, and refs. [204, 222]), suggesting that the SMC5/6 octamer is able to bind ATP. Since the SMC5/6 octamer has been shown to topologically embrace DNA [204, 309], it is possible that the ATP binding-induced power stroke by the SMC5/6 octamer serves as an intermediate to the topological loading reaction of SMC complexes onto DNA. Current models of DNA loop extrusion postulate that a transient DNA loop is inserted pseudo-topologically into the SMC lumen upon ATP binding to the complex due to a DNA clamping onto the engaged Smc ATPase heads [104, 200, 221, 233, 234, 236, 241, 242, 247, 309] which causes the power stroke [242]. Transient hinge opening in this state would lead to a topological entrapment of the DNA duplex (see Figure S6.6g-i for such a potential pathway to topological entrapment).

Since DNA twisting occurs concomitantly with loop extrusion, we conclude that DNA twisting by -0.6 turns in each step is inherent to the ATP-driven DNA loop extrusion cycle of SMC complexes. While the DNA loop extrusion mechanism still remains incompletely resolved, future models have to quantitatively consider the twist of -0.6 turns induced into the loop upon ATP binding. Based on previous structural data, we can point to a few elements that may explain the mechanism driving the induced negative supercoiling (Figure 6.5b) in a Reel & Seal model [247] (see also subsection 1.5.4) with a twist. Upon ATP binding, the HEAT-A subunit (accessory subunit I, NIPBL-Mau2 for human cohesin in Figure S6.1a) undergoes a rigid body rotation about 60° ($\Delta L_k \sim -0.17$) which forms a positively charged channel to clamp DNA onto the ATPase heads [200, 235] (Figure 6.5b step 1 to 2a, Figure S6.6a,b). While the DNA trajectory within SMC complexes in the apo state has so far not been resolved, continuous DNA binding to HEAT-A in both states may direct DNA to the binding site formed by the engaged ATPase heads and thus induce $\Delta L_k \sim -0.17$ into the extruded loop. Upon head engagement, the coiled coils unfold [238] and may spontaneously intertwine (Figure 6.5b step 2a-b) as has been observed in Molecular Dynamics simulations [239] (Figure S6.6c), by negative staining Electron Microscopy [349] (Figure S6.6d), and high-speed Atomic Force Microscopy [238] (Figure S6.6e). DNA binding to the hinge has been proposed to be essential for DNA loop extrusion [238]. If intertwining of the coiled coils occurs after DNA is bound at both the hinge and the ATPase heads, this would induce a negative linking number change of about -0.5 into the DNA loop. Such a mechanism is supported by the finding that a mutant which is defective in DNA binding at the hinge [238] shows reduced supercoiling activity [51]. These two contributions, the rigid-body rotation of HEAT-A and the spontaneous intertwining of the Smc coiled coils, therefore may additively contribute an induced linking number of -0.6 to -0.7, which is close to the observed negative linking number change of -0.6 upon ATP binding. Subsequent ATP hydrolysis disengages the ATPase heads and transfers the accumulated twist into the extruded loop, clearing the way for the next cycle of DNA loop extrusion (Figure S6.6f).

What is the impact of SMC-induced DNA supercoiling on the genome? DNA binding of CCCTC-binding factor (CTCF) demarcates topologically associated domains (TADs) *via* an orientational interaction of CTCF with cohesin [33–35, 39, 42, 97, 141, 148]. The ~2- to 3-fold higher contact frequency within TADs than across TADs [162, 304, 351, 352] might be partially explained by the barrier function of CTCF on loop progression. However, it is also plausible that supercoiling accumulates within a subset of (sub-)TADs [353] which could account for the increased intra-TAD contact frequency [352]. While the accumulation of supercoils within TADs has so far been thought to be due to transcription-mediated supercoiling [54, 352, 354], this could also be the result of cohesin-mediated loop extrusion itself. The accumulation of supercoiling within TADs can contribute to their insulating property [355] as well as to the facilitation of long-range contacts between genomic loci within a TAD, e.g. promoter-enhancer contacts [356]. The latter could explain how cohesin may drive transcription of enhancer-controlled developmental genes even in the absence of loop anchors positioned at enhancer-promoter pairs [181]. SMCs preferentially bind and loop onto positively over negatively supercoiled DNA [323, 325, 357], which could prevent SMCs from loading onto already extruded (negatively supercoiled) extruded loops. This bias could contribute to the appearance of characteristic CTCF-dependent TAD corner dots in Hi-C maps [148, 155, 249] since those depend on a fraction of TADs whose DNA is

completely contained in SMC-mediated loops [268] as observed *in vivo* [160, 161]. Similarly, extrusion of entire chromosomes without gaps between loop anchors is instrumental to compact metazoan mitotic chromosomes for sister chromatid segregation [268], which could also be mediated by the preferential recruitment to positively supercoiled, non-looped DNA. Finally, the SMC-mediated positive supercoiling of genomic regions outside loops might direct topo II to decatenate sister chromatids before relaxing supercoils [346].

Our observation that all eukaryotic SMC complexes twist DNA negatively at every loop-extrusion step establishes DNA supercoiling as an integral part of the DNA loop-extrusion mechanism. The measured quantitative loop extrusion-induced DNA twist of -0.59 ± 0.02 (mean ± SD) will help modelling and polymer simulations to capture yet finer structures of genomes [352], and will inform refined models of DNA loop extrusion with the handedness and amount of induced twist upon ATP binding.

6.4 Methods

Experimental model and subject details

Cloning and mutagenesis of cohesin forms

pLib vectors containing *homo sapiens* (hs) Smc1, Smc3-FLAG, Scc1(TEV)-HALO in which three internal tabacco etch virus (TEV) sites were inserted into Scc1, and 10xHis-STAG1 were as described [173]. pLib vectors containing Smc1 E1157Q and Smc3 E1144Q-FLAG were as described [173]. pLib vectors containing Smc1 4E (K52E, R57E, K59E, K62E) were as described [238]. A pLib vector containing FLAG-NIPBL-HALO-10xHIS and untagged Mau2 was as described [173]. A pLib vector containing 10xHis-Scc1(TEV)-HALO was generated using Gibson Assembly [358]. A pLib vector containing Scc1 with the mutations R172A, D279A and R450A, which prevent separase and caspase cleavage of Scc1 [200], and a C-terminal HALO tag was generated using Gibson Assembly [358].

pLib vectors encoding cohesin subunits were subcloned into pBig1 polycistronic expression vectors using bigBAC [359]. Smc1 and Smc3 expression cassettes were combined into pBig1a or pBig1b vectors. Scc1 and STAG1 expression cassettes were combined into pBig1a or pBig1b vectors. pBig1a Smc1, Smc3-FLAG and pBig1b Scc1(TEV)-HALO, 10xHis-STAG1 constructs were combined into a pBig2ab vector for expression of the tetrameric cohesin used in Figure 6.3 and Figure 6.4 [173]. All other tetrameric forms of cohesin were expressed by co-infection with baculoviruses generated from pBig1 Smc1, Smc3-FLAG (wild type or mutant) and pBig1 Scc1 R172A/D279A/R450A-HALO, His-STAG1 vectors.

Protein expression and purification for magnetic tweezers experiments

Wild type and EQ/EQ human cohesin, as well as NIPBL-Mau2 were expressed in and purified from *Sf9* insect cells as described previously [173]. *S. cerevisiae* condensin was expressed in and purified from *S. cerevisiae* as described previously [98]. *S. cerevisiae* SMC5/6 was expressed in and purified from *E. coli* as described previously [309].

Protein expression and purification for bulk plasmid experiments

pLib, pBig1a, pBig1b and pBig2ab vectors were transformed into DH10EmBacY cells and bacmid DNA was isolated as described [360]. *Sf9* insect cells (Thermo Fisher Scientific) were transfected with bacmid DNA using Fugene 6 (Promega). The baculovirus-containing cell supernatant (V0) was collected 96 hours after transfection and used to infect 50 ml *Sf9* cells at a density of 10^6 cells/ml. Cells were grown in Grace medium at 100 rpm and 27°C and centrifuged 72 hours after infection and the cell supernatant (V1) was then used to directly infect 750 ml expression cultures at a density of $1.2 \cdot 10^6$ cells/ml. Expression cultures were grown in Grace medium at 100 rpm and 27°C. Cells were centrifuged around 54 hours after infection, washed in PBS, frozen in liquid nitrogen and stored at -80°C.

Human topoisomerase I purification: All steps were performed at 4°C unless indicated otherwise. Around 20 ml cell pellet was thawed and lysed by Dounce homogenization in TopoI purification buffer 1 (25 mM Na_2HPO_4/NaH_2PO_4 [pH 7.5], 500 mM NaCl, 5% glycerol) supplemented with 10 mM imidazole [pH 7.5], 0.05% Tween20, 1 mM PMSF, 3 mM β-mercaptoethanol, 10 µg/ml aprotinin, 2 mM benzamidine and cOmplete EDTA-free protease inhibitor cocktail (Merck; 11873580001). After centrifugation (47000 g, 45 min),

the soluble fraction was combined with 4 ml of Toyopearl AF-chelate-650M resin (Tosoh Bioscience) pre-charged with Ni2+ ions and incubated for 2 h. Beads were washed with 3 x 10 bead volumes of TopoI purification buffer 1 supplemented with 25 mM imidazole [pH 7.5]. Bound protein was eluted with 25 ml TopoI purification buffer 2 (25 mM Na_2HPO_4/NaH_2PO_4 [pH 7.5], 150 mM NaCl, 5% glycerol, 300 mM imidazole [pH 7.5]). The eluate was then combined with 5 ml of FLAG-M2 agarose resin (Sigma; A2220) and incubated for three hours. Beads were washed with 3 x 10 bead volumes of TopoI purification buffer 3 (25 mM NaH_2PO_4/Na_2HPO_4 pH 7.5, 150 mM NaCl, 5% glycerol, 50 mM imidazole pH 7.5). Bound protein was eluted with 15 ml TopoI purification buffer 3 supplemented with 0.5 mg/ml 3xFlag peptide. Eluates were concentrated to ~0.75 ml using Vivaspin 20 100 kDa MWCO ultrafiltration units (Sartorius; VS2042), frozen in liquid nitrogen and stored at -80°C.

Method details

Synthesis of torsionally constrained dsDNA

Linear, torsionally constrained dsDNA constructs were synthesized based on a 3.6 kbp fragment of the pRL-SV40 plasmid (Promega, USA), digested with BamHI and XbaI. Afterwards, the fragment was enzymatically ligated to 600 bp dsDNA handles that contained multiple digoxigenin on one end of the dsDNA construct and multiple biotin at the other end of the dsDNA construct, as previously described [361, 362]. To synthesize these handles, a 1.2 kb fragment from pBluescript (Stratagene, USA) was amplified by PCR in the presence of Biotin-16-dUTP (Roche, Switzerland) and Digoxigenin-11-dUTP (Roche, Switzerland) in a 1:10 ratio, using forward primer 5'-GACCGAGATAGGGTTGAGTG and reverse primer 5'-CAGGGTCGGAACAGGAGAGC. Prior to enzymatic ligation *via* T4 DNA ligase (New England Biolabs, UK) overnight, the handles fragment was digested with either BamHI or XbaI. The final dsDNA construct was cleaned up from the access of handles by running on a 1% agarose gel and extracting the dsDNA construct using a gel purification kit (A9282, Promega).

Magnetic tweezers

The magnetic tweezers used in this study was described previously [217, 363]. Briefly, light transmitted through the sample was collected by a 50x oil-immersion objective (CFI Plan 50XH, Achromat; 50x; NA = 0.9, Nikon) and projected onto a 4-megapixel CMOS camera (4M60, Falcon2; Teledyne Dalsa) with a sampling frequency of 60 Hz. The applied magnetic field was generated by a pair of vertically aligned permanent neodymium-iron-boron magnets (Supermagnete GmbH, Germany), separated by a distance of 1 mm, and suspended on a motorized stage (M-126.PD2, Physik Instrumente) above the flow cell. In addition, the magnet pair could be rotated around the illumination axis by an applied DC servo step motor (C-150.PD; Physik Instrumente). Image processing of the collected light allowed to track the real-time position of both surface-attached reference beads and superparamagnetic beads coupled to the torsionally constrained dsDNA construct in three dimensions over time. Bead x, y, z position tracking was achieved with a spatial resolution of ~2 nm (ref. [217]) using a cross-correlation algorithm realized with custom-written software in LabView (2011, National Instruments Corporation) [364]. The software determined the bead positions with spectral corrections to correct for camera blur and aliasing.

Measurement of DNA loop extrusion-induced DNA twist

The flow cell preparation for the magnetic tweezers experiments used in this study has been described previously in detail [217, 363]. Briefly, polystyrene reference beads (Polysciences Europe) of 1.5 μm in diameter were diluted 1:1500 in PBS buffer (pH 7.4) and adhered to 1 M KOH treated surface of the flow cell channel as fiducial markers. Next, 0.5 mg/ml digoxigenin sheep antibody Fab fragments (Roche, Switzerland) in PBS buffer were incubated for 1 h within the flow cell channel, following incubation for 2 h of 10 mg/ml BSA (New England Biolabs, UK), diluted in PBS buffer. For each experiment, 1 pM of the torsionally constrained dsDNA construct was incubated in PBS buffer for 20 min in the flow cell channel. After washing with 500 ml PBS buffer, the addition of 100 μl streptavidin-coated superparamagnetic beads with a diameter of 1 μm (diluted 1:400 in PBS buffer; MyOne #65601 Dynabeads, Invitrogen/Life Technologies) for 5 min resulted in the attachment of the beads to biotinylated dsDNA construct. Non-attached beads were washed out with PBS buffer. Prior to conducting the force-extension experiments, we assessed whether dsDNA tethers were singly tethered by applying a high force (5 pN) and 20 negative rotations, and if they were torsionally constrained by applying 20 rotations to each direction at low force (0.3 pN). Only single and torsionally constrained DNA tethers were analyzed. All experiments were conducted at room temperature (22°C). For assessing the twist generated during the DNA loop extrusion experiments, a reference rotation curve was first acquired at 0.3 pN with negative and positive rotations of each 12 turns, with a rotation speed of 0.25 turns s^{-1}. Subsequently, the SMC proteins (Wild type cohesin: 20-100 pM; Cohesin EQ/EQ: 10-20 pM; Condensin: 0.4-1 nM; SMC5/6: 3-6 nM) in presence of 1 mM ATP or 1 mM AMP-PnP (Sigma, Netherlands) - as indicated in main text - were inserted into the flow cell channel while applying 7 pN to the dsDNA tethers. The force was then quickly lowered to a constant force of 0.3 pN to record the DNA loop extrusion steps *via* the change in bead Z-position [35, 217] for 10 minutes, which is a time where the DNA loop extrusion-activity of most SMCs ceased. To assess the degree of DNA twist that was introduced by each SMC during DNA loop extrusion, another rotation curve was recorded at 0.3 pN with negative and positive rotations of 12 turns with a rotation speed of 0.25 turns s^{-1}.

Modelling of experimentally determined cryo-EM map of apo *S. cerevisiae* cohesin

A full pseudo-atomic model of *S. cerevisiae* cohesin in the apo state (in the absence of ATP and DNA) was fitted into the cryo-EM map EMD-12880 using UCSF ChimeraX (ref. [328]). The lower part containing the ATPase heads of Smc1 and Smc3, kleisin, as well as Scc2 was taken from pdb entry 6ZZ6 (ref. [234]) and the folded coiled coils were taken from pdb entry 7OGT (ref. [365]) as done previously [291]. The EM maps and protein structures in Figure 6.5b and Figure S6.6 were depicted using UCSF ChimeraX (ref. [328]).

Human topoisomerase DNA supercoiling assay

Negatively supercoiled plasmid DNA (pH6HTN (Promega), 4 kb) was relaxed by incubating it together with human topoisomerase I for 45 min at 37°C. This mixture was then supplemented with recombinant human cohesin and NIPBL-Mau2 as indicated and incubated for a further 90 min (final reaction conditions: 20 mM NaH_2PO_4/Na_2HPO_4 [pH 7.5], 5 mM Tris [pH 7.5], 25 mM NaCl, 2.5 mM $MgCl_2$, 1 mM DTT, 0.1 mg/ml BSA (ThermoFisher

Scientific; AM2616), 2 mM ATP (Jena Biosciences; NU-1049), 0.4 nM pH6HTN, 3 nM human topoisomerase I, 0 - 32 nM cohesin (corresponding to a cohesin:plasmid molar ratio of up to 80:1) and 0 - 64 nM NIPBL-Mau2 (i.e. a two-fold molar excess relative to cohesin) in a 100 µl reaction). Reactions were stopped and purified as above. Half of the reaction volume was electrophoresed in a 0.7% agarose gel in 1x TBE buffer supplemented with 0.4 µg/ml chloroquine (Sigma-Aldrich; C6628) for 17 hours at 32 V. Gels were stained with Sybr Gold (ThermoFisher Scientific; S11494) and visualized using a ChemiDoc MP imager (Bio-Rad). For the experiment described in Figure S6.5e,f in which supercoiling by cohesin in the presence of human topoisomerase I and *E. coli* TopA was compared, negatively supercoiled plasmid DNA (pH6HTN (Promega), 4 kb) was relaxed by incubating it together with human topoisomerase I or *E. coli* TopA for 45 min or 20 min, respectively, at 37°C. These mixtures were then supplemented with recombinant human cohesin and NIPBL-Mau2 as indicated and incubated for a further 90 min (final reaction conditions: 20 mM NaH_2PO_4/Na_2HPO_4 [pH 7.5], 5 mM Tris [pH 7.5], 25 mM NaCl, 2.5 mM $MgCl_2$, 1 mM DTT, 0.1 mg/ml BSA (ThermoFisher Scientific; AM2616), 2 mM ATP (Jena Biosciences; NU-1049), 0.4 nM pH6HTN, 6 nM cohesin (corresponding to a cohesin:plasmid molar ratio of 15:1), 12 nM NIPBL-Mau2 (i.e. a two-fold molar excess relative to cohesin) and either 3 nM human topoisomerase I or 2.5 units *E. coli* TopA in a 100 µl reaction). Reactions were stopped, purified and electrophoresed as above.

6

E. coli TopA DNA supercoiling assay

Negatively supercoiled plasmid DNA (pUC19, 2.7 kb) was relaxed by incubating it together with *E. coli* TopA (New England Biolabs; M0301L) for 15 min at 37°C. This mixture was then supplemented with recombinant cohesin, NIPBL-Mau2 or PDS5A as indicated and incubated for a further 60 min (typical final reaction conditions: 25 mM NaH_2PO_4/Na_2HPO_4 [pH 7.5], 50 mM NaCl, 2.5 mM $MgCl_2$, 1 mM DTT, 0.1 mg/ml BSA (ThermoFisher Scientific; AM2616), 2 mM ATP (Jena Biosciences; NU-1010 or NU-1049), 0.6 nM pUC19, 1.25 units TopA, 15 nM cohesin and NIPBL-Mau2 at two-fold molar excess relative to cohesin in a 50 µl reaction). Reactions were stopped by the addition of sodium dodecyl sulphate (SDS; 0.33% final), EDTA (6.7 mM final) and Proteinase K (0.1 mg/ml final; Sigma-Aldrich; P6556) and incubated for 40 min at 50°C. DNA was extracted using phenol:chloroform:isoamyl alcohol (Sigma-Aldrich; P2069) and then chloroform (Sigma-Aldrich; P32211) in 5PRIME Phase Lock Gel Light tubes (Quantabio; 2302820). Purified DNA was supplemented with glycerol to 10% final concentration and trace amounts of Orange G (Sigma-Aldrich; P2069). Half of the reaction volume was electrophoresed in a 0.7% agarose gel in 1x TBE buffer for 110 min at 150 V. Gels were stained with Sybr Gold (ThermoFisher Scientific; S11494) and visualized using a ChemiDoc MP imager (Bio-Rad).

1D SMC supercoiling simulations

We used a fixed-time-step Monte Carlo algorithm for 1D simulations. Each lattice site corresponds to 1 bp and the plasmid was defined with a length of $L = 4000\,bp$, with periodic boundary conditions (lattice site 4000 equals lattice site 0). Simulations were carried out at 297 K ($k_BT = 4.1\ pN\ nm$). A time step equals one loop extrusion step, which roughly translates to 0.5 s, based on the ATPase activity of cohesin of roughly 2/protein/s (ref. [173]). The simulation was run for 1800 time steps (~15 min). For every time step, the following modules are executed:

SMCs spawn

SMC lifetime and direction lifetime assignment

SMC stepping

Linking number fixation by topoisomerase

SMC dissociation

SMCs spawn: SMCs have a uniform binding probability in the absence of any accumulated twist (i.e. for the first SMC binding to the plasmid). If twist has been accumulated in the plasmid, the SMC binds to a positively supercoiled segment with a 79% probability [323]. The binding rate of cohesin to the plasmid, k_{on}, was varied over several orders of magnitude between 10^{-5} and 10^{2} and the number of SMCs on the plasmid was counted during the simulation, which was used to evaluate the cohesin:plasmid ratio in simulations and to compare to experiments.

SMC lifetime and direction lifetime assignment Upon binding, the lifetime of the SMC is drawn from an exponential distribution with mean $1/k_{off}$. Furthermore, SMCs move directionally and can switch directions with rate constant k_{switch}=0.025 s^{-1} (roughly once or twice per minute [222] and chapter 5) and direction lifetimes are drawn from an exponential distribution with mean $1/k_{switch}$.

SMC stepping: All plasmid-bound SMCs attempt to make a step in random order during one time step of the simulation. Each SMC takes a step of $\Delta l = 150\ bp$ (ref. [35]), enlarging its loop, while a twist of $\Delta Tw = -0.6$ is added to the loop. If the segment on which the step is being made has a non-zero twist itself, a part of this twist is transferred into the loop. The relative twist per bp in a segment s is $rTw_s = Tw_s/L_s$, where Tw_s is the accumulated twist and L_s is the length of segment s. The twist transferred into the loop while the loop s is being enlarged on the expense of segment s is $\Delta Tw = -0.6 + \Delta L \cdot rTw_s$. Concomitantly, a twist of $\Delta Tw = +0.6 - \Delta L \cdot rTw_s$ is added to segment s (illustrated in Figure S6.3a). Note that region s may also be a previously extruded loop such that the loading of the next SMC produces nested loops. SMCs are allowed to traverse one another [267] and generate Z-loops while doing so. If Z-loops are generated, twist is added to the resulting Z-loop delimited by the two loop anchors defining the Z-loop. Similarly, Z-loops can be resolved by nesting loops. The loop anchors are modelled as twist diffusion barriers. After every potential move, the resulting torque τ in the loop as well as in segment s is computed *via* [362] $\tau = 2\pi k_B TC \cdot Tw/L_c$, where $C = 100\ nm$ is the twist persistence length of dsDNA (ref. [366]) and $L_c = L \cdot 0.342\ nm/bp$ denotes the contour length of the plasmid. If the torque in all concerning segments remains below the stall torque, τ_{stall}, the move is accepted and rejected otherwise. Even though not measured, SMC complexes must have a stall torque which is lower than the energy of ATP binding to the complex. Here we set $\tau_{stall} = 12\ k_B T$, slightly above the energetic cost to extrude DNA (ref. [367]). Lower values were used to emulate Smc1 4E cohesin. EQ/EQ or EQ/EQ-Smc1 4E cohesin was simulated by allowing only a single step before stalling.

Linking number fixation by topoisomerases: Topoisomerases bind to the plasmid at a rate k_{topo}. The binding position is drawn from a uniform distribution across the plasmid. Human topoisomerase I is able to relax both positive as well as negative supercoils up to a maximum of $\Delta L_k = 27$ (ref. [368]), or all linking numbers if the segment topoisomerase bound to contains fewer linking numbers. *E. coli* topoisomerase I (TopA) selectively relaxes only negative supercoils by recognizing ssDNA generated by underwinding of DNA [343, 344] and changes the linking number by exactly +1 (because it is a type IA topoisomerase that uses a strand passage mechanism; ref. [369]). In the absence of tension on the DNA, this occurs at a supercoiling density $\sigma = (L_k - L_{k0})/L_{k0} \sim -0.065$ (ref. [345]), within the range of supercoiling density values found *in vivo*[370, 371]. Simulations of experiments in the presence of TopA were thus carried out by allowing the topoisomerase to only act on segments in which $\sigma \leq -0.065$. The results were consistent within the tested range $-0.075 \leq \sigma \leq -0.05$. Note, however, that such a threshold of the supercoiling density σ appears to be absent in plasmid assays since TopA is able to fully relax negatively supercoiled DNA which suggests that TopA can relax negative supercoils even at higher supercoiling densities [51].

SMC dissociation: When a SMC lifetime reaches zero, it dissociates from the plasmid. If a Z-loop is associated with the segment, this Z-loop is first resolved upon which the loop dissolves. The loop length and its twist merges with its surround segment whose relative twist and torque is updated.

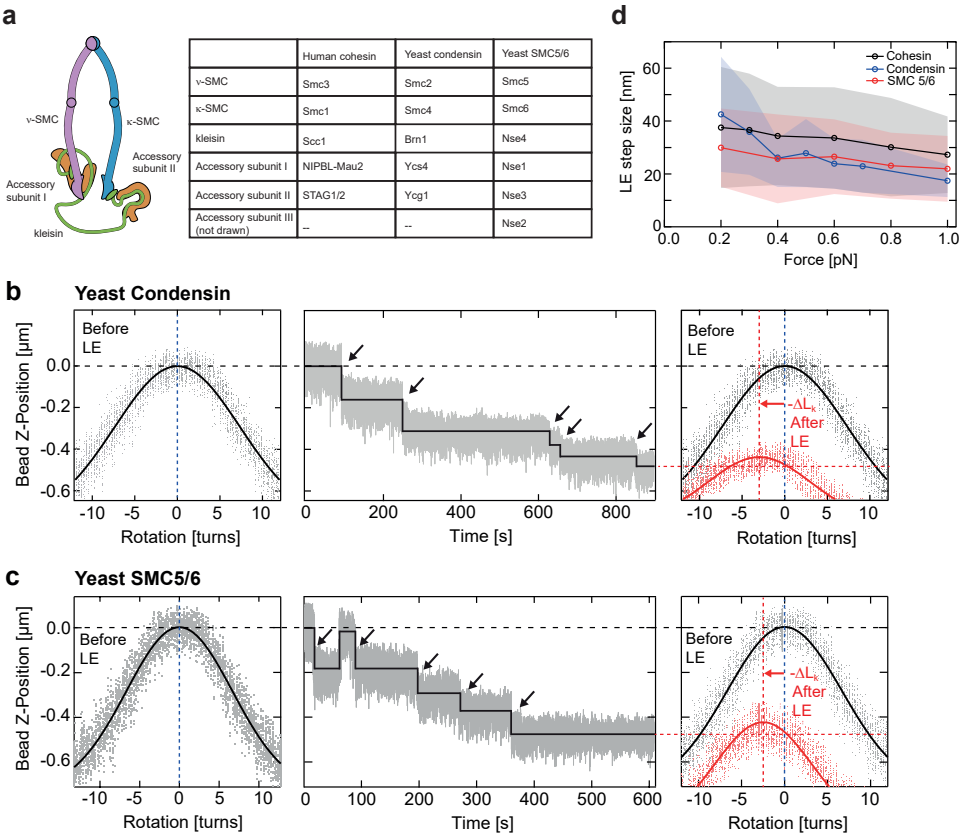
Sensitivity to parameter choices: We varied the dissociation and the topoisomerase action rate, k_{off} and k_{topo}, respectively. The results are qualitatively similar as long as the topoisomerase activity is not too low (e.g. $k_{topo} = 10^{-3}$ /step causes only 1-2 topoisomerase actions per simulation run yielding only low amounts of fixed linking numbers, see Figure S6.3b,c).

Quantification and statistical analysis

Magnetic tweezers datasets were processed and analyzed using custom-written Igor v6.37-based scripts. From our raw data, we removed traces that showed surface-adhered magnetic beads as well as dsDNA tethers where the DNA-bead attachment points were close (<100 nm) to the magnetic equator of the magnetic beads using a previously described method [372]. Tethers that detached from the surface during the measurement were also rejected from further analysis. All traces and rotation curves resulting from experiments conducted at identical conditions for each SMC were pooled and filtered to 1 Hz (moving average) for extrusion-step identification and prior to Gaussian fitting of the rotations curves. DNA loop extrusion step sizes were extracted from the change in bead Z-positions in the traces using a previously described methodology and stepfinder algorithm [217, 373]. The induced DNA twist during loop extrusion was determined by fitting a Gaussian to the bead Z-positions as a function of applied turns. The peak position of the fitted Gaussian before DNA loop extrusion C_{before} was subtracted from the peak position determined after the DNA loop extrusion experiments C_{after} which reflects the induced DNA twists $\Delta L_k = \Delta Tw = C_{after} - C_{before}$ in the unit of turns for each trace. The statistical analyses

comparing the DNA twist results for each SMC and condition were conducted using one-way analysis of variance (ANOVA) with a significance level $\alpha = 0.05$ (95% CI).

6.5 Supplementary Information



	Human cohesin	Yeast condensin	Yeast SMC5/6
ν-SMC	Smc3	Smc2	Smc5
κ-SMC	Smc1	Smc4	Smc6
kleisin	Scc1	Brn1	Nse4
Accessory subunit I	NIPBL-Mau2	Ycs4	Nse1
Accessory subunit II	STAG1/2	Ycg1	Nse3
Accessory subunit III (not drawn)	--	--	Nse2

Figure S6.1 | Nomenclature and step size distribution versus applied force for human cohesin, yeast condensin, and yeast SMC5/6. Exemplary rotation curves for yeast condensin and yeast SMC5/6
(**a**) Nomenclature for human cohesin, budding yeast condensin, and budding yeast SMC5/6. (**b**) Same as shown in Figure 6.1a-c, for yeast condensin. Black arrows in DNA loop extrusion trajectories depict single steps. (**c**) Same as (b) for yeast SMC5/6. (**d**) DNA loop extrusion step size versus DNA tension for all SMC families. Shown is the median step size as circles. Shaded areas denote the interquartile range as an indication of the spread of the data. The step size data for Condensin originates from the previous study Ryu et al. (ref. [217]). Sample size (individual steps) per force for human cohesin (black) was N = 959, 1054, 1256, 1430, 1951, 701 for increasing force; for yeast condensin (blue) N = 976, 565, 1344, 468, 605, 510, 556; and for yeast SMC5/6 (red) N = 109, 189, 206, 278, 281.

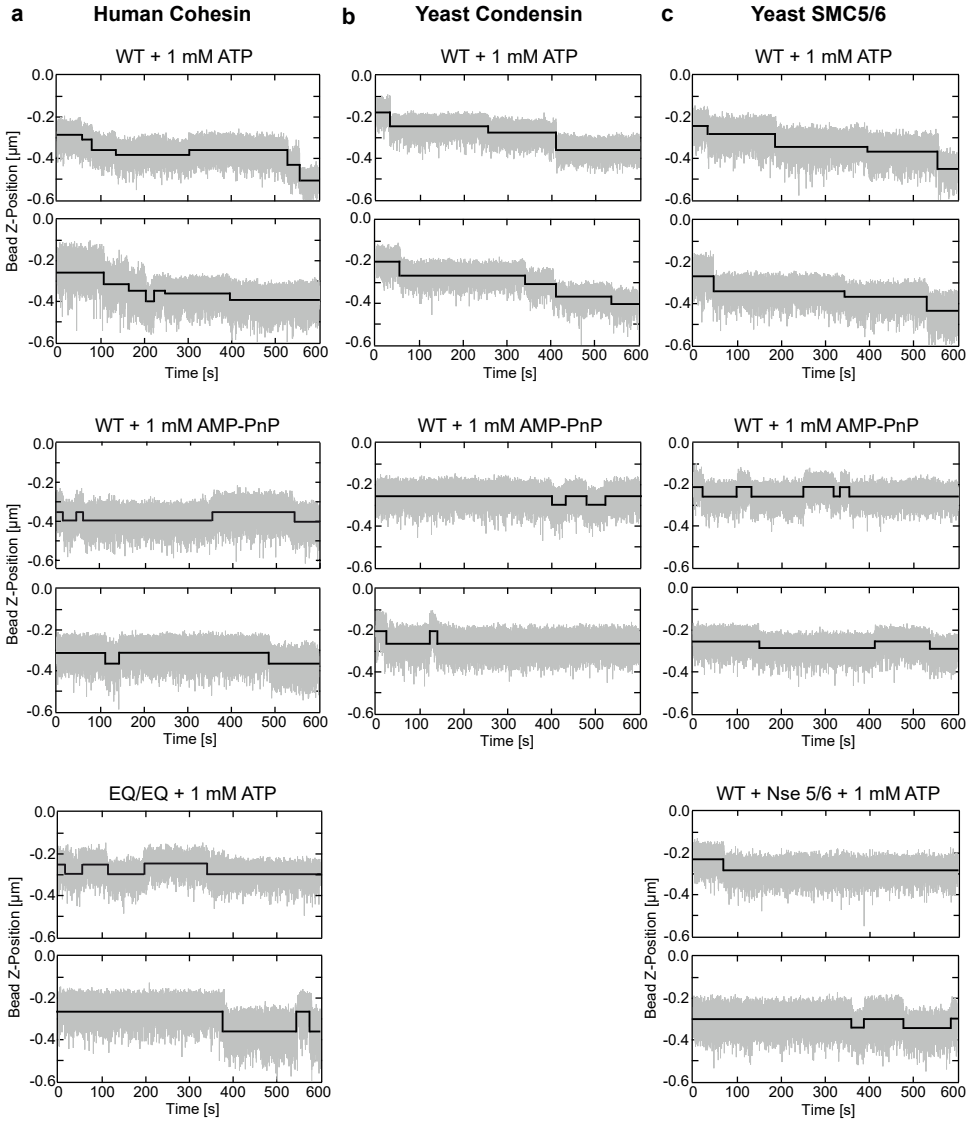


Figure S6.2 | Example DNA loop extrusion stepping trajectories for all eukaryotic SMC complexes. (**a**) Exemplary traces of cohesin-mediated DNA loop extrusion steps in the presence of ATP and wild type (WT) cohesin, AMP-PnP and WT cohesin, and ATP with the cohesin EQ/EQ mutant. (**b**) Exemplary traces for yeast condensin as in panel (a). (**c**) Exemplary traces for yeast SMC5/6 as in panel (a). Note on the bottom right panel: the SMC5/6 hexamer forms an octamer with Nse5/6. Traces show only a single downward step or repeated downward-reverse steps in the presence of AMP-PNP, the cohesin EQ/EQ mutant with ATP, or the SMC5/6 octamer in the presence of ATP.

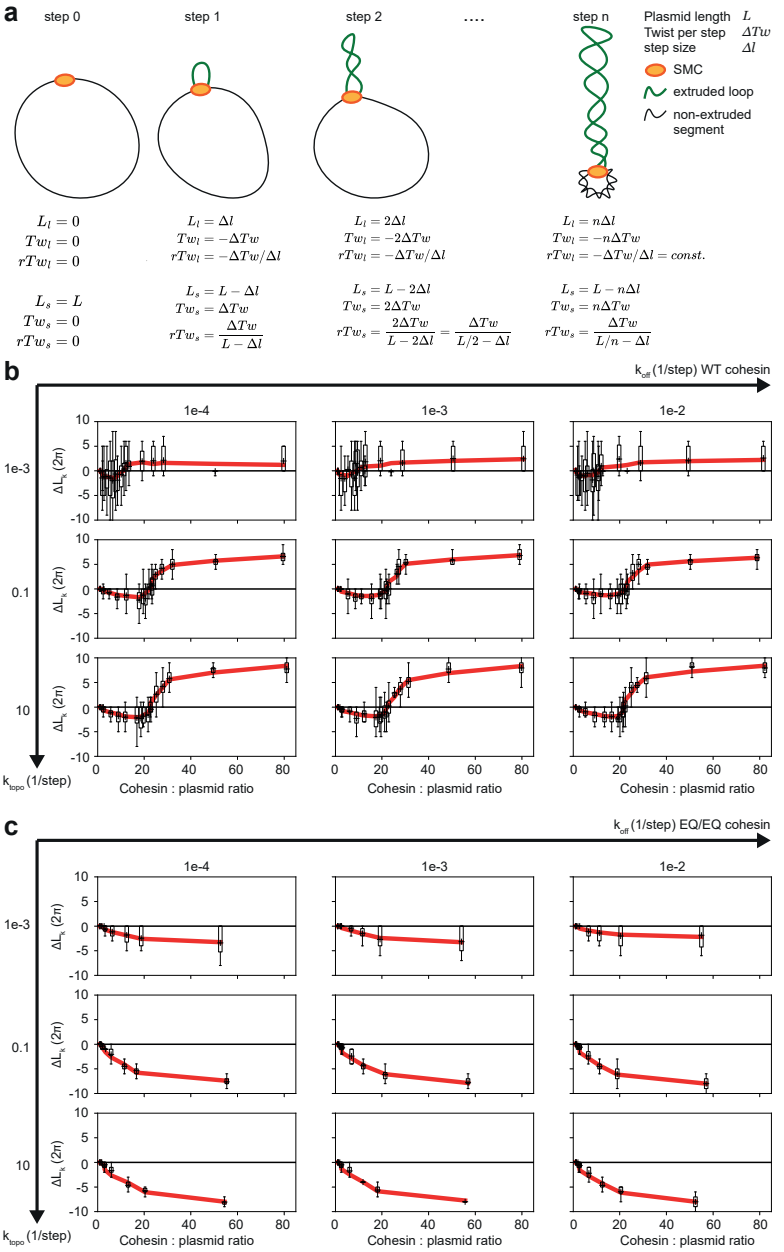


Figure S6.3 | Illustration of the supercoiling simulations and sensitivity of the simulations to parameter choices.

(**a**) Illustration of the twist (Tw), segment length (L), and relative twist (rTw) in the loop l and the plasmid segment s during the first n steps of loop extrusion by a single cohesin complex. The relative twist within the plasmid segment s increases with n while the relative twist (and thus supercoiling density and torque) within the loop remains approximately constant in this simple scenario. *Caption continues on the next page.*

Figure S6.3 | (**b**) Simulations of wild type cohesin as in Figure 6.3d for varying cohesin off-rates k_{off} and topoisomerase rates k_{topo} over two and five orders of magnitude, respectively. Bar plots represent the mean value (crosses), whiskers denote the span of the data. The red line is a guide to the eye and was constructed by applying a Savitzky-Golay filter of window length 5 and order 1. Data are from 50 independent simulations. (**c**) As in (b) except for EQ/EQ cohesin.

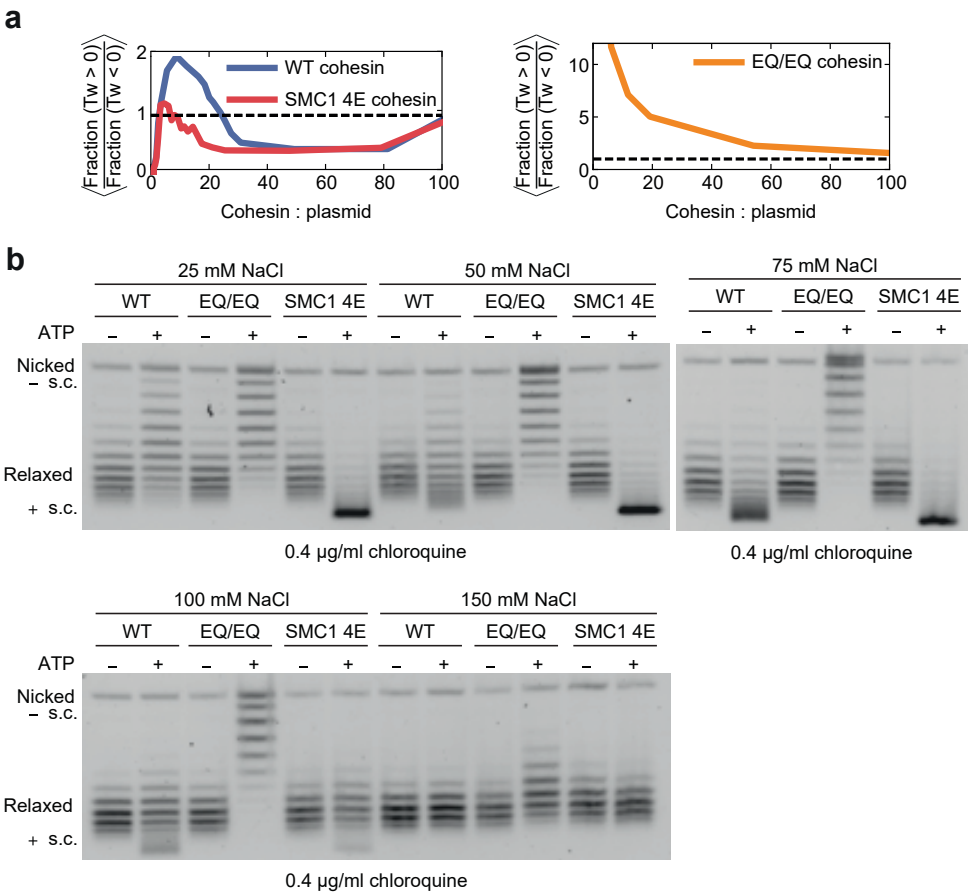


Figure S6.4 | Reason for the switch from positively to negatively supercoiled reaction products in simulations and generation of positively supercoiled plasmids by wild type cohesin at intermediate NaCl concentrations.

(**a**) At every time point, the fraction of the plasmid with a positive twist (Fraction (Tw > 0)) is divided by the fraction of the plasmid with a negative twist (Fraction (Tw < 0)). The quotient is then averaged over time and subsequently over 200 independent simulations and plotted over the number of cohesin complexes per plasmid. For wild type cohesin, the majority of the plasmid is positively supercoiled for cohesin:plasmid ratios up to ~30, while Smc1 4E cohesin barely ever produces a predominantly positively supercoiled plasmid. In contrast, EQ/EQ cohesin can only make a single step which yields that the majority of the plasmid is unlooped, carrying the positive compensatory twist from the single steps of the individual EQ/EQ cohesin complexes. (**b**) Negatively supercoiled plasmid DNA was relaxed by incubating with human topoisomerase I for 45 minutes at the indicated NaCl concentrations. The indicated forms of cohesin were added at a 15:1 cohesin:plasmid ratio along with a twofold excess of NIPBL-Mau2 relative to cohesin in the presence and absence of ATP and incubations were continued for a further 90 minutes. Reactions were then terminated and purified DNA was separated by agarose gel electrophoresis in the presence of 0.4 µg/ml chloroquine.

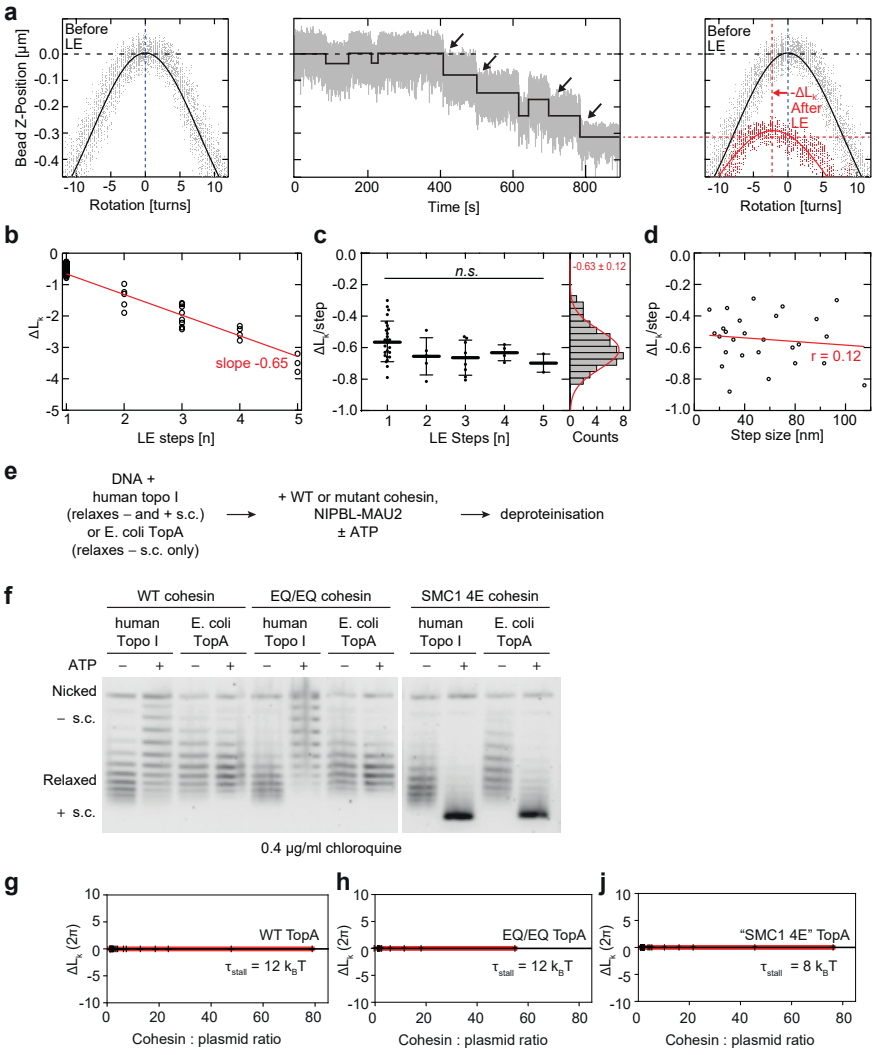


Figure S6.5 | DNA twist induced by individual DNA loop extrusion steps of the Smc1 4E mutant and *E. coli* TopA experiments and simulations.

(**a**) (Left) DNA end-to-end extension as a function of magnet rotation for a torsionally constrained 3.6 kbp DNA molecule. The DNA extension at a constant 0.3 pN force is maximal when no external rotations are applied. Upon applying positive/negative rotations, the over/underwinding of the DNA changes the linking number L_k and leads to the formation of supercoils, which are at this force symmetric for both positive and negative coiling. The solid black line depicts a Gaussian fit to the rotation curve data. (Center) Representative trajectory of Smc1 4E cohesin (20 pM Smc1 4E, 50 pM NIPBL-Mau2) showing step-wise DNA loop extrusion in the presence of 1 mM ATP at 0.3 pN. The black line depicts the fit from the step-finding algorithm. *Caption continues on next page.*

Figure S6.5 | (Right) DNA extension as a function of magnet rotation similar to (Left), conducted directly after the DNA loop extrusion experiment. The rotation curve after loop extrusion (red; solid line depicts Gaussian fit to the data) shows that the maximum DNA extension was shifted to negative magnet rotation compared to the initial rotation curve (black), caused by the uncoiling of the positive, complementary supercoils ($+\Delta L_k$) formed in the DNA molecule outside the DNA loop. The degree of negative supercoils ($-\Delta L_k$) generated by Smc1 4E residing in the loop (red arrow) is equal to the degree of positive supercoils ($+\Delta L_k$) generated outside the loop. **(b)** Change in DNA linking number, ΔL_k, over the number of steps for Smc1 4E in the presence of ATP (total N= 54). The red line represents a linear fit without offset of the change in linking number over the number of LE steps. **(c)** The same data as in (b) but divided by the number of LE steps. The histogram on the right represents all data points, fitted by a Gaussian with -0.63 ± 0.12 turns (mean ± SD). Statistical significance was assessed using ANOVA with a significance level $\alpha = 0.05$ (95% confidence interval; n.s.: p >0.05). **(d)** Change in rotation per step is plotted against the step size for traces with only a single step (N = 24). Linear fit and Pearson's correlation coefficient are shown in red. **(e)** Experiment outline. **(f)** Negatively supercoiled plasmid DNA was relaxed by incubating with human topoisomerase I or *E. coli* TopA for 45 minutes. The indicated forms of cohesin were added at a 15:1 cohesin:plasmid ratio along with a two-fold excess of NIPBL-Mau2 relative to cohesin in the presence or absence of ATP. Incubations were continued for a further 90 minutes. Reactions were then terminated and purified DNA was separated by agarose gel electrophoresis in the presence of 0.4 µg/ml chloroquine. **(g-i)** As in Figure 6.3d and Figure 6.4b,e except *E. coli* TopA was simulated instead of human topo I by allowing that TopA can only resolve negative supercoils when the supercoiling density is $\sigma < 0.065$. Note that the simulated outcome (no supercoiling) does not match the experimental outcome (positive supercoiling).

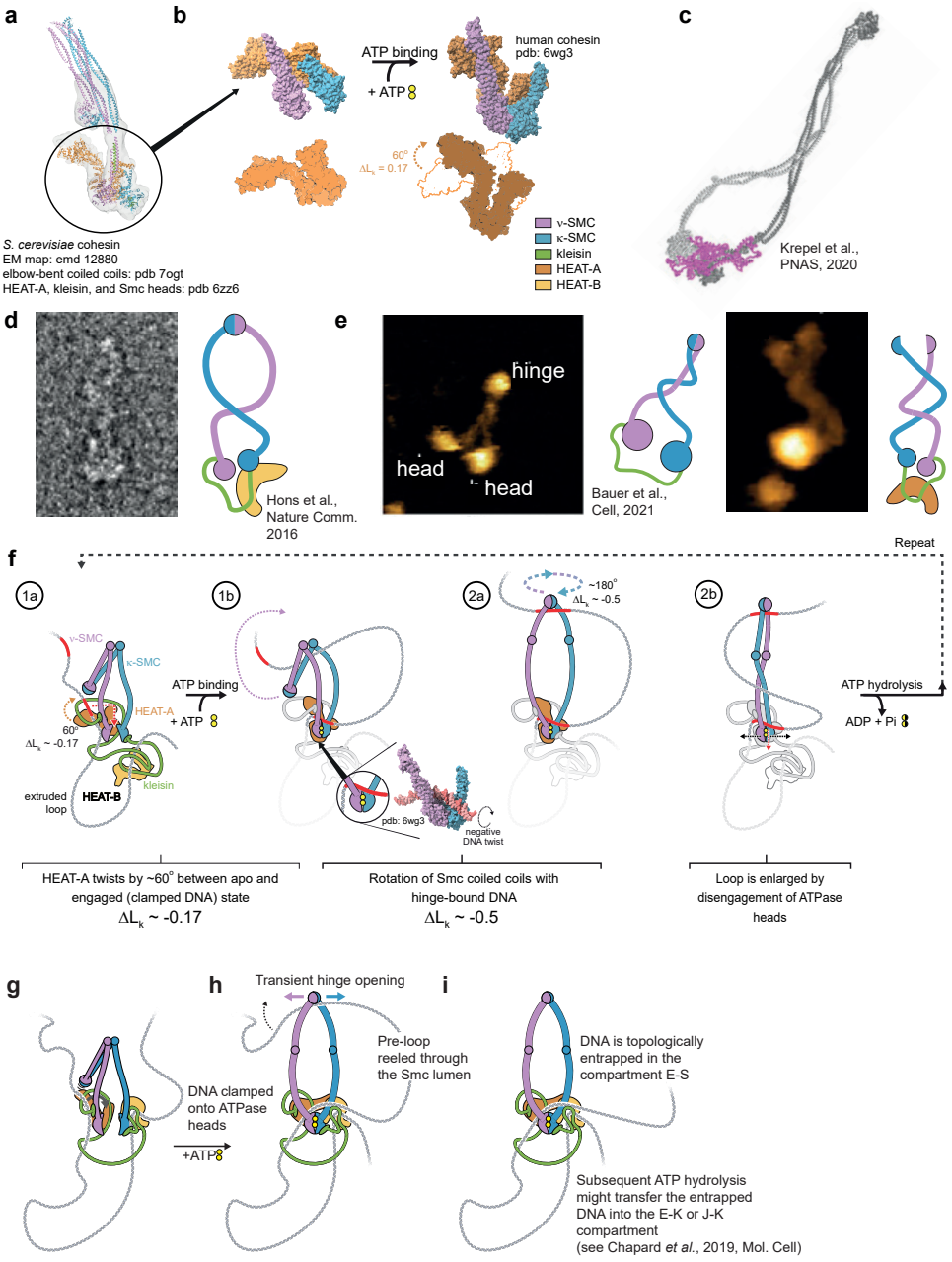


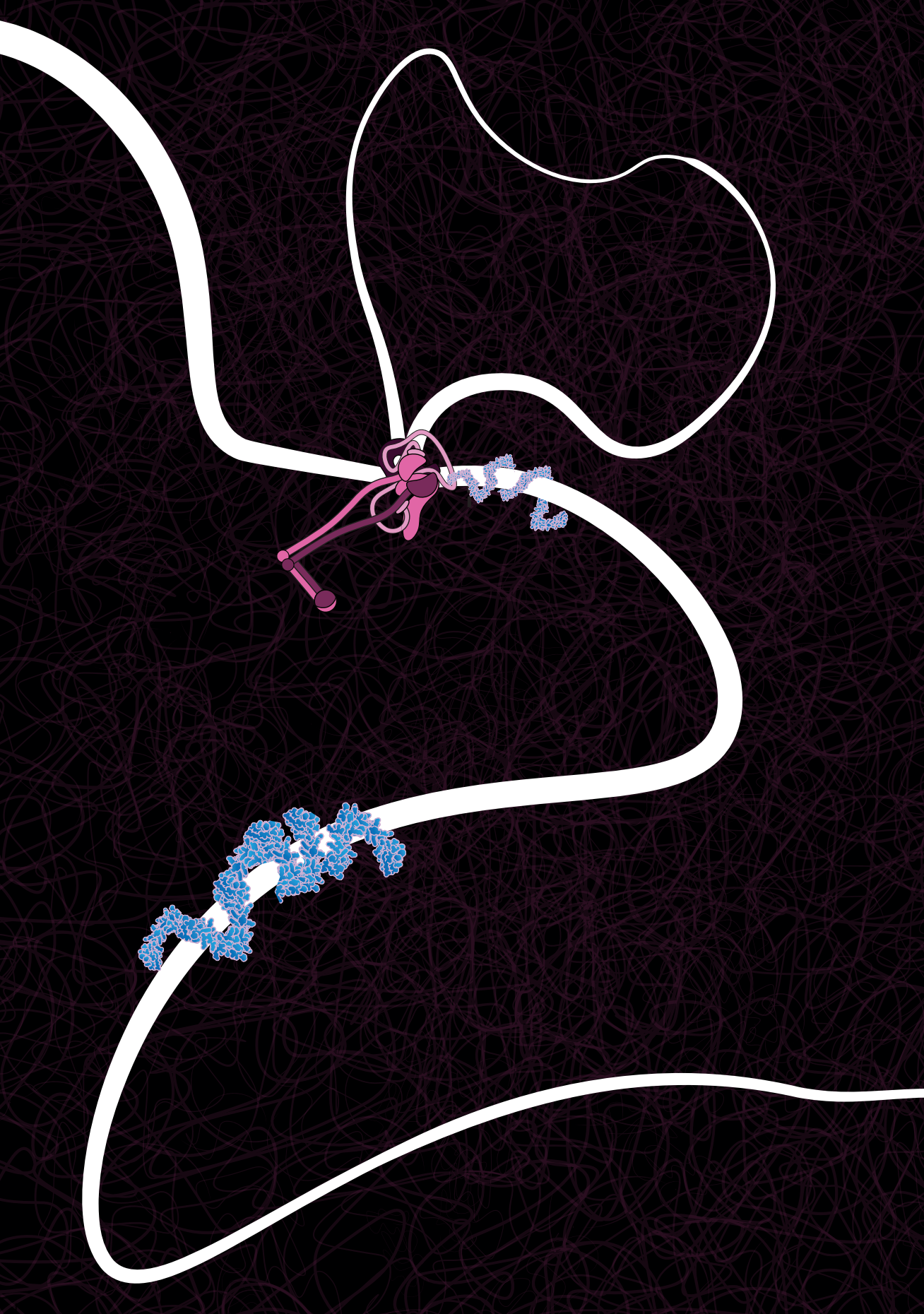
Figure S6.6 | Structural basis for twists generation upon ATP binding, potential DNA loop extrusion mechanism including twist generation, and potential pathway to topological entrapment upon ATP binding.

(**a**) Superimposition of a density map from cryo-EM data for *S. cerevisiae* cohesin [365] (EMD-12880) and a pseudo-atomic model in the apo state. (**b**) Rotation of HEAT-A between the pseudo-atomic model in the apo state shown in (a) and the ATP-bound head-engaged gripping state (pdb entry 6wg3; ref. [200]). Apo and engaged structures were aligned on the Smc3 ATPase head. HEAT-A rotates by ~ -60° ($\Delta L_k \sim -0.17$). (**c**) Conformation of human cohesin as observed by negative staining EM showing intertwined coiled coils, from ref. [349]. *Caption continues on next page.*

Figure S6.6 | (**d**) Snapshots of human cohesin imaged by high-speed AFM anecdotally showing intertwined coiled coils, adapted from ref. [238]. (**e**) Reel & Seal model [247] with a twist, incorporating an ATP binding-induced negative DNA twist into the loop. DNA is held by kleisin and HEAT-A (step 1a). Clamping of DNA on top of the engaged ATPase heads forces a small loop pseudo-topologically into the Smc lumen (step 1b). The inset to step 1b shows the twisting of the DNA which yields a negatively supercoiled loop (based on pdb entry 6wg3 (ref. [200])). Rotation of HEAT-A (if bound to DNA in the apo as well as the engaged state) induces a negative twist (see panels (a-b)). Spontaneous rotation of the coiled coils with hinge-bound DNA by ~180° (ref. [238, 239, 349]) induces a further linking number change of $\Delta L_k \sim -0.5$ into the loop (step 2a-b). ATP hydrolysis causes disengagement of the heads, merging pre-loop and already extruded loop *via* HEAT-A-kleisin latch opening. (**f**) A potential pathway to topological entrapment upon ATP binding. DNA is held at the *ν*-SMC and by accessory subunit I. (**g**) Upon ATP binding and subsequent head engagement, DNA binding on top of the ATPase heads inserts a short loop of DNA (of the length of one SMC-mediated loop extrusion step) within the loop formed by the Smc arms and the engaged Smc heads (E-S compartment). Hinge opening in this state expels the head-bound DNA segment from the Smc lumen, leading to a topological entrapment within the E-S compartment. (**h**) A single duplex of DNA is topologically embraced in the E-S compartment. Subsequent ATP hydrolysis causes the heads to disengage. The DNA is then in the S-K compartment and may be transferred to the E-K compartment in the subsequent ATP binding and head engagement. The sketched pathway is consistent with the fact that the hinge is the main DNA entry gate for topological entrapment [224, 225, 291, 374, 375] and that a cohesin hinge mutant can perform loop extrusion but is defective in cohesion establishment [139]. Furthermore, it is in line with the fact that ATP binding, and thus the head engagement, is sufficient to allow topological entrapment [234, 290, 376, 377] and that clamping of the DNA onto the engaged ATPase heads by HEAT-A is required for topological entrapment in order to insert a transient loop into the E-S compartment. Subsequent ATP hydrolysis may transfer the entrapped DNA into the S-K compartment. Note that for SMC5/6 the neck gate (the *ν*-SMC-kleisin interface) appears to be the major or only DNA entry gate [309], while DNA entry through the neck gate, as well as the hinge, can be observed for cohesin and condensin [223–225, 235, 291, 374, 375, 378]. Topological entry of DNA through the hinge is thus not the only possible entrapment mechanism upon ATP binding since an ATP binding-induced power stroke has also been proposed to enable topological DNA loading through the neck gate [309].

III

Halting DNA loop extruders



7

CTCF is a DNA-tension-dependent barrier to cohesin-mediated loop extrusion

This chapter has been published as: Iain Davidson*, Roman Barth*, Maciej Zaczek*, Jaco van der Torre, Wen Tang, Kota Nagasaka, Richard Janissen, Jacob Kerssemakers, Gordana Wutz, Cees Dekker, Jan-Michael Peters: CTCF is a DNA-tension-dependent barrier to cohesin-mediated loop extrusion, Nature. 2023 Apr;616(7958):822-827 [35].

Abstract

In eukaryotes, genomic DNA is extruded into loops by cohesin [282]. By restraining this process, the DNA-binding protein CTCF generates topologically associating domains (TADs) [33, 34] that play key roles in gene regulation and recombination during development and disease [282, 379–382]. How CTCF establishes TAD boundaries and to what extent these are permeable to cohesin is unknown [304]. To address these questions, we visualize interactions of single CTCF and cohesin molecules on DNA in vitro. We show that CTCF is sufficient to block diffusing cohesin, possibly reflecting how cohesive cohesin accumulates at TAD boundaries, as well as to block loop-extruding cohesin, reflecting how CTCF establishes TAD boundaries. CTCF functions asymmetrically, as predicted, but unexpectedly is dependent on DNA tension. Moreover, CTCF regulates cohesin's loop extrusion activity by changing its direction and by inducing loop shrinkage. Our data indicate that CTCF is not, as previously assumed, simply a barrier to cohesin-mediated loop extrusion but is an active regulator of this process, where the permeability of TAD boundaries can be modulated by DNA tension. These results reveal mechanistic principles of how CTCF controls loop extrusion and genome architecture.

7.1 Main

The folding of genomic DNA by cohesin has important roles in chromatin organization, gene regulation, and recombination [282]. Cohesin belongs to the 'structural maintenance of chromosomes' (SMC) family of ATPase complexes that can extrude DNA into loops, an activity which has been reconstituted *in vitro* for cohesin, condensin, and SMC5/6 [98, 173, 204, 212, 214, 216]. Cohesin also performs a second function by mediating sister chromatid cohesion. In individual cells, loops are located at variable positions, suggesting that loops are dynamic structures most of which are in the process of being extruded [40, 159, 160]. However, in cell-population measurements patterns emerge which reveal that most loops are formed within topologically associating domains (TADs) [39–41]. CTCF (CCCTC binding factor) is located at TAD boundaries [40, 41] and is required for their formation and for cohesin accumulation at these sites [32–34]. CTCF has unstructured N- and C- terminal regions that flank eleven zinc fingers, several of which recognize an asymmetric DNA sequence and thus position CTCF directionally on DNA [37, 38]. Most CTCF binding sites are oriented in convergent orientations so that CTCF's N-termini face the interior of TADs, suggesting that CTCF acts as an asymmetric boundary to cohesin-mediated loop extrusion [42, 97, 156]. Consistent with this possibility, CTCF's N-terminus can bind cohesin [141] and is required for TAD insulation and loop anchoring at these sites [141, 251, 252, 383]. Several mechanisms have been suggested for how CTCF might prevent loop extrusion across TAD boundaries (reviewed in [304]), namely, as a physical barrier ('roadblock'), by binding cohesin, by preventing release of cohesin from DNA, by promoting replacement of cohesin's ATPase activating subunit NIPBL by its inactive counterpart PDS5, by directly inhibiting cohesin's ATPase activity, and by promoting entrapment of DNA inside a ring structure that is formed by three of cohesin's subunits [140]. It has also been proposed that CTCF converts cohesin into an asymmetrically extruding enzyme by stalling loop extrusion on the CTCF-bound site while allowing cohesin to continue reeling DNA into the loop only from the TAD interior [141, 221, 276]. However, it remains unresolved which of these proposed mechanisms is used by CTCF and whether CTCF is sufficient for blocking loop extrusion by cohesin. Answering these questions is of great importance since CTCF is required for controlling enhancer-promoter interactions [282], cell differentiation, nuclear reprogramming [381], recombination of antigen receptor genes [379, 380], and timing of DNA replication [190], and because CTCF mutations have been implicated in tumorigenesis [382]. CTCF boundaries are also sites at which replicated DNA molecules are connected by cohesin complexes which mediate cohesion [384].

7.1.1 CTCF characterisation *in vitro*

To obtain insight into how CTCF controls cohesin, we developed *in vitro* assays in which CTCF-cohesin interactions can be visualized on DNA at the single-molecule level in real time. We first analyzed how CTCF finds its DNA consensus sequence. Consistent with previous reports [385, 386], recombinant human CTCF (Figure 7.1a) bound specifically to DNA oligonucleotides containing a single CTCF binding site in electrophoretic mobility shift assays (EMSAs) in a manner that was reduced by DNA methylation (Figure 7.1b). We introduced this CTCF binding site into linear 26.1 kb DNA molecules, tethered these at both ends to glass surfaces in flow cells, stained with Sytox Green, and imaged the DNA molecules using Highly Inclined and Laminated Optical sheet (HILO) microscopy.

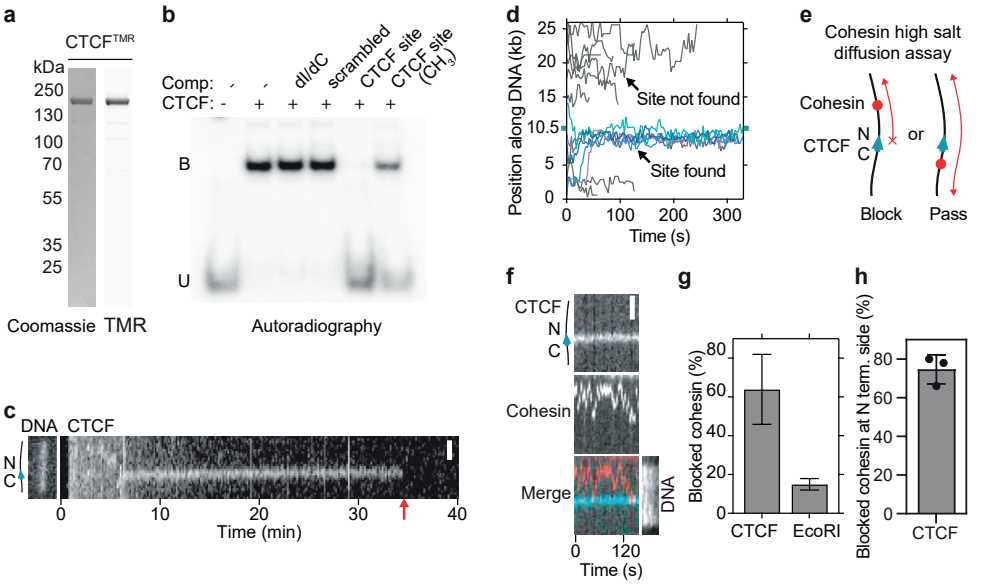


Figure 7.1 | CTCF is a directional barrier to cohesin diffusion on DNA.
(**a**) Coomassie staining of recombinant CTCF after SDS-PAGE. Tetramethylrhodamine (TMR) was visualized by epi-green excitation. (**b**) Autoradiograph of electrophoretic mobility shift assay (EMSA). CTCF was incubated with a ^{32}P-labelled DNA containing a CTCF binding site. Where indicated, reactions were supplemented with excess unlabelled competitors. dI-dC: poly(2'-deoxyinosinic-2'-deoxycytidylic acid). B: bound, U: unbound. (**c**) Example of TMR-labelled CTCF diffusing on DNA. Non-specifically bound CTCF molecules diffuse randomly and dissociate rapidly. At 5.5 min, a CTCF molecule binds DNA and diffuses until encountering the CTCF binding site at 6 min. Scale bar, 2 µm. The red arrow indicates the timepoint when CTCF bleached or dissociated. (**d**) Superposition of individual TMR-labelled CTCF-diffusion events. Events in which CTCF localized to its binding site at position 10,452 bp (cyan tick) are shown in blue (N = 6). DNA-binding events in which CTCF failed to localize to its binding site are shown in grey (N = 11). (**e**) Illustration of cohesin diffusion assay. (**f**) Example of cohesin diffusion that is blocked by CTCF. Cohesin and CTCF were labelled with Alexa660 (red) and TMR (blue), respectively. Sytox Green DNA stain was introduced into the flow cell at the end of the experiment. Scale bar, 2 µm. (**g**) Fraction of blocking events when cohesin encountered CTCF or EcoRI^{E111Q} (mean ± SD from 7 (N=264) and 3 (N=106) independent experiments, respectively). (**h**) Fraction of blocked events where cohesin diffused on the DNA between the tether point and the N-terminal side of CTCF (mean ± SD from 3 (N=48) independent experiments). In the remaining 25 % of events, cohesin diffused between the tether and the C-terminal side of CTCF. Sample sizes refer to biological replicates.

Following the introduction of fluorophore-labeled CTCF, both immobile and mobile CTCF foci were observed at various positions along the DNA (Figure 7.1c,d, Figure S7.1a). CTCF foci at the CTCF binding site were detectable for much longer than those elsewhere, where CTCF proteins often dissociated rapidly unless they arrived at the CTCF binding site while diffusing along DNA (diffusion coefficient: 0.32 ± 0.1 kb^2/s; Figure 7.1c,d, Figure S7.1b). These results indicate that CTCF finds its DNA binding site by facilitated diffusion. Most CTCF foci that were not located at the CTCF binding site were removed by a brief salt wash, unlike those at the CTCF binding site (Figure S7.1c). Fluorescence intensity and photobleaching analysis indicated that these remaining CTCF molecules were monomers (Figure S7.1d, e). Once bound to their binding sites, the mean residence time of CTCF

molecules was ~29 min (Figure S7.1f,g), which is longer than most [387–392], but not all [390], *in vivo* estimates, and longer than another recent *in vitro* measurement [393]. It is possible that additional factors, such as the action of other chromatin-bound proteins, might promote CTCF unbinding in cells.

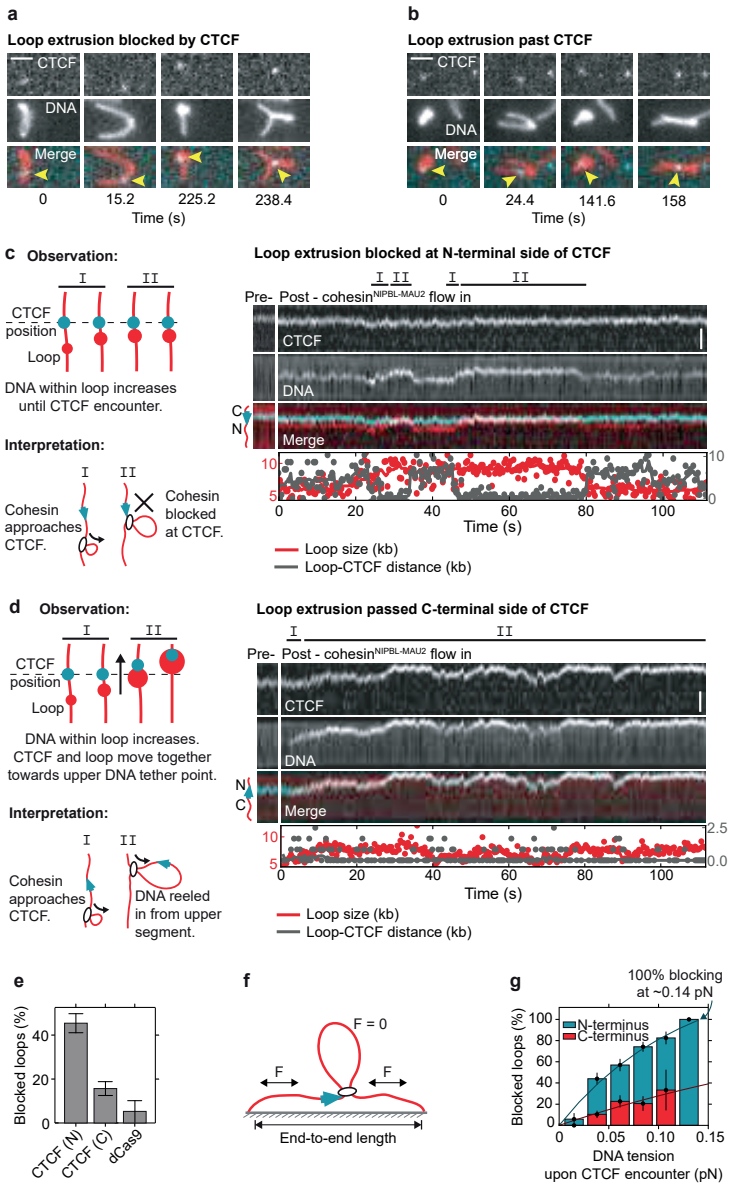


Figure 7.2 | CTCF is a direction- and tension-dependent barrier to cohesin-mediated DNA loop extrusion. (**a-b**) Examples of loop extrusion blocked by (a) or passing (b) CTCF (cyan) labeled with Janelia Fluor 646 (JF646). DNA loops (red) were visualized by Sytox Orange and perpendicular buffer flow. Scale bar, 2 µm. (**c**) Cohesin-mediated DNA loop extrusion (LE) encountering N-terminally oriented JF646-labeled CTCF (cyan). Growth of DNA loop stops upon encountering CTCF at ~28-33 s, and ~50-80 s. Scale bar, 2 µm. (**d**) Same as (c), but for a passing event. CTCF passes into the loop at 70 s, and translocates with it. Scale bar, 2 µm. (**e**) Fraction of LE events blocked upon encountering N- or C-terminally oriented CTCF or dCas9 (mean ± 95% binomial confidence interval). N = 119, 115, 19 from 13, 3, and 3 independent experiments for N-terminal, C-terminal, and dCas9 encounters, respectively. The force range between 0.04-0.08 pN was best covered and therefore chosen to compare the overall blocking efficiency (Figure S7.5c,d). (**f**) The DNA tension at the moment of encounter was calculated by the amount of DNA outside the loop and the DNA end-to-end length (see Methods section 7.3). *Caption continues on next page.*

Figure S7.0 | (g) The LE blocking probability of N- or C-terminally oriented CTCF depends on DNA tension (mean ± 95% binomial confidence interval). Solid lines are fits of the form $1-exp(-F/F0)$, which were used to compute the force at which 100% blocking is achieved (N-terminal encounters: $P_{block}(F) = 147(1-exp(-F/0.125\ pN))$; C-terminal encounters: $P_{block}(F) = 115(1-exp(-F/0.357\ pN))$). N per bin for N-terminal (n) and C-terminal (c) encounters: 0-0.015 pN: 17 (n), 12 (c); 0.015-0.026 pN: 75 (n), 77 (c); 0.026-0.05 pN: 72 (n), 53 (c); 0.05-0.072 pN: 89 (n), 34 (c); 0.096-0.119 pN: 40 (n), 6 (c); 0.119-0.142 pN: 3 (n), 0 (c). The bin for C-terminal encounters at the highest DNA tension regime is not shown due to insufficient observations (N < 3). Sample sizes refer to biological replicates from 13 independent experiments for N-terminal and 3 independent experiments for C-terminal encounters.

7.1.2 CTCF is a polar barrier to cohesin

Next, we analyzed how CTCF interacts with cohesin that diffuses along DNA. For this purpose, we used an assay in which cohesin associates with DNA in a high-salt resistant manner that is sensitive to cohesin and DNA cleavage [211], suggesting that under these conditions cohesin entraps DNA topologically and moves along DNA as has been proposed for cohesive cohesin [132, 384]. We observed that CTCF frequently blocked diffusion of recombinant human cohesin (64 ± 18 %; mean and SD), while the remaining cohesin traversed CTCF multiple times (Figure 7.1e-g, Figure S7.1h-j, Figure S7.2a-d). In contrast, $EcoRI^{E111Q}$ rarely blocked cohesin (15 ± 3 %; Figure 7.1g, Figure S7.2e). To determine the orientation of the CTCF molecules that had blocked cohesin translocation, we post-labeled the DNA molecules with a marker protein that binds to one of their ends (Figure S7.2f). This revealed that 75 ± 8 % (mean and SD) of the blocked cohesin complexes faced CTCF's N terminal side (Figure S7.1j). This can be attributed to the orientation of CTCF since inversion of its binding site reversed this blocking behavior (Figure S7.2g,h). Since diffusing cohesin binds to DNA in a manner consistent with entrapment [211], which is believed to be the interaction mode by which cohesin mediates cohesion [132], this suggests that CTCF contributes to the accumulation of cohesive cohesin at TAD boundaries [384].

7.1.3 CTCF is a polar barrier to DNA looping

To test whether CTCF also acts as a barrier to loop-extruding cohesin, we introduced a single CTCF site at position 9.7 kb in a 31.8 kb DNA, such that CTCF's N-terminus would face the longer end of the DNA. We tethered both ends of these molecules to the surfaces of flow cells and stained them with Sytox Orange. We then bound CTCF purified from HeLa cells (Figure S7.3a-e) to these DNA molecules and introduced HeLa cohesin (Figure S7.3f), recombinant NIPBL-Mau2 (Figure S7.1i), and ATP. Upon buffer flow perpendicular to the DNA axis, CTCF could be detected either near the base of (Figure 7.1.1a) or within DNA loops (Figure 7.1.1b), suggesting that it functioned as a barrier to loop extrusion in some but not all cases. To analyze this behavior quantitatively, we monitored loop extrusion in the absence of buffer flow, where loop formation results in the appearance of a bright spot on the DNA that increases in intensity over time. Tracking and quantification of loop position and size as well as of CTCF position permitted the classification of encounters of cohesin-mediated DNA loops with CTCF (Figure 7.1.1c,d; for additional examples see Figure S7.4a,b). These experiments revealed that N-terminally oriented CTCF blocked the progression of loop extrusion in 45 ± 9 % (mean ± SD) of encounters (Figure 7.1.1c,e, Figure S7.4a,b), whereas the blocking efficiency was reduced to 16 ± 7 % (mean ± SD)

when we used DNA molecules in which the orientation of the CTCF binding site had been inverted and on which cohesin therefore encountered CTCF's C-terminus (Figure 7.1.1d,e). In contrast, the control protein dCas9, which has a larger mass (180,000 Da) than CTCF-Halo-Flag (118,600 Da) blocked loop extrusion in only 5 ± 10 % (mean ± SD) of encounters (Figure 7.1.1e), consistent with the finding that cohesin can readily traverse non-interacting DNA-bound particles during loop extrusion [226]. These results indicate that monomeric CTCF, despite its relatively small mass and Stokes radius (5 nm for the N terminus of CTCF [394]), is sufficient to block loop extrusion by cohesin in a directional manner, possibly because the N-terminus of CTCF can bind cohesin [141]. Interestingly, the N- and C-terminal blocking frequencies of 45% and 16% observed in our experiments can remarkably well explain *in vivo* estimates of how frequently loops are detected between CTCF sites oriented in a convergent, divergent, or tandem manner (Figure S7.5a), suggesting that CTCF may be solely responsible for determining how frequently loops are anchored at these differently oriented sites. While performing these experiments, we also observed that loops occasionally translocated along the DNA without increasing in size (for examples, see Figure 7.1.1c,d), a behavior that is reminiscent of forms of condensin that are defective in the DNA binding site formed at the interface between the HAWK subunit Ycg1 and the kleisin Brn1 [98, 221].

7.1.4 CTCF is a DNA-tension-dependent barrier

Notably, we observed that the CTCF blocking efficiency for loop-extruding cohesin depends on the tension in the DNA that is reeled in. Since DNA molecules are tethered at both ends in our assay, loop extrusion continuously shortens the non-extruded parts of the DNA molecules and thus increases their tension, until this tension exceeds the stalling force of loop extrusion [98] (Figure 7.1.1f). We noticed that larger loops and loops extruded from DNAs with a longer end-to-end length tended to be stalled more efficiently by CTCF than those formed from less stretched DNA. Because both scenarios coincide with larger tension in the unextruded part of the DNA, we tethered DNA molecules to the surface of flow cells with various degrees of 'slack', performed CTCF loop extrusion blocking assays and calculated the tension that DNA molecules experienced when cohesin encountered CTCF. Interestingly, the efficiency of CTCF's barrier activity indeed very strongly correlated with increased DNA tension (Figure 7.1.1g). Strikingly, our data indicate that CTCF does not block loop extrusion by cohesin at all when no force is applied, while CTCF blocks loop extrusion increasingly when tension is applied to the DNA, with CTCF reaching a blocking efficiency of 100 % at ~0.14 pN. Notably, this tension is close to ~0.15 pN, the median value of the force required to stall loop extrusion itself (Figure S7.5b-d). Encounters from the C-terminal side showed a similar trend, i.e. blocked loop extrusion more frequently at higher tension, but with much lower blocking frequencies. In contrast, the ratio of blocking efficiencies of N-terminal versus C-terminal encounters (3.6 ± 0.8 -fold [mean ± SD]) was unaffected by DNA tension (Figure S7.5e). Blocking at high DNA tensions was not due to the stalling force alone, since only N-terminal but not C-terminal encounters displayed complete blocking at 0.14 pN (Figure 7.1.1g). Furthermore, in the absence of CTCF, 53 ± 16 % of DNA loops continued to grow at DNA tensions above 0.14 pN (Figure S7.5d), whereas encounters with N terminally oriented CTCF displayed complete blocking at this DNA tension (Figure 7.1.1g). We also tested other parameters that might induce the

blocking of loop extrusion by the N-terminally oriented CTCF. The time elapsed between the initiation of loop extrusion initiation and encounter with CTCF, and the loop size at the time of encounter were not significantly different between blocking and passing events (Figure S7.6a-d). Blocking events were however more frequently observed on DNA with a larger end-to-end length (Figure S7.6e,f), which can be attributed to these DNA molecules experiencing a larger DNA tension even in the absence of an extruded loop (i.e., an 'offset tension' of e.g. ~ 0.07 pN at 4 μm end-to-end length; Figure S7.6h-i), and therefore the blocking force at the encounter with CTCF is more readily reached upon loop extrusion. To test whether the blocking of cohesin-driven loop extrusion by N-terminally oriented CTCF could relate to the ability of loop extruding cohesin to 'step over' CTCF, we measured cohesin's step size during loop extrusion using magnetic tweezers. These measurements showed that cohesin, on average, takes large steps of ~40 nm (100-200 bp) on DNA and that the step size decreases when DNA tension increases (Figure S7.7a-d), as also observed for condensin [217]. We tested in simulations whether cohesin might encounter CTCF more frequently at higher DNA tension because cohesin is less likely to step over CTCF since the extrusion steps become smaller. However, this hypothesis was not supported by our simulations (Figure S7.7e,f). We therefore suspect that DNA tension increases the blocking efficiency of CTCF by other mechanisms, such as reducing the step frequency at increased tension which allows more time for CTCF-cohesin binding, or decreasing thermal fluctuations of DNA [296] which may reduce the space that CTCF has to explore to find cohesin, or that cohesin's weak motor activity can more easily overcome the low binding affinity of CTCF-cohesin interactions [141] at low DNA tension than at high tension (discussed in the legend of Figure S7.7f). Irrespective of these interpretations, our results indicate that local changes in DNA tension that could be caused by nucleosome assembly, transcription, DNA replication, supercoiling or other processes, can affect genome architecture by modulating the permeability of TAD boundaries. Since loop extrusion is sensitive to DNA tension [98] but diffusion is not, we hypothesized that the DNA tension dependence of CTCF's barrier activity might only occur upon encountering loop-extruding cohesin but not for diffusing cohesin. Indeed, we found CTCF's ability to block diffusing cohesin to be independent of DNA tension (Figure S7.5f-h). Thus, while CTCF acts as a barrier to diffusing cohesin, it can only block loop extruding cohesin at higher DNA tensions.

7.1.5 Transient loop anchoring by CTCF

To analyze the fate of loops that were blocked by CTCF, we first analyzed for how long CTCF and loops co-localize under conditions in which the loop size was constant (i.e., where loop extrusion stalled upon encounter). We frequently observed brief (tens of seconds) and repeated encounters between loops and CTCF (Figure S7.8a) as well as occasional encounters that lasted for several minutes (Figure S7.8b). The distribution of CTCF-loop interaction times following stalling events was well-described by a bi-exponential distribution, indicating the existence of two populations with mean CTCF-loop association times of 16 s and 167 s (Figure 7.3d, Figure S7.8c-m). Unlike CTCF's blocking function, the CTCF-loop association time was largely unaffected by CTCF orientation (Figure S7.8c-m). It is conceivable that the infrequent C-terminal blocking events that we observed represent occasions where cohesin in fact encountered CTCF's N-terminus after passing over its

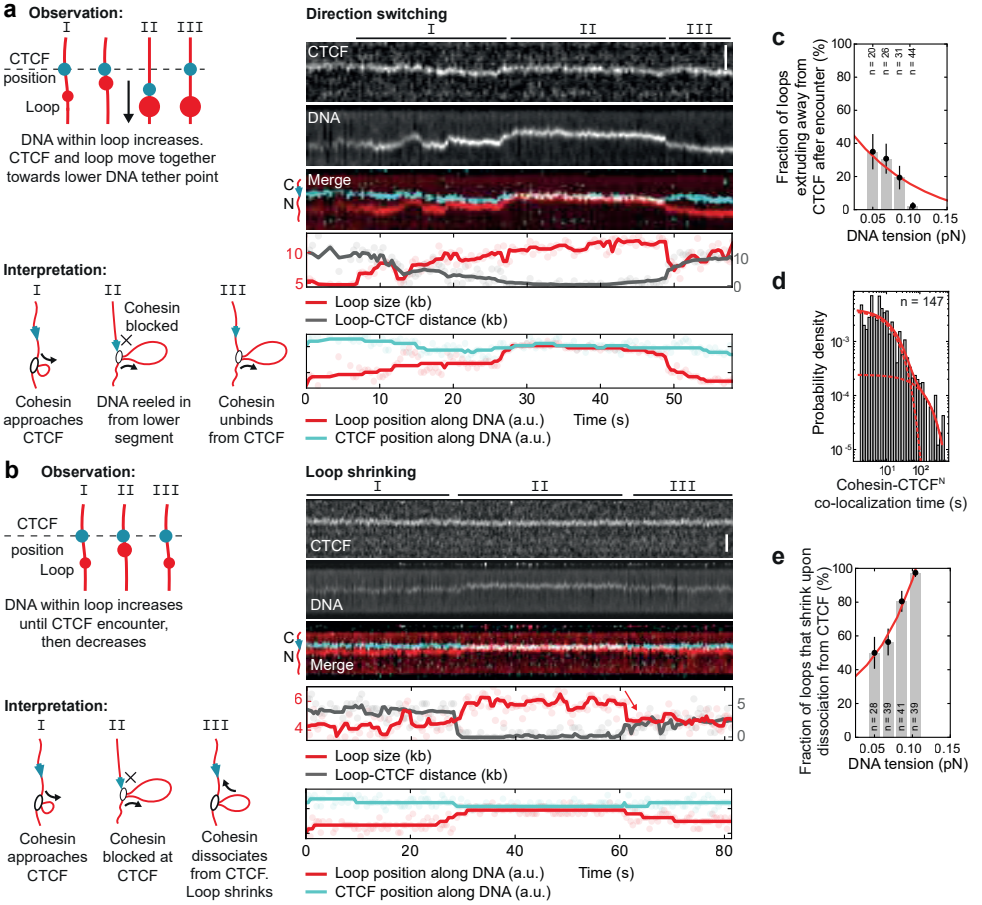


Figure 7.3 | CTCF changes the direction of cohesin-mediated loop extrusion or induces loop shrinkage, depending on the DNA tension.

(**a-b**) (Left panels) observation and interpretation illustrations of (right panels) kymographs of cohesin-mediated DNA loop extrusion encountering N-terminally oriented, JF646-labeled CTCF (cyan). DNA loops (red) were visualized by Sytox Orange. Scale bars, 2 μm. In panel (a), the growing loop encounters CTCF at 28 s. CTCF and the growing DNA loop move towards the lower DNA tether point, indicating extrusion on the side facing away from CTCF. In panel (b), growth of DNA loop stops upon encountering CTCF at ~29 s. The DNA loop shrinks following dissociation from CTCF at ~60 s. (**c**) Fraction of loops extruding away from CTCF, versus the DNA tension at the moment of encounter (mean ± 95% binomial confidence interval; 13 independent experiments). (**d**) Co-localization time of encounters between cohesin and CTCF's N-terminus. The fit denotes a 2-exponential distribution with rate constants $k_1 = 0.06\ s^{-1}$ and $k_2 = 0.006 s^{-1}$ ($\tau_1 \sim 17\ s$ and $\tau_2 \sim 167\ s$; data from 13 independent experiments). Dashed lines represent the individual components of the 2-component exponential distribution; solid line represents the final 2-component exponential distribution. (**e**) Fraction of loops which shrink upon release from CTCF versus DNA tension at the moment of encounter (mean ± 95% binomial confidence interval; 13 independent experiments). Sample sizes refer to biological replicates.

C-terminus. The results indicate that CTCF interacts with cohesin mostly transiently (> 85% of encounters lasted less than 3 minutes; Figure S7.8g), which is similar to the lifetime

that has been measured for particular loops in cells [160]. However, longer-lived loops have been predicted to exist for up to several hours [276, 395]. Since we did not observe such prolonged co-localization of CTCF and loops, additional proteins may be required to anchor loops for such long time periods, for example the PDS5 proteins, which are also required for TAD boundaries in cells [34].

7.1.6 CTCF can switch the direction of looping

It has been speculated that cohesin switches from symmetric to one-sided asymmetric extrusion at TAD boundaries at which loop 'stripes' or 'flames' have been detected in Hi-C experiments [42, 141, 221, 276]. We therefore analyzed whether a change in extrusion symmetry could be observed when cohesin encounters CTCF. However, close analysis revealed that, unlike previously assumed [173, 212, 214], cohesin rarely extrudes DNA strictly symmetrically, but instead frequently reels in DNA first from one side and then the other, switching direction multiple times (for a detailed analysis of this two-sided extrusion see chapter 5). We therefore analyzed whether CTCF can trigger a switch of the direction of loop extrusion. To investigate this, we monitored the size of DNA loops and their position relative to CTCF, following encounters that had blocked loop extrusion. At low DNA tension, we observed events where CTCF indeed switched the direction of cohesin's loop extrusion activity. Cohesin approached CTCF by reeling in the intervening DNA and then, upon encounter with CTCF, it began to reel in DNA from the other direction while remaining bound to CTCF (Figure 7.3a,c). A control experiment with gold nanoparticles that were tethered to DNA as artificial roadblocks [226], reversed the direction of loop extrusion 2.6 times less frequently at low DNA tension (Figure S7.9a), suggesting that this ability may be a specific property of CTCF. This effect can potentially explain the appearance of 'stripes' and 'flames' at TAD boundaries. Interestingly, at higher DNA tension, CTCF did not switch the direction of loop extrusion (Figure 7.3c), but instead loops tended to shrink in size upon release from CTCF (Figure 7.3b,e, Figure S7.9c-l). In most cases, loops decreased in size within a single step (i.e. within the imaging frame speed of 0.4 s, Figure S7.9c, d, g, h), but in some cases loops shrunk gradually over several seconds at a rate similar to that of loop extrusion (Figure S7.9e-j). In both cases, loops did not disrupt completely but were reduced in size by several kb and on average lost 35 % of looped DNA (Figure S7.9k,l). Such loop shrinkage could be observed with similar frequencies when cohesin collided with artificial roadblocks on DNA (Figure S7.9b), suggesting that this may be a general response of cohesin to encountering barriers on DNA, irrespective of specific binding of the roadblock to cohesin. Its physiological relevance and whether it represents a reversal of the loop extrusion mechanism or 'slippage' of DNA from the loop remains to be investigated, but it is interesting to note that the gradual shrinkage occurred at a similar rate as loop extrusion.

7.2 Discussion and conclusions

Our results indicate that CTCF molecules find their cognate binding sites by facilitated diffusion and once bound to them are sufficient as monomers to block passively diffusing cohesin complexes, which possibly reflects how DNA-entrapping cohesive cohesin accumulates at TAD boundaries [384]. CTCF is also a barrier to actively loop-extruding

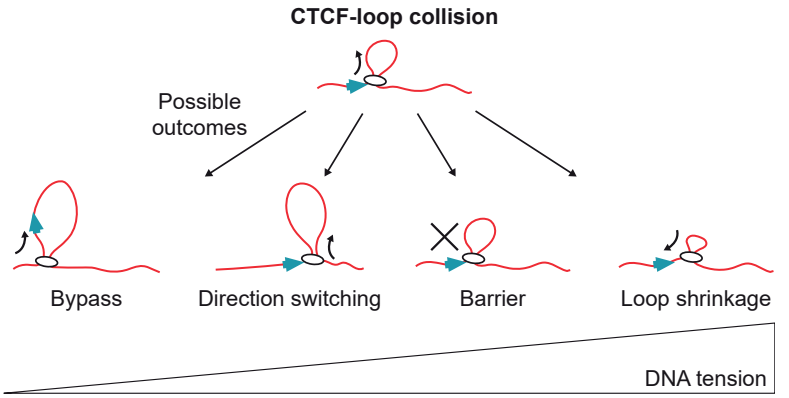


Figure 7.4 | DNA tension affects the outcome of CTCF-loop collisions.
At low DNA tension, CTCF is frequently incorporated into the growing DNA loop. At higher DNA tensions, CTCF promotes loop extrusion direction switching, blocks loop extrusion and, at highest DNA tensions, induces loop shrinkage.

cohesin, presumably reflecting how CTCF establishes TAD boundaries. As predicted from Hi-C experiments, CTCF performs this function asymmetrically with its N-terminus blocking cohesin almost four-fold more efficiently than at its C-terminus. Unexpectedly, this function is regulated by the tension of the DNA that CTCF and cohesin are bound to, implying that genomic processes that alter DNA tension will modulate the permeability of CTCF boundaries and thus the length of chromatin loops extruded by cohesin. A preprint published after submission of this manuscript reported that CTCF bound to an array of four CTCF binding sites is an impermeable barrier to DNA compaction mediated by cohesin (now published [393]). The reasons for this higher blocking activity compared to our study are unknown, but it is likely that the number of CTCF molecules that cohesin encounters could affect its ability to bypass. Furthermore, these authors employed a continuous buffer flow which induces a high degree of DNA tension (which can be estimated to be ~0.5 pN), which likely hindered passing of CTCF. The authors also reported that cohesin slowed its DNA compaction rate when encountering N-terminally oriented CTCF and accelerated when encountering C-terminally oriented CTCF. We, however, did not observe significant changes to the rate of loop extrusion when loop-extruding cohesin passed over N-terminally or C-terminally oriented CTCF (Figure S7.10b,c) or when it encountered the N-terminally oriented CTCF and then switched direction to continue extruding away from the CTCF (Figure S7.10a). Our data indicate that encounters with CTCF can alter cohesin's loop extrusion activity in at least three different ways (Figure 7.4): it can block loop extrusion, it can switch its direction, i.e. cause cohesin to reel in DNA from the opposite side as before, and it can lead to a process in which loop formation is reverted as the loop starts shrinking rather than growing. The observation that TADs detected by Hi-C are 'filled' with chromatin loops which are not anchored at both TAD boundaries may therefore not only reflect the presence of nascent loops which have not been fully extruded yet, as has been assumed so far, but also the existence of 'shrunk' loops which had already reached TAD boundaries but were switched there into a 'reverse' mode by CTCF. It is

conceivable that such a backtracking process is used as a failsafe mechanism for enabling repeated interactions between specific genomic regions in cases where these remained unproductive upon their first encounter during forward loop extrusion, for example during V(D)J recombination of antigen receptor genes [379, 380]. Although indirect effects of CTCF on cohesin, for example inhibition of WAPL or promotion of PDS5 binding at the expense of NIPBL, may enhance the establishment of a barrier to loop extrusion as detected in cell-population measurements (reviewed previously [304]), our experiments indicate that these effects are not strictly required. Together, our findings reveal that CTCF controls cohesin and therefore genome architecture *via* multiple modes. Our results will provide the basis for future mechanistic and physiological studies of CTCF's key functions in gene regulation, recombination, cell differentiation and tumorigenesis.

7.3 Methods

DNA constructs for use as substrates in cohesin diffusion assay

DNA fragments containing a single HighOC1 CTCF binding site [396] (TCAGAGTGGCG-GCCAGCAGGGGGCGCCCTTGCCAGA) were generated by PCR using Phusion Hot Start DNA polymerase (NEB #M0535S) and inserted into the plasmid pPlat (25,754 bp) at the FspAI (ThermoFisher Scientific #ER1661) restriction site in either forward or reverse complement orientation using Gibson assembly [358]. Constructs were then linearized with the restriction enzyme SpeI (New England Biolabs #R3133S) and biotinylated as previously described [211].

DNA constructs for use as substrates in loop extrusion assay

We prepared two constructs of 31.8 kb length which contain a CTCF site placed asymmetrically ~9.7 kb from one end, which allows discrimination of the orientation of the DNA construct based on the binding position of CTCF. One construct was oriented such that the N-terminus of CTCF points to toward the longer end of the DNA (plasmid #121; used for N-terminal encounters) and the motif direction of the other construct was reversed (plasmid #128; used for C-terminal encounters). Plasmid #121 was made of plasmids #64, #66, #67, #69, #118 and #71 (see Table 7.2 for a complete list of intermediate vectors and primers used). Plasmid #128 was made of plasmids #64, #66, #124, #69, #118 and #71 (see Table 7.2). Plasmids #121 and #128 were cloned using Golden Gate cloning, using BsaI-HFv2 as the type-2 restriction enzyme (NEB E1602). Intermediate vectors (#64, #66, #67, #124, #69, #118 and #71) were made using Gibson assembly and traditional (restriction enzyme based) cloning techniques (see Table 7.2) (NEB E2621 Gibson mix; NEB M0515 Q5 polymerase). Biotin-containing handles were made by a PCR reaction with primers JT337 (biotin- GAC-CGAGATAGGGTTGAGTG, IDT) and JT338 (biotin- CAGGGTCGGAACAGGAGAGC, IDT) on #18 pBluescript SK+ (Stratagene), using GoTaq 2 (Promega, M7845). This results in a 1238 bp PCR fragment, which was cleaned up with Promega Wizard™ SV Gel and PCR Cleanup System (Promega, A9282). Fresh plasmid #121 and #128 were purified using the Qiafilter plasmid midi kit (Qiagen, 12243). After purification, the plasmids were cut with both XhoI and NotI-HF and biotin handles were cut with either XhoI or NotI-HF. The digested products were mixed together with a ~10x molar excess of the biotin-handle over the linearised plasmid. Ligation was done with T4 DNA ligase (NEB, M0202L), overnight at 16oC and heat inactivated the next morning for 20 minutes at 65oC. The resulting 31.8 kb DNA construct was cleaned up using an ÄKTA pure system, with a homemade gel filtration column containing approximately 46 ml of Sephacryl S-1000 SF gel filtration media (Cytiva) in TE + 150mM $NaCl_2$. The sample was run at 0.2 ml/min and fractions of 0.5 ml were collected.

DNA constructs for use as substrates in magnetic tweezers assay

DNA constructs for magnetic tweezers experiments of 1.5 kb length were synthesized as described previously [217].

DNA constructs for protein expression

Human NIPBL with N-terminal FLAG and Halo tags and a C-terminal 10xHis tag as a tandem construct with untagged human Mau2 in pLib was as described [173]. 6xHis-Halo-EcoRI^{E111Q} and 6xHis-tetR-Halo in pLib were as described [211]. 10xHis-CTCF-Halo-Flag was cloned into pLib by combining the human CTCF ORF and the Halo-tag ORF using Gibson assembly. A C-terminal Flag-tag sequence was introduced as a 5'overhang in the reverse primer used for Halo-tag ORF amplification. To generate 10xHis-CTCF-Halo-Avi-Flag, the 10xHis-CTCF-Halo-Flag vector backbone was amplified around the end of Halo-tag sequence, at which position an Avi-tag was introduced using Gibson assembly.

Generation of radioactively labelled dsDNA probe for electrophoretic mobility shift assays (EMSA)

100 bp dsDNA fragments containing wildtype or scrambled versions of the HighOC1 CTCF binding site [396] (wt: TCAGAGTGGCGGCCAGCAGGGGGCGCCCTTGCCAGA) were prepared by overlap extension PCR: two ssDNA oligos with partially overlapping sequences were used in a PCR reaction catalyzed by Phusion Hot Start DNA Polymerase (NEB #M0535S) and purified using a PureLink PCR Purification Kit (Invitrogen #K3110002). 1 pmol of dsDNA probe was subsequently incubated with 0.5 µl [γ-^{32}P]ATP (3000Ci/mmol, 10 mCi/ml; Hartmann Analytic #SCP-301) and T4 Polynucleotide Kinase (PNK) (NEB #M0201S) in a 20 µl reaction at 37oC for 1 hour. T4 PNK was subsequently heat inactivated by incubating the reaction at 65oC for 10 min.

Generation of methylated dsDNA probe for EMSA

A 100 bp dsDNA fragment containing the HighOC1 CTCF binding site described above [396]was methylated *in vitro* using M.SssI CpG methyltransferase (NEB #M0226S) following the manufacturer's protocol. To increase methylation efficiency, four rounds of methylation, each followed by DNA purification using PureLink PCR Purification Kit (Invitrogen #K3110002), were performed. Methylation efficiency was assessed by incubating 300 ng of purified methylated DNA with 1 µl of the methylation sensitive restriction enzyme EaeI (NEB #R0508S) in 20 ul reaction containing 1x CutSmart buffer (NEB) at 37oC for 1 hour. The reaction products were resolved by electrophoresis on a 0.8% agarose gel and Ethidum bromide staining was detected using a BioRad ChemiDoc Imaging System. The final dsDNA fragment used as unlabeled, methylated competitor in Figure 7.1b was methylated with ~80% efficiency.

Generation of CTCF-Halo-Flag HeLa Kyoto cell line

HeLa Kyoto cells (RRID:CVCL_1922), a kind gift from Professor S. Narumiya, Kyoto University, Kyoto, Japan, were cultured as described previously [34]. HeLa Kyoto cells were authenticated by STR fingerprinting and tested negative for Mycoplasma contamination. The CTCF-Halo-Flag HeLa Kyoto cell line was generated by homology-directed repair using CRISPR Cas9 (D10A) paired nickase [397]. A donor plasmid comprising CTCF homology arms (719bp and 459 bp on either side of the coding sequence stop site) and Halo-Flag were cloned into plasmid pJet1.2. Cas9 guide RNA sequences were identified using an online tool (https://crispr.mit.edu; gRNA1: CACCGCAGCATGATGGACCGGTGA; gRNA2: CACCGGAGGATCATCTCGGGCGTG) and inserted into plasmid pX335 (a gift from Feng

Zhang, Addgene, 42335). HeLa Kyoto cells were transfected with donor, Cas9 nickase plasmids using Lipofectamine 2000 (Invitrogen #11668019). 7 days later, cells were labelled with Halotag TMR ligand (Promega #G8251) and sorted by flow cytometry. The clonal cell line was selected after verification of homozygous Halo-Flag insertion by PCR of genomic DNA, immunoblotting and inspection by microscopy.

Protein expression and purification

Baculoviruses for protein expression in Sf9 insect cells (Thermo Fisher Scientific) were generated as described previously [360]. Expression cultures were incubated at 27oC for 48 - 60 h after infection. Cells were centrifuged, washed in PBS, frozen in liquid nitrogen and stored at -80oC.

Recombinant CTCF protein purification: Baculovirus-infected cell pellets from cultures supplemented with 0.1 mM $ZnCl_2$ were lysed by Dounce homogenization and resuspended in CTCF lysis buffer (35 mM NaH_2PO_4/Na_2HPO_4 pH 7.4, 350 mM NaCl, 0.1 mM $ZnCl_2$, 5% glycerol, 0.05% Tween-20, 5 mM Imidazole) supplemented with 1 mM PMSF, EDTA-free cOmplete tablet (1 per 50 ml) (Roche #11873580001), 1 mM DTT and 0.001 U/µl Benzonase. The lysate was cleared by centrifugation at 18,000 g for 1 h at 4oC. The soluble fraction was incubated with NiNTA agarose (Qiagen #30230) for 1 h at 4oC and washed with CTCF buffer (35 mM NaH_2PO_4/Na_2HPO_4 pH 7.4, 150 mM NaCl, 0.1 mM $ZnCl_2$, 5% glycerol) supplemented with 1 mM DTT and 35 mM Imidazole. For the final washing step, DTT was omitted from the washing buffer. Protein was eluted with CTCF buffer supplemented with 300 mM Imidazole. The eluate was subsequently concentrated approximately two-fold using a Sartorius Vivaspin 50 kDa MWCOConcentrator (Sartorius #VS2031) and incubated with M2-FLAG agarose (Sigma Aldrich #A2220) for 90 min at 4oC. The resin was washed with CTCF buffer and incubated with fluorescent Halo ligand (TMR Promega #G8252; Alexa660 Promega #G8472,) for 15 min at room temperature. Following extensive washing with CTCF buffer, the labelled protein was eluted in CTCF buffer supplemented with 0.5 mg/ml 3xFLAG peptide. The eluate was supplemented with 1 mM DTT, concentrated 2 - 4-fold using a Sartorius Vivaspin 50 kDa MWCOConcentrator, flash-frozen and stored at -80°C.

HeLa CTCF-Halo-Flag purification: HeLa CTCF-Halo-Flag protein was purified as described for SCC1-Halo-Flag [173] except 20 mM Tris pH 7.5 was used in all CTCF purification buffers instead of 25 mM NaH_2PO_4/Na_2HPO_4 pH 7.5, and 0.1 mM $ZnCl_2$ was included in all purification buffers except for the FLAG elution buffer. HeLa CTCF was labelled with JF 646-HaloTag Ligand.

Recombinant cohesin, HeLa cohesin, NIPBL-Mau2 and EcoRI^{E111Q} protein purification: Recombinant cohesin, HeLa SCC1-Halo-Flag cohesin and recombinant NIPBL-Mau2 were purified as described [173]. EcoRI^{E111Q}-Halo and TetR-Halo were purified as described [211].

Electrophoretic mobility shift assay (EMSA)

For the competition EMSA assay 60 fmol of recombinant CTCF was mixed with 1 µg poly dI-dC (Thermo Scientific #20148E) in a 20 µl reaction containing 35 mM Tris pH 7.9, 50 mM KCl 50 mM NaCl, 5 mM $MgCl_2$, 0.1 mM $ZnCl_2$, 5% Glycerol, 1 mM DTT and 50 ng/µl BSA at room temperature for 20 min. Subsequently, 21 fmol of [γ-^{32}P]ATP (Hartmann Analytic; SCP-501) labeled dsDNA probe was added in the presence of 100x unlabeled competitors (dI-dC; wild type, scrambled or methylated CTCF oligo), and the reaction was incubated at the room temperature for additional 10 minutes. The binding reactions were loaded onto pre-run (1 hour 100 V, 10 mA, ice-cold water bath, 0.5x TBE running buffer) 4% non-denaturing acrylamide gel and the samples were resolved for 1 hour at the same conditions as the pre-run. The gel was exposed to a Phosphoscreen overnight and analyzed by phosphor imaging with a Typhoon Scanner (GE Healthcare). The images shown are representative of two independent experiments.

Recombinant CTCF single-molecule imaging characterization

CTCF flow-in, washing and imaging: Flow cells were incubated with Avidin DN (Vector Laboratories; #A3100) and DNA as described [173] except that pPlat containing a single HighOC1 CTCF binding site was used instead of λ-DNA. Flow cells were washed with 400 µl WB buffer (20 mM Tris pH 7.5, 50 mM KCL, 5 mM EDTA) supplemented with 0.1 mg/ml BSA and 10 nM Sytox Green (ThermoFisher Scientific; #S7020) or Sytox Orange (ThermoFisher Scientific; #S11368) at 50 µl/min. 100 µl recombinant CTCF-Halo (labelled with TMR in experiments shown in Figure 7.1 and in Figure S7.1a-c,f,g,j and Figure S7.2; or labelled with Alexa 660 in experiments shown in Figure S7.1d, e) was then introduced into the flow chamber at 2.5 nM final concentration in CL100 buffer (35 mM Tris pH 7.5, 100 mM KCL, 5 mM $MgCl_2$, 5% glycerol, 0.005% Tween-20, 0.1 mg/ml BSA, 1 mM TCEP) at 30 µl/min and subsequently incubated for 4 min without buffer flow. Flow cells were then washed with CL150 buffer (CL100 buffer supplemented with 50 mM KCl) at a rate of 50 µl/min to remove non-specifically bound CTCF molecules. To determine the orientation of DNA molecules following image acquisition, TMR labelled EcoRI^{E111Q}-Halo or TetR-Halo was flowed into flow cells at 2 nM or 5 nM final concentration, respectively, in EcoRI buffer (20 mM Tris 7.5, 150 mM KCl, 0.1 mg/ml BSA) supplemented with 10 nM Sytox Green at 30 µl/min, incubated for 4 min and washed with 200 ul of EcoRI buffer. All recombinant CTCF single-molecule imaging characterization and cohesin diffusion assay experiments were performed at room temperature. Unless stated otherwise, time-lapse microscopy images were acquired at 4 s intervals using a Zeiss TIRF 3 Axio Observer setup and 488 nm, 561 nm and 639 nm lasers [211]. A proto catechuic acid (PCA)/proto catechuate-3,4-dioxygenase (PCD)/Trolox oxygen scavenger system (final concentration 10 nM PCD, 2.5 mM PCA, 2 mM Trolox); was added to all buffers used during data acquisition.

Imaging the kinetics of recombinant CTCF association with DNA: To image the kinetics of CTCF association with DNA (Figure 7.1c,d, Figure S7.1a), 0.5 nM TMR labelled CTCF-Halo was introduced into flow cells in CL100 buffer at 30 µl/min. For experiments shown in Figure 7.1d and Figure S7.1a, images were acquired at 3.12 sec/frame intervals. For measurements of CTCF residence time on DNA (Figure 7.1c, Figure S7.1f,g) images

were acquired at 10.15 sec/frame intervals.

Positional analysis of recombinant CTCF on DNA: The position of recombinant CTCF on DNA was analyzed in Fiji. EcoRI or TetR mediated end-labeling was used to unambiguously assign the orientation of DNA strands tethered to the surface. The distance between the center of the mass of the fluorescent intensity signal marking the DNA end and the fluorescent signal of protein was measured, and the ratio between the measured distance and the total length of the DNA molecule was calculated as a position along the DNA in bp. Single-molecule tracking of CTCF position was performed using the custom Fiji macro 'KymoAnalysis_2.1.ijm'.

CTCF diffusion coefficient analysis: Single-molecule tracking of CTCF position was performed using the custom Fiji macro 'KymoAnalysis_2.1.ijm'. Spatial positions along the DNA molecule versus time for individual molecules were converted to basepairs by multiplying the positions in micrometers with the average number of basepairs per micrometer, i.e. with the factor (26,123 bp)/R, where R denotes the end-to-end length of the DNA molecule containing 26,123 bp. The MSD was calculated for individual traces and a linear regression in the form $MSD(\tau) = D\tau + o$ was applied to the first 10 time points (corresponding to a maximum time lag of 31.2 s). Here, D denotes the diffusion coefficient, τ is the time lag and o is an offset to correct for a finite localization uncertainty. Larger time lags were not considered for the regression in order to exclude artificial flattening of the MSD curves by reaching the DNA ends.

Recombinant CTCF photobleaching analysis: To quantify the number of recombinant Alexa 660 (A660) labeled CTCF molecules bound at a CTCF DNA binding site, A660 signals on DNA were identified in laser profile corrected images, subtracted from the local background, averaged over 10 frames and plotted in Figure S7.1e.

Determination of the residence time of recombinant CTCF on DNA: To control for fluorophore bleaching in the CTCF *in vitro* residence time experiments, the dwell-time of 'on-DNA' CTCF-HaloTMR molecules (N=140) and 'on-glass' CTCF-HaloTMR-Avi-Biotin molecules (N=142) (the latter coupled to the biotin-PEGylated glass surface *via* Avidin DN) was determined by imaging populations of these molecules in the same microfluidic flow-cell. We then performed a regression of the fluorescence lifetime distribution to an exponential function on the "on-glass" population to compute the photobleaching half-life, which was determined to be $T_{1/2on-glass} = 77.3\ min$. The "on-DNA" dataset was best described by a 2-exponential decay fit with a fixed percentage of events (97/140, 69%) that displayed rapid unbinding, which were attributed to non-specific DNA binding events based on their position along the DNA molecule. This resulted in residence times of $T_{1/2\,fast-on-DNA} = 1.2\ min$ and $T_{1/2slow-on-DNA} = 29.2\ min$, corresponding to non-specific and CTCF site-specific DNA binding events. Neither single-exponential nor 2- or 3-exponential fits in which one of the components was fixed to $T_{1/2on-glass}$ was suitable to describe the observed data. Based on this and the finding that $T_{1/2slow_on-DNA}$ was ~2.7x shorter than $T_{1/2on-glass}$ (29.2 min and 77.3 min, respectively), we concluded that the off-rate of CTCF

on-DNA was significantly faster than the fluorophore bleaching rate and thus the observed "on-DNA" dwell time of CTCF was not significantly limited by fluorophore bleaching.

HeLa CTCF single-molecule imaging characterization

CTCF flow-in, washing and imaging: Flow cells [211] were incubated with 1 mg/ml Avidin DN (Vector Laboratories) for 15 min and washed extensively with DNA buffer (20 mM Tris pH 7.5, 150 mM NaCl, 0.25 mg/ml BSA (ThermoFisher Scientific; AM2616)). 150 µl of 31.8 kb DNA containing a single CTCF site and biotinylated ends was introduced into flow cells at ~20 pM final concentration at 50 µl/min in DNA buffer supplemented with 20 nM Sytox Orange (ThermoFisher Scientific; S11368). Flow cells were washed with 400 µl of wash buffer 2 (50 mM Tris pH 7.5, 50 mM NaCl, 2.5 mM $MgCl_2$, 0.25 mg/ml BSA, 0.05% Tween-20, 20 nM Sytox Orange) at 100 µl/min, followed by 100 µl of imaging buffer (50 mM Tris pH 7.5, 50 mM NaCl, 2.5 mM $MgCl_2$, 0.25 mg/ml BSA, 0.05% Tween-20, 0.2 mg/ml glucose oxidase (Sigma; G2133), 35 mg/ml catalase (Sigma; C- 40), 9 mg/ml b-D-glucose, 2 mM Trolox (Cayman Chemical; 10011659)) and 5 mM ATP (Jena Biosciences; NU- 1010-SOL) supplemented with 20 nM Sytox Orange at 100 µl/min. Stock solutions of glucose oxidase (20 mg/ml), catalase (3.5 mg/ml) and glucose (450 mg/ml) were prepared as described [238]. JF646 labelled HeLa CTCF was then introduced into the flow chamber at 0.5 nM final concentration in 100 µl imaging buffer supplemented with 20 nM Sytox Orange at 30 µl/min. Non-specifically bound CTCF was removed by washing three times with 100 µl imaging buffer supplemented with 220 nM Sytox Orange at 100 µl/min. All HeLa CTCF single-molecule characterization and loop extrusion experiments were performed at 37oC. Time-lapse microscopy images were acquired using a Zeiss Elyra 7 with Lattice SIM^2 equipped with 561 nm and 639 nm lasers, two PCO Edge 4.2 sCMOS cameras and a 63x/1.46 Alpha Plan-Apochromat oil objective. 100 ms exposure time images were acquired sequentially for each channel at 0.4 s intervals in HILO mode.

HeLa CTCF photobleaching analysis: To quantify the number of HeLa JF646 labelled CTCF molecules bound at a CTCF DNA binding site, JF646 signals on DNA were identified in laser profile corrected images, subtracted from the local background and averaged over all frames prior to a bleaching event and plotted in Figure S7.3e. The number of bleaching steps per molecule was determined manually and indicated on Figure S7.3e. The fluorescence intensity of molecules bound at a CTCF DNA binding site that bleached in a single step was 2.2 ± 0.6 (mean ± SD).

HeLa CTCF positional analysis: The position of HeLa CTCF on DNA was analyzed as described in the section 'Determination of DNA loop size and position of single molecules' below.

Cohesin diffusion assay and image analysis: Cohesin diffusion assays were performed essentially as described [211]. CTCF was introduced into flow cells at 2 nM final concentration and incubated for 4 minutes as described in the Recombinant CTCF imaging section above. Flow cells were then washed with CL150* buffer (35 mM Tris pH 7.5, 75 mM NaCl, 75 mM KCl, 1 mM $MgCl_2$, 10% glycerol, 0.003% Tween-20, and 0.1 mg/ml BSA). Cohesin and NIPBL-Mau2 were introduced into flow cells at 0.8 - 2 nM and 2 nM, respectively

in 100 µl of CL100* buffer (35 mM Tris, pH 7.5, 50 mM NaCl, 50 mM KCl, 1 mM $MgCl_2$, 10% glycerol, 0.003% Tween-20, and 0.1 mg/ml BSA) at 30 µl/min. Flow cells were incubated for a further 4 min without buffer flow and then washed with CL250* buffer (35 mM Tris pH 7.5, 125 mM NaCl, 125 mM KCl, 1 mM $MgCl_2$, 10% glycerol, 0.003% Tween-20, and 0.1 mg/ml BSA). Cohesin and CTCF imaging was then performed in the absence of buffer flow for 160 sec at 4 sec/frame intervals. Image acquisition was repeated for 3 - 5 fields of view. DNA orientation was determined by flowing in Sytox Green and $EcoRI^{E111Q}$-Halo or TetR-Halo as described in the Recombinant CTCF imaging section above. Biotin-conjugated quantum dots QD705 (Invitrogen, #Q101163MP) or CTCF-Halo-Avi-Biotin were used as fiducial markers. CTCF-cohesin channels were aligned with TetR/$EcoRI^{E111Q}$-DNA channels using the custom-written Fiji macro 'Movement_analysis_macro_Kymo_10c_3Ch.ijm'. Each DNA molecule containing diffusing cohesin was manually examined for the presence of a single CTCF signal positioned at the regions where the CTCF binding site was introduced. DNA molecules containing multiple or non-specifically bound CTCF molecules were excluded from the analysis. The number of diffusing cohesin foci on the selected DNA molecules was determined and DNA molecules containing more than 4 mobile cohesin foci were excluded from the analysis. Cohesin behaviour on DNA was then analysed and classified as follows: 1) Cohesin diffusion blocked: a) cohesin diffuses freely along the DNA and reaches the CTCF roadblock, bounces back but does not go past the roadblock during the time of imaging. b) cohesin diffuses freely along the DNA, reaches CTCF and becomes immobilized. c) two or more cohesin molecules blocked by CTCF. 2) Cohesin passes CTCF in one direction: cohesin passes CTCF during imaging and diffuses back towards CTCF but does not pass back to the other side. 3) Cohesin passes CTCF multiple times. DNAs with the following events were also excluded from analysis: a) Cohesin diffusing or co-localizing with CTCF. b) Cohesin failing to encounter CTCF. c) Cohesin blocked by a high fluorescence intensity CTCF signal, likely a multimer. d) Cohesin or CTCF bleaches during image acquisition.

Loop extrusion assay

Perpendicular flow loop extrusion assays were performed essentially as described [173, 238]. Flow cells were incubated with 1 mg/ml Avidin DN (Vector Laboratories) for 15 min and washed extensively with DNA buffer (20 mM Tris pH 7.5, 150 mM NaCl, 0.25 mg/ml BSA (ThermoFisher Scientific; AM2616)). 40 µl of 31.8 kb DNA containing a single CTCF site and biotinylated ends was introduced into flow cells at ~ 3 pM final concentration at 15 µl/min in DNA buffer supplemented with 20 nM Sytox Orange (ThermoFisher Scientific; S11368). Flow cells were washed with 20 µl of wash buffer 1 (50 mM Tris pH 7.5, 200 mM NaCl, 1 mM $MgCl_2$, 5% glycerol, 1 mM DTT, 0.25 mg/ml BSA, 20 nM Sytox Orange) at 5 µl/min. Flow was then switched to perpendicular mode and a further 350 µl of wash buffer 1 was introduced at 100 µl/min. 400 µl of wash buffer 2 (50 mM Tris pH 7.5, 50 mM NaCl, 2.5 mM $MgCl_2$, 0.25 mg/ml BSA, 0.05% Tween-20, 20 nM Sytox Orange) was then introduced at 100 µl/min, followed by 100 µl of imaging buffer (50 mM Tris pH 7.5, 50 mM NaCl, 2.5 mM $MgCl_2$, 0.25 mg/ml BSA, 0.05% Tween-20, 0.2 mg/ml glucose oxidase (Sigma; G2133), 35 mg/ml catalase (Sigma; C- 40), 9 mg/ml b-D-glucose, 2 mM Trolox (Cayman Chemical; 10011659)) and 5 mM ATP (Jena Biosciences; NU- 1010-SOL) supplemented with 20 nM Sytox Orange at 100 µl/min. JF646 labelled CTCF was then introduced into the flow

chamber at 0.5 nM final concentration in 100 µl imaging buffer supplemented with 20 nM Sytox Orange at 30 µl/min. Non-specifically bound CTCF was removed by washing three times with 100 µl imaging buffer supplemented with 220 nM Sytox Orange at 100 µl/min. HeLa cohesin and recombinant NIPBL-Mau2 were then introduced into the flow chamber at 0.5 nM and 3.54 nM, respectively, in 250 µl imaging buffer supplemented with 220 nM Sytox Orange at 30 µl/min. For loop extrusion assays in the absence of buffer flow, flow cells were incubated with Avidin DN and washed with DNA buffer as above. DNA was introduced at 15 µl/min - 25 µl/min to vary the DNA tension. Flow cells were then washed and incubated as above without switching to perpendicular mode.

dCas9 binding to DNA

crRNA sequences were chosen at roughly 1/3 of the DNA length and at each end, two sequences were used for efficient binding of the dCas9-gRNA complex per DNA. If located at the same ends, crRNA sequences were spaced at least 2 kb apart to allow discrimination (additionally to bleaching curves) of occasional binding of two dCas9-gRNA complexes per DNA end. Binding sequences were chosen using CRISPOR (http://crispor.tefor.net/crispor.py; PAM indicated in bold):

seq7932: ACTGGACTGCGACCGGGCAG**GGG**

seq11802: CGCGGTGGAGGCAGACGTGG**CGG**

seq18967: CTGGTTATGCAGGTCGTAGT**GGG**

seq21005: GGCATACAAATATTCCATGA**AGG**

gRNA was obtained by annealing a mixture of universal 67mer Alt-R Crispr-Cas9 ATTO550-labeled tracrRNA and crRNA (IDT) matching the binding sites at 95oC for 2.5 min and slow cooling to 5oC in steps of 5oC for 2.5 min each. To couple gRNA to dCas9, 200 nM dCas9 (NEB; NEBM0652T) was mixed with 2 µM gRNA on ice in NEBuffer3.1, incubated at 37oC for 10 min and placed on ice again. To bind the dCas9-gRNA complex to DNA, DNA constructs of 31.8 kb length were used to facilitate measurements at a similar end-to-end length and force regime as for CTCF experiments. DNA was bound to the pegylated glass surface and unbound DNA was washed off with 100 µl imaging buffer. Then, 1 nM dCas9-gRNA was flushed in the flow cell and incubated for 5 min. Non-specifically bound dCas9-gRNA was removed by flushing with 100 µl imaging buffer, supplemented with 1 mg/ml heparin. Heparin was removed by washing with 100 µl imaging buffer. This typically left one to two dCas9-gRNA complexes per DNA. Loop extrusion experiments were then performed as described above with 30 pM cohesin and 75 pM NIPBL-Mau2. DNA was visualized by staining with 25 nM SytoxGreen and excited by a 488 nm laser. gRNA-ATTO550 was excited by 561 nm laser light in an alternating excitation scheme using 60x oil immersion, 1.49NA CFI APO TIRF (Nikon) objective. Emission was collected on a Photometrics Prime BSI sCMOS camera using continuous imaging and an exposure time of 100 ms per frame.

Magnetic Tweezers experiments

The magnetic tweezers instrument and experiments were conducted essentially as described [217] with minor modifications. The instrument consisted of a pair of vertically aligned, 1 mm apart, permanent neodymium-iron-boron magnets (Webcraft GmbH, Germany) that were used to generate the magnetic field [398]. The magnet pair was placed on a motorized stage (translation: Physik Instrumente #M-126.PD2; rotation: Physik Instrumente #C-150.PD) and the light of a red LED ($\lambda = 630nm$) was allowed to pass the magnet pair gap in order to illuminate the sample. Transmission was collected by a 50x oil-immersion objective (CFI Plan 50XH, Achromat; NA = 0.9, Nikon), and the bead diffraction patterns were recorded with a 4-megapixel CMOS camera (Falcon 4M60; Teledyne Dalsa, USA) at 50 Hz. The real-time tracking of the magnetic bead movement in all three dimensions was conducted with a LabView 2011-based (National Instruments, USA) control software described and published previously [364, 399]. Surface-adhered 1.5 µm polystyrene reference beads (PolySciences, Germany) were used as a reference to correct for instrumental drift occurring during measurements. 100-200 beads could be tracked simultaneously in one field of view with a spatial resolution of ~2 nm for the used 1.5 kb long dsDNA tethers [217]. The flow cell and DNA tethering were prepared as described previously [217]. In brief, the reference beads were diluted 1:1500 in PBS buffer (pH 7.4; Sigma Aldrich, USA) and then adhered (~5 min) to the cover glass surface of the flow cell. After washing non-adhered beads out with PBS, sheep digoxigenin antibodies (Roche, Switzerland) at a concentration of 0.1 mg/ml were incubated for 1 h within the flow cell, following a 500 µl washing with PBS and 2 h incubation of 10 mg/ml BSA (New England Biolabs, UK) diluted in PBS (pH 7.4) buffer. After washing with 500 µl PBS buffer, 1 pM of the 1.5 kb linear dsDNA construct was incubated in PBS buffer for 20 min in the flow cell. After washing again with 500 µl PBS buffer, streptavidin-coated superparamagnetic beads (DynaBeads MyOne, LifeTechnologies, USA; prior diluted 1:400 from stock in PBS buffer) with a diameter of 1 µm were added resulting in the attachment of the beads to the surface-tethered dsDNA constructs after ~5 min; unbound beads were washed out afterwards with PBS buffer. Prior to cohesin loop extrusion experiments, the quality of tethered dsDNA constructs was assessed by applying a combination of zero and high force (8 pN), and 30 rotations to each direction at high force: only tethers with singly bound dsDNA and correct DNA end-to-end lengths were used for the subsequent single-molecule experiments. After washing the flow cell with cohesin reaction buffer (40 mM Tris-HCl pH 7.5, 50 mM NaCl, 2.5 mM $MgCl_2$, 1 mM DTT, 0.25 mg/ml BSA, 0.05% Tween-20), 0.1 nM cohesin and 0.25 nM NIPBL-Mau2 were introduced in cohesin buffer supplemented with 2 mM ATP to stretched dsDNA tethers at high force (8 pN). For force-titration experiments (Figure S7.7), the force was lowered in individual experiments to 1, 0.8, 0.6, 0.4, 0.3, 0.2, and 0.1 pN, and maintained for 10 min. All magnetic tweezers experiments were performed at room temperature. The Z-bead position over time was extracted with custom-written scripts in IGOR Pro (v6.37, Wavemetrics, USA), as previously described [217],64 and a custom-written automated step detection algorithm (MatLab, MathWorks, USA) was applied to the individual traces as described [217, 373], in order to extract individual loop extrusion step sizes. Step sizes measured at the same condition from different traces and experiments were pooled and converted into basepairs [217] to construct the distribution of cohesin step sizes in dependence of force (Figure S7.7c,d).

Loop extrusion image analysis

Fluorescence images were cropped to regions of interest spanning single stretched DNA molecules. Cropped images were then further analysed in a custom written python-based as described 46 which allows semi-automated processing and inspection of the data. Images were median-filtered and kymographs were constructed along the long axis of the stretched DNA molecule.

End-to-end length of DNA: The ends of the DNA molecule were determined by a 'peak peeling' algorithm: a time span of the kymograph in which no loop extrusion occurs (usually before flushing in of cohesin) is selected and temporally averaged, which yields an intensity profile of DNA along its long axis. Gaussian peaks of width roughly equal to the full-width-half-maximum (FWHM) of the microscope's point spread function (here 300 nm) were placed at the position of maximum intensity. The height of the Gaussian equals the maximum intensity of the profile. The Gaussian was then subtracted from the intensity profile. The same procedure was iteratively applied on the new profile until less than 10% of the original area under the intensity profile remained. The location of the outermost peaks corresponded to the ends of the DNA and the difference between the peaks corresponded to the end-to-end length.

Determination of DNA loop size and position of single molecules: Peaks in each frame corresponding to DNA loops or single molecules (CTCF or dCas9, from here on referred to as CTCF) were detected in every frame, if present, using the scipy.find_peaks algorithm [299]. Detected spots were then connected using the trackpy package [400] which allows for tracking of loop and single-molecule positions over time. The intensity of the loop (I_{loop}) is computed as the summation of 7 pixels surrounding the peak position, corrected for the amount of DNA outside of the loop that falls within this window (a more detailed description of this procedure and the associated error is described in the error estimation of the DNA tension; see below). The size of the loop L_{loop} (Figure S7.10d) is computed by the fraction of intensity attributed to the loop, I_{loop}, in relation to the total intensity along the entire DNA molecule, I, and the DNA length, $L = 31.8\ kbp$,

$$L_{loop} = \frac{I_{loop}}{I} \cdot L. \tag{7.1}$$

The position of the loop, x_{loop} (in bp), is computed relative to one randomly chosen DNA end and thus depends on the integrated DNA intensity between the chosen end to the loop position (the lead DNA I_{lead}; see Figure S7.10e),

$$x_{loop} = \frac{I_{lead}}{I} \cdot L. \tag{7.2}$$

The loop-CTCF distance is computed similarly by measuring the DNA intensity between loop and CTCF, $I_{loop-CTCF}$. The integration of the DNA intensity between loop and CTCF position was corrected for half of the loop size since loops appear as Gaussian-shaped foci and overlap the DNA connecting loop and CTCF (more details are described below). The loop-CTCF distance is

$$d_{loop-CTCF} = \frac{I_{loop-CTCF}}{I} \cdot L. \tag{7.3}$$

Table 7.1 | Correction factor α and persistence length L_p for varying concentrations of Sytox Orange (SxO) and Sytox Green (SxG).

SxO/SxG concentration [nM]	Correction factor α	Persistence length L_p [nm]
0	1	46.1
10	1.0258	41.9
50	1.0523	36
100	1.0649	35.1
200	1.0948	37.1
500	1.3829	37.2

The binding position of CTCF was measured relative to the DNA's end points. Given the lower (x_{down}) and upper (x_{up}) DNA end points, the tracked CTCF position in physical space, x_{CTCF}, is converted to its sequence position L_{CTCF} *via*

$$L_{CTCF} = \frac{x_{CTCF} - x_{down}}{x_{up} - x_{down}} \cdot L, \tag{7.4}$$

where $x_{up} - x_{down}$ corresponds to the DNA end-to-end length R.

MSD calculation: From the position of the CTCF over time, the MSD is computed over a moving window, whose length w was roughly adjusted based on the time before and after encounter of the CTCF by cohesin (usually between 21 and 51 frames, corresponding to ~8 s to ~20 s) as

$$\mathrm{MSD}(t) = \frac{1}{w} \sum_{\tau=t-\frac{w}{2}}^{t+\frac{w}{2}} |x_{CTCF}(\tau+1) - x_{CTCF}(\tau)|^2 \tag{7.5}$$

Force calculation: The DNA tension within the loop is zero in the absence of buffer flow. In the fraction of DNA outside the loop, the tension depends on the end-to-end length of the DNA. The amount of DNA outside the loop, $L_{nonloop}$ is computed by subtracting the loop size from the length of the DNA construct: $L_{nonloop} = L - L_{loop}$. Its contour length $L^c_{nonloop}$ is computed as

$$L^c_{nonloop} = \alpha L_{nonloop} \cdot 0.342 \ \frac{\mathrm{nm}}{\mathrm{bp}}, \tag{7.6}$$

accounting for the distance between base pairs of 0.324 *nm* and α is a correction factor thereof to account for the slightly different contour length of DNA molecules when bound by Sytox Orange [401]. Values of α were measured by generating force-extension curves of 10 kbp DNA constructs using Magnetic Tweezers [217]and fit to the worm-like chain (WLC) model which is also used to compute the DNA tension [402] (see below; Table 7.2).

The relative extension, that is the extension of a DNA molecule relative to its contour length, of the DNA outside of the loop is computed as

$$r = \frac{R}{L^c_{nonloop}}, \tag{7.7}$$

where $R = x_{up} - x_{down}$ is the end-to-end length of the tethered DNA molecule (Figure S7.10d). The empirically determined and well-established force-extension relationship is used to directly convert the relative extension r of a DNA molecule of length $L_{nonloop}$ and contour length $L^c_{nonloop}$ tethered to an extension (end-to-end length) R [402]:

$$F = \frac{k_B T}{L_p} \left(\frac{1}{4(1-r)^2} - \frac{1}{4} + \sum_{i=1}^{7} a_i r^i \right) \tag{7.8}$$

where $k_B T = 1.3806503 \cdot 10^{-23} \cdot 310\ K \cdot 10^{-18}\ pN \cdot \mu m$, L_p is the persistence length of DNA at the respective concentration of Sytox Orange (Table 7.2), $a_1 = 1$, $a_2 = -0.5164228$, $a_3 = -2.737418$, $a_4 = 16.07497$, $a_5 = -38.87607$, $a_6 = 39.49944$, $a_7 = -14.17718$. For an error estimation of these DNA tension measurements, see below.

Smoothing of loop size, loop-CTCF distance and MSD traces: Traces of loop size and loop-CTCF distance were filtered by an edge-preserving Chung-Kennedy filter [403]with $N = 4,\ 8,\ 16,$ and 32 sampling window lengths, weighting parameter $M = 10$ and sharpness parameter $P = 40$ or a median filter with window size 21 frames.

CTCF-loop co-localization analysis: The duration of CTCF-loop co-localization was identified on smoothed (Chung-Kennedy-filtered) time traces of loop-CTCF distances. Connected stretches at which the loop and CTCF colocalize ($d_{loop-CTCF} = 0$) were identified and components with a gap of only one frame were merged to avoid over-segmenting connected components due to noise in images and erroneous tracking of loop and CTCF positions.

Determination of loop extrusion rates: Loop extrusion rates were determined as described previously [98]. Due to the short time between cohesin flush-in and encounter with CTCF it was not always possible to measure the loop extrusion rate before CTCF encounter. For this reason, we also compared the LE rate after encounter with CTCF with the LE rate measured in the absence of CTCF (Figure S7.10).

Determination of loop shrinkage rate, time span and size of slipped loop: Dissociation events were identified by the time point at which the loop-CTCF distance increased above 0 pixels. To avoid confounding effects of loop release with subsequent re-initiation of loop extrusion, the end point of the loop shrinkage period was marked manually at the time point at which the loop stops shrinking. The loop size before the slippage event was computed as an average over the loop size in the 5 frames (2 seconds) before the slippage event (over which the loop size remained constant). The loop size after the slipping event was based on the average of 1-5 frames, depending on if the loop immediately grew again after slipping (then only the last frame in which the loop still decreases in size can be used) or retained its size (then the loop size was averaged over up to 5 frames). The difference between the loop size before and after slipping corresponds to the fraction of the loop which was lost upon dissociation. The loop shrinkage rate was determined by a linear fit of the loop size between the dissociation time point between loop and cohesin and the time point at which the looped size stopped decreasing.

Fitting of CTCF-loop co-localization time distributions: Distributions of CTCF-loop co-localization times for N- and C-terminal encounters were fit to mono-, bi-exponential and lognormal distributions by a Maximum Likelihood Estimation routine using *scipy.optimize.minimize* [299]. The log-likelihood $\log(\mathcal{L})$ and the number of model parameters k, as well as the number of data points n was used to compute the Bayesian information criterion (BIC) for each of the models and distributions as displayed in Figure S7.8e

$$\text{BIC} = k \cdot \log(n) - 2\mathcal{L}. \tag{7.9}$$

Measuring the stalling force of cohesin: The stalling force of cohesin was measured on DNA molecules without binding protein (i.e. without CTCF, dCas9 or EcoRI) which were stretched to > 50% of its contour length (contour length 11.6 µm for a DNA length of 31.8 kbp in imaging buffer containing 100 nM Sytox Orange). On such stretched molecules, the stalling force is reached before the cohesin reaches one end of the DNA, thus eliminating confounding effects from the DNA ends. For kymographs with multiple extrusion, slipping and direction change events (as e.g. in Figure S7.4a,b), the highest force value (as long as it did not approach one of the DNA ends) during the acquisition time was considered as the stalling force. Note that, in contrast to the stalling forces measured using the fluorescence single-molecule assay, cohesin-mediated loop extrusion steps measured using Magnetic Tweezers (see below) were still detectable at forces up to 1 pN. At this high force, only a single or a few steps were observed per loop extrusion event. Since this would result in loop sizes $\leq$ 1 kbp, they would thus not be visible in the fluorescence assay. Additionally, cohesin molecules may stochastically halt loop extrusion even on DNA with low end-to-end distance and thus at very low tension (see e.g. Figure S7.4a, timepoint 62 s). Thus, a mixture of cohesin molecules was measured, some of which stopped extruding due to reaching the stalling force and some stopped for other reasons. Therefore, the stalling force shown in Figure S7.5b is thus a conservative estimate of the stalling force, but is comparable to encounters with CTCF and dCas9 since these data points were acquired under identical conditions.

Simulating the encounter probability of cohesin and CTCF, given force-dependent cohesin step sizes

A 10 kb stretch of DNA was simulated on which CTCF was assumed to be positioned 7 kb from one end. The cohesin binding site was uniformly sampled along the DNA length. For each force value, step sizes were sampled from the empirically obtained distribution as measured by magnetic tweezers experiments. The simulations were repeated 500 times for every force value and events in which cohesin came within 50 bp of CTCF were counted as encounters, which constitutes a conservative threshold for the interaction distance between cohesin and CTCF.

DNA tension measurement error estimation

The error σ_F of the DNA tension F was computed via error propagation:

$$\begin{aligned}\sigma_F^2 &= \left(\frac{\partial F}{\partial L_p}\sigma_{L_p}\right)^2 + \left(\frac{\partial F}{\partial T}\sigma_T\right)^2 + \left(\frac{\partial F}{\partial r}\sigma_r\right)^2 \\ &= \left(-F\frac{\sigma_{L_p}}{L_p}\right)^2 + \left(F\frac{\sigma_T}{T}\right)^2 + \left(\frac{k_BT}{L_p}\left(\frac{1}{2(1-r)^3} + \sum_{i=1}^{7} i a_i r^{i-1}\right)\sigma_r\right)^2\end{aligned}$$

where $\frac{\partial F}{\partial x_i}$ denote partial derivatives of F with respect to the variables x_i. Here, the variables x_i represent the temperature T, persistence length L_p and the relative extension r. We estimate the contribution on the temperature T, persistence length L_p and the relative extension r separately. Here, we perform a 'worst case' analysis on the resulting error, i.e. we conservatively estimate the uncertainties on the contributing parameters and establish a reasonable upper bound on the resulting error. Persistence length: Fitting of average force extension curves ± one standard deviation yields a persistence length of $L_p = 37.1\ nm$ and $\sigma_{L_p} = 2.1\ nm$. Temperature: The experiments were temperature controlled and we assume a deviation of the set temperature within 1 K, i.e. $\sigma_T = 1\ K$. Relative extension: The relative extension is the fraction of the DNA end-to-end length R and the contour length of the DNA outside the loop $L^c_{nonloop}$

$$r = \frac{R}{L^c_{nonloop}}, \tag{7.10}$$

where $L^c_{nonloop} = \alpha \cdot 0.342\ \frac{nm}{bp} \cdot\ L_{nonloop}$. Its error is computed as

$$\sigma_r^2 = \left(\frac{\sigma_R}{L^c_{nonloop}}\right)^2 + \left(\frac{\sigma_{L^c_{nonloop}} R}{\left(L^c_{nonloop}\right)^2}\right)^2, \tag{7.11}$$

where $\sigma_{L^c_{nonloop}} = \alpha \cdot 0.342\ \frac{nm}{bp}\sigma_{Lnonloop}$. The uncertainty in end-to-end length σ_R is estimated to be in the order of the point spread function (PSF) width due to convolution of the DNA ends with the PSF. The width of the microscope PSF was determined to be $\sigma_{PSF} = 180 \pm 13\ nm$ (Figure S7.10g,h) on single emitters. Due to a potential contribution of DNA intensity beyond the tether points due to the flexibility of DNA (Figure S7.10f; this contribution diminishes after averaging the DNA profile over multiple frames), we conservatively set $\sigma_R = 250\ nm$.

The amount of non-extruded DNA is computed from the total DNA length ($L = 31.8\ kb$) and the intensity of the extruded loop I_{loop} with respect to the total DNA intensity I, i.e. $L_{nonloop} = L\left(1 - \frac{I_{loop}}{I}\right)$. The uncertainty of the DNA amount outside the loop is thus

$$\sigma^2_{L_{nonloop}} = \left(L\frac{\sigma_{I_{loop}}}{I}\right)^2 + \left(-L\frac{I_{loop}\sigma_I}{I^2}\right)^2. \tag{7.12}$$

The total DNA intensity I is constant throughout the experiment due to exchanging Sytox Orange molecules, making DNA photobleaching for the duration of the experiment negligible. Furthermore, several thousand images are acquired, effectively diminishing the uncertainty of the total DNA intensity I. We thus set $\sigma_I \approx 0$. The uncertainty of the loop intensity can be estimated as follows: An illustration of the DNA intensity profile along its long axis is shown in Figure S7.10f. The extruded loop appears as a protrusion on top of the DNA intensity stemming from the non-extruded fraction of the DNA construct. To quantify the loop intensity I_{loop}, we first find the position at the maximum intensity, x_{loop}. We then integrate the intensity profile within a 7-pixel window ($w = 7$) around x_{loop}. However, not only DNA within the loop but also non-extruded DNA contributes to this integral since DNA in the loop and outside cannot be spatially distinguished. We thus correct the integral by measuring the average DNA intensity per pixel $\langle \frac{\partial I}{\partial x} \rangle$ at least two pixels ($\approx$ 200 *nm*) away from the loop and the DNA ends, i.e. without extruded loop. The loop intensity I_{loop} is thus computed as

$$I_{loop} = \int_{x_{loop}-w/2}^{x_{loop}+w/2} I(x)dx - w\langle\frac{\partial I}{\partial x}\rangle. \tag{7.13}$$

The first term describes the integration over a window of size w around the peak maximum position x_{loop} and the second term describes the correction due to the contribution of the DNA intensity from the non-extruded DNA to the integral. The peak is roughly Gaussian-shaped due to the convolution of the loop intensity with the PSF of the microscope. The integral over the peak can be solved explicitly in this case:

$$\begin{aligned} I_{loop} &= \frac{A}{\sqrt{2\pi}\sigma} \int_{x_{loop}-\frac{w}{2}}^{x_{loop}+\frac{w}{2}} e^{-\frac{\left(x-x_{loop}\right)^2}{2\sigma^2}} dx \\ &= \frac{A}{2}\left(\mathrm{erf}\left(\frac{w}{2\sqrt{2}\sigma}\right) - \mathrm{erf}\left(\frac{-w}{2\sqrt{2}\sigma}\right)\right) \end{aligned}$$

A is a normalization constant, erf () is the Error function, and σ is the width of the peak corresponding to the extruded loop. An integral over the entire Gaussian (from $-\infty$ to ∞ instead of $x_{loop} - \frac{w}{2}$ to $x_{loop} + \frac{w}{2}$) yields A. The error by truncating the integral outside the range $[x_{loop} - \frac{w}{2},\ x_{loop} + \frac{w}{2}]$ is thus

$$\sigma_{I_{loop}} = I_{loop}\left(1 - \frac{1}{2}\left(\mathrm{erf}\left(\frac{w}{2\sqrt{2}\sigma}\right) - \mathrm{erf}\left(\frac{-w}{2\sqrt{2}\sigma}\right)\right)\right) \tag{7.14}$$

$\sigma_{I_{loop}}$ is monotonic and increases for larger values of σ. To estimate the width of the Gaussian, we consider a loop of size 10 kb, i.e. $L_{loop} = 10\ kb$, which is close to the maximum loop size in our experiments on the DNA template of $L = 31.8\ kb$ (94% of loop sizes are below 10 kb), yielding the maximum encountered value for the loop peak width σ. The intensity from this signal is convolved with the microscope PSF. An upper bound of the physical dimensions of the unconstrained loop (there is no tension acting on DNA in the loop) can be estimated by the radius of gyration of a circular DNA molecule $R_g = \sqrt{\frac{Nb^2}{12}}$.

The quantity $b = 2L_p$ is the Kuhn length and the chain is made of $N = \frac{L_{loop} \cdot 0.342 \ \frac{nm}{bp}}{b} \approx 49$ segments, yielding $R_g \approx 140\ nm$. For simplicity, we assume that both the PSF and the DNA distribution of the loop are Gaussian-shaped. The convolution of the loop with the PSF is then described by a Gaussian with width $\sigma = \sqrt{R_g^2 + \sigma_{PSF}^2} \approx 230\ nm$, with $\sigma_{PSF} = 180 \pm 13\ nm$ (Figure S7.10g,h). For a loop size $L_{loop} = 10\ kb$, the relative error of the loop intensity is thus $\sigma_{I_{loop}} \approx 15\%$, while$\sigma_{I_{loop}} \approx 9\%$ for a loop size of 3 kb.

We can express the error of the length of the non-extruded DNA irrespective of the absolute intensity value since intensity and DNA content scale linearly under the assumption of a constant density of Sytox Orange dye molecules along the DNA.

$$\begin{aligned}
\sigma_{L_{nonloop}} &= L \frac{\sigma_{I_{loop}}}{I} \\
&= L \frac{I_{loop}}{I} \left(1 - \frac{1}{2}\left(\mathrm{erf}\left(\frac{w}{2\sqrt{2}\sigma}\right) - \mathrm{erf}\left(\frac{-w}{2\sqrt{2}\sigma}\right)\right)\right) \\
&= L_{loop} \left(1 - \frac{1}{2}\left(\mathrm{erf}\left(\frac{w}{2\sqrt{2}\sigma}\right) - \mathrm{erf}\left(\frac{-w}{2\sqrt{2}\sigma}\right)\right)\right)
\end{aligned}$$

We can thus evaluate the uncertainty of the relative extension, σ_r by using the determined values for σ_R and $\sigma^c_{L_{nonloop}} = \alpha \cdot 0.342 \frac{nm}{bp} \cdot \sigma_{L_{noonloop}}$. Illustratively, the absolute and relative error is evaluated for fixed end-to-end length R and fixed loop size L_{loop} in Figure S7.10i,j. For relevant end-to-end length values and loop sizes, the relative error of the DNA tension measurement is $< 17\%$.

Computation of the combinatorial probability of CTCF-CTCF loops

The probability to observe CTCF-CTCF loops with the two respective CTCF sites in a convergent (><), tandem (» and «), or divergent (<>) manner, as reported previously from Hi-C data [34, 103, 155, 303], was obtained from the loop extrusion stalling probability upon encounter with CTCF on its N-terminal (P^N) and C-terminal (P^C) side in the force range 0.04-0.08 pN, as shown in Figure 7.1.1e:

$$\begin{aligned}
P(><) &= P^N P^N / A \\
P(\gg \ \vee \ \ll) &= P^N P^C / A \\
P(<>) &= P^C P^C / A,
\end{aligned}$$

where A is a normalization constant, i.e. $A = P(><) + P(\gg \vee \ll) + P(<>)$.

Statistical analysis and reproducibility

Statistical analysis was performed using GraphPad Prism v.9.4.1 or Python v3.7.7 using scipy v1.5.2, numpy v1.21.6, and statsmodels v0.12.2. No statistical methods were used to determine sample size. Experiments were not randomised and the investigators were not blinded to allocation. Figures were assembled using Adobe Illustrator 27.2. All experiments were performed at least twice with consistent results.

7.4 Supplementary Information

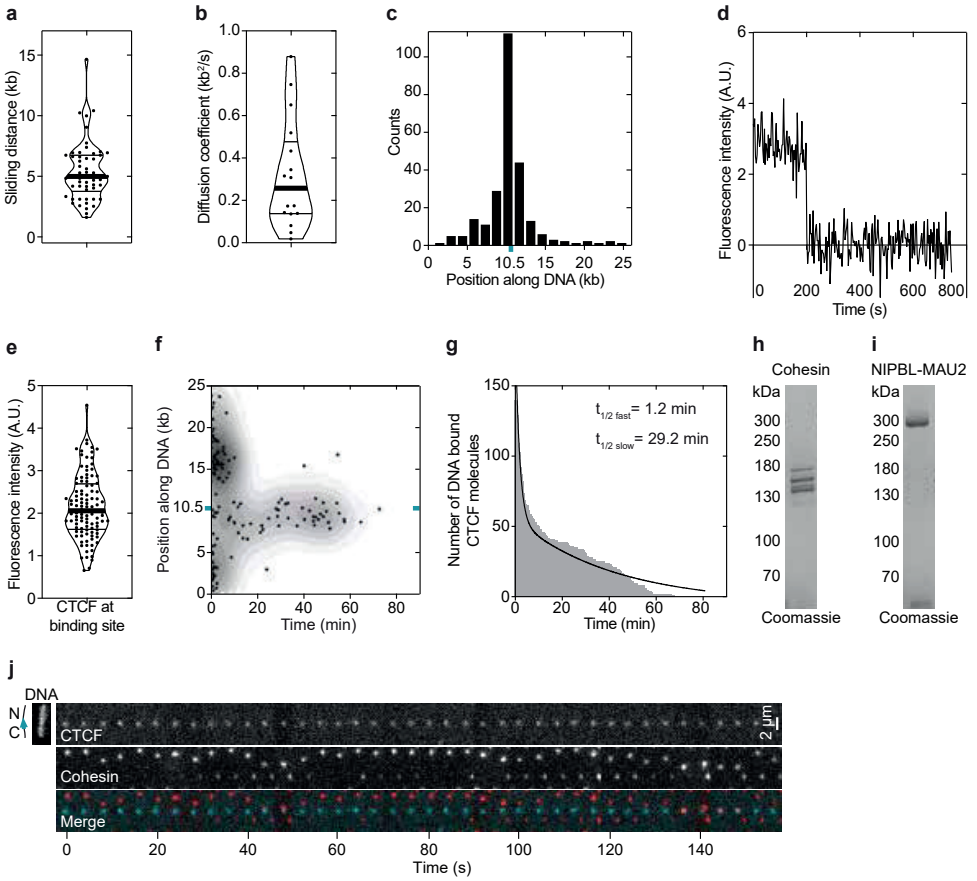


Figure S7.1 | Recombinant CTCF characterization.
(**a**) Distance (kb) travelled by TMR labelled CTCF molecules while diffusing before encountering the CTCF binding site or dissociating. The thick line denotes the median; thin lines denote quartiles. N = 54. (**b**) Diffusion coefficient of diffusing TMR labelled CTCF molecules. The thick line denotes the median; thin lines denote quartiles. N = 17. (**c**) Position of DNA bound TMR labelled CTCF following a brief wash step. The CTCF binding site (cyan tick) is at position 10,452 bp out of 26,123 bp. N = 251. The orientation of the DNA was determined using end-labelling by TetR as shown in Figure S7.2f. (**d**) Time trace of Alexa 660 (A660)-labeled CTCF signal bound at its DNA binding site bleaching in one step. (**e**) Fluorescence intensity of A660 labelled CTCF signals at the CTCF binding site. N = 104. The thick line denotes the median; thin lines denote quartiles. (**f**) Residence time of TMR labelled CTCF on DNA. The CTCF binding site (cyan tick) is at position 10,452 bp out of 26,123 bp. N = 140. (**g**) Residence time of TMR-labeled CTCF on DNA from (f) plotted as a histogram. Bi-exponential decay curve was fitted using Prism. (**h-i**) Coomassie staining of recombinant cohesin and NIPBL-Mau2 after SDS-PAGE. (**j**) Example of cohesin diffusion blocked by CTCF. Cohesin and CTCF were labelled with A660 and TMR, respectively. Sytox Green DNA stain was introduced into the flow cell at the end of the experiment. This data is identical to main Figure 7.1f except it is formatted as a montage rather than as a kymograph.

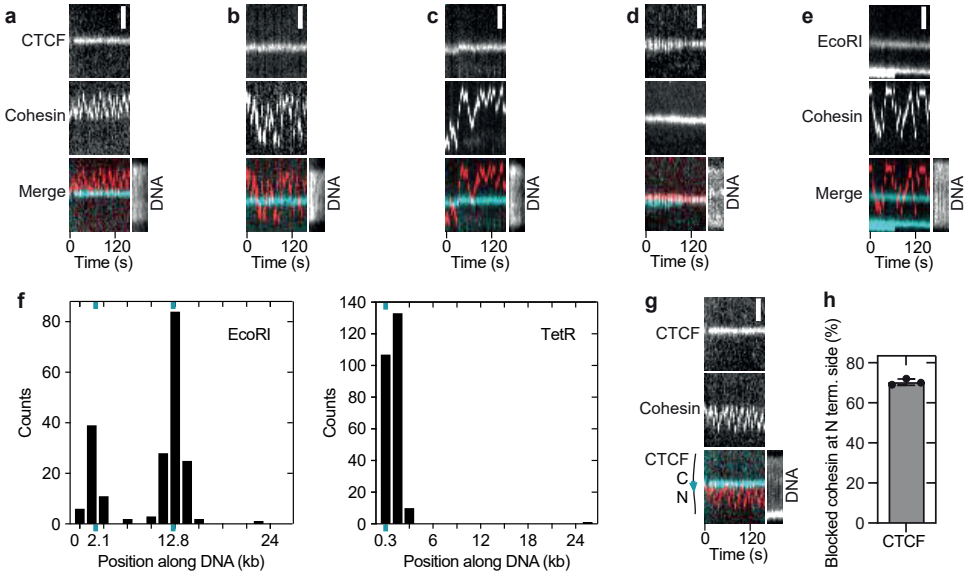


Figure S7.2 | Cohesin diffusion assay characterization.
(**a-d**) Examples of cohesin diffusion on DNAs with CTCF bound at its binding site. Cohesin was labelled with Alexa 660 (red). CTCF was labeled with tetramethylrhodamine (TMR) (cyan). Sytox Green DNA stain was introduced into the flow cell at the end of the experiment. Scale bar, 2 μm. (a) Example of cohesin diffusion blocked by CTCF. (b) Example of cohesin diffusing past CTCF multiple times within the imaging timeframe. (c) Example of cohesin diffusing past CTCF in one direction only. Example of cohesin diffusing past CTCF in one direction only. This behavior was observed very infrequently (2 ± 3 % of N = 264 events). This could be because cohesin-CTCF encounters were recorded after the system has reached equilibrium and so all the single-pass events had occurred before we could image them. If so, is unknown why some cohesin molecules behaved differently i.e. were able to pass CTCF multiple times (Figure S7.2b). (d) Example of cohesin-CTCF colocalization. (**e**) Example of cohesin diffusing past TMR-labeled EcoRI^{E111Q}. (**f**) Positions of DNA bound (left) Janelia Fluor 646-labeled EcoRI^{E111Q} and (right) TMR-labeled TetR, which were flowed into flow cells at the end of diffusion experiments to determine the DNA orientation and hence the orientation of the CTCF binding site at position 10,452 bp. EcoRI restriction sites were present at positions 2,177 bp and 12,802 bp out of 26,123 bp. N = 201. Six TetO sequences were present at positions 40-274 bp. N = 251. (**g**) As in main Figure 7.1f, except using a DNA in which the CTCF site was inverted. (**h**) Fraction of blocked events that diffused on the DNA between the tether point and the N terminal side of CTCF using the DNA template as used in (g) (mean ± SD (N=48) from 3 independent experiments).

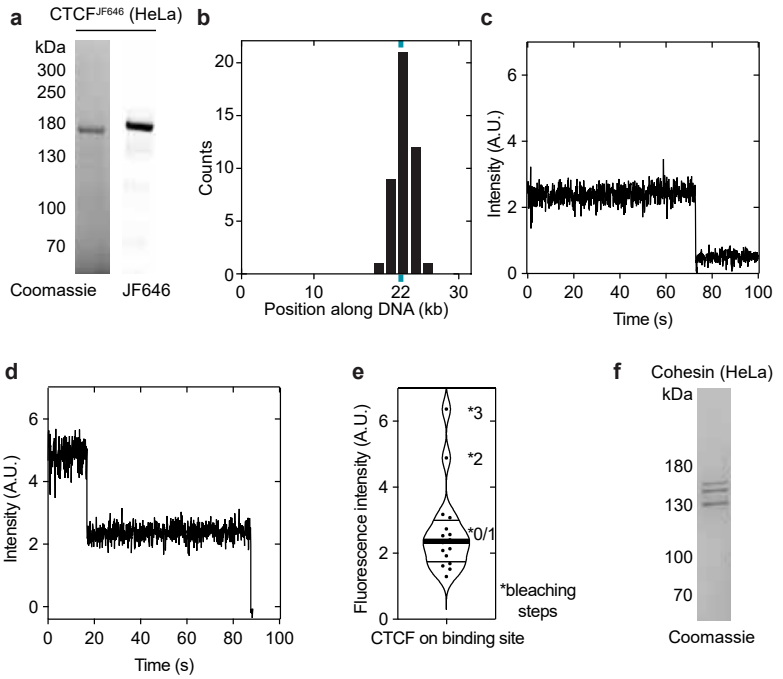


Figure S7.3 | HeLa CTCF characterization.
(**a**) Coomassie staining of HeLa CTCF after SDS-PAGE. JF646 was visualized by epi-red excitation. (**b**) Position of DNA bound JF646-labeled CTCF following a wash step with a buffer supplemented with 220 nM Sytox Orange. The CTCF binding site (cyan tick) is at position 9,667 bp out of 31,767 bp. N = 251. (**c**) Time trace of JF646-labeled CTCF signal bound at its DNA binding site bleaching in one step. (**d**) Time trace of JF646-labeled CTCF signal bound at its DNA binding site bleaching in two steps. (**e**) Fluorescence intensity of JF646-labeled CTCF signals at the CTCF binding site. The thick line denotes the median; thin lines denote quartiles. N = 16. (**f**) Coomassie staining of HeLa cohesin after SDS-PAGE.

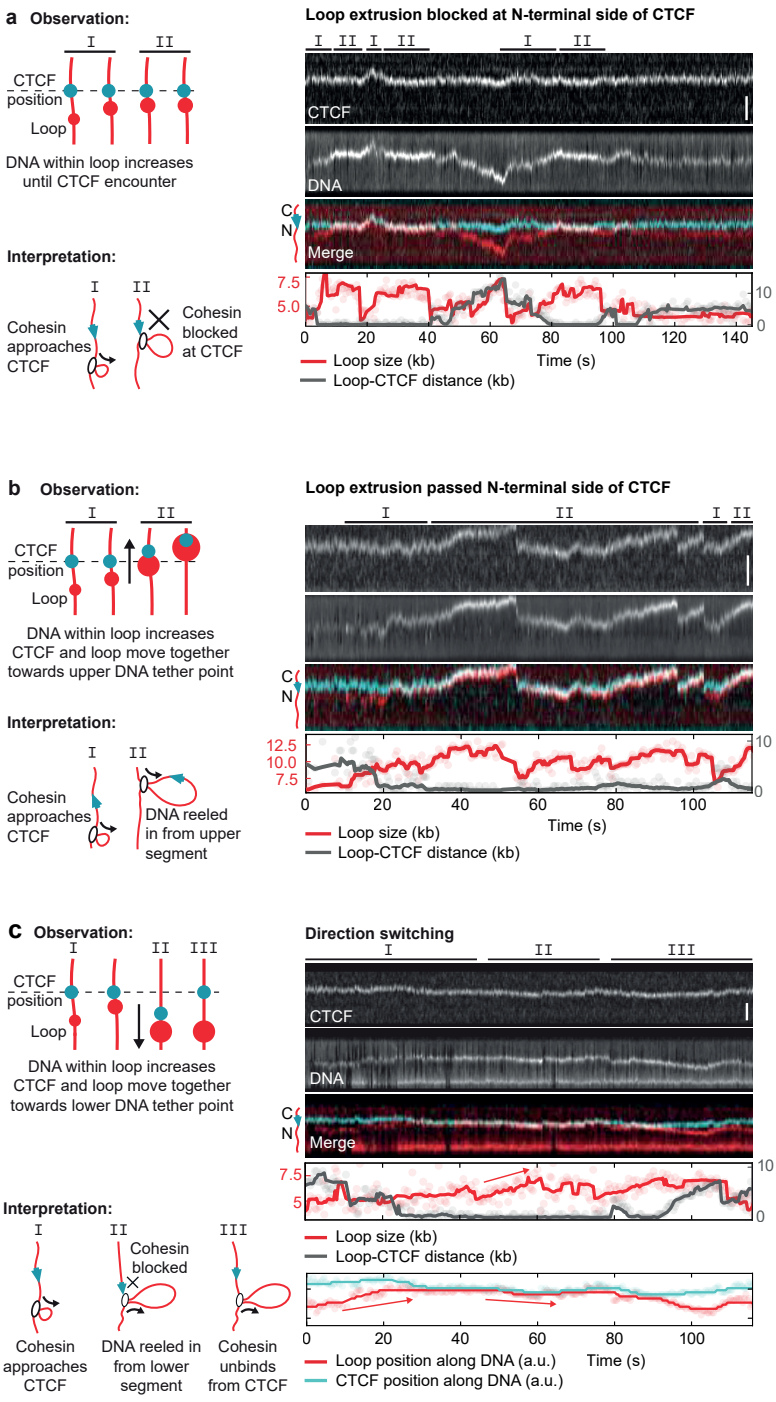


Figure S7.4 | Additional examples of loop extrusion blocking, passing and direction switching upon encountering CTCF.

(**a-c**) (Left panels) observation and interpretation illustrations of (right panels) kymographs of cohesin-mediated DNA loop extrusion encountering N-terminally oriented CTCF (cyan) labeled with Janelia Fluor 646 (JF646).

Caption continues on next page.

Figure S7.4 | DNA loops were visualized by Sytox Orange stain. Scale bar, 2 μm. (a) Growth of the DNA loop stops upon encountering CTCF at timepoints ~ 12 - 18 s, 22 - 40 s and 82 - 95 s. (b) The DNA loop continues to grow upon encountering CTCF at 31 s, and CTCF passes into the loop and translocates with it. (c) The growing loop encounters CTCF at 28 s. CTCF and the growing DNA loop move towards the lower DNA tether point, indicating extrusion on the side facing away from CTCF.

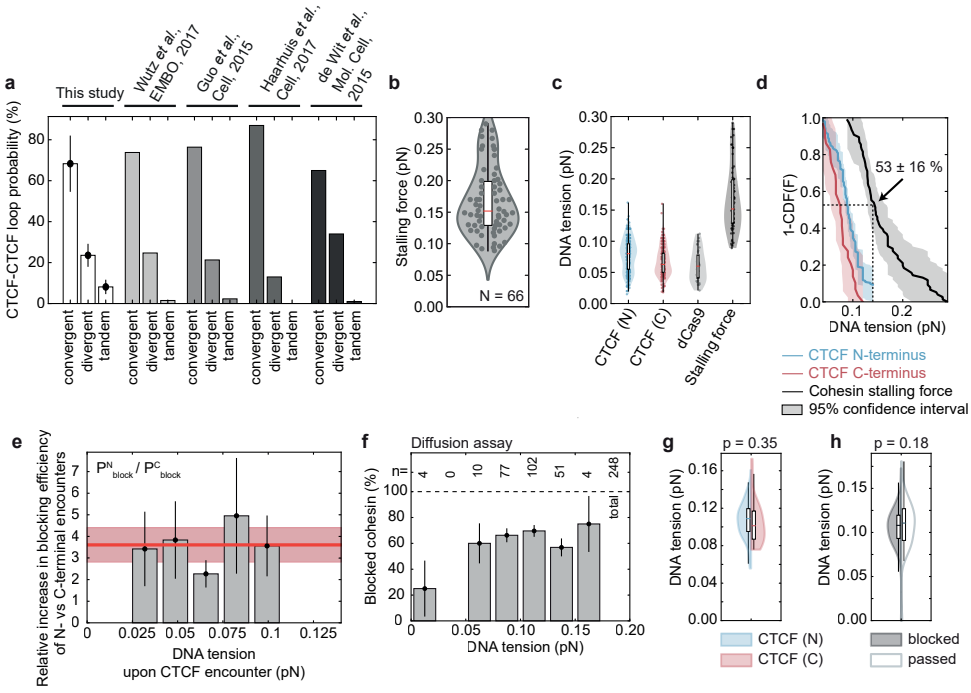


Figure S7.5 | Stalling force of cohesin and force sampling for encounters with the N-/C-terminus of CTCF and dCas9.

(**a**) Combinatorial loop extrusion blocking efficiency at a pair of CTCF sites oriented in a convergent (><), tandem (» and «), and divergent (<>) manner. The percentages were obtained by multiplying the blocking probability of N- and C-terminal encounters in the force range 0.04-0.08 pN, as depicted in Figure 7.1.1e, and normalizing to 100% (see Methods section 7.3). Bar heights denote mean values. Error bars denote the error propagation after multiplication, given the 95% binomial confidence interval as depicted in Figure 7.1.1e. The relative fraction of CTCF-anchored loops that we obtained from the single-molecule experiments are compared to published values extracted from Hi-C data [34, 103, 155, 303]. (**b**) Stalling force of cohesin. horizontal line median; boxes extend to the quartiles and the whiskers show the range of the data (median-1.5*interquartile range (IQR); median+1.5·IQR). Data from 2 independent experiments. (**c**) The DNA tension measured at encounters of loop-extruding cohesin with the N- and C-terminus of CTCF and dCas9. The stalling force values from panel (b) is shown for comparison. N = 297, 184, 37, 66 for CTCF (N), CTCF (C), dCas9 and the stalling force measurements, respectively. (**d**) The empirical survival function (1-CDF) of the data shown in panel c. Thick line represents the mean. shaded areas represent 95% confidence intervals. At the DNA tension of complete stalling at the CTCF N-terminus, 0.14 pN, the survival function decays to 53 ± 16 %, i.e. if loops would be halted alone by reaching the stalling force, one would expect ~53% of loops to exceed the DNA tension of 0.14 pN, which was not observed (compare blue line for stalling at the CTCF N-terminus and Figure 7.1.1g). (**e**) Ratio of the N-terminal and C-terminal blocking probabilities. N-terminal encounters block loop extrusion 3.6 ± 0.8 -fold (The bar height denotes the mean, error bars denote the error propagation after multiplication, given the 95% binomial confidence interval as depicted in Figure 7.1.1g) more often than encounters from CTCF's C-terminal side, independently of DNA tension. N per bin for N-terminal (n) and C-terminal (c) encounters: 0.025-0.0415pN: 70 (n), 72 (c); 0.0415-0.058 pN: 81 (n), 67 (c); 0.058-0.075 pN: 84 (n), 30 (c); 0.075-0.091 pN: 20 (n), 14 (c); 0.091-0.1075 pN: 40 (n), 6 (c); 0.119-0.142 pN: 3 (n), 0 (c). Sample sizes refer to biological replicates. (**f**) Fraction of blocked molecules in the cohesin diffusion assay as a function of DNA tension (note that the DNA tension is constant in diffusion assays since no DNA loop is being extruded). The bar height denotes the mean, error bars denote the error propagation after multiplication, given the 95% binomial confidence interval. (**g**) DNA tension of DNA molecules on which diffusing cohesin was blocked by N-terminally oriented CTCF (left; N = 74 from 2 independent experiments) or by C-terminally oriented CTCF (right; N = 27 from 5 independent experiments). Statistical significance was assessed by a 2-sided 2-sample Kolmogorov-Smirnoff test. (**h**) Violin plot of DNA tension for DNA molecules on which diffusing cohesin was blocked by CTCF (left; N = 161 from 7 independent experiments) or repeatedly passed CTCF (right; N = 88 from 7 independent experiments). Statistical significance was assessed by a 2-sample Kolmogorov-Smirnoff test. Thick horizontal lines on boxplots denote median values, the box extends from the lower to upper quartile values and whisker limits denote the range of data within 1.5 times the interquartile range from the median.

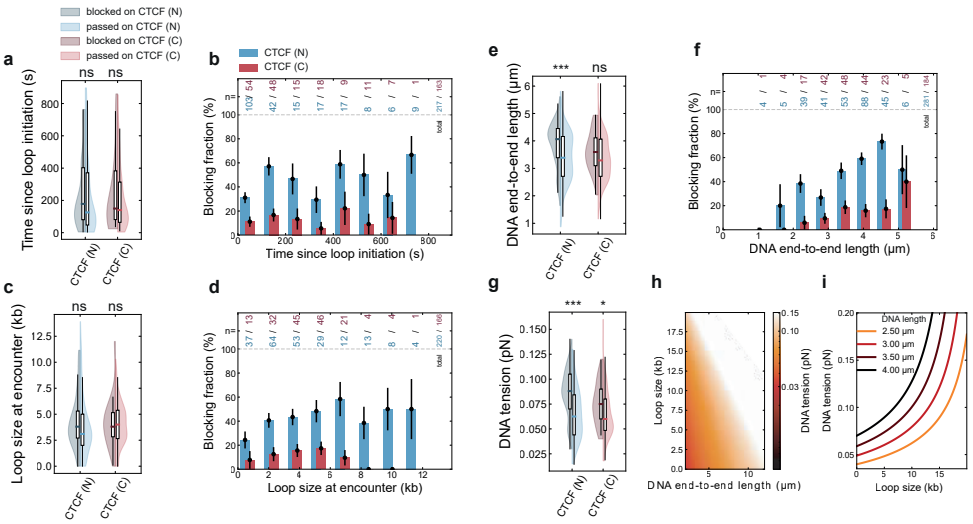


Figure S7.6 | Effect of time since loop extrusion initiation, loop size, DNA end to end length and DNA tension on the loop extrusion blocking probability of CTCF.
(**a**) Time since loop extrusion initiation and N-terminal (blue) and C-terminal (red) CTCF encounters for events which blocked (left part of violin, dark shading) and did not block loop extrusion (right part of violin, light shading). The horizontal line is the median; boxes extend to the quartiles and the whiskers show the range of the data (median-1.5· interquartile range (IQR); median+1.5·IQR). NTD: p = 0.14; CTD: p = 0.89. (**b**) CTCF blocking fraction for N-terminal (blue) and C-terminal (red) encounters for binned times between loop initiation and CTCF encounter. The number of data points per bin is shown on top. Error bars on bar plots denote 95% confidence intervals. (**c**) d, as for (a, b) but for the loop size at encounter (NTD: p = 0.08; CTD: p = 0.90). Error bars on bar plots denote 95% confidence intervals. (**e**) f, as for (a, b) but for the DNA end-to-end length (NTD: p = $5.05 \cdot 10^{-5}$; CTD: p = 0.09). Error bars on bar plots denote 95% confidence intervals. DNAs with a higher end-to-end length are under higher DNA tension due to entropic effects. (**g**) as for (a) but for the DNA tension (NTD: p = $1.38 \cdot 10^{-9}$; CTD: p = 0.02). For a binned representation of the CTCF blocking probability against DNA tension, see Figure 7.1.1g. Thick horizontal lines on boxplots denote median values, the box extends from the lower to upper quartile values and whisker limits denote the range of data within 1.5 times the interquartile range from the median. Error bars on bar plots denote 95% confidence intervals. (**h**) Calculated DNA tension for values of DNA end-to-end length and loop size. The color scale shows white for DNA tension values of *geq* 0.15 pN (see Methods section 7.3). (**i**) Cross-sections through the two-dimensional representation in (h) for specific values of DNA end-to-end length. Even without an extruded loop (loop size = 0 kb), the tethering of the DNA to the surface at the given end-to-end lengths contributes to the DNA tension. For example, a 31.8 kb DNA construct tethered with an end-to-end length of 4 µm (black line) results in a DNA tension of ~0.07 pN. Statistical significance was assessed by a two-sided Mann-Whitney U test without multiple comparison adjustments.

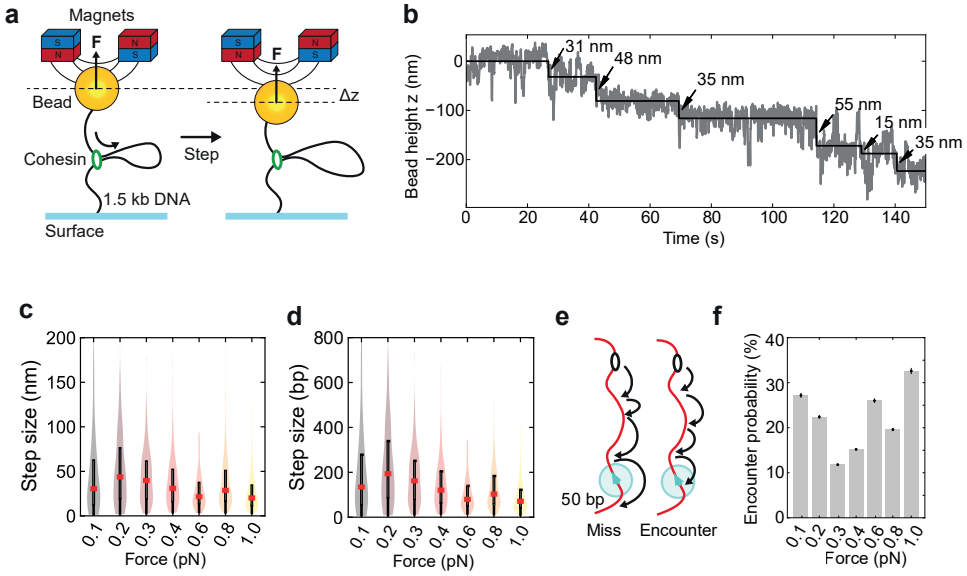


Figure S7.7 | The force-dependent step size of cohesin loop extrusion does not solely explain the observed force dependence of CTCF blocking loop extrusion.
(**a**) Magnetic tweezers setup to observe individual loop extrusion steps by human cohesin, depending on the applied force, based on [217]. The change in bead height Δz corresponds to steps by cohesin. (**b**) Example magnet tweezers trace showing stepwise changes in bead height in the presence of cohesin, NIPBL-Mau2 and ATP. The line denotes steps fitted using the step-finding algorithm. (**c**) Step sizes in nanometers as measured by magnetic tweezers experiments, for various applied forces ranging from 0.1 pN to 1 pN. The horizontal line is the median; boxes extend to the quartiles and the whiskers show the range of the data (median-1.5· interquartile range (IQR); median+1.5·IQR). N = 100, 128, 168, 116, 148, 338, 270 from left to right from 2 independent experiments. (**d**) Step sizes versus force from (c), but converted to base pairs. The median, quartiles and data range are shown as described in (c). (**e**) Simulation setup: starting from a randomly chosen binding position along DNA, cohesin takes steps along DNA, which are sampled from the measured step size distribution. An 'encounter' is considered if cohesin comes within 50 bp of CTCF. Under the lenient assumption that the CTCF N-terminus is unstructured and may be approximated by a freely jointed chain, its radius of gyration R_G is estimated using the $N_K = 268$ amino acids from the N-terminus to zinc finger 126, with a contour length of $l_K \sim 0.4\ nm$ per amino acid [404], resulting in $R_G = 2/6 N_k l_K \sim 7\ nm$ [405]. This distance corresponds to roughly 20 bp, given the contour length of a basepair of 0.3 nm. A threshold of 50 bp was thus conservatively chosen because the CTCF N-terminus may be as long as 14 nm but is likely more compact due to folding of the CTCF N-terminus. The simulations thus likely represent an upper limit of the encounter probability. (**f**) Simulated encounter probability of cohesin and CTCF (mean ± 95% binomial confidence interval; N = 500 independent simulations). Note that the encounter probability does not exceed ~40%, even at the smallest step size distribution (measured at 1 pN). In contrast, the blocking probability of N-terminal encounters of cohesin and CTCF increases from 0 to 100% within 0-0.14 pN (Figure 7.1.1g). Force-dependent step sizes of cohesin can thus not solely explain the observed N-terminal blocking probability. We therefore suspect that DNA tension increases the blocking efficacy of CTCF by other mechanisms, such as by reducing not only cohesin's step size but also the frequency with which it takes steps, thus providing more time for CTCF and cohesin to bind to each other; or by reducing thermal fluctuations of DNA [296], which could reduce the space that CTCF has to explore to find cohesin. It is also conceivable that cohesin's 'motor' activity can overcome the low 1 μM binding affinity of CTCF-cohesin interactions [141] more easily at low DNA tension than at high tensions, which are close to the stalling force of loop extrusion, and at which cohesin has to generate higher forces to extrude DNA. Finally, DNA tension could also change cohesin's responsiveness to CTCF by influencing how cohesin performs loop extrusion [242].

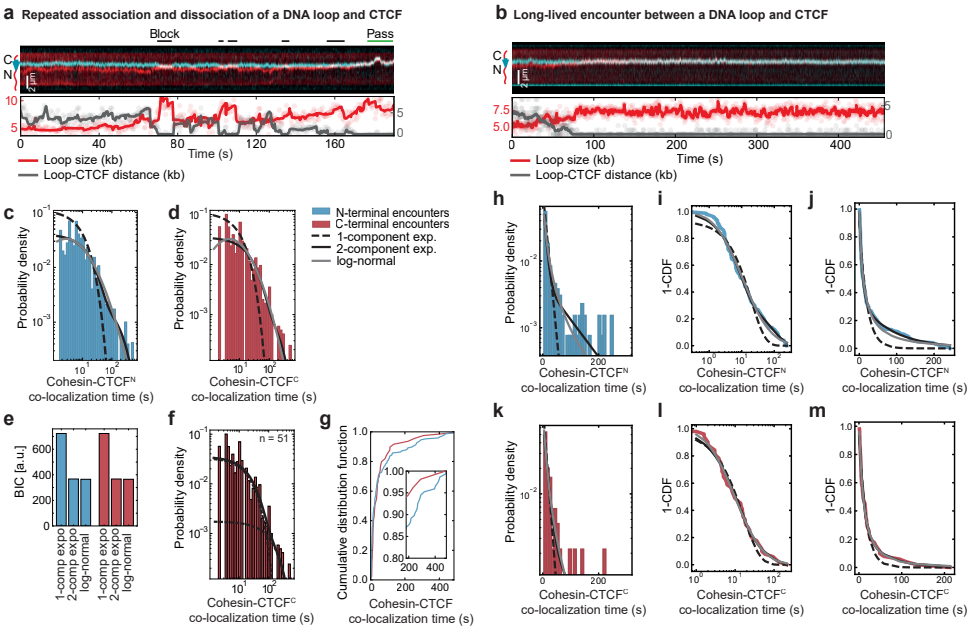


Figure S7.8 | Cohesin-CTCF residence time characterization.

(**a**) Example of repeated approaching of CTCF (cyan) by cohesin, blocking of further loop extrusion and dissociation of the cohesin-CTCF interaction. Cohesin passes CTCF at the end of the kymograph. DNA loops (red) were visualized by Sytox Orange stain. Scale bar, 2 µm. (**b**) Example of a growing loop encountering CTCF (cyan), stalling and co-localizing until the end of image acquisition. DNA loops (red) were visualized by Sytox Orange stain. Scale bar, 2 µm. (**c**) Co-localization times of cohesin for the encounters from the N- and (**d**) C-terminal side of CTCF (N = 147 and N = 51 for N- and C-terminal encounters, respectively). The distributions are fitted to a mono-exponential, bi-exponential and log-normal distribution. e, Bayesian Information Criterion (BIC) for the three models on N- and C-terminal encounters. Notably, both a 2-exponential as well as a lognormal distribution fit the distributions equally well. The parameters of the lognormal fits of the form $(x\sigma\sqrt{2\pi})^{-1}\exp(ln(x)-\mu)^2/(2\sigma^2)$ are $\mu = 3\ s$, $\sigma = 1.5\ s$ for N-terminal and $\mu = 3\ s$, $\sigma = 1.3\ s$ for C-terminal encounters. (**f**) The residence time of encounters between cohesin and CTCF's C-terminus is well described by a bi-exponential distribution with rate constants $k_1 = 0.04\ s^{-1}$ and $k_2 = 0.01\ s{-}1$ ($\tau_1 \sim 25\ s$ and $\tau_2 \sim 100\ s$). (**g**) Cumulative distribution function of the cohesin-CTCF co-localization time for N- (blue) and C-terminal (red) encounters. Inset: magnified view of co-localization times *geq* 3 min. (**h**) data from panel c on a linear x-axis. (**i**) The data in panels c and h plotted as 1-CDF (Cumulative Distribution Function) on logarithmic x-axis. (**j**) the data in panel (i) on a linear x-axis. k-m, as panels (h-j) for encounters of cohesin with C-terminally oriented CTCF.

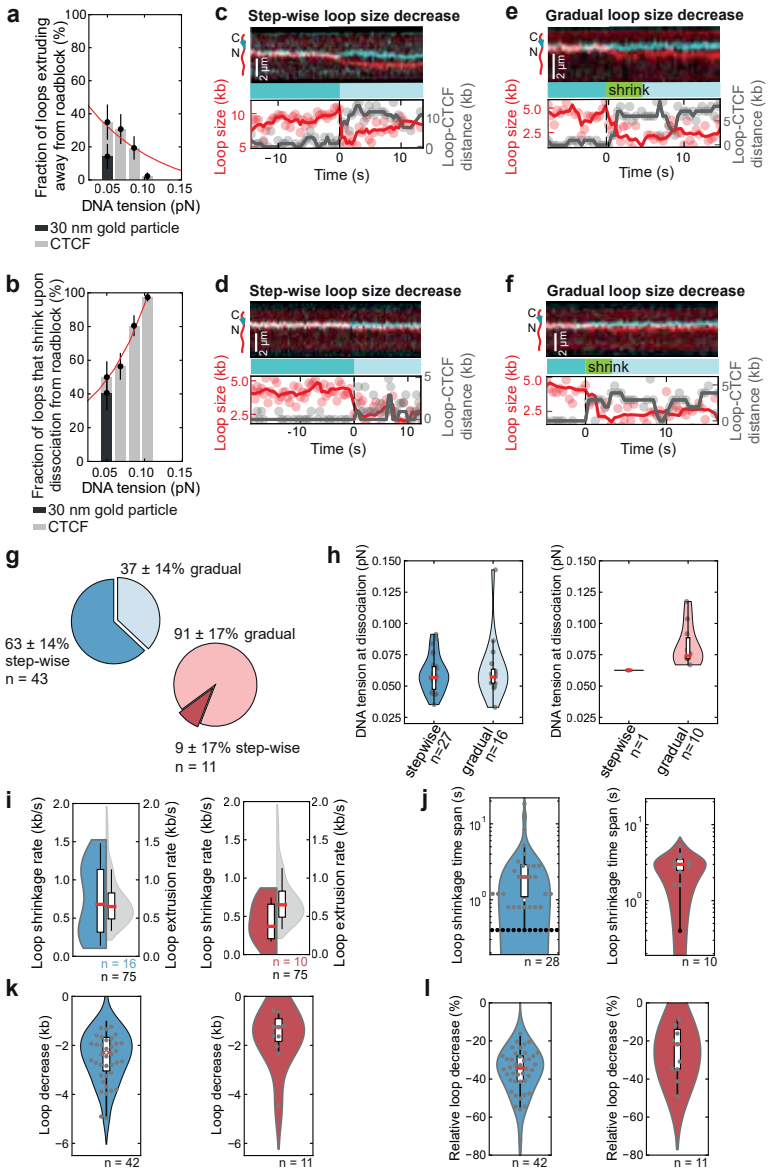


Figure S7.9 | Characterisation of direction switching and loop shrinkage following encounters between cohesin and CTCF or gold nanoparticles.

(**a**) The fraction of loops extruding on the side facing away from CTCF (grey bars) or 30 nm gold nanoparticles (black bar; 14 ± 8 % [mean ± 95% binomial confidence interval]). CTCF data is replotted from Figure 7.3c. Encounters with gold nanoparticles over a force range of 0.02-0.05 pN were reanalyzed from [226] (N = 21 biological replicates from 2 independent experiments). (**b**) The fraction of loops which shrink upon release from CTCF (grey bars) or 30 nm gold nanoparticles (black bar; 41 ± 10 % [mean ± 95% binomial confidence interval]) versus DNA tension at the moment of encounter. CTCF data is replotted from Figure 7.3e. Encounters with gold nanoparticles over a force range of 0.02-0.05 pN were reanalyzed from [226] (N = 22 biological replicates from 2 independent experiments). (**c-d**) Examples of step-wise and (**e-f**) continuous loop shrinkage upon dissociation of cohesin from CTCF. Scale bar, 2 µm. *Caption continues on next page.*

Figure S7.9 | (**g**) The fraction of step-wise and continuous loop shrinkage for encounters from the N-terminal (blue) and C-terminal (red) side (mean ± 95% binomial confidence interval). (**h**) DNA tension for loops which shrink step-wise or gradually. There is no statistically significant difference in DNA tension between the two modes ($p>0.05$, 2-sided 2-sample Kolmogorov-Smirnov test). (**i**) Loop shrinkage rate, in comparison to cohesin loop extrusion rate (grey), and (**j**) distribution of shrinkage time spans. Black dots represent step-wise shrinkage events that happen within one imaging time interval, i.e. 0.4 s. (**k**) Absolute and (**l**) relative loop size decrease for N- and C-terminal encounters in blue and red, respectively. Thick horizontal lines on boxplots denote median values, the box extends from the lower to upper quartile values and whisker limits denote the range of data within 1.5 times the interquartile range from the median. Data for N-/C-terminal encounters were collected from 13 and 3 independent measurements, respectively.

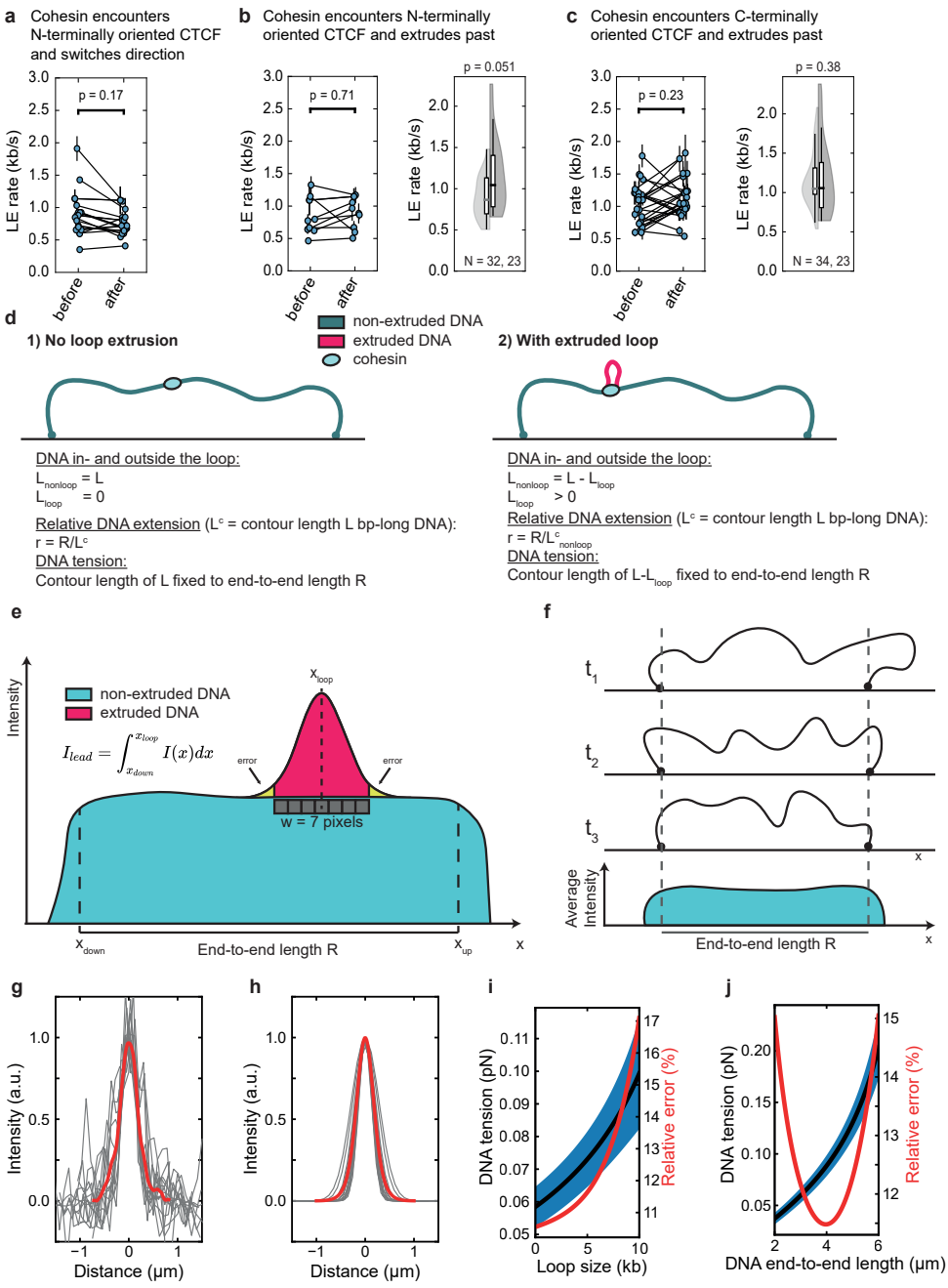


Figure S7.10 | The loop extrusion rate does not change after encounter of cohesin with CTCF.
(**a**) Loop extrusion (LE) rate before and after encounter with N-terminally oriented CTCF when cohesin was blocked at CTCF and then switched extrusion direction to extrude away from it (see e.g. Figure 7.3a). Statistical significance was assessed by a 2-sided Wilcoxon rank-sum test (mean ± SEM; N = 22). *Caption continues on next page.*

Figure S7.10 | (**b**) As for (a) but for events where cohesin passed over N-terminally oriented CTCF (mean ± SEM; N = 9). For events where the time between onset of LE and encounter with CTCF was too short to measure the LE rate, the LE rate was determined after passage and compared to the LE rate in the absence of CTCF (split violin plot on the right). For the latter, statistical significance was assessed by a 2-sided 2-sample Kolmogorov-Smirnoff test. (**c**) as for (b) for events where cohesin passed over C-terminally oriented CTCF (mean ± SEM; N = 38). Error bars on individual data points denote the standard deviations of determined loop extrusion rates in moving 11-frame windows (4.4 s) during the duration of loop extrusion before encounter. Thick horizontal lines on boxplots denote median values, the box extends from the lower to upper quartile values and whisker limits denote the range of data within 1.5 times the interquartile range from the median. Sample sizes (N) refer to biological replicates from 13 independent experiments for N-terminal and 3 independent experiments for C-terminal encounters. (**d-j**), Illustrations of loop size and DNA tension determination, and DNA tension error estimation. (d) In the absence of loop extrusion, a DNA molecule of length L bp is tethered to a surface with end-to-end distance R. The relative DNA extension is thus computed as the ratio of the end-to-end distance R and the contour length of the entire DNA molecule. In the presence of an extruded DNA loop, the non-extruded part of the DNA is effectively shortened by the loop size L_{loop}, while the DNA inside the loop does not experience any tension in the absence of buffer flow. The relative extension is now computed as ratio of the (constant) end-to-end length R and the contour length of the non-extruded DNA, which has the size $L_{nonloop}$ = L - L_{loop}. An increasing loop progressively shortens the non-extruded part of the DNA, giving rise to an increasing tension on the non-extruded DNA due to a fixed end-to-end length R. (e) Illustration of the DNA intensity profile along its long axis. The DNA intensity appears slightly larger than its real end-to-end length due to convolution of the DNA intensity with the microscope point spread function, which is corrected for by the peak peeling algorithm to determine the DNA end-to-end length. The lead intensity is determined as the integrated intensity between one of the DNA ends and the loop position (which corresponds to the peak position of the looped DNA intensity). A 7-pixel window around the loop position is summed and corrected by the intensity contributing in this window from the non-extruded DNA. However, the loop intensity is underestimated due to truncation of the integration outside the range $[x_{loop} - w/2, x_{loop} + w/2]$ (yellow area). (f) At low end-to-end length, the flexibility of DNA might yield a DNA intensity beyond the DNA's tether points. (g) Cross-sections around point-emitters (gray; N = 15) centered at their maximum value and mean trace (red). (h) Gaussian fits the single traces (gray) and mean fit (red) centered at their maximum value. The average Gaussian width was found to be σ = 180 ± 13 nm. (i) DNA tension with absolute (black line and blue area correspond to mean and error of the DNA tension) and relative (red) error of the DNA tension over loop sizes from L_{loop} = 0 to 10 kb at fixed end-to-end length R = 3.5 µm. Error bars denote the estimated error, represented also as red line separately on the right y-axis. (j) Analogous to (i) for a fixed loop size of L_{loop} = 5 kb and varying end-to-end length R. Error bars denote the estimated error, represented also as red line separately on the right y-axis.

Table 7.2 | Cloning scheme: Plasmid (number and name), Cloning reaction, templates used and primers. *Table continues on next page.*

Supplementary Table 1. Cloning scheme: Plasmid (number and name), Cloning reaction, templates used and primers

#64	pBlueScript-GoldenGate			
KLD cloning	Primers	Sequence	Template	
PCR (Q5 NEB E0555L)	JT290	CGACAAGGTCTCGTCCACTAGTTCTAGAGCGG	#61	pBlueScript-noBsaI
	JT291	CGACAAGGTCTCGTCGTATATCGATACCGTCGACCT		
KLD reaction	NEB M0554S			

#61	pBlueScript-noBsaIsite			
KLD cloning				
PCR (Q5 NEB E0555L)	JT288	GGACCCACGCTCACCGGCTC	#18	pBlueScriptII SK+ (stratagene)
	JT289	CGCGGTATCATTGCAGCACTG		
KLD reaction	NEB M0554S			

#66	pGGA-JT292JT293			
Blunt ligation	Primers	Sequence	Template	
PCR (Q5 NEB E0555L)	JT292	CGACAAGGTCTCGACGAAATGCGCGTATGGGGATGG	Lambda	NEB N3011S
	JT293	CGACAAGGTCTCGGGCTCCGTGGGCCAGGTGGT		
PCR (Q5 NEB E0555L)	JT403	CCTGTAGTCTTCTTAATTAAGACGTCAG	pGGAselect	NEB N0309AAVIAL
	JT401	GTACCAAGTCTTCGAATTCGGATC		
Blunt ligation	Enzymes used: T4 PNK (NEB M0201), DPNI (NEB R0176), T4 DNA Ligase (NEB M0202)			

#67	pGGA-JT294JT295			
Blunt ligation	Primers	Sequence	Template	
PCR (Q5 NEB E0555L)	JT294	CGACAAGGTCTCCAGCCGCGACTTACCATGTAT		
	JT295	CGACAAGGTCTCCATACCGCGATGGTTGGAGTTCCA	**Template GeneBlock34 (IDT):**	

CCGCGACTTACCATGTATCTCGTGCGGAACGCTCACGTCTGCTCAGTGGGATGCCGGACATTACCGGACAACTGCTGCGGCACCTCAACTCCGATTTAATGAACGCAATATTCACAAGCAATGCGTGGTGTGCAACCAGCACA
AAAGCGGAAATCTCGTTCCGTATCGCGTCGAACTGATTAGCCGCATCGGGCAGGAAGCAGTAGACGAAATCGAATCAAACCATAACCGCCATCGCTGGACTATCGAAGAGTGCAAGGCGATCAAGGCAGAGTACCAACAG
AAACTCAAAGACCTGCGAAATAGCAGAAGTGAGGCCGCATGACGTTCTCAGTAAAAACCATTCCAGACATGCTCGTTGAAGCATACGGAAATCAGACAGAAGTAGCACGCAGACTGAAATGTAGTCGCGGTACGGTCAGA
AAATACGTTGATGATAAAGACGGGAAAATGCACGCCATCGTCAACGACGTTCTCATGGTTCATCGCGGATGGAGTGAAAGAGATGCGCTATTACGAAAAAATTGATGGCAGCAAATACCGAAATATTTGGGTAGTTGGCGA
TCTGCACGGATGCTACACGAACCTGATGAACAAACTGGATACGATTGGATTCGACAACAAAAAAGACCTGCTTATCTCGGTGGGCGATTTGGTTGATCGTGGTGCAGAGAACGTTGAATGCCTGGAATTAATCACATTCCCCT
GGTTCAGAGCTGTACGTGGAAACCATGAGCAAATGATGATTGATGGCTTATCAGAGCGTGGAAACGTTAATCACTGGCTGCTTAATGGCGGTGGCTGGTTCTTTAATCTCGATTACGACAAAGAAATTCTGGCTAAAGCTCTTG
CCCATAAAGCAGATGAACTTCCGTTAATCATCGAACTGGTGAGCAAAGATAAAAAATATGTTATCTGCCACGCCGATTATCCCTTTGACGTATACGAGTTTGGAAAGCCAGTTGATCATCAGCAGGTAATCTGGAACCGCGAA
CGAATCAGCAACTCACAAAACGGGATCGTGTGGCCACCAGGGGGCGCTAAAAGAAATCAAAGGCGCGGACACGTTCATCTTTGGTCATACGCCAGCAGTGAAACCACTCAAGTTTGCCAACCAAATGTATATCGATACCGG
CGCAGTGTTCTGCGGAAACCTAACATTGATTCAGGTACAGGGAGAAGGCGCATGAGACTCGAAAGCGTAGCTAAATTTCATTCGCCAAAAAGCCCGATGATGAGCGACTCACCACGGGCCACGGCTTCTGACTCTCTTTCCG
GTACTGATGTGATGGCTGCTATGGGGATGGCGCAATCACAAGCCGGATTCGGTATGGCTGCATTCTGCGGTAAGCACGAACTCAGCCAGAACGACAAACAAAAGGCTATCAACTATCTGATGCAATTTGCACACAAGGTATC
GGGGAAATACCGTGGTGTGGCAAAGCTTGAAGGAAATACTAAGGCAAAGGTACTGCAAGTGCTCGCAACATTCGCTTATGCGGATTATTGCCGTAGTGCCGCGACGCCGGGGGCAAGATGCAGAGATTGCCATGGTACAG
GCCGTGCGGTTGATATTGCCAAAACAGAGCTGTGGGGGAGAGTTGTCGAGAAAGAGTGCGGAAGATGCAAAGGCGTCGGCTATTCAAGGATGCCAGCAAGCGCAGCATATCGCGCTGTGACGATGCTAATCCCCAAACCTT
ACCCAACCCACCTGGTCACGCACTGTTAAGCCGCTGTATGACGCTCTGGTGGTGCAATGCCACAAAGAAGAGTCAATCGCAGACAACATTTTGAATGCGGTCACACGTTAGCAGCATGATTGCCACGGATGGCAACATATTA
ACGGCATGATATTGACTTATTGAATAAAATTGGGTAAATTTGACTCAACGATGGGTTAATTCGCTCGTTGTGGTAGTGAGATGAAAAGAGGCGGCGCTTACTACCGATTCCGCCTAGTTGGTCACTTCGACGTATCGTCTGGAA
PCR (Q5 NEB E0555L)	JT403	CCTGTAGTCTTCTTAATTAAGACGTCAG	pGGAselect	NEB N0309AAVIAL
	JT401	GTACCAAGTCTTCGAATTCGGATC		
Blunt ligation	Enzymes used: T4 PNK (NEB M0201), DPNI (NEB R0176), T4 DNA Ligase (NEB M0202)			

#69	pGGA-JT310JT314JT311JT315			
Gibson Assembly	Primers	Sequence	Template	
PCR (Q5 NEB E0555L)	JT310	TCGAGGGATCCGAATTCGAAGACTTGGTACCGACAAGGTCTCCGTATATGAAA	Lambda	NEB N3011S
	JT314	GCTAACCAATTCCTAGGCAGGTCATTG		
PCR (Q5 NEB E0555L)	JT315	CACTGTTGCCAATGACCTGCCTAG	Lambda	NEB N3011S
	JT311	TTCTGACGTCTTAATTAAGAAGACTACAGGCGACAAGGTCTCCCTGACGCTGGCATTCGCATC		
PCR (Q5 NEB E0555L)	JT403	CCTGTAGTCTTCTTAATTAAGACGTCAG	pGGAselect	NEB N0309AAVIAL
	JT401	GTACCAAGTCTTCGAATTCGGATC		
Gibson Assembly	NEB E5520 (NEB HiFi mix)			

#118	pGGA-JT298JT299 noCTCF			
KLD cloning	Primers	Sequence	Template	
PCR (Q5 NEB E0555L)	NH3	TGTGCTCAACAGACGTTTACTGTTC	#65 pGGA-JT298JT299	
	NH4	TCCTGTAATAAGCAGGGCCAG		
KLD reaction	NEB M0554S			

#65	pGGA-JT298JT299			
Blunt ligation	Primers	Sequence	Template	
PCR (Q5 NEB E0555L)	JT298	CGACAAGGTCTCGTCAGACGCGCCGGACGCTACCAGC		
	JT299	CGACAAGGTCTCGACTTCAAATCTTTCTGTATGAAGATTTGAGCACGTTGGCC	**Template GeneBlock35 (IDT):**	

GCGCCGGACGCTACCAGCTTCTTTCCCGTTGGTGGGATGCCTACCGCAAGCAGCTTGGCCTGAAAGACTTCTCTCCGAAAAGTCAGGACGCTGTGGCATTGCAGCAGATTAAGGAGCGTGGCGCTTTACCTATGATTGATCGTG
GTGATATCCGTCAGGCAATCGACCGTTGCAGCAATATCTGGGCTTCACTGCCGGGCGCTGGTTATGGTCAGTTCGAGCATAAGGCTGACAGCCTGATTGCAAAATTCAAAGAAGCGGGCGGAACGGTCAGAGAGATTGATGT
ATGAGCAGAGTCACCGCGATTATCTCCGCTCTGGTTATCTGCATCATCGTCTGCCTGTCATGGGCTGTTAATCATTACCGTGATAACGCCATTACCTACAAAGCCCAGCGCGACAAAAATGCCAGAGAACTGAAGCTGGCGAA
CGCGGCAATTACTGACATGCAGATGCGTCAGCGTGATGTTGCTGCGCTCGATGCAAAATACACGAAGGAGTTAGCTGATGCTAAAGCTGAAAATGATGCTCTGCGTGATGATGTTGCCGCTGGTCGTCGTCGGTTGCACATCA
AAGCAGTCTGTCAGTCAGTGCGTGAAGCCACCACCGCCTCCGGCGTGGATAATGCAGCCTCCCCCCGACTGGCAGACACCGCTGAACGGGATTATTTCACCCTCAGAGAGAGGCTGATCACTATGCAAAAACAACTGGAAG
GAACCCAGAAGTATATTAATGAGCAGTGCAGATAGAGTTGCCCATATCGATGGGCAACTCATGCAATTATTGTGAGCAATACACACGCGCTTCCAGCGGAGTATAAATGCCTAAAGTAATAAAACCGAGCAATCCATTTACG
AATGTTTGCTGGGTTTCTGTTTTAACAACATTTTCTGCGCCCGCCACAAATTTTGGCTGCATCGACAGTTTTCTTCTGCCCAATTCCAGAAACGAAGAAATGATGGGTGATGGTTTCCTTTGGTGCTACTGCTGCCGGTTTGTTTTGAA
CAGTAAACGTCTGTTGAGCACATAGCGCCCCCTGGTGGCCATCCTGTAATAAGCAGGGCCAGCGCAGTAGCGAGTAGCATTTTTTTCATGGTGTTATTCCCGATGCTTTTTGAAGTTCGCAGAATCGTATGTGTAGAAAATTAAA
CAAACCCTAAACAATGAGTTGAAATTTCATATTGTTAATATTTATTAATGTATGTCAGGTGCGATGAATCGTCATTGTATTCCCGGATTAACTATGTCCACAGCCCTGACGGGGAACTTCTCTGCGGGAGTGTCCGGGAATAATT
AAAACGATGCACACAGGGTTTAGCGCGTACACGTATTGCATTATGCCAACGCCCCGGTGCTGACACGGAAGAAACCGGACGTTATGATTTAGCGTGGAAAGATTTGTGTAGTGTTCTGAATGCTCTCAGTAAATAGTAATGA
ATTATCAAAGGTATAGTAATATCTTTTATGTTCATGGATATTTGTAACCCATCGGAAAACTCCTGCTTTAGCAAGATTTTCCCTGTATTGCTGAAATGTGATTTCTCTTGATTTCAACCTATCATAGGACGTTTCTATAAGATGCGTG
TTTCTTGAGAATTTAACATTTACAACCTTTTAAGTCCTTTTATTAACACGGTGTTATCGTTTTCTAACACGATGTGAATATTATCTGTGGCTAGATAGTAAATATAATGTGAGACGTTGTGACGTTTAGTTCAGAATAAAACAATT
CACAGTCTAAATCTTTTCGCACTTGATCGAATATTTCTTTAAAAATGGCAACCTGAGCCATTGGTAAAACCTTCCATGTGATACGAGGGCGCGTAGTTTGCATTATCGTTTTTATCGTTTCAATCTGGTCTGACCTCCTTGTGTTTTGT
TGATGATTTATGTCAAATATTAGGAATGTTTTCACTTAATAGTATTGGTTGCGTAACAAAGTGCGGTCCTGCTGGCATTCTGGAGGGAAATACAACCGACAGATGTATGTAAGGCCAACGTGCTCAAATCTTCATACAGAAAGA
PCR (Q5 NEB E0555L)	JT403	CCTGTAGTCTTCTTAATTAAGACGTCAG	pGGAselect	NEB N0309AAVIAL
	JT401	GTACCAAGTCTTCGAATTCGGATC		
Blunt ligation	Enzymes used: T4 PNK (NEB M0201), DPNI (NEB R0176), T4 DNA Ligase (NEB M0202)			

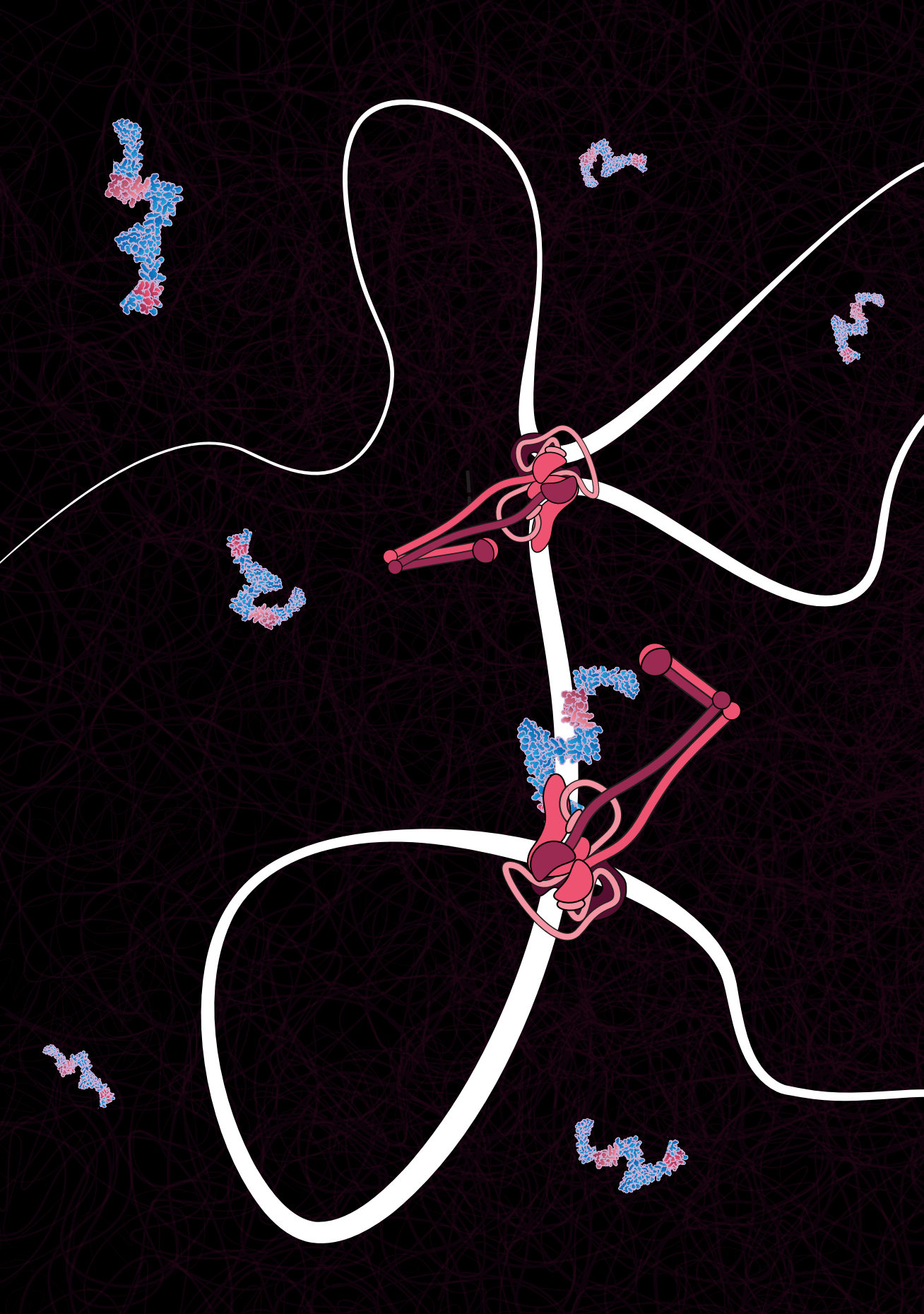
#71	pGGA-JT312JT316JT313JT317			
Gibson Assembly	Primers	Sequence	Template	
PCR (Q5 NEB E0555L)	JT312	TCGAGGGATCCGAATTCGAAGACTTGGTACCGACAAGGTCTCGAAGTAAGAG	Lambda	NEB N3011S
	JT316	TTCACATAAAACATTTTGCATCAGCGAC		
PCR (Q5 NEB E0555L)	JT313	TTCTGACGTCTTAATTAAGAAGACTACAGGCGACAAGGTCTCGTGGATCCGGC	Lambda	NEB N3011S
	JT317	GTGGTCAGGTTGTGGTGATTGGTC		
PCR (Q5 NEB E0555L)	JT403	CCTGTAGTCTTCTTAATTAAGACGTCAG	pGGAselect	NEB N0309AAVIAL
	JT401	GTACCAAGTCTTCGAATTCGGATC		
Gibson Assembly	NEB E5520 (NEB HiFi mix)			

Table 7.2 | *continued.*

#124	pGGA-CTCFdiv			
KLD cloning	Primers	Sequence	Template	
PCR (Q5 NEB E0555L)	JT444	TAGCGCCCCCTGGTGGCCAAAAGAAATCAAAGGCGCGGAC	#67	pGGA-JT294JT295
	NH2	CACGATCCCGTTTTGTGAGTTG		
KLD reaction	NEB M0554S			

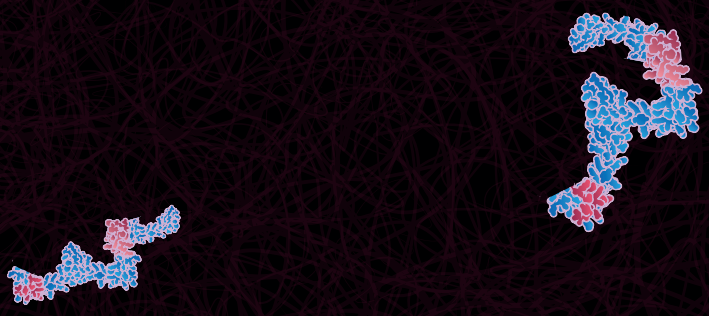
#121	pBS-30-1xCTCF
Backbone	#64 pBlueScript-GoldenGate
Insert	#66 pGGA-JT292JT293
Insert	#67 pGGA-JT294JT295
Insert	#69 pGGA-JT310JT314JT311JT315
Insert	#118 pGGA-JT298JT299 noCTCF
Insert	#71 pGGA-JT312JT316JT313JT317
Golden Gate reaction	NEB E1601

#128	pBS-30-1xCTCFdiv
Backbone	#64 pBlueScript-GoldenGate
Insert	#66 pGGA-JT292JT293
Insert	#124 pGGA-CTCFdiv
Insert	#69 pGGA-JT310JT314JT311JT315
Insert	#118 pGGA-JT298JT299 noCTCF
Insert	#71 pGGA-JT312JT316JT313JT317
Golden Gate reaction	NEB E1601



8

Two CTCF motifs impede cohesin-mediated DNA loop extrusion



This chapter is in preparation for publication: Roman Barth*, Richard Janissen*, Laura Muras*, Jaco van der Torre, Eli van der Sluis, Ashmiani van den Berg, Iain F. Davidson, Jan-Michael Peters, and Cees Dekker: Two CTCF motifs impede cohesin-mediated DNA loop extrusion.

Abstract

The cohesin motor enzyme extrudes DNA into loops that are encoded in the genome through binding of the CTCF protein that stalls DNA loop extrusion (LE) when cohesin encounters its N-terminal region (NTR). It remains unclear, however, what protein-protein interactions underlie this stalling. Here, we use single-molecule magnetic tweezers and fluorescence assays to resolve the effect of fragments of the CTCF NTR on cohesin's LE activity. We find that cohesin frequently fails to perform complete LE step cycles when exposed to the YDF motif of the CTCF NTR. The YDF motif is furthermore observed to convert cohesin into a unidirectional extruder. Independently, the KTYQR motif near the CTCF N-terminus reduces cohesin's LE activity by impeding LE stepping, while not affecting its LE directionality. We thus identify two CTCF motifs that impede cohesin-mediated DNA loop extrusion through distinct mechanisms – highlighting the multifaceted ways employed by CTCF to modulate loop extrusion.

This chapter is a progress report and is currently being developed further before the corresponding manuscript will be submitted for publication.

The spatial organization of the genome crucially impacts many nuclear processes such as replication, transcription, and DNA repair [1, 282]. Structural Maintenance of Chromosomes (SMC) proteins are one of the main contributors that fold genomes. All eukaryotic members of the SMC family, condensin, cohesin, and SMC5/6, share the ability to progressively extrude DNA loops, utilizing ATP as power source [98, 138, 173, 204, 212, 214–216]. These SMC proteins exhibit a common architecture consisting of two Smc subunits that heterodimerize at their hinge region and harbor ATP-binding domains (Figure 8.1a) at the ATPase 'heads' that are mutually connected *via* a flexible kleisin linker to which two subunits attach (NIPBL and STAG for cohesin). Cohesin is the dominant SMC protein to extrude loops during interphase [160, 161, 214]. Genomic contact maps of cell populations, generated by Chromosome Conformation Capture ('3C') technology, revealed genomic regions spanning several hundred kbps on average. These regions exhibit enhanced self-interaction and modest insulation to neighboring domains, forming so-called topologically associated domains (TADs). TADs critically depend on cohesin and are weakened by the depletion of CCCTC-binding factor (CTCF) [33, 153].

The borders of TADs are often demarcated by co-localization of cohesin and CTCF [32–34, 39, 41, 304] CTCF-anchored DNA loops [32–34, 39, 41, 148, 153, 158]. Strikingly, the majority of CTCF binding sites colocalizing with cohesin-dependent loops are oriented such that the N-terminal region of the CTCF protein points toward the interior of the loop/TAD [34, 103, 148, 155, 249, 303], suggesting that loop progression by cohesin is specifically inhibited by the CTCF N-terminus. Previous single-molecule *in vitro* experiments demonstrated that CTCF alone is sufficient to halt DNA loop extrusion (LE) by cohesin and that cohesin approaching CTCF from its N-terminal side is indeed blocked more efficiently than when approached from CTCF's C-terminal side [35, 393] (chapter 7). The interaction between cohesin and CTCF [35, 150, 151, 251, 388, 406] is, at least partially, mediated by a YxF motif [141] that is located at CTCF's N-terminal region (CTCF NTR; Figure 8.1a). This motif binds to a 'conserved essential surface' (CES) formed by STAG1/2 and kleisin of the cohesin complex [141, 248, 250, 407, 408, 408, 409, 409, 410]. Mutation of the YxF motif strongly diminishes CTCF-anchored loops [141]. However, cohesin enrichment at CTCF binding sites is less affected [141]. This suggests that cohesin-CTCF co-localization does not necessarily give rise to CTCF-anchored loops [141, 252, 406, 411, 412]. Although the mechanism by which CTCF inhibits DNA loop extrusion by cohesin remains 'a mechanistic mystery' [411], several hypotheses exist. These include nucleosome positioning around CTCF binding sites, clustering of CTCF *in vivo*, or indirect inhibition of cohesin's ATPase activity through regulatory proteins such as PDS5 and ESCO1. Notably, however, most of the proposed mechanisms involve changes to the *in vivo* micro-environment of DNA-bound CTCF or involve regulatory proteins, while both of these were absent in *in vitro* experiments on CTCF-cohesin interactions [35, 393] (chapter 7). It thus remains an important unresolved question how CTCF regulates cohesin-driven DNA loop extrusion.

To probe the NTR for potential motifs besides the YxF motif and to understand how CTCF interferes with DNA loop extrusion by cohesin, we directly measured the effects of protein-protein interactions between cohesin and CTCF on DNA loop extrusion. We

dissected the CTCF N-terminus into individual fragments and measured their impact on cohesin's DNA loop extrusion dynamics using *in vitro* single-molecule magnetic tweezers and fluorescence visualization assays. Fluorescence HILO microscopy [98] allows us to measure extrusion characteristics such as the directionality of extrusion [222] (chapter 5), while magnetic tweezers are a sensitive single-molecule technique that can resolve single loop-extrusion steps of cohesin and other SMCs [35, 50] (chapter 6, chapter 7). The two assays thus probe complementary aspects of cohesin-mediated DNA loop extrusion and its response to various CTCF NTR fragments. We find that the presence of the CTCF N-terminal region (NTR) reduced the frequency with which cohesin engages in active DNA loop extrusion phases. MT experiments suggest that cohesin frequently fails to complete DNA loop extrusion steps when exposed to the YDF motif. However, this effect vanished when a second motif, KTYQR, positioned at amino acids 23-27 of CTCF, was present. Remarkably, the presence of the YDF motif caused cohesin to predominantly extrude DNA loops unidirectionally as observed by fluorescence microscopy. We furthermore show that the KTYQR motif, previously only implicated in binding PDS5A [252], also directly impedes cohesin's stepping activity while it does not affect its LE directionality. These results untangle the individual and combined effects of the two motifs on cohesin's DNA loop extrusion dynamics, thus elucidating how CTCF stalls cohesin.

8.1 CTCF's NTR impedes DNA loop extrusion activity of cohesin

To probe cohesin's DNA loop extrusion behavior in response to the CTCF NTR and CTR (Figure S8.1a) on the single-step level, we turned to magnetic tweezers (MT). This method is capable to detect individual LE steps of SMC complexes [35, 50] (chapter 6, chapter 7). Short (1.5 kbp) dsDNA molecules were tethered between a superparamagnetic 1 μm bead and an anti-digoxygenin-functionalized glass surface [217]. To apply force, a pair of permanent magnets was positioned above the flow cell (Figure 8.1b; Methods). 100 pM purified human cohesin^{STAG1} (henceforth simply denoted as 'cohesin' unless otherwise stated), 250 pM NIPBL-Mau2, and 1 mM ATP were flushed into the flow cell together with a CTCF fragment as indicated, at an applied force of 7 pN to prevent compaction of the DNA molecules during flush-in (Methods). After stopping the flow, the applied force was rapidly lowered to 0.3 pN, upon which the DNA end-to-end length changed stepwise [35, 50] (chapter 6, chapter 7). To characterize cohesin's stepping characteristics, we developed a step-finding algorithm (Methods; Figure S8.1b,c) which more accurately extracted steps compared to a method we previously employed to identify SMC-induced LE steps (Methods; Figure S8.1d-f [373].

MT traces typically showed three distinct signatures (Figure 8.1c). Consecutive downward steps (> 1 step; green), a single downward step which was followed by a single upward step of the same (± 25%) size (blue; Methods; [217]), and a large upward step exceeding the previous downward step in size (yellow). We refer to these three signatures as consecutive steps, reversals (the combination of down- and upward steps), and loop loss, respectively. Reversals occurred about as frequently as consecutive downward steps and loop loss was observed only infrequently (e.g. Figure 8.1d; 0 μM NTR). Consecutive steps

relate to cohesin's ability to extrude loops of DNA in a step-wise fashion. When using SMC mutants that can bind but are unable to hydrolyse ATP (EQ/EQ mutant) or when using non-hydrolyzable ATP analogous, SMCs cannot take consecutive LE steps. Instead, only reversals (and associated DNA twist) are observed in these cases [50] (chapter 6). Current working models of SMC-mediated DNA loop extrusion propose that DNA is reeled into the Smc lumen upon ATP binding, constituting a single downward step [241, 242, 247]. In a recent preprint, Yamauchi et al. simulated trajectories of loop extrusion by all-atom molecular dynamics simulations [413]. In the simulated trajectories, DNA was reeled through the Smc lumen, as predicted [241, 242, 247]. Notably, however, this yielded only a small fraction (45/770 runs) of successful LE steps, i.e. cases in which the DNA segment that was threaded through the lumen was merged with the already extruded loop. In most cases, DNA slipped out of the lumen. In agreement with the SMC stepping observed in MT experiments, the authors did not detect a significantly different step size for successful and unsuccessful LE steps [50, 413] (chapter 6). Based on these findings, our working model considers reversals as attempted but unsuccessful loop extrusion steps (Figure S8.1g).

In experiments where we added increasing concentrations of the CTCF NTR to cohesin in MT experiments, we observed an overall decrease in activity, i.e., fewer steps per trace during the fixed acquisition time of 12 min (Figure 8.1d; Figure S8.1h,i). Both consecutive steps and reversals were equally affected as the ratio of their relative contributions remained approximately constant (Figure S8.1k). In similar experiments on *Saccharomyces cerevisiae* condensin, the addition of CTCF's NTR did not affect the LE activity (Figure S8.1j,k; Figure 8.1f; Methods), indicating that the observed impact on human cohesin was specific and not an artefact arising from crowding or contaminations of the NTR purification.

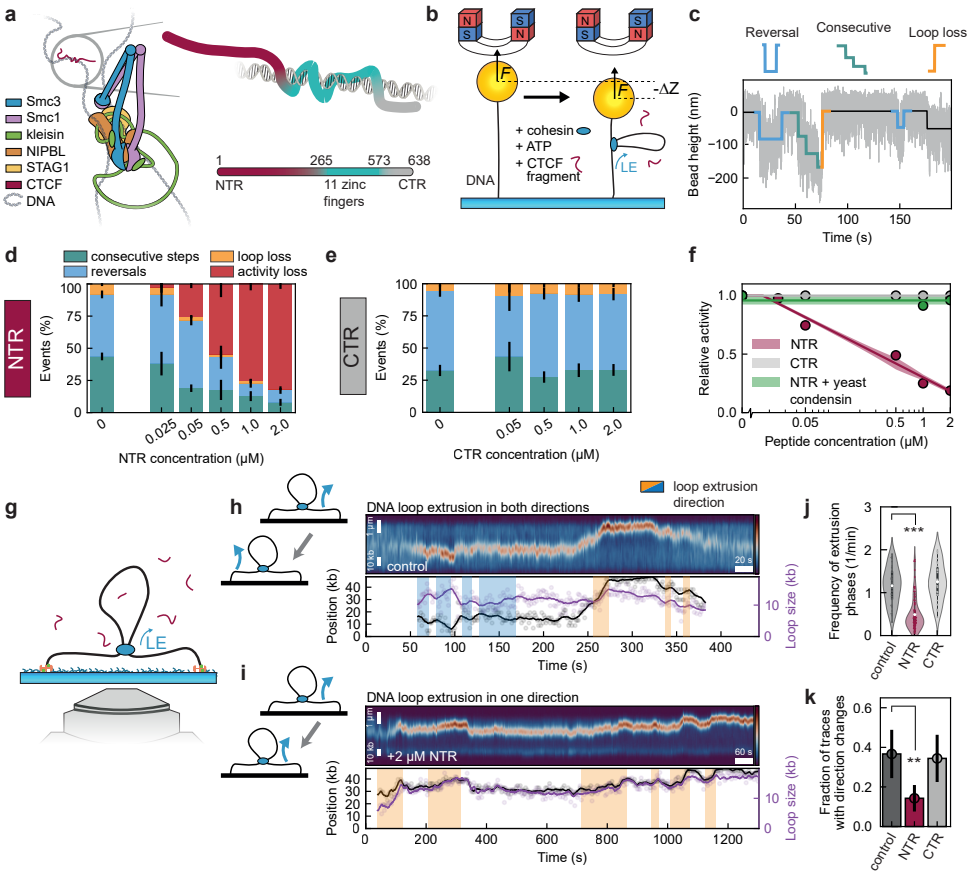


Figure 8.1 | The N-terminal region of CTCF inhibts DNA loop extrusion by cohesin.
(**a**) Illustration of a DNA-bound human cohesin complex and CTCF (based on PDB 7W1M [393]). The magnification (right) shows the N-terminal region (NTR; red) and C-terminal region (CTR; grey) of CTCF, separated by 11 zinc fingers (cyan) that bind to DNA. (**b**) Magnetic Tweezers (MT) setup used to observe individual DNA loop extrusion (LE) steps by SMC protein complexes [35, 50] (chapter 6, chapter 7). 100 pM cohesin and 250 pM NIPBL-Mau2 are used for all MT experiments. (**c**) Example MT trace illustrating the classification of identified stepping events into reversals (blue), consecutive LE steps (green), and loop loss (yellow). Raw data is shown in light grey, the black line shows the resulting fit from the stepfinding algorithm (Methods). (**d**) The fraction of stepping events observed in MT traces classified into consecutive LE steps (green; downwards), reversals (blue; down-up step combination; Methods), and loop loss (yellow; upwards) with increasing concentrations of CTCF NTR. Loss in activity (red) denotes that proportionally fewer steps are observed compared to the control measurement ([NTR] = 0 μM). Error bars denote 95% binomial confidence intervals (CI). N = 1097, 109, 537, 46, 75, 77 steps from 193, 19, 114, 16, 51, 103 tethers, respectively, for increasing CTCF NTR concentration. (**e**) As described in (d) but the CTCF CTR was added in increasing concentrations. Error bars denote 95% binomial confidence intervals (CI). N = 445, 83, 466, 409, 609 from 114, 51, 62, 48, 96 tethers, respectively, for increasing CTCF CTR concentration. (**f**) The relative stepping activity of cohesin, compared to the corresponding control ([peptide] = 0 μM), for increasing concentrations of CTCF NTR, CTR, and NTR in combination with yeast condensin. The data corresponds to the sum of consecutive LE steps, reversals, and loop loss (1-the red bar height) of panels (d), (e), and Figure S8.1j, respectively. The solid line depicts a piecewise linear fit (constant until a peptide concentration x_0 and linear fit with negative slope thereafter). The shaded error bar shows the interpolated error of the individual data points. (**g**) Fluorescence visualization assay to observe cohesin-mediated DNA loop extrusion on individual strands of λ-DNA that were tethered to a pegylated glass surface *via* biotin:streptavidin binding. *Caption continues on next page.*

Figure 8.1 | (**h**) Example kymograph of DNA loop extrusion by cohesin that extrudes sequentially towards both directions. The loop position and size are shown below (black/purple points). Solid lines depict the raw data processed by a Savitzky–Golay filter of order 1 and window size of 51 frames (10.2 s). Segmented extrusion phases (Methods and chapter 5 [222]) are shaded yellow for extrusion upwards and blue for extrusion downwards. DNA was stained by 100 nM SYTOXOrange. (**i**) As in (h) but for a cohesin molecule extruding consistently towards one direction (upwards here). (**j**) Frequency with which extrusion phases occur in the absence or the presence of 2 μM CTCF NTR or CTR. White dots denote the mean, the box shows the quartiles of the data, and whiskers extend to 1.5·interquartile range. Statistical significance was assessed by a Kruskal–Wallis test (***: $p < 0.001$ and the Cliff's Delta statistic is ≤ 0.33). N = 45, 91, 40 DNA molecules from left to right. (**k**) Fraction of traces that show direction changes, i.e. extrusion in both directions, at least once during the trace for cohesin in the absence (control) or the presence of 2 μM CTCF NTR or CTR. Error bars denote the 95% binomial confidence interval. Statistical significance was assessed by a two-sided Fisher's exact test (**: $p < 0.01$ and the 95% odds ratio CI does not contain 1). N = 58, 103, 57 DNA molecules from left to right.

We also tested whether the CTR of CTCF affected cohesin since a previous report identified an interaction of the CTR with cohesin^{STAG2} [414]. Neither cohesin's stepping activity nor the fraction of observed consecutive steps and reversals was affected (Figure 8.1e,f and Figure S8.1k). Even though the authors of ref. [414] did not test for interactions between CTCF and cohesin^{STAG1} (instead of cohesin^{STAG2}) these results are in line with other recent studies that consistently found that the NTR, not the CTR, predominantly interacts with cohesin [35, 141, 251, 252, 393, 415] (chapter 7). We conclude that the CTCF NTR reduces the stepping activity of cohesin in a concentration-dependent manner (Figure 8.1f).

The MT findings were confirmed by our visualization assay. We tested whether the CTCF NTR or CTR would impact the characteristics of DNA loop extrusion by cohesin on long (48.5 kbp) SYTOXOrange-stained λ-DNA that is visualized by fluorescence HILO microscopy (Figure 8.1g; Methods) [98]. While MT traces of cohesin in the presence of the NTR exhibit a strong decrease in the number of consecutive steps and reversals (Figure 8.1d,f), cohesin still extruded loops on λ-DNA (Figure 8.1h,i; Figure S8.1l). However, the frequency with which cohesin engaged in active extrusion phases decreased roughly 3-fold in the presence of the CTCF NTR, but not for the CTR (comparing median values; Figure 8.1j; Methods). Once engaged in an extrusion phase, cohesin however extruded at a rate indistinguishable from the LE rate in the absence of any CTCF fragment (Figure S8.1m). The data thus indicate that the CTCF NTR inhibits cohesin-mediated LE by temporarily preventing cohesin from engaging in active loop-extrusion phases.

8.2 The CTCF NTR transforms cohesin into a unidirectional extruder

Human cohesin is, like the other eukaryotic SMC complexes, an asymmetric loop extruder [222] (chapter 5). However, cohesin can extrude DNA alternatingly from both sides, making it a bidirectional extruder (Figure 8.1h). This contrasts yeast condensin, which is a strictly unidirectional extruder, likely owing to the strong DNA binding affinity of condensin's 'safety belt' formed by Brn1 and Ycg1 [210, 221, 222] (chapter 5). When supplied with a two-fold excess of NIPBL-Mau2 (100 pM cohesin, 200 pM NIPBL-Mau2), cohesin extruded DNA bidirectionally in 35 ± 12% of the cases (Figure 8.1h,k and Figure S8.1l) and unidirectionally

in 65 ± 12% of the cases (Figure 8.1k; Figure S8.1l) as previously observed [222] (chapter 5). However, upon adding 2 μM CTCF NTR, the fraction of strictly unidirectionally extruding cohesin molecules increased to 87 ± 6% (Figure 8.1i,k). By contrast, this effect was absent when cohesin was exposed to the CTCF CTR (Figure 8.1k). Unexpectedly, the CTCF NTR thus renders cohesin to be a mostly unidirectional extruder.

8.3 The CTCF NTR impacts cohesin-driven DNA loop extrusion beyond the YDF motif

The CTCF NTR contains a conserved YDF motif at position 226-228 (Figure 8.3a, Figure S8.2a) which binds to the conserved essential surface (CES) formed by STAG2 and kleisin [141, 248, 407–410] and an analogous surface formed by STAG1 and kleisin [250]. To disentangle the individual contribution of the YDF motif to the CTCF NTR's interference with loop extrusion by cohesin, we used a very short fragment of the NTR containing the YDF motif (position 222-232; henceforth referred to as 'YDF peptide') as well as a version of this peptide in which the YDF motif was mutated to ADA (Figure 8.3b; henceforth referred to as 'ADA peptide'). Magnetic tweezers traces displayed an activity decrease of ~30% in the presence of 2 μM YDF peptide (Figure 8.3c-f). Furthermore, while the fraction of consecutive steps decreased, the fraction of reversals remained approximately constant (Figure 8.3e). In other words, the ratio of reversals to consecutive steps increased with increasing concentrations of the YDF peptide (Figure 8.3g). These effects were not observed when increasing concentrations of the ADA peptide were added instead of the YDF peptide (Figure S8.3b, Figure 8.3f,g).

To probe the effect of the YDF motif in the context of the full CTCF NTR, we also conducted MT experiments with two variants of the CTCF NTR in which the YDF motif was mutated to FDF (NTR^{FDF}) or ADA (NTR^{ADA}) (Figure 8.3b, Figure S8.3a,c,d). Similar to the CTCF NTR, both fragments caused an overall reduction in cohesin's stepping activity (Figure 8.3h). However, the effect became progressively weaker and noticeable only at higher concentrations after the mutation of Y226F and the combined mutation of Y226A and F228A. The diminishing effect of these mutations on cohesin can be attributed to the lower binding affinity (as compared to a short peptide comprising the YDF motif [141]), similar to reports of the FGF binding motif found in WAPL [407] and the ADA motif to CES^{STAG2} [141]. The fraction of consecutive steps and reversals were approximately equally reduced upon the addition of the NTR^{FDF} and the NTR^{ADA} (Figure 8.3i), similar to observations for the wild-type NTR (Figure S8.1k).

Summing up, the wild-type CTCF NTR reduces the occurrence of steps the most, followed by NTR^{FDF} and the YDF motif alone. Mutagenesis of Y226A and F228A (NTR^{ADA}) largely, but not entirely, alleviates the inhibition of cohesion by CTCF (Figure 8.3j). When exposed to the YDF motif alone, cohesin fails more frequently to take consecutive loop extrusion steps (Figure 8.3e) but maintains the frequency of reversals (Figure 8.3g,k). The data indicates that the strong reduction in cohesin activity observed in the presence of the CTCF NTR can *not* be solely attributed to the YDF motif. Firstly, the YDF peptide alone affects cohesin's stepping activity to a lesser extent than the wild-type NTR. Secondly, the

mutagenesis of Y226A and F228A (NTR^{ADA}) still retains an effect on cohesin. Lastly, the ratio of reversals to consecutive steps is strongly influenced by the YDF peptide but not by the full NTR, suggesting that this impact of the YDF motif on cohesin is masked in the wild-type NTR.

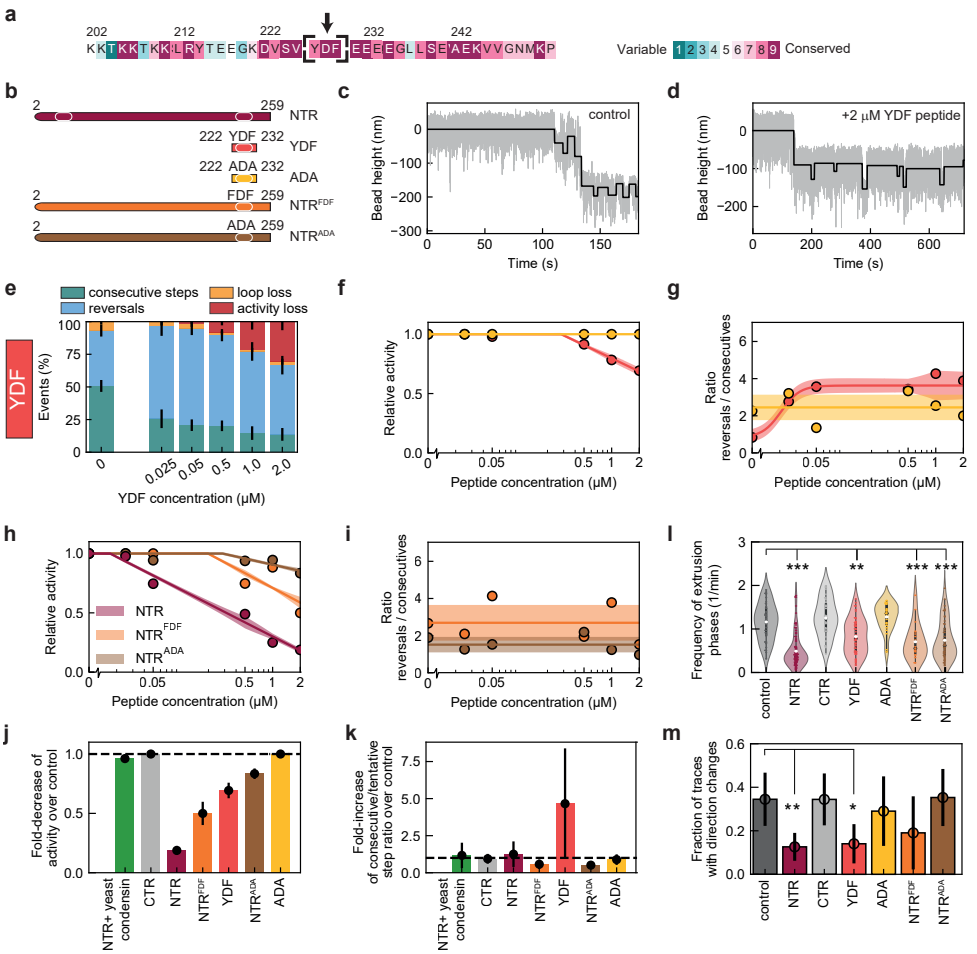


Figure 8.2 | CTCF's YDF motif yields more reversals and renders loop extrusion unidirectional.
(**a**) ConSurf [416, 417] conservation annotation of the human CTCF sequence surrounding the YDF motif. (**b**) Illustration of CTCF fragments used: short YDF and ADA peptides as well as CTCF NTR and corresponding mutations of the YDF motif. (**c**) Example MT trace of cohesin in the absence of any CTCF fragment. The light grey line denotes the raw data and the solid black line is the resulting stepfinder fit. (**d**) As in (c) for a trace in the presence of 2 μM YDF peptide. (**e**) The fraction of stepping events observed in MT traces classified into consecutive LE steps (green; downwards), reversals (blue; down-up step combination; Methods), and loop loss (yellow; upwards) with increasing concentrations of the YDF peptide. Error bars denote 95% binomial confidence intervals (CI). N = 463, 203, 337, 344, 139, 132 steps from 155, 97, 133, 125, 41, 44 tethers, respectively, for increasing YDF peptide concentration. (**f**) The relative stepping activity of cohesin, compared to the control ([peptide] = 0 μM), for increasing concentrations of the YDF and the ADA peptides. The data corresponds to the sum of consecutive LE steps, reversals, and loop loss (1-the red bar height) panel (e) and Figure S8.3b, respectively. The solid line depicts a piecewise linear fit (constant until a peptide concentration x_0 and linear fit with negative slope thereafter). The shaded error bar shows the interpolated error of the individual data points. (**g**) The ratio of reversal and consecutive steps for cohesin with increasing concentrations of the YDF and the ADA peptides (data shown in panel (e) and Figure S8.3b, respectively). No dependence of the reversal/consecutive ratio was noted for the ADA peptide, while the ratio increases for increasing concentrations of the YDF peptide. For the YDF peptide, the data was fitted to a function of the form $\frac{-A}{1+e^{-log(x/x_0)/k}} + O$ and serves as a guide for the eye. For the ADA peptide, the horizontal line and the shaded error bar denote the mean and SD of the individual data points.
Caption continues on next page.

Figure 8.2 | (**h**) As in panel (f) for cohesin but in the presence of increasing concentrations of CTCF NTR^{FDF} and NTR^{ADA}. The corresponding data is shown in Figure S8.3c,d. (**i**) As in panel (g) for cohesin but in the presence of increasing concentrations of NTR^{FDF} and NTR^{ADA}. The corresponding data is shown in Figure S8.3c,d. (**j**) Overview of the activity loss of cohesin quantified as the fraction of the number of steps that were not observed with 2 μM peptide (or the highest measured concentration) compared to the control in the absence of the peptide. Data points and error bars are the same as shown in Figure 8.1d-e, Figure S8.1j, Figure 8.3e, and Figure S8.3b-d. (**k**) Overview of the ratio of reversals vs. consecutive steps compared to the control. Data points are the same as shown in Figure S8.1k, Figure 8.3g but each data point was divided by the ratio of the respective control measurement and the displayed error was calculated according to error propagation. (**l**) Frequency with which DNA extrusion phases occur in the absence or the presence of 2 μM of the indicated CTCF fragment. White dots denote the mean, the box shows the quartiles of the data, and whiskers extend to 1.5·interquartile range. Statistical significance was assessed by a Kruskal–Wallis test (**: $p < 0.01$; ***: $p < 0.001$ and the Cliff's Delta statistic is ≤ 0.33). N = 45, 91, 40, 52, 25, 19, 49, DNA molecules from left to right. (**m**) Fraction of traces that show direction changes in the absence (control) or the presence of 2 μM of the indicated CTCF fragment. The bar height denotes the fraction of traces and the error bar denotes the 95% binomial confidence interval. Statistical significance was assessed by a two-sided Fisher's exact test (*: $p < 0.05$; **: $p < 0.01$ and the 95% odds ratio CI does not contain 1). N = 58, 103, 57, 57, 32, 21, 51 DNA molecules from left to right.

8.4 The YDF motif transforms cohesin into a unidirectional extruder

Next, we tested the impact of the YDF motif on the dynamics and directionality of loop extrusion by cohesin on λ-DNA using fluorescence HILO microscopy. The frequency with which cohesin engaged in active extrusion phases was reduced by about 1.5-fold upon adding the YDF peptide and roughly 2-fold upon adding the NTR^{FDF} and NTR^{ADA} fragments (Figure 8.3l). However, the rate at which DNA was extruded was indistinguishable from control measurements (Figure S8.3e). The fact that the NTR^{ADA} fragment but not the ADA peptide affected the frequency of extrusion phases confirms that cohesin is not solely affected by CTCF's YDF motif.

Strikingly, the addition of the YDF motif alone rendered cohesin to extrude mostly unidirectionally, mimicking the effect of the wild-type NTR (Figure 8.3m). The Y226F mutation caused an intermediate (yet not statistically significant) effect, while the Y226A/F228A mutations as well as the ADA peptide entirely alleviated the effect of the YDF peptide and the wild-type NTR on cohesin's loop extrusion directionality. These results indicate that the YDF motif alone converts cohesin into a predominantly unidirectional extruder. However, the YDF motif cannot fully account for the reduced frequency with which cohesin engages in active loop extrusion phases.

8.5 The N-terminal KTYQR motif inhibits DNA loop extrusion by cohesin

Nora et al. [252] scanned CTCF for regions that affect the insulation of TADs *in vivo* by generating mutant cell lines of CTCF truncations and deletions and performing 5C analysis. Regions that affected TAD insulation the most encompassed the YDF motif, an RNA-binding and zinc finger domain (NTR264-288), and the N-terminal region NTR1-89. We focused on examining NTR1-89 (since the zinc finger-containing fragment NTR264-288

can bind to DNA in our assays). In this region, the conserved KTYQR motif at position 23-27 (Figure 8.3a) constitutes a binding site to PDS5A [252]. Even though PDS5A was not included in our experiments, we tested whether a short fragment containing the KTYQR motif (amino acids 16-31, henceforth referred to as KTYQR peptide) would affect DNA loop extrusion by cohesin.

Upon addition of KTYQR peptide, cohesin exhibited a reduced stepping activity (Figure 8.3c-f) but no change in the ratio of consecutive steps and reversals (Figure 8.3g), reminiscent of the effect of wild-type NTR on cohesin (Figure 8.1d,f). To probe the effects of the NTR in the absence of the KTYQR motif on cohesin, we expressed and purified an N-terminal truncation of the NTR ($NTR^{\Delta 1\text{-}89}$) as well as the NTR in which the KTYQR motif was mutated to alanines ($NTR^{23\text{-}27A}$; Figure 8.3b and Figure S8.4a). Both constructs did not affect cohesin's overall stepping activity (Figure 8.3f, Figure S8.4b,c), but caused cohesin to take proportionally more reversals than consecutive steps (Figure 8.4g). The absence of the KTYQR motif in CTCF's NTR thus mimics the effect of the YDF motif on cohesin (Figure 8.3g).

We conclude that both CTCF's YDF and KTYQR motifs distinctly affect cohesin's ability to take consecutive loop extrusion steps. In isolation, both decrease cohesin's overall activity (Figure 8.3h). However, only the YDF motif selectively hampers cohesin's ability to take consecutive LE steps while the occurrence of reversals remains unaffected (Figure 8.3e and 3I). In the absence of the KTYQR motif in CTCF's NTR (i.e., for $NTR^{\Delta 1\text{-}89}$ and $NTR^{23\text{-}27A}$), consecutive steps were also suppressed, mimicking the effect of the YDF motif.

The data indicates that the effects of the YDF and KTYQR motifs are not simply additive. While the wild-type NTR inhibits 81 ± 4 % of cohesin-induced LE steps, the YDF and KTYQR motifs in isolation inhibit 31 ± 7 % and 34 ± 7 % of steps, respectively (and thus the sum of the impact of the YDF and KTYQR motifs adds up to 64 ± 10 %). Notably, while the YDF peptide selectively suppresses consecutive LE steps, the occurrence of both consecutives and reversals is reduced upon the addition of the wild-type NTR which comprises both the YDF and KTYQR motifs (Figure 8.3i). Mutagenesis of the YDF motif within the NTR (NTR^{FDF} and NTR^{ADA}) partially alleviates the strong impact of the wild-type NTR but retains a decrease in cohesin activity.

The frequency of active loop extrusion phases by cohesin on λ-DNA was roughly 2-fold reduced in the presence of 2 μM KTYQR peptide (Figure 8.3j), similar to the reduction in LE frequency in response to mutagenesis of the YDF motif (NTR^{FDF} and NTR^{ADA}). However, deletion of the N-terminal 89 amino acids ($NTR^{\Delta 1\text{-}89}$) and mutagenesis of the KTYQR motif ($NTR^{23\text{-}27A}$) caused only a modest decrease in the extrusion frequency of 20% and 25%, respectively (comparing median values). The directionality of DNA loop extrusion by cohesin was found to be unaffected by the KTYQR peptide (Figure 8.3k), i.e. bi-directional loop extrusion was observed. In line with this finding, the N-terminal truncation ($NTR^{\Delta 1\text{-}89}$) and $NTR^{23\text{-}27A}$ made cohesin a unidirectional extruder, similar to the wild-type NTR and

the YDF peptide. We conclude that the YDF but not the KTYQR motif affects cohesin's LE directionality.

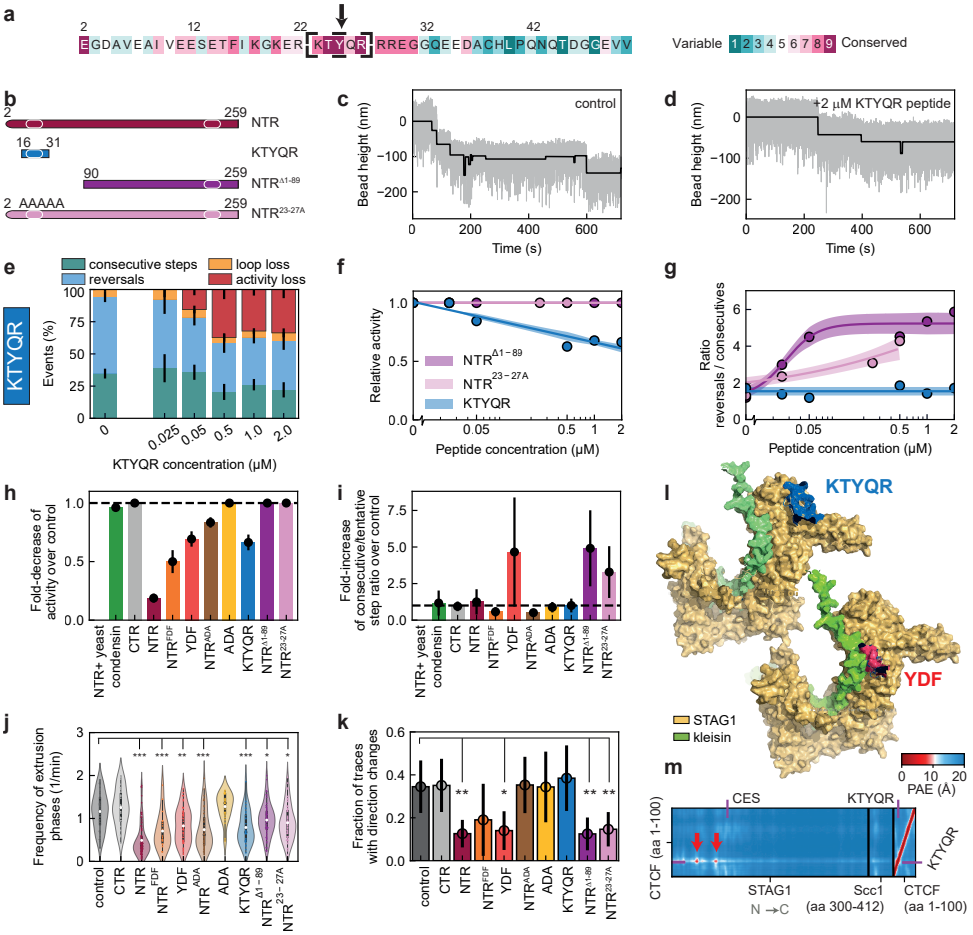


Figure 8.3 | CTCF's KTYQR motif reduces cohesin's stepping activity.
(**a**) ConSurf [416, 417] conservation annotation of the human CTCF sequence surrounding the KTYQR motif. (**b**) Illustration of the short KTYQR peptide as well as CTCF NTR and corresponding truncation and mutation of the KTYQR motif therein. (**c**) Example MT trace of cohesin in the absence of any CTCF fragment. The light grey line denotes the raw data and the solid black line is the resulting stepfinder fit. (**d**) As in (c) but for a trace in the presence of 2 μM KTYQR peptide. (**e**) The fraction of stepping events observed in MT traces classified into consecutive LE steps (green; downwards), reversals (blue; down-up step combination; Methods), and loop loss (yellow; upwards) with increasing concentrations of the KTYQR peptide. Error bars denote 95% binomial confidence intervals (CI). N = 599, 90, 215, 101, 257, 126 steps from 295, 68, 90, 49, 185, 64 tethers, respectively, for increasing peptide concentration. (**f**) The relative stepping activity of cohesin, compared to the control ([peptide] = 0 μM), for increasing concentrations of CTCF NTR$^{\Delta 1\text{-}89}$, NTR$^{23\text{-}27A}$, and the KTYQR peptide. The data to the sum of consecutive LE steps, reversals, and loop loss (1-the red bar height) of panel (e) and Figure S8.4b,c, respectively. The solid line depicts a piecewise linear fit (constant until a peptide concentration x_0 and linear fit with negative slope thereafter). The shaded error bar shows the interpolated error of the individual data points. (**g**) The ratio of reversal and consecutive steps for cohesin with increasing concentrations of CTCF NTR$^{\Delta 1\text{-}89}$, NTR$^{23\text{-}27A}$, and the KTYQR peptide (data shown in panel (e) and Figure S8.4b,c, respectively). No dependence of the reversal/consecutive ratio (Methods) was noted for the KTYQR peptide while the ratio increases for increasing concentrations of the CTCF NTR$^{\Delta 1\text{-}89}$ and NTR$^{23\text{-}27A}$ fragments. For the CTCF NTR$^{\Delta 1\text{-}89}$ and NTR$^{23\text{-}27A}$ fragments, the data was fitted to a function form $\frac{-A}{1+e^{-\log(x/x_0)/k}} + O$ and serves as a guide for the eye. For the KTYQR peptide, the horizontal line and the shaded error bar denote the mean and SD of the individual data points. *Caption continues on next page.*

Figure 8.3 | (**h**) Overview of the activity loss of cohesin quantified as the fraction of the number of steps that were not observed with 2 μM peptide (or the highest measured concentration) compared to the control in the absence of the peptide. Data points and error bars are the same as shown in Figure 8.1d-e, Figure S8.1j, Figure 8.3e, Figure S8.3b-d, Figure 8.3e, and Figure S8.4b,c. (**i**) Overview of the ratio of reversals vs. consecutive steps compared to the control. Data points are the same as shown in Figure S8.1k, Figure 8.3g, Figure 8.3g but each data point was divided by the ratio of the respective control measurement and the displayed error was calculated according to error propagation. (**j**) Frequency with which extrusion phases occur in the absence or presence of 2 μM of the indicated CTCF fragment. White dots denote the mean, the box shows the quartiles of the data, and whiskers extend to 1.5·interquartile range. Statistical significance was assessed by a Kruskal–Wallis test (**: $p < 0.01$; ***: $p < 0.001$ and the Cliff's Delta statistic is *le* 0.33). N = 45, 40, 91, 19, 52, 49, 25, 29, 62, 58 DNA molecules from left to right. (**k**) Fraction of traces that show direction changes in the absence (control) or the presence of 2 μM of the indicated CTCF fragment. The bar height denotes the fraction of traces and the error bar denotes the 95% binomial confidence interval. Statistical significance was assessed by a two-sided Fisher's exact test (*: $p < 0.05$; **: $p < 0.01$ and the 95% odds ratio CI does not contain 1). N = 58, 57, 103, 21, 57, 51, 32, 39, 72, 75 DNA molecules from left to right. (**l**) AlphaFold prediction of STAG1, kleisin fragment amino acids 300-412, and CTCF fragment #1 (spanning amino acids 1-100), containing the KTYQR motif which is shown in blue, or CTCF fragment #3 (spanning amino acids 180-264), containing the YDF motif. The KTYQR/YDF motif of one representative prediction is shown. See Figure S8.4f for the placement of the motifs in 25 independent predictions. (**m**) Predicted Alignment Error (PAE) map of CTCF fragment #1, STAG1, and kleisin fragment amino acids 300-412. The position of the KTYQR motif is indicated by the purple tick.

8.6 An AlphaFold screen identifies a binding pocket of the KTYQR motif on STAG1/2

The MT data strongly suggest that CTCF's KTYQR motif interacts directly with cohesin. Notably, our experiments did not include PDS5A, which was previously identified as a KTYQR-binding protein [252]. To identify potential binding targets of the KTYQR motif on cohesin, we computationally screened cohesin subunits and fragments of the CTCF NTR for interactions using AlphaFold-Multimer [311, 418]. To this end, we segmented cohesin into various subunits (NIPBL-Mau2, STAG1/2, ATPase heads, and the hinge including respective segments of Scc1; see Methods; Table 8.2 and Table 8.3) and segmented CTCF into three overlapping fragments (#1: 1-100 (including KTYQR); #2: 90-190; #3 180-264 (including YDF); Table 8.2).

AlphaFold only predicted interactions between STAG1/2 and CTCF fragments #1, between STAG1/2 and fragment #3 (Figure 8.3l,m Figure S8.4e-h), as well as between the Smc hinge and fragment #3 (Figure S8.4i). Interactions between STAG1/2 and fragment #3 were mediated between the CES and the YDF motif (Figure S8.4e) as previously experimentally resolved [141]. In 16/25 and 25/25 models for STAG1 and STAG2, respectively, the KTYQR motif was placed N-terminally of the CES on a conserved surface (ConSurf conservation scale 8/9 [416, 417]; Figure S8.4e-g). Even though the predicted local distance difference test (pLDDT, a measure of local confidence [311]) remained below 50 (typical for intrinsically disordered regions [311]), AlphaFold placed the KTYQR motif onto STAG1 with low structural variation in 16/25 models (Figure S8.4h). Unfortunately, AlphaFold does not permit to predict binding affinities of the KTYQR and YDF motifs on STAG1/2. It is possible that KTYQR motif binds only transiently or with low affinity to STAG1/2, which would be challenging to detect in bulk [141] or with *in vivo* binding assays [252].

Beyond interactions of the KTYQR and YDF motifs with STAG1/2, AlphaFold predicted an interaction between the YDF motif and the Smc hinge in 18/25 models (Figure S8.4i-k). This unexpected finding implies that the YDF motif, while it can bind STAG1/2, potentially also modulates loop extrusion by cohesin as well as the opening of the Smc hinge to modulate cohesion [224, 291, 375, 419]. Notably, Nagasaka et al. [139] described a separation-of-function mutant ($SMC1^{K540D/K637D/K648D}$; henceforth referred to as $SMC1^{3D}$) which supports DNA loop extrusion but is defective in cohesion, similar to other hinge mutants previously identified in yeast [291]. Curiously, $SMC1^{3D}$ was also deficient in forming CTCF-anchored loops, but not TADs, *in vivo*. It is tempting to speculate that an interaction between the cohesin hinge and YDF, as predicted by AlphaFold, contributes to the formation of CTCF-anchored loops *via* a direct YDF-hinge interaction. Possibly, this could also modulate cohesin's tendency to undergo unsuccessful LE steps (reversals) because of a change in the DNA binding affinity of the Smc hinge which is crucial to support loop extrusion on λ-DNA [238].

8.7 Conclusions

CTCF halts DNA loop extrusion by cohesin [32–35, 39, 41, 153, 153, 158, 393] (chapter 7), which is largely mediated by YDF-CES binding of STAG1/2 *in vivo* [141]. However, *how* CTCF affects cohesin remained so far unclear. To answer this, we studied the effects of the CTCF NTR as well as the individual contributions of the YDF and KTYQR motifs found therein on DNA loop extrusion by cohesin *in vitro* (Figure 8.4). The YDF motif causes cohesin to be less active compared to control experiments, strongly reduces the occurrence of consecutive loop-extrusion steps, and renders LE unidirectional. The KTYQR motif reduces cohesin's stepping activity. The full NTR displays a strong activity decrease and renders extrusion unidirectional. In contrast to the YDF motif, however, it reduces reversals and consecutive steps similarly. This shows that the effect of the full NTR lies beyond the contribution of the individual YDF and KTYQR motifs, as effects that are prevalent for one motif in isolation appear to be modulated in the context of the whole NTR. Mutating the YDF motif of the NTR has a similar effect on cohesin as the KTYQR peptide alone. However, truncating or mutating the KTYQR motif within the NTR does not entirely recapitulate the effect of the YDF peptide in isolation.

Because we did not observe any effects of the CTCF CTR on cohesin, we conclude that direct protein-protein interactions between cohesin and CTCF are mediated by CTCF's NTR and not its CTR. However, we have previously observed a non-zero probability of CTCF halting DNA loop extrusion by cohesin when CTCF is approached from its C-terminal side *in vitro* [35] (chapter 7). The data presented here suggests that, even though CTCF was approached from its C-terminal side, cohesin interacted with CTCF's N-terminus in these instances. In support of this hypothesis, the residence times of cohesin and CTCF were indistinguishable between N- and C-terminal encounters [35] (chapter 7).

The YDF motif renders cohesin to predominantly extrude DNA unidirectionally. This observation possibly relates to our previous finding that cohesin, if not halted and 'pushed back' by CTCF, extrudes away from CTCF while remaining anchored there [35] (chapter 7).



Figure 8.4 | Overview of the effects of the experimentally tested CTCF NTR fragments on cohesin-mediated DNA loop extrusion.

The anchoring of cohesin *via* the CES-YDF interaction could ensure that cohesin does not repeatedly attempt to pass the already encountered CTCF molecule but instead extrudes toward the other CTCF binding site delineating a cohesin-dependent TAD or CTCF-CTCF loop.

Further experiments are currently being conducted to pinpoint the mechanism by which the NTR and the YDF/KTYQR motifs cause the effects observed here. First, previous studies did not observe an association between STAG1/2 and the KTYQR motif in bulk and *in vivo* experiments [141, 252]. Because the interaction might be transient, single-molecule binding studies (e.g. Fluorescence Correlation Spectroscopy) of a fluorescently labeled KTYQR peptide will be conducted. Furthermore, it remains unclear how the YDF and KTYQR motifs modulate the overall cohesin stepping activity in MT traces. We are currently conducting measurements of cohesin's ATPase rate in the absence and presence of the YDF and KTYQR motifs to elucidate whether these motifs directly affect cohesin's ability to hydrolyse ATP. It also remains unclear why and how the individual contributions of the YDF and KTYQR motifs are modulated in the wild-type NTR. To clarify whether this originates from the polymeric nature of the CTCF NTR (both motifs lie on the same peptide), we are conducting MT experiments in which the YDF and KTYQR peptides are added simultaneously. This experiment could also indicate contributions of the NTR beyond the YDF and KTYQR motifs that we cannot readily exclude.

Fluorescence experiments showed that the YDF motif strongly impacts cohesin's extrusion directionality. Two possible molecular mechanisms are conceivable to explain this observation. Since we previously observed that the relative abundance of NIPBL-Mau2 to cohesin affects cohesin's directionality [222] (chapter 5), it is possible that the YDF motif modulates NIPBL-Mau2 turnover on cohesin which would 'lock' cohesin in one direction. However, to switch extrusion directions, DNA binding at both the motor (analogous to NIPBL-Mau2) as well as the 'anchor' side (analogous to STAG1) has to be weakened. However, the AlphaFold screen did not predict any interaction between the YDF motif and NIPBL-Mau2.

Instead, the interaction between the YDF motif and STAG1/2 [141] could enhance the affinity of STAG1/2-Scc1 to DNA, preventing direction changes. This hypothesis is in line with experiments in which the anchor subunit of yeast condensin (Ycg1) was deleted, which effectively lowered condensin's DNA binding strength at the anchor side and caused yeast condensin to extrude DNA bidirectionally [221]. To test this hypothesis, we are thus conducting DNA binding strength measurements of STAG1-Scc1 in the absence and presence of the YDF motif. Lastly, the prediction that the YDF motif binds the Smc hinge could relate to the reduced ability of CTCF to halt cohesin *in vivo* upon mutation of three basic residues in the Smc hinge ($SMC1^{3D}$) [139] if the $SMC1^{3D}$ indeed shows a reduced affinity to the YDF motif. However, these hypotheses require to be experimentally tested.

Our study reveals that CTCF's N-terminal region modulates cohesin's DNA loop extrusion activity and directionality through at least its YDF and KTYQR motifs via direct protein-protein interactions. However, it remains unclear how and at which stage these and possibly other cohesin-interacting motifs affect the DNA loop extrusion cycle. Experiments designed to capture the temporal sequence of conformational changes leading to loop extrusion are needed to answer this question in the future.

8.8 Methods

Human cohesin and NIPBL-Mau2 expression and purification

Wild-type human cohesin and NIPBL-Mau2 were expressed in and purified from *Sf9* insect cells as described previously [173].

Yeast condensin expression and purification

Wild-type yeast condensin was expressed and purified from *S. cerevisiae* as described previously [98].

Expression and purification of CTCF fragments

Various fragments of human CTCF (Uniprot #P49711, see Table 8.1) were cloned into a pMAL-C5X derived plasmid, resulting in MBP fusions with a flexible linker and a 3C protease site at the fusion point (MBP-NSSSNNNNNNNNNNLGIEGRISHMSMGGRDIVDGSEF-LEVLFQ ↓GP-CTCF). Plasmids were transformed into *E. coli* ER2566 cells (New England Biolabs, fhuA2 lacZ::T7 gene1 [lon] ompT gal sulA11 R(mcr73::miniTn10–TetS)2 [dcm] R(zgb-210::Tn10–TetS) endA1 Δ(mcrCmrr)114::IS10) containing a Rosetta ™-2 plasmid (Novagen), and selected on LB media containing 100 µg/ml ampicillin and 34 µg/ml chloramphenicol. For expression, overnight pre-cultures were diluted 1:100 into 1 liter prewarmed LB with fresh antibiotics, cultures were grown in baffled flasks shaking at 200 rpm at 37°C until the OD600 reached ~0.6, and expression was induced by the addition of IPTG to a final concentration of 1 mM. After 3 hours of induction, the cells were harvested by centrifugation (8 min, 4500 rpm, JLA8.1000 rotor), washed in 50 ml PBS, and resuspended in buffer A (50 mM Tris-HCl pH 7.5, 300 mM NaCl). Cells were lysed using a French Press (Constant Systems) at 20 kpsi, 4°C and unbroken cells, debris and aggregates were pelleted in a Ti45 rotor (30 min, 40.000 rpm, 4°). The lysate was applied to 2 ml prewashed amylose resin (NEB #E8021) and incubated for one hour while rotating at 4°C. Subsequently, the resin was washed with 50 ml of buffer A and finally, CTCF fragments were eluted with 15 ml of buffer A supplemented with 1 mM β-mercaptoethanol and homemade 3C protease. CTCF fragments were concentrated using a Vivaspin centrifugal concentrator (10 kDa cut-off) and further purified by size exclusion chromatography (SEC) on a Superdex 200 Increase 10/300 column pre-equilibrated with buffer B (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 5% (w/v) glycerol, 0.05 mM TCEP).

To purify the CTCF CTR (10xHis-CTCF(578-727)-TEV-Halo) and NTR^{ADA} (10xHis-CTCF(2-259)-TEV-Halo), baculovirus-infected insect *Sf9* cell pellets from cultures supplemented with 0.1 mM $ZnCl_2$ were lysed by Dounce homogenization and resuspended in CTCF buffer (35 mM Hepes pH 7.5, 350 mM NaCl, 5% glycerol) supplemented with 0.1 mM $ZnCl_2$, 0.05% Tween-20, 5 mM imidazole, 1 mM PMSF, EDTA-free cOmplete tablet (1 per 50 ml) (Roche, 11873580001), 3 mM betamercaptoethanol, 10 µg/ml aprotinin, 2 mM benzamidine and benzonase. The lysate was cleared by centrifugation at 18,000g for 45 min at 4°C. The soluble fraction was incubated with NiNTA agarose (Qiagen, 30230) for 90 min at 4°C and washed first with CTCF buffer supplemented with 0.1 mM $ZnCl_2$, 0.01 % Tween-20, 35 mM imidazole, 1 mM PMSF and 2 mM benzamidine and then with CTCF buffer supplemented with 35 mM imidazole only. Protein was eluted with CTCF buffer supplemented with 300 mM imidazole. The eluate was concentrated using a Vivaspin 30

kDa MWCO concentrator (Sartorius) and buffer was exchanged using a Zeba 7 kDa MWCO desalting column equilibrated in CTCF buffer. Protein was then flash-frozen and stored at -80°C.

In addition to these longer protein fragments, smaller fragments were synthesized in-house or ordered from Genscript (Table 8.1). Peptides were synthesized on a Liberty Blue peptide synthesizer (CEM) using standard Fmoc chemistry. For each amino acid cycle, a 4 min coupling with DIC/Oxyma was performed. Peptides were purified on a Phenomenex Luna C18(2) using a 2-45% in 45 min 0,1%TFA/ACN+0,1%TFA gradient. The identity of the peptide was confirmed using MALDI-MS (4800 MALDI TOF/TOF, Sciex)

Synthesis of dsDNA construct for fluorescence-based loop extrusion experiments

The synthesis of the dsDNA construct based on λ phage DNA was described previously [98]. Briefly, single-stranded DNA ends of λ phage DNA (D1521, Promega) were filled with DNA polymerase I large fragment (M0210, New England Biolabs) lacking exonuclease activity, in the presence of dTTP, dCTP, dGTP, and biotin-14-dATP (19524016, Invitrogen). The DNA construct was purified from excess nucleotides and polymerase using a PCR clean-up kit (20021, Qiagen) and stored at −20°C until use. Once an aliquot was thawed, it was kept at 4°C.

Synthesis of dsDNA construct for magnetic tweezers loop extrusion experiments

Singly biotinylated, linear dsDNA constructs with a length of 1.5 kbp were synthesized via PCR and enzymatically ligated to digoxigenin-enriched DNA handles, as described previously [98]. Briefly, a biotinylated DNA fragment of 1.5 kbp length was produced by using biotin-labeled forward and reverse primers that contain a BsaI restriction site on pBluescript II SK+. To create the digoxigenin-enriched handle, a 485 bp fragment from pBluescript II SK+ (Stratagene, Agilent Technologies Inc., USA) was amplified by PCR in the presence of 1:5 digoxigenin-11-dUTP:dTTP (Jena Bioscience, Germany). Prior to ligations of the DNA fragment and handle, the amplicons were digested with the non-palindromic restriction enzyme BsaI-HFv2 (New England Biolabs, UK). The ligation of the DIG-handle and biotinylated DNA fragment was carried out overnight using T4 DNA ligase (New England Biolabs, UK). The final dsDNA construct was cleaned up from the access of handle by running on a 1% agarose gel and extracting the dsDNA construct using a gel purification kit (A9282, Promega) and stored at -20°C until use.

Measuring DNA loop extrusion activity and steps with magnetic tweezers

The magnetic tweezers instrument and experimental methodology used in this study was previously described in detail [217]. Bead x, y, z position tracking was achieved with a spatial resolution of ~2 nm. The concentration of cohesin was 100 pM together with 250 pM NIPBL-Mau2; the varying concentrations of the different CTCF fragments that were additionally added are denoted within the manuscript text. Experiments were conducted at 22.3°C in a buffer containing 40 mM Tris pH 7.5, 40 mM NaCl, 2.5 mM $MgCl_2$, 1 mM DTT,

1 mM ATP, 0.05% Tween-20, 0.25 mg/ml BSA. Experiments with yeast condensin were conducted with 1.5-2 nM condensin at 22.3°C in a buffer containing 40 mM Tris pH 7.5, 40 mM NaCl, 2.5 mM $MgCl_2$, 1 mM DTT, 1 mM ATP, 0.25 mg/ml BSA. Data was acquired for 12 min after protein flush-in and lowering the applied force from 7 pN to 0.3 pN. Datasets were processed using custom-written Igor v6.37-based scripts, removing traces that showed surface-adhered magnetic beads, ruptured tethers, as well as tethers containing more than one dsDNA [217]. The bead Z-positions of all traces conducted at identical conditions were pooled and filtered to 1 Hz (moving average) for subsequent quantitative loop extrusion-step analysis. The step-finding procedure is based on the quantitative method described by Smith in 1998 for detecting edges in noisy time-series data [420]. The algorithm was first made accessible for utilization in the Python programming environment within the framework of the stepfinder Python package developed for the identification of steps in one-dimensional data with low signal-to-noise ratio [421] and subsequently refined to suit the specific requirements of single-molecule force spectroscopy experiments. In brief, the algorithm enables the identification of steps with consistent reliability despite variable noise levels in the data as follows: at each time point x_i, the difference between the forward and backward estimation $\Delta_i = X_{i+} - X_{i-}$) of a Chung-Kennedy filter [403] applied to the time-series data is compared to the local root-mean-square (RMS) noise, calculated from the weighted moving variances $w_{i\pm}s_{i\pm}$) over the same window:

$$Y_i = \frac{\Delta_i}{\sqrt{w_{i+}s_{i+} + w_{i-}s_{i-}}},$$

where the weights w correspond to the normalized switching factors of the Chung-Kennedy filter and are defined as

$$w_{i\pm} = \frac{s_{i\pm}^{-p}}{s_{i\pm}^{-p} + s_{i-}^{-p}},$$

with p being the nonlinearity factor (edginess) proposed by Chung and Kennedy. In practice, $p = 20$ was chosen.

8

The noise component in the denominator is a function of the experimental conditions, e.g. magnetic tweezer force and DNA length. It consists of the weighted moving variance calculated to the mean of the data and averaged over a window and does not exhibit peaks at step locations, unlike the standard deviation. When an edge is encountered in the data, the output function Y responds with a triangular waveform peaking at the location of the edge itself. The benefit of this edge detector is that its probability distribution is known, allowing for the calculation of confidence limits for the presence and absence of a step at each time point [420]. This relationship is leveraged to define an initial threshold for the detection of a peak in Y

$$y_c = \frac{2}{3} \cdot \frac{\text{minimum step size}}{\text{local noise}} \tag{8.1}$$

which ensures that the probability P_T for the detection of a true step (efficiency) is maintained at a constant level. It can be shown that assuming an SNR of 1 and thus choosing $y_c = 2/3$ leads to the detection of steps with an efficacy of 95%.
In addition, a downstream rejection procedure was applied to verify whether Equation 8.1

is satisfied for each individual detected step. For this purpose, the local step noise was computed within an interval around the index of the event, which was determined by the minimum allowed step spacing corresponding to two times the filter window. Based on this noise component and y_c, a dynamically adapting acceptance threshold value (in units of the input signal) for the current slice of the time trace was back-calculated corresponding to the minimum allowed step size in Equation 8.1. Above-threshold steps were considered valid, while below-threshold steps were rejected and the neighboring plateaus were merged. After each rejection, the quality of the remaining events was reassessed in an iterative procedure until the stepfinder output consisted of valid steps only (see figure S1). To address the problem that statistical predictors such as the probability of true and false outputs and location accuracy of the stepfinder are also a function of the filter window width W, the detector was run with a sequence of 50 window widths between 0.5 s and 5.0 s and the results were analysed. After determining the value W_{opt} which yielded the minimum standard deviation of the step mass (step size over noise for each event) and the maximum ratio of valid over rejected steps, the final window size was conservatively set to $\frac{4}{5}W_{opt}$) and the final step-finding analysis was conducted. This procedure, named *LowForceAutoStepfinder* (LFAS) was evaluated against a previously employed method to identify steps in MT traces of SMC-mediated DNA loop extrusion (*AutoStepfinder*) [217]. To this end, we computed the true positive rate (TPR) of detected steps using step validation data as conducted previously [217]. In brief, 1.5 kbp-long dsDNA was tethered between the glass surface and magnetic bead as described above and a force of 0.3 pN or 1 pN was applied. The piezo holding the objective was set to step up or down every 10 seconds for a range of step sizes (Figure S8.1d,e). Both stepfinder algorithms were applied to these traces. These datasets were then compared to the known steps from the piezo motion. In line with the previous stepfinder evaluation [217], a step detection was judged correct if there was a piezo step nearby within 10% (~1 s) of the expected dwell time and 30% of the expected step size. We found that the LFAS has a higher TPR at smaller forces (in particular at 0.3 pN used in this study) and smaller induced step size (Figure S8.1f) and was thus used throughout this study.

Measuring DNA loop extrusion activity with fluorescence

DNA loop extrusion experiments and imaging were performed as described previously [222] (chapter 5). DNA Loop extrusion reactions by human cohesin were carried out in a buffer containing 40 mM Tris-HCl pH 7.5, 30 mM NaCl, 2.5 mM $MgCl_2$, 2.5% D-glucose, 2 mM Trolox, 10 nM catalase, 18.75 nM glucose oxidase, 100 nM Sytox Orange, 0.5 mg/ml BSA, 1 mM DTT, 1 mM ATP with 100 pM cohesin and 200 pM NIPBL-Mau2 at 37°C. All data were acquired using an exposure time of 100 ms.

Quantification and segmentation of kymographs

Kymographs were quantified and segmented as described in [222] (chapter 5). In brief, images were cropped and processed by a median filter with a window size of 5 frames and the background was subtracted using a Top-Hat filter of size 10 pixels. The intensity of the DNA loop is normalized to the intensity along the entire DNA molecule and multiplied by the known length of the DNA molecule (48.5 kbp). The DNA loop position was determined as the relative position of the loop from one end of the DNA. Segmentation of cohesin-

mediated DNA loop extrusion phases was performed semi-automatically under assistance of a change point detection algorithm that was subsequently manually curated since an entirely automatic segmentation was not robust and returns spurious results on some traces or on parts of traces. After filtering, the change point detection was performed using a window-based change point detection algorithm [422], which were subsequently merged and classified based on the error estimation of the assay [35, 222] (chapter 7, chapter 5). The direction of each DNA loop extrusion and DNA loop slipping phase was subsequently annotated by subjecting the DNA loop position over time within each segment to a linear regression and classifying the segment as direction1/direction2 based on the slope of the linear fit. Note that the nomenclature of the direction is arbitrary. Here, the first direction to which the loop travels is called direction 1. A loop extrusion trace was labeled as unidirectional if all DNA loop extrusion phases in the trace pointed in the same direction and as bidirectional otherwise. The frequency of DNA loop extrusion phases was computed by dividing the number of DNA loop extrusion phases by the time between the start of the first and last DNA loop extrusion phase of every trace. The direction switching frequency was computed by dividing the number of direction switches by the time between the start of the first and last DNA loop extrusion phases of every trace. The displayed DNA loop size and position traces (Figure 8.1h-i, Figure S8.1l) were filtered using a Savitzky–Golay filter with a window length of 51 frames (5.1 s) and order 1 (implemented as part of the *scipy* package [299]).

Fitting of activity reduction and reversal/consecutive step ratio data

To fit the activity reduction in dependence of the peptide concentrations from MT experiments, the correlation between the peptide concentrations and the activity reduction data was first assessed using Spearman's correlation coefficient (function scipy.stats.spearmanr of the scipy package [299]). If the correlation coefficient was below 0.8, the data points were not subject to a fit, but described by a horizontal line with a height equal to the average value of the data points and the error bar equals the ±SD of the data points (e.g. CTR and NTR+condensin data in Figure 8.1f). If the Spearman's correlation coefficient exceeded 0.8, a piecewise linear model was fitted to the data that consisted of a horizontal line up to a concentration x_0 and a line with a negative slope thereafter in logarithmic space (e.g. NTR data in Figure 8.1f), i.e.

$$y(x) = \begin{cases} B & x \leq x_0 \\ B - m\log(x/x_0) & x > x_0 \end{cases}$$

where B is an offset, m denotes the slope of the line in logarithmic space, and x_0 denotes the concentrations at which the horizontal line transitions into the line with a negative slope. To fit the ratio of reversal to consecutive steps in dependence of peptide concentrations, the same approach was used, but the data points were fitted to the following function if the Spearman's correlation coefficient exceeded 0.8 (e.g. YDF data in Figure 8.3g):

$$y(x) = B + \frac{L}{1 + (x/x_0)^{-k}}$$

This function is a logistic function with x and x_0 on a logarithmic scale. $L+B$ denotes the supremum of the function, and k is the logistic growth rate.

AlphaFold predictions and structure analyses

Structure predictions of cohesin subunits and CTCF fragments were performed with AlphaFold v2 [311, 418], downloaded from `https://github.com/google-deepmind/alphafold`. We used the default parameters except that all models were relaxed with a maximum number of 20 outer and 200 inner iterations. The sequences of the individual cohesin and CTCF fragments are listed in Table 8.2 and the individual chains as input to the AlphaFold-multimer pipeline are listed in Table 8.3. After the structure prediction, individual predictions of a combination of cohesin subunits and CTCF fragments were aligned on the NIPBL, STAG1, STAG2, or Smc3, respectively, using custom scripts using PyMOL [423]. The structures were also rendered in PyMOL. The structural similarity of individual amino acids within CTCF fragments #1-3 (and scrambles thereof) was computed as follows: the α-carbon (Cα) coordinates were extracted from the PDB files and the distance of each amino acid in prediction pi to the corresponding amino acid in prediction p_j was computed, where $i \neq j$. Subsequently, the median distance of each amino acid in the prediction p_i to all other predictions was computed and displayed as one individual line (the median distance of each amino acid in prediction p_i to all other predictions) in Figure S8.4g.

8.9 Supplementary Information

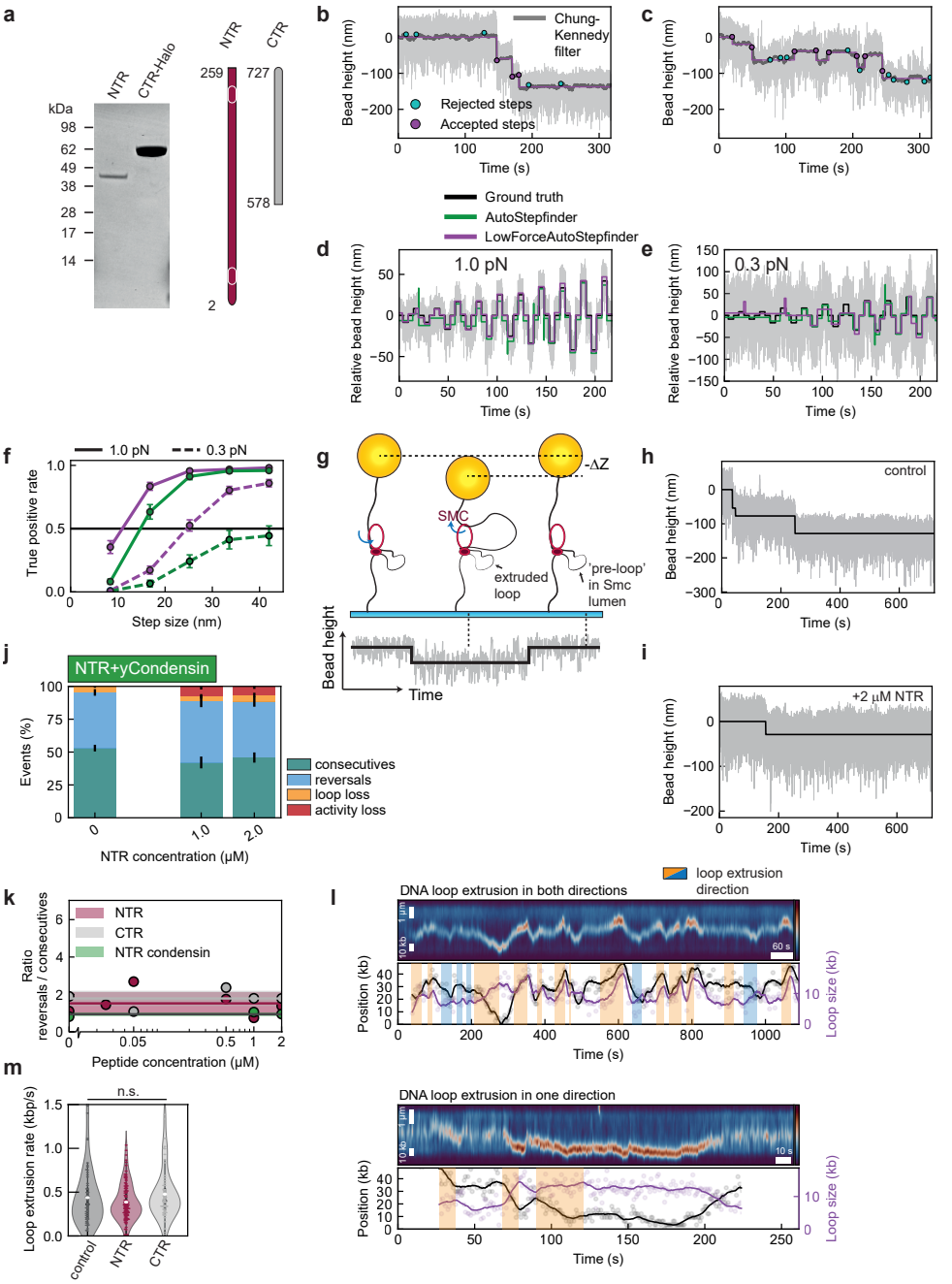


Figure S8.1 | SDS-PAGE of CTCF NTR and CTR, exemplary MT traces, yeast condensin stepping activity with NTR, and additional examples of cohesin-mediated DNA loop extrusion as observed using the fluorescence assay.

(**a**) Coomassie staining of CTCF NTR and CTR fragments after SDS-PAGE. (**b,c**) Depiction of the step detection output (magenta line) for two representative MT trajectories at 0.3 pN. Edges in the trace are detected by calculating an output function for each time point based on the difference between forward and backward averages divided by the standard deviation of the local noise component (Methods). A variable threshold maintains a constant detection efficiency (probability of true output events) of 95%. Above-threshold peaks in the output are represented by magenta dots, cyan dots depict rejected below-threshold steps. *Caption continues on next page.*

Figure S8.1 | (**d,e**) Example MT traces of piezo-induced ΔZ steps measured on 1.5 kbp DNA at 1.0 pN and 0.3 pN stretching force, respectively, serving as a means to evaluate the *LowForceAutoStepfinder* (LFAS, magenta) with the *AutoStepfinder* (AS, green) used and validated in [217]. (**f**) Evaluation of the step detection sensitivity and detection limits of the step-finding methods on piezo-induced stepping trajectories using the true positive rate (Methods). (**g**) Illustration of a failed attempt of an SMC protein complex to conclude a LE step cycle with an exemplary MT trace displaying a reversal step in our working model (same trace as shown in Figure 8.1c). DNA is reeled through the Smc lumen which causes a step. Failure to stabilize the pre-loop and to merge it with the already-extruded loop causes reversal of the bead height to the original level. (**h**) Example MT trace of cohesin in the absence of any CTCF fragment. The light grey line denotes the raw data and the solid black line is the resulting stepfinder fit. (**i**) As in (h) for a trace in the presence of 2 μM CTCF NTR. (**j**) Fractions of stepping events classified as consecutive LE steps, reversals, and loop loss for MT traces of DNA loop extrusion by yeast condensin with increasing concentrations of CTCF NTR. Error bars denote 95% binomial confidence intervals (CI). N = 1581, 454, 636 steps from 171, 50, 77 tethers, respectively, for increasing CTCF NTR concentration. (**k**) The ratio of reversal and consecutive steps for cohesin with increasing concentrations of CTCF NTR and CTR (corresponding data shown in Figure 8.1d,e) and yeast condensin with CTCF NTR (data shown in panel (j)). No dependence of the reversal/consecutive ratio (Methods) was noted. Horizontal lines and the shaded error bars denote the mean and SD of the individual data points, respectively. (**l**) Kymographs and loop position/size quantification of additional examples of cohesin-mediated DNA loop extrusion as in Figure 8.1h,i. (**m**) Loop extrusion rate within segmented loop extrusion phases of cohesin molecules in the presence of 2 μM of the indicated CTCF fragment. Statistical significance was assessed by a Kruskal–Wallis test. N = 57, 110, 61 from left to right.

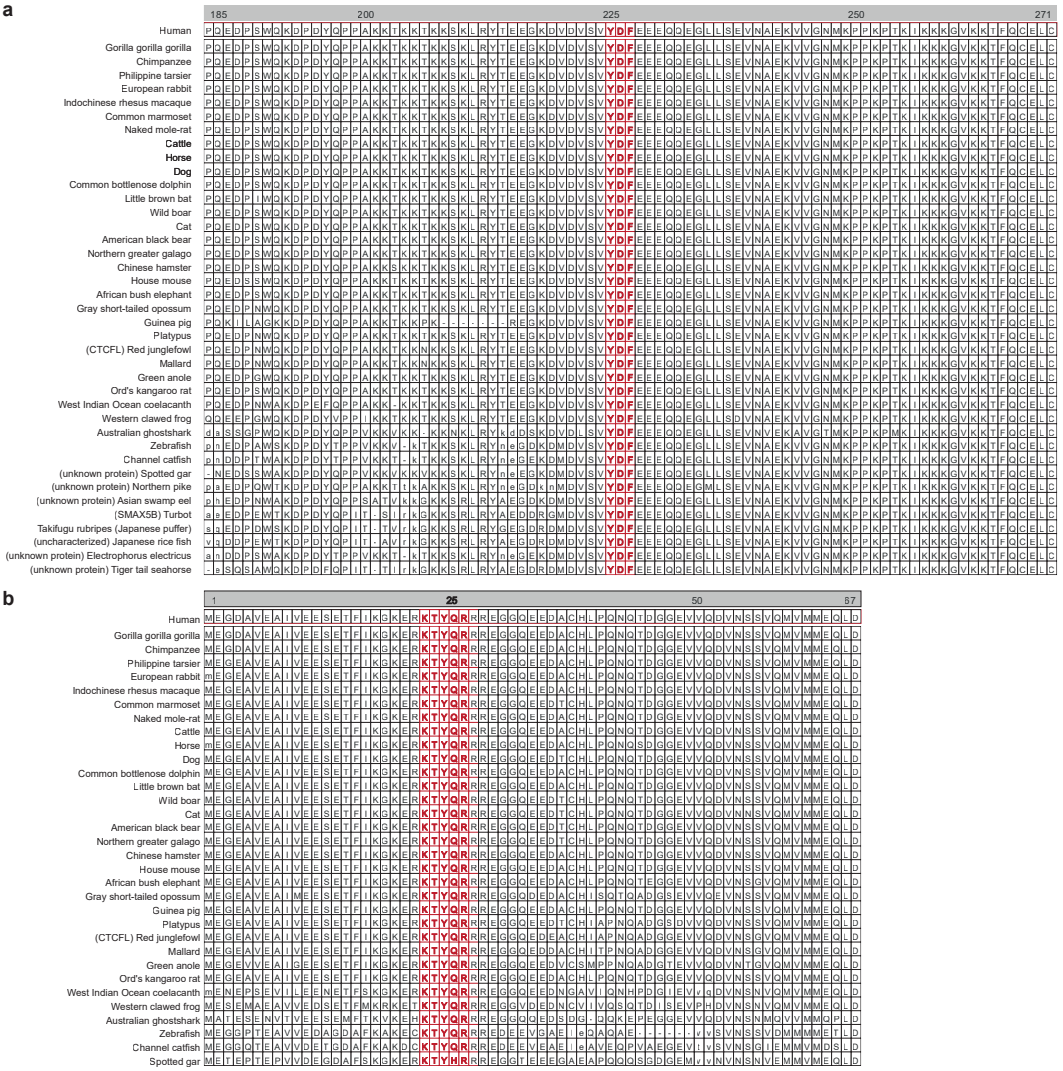


Figure S8.2 | Sequence homology assessment *via* SLiMSearch of CTCF fragments surrounding the YDF and KTYQR motifs of human CTCF.
(**a**) SLiMSearch4 [424] of the YDF motif (motif regular expression used was [PF-CAVIYL][FY][GDEN]F.{0,1}[DANE].{0,1}[DE] as described in [141]) with default parameters. Sequences were subsequentially aligned using ProViz [425] within the metazoa database. (**b**) As in (a) but surrounding the KTYQR motif (motif regular expression used was KTYQR).

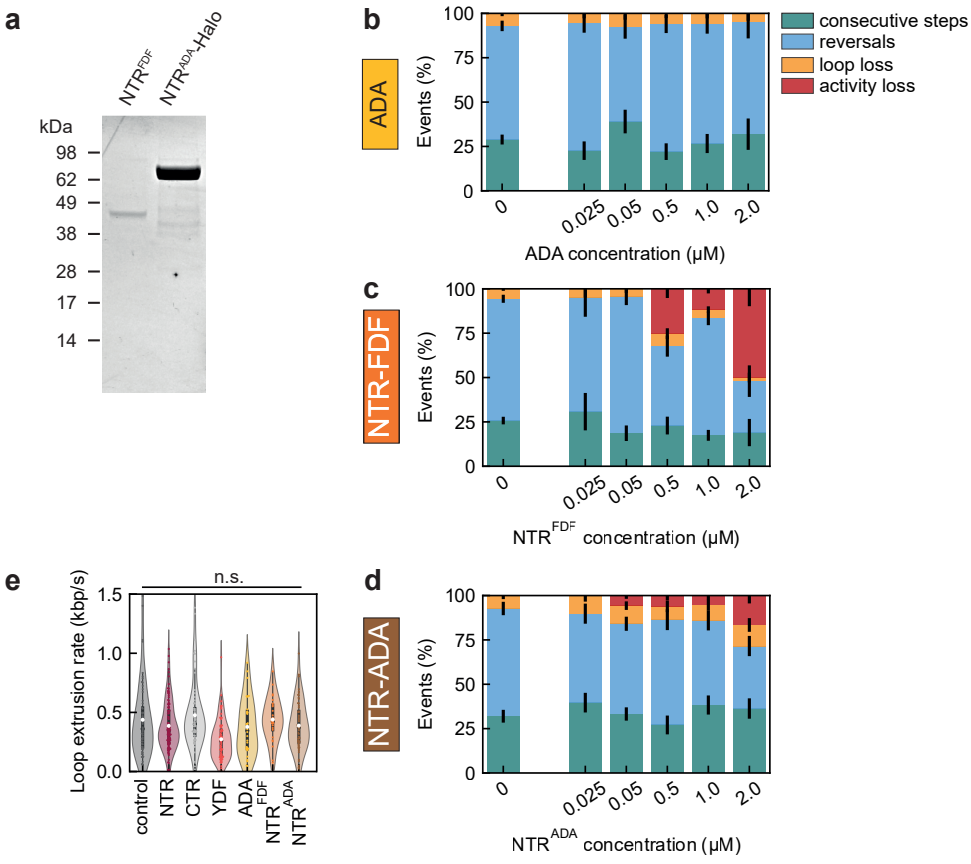


Figure S8.3 | SDS-PAGE of NTR^{FDF} and NTR^{ADA} and cohesin stepping activity in presence of the ADA peptide, NTR^{FDF}, and NTR^{ADA}.
(**a**) Coomassie staining of CTCF NTR^{FDF} and NTR^{ADA} fragments after SDS-PAGE. (**b**) The fraction of stepping events observed in MT traces classified into consecutive LE steps (green; downwards), reversals (blue; down-up step combination; Methods), and loop loss (yellow; upwards) with increasing concentrations of ADA peptide. Error bars denote 95% binomial confidence intervals (CI). N = 1060, 297, 237, 366, 357, 145 steps from 276, 79, 66, 72, 41, 77 tethers, respectively, for increasing peptide concentration. (**c**) As for (b) but with increasing concentrations of CTCF NTR^{FDF}. N = 1604, 104, 360, 202, 932, 50 steps from 432, 48, 69, 62, 223, 79 tethers, respectively, for increasing peptide concentration. (**d**) As for (b) but with increasing concentrations of CTCF NTR^{ADA}. N = 644, 333, 573, 253, 296, 225 steps from 142, 69, 134, 62, 72, 62 tethers, respectively, for increasing peptide concentration. (**e**) Loop extrusion rate within segmented loop extrusion phases of cohesin molecules in the presence of 2 μM of the indicated CTCF fragment. Statistical significance was assessed by a Kruskal–Wallis test. N = 57, 110, 61, 72, 32, 25, 57 from left to right.

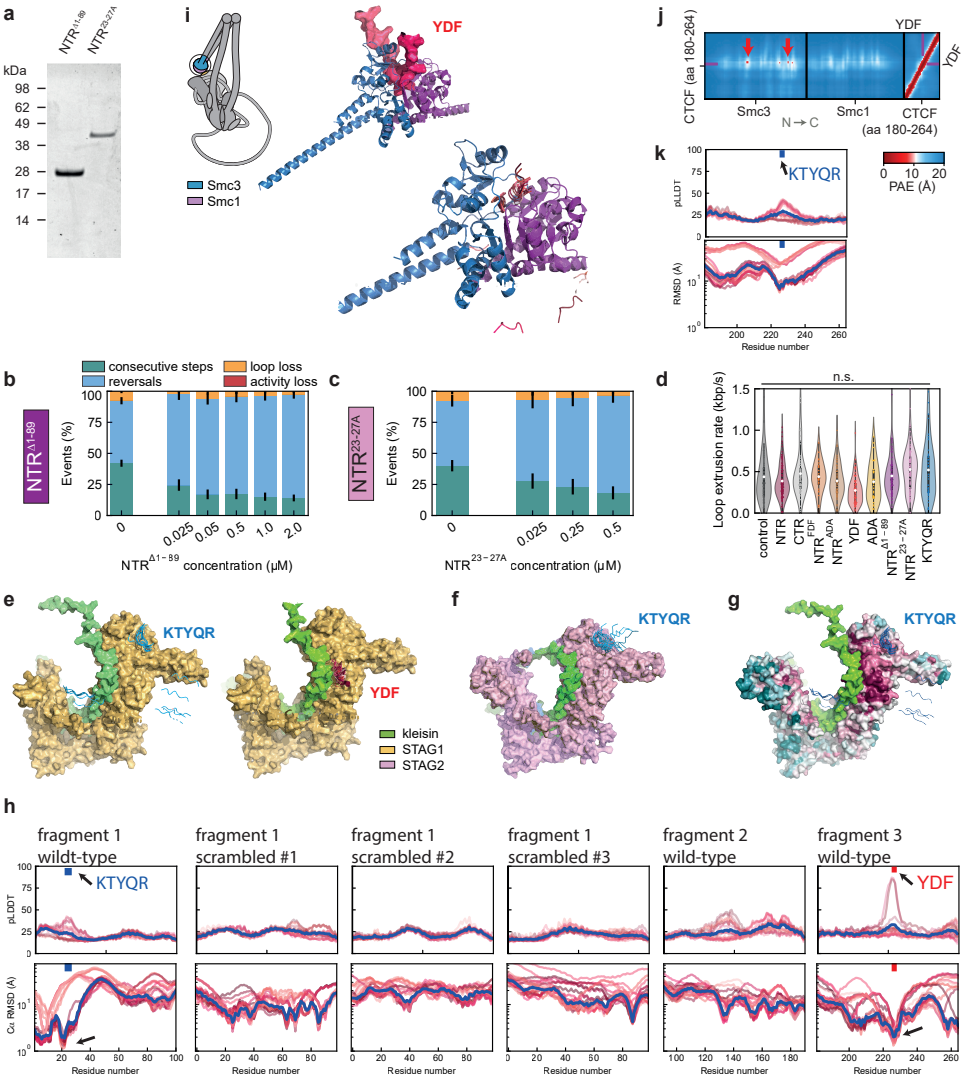


Figure S8.4 | SDS-PAGE of NTR$^{\Delta 1\text{-}89}$ and NTR$^{23\text{-}27A}$, cohesin stepping activity in presence of the NTR$^{\Delta 1\text{-}89}$ and NTR$^{23\text{-}27A}$ fragments, and AlphaFold predictions.

(**a**) Coomassie staining of CTCF NTR$^{\Delta 1\text{-}89}$ and NTR$^{23\text{-}27A}$ fragments after SDS-PAGE. (**b**) The fraction of stepping events observed in MT traces classified into consecutive LE steps (green; downwards), reversals (blue; down-up step combination; Methods), and loop loss (yellow; upwards) with increasing concentrations of CTCF NTR$^{\Delta 1\text{-}89}$. Error bars denote 95% binomial confidence intervals (CI). N = 1155, 355, 429, 608, 911, 860 steps from 277, 70, 52, 94, 137, 154 tethers, respectively, for increasing peptide concentration. (**c**) As for (b) but with increasing concentrations of CTCF NTR$^{23\text{-}27A}$. N = 444, 292, 522, 329 steps from 94, 37, 40, 38 tethers, respectively, for increasing peptide concentration. (**d**) Frequency of extrusion direction change in the absence or the presence of 2 μM of the indicated CTCF fragment. White dots denote the mean, the box shows the quartiles of the data, and whiskers extend to 1.5·interquartile range. Statistical significance was assessed by a Kruskal–Wallis test (***: $p < 0.001$ and the Cliff's Delta statistic is ≤ 0.33). N = 20, 20, 13, 4, 8, 18, 11, 15, 9, 11 DNA molecules from left to right. (**e**) Loop extrusion rate within segmented loop extrusion phases of cohesin molecules in the presence of 2 μM of the indicated CTCF fragment. Statistical significance was assessed by a Kruskal–Wallis test. N = 57, 110, 61, 25, 57, 72, 32, 76, 79, 46 from left to right. *Caption continues on next page.*

Figure S8.4 | (**f**) AlphaFold predictions of STAG1 with a kleisin fragment and a CTCF fragment spanning residues amino acids 1-100 containing the KTYQR motif (left) or a CTCF fragment spanning residues amino acids 180-264 containing the YDF motif (right). All 25 predictions were aligned using PyMOL [423] onto the STAG1 structure and were virtually identical to one another. Only the KTYQR motif (blue) or YDF motif (red) ± 4 amino acids is shown for each CTCF fragment chain for clarity. The placement of the KTYQR and YDF motif resulting from the 25 individual runs is shown individually. (**g**) As for (e) but with STAG2 instead of STAG1. (**h**) Predicted local distance difference test (pLDDT; scale from 0 to 100) values for 25 individual runs (red-shaded lines; the blue line denotes the median) over CTCF fragment #1 (amino acids 1-100), 3 randomly scrambled versions of this sequence, CTCF fragment #2 (amino acids 90-190), and CTCF fragment #3 (amino acids 180-264). The positions of the KTYQR and YDF motifs are annotated (blue and red arrows, respectively). Note that a low pLDDT value is expected since the CTCF NTR is predominantly intrinsically disordered [426]. The bottom row shows the average distance of each amino acid within each CTCF fragment to the same amino acid in the same fragments of the other 24 runs. (**i**) AlphaFold predictions of the Smc hinge and CTCF fragment #3 (containing the YDF motif). The 25 predictions were aligned on Smc3 and the 25 individual placements of the YDF motif are shown as red ribbons. (**j**) Predicted Alignment Error (PAE) map of the Smc3-Smc1 hinge and CTCF fragment #3. The position of the YDF motif is indicated by the purple tick. (**k**) As panel (h) for CTCF fragment #3 in association with the Smc hinge.

Table 8.1 | Peptide sequences of CTCF fragments. YxF and KTYQR motifs, as well as mutations thereof, are marked in bold.

Peptide	Sequence	Position	Provider
YDF	DVSV**YDF**EEE	222-231	In-house
ADA	DVSV**ADA**EEE	222-231	In-house
KTYQR	FIKGKER**KTYQR**RREG	17-32	GenScript
NTR	EGDAVEAIVEESETFIKGKER**KTYQR**RREGGQEEDACH LPQNQTDGGEVVQDVNSSVQMVMMEQLDPTLLQMKT EVMEGTVAPEAEAAVDDTQIITLQVVNMEEQPINIGEL QLVQVPVPVTVPVATTSVEELQGAYENEVSKEGLAESE PMICHTLPLPEGFQVVKVGANGEVETLEQGELPPQEDP SWQKDPDYQPPAKKTKKTKKSKLRYTEEGKDVDVSV **YDF**EEEQQEGLLSEVNAEKVVGNMKPPKPTKIKK	2-259	In-house
CTR	AGPDGVEGENGGETKKSKRGRKRKMRSKKEDSSDSEN AEPDLDDNEDEEEPAVEIEPEPEPQPVTPAPPPAKKR RGRPPGRTNQPKQNQPTAIIQVEDQNTGAIENIIVEV KKEPDAEPAEGEEEEAQPAATDAPNGDLTPEMILSMM DR	578-727	In-house
NTR^{ADA}	EGDAVEAIVEESETFIKGKER**KTYQR**RREGGQEEDACH LPQNQTDGGEVVQDVNSSVQMVMMEQLDPTLLQMKT EVMEGTVAPEAEAAVDDTQIITLQVVNMEEQPINIGEL QLVQVPVPVTVPVATTSVEELQGAYENEVSKEGLAESE PMICHTLPLPEGFQVVKVGANGEVETLEQGELPPQEDP SWQKDPDYQPPAKKTKKTKKSKLRYTEEGKDVDVSV **ADA**EEEQQEGLLSEVNAEKVVGNMKPPKPTKIKK	2-259	In-house
NTR^{FDF}	EGDAVEAIVEESETFIKGKER**KTYQR**RREGGQEEDACH LPQNQTDGGEVVQDVNSSVQMVMMEQLDPTLLQMKT EVMEGTVAPEAEAAVDDTQIITLQVVNMEEQPINIGEL QLVQVPVPVTVPVATTSVEELQGAYENEVSKEGLAESE PMICHTLPLPEGFQVVKVGANGEVETLEQGELPPQEDP SWQKDPDYQPPAKKTKKTKKSKLRYTEEGKDVDVSV **FDF**EEEQQEGLLSEVNAEKVVGNMKPPKPTKIKK	2-259	In-house
NTR$^{\Delta 1\text{-}89}$	VDDTQIITLQVVNMEEQPINIGELQLVQVPVPVTVPV ATTSVEELQGAYENEVSKEGLAESEPMICHTLPLPEG FQVVKVGANGEVETLEQGELPPQEDPSWQKDPDYQPP AKKTKKTKKSKLRYTEEGKDVDVSV**YDF**EEEQQEGLL SEVNAEKVVGNMKPPKPTKIKK	90-259	In-house
NTR$^{23\text{-}27A}$	EGDAVEAIVEESETFIKGKER**AAAAA**RREGGQEEDACH LPQNQTDGGEVVQDVNSSVQMVMMEQLDPTLLQMKT EVMEGTVAPEAEAAVDDTQIITLQVVNMEEQPINIGEL QLVQVPVPVTVPVATTSVEELQGAYENEVSKEGLAESE PMICHTLPLPEGFQVVKVGANGEVETLEQGELPPQEDP SWQKDPDYQPPAKKTKKTKKSKLRYTEEGKDVDVSV **YDF**EEEQQEGLLSEVNAEKVVGNMKPPKPTKIKK	2-259	In-house

Table 8.2 | Sequences of CTCF fragments and cohesin subunits used for AlphaFold predictions. All subunits refer to the human sequence. The Smc1/3 head was assembled by the N- and C-termini of Smc1 and Smc3, respectively, linked by a flexible linker consisting of 32 Glycines (32xG).

Name	Position
Scc1 fragment binding to NIPBL	143-181
Scc1 fragment binding to STAG1/2	301-413
Scc1 fragment binding to Smc3	1-100
Scc1 fragment binding to Smc1	531-631
NIPBL	1191-2641
Mau2	1-613
STAG1	86-971
STAG2	77-1064
Smc1 head	(1-240)-32xG-(994-1233)
Smc3 head	(1-201)-32xG-(978-1217)
Smc1 hinge	481-711
Smc3 hinge	451-691
CTCF fragment #1	1-100
CTCF fragment #2	90-190
CTCF fragment #3	180-264

Table 8.3 | Components of cohesin subunits used in AlphaFold predictions.

Name	Components (nomenclature as inTable 8.2
NIPBL-Mau2	NIPBL, Mau2, Scc1 fragment binding to NIPBL
STAG1	STAG1, Scc1 fragment binding to STAG1/2
STAG2	STAG2, Scc1 fragment binding to STAG1/2
ATPase heads	Smc1 head, Smc3 head, Scc1 fragment binding to Smc1, Scc1 fragment binding to Smc3
Hinge	Smc1 hinge, Smc3 hinge

IV

Outlook and summary



9

Open questions and future directions



Abstract

In this thesis, we utilized single-molecule fluorescence and magnetic tweezers to study the intricacies of DNA loop extrusion by SMC complexes in a controlled in vitro environment. We observed, described, and analyzed the behaviour of SMCs which yielded detailed yet important contributions to the resolution of the DNA loop extrusion mechanism as well as to the physiological impact on genome structure. In this final chapter, I formulate and discuss major outstanding questions in the understanding of DNA loop extrusion and consequently SMC and genome biology.

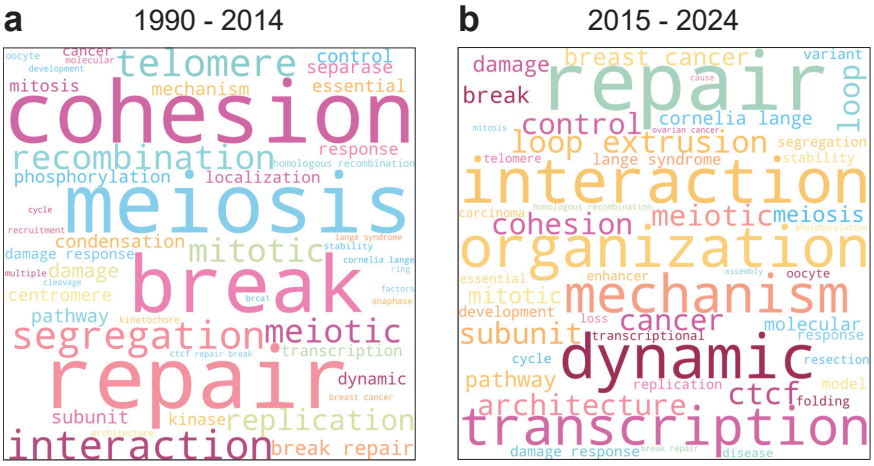


Figure 9.1 | The focus of the SMC field has partially shifted in the last decade.
(**a**) A word cloud generated from the titles of PubMed entries using one or multiple keywords of the list "structural maintenance of chromosomes, cohesin, condensin, smc5/6, rad50, recn, coalescin, mukbef, mksbef, bssmc, loop extrusion" from 1990 to 2024. A maximum of 50 words are displayed for publications between the years 1990 - 2014. (**b**) As panel (a) but for publications appearing between 2015 - 2024.

SMC proteins were initially discovered for their role in proper chromosome segregation in budding yeast. Initial research after identification of the heterodimerization of cohesin, condensin, and SMC5/6 therefore focused on sister chromatid cohesion and segregation, meiosis, as well as DNA damage and its repair (Figure 9.1a). After Nasmyth's concrete hypothesis in 2001 that SMC proteins might extrude DNA loops [118], not a single publication concerning DNA loop extrusion and SMC proteins has been listed on PubMed for more than a decade. Only in 2012, Alipour & Marko theoretically describe a "hypothetical DNA-loop-extruding enzyme machines with general properties consistent with those of condensins" [111], based on the notion in the SMC field that SMCs are somehow involved in stabilizing loop structures by bridging (e.g. [151]). Alipour and Marko reasoned that random polymer motion leading to looping would only efficiently form small and unspecific loops, thus (once again) postulating a DNA loop extrusion mechanism. From 2015 onwards, starting with Wang et al. [427] and Sanborn et al. [97], DNA loop extrusion became a topic of major interest for the SMC and the genome biology field.

Accordingly, the research focus has significantly shifted: the last decade of SMC research was dominated by questions relating to the organization and architecture of the genome based on the activity of SMCs. The word cloud shown in Figure 9.1b demonstrates that SMC research slowly becomes more translational as the (mis)function of SMCs in disease, often in cancer, is interrogated. The keyword "dynamic" illustrates that the field has come to appreciate the unsteady nature of DNA extrusion and the genome *per se*, which contrasts earlier years where SMC proteins were thought to form rather static DNA-bridging structures. Furthermore, beyond loop extrusion as a means to shape the genome, its functional role in transcription and during development is highlighted more and more. Nevertheless, DNA damage and repair as well as cohesion remained topics of active research. Finally, as exemplified by this thesis, research is also increasingly dedicated towards the

resolution of the mechanism of loop extrusion and interactions of SMCs with other nuclear proteins, in particular CTCF (Figure 9.1b). Below, we formulate some key open questions in the field, with an emphasis on questions that relate to the insights obtained in this thesis.

9.1 Is DNA temporarily pseudotopologically embraced during loop extrusion?

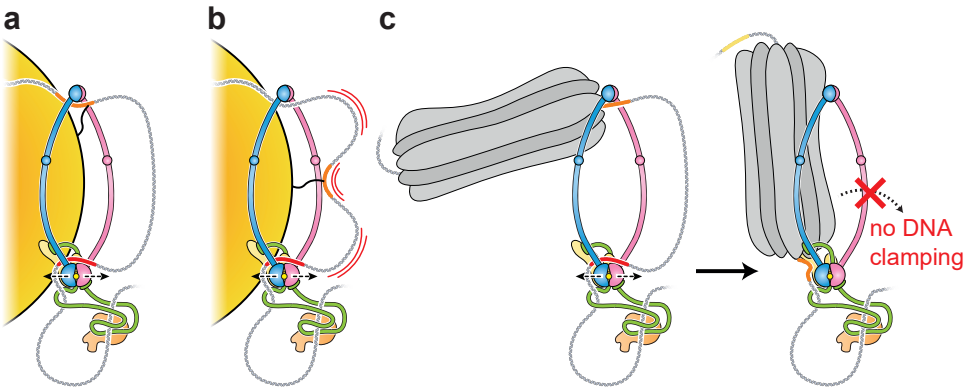


Figure 9.2 | Large DNA-encircling roadblocks are predicted to halt loop extrusion.
(**a**) Within the Reel & Seal model [247], a PEG-dCas9-DNA-bound roadblock is likely to encounter the SMC hinge. (**b**) Reeling DNA pseudotopologically past the roadblock attachment point would cause DNA to bend additionally as long as the roadblock is too large to fit through the Smc lumen. (**c**) A large DNA-encircling roadblock (here drawn as an oversized MCM complex) would also be encountered by the SMC hinge. The DNA segment in front of the roadblock (orange DNA segment) could be moved to the ATPase heads by ATP hydrolysis and concomitant disengagement of the ATPase heads. However, the DNA-encircling roadblock would have to be threaded through the Smc lumen at the next ATP binding and DNA clamping event. If the roadblock is larger than the Smc lumen, this step is prohibited, which in turn halts loop extrusion. Note that this holds true also for a very stiff downstream segment of DNA.

In chapter 3, we bound biological roadblocks (nucleosomes, RNA polymerase, and dCas9) to DNA. Because these proteins are small enough to fit through the Smc lumen and were ineffective in halting loop extrusion (Figure 3.2), we then used dCas9-bound nanoparticles of sizes up to 200 nm (Figure 3.4). However, it is possible that tethering these nanoparticles to DNA *via* a PEG-linker and dCas9 helped condensin and cohesin to pass these roadblocks. Because DNA has to bend considerably on the length scale of the Smc lumen (a segment of ~150 bp [35, 217] has to bend into a semi-circle of diameter ~35 nm) the roadblock is most likely to encounter the SMC at the hinge (Figure 9.2a,b). A roadblock which encircles DNA (such as MCM) would also encounter the SMC hinge and could be transported close to the ATPase heads upon ATP hydrolysis within the Reel & Seal model [247]. However, upon the next ATP binding and ATPase head engagement event, the roadblock-covered DNA segment would have to be threaded pseudotopologically through the Smc lumen. If the roadblock is too large to fit through the lumen, DNA cannot be clamped onto the ATPase heads, which would stall DNA loop extrusion efficiently (Figure 9.2c). In fact, not only big roadblocks without linker to their DNA attachments would be predicted to stall loop extrusion, but also stiff DNA filaments. As long as the downstream protein-covered

DNA segment is sufficiently stiff to oppose bending of the DNA into a semi-circle of ~35 nm diameter, loop extrusion would come to a standstill. If such roadblocks or stiff DNA segments could still be reeled into the loop, this would argue against a temporary pseudotopological embracement of DNA within the Smc lumen. It is difficult to design such a large DNA-embracing roadblock (the MCM complex is smaller (20 nm long along DNA [428]) and does not impose an impermeable barrier to loop extrusion despite it containing a YxF motif like CTCF [191]), even though emerging technologies to design large cyclic protein homo-oligomers [429, 430] and DNA binding proteins [431] could be used to design such roadblocks *de novo*. Alternatively, arrays of DNA-binding proteins such as Rap1 [272], nucleoprotein filament-forming proteins or even bundles of DNA designed by DNA origami technology could be used to stiffen DNA. Ideally, DNA would remain accessible (and not protein-covered) to discern the effects of limited accessibility and stiffness. In this regard, *de novo* designed large protein roadblocks offer the advantage of having a small DNA footprint, thus not affecting DNA accessibility.

9.2 What determines whether a loop extruder can switch directions?

The results presented in chapter 5 showed that human cohesin and yeast SMC5/6, but not yeast condensin, do frequently switch the direction of loop extrusion. A striking difference between yeast condensin and the remaining SMCs is the safety belt formed by condensin's Ycg1 subunit together with its kleisin Brn1 [98, 210] which has so far not been found in cohesin nor in SMC5/6 [228, 230, 247]. Disruption of the safety belt in yeast condensin caused condensin to extrude, like cohesin and SMC5/6, bidirectionally [221]. It would be of interest to test whether introducing residues similar to condensin's safety belt in bidirectional extruders transforms them into unidirectional extruders.

Furthermore, it is unclear at which step during the loop extrusion cycle DNA strand exchange occurs. Because DNA strand exchange requires the temporary dissociation of DNA from the anchor side (HEAT-B and corresponding kleisin residues) and the motor side (involving HEAT-A), the bidirectional extrusion of cohesin postulates that STAG1/2, in contrast to Ycg1, does not permanently bind DNA which might also contribute to loop slipping and diffusion. A FRET-based protein-DNA sensor [432] would permit to measure the binding of intercalator-stained DNA to each individual DNA binding site on SMCs. This strategy might also permit to determine whether extrusion direction switches coincide with a temporary DNA unbinding at the anchor side. Furthermore, three-color FRET [433] between DNA and two DNA binding sites on the SMC would permit to determine whether DNA handover between the various binding sites involves unbinding of DNA from one binding site and subsequent binding of DNA by another side or whether both DNA binding sites are bound to DNA simultaneously to ensure that the loop cannot be lost during DNA handover. Potentially, such experiments could also hint towards whether DNA strand exchanges, and thus direction changes, happen (accidentally) when DNA dissociates simultaneously from both the motor and the anchor side or whether at least one DNA binding site remains bound to DNA at all times to prevent loop loss.

9.3 Which conformational changes account for the 'step' in the loop extrusion cycle?

Magnetic tweezers have been proven powerful to resolve individual DNA loop extrusion steps [35, 217] (and chapter 6). On the other hand, combinations of FRET pairs at various locations on cohesin [238] were invaluable to determining which subunits are in proximity in response to ATP and/or DNA. The combination of both approaches could yield critical insights. The synchronized measurement of the FRET efficiency between any two locations and the bead height would permit to relate conformational changes, as probed by the FRET pairs, to the loop extrusion step. Such measurements would not only be able to determine the order of events but also reveal the absolute timing before and after a loop extrusion step. FRET pairs located at the coiled coils or between hinge and HEAT-A could furthermore reveal whether arm zipping and unfolding of the coiled coils at the elbow is part of the loop extrusion cycle or not. Similarly, these measurements could also be extended to probe the timing of the various SMC-DNA interactions *via* FRET between DNA binding sites and intercalator-stained DNA with respect to the time point of the loop extrusion step.

9.4 How do dimers of SMCs work and what are the minimal structural requirements for loop extrusion?

The SMC field would benefit from the detailed characterisation of DNA loop extrusion by bacterial SMC complexes on the single-molecule level *in vitro*. Because bacterial SMCs are predicted to be dimers and zip up the left and right arm of the bacterial chromosome [120, 126], it is unclear whether these complexes contain an anchor side or whether two motor domains suffice. It is also of interest to measure the twist induced into the loop by a dimeric SMC as a means to determine with more certainty how twist is generated by SMC complexes: Would dimeric SMC proteins generate -1.2 turns/step (double the twist of a monomeric SMC), the same twist as a monomeric SMC or would the twist from both motor domains cancel? More generally, it is unclear whether loop extrusion models currently developed primarily for condensin and cohesin (see chapter 1) are applicable to bacterial SMCs and dimeric SMCs in general. It is possible that, even though bacterial SMCs are dimers *in vivo*, monomers of bacterial SMCs can be functionally reconstituted *in vitro*, similar to SMC5/6 which extrudes DNA symmetrically as a dimer [204] but asymmetrically as a monomer (chapter 5 and [222]). Along these lines, it would also be of interest to reconstitute DNA loop extrusion by some archaeal SMC members whose Hi-C data indicates SMC-dependent loop and domain formation [434]. Furthermore, distant members of the SMC family include but are not limited to the DNA-repair protein Rad50 which assembles into the Mre11-Rad50-Nbs1 (MRN) complex [435], coalescin (ClsN) identified from the archaeon *Sulfolobus* [436], RecN [437, 438], and Wadjet SMCs/MksBEF (MukBEF-like SMC proteins) [62, 439, 440]. Next to bacterial SMC complexes, none of these mentioned complexes have been unequivocally shown to extrude DNA. Rad50 [441], part of the MRN complex, as well as the archaeal ClsN [436] contain (putative) zink hook motifs without known DNA affinity. Instead, MRN dimerizes *via* its zinc hook motif in Rad50, possibly to tether two DNA strands [442]. By contrast, the hinge dimerization domain can bind DNA

[237, 238, 443–446], which is essential for its DNA loop extrusion [238] and supercoiling [50, 51] activity (chapter 6). While the zink hook of Rad50 may be functionally replaced by the MukB hinge domain to maintain dimerization of Rad50 [447], it would be interesting to know whether the hinge domain of a loop-extruding SMC may be replaced by a zinc hook. Based on the available literature described so far, failure to bind DNA at the zinc hook-replaced hinge is predicted to alleviate loop extrusion. Consequently, zinc-hook-containing members of the SMC superfamily would be predicted to be unable to extrude DNA loops. Other members of the SMC superfamily, like the bacterial RecN, contain only short coiled coils which may be too short to permit the large-scale conformational changes thought to underlie DNA loop extrusion. The diversity among the SMC superfamily across the tree of life offers a pathway to determine strict requirements of loop extrusion on SMCs. Hints for DNA loop extrusion by members of the SMC family may be obtained by searching the genomes of diverse species for SMC-mediated loops (as e.g. in [434, 448]) at the increasing resolution of chromosome conformation capture techniques and loop detection algorithms [449]. Furthermore, biochemical reconstitution of candidate SMCs and analysis of their loop-extrusion characteristics such as the loop-extrusion rate, stalling force, step size, induced twist, and directionality, as outlined in this thesis, bears the potential to formulate dependencies of the loop-extrusion dynamics to the properties of the SMC member and its subunits.

9.5 How are SMC complexes regulated *in vivo* by their interactome?

Although all eukaryotic SMC complexes perform loop extrusion with comparable characteristics, their function *in vivo* is highly divergent. Likely, interactions of a wide range of proteins with SMC complexes as well as post-translational modifications (PTMs) determine their fate throughout the cell cycle [450–452]. While the role of interacting partners, as well as PTMs in the establishment and maintenance of sister chromatid cohesion, is getting more explored [127, 136, 453, 454], this is only starting to be revealed for loop-extruding SMC complexes [141, 455]. The most studied interaction partners of extruding cohesin are CTCF [35, 141] and other YxF-containing proteins [191, 250]. However, other interactions of SMCs appear to be of great importance such as the extensive network of SMCs with the spliceosome [451] which impacts the regulation of proteins even when gene regulation at the transcriptional level is barely affected [456]. To understand the varied roles of all SMCs *in vivo* it will be important to identify interacting partners, e.g. by massive computational screens using AlphaFold [311, 457] to obtain candidates. Single-molecule assays such as those presented and utilized in this thesis hold the promise to reveal the impact of SMC interactors and PTMs on the dynamics and stability of DNA loop extrusion.

9.6 Why so complicated?

SMC complexes are relatively large assemblies of mostly comparatively large subunits (e.g. all subunits of human cohesin and condensin except the kleisin contain more than 1000 amino acids) with little symmetry. The SMC can undergo a range of large conformational changes and proves to be flexible, with multiple subunit interactions which vary throughout

the ATPase cycle. At its core, DNA loop extrusion is simply described as an anchored translocation. The motor side of the SMC (involving HEAT-A) translocates along DNA while the anchor side remains stably bound to DNA (involving HEAT-B; neglecting loop diffusion and slipping events). Other DNA translocators appear to employ a comparatively simpler translocation mechanism: for example, FtsK in *E. coli* and SpoIIIE in *B. subtilis* are homo-hexamers which exploit a treadmilling-like mechanism for translocation based on the varying affinity of each monomer to DNA, depending on the nucleotide state [458, 459]. Transcription elongation may occur by a single RNA polymerase and helicases can act as a single subunit. Other translocators can also be fast (FtsK translocates with up to 50 kbp/s [460]) and can generally withstand a much larger force than SMCs (typically several tens of pN [461, 462]). One may thus wonder why SMC complexes evolved such a complicated mechanism to translocate DNA. What is the advantage compared to a simpler translocation mechanism?

Of course, there are also striking differences between DNA loop extrusion by SMC complexes and DNA translocation by other proteins. For instance, these translocators move in steps of 1-2 bp [463, 464]. Therefore, translocation tends to stall or reverse upon encountering tightly bound roadblocks on DNA or evict DNA almost entirely when trains of translocators process the same segment of DNA [465–467]. This is in stark contrast to SMC complexes which take steps of up to >100 bp and easily overcome even the most tightly bound and very large roadblocks (chapter 3 and [272]). The underlying cause of the complicated (nontopological) loop-extrusion mechanism might thus be to alter the genome architecture concomitant with all other nuclear processes and without stripping the looped DNA from all other proteins.

Summary

Faithfully handling, processing, maintaining, and segregating genetic material is a delicate task for every cell. Structural maintenance of chromosomes (SMC) proteins play a role in many aspects of life that occur on genomes: DNA repair, sister chromatid cohesion and segregation, chromosome condensation, and gene regulation. Two basic working principles of SMCs have been identified: the topological embracing of two strands of DNA and DNA loop extrusion. I devoted the last four years to understanding, at least in parts, the latter. We employed various strategies to scrutinize some basic and some not-so-basic aspects of loop extrusion with the overarching goal of arriving at a DNA loop extrusion mechanism which can explain all available data.

In chapter 1, we provided a broad introduction to the basic principles in genome biology with a focus on setting the stage for the layers that SMC complexes add to the organization of genomes across the tree of life. We introduced the activity of SMCs during cell division and interphase after which we delved into the description of SMC biology from a biophysical perspective: loop extrusion rates, stalling forces, directionality, and topology, leading us to a discussion of the currently proposed model of DNA loop extrusion which culminated in the proposal of the so-called Real & Seal model. Finally, we introduced common and so far well-studied interaction partners as well as more generic roadblocks which have the potential to halt DNA loop extrusion.

In chapter 3, we attempted to halt DNA loop extruders by the sheer size of roadblocks in their way - and failed to do so. The last resort in the attempt to effectively halt condensin from extruding past a single nanoparticle was to make the particle much larger than the ring-shaped SMC itself. Yet even this roadblock could be passed into the loop. Apparently, size does not always matter. These surprising results could be interpreted in two ways: either an interface of the SMC ring opened upon roadblock passage or DNA loop extrusion occurs nontopologically. To discriminate between these options, we turned to an engineered version of human cohesin whose Smc-kleisin were fused and its Smc-Smc interface crosslinked. Contrary to expectations, this 'single-chain cohesin' version was *also* able to pass 200 nm big roadblocks into the loop. The interpretation was thus clear: DNA loop extrusion is, contrary to sister chromatid cohesion, a process during which SMCs are only nontopologically bound to DNA, i.e. they are not inserted through the SMC ring lumen. During the preparation and revision of the accompanying publication, pseudotopological models for loop extrusion were proposed. Prompted by these new ideas, we identified several predictions of the proposed pseudotopological models and tested those against our experimental data in chapter 4. We found that the predictions made by pseudotopological models stand in contrast to our data. Instead, they are consistent with a nontopological mode of loop extrusion.

In chapter 5, we took a different approach to characterise loop extrusion by various SMC

complexes. Previous demonstrations of loop extrusion by cohesin and SMC5/6 reported that those complexes reel in DNA from both sides simultaneously, while yeast condensin extruded DNA strictly from one side only. Was it possible that different eukaryotic SMC complexes exploit different mechanisms for loop extrusion, one where DNA loops are extruded simultaneously from both sides and one where DNA loops are extruded only from one side? To answer this question, we closely monitored the extrusion dynamics of condensin, cohesin, and SMC5/6 side by side and in simple environments. All extruding complexes were monomeric and no tension from buffer flow acted on DNA. In these conditions, it became evident that loop extrusion is in fact frequently interspersed by phases during which SMC can diffuse along together with their extruded loop and that loops can gradually shrink in size. However, the extrusion phases *per se* were well characterised as asymmetric extrusion in all cases. In the absence of buffer flow, the direction of this asymmetric extrusion can switch over time, which effectively resembles 'apparent symmetric extrusion' when averaged over sufficiently long times. Focusing from here onwards on cohesin, we could directly visualize an exchange of its dynamics subunit NIPBL in between extrusion phases. Here, we noted that reversals of the extrusion direction overwhelmingly corresponded to an exchange of NIPBL on cohesin. This led us to propose that on cohesin, NIPBL acts as a molecular gearbox whose dissociation and subsequent association to cohesin appears to be required for extrusion direction switches. We concluded our findings with a discussion of molecular determinants for direction switches in all SMC complexes and a potential "DNA strand exchange" model which could enable direction switches within the Real & Seal model.

In chapter 6, we revisit an old finding with major implications for the loop extrusion mechanism and genome architecture: DNA supercoiling by SMC complexes. Here, we used magnetic tweezers to measure the twist during individual DNA loop extrusion steps by SMC complexes. Like the step itself, we found that also DNA twist is induced upon ATP binding. Strikingly, for all three SMC complexes, we found the same dependency on ATP binding, the same handedness, and even the same magnitude of the twist which is induced into the loop in each step: -0.6 turns. This quantity is thus conserved among the eukaryotic SMCs and is an integral part of the loop extrusion mechanism, which has so far not been considered. Previously, SMC-induced supercoiling was analyzed in bulk assays using plasmids and topoisomerase. While such assays recapitulate the concerted action of several cohesin complexes and topoisomerases on topologically constrained DNA, the experiments were difficult to interpret. To reconcile bulk plasmid and magnetic tweezers data, we computationally mapped how the aforementioned plasmid assay would behave assuming solely that each loop extrusion step induces -0.6 turns into the loop and consequently +0.6 turns in the DNA outside the loop as a compensatory twist. This requirement was sufficient to largely reconcile the experimental plasmid assay results, lending further support to our magnetic tweezers data.

In the third part of this thesis, we turn towards understanding whether and how CTCF is a roadblock to cohesin-mediated loop extrusion. In chapter 7, we reconstituted DNA-bound CTCF for obstacle bypass experiments. We found that diffusing cohesin (without an extruded DNA loop) is frequently halted when it encounters CTCF from its N-terminal side. As anticipated by chromosome conformation capture, we could unequivocally confirm

that CTCF is indeed also a barrier to loop extrusion by cohesin when encountered from its N-terminal side. Unexpectedly, however, we noted a strong dependency of CTCF's blocking efficiency with the DNA tension at the moment of the encounter, a phenomenon for which an explanation remains pending to date. Furthermore, we observed a wide range of possible outcomes of cohesin-CTCF encounters. For instance, cohesin complexes could reverse their direction of extrusion upon CTCF encounter, preferentially after encounters during which the DNA tension was comparatively low. Alternatively, cohesin complexes could attempt several times to approach and overcome CTCF after partial loop loss caused by the cohesin-CTCF encounter. The latter scenario was more common at higher DNA tension. These results illustrate the dynamic nature of DNA loop extrusion, an aspect which is hard to impossible to note in live-cell or chromatin conformation capture experiments.

In chapter 8, we delved into distinct modes by which CTCF interferes with cohesin to halt loop extrusion. We segmented CTCF's N terminal region into various fragments and exposed cohesin to the resulting peptides in solution. Using magnetic tweezers, we monitored the frequency of consecutive loop extrusion steps and forward-reverse step combinations ('reversals'). When exposed to peptides containing the YxF motif (residues 226-228), cohesin tended to fail to engage in consecutive loop extrusion steps. This observation suggests that the YxF motif interferes with a stage of the loop extrusion cycle occurring after ATP binding, often leading to a failure in enlarging the extruded loop. Furthermore, loops visualized using single-molecule fluorescence tended to be extruded in a unidirectional manner, resembling the unidirectional extrusion of yeast condensin owing to its strong DNA binding at the safety belt (chapter 5 [98, 221]). It is thus possible that the binding of the YxF motif to cohesin strengthens its DNA-binding affinity, thus effectively reducing the probability of switching extrusion directions. Apart from the known YxF motif, we identified an annotated Pds5 binding site located at CTCF's N terminus (KTYQR motif; residues 16-31) that also impacts DNA loop extrusion by cohesin. Unlike the YxF motif, the KTYQR motif impairs the overall stepping activity of cohesin but does not affect the frequency of consecutive or reversal steps nor the directionality of loop extrusion by cohesin.

Finally, the thesis is concluded with a brief retrospective on the main topics of interest of SMC research throughout the last three decades. My own research within the realm of SMC proteins started around 2019, at a time of great excitement about the demonstration of DNA loop extrusion. It is gratifying to see how the field rapidly shifted gears and adapted to these thrilling findings. In the final chapter 9, I discuss some perspectives for SMC research on the single-molecule level. In particular, I propose that conserved and varying aspects of loop extrusion may be revealed by exploiting the diversity among the SMC superfamily across the tree of life, in particular by dimers of SMCs. Furthermore, I discuss critical experiments which have the potential to reveal the conformational changes which proceed and succeed the loop extrusion step *via* a combination of smFRET and magnetic tweezers, and/or by an assay to probe the FRET efficiency between the DNA binding sites of SMCs and DNA. Finally, I question why SMCs appear to exploit a comparatively complex mechanism for loop extrusion when it is at its core simply an anchored translocation. I conclude that loop extrusion has to efficiently enable roadblock passage *in vivo* in order to shape chromosomes without disturbing all other processes on DNA. The large ring-like

structure appears to be ideally suited *both* for sister chromatid cohesion as well as for nontopological loop extrusion.

Samenvatting

Betrouwbaar omgaan met genetisch materiaal, en het verwerken, onderhouden en scheiden ervan is een delicate taak voor elke cel. Structurele onderhoudseiwitten van chromosomen (SMC-eiwitten) spelen een rol in veel aspecten van genomen: DNA-reparatie, zusterchromatidecohesie en -segregatie, chromosoomcondensatie en genregulatie. Twee basis principes van SMC's zijn duidelijk egeworden: het topologisch omhelzen van twee strengen DNA en DNA-lusextrusie. Ik heb de afgelopen vier jaar gewijd aan het (gedeeltelijk) begrijpen van het laatste. We hebben verschillende strategieën toegepast om enkele basis- en niet-zo-basisaspecten van lusextrusie te onderzoeken met als overkoepelend doel te komen tot een mechanisme van DNA-lusextrusie dat alle beschikbare gegevens kan verklaren.

In hoofdstuk 1 geef ik een brede introductie tot de basisprincipes in de organisatie van de genoom met een focus op wat SMC-complexen toevoegen aan de organisatie van genomen. We beschrijven de activiteit van SMC's tijdens celverdeling en interfase, waarna we ingingen op de beschrijving van SMC-biologie vanuit een biofysisch perspectief: lusextrusiesnelheden, stopkrachten, richting en topologie, wat leidde tot een discussie over het momenteel voorgestelde model van DNA-lusextrusie dat uitmondde in het voorstel van het zogenaamde Real Seal-model. Ten slotte introduceerden we veelvoorkomende en tot nu toe goed bestudeerde interactiepartners, evenals meer generieke obstakels die de potentie hebben om DNA-lusextrusie te stoppen.

In hoofdstuk 3 hebben we op naïeve wijze geprobeerd SMC's te stoppen door alleen de grootte van obstakels op hun pad te doen toenemen - en dit faalde. Het laatste redmiddel in onze pogingen om condensin te doen stoppen met lusvorming voorbij een enkel nanodeeltje was om het deeltje veel groter te maken dan de ringvormige SMC zelf. Toch kon zelfs dit obstakel in de lus worden gepasseerd! Blijkbaar doet grootte er niet altijd toe. Deze verrassende resultaten konden op twee manieren worden geïnterpreteerd: ofwel een koppelpunt van de SMC-ring opende bij het passeren van het obstakel, ofwel DNA-lusextrusie vond plaats op een niet-topologische manier. Om deze opties te onderscheiden, richtten we ons op een gemodificeerde versie van menselijke cohesin waarvan de Smc-kleisines werden gefuseerd en de Smc-Smc-interface covalent werd gebonden. In tegenstelling tot de verwachtingen kon zelfs deze ' enkelketen-cohesin'-versie obstakels van 200 nm passeren in de lus. De interpretatie was dus duidelijk: DNA-lusextrusie is, in tegenstelling tot zusterchromatidecohesie, een proces waarbij DNA *niet*-topologisch gebonden is aan het SMC complex. Tijdens de voorbereiding en revisie van de bijbehorende publicatie werden pseudotopologische modellen voor lusextrusie voorgesteld. Aangemoedigd door deze nieuwe ideeën, onderzochten we verschillende voorspellingen van de voorgestelde pseudotopologische modellen en testten die tegen onze experimentele gegevens in hoofdstuk 4. We ontdekten dat de voorspellingen gedaan door pseudotopologische modellen niet consistent waren met onze gegevens. In plaats daarvan zijn ze consistent met een

niet-topologische modus van lusextrusie.

In hoofdstuk 5 namen we een andere benadering om lusextrusie door verschillende SMC-complexen te karakteriseren. Eerdere demonstraties van lusextrusie door cohesin en SMC5/6 hadden laten zien dat die complexen DNA symmetrisch van beide kanten tegelijkertijd een lus vormden, terwijl gistcondensine alleen de lus met DNA van één kant vergroot. Was het mogelijk dat verschillende eukaryotische SMC-complexen verschillende mechanismen benutten voor lusextrusie? Om deze vraag te beantwoorden, hebben we de lusvormingsdynamiek van condensin, cohesin en SMC5/6 nauwlettend gevolgd, zij aan zij en in goed gedefinieerde omgevingen. Alle lusvormende complexen waren monomeer en er was geen mechanische spanning van bufferstroming op het DNA. Onder deze omstandigheden werd het duidelijk dat lusextrusie in feite vaak wordt onderbroken door fasen waarin SMC's samen met de gevormde lus kunnen diffunderen en waarin lussen geleidelijk in grootte kunnen afnemen. De extrusiefasen *per se* werden goed gekarakteriseerd als asymmetrische extrusie. De richting van deze asymmetrische extrusie kan veranderen, wat effectief lijkt op symmetrische extrusie wanneer dit wordt gemiddeld over voldoende lange tijd. Voor cohesin konden we rechtstreeks een uitwisseling van zijn dynamische subeenheid NIPBL visualiseren tussen de verschillende extrusiefasen. Meer specifiek merkten we op dat omkeringen van de extrusierichting vrijwel altijd overeenkwamen met een uitwisseling van NIPBL op het cohesin. Dit leidde ons tot de hypothese dat NIPBL op cohesin fungeert als een moleculaire versnellingsbak waarvan de dissociatie en daaropvolgende associatie met cohesin nodig lijken te zijn voor omkeringen van de extrusierichting. Tot slot beschreven we de moleculaire bepalende factoren voor richtingsschakelingen in alle SMC-complexen en een potentieel "DNA-strengenuitwisseling"model dat richtingswijzigingen mogelijk zou kunnen maken binnen het Real Seal-model.

In hoofdstuk 6 gaan we terug naar een oude bevinding met grote implicaties voor het lusextrusiemechanisme en de genomische architectuur: DNA-supercoiling door SMC-complexen. Hier gebruikten we magnetische pincetten om de torsie te meten tijdens individuele DNA-lusextrusiestappen door SMC-complexen. Net als de stap zelf ontdekten we dat DNA-torsie ook wordt veroorzaakt door ATP-binding. Opvallend genoeg vonden we voor alle drie types SMC complexen dezelfde afhankelijkheid van ATP-binding, dezelfde richting, en zelfs dezelfde grootte van de torsie die in elke stap in de lus wordt veroorzaakt: -0.6 windingen. Deze hoeveelheid is dus geconserveerd bij de eukaryotische SMC's en een integraal onderdeel van het lusextrusiemechanisme, dat tot nu toe niet in overweging is genomen. Eerder werden bulktests met behulp van plasmiden en topoisomerase gebruikt om de door SMC's veroorzaakte windingen te analyseren, wat onverwacht gedrag vertoonde en moeilijk te interpreteren was. Om bulk plasmide- en magnetische pincetgegevens met elkaar in overeenstemming te brengen, hebben we numeriek in simulaties in kaart gebracht hoe de eerder genoemde plasmidetest zich zou gedragen, waarbij alleen werd aangenomen dat elke lusextrusiestap -0.6 windingen in de lus induceert en bijgevolg +0.6 windingen in het DNA buiten de lus als een compenserende draai. De resultaten waren in goede overeenstemming met de experimentele plasmidetestresultaten, wat onze magnetische pincetgegevens ondersteunt.

In het derde deel van deze these richten we ons op het begrijpen of en hoe CTCF een obstakel vormt voor cohesin-afhankelijke lusextrusie. In hoofdstuk 7 hebben we DNA-

gebonden CTCF gereconstitueerd voor obstakelpassage-experimenten. We ontdekten dat diffunderende cohesin (zonder lus) vaak wordt gestopt wanneer het CTCF tegenkomt vanaf de N-terminale zijde. Zoals verwacht door chromosoomconformatievangst (chromosome conformation capture) experimenten, konden we ondubbelzinnig bevestigen dat CTCF inderdaad een barrière vormt voor lusextrusie door cohesin wanneer het vanaf de N-terminale zijde wordt tegengekomen. Echter, tot onze verrassing merkten we een sterke afhankelijkheid op van de blokkerende efficiëntie van CTCF met de DNA-spanning op het moment van de ontmoeting, een fenomeen waarvoor tot op heden een verklaring ontbreekt. Verder observeerden we een breed scala aan mogelijke uitkomsten van cohesin-CTCF-ontmoetingen. Bijvoorbeeld, cohesincomplexen konden hun extrusierichting omkeren bij het tegenkomen van CTCF, bij voorkeur na ontmoetingen waarbij de DNA-spanning relatief laag was. Alternatief konden cohesincomplexen meerdere keren proberen om CTCF te benaderen en te overwinnen na een gedeeltelijk verlies van de DNA lus veroorzaakt door de cohesin-CTCF-ontmoeting. Het laatstgenoemde scenario kwam vaker voor bij hogere DNA-spanning. Deze resultaten illustreren de dynamische aard van DNA-lusextrusie, een aspect dat vrijwel onmogelijk is om te meten in experimenten met levende cellen of chromatineconformatievangst.

In hoofdstuk 8 doken we in verschillende manieren waarop CTCF interfereert met cohesin om lusextrusie te stoppen. We hebben het N-terminale gebied van CTCF opgesplitst in verschillende fragmenten en cohesin blootgesteld aan de resulterende peptiden in oplossing. Met behulp van magnetische pincetten hebben we de frequentie van opeenvolgende lusextrusiestappen en voorwaartse-achterwaartse stapcombinaties ('omkeringen') gecontroleerd. Wanneer cohesine werd blootgesteld aan peptiden die het YxF-motief (residuen 226-228) bevatten, was het vaak niet in staat om opeenvolgende stappen van lusuitdrijving uit te voeren. Deze waarneming suggereert dat het YxF-motief interfereert met een stadium van de lusextrusiecyclus dat optreedt na ATP-binding, wat vaak leidt tot mislukken van het vergroten van de geëxtrudeerde lus. Bovendien bleken lussen gevisualiseerd met behulp van enkel-molecuul fluorescentie vaak eenzijdig te worden geëxtrudeerd, wat lijkt op de eenzijdige extrusie van gist-condensin vanwege zijn sterke DNA-binding aan de veiligheidsgordel (hoofdstuk 5 [98, 221]). Het is dus mogelijk dat de binding van het YxF-motief aan cohesin de DNA-bindingsaffiniteit ervan versterkt, waardoor effectief de waarschijnlijkheid van het wisselen van extrusierichtingen wordt verminderd. Naast het bekende YxF-motief hebben we een geannoteerde Pds5-bindingsplaats geïdentificeerd die zich bevindt aan het N-terminus van CTCF (KTYQR-motief; residuen 16-31), die de DNA-lusextrusie door cohesin onafhankelijk beïnvloedt. In tegenstelling tot het YxF-motief, belemmert het KTYQR-motief de algehele activiteit van cohesin, maar beïnvloedt het niet de frequentie van opeenvolgende of omkeringstappen, noch de richting van lusextrusie door cohesin.

Tot slot wordt het proefschrift afgesloten met een korte terugblik op de belangrijkste onderwerpen van SMC-onderzoek gedurende de afgelopen drie decennia. Mijn eigen onderzoek binnen het domein van SMC-eiwitten begon rond 2019, op een moment van grote opwinding over de demonstratie van DNA-lusextrusie. Het is bevredigend om te zien hoe het veld snel van koers veranderde en zich aanpaste aan deze opwindende bevindingen, waarmee de basis werd gelegd voor dit proefschrift. In het laatste hoofdstuk 9 bespreek ik

enkele perspectieven voor SMC-onderzoek op het niveau van enkele moleculen. In het bijzonder stel ik voor dat zowel geconserveerde als verschillende aspecten van lus extrusie onthuld kunnen worden door het analyseren van de diversiteit van de SMC superfamilie in verschillende domeinen van het leven. Verder bespreek ik kritische experimenten die de potentie hebben om de conformationele veranderingen die voorafgaan en volgen op de lusextrusiestap te onthullen *via* een combinatie van smFRET en magnetische pincetten en/of door een assay om de FRET-efficiëntie tussen de DNA-bindingsplaatsen van SMC's en DNA. Tot slot vraag ik me af waarom SMC's een relatief complex mechanisme lijken te benutten voor lusextrusie wanneer het in wezen gewoon een verankerde translocatie is. Ik concludeer dat lusextrusie efficiënt moet kunnen zorgen voor passage van obstakels *in vivo* om chromosomen vorm te geven zonder alle andere processen op DNA te verstoren. De grote ringvormige structuur lijkt ideaal geschikt te zijn zowel voor zusterchromatidecohesie als voor niet-topologische lusextrusie.

Zusammenfassung

Die sorgfältige Handhabung, Verarbeitung, Pflege und Segregation genetischen Materials ist eine heikle Aufgabe für jede Zelle. SMC Proteine spielen eine Rolle in vielen Aspekten des Lebens, die auf Genomen auftreten: DNA-Reparatur, Zusammenhalt und Trennung von Schwesterchromatiden, Chromosomenkondensation und Genregulierung. Zwei grundlegende Arbeitsprinzipien von SMCs sind bekannt: topologisches Umfassen von zwei DNA-Strängen und DNA-Loop-Extrusion. Ich habe die letzten vier Jahre darauf verwendet, Letzteres zumindest teilweise zu verstehen. Wir haben verschiedene Strategien eingesetzt, um einige grundlegende und weniger grundlegende Aspekte von Loop-Extrusion zu untersuchen, mit dem allumfassenden Ziel, einen DNA-Loop-Extrusionsmechanismus zu entwickeln, der alle verfügbaren Daten erklären kann.

In Kapitel 1 haben wir eine breit gefächerte Einführung in die grundlegenden Prinzipien der Genombiologie gegeben, wobei der Schwerpunkt auf der Schaffung des Rahmens für all die Ebenen lag, die SMC-Komplexe zur Organisation von Genomen in allen Ästen des Baum des Lebens beitragen. Wir haben die Aktivität von SMCs während der Zellteilung und Interphase vorgestellt, wonach wir uns mit der Beschreibung der SMC-Biologie aus biophysikalischer Sicht befasst haben: Loop-Extrusionsraten, Blockierkräfte, Richtung und Topologie, was uns zu einer Diskussion der derzeit vorgeschlagenen Modelle der DNA-Loop-Extrusion führte, das in der Darlegung des "Real & Seal"-Modells gipfelte. Schließlich haben wir gemeinsame und bisher gut erforschte Interaktionspartner sowie allgemeinere Blockaden eingeführt, die das Potenzial haben, DNA-Loop-Extrusion zu stoppen.

In Kapitel 3 haben wir versucht, DNA-Loop-Extruder durch die schiere Größe von Blockaden in ihrem Weg zu stoppen - und sind gescheitert. Die letzte Zuflucht im Versuch, Condensin effektiv daran zu hindern, über ein einzelnes Nanopartikel hinauszuextrudieren, bestand darin, das Partikel größer zu machen als einen Kreis mit einem Umfang gleich der Konturlänge der Peptidkette, die die ringförmige SMC-Struktur bildet. Doch selbst diese Blockade konnte in den Loop eindringen. Offensichtlich spielt die Größe doch nicht immer eine Rolle. Diese überraschenden Ergebnisse könnten auf zwei Arten interpretiert werden: Entweder öffnete sich eine Schnittstelle des SMC-Rings beim Passieren der Blockade, oder die DNA-Loop-Extrusion erfolgt nontopologisch. Um zwischen diesen Optionen zu unterscheiden, wandten wir uns einer modifizierten Version von menschlichem Cohesin zu, dessen Smc-Kleisin- und Smc-Smc-Schnittstellen vernetzt werden können. Entgegen der Annahme in der Fachwelt konnte auch diese Version von 'Single-Chain Cohesin' 200 nm große Blockaden in den Loop passieren lassen. Die Interpretation war daher klar: Die DNA-Loop-Extrusion ist im Gegensatz zur Zusammenhaltung von Schwesterchromatiden ein Prozess, bei dem SMCs nur nontopologisch an DNA gebunden sind. Während der Vorbereitung und Überarbeitung der begleitenden Veröffentlichung wurden pseudotopologische Modelle für die Loop-Extrusion vorgeschlagen. Angeregt durch diese neuen Ideen

identifizierten wir in Kapitel 4 mehrere Vorhersagen dieser vorgeschlagenen pseudotopologischen Modelle und testeten sie anhand unserer experimentellen Daten. Wir stellten fest, dass die Vorhersagen der pseudotopologischen Modelle im Widerspruch zu unseren Daten stehen. Stattdessen stimmen sie mit einem nontopologischen Modus der Loop-Extrusion überein.

In Kapitel 5 gingen wir einen anderen Weg, um die Loop-Extrusion durch verschiedene SMC-Komplexe zu charakterisieren. Frühere Demonstrationen von Loop-Extrusion durch Cohesin und SMC5/6 berichteten, dass diese Komplexe DNA gleichzeitig von beiden Seiten einholt, während *S. cerevisiae* Condensin DNA nur von einer Seite extrudierte. War es möglich, dass verschiedene eukaryotische SMC-Komplexe unterschiedliche Mechanismen für die Loop-Extrusion nutzen, bei denen die Loop-Extrusion keine Richtung hat? Um diese Frage zu beantworten, nahmen wir die Extrusionsdynamik von Condensin, Cohesin und SMC5/6 Seite an Seite und in relative einfachen, kontrollierten Umgebungen auf. Alle extrudierenden Komplexe waren Monomere, und keine Spannung von dem kontinuierlichen Fluss von Puffer wirkte auf die DNA. Unter diesen Bedingungen wurde deutlich, dass die Loop-Extrusion tatsächlich oft von Phasen durchsetzt ist, in denen SMCs zusammen mit ihrer extrudierten DNA diffundieren können und dass die extrudierte DNA allmählich in der Größe abnehmen können. Die Extrusionsphasen selbst konnten jedoch überzeugend als asymmetrische Extrusion charakterisiert werden. In Abwesenheit von Pufferfluss kann die Richtung dieser asymmetrischen Extrusion wechseln, was effektiv einer symmetrischen Extrusion entspricht, wenn sie über (ausreichend lange) Zeit gemittelt wird. Ab diesem Zeitpunkt konzentrierten wir uns auf Cohesin und konnten direkt einen Austausch seiner dynamischen Untereinheit NIPBL zwischen den Extrusionsphasen visualisieren. Genauer gesagt stellten wir fest, dass Umkehrungen der Extrusionsrichtung überwiegend einem Austausch von NIPBL an Cohesin entsprachen. Dies führte uns dazu, vorzuschlagen, dass NIPBL auf Cohesin als molekulares Getriebe wirkt, dessen Dissoziation und anschließende Assoziation mit Cohesin anscheinend für Richtungswechsel der Extrusion erforderlich ist. Wir schlossen unsere Ergebnisse mit einer Diskussion molekularer Determinanten für Richtungswechsel in allen SMC-Komplexen und einem potenziellen "DNA-Strangaustausch"-Modell ab, das Richtungswechsel im Rahmen des Real & Seal-Modells ermöglichen könnte.

In Kapitel 6 gehen wir auf eine alte Erkenntnis mit großen Auswirkungen auf den Loop-Extrusionsmechanismus und die Genomarchitektur zurück: Supercoiling von DNA durch SMC-Komplexe. Zuvor eingesetzte biochemische Assays mit Plasmiden und Topoisomerasen zur Visualisierung der durch die SMC verursachten Supercoils wiesen unerwartetes Verhalten auf und waren schwer zu interpretieren (auf das wir später im Kapitel hinweisen). Stattdessen verwendeten wir Magnetic Tweezers, um das Supercoilen von DNA während einzelner DNA-Loop-Extrusionsschritte durch SMC-Komplexe zu messen. Wie der Schritt selbst stellten wir fest, dass die DNA Supercoiling auch bei ATP-Bindung induziert wird. Bemerkenswerterweise fanden wir dieselbe Abhängigkeit von der ATP-Bindung, dieselbe Händigkeit und sogar dieselbe Größe des Supercoilings, die in jedem Schritt in den Loop induziert wird: -0,6 Umdrehungen. Diese Größe ist also unter den eukaryotischen SMCs konserviert und ein integraler Bestandteil des Loop-Extrusionsmechanismus, der bisher nicht berücksichtigt wurde. Wir validierten unsere Einzelmolekülergebnisse, indem wir

simulierten, wie sich das oben genannte Plasmid-Assay verhalten hätte, wenn in der Tat jeder Loop-Extrusionsschritt allein -0,6 Umdrehungen in den Loop und folglich +0,6 Umdrehungen in die DNA außerhalb des Loops induziert. Diese Anforderung reichte aus, um die experimentellen Plasmid-Assay-Ergebnisse weitgehend zu erklären und unsere Daten zu unterstützen.

Im dritten Teil dieser Arbeit wenden wir uns der Frage zu, ob und wie CTCF eine Blockade für die durch Cohesin vermittelte Loop-Extrusion ist. In Kapitel 7 rekonstituierten wir CTCF, das an DNA gebunden ist, für Experimente in denen Cohesin versucht, an CTCF vorbei zu kommen. Wir stellten fest, dass diffundierendes Cohesin (ohne Loop) häufig gestoppt wird, wenn es CTCF von seiner N-terminalen Seite trifft. Wie von Lebendzellenexperimenten erwartet, konnten wir bestätigen, dass CTCF tatsächlich auch eine Barriere für die Loop-Extrusion durch Cohesin ist, wenn es von seiner N-terminalen Seite her getroffen wird. Unerwarteterweise jedoch bemerkten wir eine starke Abhängigkeit der Blockiereffizienz von CTCF mit der DNA-Spannung zum Zeitpunkt der Begegnung. Ein Phänomen, für das bisher keine Erklärung vorliegt. Darüber hinaus beobachteten wir eine Vielzahl möglicher Ergebnisse von Cohesin-CTCF-Begegnungen. Zum Beispiel konnten Cohesin-Komplexe ihre Extrusionsrichtung bei der Begegnung mit CTCF umkehren, bevorzugt nach Begegnungen, bei denen die DNA-Spannung vergleichsweise niedrig war. Alternativ konnten Cohesin-Komplexe mehrmals versuchen, sich CTCF zu nähern und zu überwinden, nachdem durch die Cohesin-CTCF-Begegnung ein teilweiser Loopverlust verursacht wurde. Dieses Szenario trat bei höherer DNA-Spannung häufiger auf. Diese Ergebnisse veranschaulichen die dynamische Natur der DNA-Loop-Extrusion, ein Aspekt, der schwer bis unmöglich in Lebendzellenexperimenten zu erfassen ist.

In Kapitel 8 untersuchten wir den Mechanismus, durch den CTCF mit Cohesin interferiert, um die Extrusion von Loops zu stoppen. Wir segmentierten den N-terminalen Bereich von CTCF in verschiedene Teile und setzten Cohesin den resultierenden Peptiden in Lösung aus. Mit Hilfe von Magnetic Tweeezers beobachteten wir die Häufigkeit aufeinanderfolgender Loop-Extrusionsschritte und Vorwärts-Rückwärts-Schrittkombinationen ('provisorische' Schritte). Wenn Peptiden ausgesetzt, die das YxF-Motiv (Aminosäuren 226-228) enthalten, neigte Cohesin dazu, nicht mehr konsekutive Loop-Extrusionschritte zu nehmen. Diese Beobachtung deutet darauf hin, dass das YxF-Motiv eine Phase des Loop-Extrusionszyklus stört, die nach der ATP-Bindung auftritt und häufig dazu führt, dass der extrudierte Loop nicht vergrößert wird. Darüber hinaus neigten Loops, die mit Einzelmolekülfluoreszenz sichtbar gemacht wurden, dazu, unidirektional extrudiert zu werden, was der unidirektionalen Extrusion von *S. cerevisiae* Condensin aufgrund seiner starken DNA-Bindung am 'safety belt' ähnelt (Kapitel 5 [98, 221]). Es ist also möglich, dass die Bindung des YxF-Motivs an Cohesin dessen DNA-Bindungsaffinität stärkt und damit die Wahrscheinlichkeit eines Wechsels der Extrusionsrichtung effektiv verringert. Neben dem bekannten YxF-Motiv identifizierten wir eine annotierte Pds5-Bindungsstelle am N-Terminus von CTCF (KTYQR-Motiv; Aminosäuren 16-31), die die Extrusion des DNA Loops durch Cohesin beeinflusst. Im Gegensatz zum YxF-Motiv beeinträchtigt das KTYQR-Motiv die gesamte Stepping-Aktivität von Cohesin, hat aber ansonsten keinen Einfluss auf die Häufigkeit von konsekutiven Schritten oder die Richtung der Loop-Extrusion durch Cohesin.

Schließlich wird die Arbeit mit einer kurzen Rückschau auf die wichtigsten Themen der

SMC-Forschung der letzten drei Jahrzehnte abgeschlossen. Meine eigene Forschung im Bereich der SMC Proteine begann etwa 2019, zu einer Zeit großer Aufregung über die Demonstration der DNA-Loop-Extrusion. Es ist erfreulich zu sehen, wie das Feld schnell in Fahrt kam und sich diesen aufregenden Erkenntnissen anpasste. Im abschließenden Kapitel 9 diskutiere ich einige Perspektiven für die SMC-Forschung auf der Einzelmolekülebene. Insbesondere schlage ich vor, dass konservierte und unterschiedliche Aspekte der Loop-Extrusion durch Ausnutzung der Vielfalt der SMC-Superfamilie über den Lebensbaum hinweg aufgedeckt werden können, insbesondere durch Dimere von SMCs. Darüber hinaus diskutiere ich kritische Experimente, die das Potenzial haben, die Konformationsveränderungen aufzudecken, die dem Loop-Extrusionsschritt vorausgehen und folgen. Eine Kombination aus smFRET und magnetic tweezers und/oder durch ein Assay zur Untersuchung der FRET-Effizienz zwischen den DNA-Bindungsstellen von SMCs und DNA selbst hat das Potenzial, diese Konformationsveränderungen zeitlich und räumlich zu messen. Schließlich stelle ich die Frage, warum SMCs offenbar einen vergleichsweise komplexen Mechanismus für die Loop-Extrusion nutzen, wenn es im Kern einfach eine Translokation mit gleichzeitiger Verankerung ist. Ich schließe darauf, dass Loop-Extrusion *in vivo* die effiziente Passage von Hindernissen gewährleisten muss, um Chromosomen ihre Form zu verleihen, ohne alle anderen Prozesse auf der DNA zu stören. Die große ringförmige Struktur von SMCs scheint sowohl für den Zusammenhalt von Schwesterchromatiden als auch für die nontopologische Loop-Extrusion ideal geeignet zu sein.

References

[1] Eugene Kim, Roman Barth, and Cees Dekker. Looping the Genome with SMC Complexes. *Annual Review of Biochemistry*, 92(1):15–41, June 2023.

[2] Jaume Pellicer, Michael F. Fay, and Ilia J. Leitch. The largest eukaryotic genome of them all? *Botanical Journal of the Linnean Society*, 164(1):10–15, September 2010.

[3] Gordon M. Bennett and Nancy A. Moran. Small, Smaller, Smallest: The Origins and Evolution of Ancient Dual Symbioses in a Phloem-Feeding Insect. *Genome Biology and Evolution*, 5(9):1675–1688, September 2013.

[4] Tom Misteli. The Self-Organizing Genome: Principles of Genome Architecture and Function. *Cell*, 183(1):28–45, October 2020.

[5] Jan Soroczynski and Viviana I. Risca. Technological advances in probing 4D genome organization. *Current Opinion in Cell Biology*, 84:102211, October 2023.

[6] Job Dekker, Karsten Rippe, Martijn Dekker, and Nancy Kleckner. Capturing chromosome conformation. *Science*, 295(5558):1306–1311, 2002.

[7] Erez Lieberman-Aiden, Nynke L. Van Berkum, Louise Williams, Maxim Imakaev, Tobias Ragoczy, Agnes Telling, Ido Amit, Bryan R. Lajoie, Peter J. Sabo, Michael O. Dorschner, Richard Sandstrom, Bradley Bernstein, M. A. Bender, Mark Groudine, Andreas Gnirke, John Stamatoyannopoulos, Leonid A. Mirny, Eric S. Lander, and Job Dekker. Comprehensive Mapping of Long-Range Interactions Reveals Folding Principles of the Human Genome. *Science*, 326(5950):289–293, October 2009.

[8] Jian-Feng Xiang and Victor G Corces. Regulation of 3D chromatin organization by CTCF. *Current Opinion in Genetics & Development*, 67:33–40, April 2021.

[9] Andreas Bolzer, Gregor Kreth, Irina Solovei, Daniela Koehler, Kaan Saracoglu, Christine Fauth, Stefan Müller, Roland Eils, Christoph Cremer, Michael R. Speicher, and Thomas Cremer. Three-dimensional maps of all chromosomes in human male fibroblast nuclei and prometaphase rosettes. *PLoS Biology*, 3(5):0826–0842, 2005.

[10] T. Cremer and C. Cremer. Chromosome territories, nuclear architecture and gene regulation in mammalian cells. *Nature Reviews Genetics*, 2(4):292–301, 2001.

[11] Irina Solovei, Katharina Thanisch, and Yana Feodorova. How to rule the nucleus: Divide et impera. *Current Opinion in Cell Biology*, 40:47–59, June 2016.

[12] Tim J. Stevens, David Lando, Srinjan Basu, Liam P. Atkinson, Yang Cao, Steven F. Lee, Martin Leeb, Kai J. Wohlfahrt, Wayne Boucher, Aoife O'Shaughnessy-Kirwan, Julie Cramard, Andre J. Faure, Meryem Ralser, Enrique Blanco, Lluis Morey, Miriam Sansó, Matthieu G.S. Palayret, Ben Lehner, Luciano Di Croce, Anton Wutz, Brian Hendrich, Dave Klenerman, and Ernest D. Laue. 3D structures of individual mammalian genomes studied by single-cell Hi-C. *Nature*, 544(7648):59–64, April 2017.

[13] Martin Falk, Yana Feodorova, Natalia Naumova, Maxim Imakaev, Bryan R. Lajoie, Heinrich Leonhardt, Boris Joffe, Job Dekker, Geoffrey Fudenberg, Irina

Solovei, and Leonid A. Mirny. Heterochromatin drives compartmentalization of inverted and conventional nuclei. *Nature*, 570(7761):395–399, 2019.

[14] Adam G. Larson, Daniel Elnatan, Madeline M. Keenen, Michael J. Trnka, Jonathan B. Johnston, Alma L. Burlingame, David A. Agard, Sy Redding, and Geeta J. Narlikar. Liquid droplet formation by HP1α suggests a role for phase separation in heterochromatin. *Nature*, 547(7662):236–240, 2017.

[15] Amy R. Strom, Alexander V. Emelyanov, Mustafa Mir, Dmitry V. Fyodorov, Xavier Darzacq, and Gary H. Karpen. Phase separation drives heterochromatin domain formation. *Nature*, 547(7662):241–245, 2017.

[16] Hannah L. Harris, Huiya Gu, Moshe Olshansky, Ailun Wang, Irene Farabella, Yossi Eliaz, Achyuth Kalluchi, Akshay Krishna, Mozes Jacobs, Gesine Cauer, Melanie Pham, Suhas S. P. Rao, Olga Dudchenko, Arina Omer, Kiana Mohajeri, Sungjae Kim, Michael H. Nichols, Eric S. Davis, Dimos Gkountaroulis, Devika Udupa, Aviva Presser Aiden, Victor G. Corces, Douglas H. Phanstiel, William Stafford Noble, Guy Nir, Michele Di Pierro, Jeong-Sun Seo, Michael E. Talkowski, Erez Lieberman Aiden, and M. Jordan Rowley. Chromatin alternates between A and B compartments at kilobase scale for subgenic organization. *Nature Communications*, 14(1):3303, June 2023.

[17] Won Ki Cho, Jan Hendrik Spille, Micca Hecht, Choongman Lee, Charles Li, Valentin Grube, and Ibrahim I. Cisse. Mediator and RNA polymerase II clusters associate in transcription-dependent condensates. *Science*, 361(6400):412–415, 2018.

[18] Ibrahim I. Cisse, Ignacio Izeddin, Sebastien Z. Causse, Lydia Boudarene, Adrien Senecal, Leila Muresan, Claire Dugast-Darzacq, Bassam Hajj, Maxime Dahan, and Xavier Darzacq. Real-time dynamics of RNA polymerase II clustering in live human cells. *Science*, 341(6146):664–667, 2013.

[19] Rachel Patton McCord, Noam Kaplan, and Luca Giorgetti. Chromosome Conformation Capture and Beyond: Toward an Integrative View of Chromosome Structure and Function. *Molecular Cell*, 77(4):688–708, 2020.

[20] Miguel R Branco and Ana Pombo. Intermingling of Chromosome Territories in Interphase Suggests Role in Translocations and Transcription-Dependent Associations. *PLoS Biology*, 4(5):e138, April 2006.

[21] Takashi Nagano, Yaniv Lubling, Tim J. Stevens, Stefan Schoenfelder, Eitan Yaffe, Wendy Dean, Ernest D. Laue, Amos Tanay, and Peter Fraser. Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. *Nature*, 502(7469):59–64, October 2013.

[22] Shovamayee Maharana, K. Venkatesan Iyer, Nikhil Jain, Mallika Nagarajan, Yejun Wang, and G. V. Shivashankar. Chromosome intermingling—the physical basis of chromosome organization in differentiated cells. *Nucleic Acids Research*, 44(11):5148–5160, June 2016.

[23] Roger D. Kornberg. Chromatin Structure: A Repeating Unit of Histones and DNA: Chromatin structure is based on a repeating unit of eight histone molecules and about 200 DNA base pairs. *Science*, 184(4139):868–871, May 1974.

[24] Karolin Luger, Armin W. Mäder, Robin K. Richmond, David F. Sargent, and Timothy J. Richmond. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature*, 389(6648):251–260, 1997.

[25] C.L.F. Woodcock, J.P. Safer, and J.E. Stanchfield. Structural repeating units in chromatin. *Experimental Cell Research*, 97(1):101–110, January 1976.

[26] R. Barth, K. Bystricky, and H. A. Shaban. Coupling chromatin structure and dynamics by live super-resolution imaging. *Science Advances*, 6(27):eaaz2196, July 2020.

[27] Maria Aurelia Ricci, Carlo Manzo, María Filomena García-Parajo, Melike Lakadamyali, and Maria Pia Cosma. Chromatin fibers are formed by heterogeneous groups of nucleosomes in vivo. *Cell*, 160(6):1145–1158, 2015.

[28] Pablo Aurelio Gómez-García, Stephanie Portillo-Ledesma, Maria Victoria Neguembor, Martina Pesaresi, Walaa Oweis, Talia Rohrlich, Stefan Wieser, Eran Meshorer, Tamar Schlick, Maria Pia Cosma, and Melike Lakadamyali. Mesoscale Modeling and Single-Nucleosome Tracking Reveal Remodeling of Clutch Folding and Dynamics in Stem Cell Differentiation. *Cell Reports*, 34(2):108614, January 2021.

[29] Stephanie Portillo-Ledesma, Lucille H. Tsao, Meghna Wagley, Melike Lakadamyali, Maria Pia Cosma, and Tamar Schlick. Nucleosome Clutches are Regulated by Chromatin Internal Parameters. *Journal of Molecular Biology*, 433(6):166701, March 2021.

[30] Jianquan Xu, Hongqiang Ma, Jingyi Jin, Shikhar Uttam, Rao Fu, Yi Huang, and Yang Liu. Super-Resolution Imaging of Higher-Order Chromatin Structures at Different Epigenomic States in Single Mammalian Cells. *Cell Reports*, 24(4):873–882, 2018.

[31] Claire Hoencamp, Olga Dudchenko, Ahmed M. O. Elbatsh, Sumitabha Brahmachari, Jonne A. Raaijmakers, Tom Van Schaik, Ángela Sedeño Cacciatore, Vinícius G. Contessoto, Roy G. H. P. Van Heesbeen, Bram Van Den Broek, Aditya N. Mhaskar, Hans Teunissen, Brian Glenn St Hilaire, David Weisz, Arina D. Omer, Melanie Pham, Zane Colaric, Zhenzhen Yang, Suhas S. P. Rao, Namita Mitra, Christopher Lui, Weijie Yao, Ruqayya Khan, Leonid L. Moroz, Andrea Kohn, Judy St. Leger, Alexandria Mena, Karen Holcroft, Maria Cristina Gambetta, Fabian Lim, Emma Farley, Nils Stein, Alexander Haddad, Daniel Chauss, Ayse Sena Mutlu, Meng C. Wang, Neil D. Young, Evin Hildebrandt, Hans H. Cheng, Christopher J. Knight, Theresa L. U. Burnham, Kevin A. Hovel, Andrew J. Beel, Pierre-Jean Mattei, Roger D. Kornberg, Wesley C. Warren, Gregory Cary, José Luis Gómez-Skarmeta, Veronica Hinman, Kerstin Lindblad-Toh, Federica Di Palma, Kazuhiro Maeshima, Asha S. Multani, Sen Pathak, Liesl Nel-Themaat, Richard R. Behringer, Parwinder Kaur, René H. Medema, Bas Van Steensel, Elzo De Wit, José N. Onuchic, Michele Di Pierro, Erez Lieberman Aiden, and Benjamin D. Rowland. 3D genomics across the tree of life reveals condensin II as a determinant of architecture type. *Science*, 372(6545):984–989, May 2021.

[32] Georg A. Busslinger, Roman R. Stocsits, Petra Van Der Lelij, Elin Axelsson, Antonio Tedeschi, Niels Galjart, and Jan Michael Peters. Cohesin is positioned in mammalian genomes by transcription, CTCF and Wapl. *Nature*, 544(7651):503–507, 2017.

[33] Elphège P. Nora, Anton Goloborodko, Anne Laure Valton, Johan H. Gibcus, Alec Uebersohn, Nezar Abdennur, Job Dekker, Leonid A. Mirny, and Benoit G. Bruneau. Targeted Degradation of CTCF Decouples Local Insulation of Chromosome Domains from Genomic Compartmentalization. *Cell*, 169(5):930–944.e22, 2017.

[34] Gordana Wutz, Csilla Várnai, Kota Nagasaka, David A Cisneros, Roman R Stocsits, Wen Tang, Stefan Schoenfelder, Gregor Jessberger, Matthias Muhar, M Julius Hossain, Nike Walther, Birgit Koch, Moritz Kueblbeck, Jan Ellenberg, Johannes Zuber, Peter Fraser, and Jan-Michael Peters. Topologically associating domains and chromatin loops depend on cohesin and are regulated by CTCF, WAPL, and PDS5 proteins. *The EMBO Journal*, 36(24):3573–3599, December 2017.

[35] Iain F. Davidson, Roman Barth, Maciej Zaczek, Jaco Van Der Torre, Wen Tang, Kota Nagasaka, Richard Janissen, Jacob Kerssemakers, Gordana Wutz, Cees Dekker, and Jan-Michael Peters. CTCF is a DNA-tension-dependent barrier to cohesin-mediated loop extrusion. *Nature*, 616(7958):822–827, April 2023.

[36] Abraham S. Weintraub, Charles H. Li, Alicia V. Zamudio, Alla A. Sigova, Nancy M. Hannett, Daniel S. Day, Brian J. Abraham, Malkiel A. Cohen, Behnam Nabet, Dennis L. Buckley, Yang Eric Guo, Denes Hnisz, Rudolf Jaenisch, James E. Bradner, Nathanael S. Gray, and Richard A. Young. YY1 Is a Structural Regulator of Enhancer-Promoter Loops. *Cell*, 171(7):1573–1588.e28, 2017.

[37] Hideharu Hashimoto, Dongxue Wang, John R. Horton, Xing Zhang, Victor G. Corces, and Xiaodong Cheng. Structural Basis for the Versatile and Methylation-Dependent Binding of CTCF to DNA. *Molecular Cell*, 66(5):711–720.e3, June 2017.

[38] Maolu Yin, Jiuyu Wang, Min Wang, Xinmei Li, Mo Zhang, Qiang Wu, and Yanli Wang. Molecular mechanism of directional CTCF recognition of a diverse range of genomic sites. *Cell Research*, 27(11):1365–1377, November 2017.

[39] Jesse R. Dixon, Siddarth Selvaraj, Feng Yue, Audrey Kim, Yan Li, Yin Shen, Ming Hu, Jun S. Liu, and Bing Ren. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature*, 485(7398):376–380, May 2012.

[40] Ilya M. Flyamer, Johanna Gassler, Maxim Imakaev, Hugo B. Brandão, Sergey V. Ulianov, Nezar Abdennur, Sergey V. Razin, Leonid A. Mirny, and Kikuë Tachibana-Konwalski. Single-nucleus Hi-C reveals unique chromatin reorganization at oocyte-to-zygote transition. *Nature*, 544(7648):110–114, 2017.

[41] Elphège P. Nora, Bryan R. Lajoie, Edda G. Schulz, Luca Giorgetti, Ikuhiro Okamoto, Nicolas Servant, Tristan Piolot, Nynke L. Van Berkum, Johannes Meisig, John Sedat, Joost Gribnau, Emmanuel Barillot, Nils Blüthgen, Job Dekker, and Edith Heard. Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature*, 485(7398):381–385, 2012.

[42] Geoffrey Fudenberg, Maxim Imakaev, Carolyn Lu, Anton Goloborodko, Nezar Abdennur, and Leonid A. Mirny. Formation of Chromosomal Domains by Loop Extrusion. *Cell Reports*, 15(9):2038–2049, 2016.

[43] Johannes Nuebler, Geoffrey Fudenberg, Maxim Imakaev, Nezar Abdennur, and Leonid A. Mirny. Chromatin organization by an interplay of loop extrusion and compartmental segregation. *Proceedings of the National Academy of Sciences of the United States of America*, 115(29):E6697–E6706, 2018.

[44] Mattia Conte, Luca Fiorillo, Simona Bianco, Andrea M. Chiariello, Andrea Esposito, and Mario Nicodemi. Polymer physics indicates chromatin folding variability across single-cells results from state degeneracy in phase separation. *Nature Communications*, 11(1):3289, July 2020.

[45] Tung B. K. Le, Maxim V. Imakaev, Leonid A. Mirny, and Michael T. Laub. High-Resolution Mapping of the Spatial Organization of a Bacterial Chromosome. *Science*, 342(6159):731–734, November 2013.

[46] Amaury Bignaud, Charlotte Cockram, Céline Borde, Justine Groseille, Eric Allemand, Agnès Thierry, Martial Marbouty, Julien Mozziconacci, Olivier Espéli, and Romain Koszul. Transcription-induced domains form the elementary constraining building blocks of bacterial chromosomes. *Nature Structural & Molecular Biology*, January 2024.

[47] Tung Bk Le and Michael T Laub. Transcription rate and transcript length drive formation of chromosomal interaction domain boundaries. *The EMBO Journal*, 35(14):1582–1595, July 2016.

[48] Keiji Kimura and Tatsuya Hirano. ATP-dependent positive supercoiling of DNA by 13S condensin: A biochemical implication for chromosome condensation. *Cell*, 90(4):625–634, 1997.

[49] Belén Martínez-García, Sílvia Dyson, Joana Segura, Alba Ayats, Erin E Cutts, Pilar Gutierrez-Escribano, Luís Aragón, and Joaquim Roca. Condensin pinches a short negatively supercoiled DNA loop during each round of ATP usage. *The EMBO Journal*, 42(3):e111913, February 2023.

[50] Richard Janissen, Roman Barth, Iain F. Davidson, Michael Taschner, Stephan Gruber, Jan-Michael Peters, and Cees Dekker. All eukaryotic SMC proteins induce a twist of -0.6 at each DNA-loop-extrusion step, March 2024.

[51] Iain F. Davidson, Roman Barth, Sabrina Horn, Richard Janissen, Kota Nagasaka, Gordana Wutz, Roman R. Stocsits, Benedikt Bauer, Cees Dekker, and Jan-Michael Peters. Cohesin supercoils DNA during loop extrusion, March 2024.

[52] J Vinograd, J Lebowitz, R Radloff, R Watson, and P Laipis. The twisted circular form of polyoma viral DNA. *Proceedings of the National Academy of Sciences*, 53(5):1104–1111, May 1965.

[53] Hugh-G. Patterton and Claus Von Holt. Negative Supercoiling and Nucleosome Cores. *Journal of Molecular Biology*, 229(3):623–636, February 1993.

[54] Richard Janissen, Roman Barth, Minco Polinder, Jaco van der Torre, and Cees Dekker. Single-molecule visualization of twin-supercoiled domains generated during transcription. *Nucleic Acids Research*, page gkad1181, December 2023.

[55] L. F. Liu and J. C. Wang. Supercoiling of the DNA template during transcription. *Proceedings of the National Academy of Sciences of the United States of America*, 84(20):7024–7027, 1987.

[56] Ziyuan Wang and Peter Dröge. Long-range effects in a supercoiled DNA domain generated by transcription in vitro 1 1Edited by M. Yaniv. *Journal of Molecular Biology*, 271(4):499–510, August 1997.

[57] Hai-Young Wu, Shihua Shyy, James C. Wang, and Leroy F. Liu. Transcription generates positively and negatively supercoiled domains in the template. *Cell*, 53(3):433–440, May 1988.

[58] Nick Gilbert and James Allan. Supercoiling in DNA and chromatin. *Current Opinion in Genetics and Development*, 25(1):15–21, April 2014.

[59] J. F. Marko and E. D. Siggia. Statistical mechanics of supercoiled DNA. *Physical Review E*, 52(3):2912–2938, September 1995.

[60] T.R. Strick, J.-F. Allemand, D. Bensimon, and V. Croquette. Behavior of Supercoiled DNA. *Biophysical Journal*, 74(4):2016–2028, April 1998.

[61] Karl Drlica. Control of bacterial DNA supercoiling. *Molecular Microbiology*, 6(4):425–433, February 1992.

[62] Zoya M. Petrushenko, Weifeng She, and Valentin V. Rybenkov. A new family of bacterial condensins. *Molecular Microbiology*, 81(4):881–896, August 2011.

[63] Hertwig Oscar. *Lehrbuch Der Entwicklungsgeschichte Des Menschen Und Der Wirbeltiere*. 1890.

[64] A. D. Riggs. DNA methylation and late replication probably aid cell memory, and type I DNA reeling could aid chromosome folding and enhancer function. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 326(1235):285–297, 1990.

[65] Zellsubstanz Flemming. *Kern Und Zellteilung*. 1882.

[66] Albrecht Kossel. *Ueber Die Chemische Zusammensetzung Der Peptone*, volume 3. 1879.

[67] James R. Paulson and U. K. Laemmli. The structure of histone-depleted metaphase chromosomes. *Cell*, 12(3):817–828, November 1977.

[68] Susanne Leidescher, Johannes Ribisel, Simon Ullrich, Yana Feodorova, Erica Hildebrand, Alexandra Galitsyna, Sebastian Bultmann, Stephanie Link, Katharina Thanisch, Christopher Mulholland, Job Dekker, Heinrich Leonhardt, Leonid Mirny, and Irina Solovei. Spatial organization of transcribed eukaryotic genes. *Nature Cell Biology*, 24(3):327–339, March 2022.

[69] H. MacGregor. *Lampbrush Chromosomes*, volume 36. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013.

[70] E. Englesberg, C. Squires, and F. Meronk. The L-arabinose operon in Escherichia coli B-r: A genetic demonstration of two functional states of the product of a regulator gene. *Proceedings of the National Academy of Sciences of the United States of America*, 62(4):1100–1107, April 1969.

[71] Hans Peter Müller, JoséM M. Sogo, and Walter Schaffner. An enhancer stimulates transcription in Trans when attached to the promoter via a protein bridge. *Cell*, 58(4):767–777, August 1989.

[72] C. Wood and S. Tonegawa. Diversity and joining segments of mouse immunoglobulin heavy chain genes are closely linked and in the same orientation: Implications for the joining mechanism. *Proceedings of the National Academy of Sciences of the United States of America*, 80(10):3030–3034, May 1983.

[73] K S Matthews. DNA looping. *Microbiological Reviews*, 56(1):123–136, March 1992.

[74] Leonid A. Mirny. The fractal globule as a model of chromatin architecture in the cell. *Chromosome Research*, 19(1):37–51, 2011.

[75] F. W. Studier and P. K. Bandyopadhyay. Model for how type I restriction enzymes select cleavage sites in DNA. *Proceedings of the National Academy of Sciences of the United States of America*, 85(13):4677–4681, July 1988.

[76] Ralf Seidel, John van Noort, Carsten van der Scheer, Joost G P Bloom, Nynke H Dekker, Christina F Dutta, Alex Blundell, Terence Robinson, Keith Firman, and Cees Dekker. Real-time observation of DNA translocation by the type I restriction modification enzyme EcoR124I. *Nature Structural & Molecular Biology*, 11(9):838–843, September 2004.

[77] V. L. Larionov, T. S. Karpova, N. Y. Kouprina, and G. A. Jouravleva. A mutant of Saccharomyces cerevisiae with impaired maintenance of centromeric plasmids. *Current Genetics*, 10(1):15–20, October 1985.

[78] A. V. Strunnikov, V. L. Larionov, and D. Koshland. SMC1: An essential yeast gene encoding a putative head-rod-tail protein is required for nuclear division and defines a new ubiquitous protein family. *Journal of Cell Biology*, 123(6 II):1635–1648, December 1993.

[79] H. Niki, A. Jaffe, R. Imamura, T. Ogura, and S. Hiraga. The new gene mukB codes for a 177 kd protein with coiled-coil domains involved in chromosome partitioning of E. coli. *EMBO Journal*, 10(1):183–193, 1991.

[80] H. Niki, R. Imamura, M. Kitaoka, K. Yamanaka, T. Ogura, and S. Hiraga. E.coli MukB protein involved in chromosome partition forms a homodimer with a rod-and-hinge structure having DNA binding and ATP/GTP binding activities. *EMBO Journal*, 11(13):5101–5109, December 1992.

[81] Alexander V. Strunnikov, Eileen Hogan, and Douglas Koshland. SMC2, a Saccharomyces cerevisiae gene essential for chromosome segregation and condensation, defines a subgroup within the SMC family. *Genes and Development*, 9(5):587–599, March 1995.

[82] Tatsuya Hirano and Timothy J. Mitchison. A heterodimeric coiled-coil protein required for mitotic chromosome condensation in vitro. *Cell*, 79(3):449–458, November 1994.

[83] John A. Schmiesing, Alexander R. Ball, Heather C. Gregson, Janet M. Alderton, Sharleen Zhou, and Kyoko Yokomori. Identification of two distinct human SMC protein complexes involved in mitotic chromosome dynamics. *Proceedings of the National Academy of Sciences of the United States of America*, 95(22):12906–12911, October 1998.

[84] Ana Losada, Michiko Hirano, and Tatsuya Hirano. Identification of *Xenopus* SMC protein complexes required for sister chromatid cohesion. *Genes & Development*, 12(13):1986–1997, July 1998.

[85] Rainer P. Birkenbihl and Suresh Subramani. The rad21 Gene Product of Schizosaccharomyces pombe Is a Nuclear, Cell Cycle-regulated Phosphoprotein. *Journal of Biological Chemistry*, 270(13):7703–7711, March 1995.

[86] L. Carramolino, B.C. Lee, A. Zaballos, A. Peled, I. Barthelemy, Y. Shav-Tal, I. Prieto, P. Carmi, Y. Gothelf, G. González De Buitrago, M. Aracil, G. Márquez, J.L. Barbero, and D. Zipori. SA-1, a nuclear protein encoded by one member of a novel gene family: Molecular cloning and detection in hemopoietic organs. *Gene*, 195(2):151–159, August 1997.

[87] Attila Tóth, Rafal Ciosk, Frank Uhlmann, Marta Galova, Alexander Schleiffer, and Kim Nasmyth. Yeast cohesin complex requires a conserved protein, Eco1p (Ctf7), to establish cohesion between sister chromatids during DNA replication. *Genes & development*, 13(3):320–333, 1999.

[88] Tatsuya Hirano, Ryuji Kobayashi, and Michiko Hirano. Condensins, chromosome condensation protein complexes containing XCAP- C, XCAP-E and a Xenopus homolog of the Drosophila Barren protein. *Cell*, 89(4):511–521, May 1997.

[89] A. R. Lehmann, M. Walicka, D. J. F. Griffiths, J. M. Murray, F. Z. Watts, S. McCready, and A. M. Carr. The *rad18* Gene of *Schizosaccharomyces pombe* Defines a New Subgroup of the SMC Superfamily Involved in DNA Repair. *Molecular and Cellular Biology*, 15(12):7067–7080, December 1995.

[90] M. I. Fousteri. A novel SMC protein complex in Schizosaccharomyces pombe contains the Rad18 DNA repair protein. *The EMBO Journal*, 19(7):1691–1702, April 2000.

[91] Yoko Fujioka, Yukio Kimata, Koji Nomaguchi, Kenji Watanabe, and Kenji Kohno. Identification of a Novel Non-structural Maintenance of Chromosomes (SMC) Component of the SMC5-SMC6 Complex Involved in DNA Repair. *Journal of Biological Chemistry*, 277(24):21585–21591, June 2002.

[92] Stephanie Pebernard, W. Hayes McDonald, Yelena Pavlova, John R. Yates, and Michael N. Boddy. Nse1, Nse2, and a Novel Subunit of the Smc5-Smc6 Complex, Nse3, Play a Crucial Role in Meiosis. *Molecular Biology of the Cell*, 15(11):4866–4876, November 2004.

[93] Stephanie Pebernard, James Wohlschlegel, W. Hayes McDonald, John R. Yates, and Michael N. Boddy. The Nse5-Nse6 Dimer Mediates DNA Repair Roles of the Smc5-Smc6 Complex. *Molecular and Cellular Biology*, 26(5):1617–1630, March 2006.

[94] Emily A. Andrews, Jan Palecek, John Sergeant, Elaine Taylor, Alan R. Lehmann, and Felicity Z. Watts. Nse2, a Component of the Smc5-6 Complex, Is a SUMO Ligase Required for the Response to DNA Damage. *Molecular and Cellular Biology*, 25(1):185–196, January 2005.

[95] Bin Hu, Chunyan Liao, Stefan H. Millson, Mehdi Mollapour, Chrisostomos Prodromou, Laurence H. Pearl, Peter W. Piper, and Barry Panaretou. Qri2/Nse4, a component of the essential Smc5/6 DNA repair complex. *Molecular Microbiology*, 55(6):1735–1750, March 2005.

[96] Gary Z. Wang and Stephen P. Goff. Regulation of Yin Yang 1 by tyrosine phosphorylation. *Journal of Biological Chemistry*, 290(36):21890–21900, 2015.

[97] Adrian L. Sanborn, Suhas S.P. Rao, Su Chen Huang, Neva C. Durand, Miriam H. Huntley, Andrew I. Jewett, Ivan D. Bochkov, Dharmaraj Chinnappan, Ashok Cutkosky, Jian Li, Kristopher P. Geeting, Andreas Gnirke, Alexandre Melnikov, Doug McKenna, Elena K. Stamenova, Eric S. Lander, and Erez Lieberman Aiden. Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. *Proceedings of the National Academy of Sciences of the United States of America*, 112(47):E6456–E6465, 2015.

[98] Mahipal Ganji, Indra A. Shaltiel, Shveta Bisht, Eugene Kim, Ana Kalichava, Christian H. Haering, and Cees Dekker. Real-time imaging of DNA loop extrusion by condensin. *Science*, 360(6384):102–105, April 2018.

[99] Jonathan N. Wells, Thomas G. Gligoris, Kim A. Nasmyth, and Joseph A. Marsh. Evolution of condensin and cohesin complexes driven by replacement of Kite by Hawk proteins. *Current Biology*, 27(1):R17–R18, 2017.

[100] Sotaro Kikuchi, Dominika M. Borek, Zbyszek Otwinowski, Diana R. Tomchick, and Hongtao Yu. Crystal structure of the cohesin loader Scc2 and insight into cohesinopathy. *Proceedings of the National Academy of Sciences of the United States of America*, 113(44):12444–12449, 2016.

[101] Naomi J. Petela, Thomas G. Gligoris, Jean Metson, Byung-Gil Lee, Menelaos Voulgaris, Bin Hu, Sotaro Kikuchi, Christophe Chapard, Wentao Chen, Eeson Rajendra, Madhusudhan Srinivisan, Hongtao Yu, Jan Löwe, and Kim A. Nasmyth. Scc2 Is a Potent Activator of Cohesin's ATPase that Promotes Loading by Binding Scc1 without Pds5. *Molecular Cell*, 70(6):1134–1148.e7, June 2018.

[102] Johannes Stigler, Gamze Çamdere, Douglas E. Koshland, and Eric C. Greene. Single-Molecule Imaging Reveals a Collapsed Conformational State for DNA-Bound Cohesin. *Cell Reports*, 15(5):988–998, 2016.

[103] Judith H.I. Haarhuis, Robin H. van der Weide, Vincent A. Blomen, J. Omar Yáñez-Cuna, Mario Amendola, Marjon S. van Ruiten, Peter H.L. Krijger, Hans Teunissen, René H. Medema, Bas van Steensel, Thijn R. Brummelkamp, Elzo de Wit, and Benjamin D. Rowland. The Cohesin Release Factor WAPL Restricts Chromatin Loop Extension. *Cell*, 169(4):693–707.e14, 2017.

[104] You Yu, Shibai Li, Zheng Ser, Huihui Kuang, Thane Than, Danying Guan, Xiaolan Zhao, Dinshaw J Patel, and Hongtao Yu. Cryo-EM structure of DNA-bound Smc5 / 6 reveals DNA clamping enabled by multi-subunit conformational changes. pages 1–10, 2022.

[105] Jianhua Song, Andrea Lafont, Jingrong Chen, Frank M. Wu, Katsuhiko Shirahige, and Susannah Rankin. Cohesin acetylation promotes sister chromatid cohesion only in association with the replication machinery. *Journal of Biological Chemistry*, 287(41):34325–34336, October 2012.

[106] Tom Rolef Ben-Shahar, Sebastian Heeger, Chris Lehane, Philip East, Helen Flynn, Mark Skehel, and Frank Uhlmann. Eco1-dependent cohesin acetylation during establishment of sister chromatid cohesion. *Science*, 321(5888):563–566, July 2008.

[107] Elçin Ünal, Jill M. Heidinger-Pauli, Woong Kim, Vincent Guacci, Itay Onn, Steven P. Gygi, and Douglas E. Koshland. A Molecular Determinant for the Establishment of Sister Chromatid Cohesion. *Science*, 321(5888):566–569, July 2008.

[108] Jinglan Zhang, Xiaomin Shi, Yehua Li, Beom Jun Kim, Junling Jia, Zhiwei Huang, Tao Yang, Xiaoyong Fu, Sung Yun Jung, Yi Wang, Pumin Zhang, Seong Tae Kim, Xuewen Pan, and Jun Qin. Acetylation of Smc3 by Eco1 Is Required for S Phase Sister Chromatid Cohesion in Both Human and Yeast. *Molecular Cell*, 31(1):143–151, July 2008.

[109] J. B. Rattner and C. C. Lin. Radial loops and helical coils coexist in metaphase chromosomes. *Cell*, 42(1):291–296, August 1985.

[110] Noriko Saitoh, Ilya G. Goldberg, Edgar R. Wood, and William C. Earnshaw. ScII: An abundant chromosome scaffold protein is a member of a family of putative ATPases with an unusual predicted tertiary structure. *Journal of Cell Biology*, 127(2):303–318, October 1994.

[111] Elnaz Alipour and John F. Marko. Self-organization of domain structures by DNA-loop-extruding enzymes. *Nucleic Acids Research*, 40(22):11202–11212, December 2012.

[112] Anton Goloborodko, Maxim V. Imakaev, John F. Marko, and Leonid Mirny. Compaction and segregation of sister chromatids via active loop extrusion. *eLife*, 5(MAY2016), May 2016.

[113] Anton Goloborodko, John F. Marko, and Leonid A. Mirny. Chromosome Compaction by Active Loop Extrusion. *Biophysical Journal*, 110(10):2162–2168, 2016.

[114] Toru Hirota, Daniel Gerlich, Birgit Koch, Jan Ellenberg, and Jan Michael Peters. Distinct functions of condensin I and II in mitotic chromosome assembly. *Journal of Cell Science*, 117(26):6435–6445, December 2004.

[115] Takao Ono, Ana Losada, Michiko Hirano, Michael P. Myers, Andrew F. Neuwald, and Tatsuya Hirano. Differential contributions of condensin I and condensin II to mitotic chromosome architecture in vertebrate cells. *Cell*, 115(1):109–121, October 2003.

[116] Johan H. Gibcus, Kumiko Samejima, Anton Goloborodko, Itaru Samejima, Natalia Naumova, Johannes Nuebler, Masato T. Kanemaki, Linfeng Xie, James R. Paulson, William C. Earnshaw, Leonid A. Mirny, and Job Dekker. A pathway for mitotic chromosome formation. *Science*, 359(6376), 2018.

[117] Nike Walther, M. Julius Hossain, Antonio Z. Politi, Birgit Koch, Moritz Kueblbeck, Øyvind Odegard-fougner, Marko Lampe, and Jan Ellenberg. A quantitative map of human Condensins provides new insights into mitotic chromosome architecture. *Journal of Cell Biology*, 217(7):2309–2328, 2018.

[118] K. Nasmyth. Disseminating the genome: Joining, resolving, and separating sister chromatids during mitosis and meiosis. *Annual Review of Genetics*, 35:673–745, 2001.

[119] Stephan Gruber and Jeff Errington. Recruitment of Condensin to Replication Origin Regions by ParB/SpoOJ Promotes Chromosome Segregation in B. subtilis. *Cell*, 137(4):685–696, May 2009.

[120] Xindan Wang, Hugo B. Brandão, Tung B.K. Le, Michael T. Laub, and David Z. Rudner. Bacillus subtilis SMC complexes juxtapose chromosome arms as they travel from origin to terminus. *Science*, 355(6324):524–527, February 2017.

[121] Xindan Wang, Anna C. Hughes, Hugo B. Brandão, Benjamin Walker, Carrie Lierz, Jared C. Cochran, Martha G. Oakley, Andrew C. Kruse, and David Z. Rudner. In Vivo Evidence for ATPase-Dependent DNA Translocation by the Bacillus subtilis SMC Condensin Complex. *Molecular Cell*, 71(5):841–847.e5, 2018.

[122] Christos Gogou, Aleksandre Japaridze, and Cees Dekker. Mechanisms for Chromosome Segregation in Bacteria. *Frontiers in Microbiology*, 12, June 2021.

[123] Suckjoon Jun and Bela Mulder. Entropy-driven spatial organization of highly confined polymers: Lessons for the bacterial chromosome. *Proceedings of the National Academy of Sciences of the United States of America*, 103(33):12388–12393, August 2006.

[124] Aleksandre Japaridze, Raman Van Wee, Christos Gogou, Jacob W. J. Kerssemakers, Daan F. Van Den Berg, and Cees Dekker. MukBEF-dependent chromosomal organization in widened Escherichia coli. *Frontiers in Microbiology*, 14:1107093, March 2023.

[125] Sophie Nolivos, Amy L. Upton, Anjana Badrinarayanan, Julius Müller, Katarzyna Zawadzka, Jakub Wiktor, Amber Gill, Lidia Arciszewska, Emilien Nicolas, and David Sherratt. MatP regulates the coordinated action of topoisomerase IV and MukBEF in chromosome segregation. *Nature Communications*, 7(1):10466, April 2016.

[126] Jarno Mäkelä and David J Sherratt. Organization of the Escherichia coli Chromosome by a MukBEF Axial Core. *Molecular Cell*, 78(2):250–260.e5, 2020.

[127] Madhusudhan Srinivasan, Marco Fumasoni, Naomi J. Petela, Andrew Murray, and Kim A. Nasmyth. Cohesion is established during dna replication utilising chromosome associated cohesin rings as well as those loaded de novo onto nascent dnas. *eLife*, 9:1–27, June 2020.

[128] Tatsuro S. Takahashi, Pannyun Yiu, Michael F. Chou, Steven Gygi, and Johannes C. Walter. Recruitment of Xenopus Scc2 and cohesin to chromatin requires the pre-replication complex. *Nature Cell Biology*, 6(10):991–996, October 2004.

[129] Tatsuro S. Takahashi, Abhijit Basu, Vladimir Bermudez, Jerard Hurwitz, and Johannes C. Walter. Cdc7–Drf1 kinase links chromosome cohesion to the initiation of DNA replication in *Xenopus* egg extracts. *Genes & Development*, 22(14):1894–1905, July 2008.

[130] Ge Zheng, Mohammed Kanchwala, Chao Xing, and Hongtao Yu. MCM2–7-dependent cohesin loading during S phase promotes sister-chromatid cohesion. *eLife*, 7, April 2018.

[131] Stephan Gruber, Christian H. Haering, and Kim Nasmyth. Chromosomal cohesin forms a ring. *Cell*, 112(6):765–777, March 2003.

[132] Christian H. Haering, Ana Maria Farcas, Prakash Arumugam, Jean Metson, and Kim Nasmyth. The cohesin ring concatenates sister DNA molecules. *Nature*, 454(7202):297–301, July 2008.

[133] Itay Onn, Jill M. Heidinger-Pauli, Vincent Guacci, Elçin Ünal, and Douglas E. Koshland. Sister Chromatid Cohesion: A Simple Concept with a Complex Reality. *Annual Review of Cell and Developmental Biology*, 24(1):105–129, November 2008.

[134] J.-M. Peters and T. Nishiyama. Sister Chromatid Cohesion. *Cold Spring Harbor Perspectives in Biology*, 4(11):a011130–a011130, November 2012.

[135] Yasuto Murayama, Catarina P. Samora, Yumiko Kurokawa, Hiroshi Iwasaki, and Frank Uhlmann. Establishment of DNA-DNA Interactions by the Cohesin Ring. *Cell*, 172(3):465–477.e15, January 2018.

[136] Masashi Minamino, Céline Bouchoux, Berta Canal, John F.X. Diffley, and Frank Uhlmann. A replication fork determinant for the establishment of sister chromatid cohesion. *Cell*, 186(4):837–849.e11, February 2023.

[137] Pilar Gutierrez-Escribano, Matthew D. Newton, Aida Llauró, Jonas Huber, Loredana Tanasie, Joseph Davy, Isabel Aly, Ricardo Aramayo, Alex Montoya, Holger

Kramer, Johannes Stigler, David S. Rueda, and Luis Aragon. A conserved ATP- And Scc2/4-dependent activity for cohesin in tethering DNA molecules. *Science Advances*, 5(11), 2019.

[138] Minzhe Tang, Georgii Pobegalov, Hideki Tanizawa, Zhuo A. Chen, Juri Rappsilber, Maxim Molodtsov, Ken-ichi Noma, and Frank Uhlmann. Establishment of dsDNA-dsDNA interactions by the condensin complex. *Molecular Cell*, page S1097276523007463, October 2023.

[139] Kota Nagasaka, Iain F. Davidson, Roman R. Stocsits, Wen Tang, Gordana Wutz, Paul Batty, Melanie Panarotto, Gabriele Litos, Alexander Schleiffer, Daniel W. Gerlich, and Jan-Michael Peters. Cohesin mediates DNA loop extrusion and sister chromatid cohesion by distinct mechanisms. *Molecular Cell*, 83(17):3049–3063.e6, September 2023.

[140] Yu Liu and Job Dekker. CTCF–CTCF loops and intra-TAD interactions show differential dependence on cohesin ring integrity. *Nature Cell Biology*, October 2022.

[141] Yan Li, Judith H.I. Haarhuis, Ángela Sedeño Cacciatore, Roel Oldenkamp, Marjon S. van Ruiten, Laureen Willems, Hans Teunissen, Kyle W. Muir, Elzo de Wit, Benjamin D. Rowland, and Daniel Panne. The structural basis for cohesin–CTCF-anchored loops. *Nature*, 578(7795):472–476, 2020.

[142] George Cameron, Dominika Gruszka, Sherry Xie, Kim A Nasmyth, Madhusudhan Srinivasan, and Hasan Yardimci. Sister chromatid cohesion establishment during DNA replication termination. Preprint, Biophysics, September 2022.

[143] Torahiko L. Higashi, Megumi Ikeda, Hiroshi Tanaka, Takuro Nakagawa, Masashige Bando, Katsuhiko Shirahige, Yumiko Kubota, Haruhiko Takisawa, Hisao Masukata, and Tatsuro S. Takahashi. The Prereplication Complex Recruits XEco2 to Chromatin to Promote Cohesin Acetylation in Xenopus Egg Extracts. *Current Biology*, 22(11):977–988, June 2012.

[144] Rene Ladurner, Emanuel Kreidl, Miroslav P Ivanov, Heinz Ekker, Maria Helena Idarraga-Amado, Georg A Busslinger, Gordana Wutz, David A Cisneros, and Jan-Michael Peters. Sororin actively maintains sister chromatid cohesion. *The EMBO Journal*, 35(6):635–653, March 2016.

[145] Tomoko Nishiyama, Rene Ladurner, Julia Schmitz, Emanuel Kreidl, Alexander Schleiffer, Venugopal Bhaskara, Masashige Bando, Katsuhiko Shirahige, Anthony A. Hyman, Karl Mechtler, and Jan-Michael Peters. Sororin Mediates Sister Chromatid Cohesion by Antagonizing Wapl. *Cell*, 143(5):737–749, November 2010.

[146] Sara Cuylen, Jutta Metz, and Christian H. Haering. Condensin structures chromosomal DNA through topological links. *Nature Structural and Molecular Biology*, 18(8):894–901, 2011.

[147] Kazuhisa Kinoshita, Yuko Tsubota, Shoji Tane, Yuuki Aizawa, Ryota Sakata, Kozo Takeuchi, Keishi Shintomi, Tomoko Nishiyama, and Tatsuya Hirano. A loop extrusion–independent mechanism contributes to condensin I–mediated chromosome shaping. *Journal of Cell Biology*, 221(3):e202109016, March 2022.

[148] Suhas S.P. Rao, Miriam H. Huntley, Neva C. Durand, Elena K. Stamenova, Ivan D. Bochkov, James T. Robinson, Adrian L. Sanborn, Ido Machol, Arina D. Omer, Eric S. Lander, and Erez Lieberman Aiden. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell*, 159(7):1665–1680, 2014.

[149] Suzana Hadjur, Luke M. Williams, Natalie K. Ryan, Bradley S. Cobb, Tom Sexton, Peter Fraser, Amanda G. Fisher, and Matthias Merkenschlager. Cohesins form

chromosomal cis-interactions at the developmentally regulated IFNG locus. *Nature*, 460(7253):410–413, July 2009.

[150] Vania Parelho, Suzana Hadjur, Mikhail Spivakov, Marion Leleu, Stephan Sauer, Heather C. Gregson, Adam Jarmuz, Claudia Canzonetta, Zoe Webster, Tatyana Nesterova, Bradley S. Cobb, Kyoko Yokomori, Niall Dillon, Luis Aragon, Amanda G. Fisher, and Matthias Merkenschlager. Cohesins Functionally Associate with CTCF on Mammalian Chromosome Arms. *Cell*, 132(3):422–433, February 2008.

[151] Kerstin S. Wendt, Keisuke Yoshida, Takehiko Itoh, Masashige Bando, Birgit Koch, Erika Schirghuber, Shuichi Tsutsumi, Genta Nagae, Ko Ishihara, Tsuyoshi Mishiro, Kazuhide Yahata, Fumio Imamoto, Hiroyuki Aburatani, Mitsuyoshi Nakao, Naoko Imamoto, Kazuhiro Maeshima, Katsuhiko Shirahige, and Jan-Michael Peters. Cohesin mediates transcriptional insulation by CCCTC-binding factor. *Nature*, 451(7180):796–801, 2008-02-30.

[152] Johanna Gassler, Hugo B Brandão, Maxim Imakaev, Ilya M Flyamer, Sabrina Ladstätter, Wendy A Bickmore, Jan-Michael Peters, Leonid A Mirny, and Kikuë Tachibana. A mechanism of cohesin-dependent loop extrusion organizes zygotic genome architecture. *The EMBO Journal*, 36(24):3600–3618, December 2017.

[153] Suhas S.P. Rao, Su Chen Huang, Brian Glenn St Hilaire, Jesse M. Engreitz, Elizabeth M. Perez, Kyong Rim Kieffer-Kwon, Adrian L. Sanborn, Sarah E. Johnstone, Gavin D. Bascom, Ivan D. Bochkov, Xingfan Huang, Muhammad S. Shamim, Jaeweon Shin, Douglass Turner, Ziyi Ye, Arina D. Omer, James T. Robinson, Tamar Schlick, Bradley E. Bernstein, Rafael Casellas, Eric S. Lander, and Erez Lieberman Aiden. Cohesin Loss Eliminates All Loop Domains. *Cell*, 171(2):305–320.e24, 2017.

[154] Wibke Schwarzer, Nezar Abdennur, Anton Goloborodko, Aleksandra Pekowska, Geoffrey Fudenberg, Yann Loe-Mie, Nuno A Fonseca, Wolfgang Huber, Christian H. Haering, Leonid Mirny, and Francois Spitz. Two independent modes of chromatin organization revealed by cohesin removal. *Nature*, 551(7678):51–56, November 2017.

[155] Elzo de Wit, Erica S.M. Vos, Sjoerd J.B. Holwerda, Christian Valdes-Quezada, Marjon J.A.M. Verstegen, Hans Teunissen, Erik Splinter, Patrick J. Wijchers, Peter H.L. Krijger, and Wouter de Laat. CTCF Binding Polarity Determines Chromatin Looping. *Molecular Cell*, 60(4):676–684, November 2015.

[156] Michael H. Nichols and Victor G. Corces. A CTCF Code for 3D Genome Architecture. *Cell*, 162(4):703–705, August 2015.

[157] Antonio Tedeschi, Gordana Wutz, Sébastien Huet, Markus Jaritz, Annelie Wuensche, Erika Schirghuber, Iain Finley Davidson, Wen Tang, David A. Cisneros, Venugopal Bhaskara, Tomoko Nishiyama, Alipasha Vaziri, Anton Wutz, Jan Ellenberg, and Jan-Michael Peters. Wapl is an essential regulator of chromatin structure and chromosome segregation. *Nature*, 501(7468):564–568, September 2013.

[158] Matteo Vietri Rudan, Christopher Barrington, Stephen Henderson, Christina Ernst, Duncan T. Odom, Amos Tanay, and Suzana Hadjur. Comparative Hi-C Reveals that CTCF Underlies Evolution of Chromosomal Domain Architecture. *Cell Reports*, 10(8):1297–1309, March 2015.

[159] Ks Beckwith, Ø Ødegård-Fougner, Nr Morero, C Barton, F Schueder, W Tang, S Alexander, Jm Peters, R Jungmann, E Birney, and J Ellenberg. Visualization of loop extrusion by DNA nanoscale tracing in single human cells. Preprint, Cell Biology, April 2021.

[160] Michele Gabriele, Hugo B. Brandão, Simon Grosse-Holz, Asmita Jha, Gina M. Dailey, Claudia Cattoglio, Tsung Han S. Hsieh, Leonid Mirny, Christoph Zechner, and Anders S. Hansen. Dynamics of CTCF-

and cohesin-mediated chromatin looping revealed by live-cell imaging. *Science*, 376(6592):476–501, April 2022.

[161] Pia Mach, Pavel I. Kos, Yinxiu Zhan, Julie Cramard, Simon Gaudin, Jana Tünnermann, Edoardo Marchi, Jan Eglinger, Jessica Zuin, Mariya Kryzhanovska, Sebastien Smallwood, Laurent Gelman, Gregory Roth, Elphège P. Nora, Guido Tiana, and Luca Giorgetti. Cohesin and CTCF control the dynamics of chromosome folding. *Nature Genetics*, 54(12):1907–1918, December 2022.

[162] Elizabeth H. Finn, Gianluca Pegoraro, Hugo B. Brandão, Anne Laure Valton, Marlies E. Oomen, Job Dekker, Leonid Mirny, and Tom Misteli. Extensive Heterogeneity and Intrinsic Variation in Spatial Genome Organization. *Cell*, 176(6):1502–1515.e10, 2019.

[163] Emily M. Smith, Bryan R. Lajoie, Gaurav Jain, and Job Dekker. Invariant TAD Boundaries Constrain Cell-Type-Specific Looping Interactions between Promoters and Distal Elements around the CFTR Locus. *American Journal of Human Genetics*, 98(1):185–201, January 2016.

[164] Malte Spielmann, Darío G. Lupiáñez, and Stefan Mundlos. Structural variation in the 3D genome. *Nature Reviews Genetics*, 19(7):453–467, July 2018.

[165] Jessica Zuin, Gregory Roth, Yinxiu Zhan, Julie Cramard, Josef Redolfi, Ewa Piskadlo, Pia Mach, Mariya Kryzhanovska, Gergely Tihanyi, Hubertus Kohler, Mathias Eder, Christ Leemans, Bas van Steensel, Peter Meister, Sebastien Smallwood, and Luca Giorgetti. Nonlinear control of transcription through enhancer–promoter interactions. *Nature*, 604(7906):571–577, April 2022.

[166] Boryana Doyle, Geoffrey Fudenberg, Maxim Imakaev, and Leonid A. Mirny. Chromatin Loops as Allosteric Modulators of Enhancer-Promoter Interactions. *PLoS Computational Biology*, 10(10), 2014.

[167] Orsolya Symmons, Leslie Pan, Silvia Remeseiro, Tugce Aktas, Felix Klein, Wolfgang Huber, and François Spitz. The Shh Topological Domain Facilitates the Action of Remote Enhancers by Reducing the Effects of Genomic Distances. *Developmental Cell*, 39(5):529–543, December 2016.

[168] Jan Krefting, Miguel A. Andrade-Navarro, and Jonas Ibn-Salem. Evolutionary stability of topologically associating domains is associated with conserved gene regulation. *BMC Biology*, 16(1):1–12, August 2018.

[169] Jeffrey M. Alexander, Juan Guan, Bingkun Li, Lenka Maliskova, Michael Song, Yin Shen, Bo Huang, Stavros Lomvardas, and Orion D. Weiner. Live-cell imaging reveals enhancer-dependent sox2 transcription in the absence of enhancer proximity. *eLife*, 8, May 2019.

[170] Nezha S. Benabdallah, Iain Williamson, Robert S. Illingworth, Lauren Kane, Shelagh Boyle, Dipta Sengupta, Graeme R. Grimes, Pierre Therizols, and Wendy A. Bickmore. Decreased Enhancer-Promoter Proximity Accompanying Enhancer Activation. *Molecular Cell*, 76(3):473–484.e7, 2019.

[171] Yad Ghavi-Helm, Aleksander Jankowski, Sascha Meiers, Rebecca R. Viales, Jan O. Korbel, and Eileen E.M. Furlong. Highly rearranged chromosomes reveal uncoupling between genome topology and gene expression. *Nature Genetics*, 51(8):1272–1282, August 2019.

[172] Iain Williamson, Lauren Kane, Paul S. Devenney, Ilya M. Flyamer, Eve Anderson, Fiona Kilanowski, Robert E. Hill, Wendy A. Bickmore, and Laura A. Lettice. Developmentally regulated Shh expression is robust to TAD perturbations. *Development*, January 2019.

[173] Iain F. Davidson, Benedikt Bauer, Daniela Goetz, Wen Tang, Gordana Wutz, and Jan Michael Peters. DNA loop extrusion by human cohesin. *Science*, 366(6471):1338–1345, December 2019.

[174] Hugo B. Brandão, Payel Paul, Aafke A. van den Berg, David Z. Rudner, Xindan Wang, and Leonid A. Mirny. RNA polymerases as moving barriers to condensin loop extrusion. *Proceedings of the National Academy of Sciences of the United States of America*, 116(41):20489–20499, 2019.

[175] Edward J. Banigan, Wen Tang, Aafke A. Van Den Berg, Roman R. Stocsits, Gordana Wutz, Hugo B. Brandão, Georg A. Busslinger, Jan-Michael Peters, and Leonid A. Mirny. Transcription shapes 3D chromatin organization by interacting with loop extrusion. *Proceedings of the National Academy of Sciences*, 120(11):e2210480120, March 2023.

[176] Mattia Conte, Ehsan Irani, Andrea M. Chiariello, Alex Abraham, Simona Bianco, Andrea Esposito, and Mario Nicodemi. Loop-extrusion and polymer phase-separation can co-exist at the single-molecule level to shape chromatin folding. *Nature Communications*, 13(1):4070, July 2022.

[177] Viraat Y. Goel, Miles K. Huseyin, and Anders S. Hansen. Region Capture Micro-C reveals coalescence of enhancers and promoters into nested microcompartments. *Nature Genetics*, 55(6):1048–1056, June 2023.

[178] Ryanggeun Lee, Moo-Koo Kang, Yong-Jin Kim, Bobae Yang, Hwanyong Shim, Sugyung Kim, Kyungwoo Kim, Chul Min Yang, Byeong-gyu Min, Woong-Jae Jung, Eun-Chong Lee, Jung-Sik Joo, Gunhee Park, Won-Ki Cho, and Hyoung-Pyo Kim. CTCF-mediated chromatin looping provides a topological framework for the formation of phase-separated transcriptional condensates. *Nucleic Acids Research*, 50(1):207–226, January 2022.

[179] Shreeta Chakraborty, Nina Kopitchinski, Zhenyu Zuo, Ariel Eraso, Parirokh Awasthi, Raj Chari, Apratim Mitra, Ian C. Tobias, Sakthi D. Moorthy, Ryan K. Dale, Jennifer A. Mitchell, Timothy J. Petros, and Pedro P. Rocha. Enhancer–promoter interactions can bypass CTCF-mediated boundaries and contribute to phenotypic robustness. *Nature Genetics*, 55(2):280–290, February 2023.

[180] Niels J. Rinzema, Konstantinos Sofiadis, Sjoerd J. D. Tjalsma, Marjon J. A. M. Verstegen, Yuva Oz, Christian Valdes-Quezada, Anna-Karina Felder, Teodora Filipovska, Stefan Van Der Elst, Zaria De Andrade Dos Ramos, Ruiqi Han, Peter H. L. Krijger, and Wouter De Laat. Building regulatory landscapes reveals that an enhancer can recruit cohesin to create contact domains, engage CTCF sites and activate distant genes. *Nature Structural & Molecular Biology*, 29(6):563–574, June 2022.

[181] Julia A. Horsfield. Full circle: A brief history of cohesin and the regulation of gene expression. *The FEBS Journal*, page febs.16362, January 2022.

[182] Lea Kiefer, Anna Chiosso, Jennifer Langen, Alex Buckley, Simon Gaudin, Sandy M. Rajkumar, Gabrielle Isabelle F. Servito, Elizabeth S. Cha, Akshara Vijay, Albert Yeung, Adan Horta, Michael H. Mui, and Daniele Canzio. WAPL functions as a rheostat of Protocadherin isoform diversity that controls neural wiring. *Science*, 380(6651):eadf8440, June 2023.

[183] Julia A. Horsfield, Sasha H. Anagnostou, Jimmy Kuang-Hsien Hu, Kitty Hsiao Yu Cho, Robert Geisler, Graham Lieschke, Kathryn E. Crosier, and Philip S. Crosier. Cohesin-dependent regulation of Runx genes. *Development*, 134(14):2639–2649, July 2007.

[184] Denes Hnisz, Brian J. Abraham, Tong Ihn Lee, Ashley Lau, Violaine Saint-André, Alla A. Sigova, Heather A. Hoke, and Richard A. Young. Super-Enhancers in the Control of Cell Identity and Disease. *Cell*, 155(4):934–947, November 2013.

[185] Chengqi Lin, Alexander S. Garruss, Zhuojuan Luo, Fengli Guo, and Ali Shilatifard.

The RNA Pol II Elongation Factor Ell3 Marks Enhancers in ES Cells and Primes Future Gene Activation. *Cell*, 152(1-2):144–156, January 2013.

[186] Rosa J. Stolper, Felice H. Tsang, Emily Georgiades, Lars L.P. Hansen, Damien J. Downes, Caroline L. Harrold, Jim R. Hughes, Robert A. Beagrie, Benjamin Davies, Mira T. Kassouf, and Douglas R. Higgs. Loop extrusion by cohesin plays a key role in enhancer-activated gene expression during differentiation. Preprint, Molecular Biology, September 2023.

[187] Tyrone Ryba, Ichiro Hiratani, Junjie Lu, Mari Itoh, Michael Kulik, Jinfeng Zhang, Thomas C. Schulz, Allan J. Robins, Stephen Dalton, and David M. Gilbert. Evolutionarily conserved replication timing profiles predict long-range chromatin interactions and distinguish closely related cell types. *Genome Research*, 20(6):761–770, June 2010.

[188] Benjamin D. Pope, Tyrone Ryba, Vishnu Dileep, Feng Yue, Weisheng Wu, Olgert Denas, Daniel L. Vera, Yanli Wang, R. Scott Hansen, Theresa K. Canfield, Robert E. Thurman, Yong Cheng, Günhan Gülsoy, Jonathan H. Dennis, Michael P. Snyder, John A. Stamatoyannopoulos, James Taylor, Ross C. Hardison, Tamer Kahveci, Bing Ren, and David M. Gilbert. Topologically associating domains are stable units of replication-timing regulation. *Nature*, 515(7527):402–405, November 2014.

[189] Yongzheng Li, Boxin Xue, Mengling Zhang, Liwei Zhang, Yingping Hou, Yizhi Qin, Haizhen Long, Qian Peter Su, Yao Wang, Xiaodong Guan, Yanyan Jin, Yuan Cao, Guohong Li, and Yujie Sun. Transcription-coupled structural dynamics of topologically associating domains regulate replication origin efficiency. *Genome Biology*, 22(1):206, December 2021.

[190] Daniel J. Emerson, Peiyao A. Zhao, Ashley L. Cook, R. Jordan Barnett, Kyle N. Klein, Dalila Saulebekova, Chunmin Ge, Linda Zhou, Zoltan Simandi, Miriam K. Minsk, Katelyn R. Titus, Weitao Wang, Wanfeng Gong, Di Zhang, Liyan Yang, Sergey V. Venev, Johan H. Gibcus, Hongbo Yang, Takayo Sasaki, Masato T. Kanemaki, Feng Yue, Job Dekker, Chun-Long Chen, David M. Gilbert, and Jennifer E. Phillips-Cremins. Cohesin-mediated loop anchors confine the locations of human replication origins. *Nature*, June 2022.

[191] Bart J. H. Dequeker, Matthias J. Scherr, Hugo B. Brandão, Johanna Gassler, Sean Powell, Imre Gaspar, Ilya M. Flyamer, Aleksandar Lalic, Wen Tang, Roman Stocsits, Iain F. Davidson, Jan-Michael Peters, Karl E. Duderstadt, Leonid A. Mirny, and Kikuë Tachibana. MCM complexes are barriers that restrict cohesin-mediated loop extrusion. *Nature*, 606(7912):197–203, June 2022.

[192] Kristian Jeppsson, Toyonori Sakata, Ryuichiro Nakato, Stefina Milanova, Katsuhiko Shirahige, and Camilla Björkegren. Cohesin-dependent chromosome loop extrusion is limited by transcription and stalled replication forks. *Science Advances*, 8(23), June 2022.

[193] Camilla Sjögren and Kim Nasmyth. Sister chromatid cohesion is required for postreplicative double-strand break repair in Saccharomyces cerevisiae. *Current Biology*, 11(12):991–995, June 2001.

[194] Giacomo De Piccoli, Felipe Cortes-Ledesma, Gregory Ira, Jordi Torres-Rosell, Stefan Uhle, Sarah Farmer, Ji-Young Hwang, Felix Machin, Audrey Ceschia, Alexandra McAleenan, Violeta Cordon-Preciado, Andrés Clemente-Blanco, Felip Vilella-Mitjana, Pranav Ullal, Adam Jarmuz, Beatriz Leitao, Debra Bressan, Farokh Dotiwala, Alma Papusha, Xiaolan Zhao, Kyungjae Myung, James E. Haber, Andrés Aguilera, and Luis Aragón. Smc5–Smc6 mediate DNA double-strand-break repair by promoting sister-chromatid recombination. *Nature Cell Biology*, 8(9):1032–1034, September 2006.

[195] Patrick Ryan Potts, Matthew H Porteus, and Hongtao Yu. Human SMC5/6 complex promotes sister chromatid homologous recombination by recruiting the SMC1/3 cohesin complex to double-strand breaks. *The EMBO Journal*, 25(14):3377–3388, July 2006.

[196] Coline Arnould, Vincent Rocher, Anne Laure Finoux, Thomas Clouaire, Kevin Li, Felix Zhou, Pierre Caron, Philippe E. Mangeot, Emiliano P. Ricci, Raphaël Mourad, James E. Haber, Daan Noordermeer, and Gaëlle Legube. Loop extrusion as a mechanism for formation of DNA damage repair foci. *Nature*, 590(7847):660–665, February 2021.

[197] Lena Ström, Charlotte Karlsson, Hanna Betts Lindroos, Sara Wedahl, Yuki Katou, Katsuhiko Shirahige, and Camilla Sjögren. Postreplicative formation of cohesion is required for repair and induced by a single DNA break. *Science*, 317(5835):242–245, July 2007.

[198] Elçin Ünal, Jill M. Heidinger-Pauli, and Douglas Koshland. DNA double-strand breaks trigger genome-wide sister-chromatid cohesion through Eco1 (Ctf7). *Science*, 317(5835):245–248, July 2007.

[199] Aaron Alt, Hung Q. Dang, Owen S. Wells, Luis M. Polo, Matt A. Smith, Grant A. McGregor, Thomas Welte, Alan R. Lehmann, Laurence H. Pearl, Johanne M. Murray, and Antony W. Oliver. Specialized interfaces of Smc5/6 control hinge stability and DNA association. *Nature Communications*, 8(1):14011, April 2017.

[200] Zhubing Shi, Haishan Gao, Xiao Chen Bai, and Hongtao Yu. Cryo-EM structure of the human cohesin-NIPBL-DNA complex. *Science*, 368(6498):1454–1459, June 2020.

[201] Jin H. Yang, Hugo B. Brandão, and Anders S. Hansen. DNA double-strand break end synapsis by DNA loop extrusion. *Nature Communications*, 14(1):1913, April 2023.

[202] Robert Shroff, Ayelet Arbel-Eden, Duane Pilch, Grzegorz Ira, William M. Bonner, John H. Petrini, James E. Haber, and Michael Lichten. Distribution and Dynamics of Chromatin Modification Induced by a Defined DNA Double-Strand Break. *Current Biology*, 14(19):1703–1711, October 2004.

[203] Aurèle Piazza, Hélène Bordelet, Agnès Dumont, Agnès Thierry, Jérôme Savocco, Fabien Girard, and Romain Koszul. Cohesin regulates homology search during recombinational DNA repair. *Nature Cell Biology*, 23(11):1176–1186, November 2021.

[204] Biswajit Pradhan, Takaharu Kanno, Miki Umeda Igarashi, Mun Siong Loke, Martin Dieter Baaske, Jan Siu Kei Wong, Kristian Jeppsson, Camilla Björkegren, and Eugene Kim. The Smc5/6 complex is a DNA loop-extruding motor. *Nature*, 616(7958):843–848, April 2023.

[205] Jan-Michael Peters. How DNA loop extrusion mediated by cohesin enables V(D)J recombination. *Current Opinion in Cell Biology*, 70:75–83, June 2021.

[206] Keiji Kimura, Valentin V. Rybenkov, Nancy J. Crisona, Tatsuya Hirano, and Nicholas R. Cozzarelli. 13S condensin actively reconfigures DNA by introducing global positive writhe: Implications for chromosome condensation. *Cell*, 98(2):239–248, 1999.

[207] Terence R. Strick, Tatsuhiko Kawaguchi, and Tatsuya Hirano. Real-time detection of single-molecule DNA compaction by condensin I. *Current Biology*, 14(10):874–880, May 2004.

[208] Jorine M Eeftens, Shveta Bisht, Jacob Kerssemakers, Marc Kschonsak, Christian H Haering, and Cees Dekker. Real-time detection of condensin-driven DNA compaction reveals a multistep binding mechanism. *The EMBO Journal*, 36(23):3448–3457, December 2017.

[209] Tsuyoshi Terakawa, Shveta Bisht, Jorine M. Eeftens, Cees Dekker, Christian H. Haering, and Eric C. Greene. The condensin complex is a mechanochemical motor that translocates along DNA. *Science*, 358(6363):672–676, November 2017.

[210] Marc Kschonsak, Fabian Merkel, Shveta Bisht, Jutta Metz, Vladimir Rybin, Markus Hassler, and Christian H. Haering. Structural Basis for a Safety-Belt Mechanism That Anchors Condensin to Chromosomes. *Cell*, 171(3):588–600.e24, October 2017.

[211] Iain F Davidson, Daniela Goetz, Maciej P Zaczek, Maxim I Molodtsov, Pim J Huis in 't Veld, Florian Weissmann, Gabriele Litos, David A Cisneros, Maria Ocampo-Hafalla, Rene Ladurner, Frank Uhlmann, Alipasha Vaziri, and Jan-Michael Peters. Rapid movement and transcriptional relocalization of human cohesin on DNA. *The EMBO Journal*, 35(24):2671–2685, 2016.

[212] Yoori Kim, Zhubing Shi, Hongshan Zhang, Ilya J. Finkelstein, and Hongtao Yu. Human cohesin compacts DNA by loop extrusion. *Science*, 366(6471):1345–1349, 2019.

[213] Pilar Gutierrez-Escribano, Silvia Hormeño, Julene Madariaga-Marcos, Roger Solé-Soler, Francis J. O'Reilly, Kyle Morris, Clara Aicart-Ramos, Ricardo Aramayo, Alex Montoya, Holger Kramer, Juri Rappsilber, Jordi Torres-Rosell, Fernando Moreno-Herrero, and Luis Aragon. Purified Smc5/6 Complex Exhibits DNA Substrate Recognition and Compaction. *Molecular Cell*, 80(6):1039–1054.e6, December 2020.

[214] Stefan Golfier, Thomas Quail, Hiroshi Kimura, and Jan Brugués. Cohesin and condensin extrude DNA loops in a cell-cycle dependent manner. *eLife*, 9:1–34, 2020.

[215] Torahiko L. Higashi, Georgii Pobegalov, Minzhe Tang, Maxim I. Molodtsov, and Frank Uhlmann. A brownian ratchet model for dna loop extrusion by the cohesin complex. *eLife*, 10:2021.02.14.431132, February 2021.

[216] Muwen Kong, Erin E. Cutts, Dongqing Pan, Fabienne Beuron, Thangavelu Kaliyappan, Chaoyou Xue, Edward P. Morris, Andrea Musacchio, Alessandro Vannini, and Eric C. Greene. Human Condensin I and II Drive Extensive ATP-Dependent Compaction of Nucleosome-Bound DNA. *Molecular Cell*, 79(1):99–114.e9, 2020.

[217] Je-Kyung Ryu, Sang-Hyun Rah, Richard Janissen, Jacob W J Kerssemakers, Andrea Bonato, Davide Michieletto, and Cees Dekker. Condensin extrudes DNA loops in steps up to hundreds of base pairs that are generated by ATP binding events. *Nucleic Acids Research*, 50(2):820–832, January 2022.

[218] Ngat T. Tran, Michael T. Laub, and Tung B.K. Le. SMC Progressively Aligns Chromosomal Arms in Caulobacter crescentus but Is Antagonized by Convergent Transcription. *Cell Reports*, 20(9):2057–2071, 2017.

[219] Edward J. Banigan and Leonid A. Mirny. Loop extrusion: Theory meets single-molecule experiments. *Current Opinion in Cell Biology*, 64:124–138, June 2020.

[220] Yusuke Kato, Takuya Miyakawa, and Masaru Tanokura. Overview of the mechanism of cytoskeletal motors based on structure. pages 571–581, 2018.

[221] Indra A. Shaltiel, Sumanjit Datta, Léa Lecomte, Markus Hassler, Marc Kschonsak, Sol Bravo, Catherine Stober, Jenny Ormanns, Sebastian Eustermann, and Christian H. Haering. A hold-and-feed mechanism drives directional DNA loop extrusion by condensin. *Science*, 376(6597):1087–1094, June 2022.

[222] Roman Barth, Iain F. Davidson, Jaco Van Der Torre, Michael Taschner, Stephan Gruber, Jan-Michael Peters, and Cees Dekker. SMC motor proteins extrude DNA asymmetrically and contain a direction switch. Preprint, Molecular Biology, December 2023.

[223] Yasuto Murayama and Frank Uhlmann. Biochemical reconstitution of topological DNA binding by the cohesin ring. *Nature*, 505(7483):367–371, 2014.

[224] James E Collier and Kim A Nasmyth. DNA passes through cohesin's hinge as well as its Smc3–kleisin interface. *eLife*, 11:e80310, September 2022.

[225] Kim Nasmyth. Cohesin: A catenase with separate entry and exit gates? *Nature Cell Biology*, 13(10):1170–1177, October 2011.

[226] Biswajit Pradhan, Roman Barth, Eugene Kim, Iain F. Davidson, Benedikt Bauer, Theo van Laar, Wayne Yang, Je-Kyung Ryu, Jaco van der Torre, Jan-Michael Peters, and Cees Dekker. SMC complexes can traverse physical roadblocks bigger than their ring size. *Cell Reports*, 41(3):111491, October 2022.

[227] Roel Oldenkamp and Benjamin D. Rowland. A walk through the SMC cycle: From catching DNAs to shaping the genome. *Molecular Cell*, 82(9):1616–1630, May 2022.

[228] Frank Bürmann and Jan Löwe. Structural biology of SMC complexes across the tree of life. *Current Opinion in Structural Biology*, 80:102598, June 2023.

[229] Frank Bürmann, Byung Gil Lee, Thane Than, Ludwig Sinn, Francis J. O'Reilly, Stanislau Yatskevich, Juri Rappsilber, Bin Hu, Kim Nasmyth, and Jan Löwe. A folded conformation of MukBEF and cohesin. *Nature Structural and Molecular Biology*, 26(3):227–236, 2019.

[230] Sumanjit Datta, Léa Lecomte, and Christian H Haering. Structural insights into DNA loop extrusion by SMC protein complexes. *Current Opinion in Structural Biology*, 65:102–109, December 2020.

[231] Shige H. Yoshimura, Kohji Hizume, Akiko Murakami, Takashi Sutani, Kunio Takeyasu, and Mitsuhiro Yanagida. Condensin Architecture and Interaction with DNA. *Current Biology*, 12(6):508–513, March 2002.

[232] Stephen T Hallett, Isabella Campbell Harry, Pascale Schellenberger, Lihong Zhou, Nora B Cronin, Jonathan Baxter, Thomas J Etheridge, Johanne M Murray, and Antony W Oliver. Cryo-EM structure of the Smc5/6 holo-complex. *Nucleic Acids Research*, 50(16):9505–9520, September 2022.

[233] Marie Laure Diebold-Durand, Hansol Lee, Laura B. Ruiz Avila, Haemin Noh, Ho Chul Shin, Haeri Im, Florian P. Bock, Frank Bürmann, Alexandre Durand, Alrun Basfeld, Sihyun Ham, Jérôme Basquin, Byung Ha Oh, and Stephan Gruber. Structure of Full-Length SMC and Rearrangements Required for Chromosome Organization. *Molecular Cell*, 67(2):334–347.e5, 2017.

[234] James E. Collier, Byung Gil Lee, Maurici Brunet Roig, Stanislav Yatskevich, Naomi J. Petela, Jean Metson, Menelaos Voulgaris, Andres Gonzalez Llamazares, Jan Löwe, and Kim A. Nasmyth. Transport of DNA within cohesin involves clamping on top of engaged heads by SCC2 and entrapment within the ring by SCC3. *eLife*, 9:1–36, 2020.

[235] Torahiko L. Higashi, Patrik Eickhoff, Joana S. Sousa, Julia Locke, Andrea Nans, Helen R. Flynn, Ambrosius P. Snijders, George Papageorgiou, Nicola O'Reilly, Zhuo A. Chen, Francis J. O'Reilly, Juri Rappsilber, Alessandro Costa, and Frank Uhlmann. A Structure-Based Mechanism for DNA Entry into the Cohesin Ring. *Molecular Cell*, 79(6):917–933.e9, 2020.

[236] Byung Gil Lee, James Rhodes, and Jan Löwe. Clamping of DNA shuts the condensin neck gate. *Proceedings of the National Academy of Sciences of the United States of America*, 119(14):1–10, 2022.

[237] Je-kyung Ryu, Allard J Katan, Eli O Van Der Sluis, Thomas Wisse, Ralph De Groot, Christian H Haering, and Cees Dekker. The condensin holocomplex cycles dynamically between open and collapsed states. *Nature Structural &*

Molecular Biology, 27(December):1134–1141, 2020.

[238] Benedikt W. Bauer, Iain F. Davidson, Daniel Canena, Gordana Wutz, Wen Tang, Gabriele Litos, Sabrina Horn, Peter Hinterdorfer, and Jan Michael Peters. Cohesin mediates DNA loop extrusion by a "swing and clamp" mechanism. *Cell*, 184(21):5448–5464, 2021.

[239] Dana Krepel, Aram Davtyan, Nicholas P. Schafer, Peter G. Wolynes, and José N. Onuchic. Braiding topology and the energy landscape of chromosome organization proteins. *Proceedings of the National Academy of Sciences*, 117(3):1468–1477, January 2020.

[240] Andrea Bonato, Jae-Won Jang, Kyoung-Wook Moon, Davide Michieletto, and Je-Kyung Ryu. Spontaneous Unidirectional Loop Extrusion Emerges from Symmetry Breaking of SMC Extension. 2023.

[241] John F. Marko, Paolo De Los Rios, Alessandro Barducci, and Stephan Gruber. DNA-segment-capture model for loop extrusion by structural maintenance of chromosome (SMC) protein complexes. *Nucleic Acids Research*, 47(13):6956–6972, 2019.

[242] Stefanos K Nomidis, Enrico Carlon, Stephan Gruber, and John F Marko. DNA tension-modulated translocation and loop extrusion by SMC complexes revealed by molecular dynamics simulations. *Nucleic Acids Research*, 50(9):4974–4987, 2022.

[243] Byung Gil Lee, Fabian Merkel, Matteo Allegretti, Markus Hassler, Christopher Cawood, Léa Lecomte, Francis J. O'Reilly, Ludwig R. Sinn, Pilar Gutierrez-Escribano, Marc Kschonsak, Sol Bravo, Takanori Nakane, Juri Rappsilber, Luis Aragon, Martin Beck, Jan Löwe, and Christian H. Haering. Cryo-EM structures of holo condensin reveal a subunit flip-flop mechanism. *Nature Structural and Molecular Biology*, 27(8):743–751, 2020.

[244] Roman Barth, Biswajit Pradhan, Eugene Kim, Iain F. Davidson, Jaco Van Der Torre, Jan-Michael Peters, and Cees Dekker. Testing pseudotopological and nontopological models for SMC-driven DNA loop extrusion against roadblock-traversal experiments. *Scientific Reports*, 13(1):8100, May 2023.

[245] Michiko Hirano and Tatsuya Hirano. Positive and negative regulation of SMC–DNA interactions by ATP and accessory proteins. *The EMBO Journal*, 23(13):2664–2673, July 2004.

[246] Georgii Pobegalov, Lee-Ya Chu, Jan-Michael Peters, and Maxim I. Molodtsov. Single cohesin molecules generate force by two distinct mechanisms. *Nature Communications*, 14(1):3946, July 2023.

[247] Cees Dekker, Christian H. Haering, Jan-Michael Peters, and Benjamin D. Rowland. How do molecular motors fold the genome? *Science*, 382(6671):646–648, November 2023.

[248] Eric D. Rubio, David J. Reiss, Piri L. Welcsh, Christine M. Disteche, Galina N. Filippova, Nitin S. Baliga, Ruedi Aebersold, Jeffrey A. Ranish, and Anton Krumm. CTCF physically links cohesin to chromatin. *Proceedings of the National Academy of Sciences of the United States of America*, 105(24):8309–8314, June 2008.

[249] Zhonghui Tang, Oscar Junhong Luo, Xingwang Li, Meizhen Zheng, Jacqueline Jufen Zhu, Przemyslaw Szalaj, Pawel Trzaskoma, Adriana Magalska, Jakub Wlodarczyk, Blazej Ruszczycki, Paul Michalski, Emaly Piecuch, Ping Wang, Danjuan Wang, Simon Zhongyuan Tian, May Penrad-Mobayed, Laurent M. Sachs, Xiaoan Ruan, Chia Lin Wei, Edison T. Liu, Grzegorz M. Wilczynski, Dariusz Plewczynski, Guoliang Li, and Yijun Ruan. CTCF-Mediated Human 3D Genome Architecture Reveals Chromatin Topology for Transcription. *Cell*, 163(7):1611–1627, 2015.

[250] Alberto García-Nieto, Amrita Patel, Yan Li, Roel Oldenkamp, Leonardo Feletto, Joshua J. Graham, Laureen Willems, Kyle W. Muir, Daniel Panne, and Benjamin D. Rowland. Structural basis of centromeric cohesion protection. *Nature Structural & Molecular Biology*, 30(6):853–859, June 2023.

[251] Elena M. Pugacheva, Naoki Kubo, Dmitri Loukinov, Md Tajmul, Sungyun Kang, Alexander L. Kovalchuk, Alexander V. Strunnikov, Gabriel E. Zentner, Bing Ren, and Victor V. Lobanenkov. CTCF mediates chromatin looping via N-terminal domain-dependent cohesin retention. *Proceedings of the National Academy of Sciences of the United States of America*, 117(4):2020–2031, January 2020.

[252] Elphège P. Nora, Laura Caccianini, Geoffrey Fudenberg, Kevin So, Vasumathi Kameswaran, Abigail Nagle, Alec Uebersohn, Bassam Hajj, Agnès Le Saux, Antoine Coulon, Leonid A. Mirny, Katherine S. Pollard, Maxime Dahan, and Benoit G. Bruneau. Molecular basis of CTCF binding polarity in genome folding. *Nature Communications*, 11(1):5612, December 2020.

[253] Jennifer M. Luppino, Daniel S. Park, Son C. Nguyen, Yemin Lan, Zhuxuan Xu, Rebecca Yunker, and Eric F. Joyce. Cohesin promotes stochastic domain intermingling to ensure proper regulation of boundary-proximal genes. *Nature Genetics*, 52(8):840–848, 2020.

[254] Xheni Karaboja, Zhongqing Ren, Hugo B. Brandão, Payel Paul, David Z. Rudner, and Xindan Wang. XerD unloads bacterial SMC complexes at the replication terminus. *Molecular Cell*, 81(4):756–766.e8, February 2021.

[255] Michael H. Kagey, Jamie J. Newman, Steve Bilodeau, Ye Zhan, David A. Orlando, Nynke L. Van Berkum, Christopher C. Ebmeier, Jesse Goossens, Peter B. Rahl, Stuart S. Levine, Dylan J. Taatjes, Job Dekker, and Richard A. Young. Mediator and cohesin connect gene expression and chromatin architecture. *Nature*, 467(7314):430–435, 2010.

[256] Shyam Ramasamy, Abrar Aljahani, Magdalena A. Karpinska, T. B. Ngoc Cao, Taras Velychko, J. Neos Cruz, Michael Lidschreiber, and A. Marieke Oudelaar. The Mediator complex regulates enhancer-promoter interactions. *Nature Structural & Molecular Biology*, 30(7):991–1000, July 2023.

[257] Mark Mattingly, Chris Seidel, Sofía Muñoz, Yan Hao, Ying Zhang, Zhihui Wen, Laurence Florens, Frank Uhlmann, and Jennifer L. Gerton. Mediator recruits the cohesin loader Scc2 to RNA Pol II-transcribed genes and promotes sister chromatid cohesion. *Current Biology*, June 2022.

[258] Jonathan A. Beagan, Michael T. Duong, Katelyn R. Titus, Linda Zhou, Zhendong Cao, Jingjing Ma, Caroline V. Lachanski, Daniel R. Gillis, and Jennifer E. Phillips-Cremins. YY1 and CTCF orchestrate a 3D chromatin looping switch during early neural lineage commitment. *Genome Research*, 27(7):1139–1152, July 2017.

[259] Xuan Pan, Madhusudhan Papasani, Yi Hao, Marco Calamito, Fang Wei, William J. Quinn, Arindam Basu, Junwen Wang, Suchita Hodawadekar, Kristina Zaprazna, Huifei Liu, Yang Shi, David Allman, Michael Cancro, and Michael L. Atchison. YY1 controls Igκ repertoire and B-cell development, and localizes with condensin on the Igκ locus. *EMBO Journal*, 32(8):1168–1182, 2013.

[260] Sven Heinz, Lorane Texari, Michael G.B. Hayes, Matthew Urbanowski, Max W. Chang, Ninvita Givarkes, Alexander Rialdi, Kris M. White, Randy A. Albrecht, Lars Pache, Ivan Marazzi, Adolfo García-Sastre, Megan L. Shaw, and Christopher Benner. Transcription Elongation Can Affect Genome 3D Structure. *Cell*, 174(6):1522–1536.e22, 2018.

[261] Kiwon Lee, Chris C.S. Hsiung, Peng Huang, Arjun Raj, and Gerd A. Blobel. Dynamic enhancer–gene body contacts during transcription elongation. *Genes and Development*, 29(19):1992–1997, 2015.

[262] Cheri A. Schaaf, Hojoong Kwak, Amanda Koenig, Ziva Misulovin, David W. Gohara, Audrey Watson, Yanjiao Zhou, John T. Lis, and Dale Dorsett. Genome-Wide Control of RNA Polymerase II Activity by Cohesin. *PLoS Genetics*, 9(3), 2013.

[263] Hai Pan, Miao Jin, Ashwin Ghadiyaram, Parminder Kaur, Henry E. Miller, Hai Minh Ta, Ming Liu, Yanlin Fan, Chelsea Mahn, Aparna Gorthi, Changjiang You, Jacob Piehler, Robert Riehn, Alexander J.R. Bishop, Yizhi Jane Tao, and Hong Wang. Cohesin SA1 and SA2 are RNA binding proteins that localize to RNA containing regions on DNA. *Nucleic Acids Research*, 48(10):5639–5655, June 2021.

[264] Xiangduo Kong, Yen-Yun Chen, Jianhuang Lin, Ebony Flowers, Eric Van Nostrand, Steven M. Blue, Jonathan Chau, Christopher I-Hsing Ma, Isaiah Mohr, Ryan Thai, Chengguo Yao, Alexander R. Ball, Richard Chien, Shimako Kawauchi, Rosaysela Santos, Anne L. Calof, Arthur D. Lander, Yongsheng Shi, Mitsuru Okuwaki, Gene W. Yeo, and Kyoko Yokomori. The cohesin loader NIPBL interacts with pre-ribosomal RNA and treacle to regulate ribosomal RNA synthesis. Preprint, Molecular Biology, June 2019.

[265] Anna Anchimiuk, Virginia S. Lioy, Florian Patrick Bock, Anita Minnen, Frederic Boccard, and Stephan Gruber. A low smc flux avoids collisions and facilitates chromosome organization in bacillus subtilis. *eLife*, 10, August 2021.

[266] Hugo B. Brandão, Zhongqing Ren, Xheni Karaboja, Leonid A. Mirny, and Xindan Wang. DNA-loop-extruding SMC complexes can traverse one another in vivo. *Nature Structural and Molecular Biology*, 28(8):642–651, August 2021.

[267] Eugene Kim, Jacob Kerssemakers, Indra A. Shaltiel, Christian H. Haering, and Cees Dekker. DNA-loop extruding condensin complexes can traverse one another. *Nature*, 579(7799):438–442, March 2020.

[268] Edward J Banigan, Aafke A van den Berg, Hugo B Brandão, John F Marko, and Leonid A Mirny. Chromosome organization by one-sided and two-sided loop extrusion. *eLife*, 9:e53558, April 2020.

[269] Sofía Muñoz, Masashi Minamino, Corella S. Casas-Delucchi, Harshil Patel, and Frank Uhlmann. A Role for Chromatin Remodeling in Cohesin Loading onto Chromosomes. *Molecular Cell*, 74(4):664–673.e5, May 2019.

[270] Jason Otterstrom, Alvaro Castells-Garcia, Chiara Vicario, Pablo A. Gomez-Garcia, Maria Pia Cosma, and Melike Lakadamyali. Super-resolution microscopy reveals how histone tail acetylation affects DNA compaction within nucleosomes in vivo. *Nucleic Acids Research*, 47(16):8470–8484, July 2019.

[271] Thomas M. Guérin, Claire Béneut, Natalja Barinova, Virginia López, Luciana Lazar-Stefanita, A. Deshayes, Agnès Thierry, Romain Koszul, Karine Dubrana, and Stéphane Marcand. Condensin-Mediated Chromosome Folding and Internal Telomeres Drive Dicentric Severing by Cytokinesis. *Molecular Cell*, 75(1):131–144.e3, July 2019.

[272] Brian T. Analikwu, Alice Deshayes, Jaco Van Der Torre, Thomas Guérin, Allard J. Katan, Claire Béneut, Roman Barth, Jamie Phipps, Vittore Scolari, Xavier Veaute, Christopher Barrington, Didier Busso, Frank Uhlmann, Karine Dubrana, Stefano Mattarocci, Cees Dekker, and Stéphane Marcand. Telomere protein arrays stall DNA loop extrusion by condensin. Preprint, Molecular Biology, October 2023.

[273] Christopher T. Cummings and M. Jordan Rowley. Implications of Dosage Deficiencies in CTCF and Cohesin on Genome Organization, Gene Expression, and Human Neurodevelopment. *Genes*, 13(4):583, March 2022.

[274] Kyoung-Wook Moon and Je-Kyung Ryu. Current working models of SMC-driven DNA-loop extrusion. *Biochemical Society Transactions*, 51(5):1801–1810, October 2023.

[275] Torahiko L. Higashi and Frank Uhlmann. SMC complexes: Lifting the lid on loop extrusion. *Current Opinion in Cell Biology*, 74:13–22, February 2022.

[276] Laura Vian, Aleksandra Pękowska, Suhas S.P. Rao, Kyong Rim Kieffer-Kwon, Seolkyoung Jung, Laura Baranello, Su Chen Huang, Laila El Khattabi, Marei Dose, Nathanael Pruett, Adrian L. Sanborn, Andres Canela, Yaakov Maman, Anna Oksanen, Wolfgang Resch, Xingwang Li, Byoungkoo Lee, Alexander L. Kovalchuk, Zhonghui Tang, Steevenson Nelson, Michele Di Pierro, Ryan R. Cheng, Ido Machol, Brian Glenn St Hilaire, Neva C. Durand, Muhammad S. Shamim, Elena K. Stamenova, José N. Onuchic, Yijun Ruan, Andre Nussenzweig, David Levens, Erez Lieberman Aiden, and Rafael Casellas. The Energetics and Physiological Impact of Cohesin Extrusion. *Cell*, 173(5):1165–1178.e20, May 2018.

[277] David L. Spector. The dynamics of chromosome organization and gene regulation. *Annual Review of Biochemistry*, 72(1):573–608, June 2003.

[278] Natalia Naumova, Maxim Imakaev, Geoffrey Fudenberg, Ye Zhan, Bryan R. Lajoie, Leonid A. Mirny, and Job Dekker. Organization of the mitotic chromosome. *Science*, 342(6161):948–953, 2013.

[279] Frank Uhlmann. SMC complexes: From DNA to chromosomes. *Nature Reviews Molecular Cell Biology*, 17(7):399–412, July 2016.

[280] Siheng Xiang and Douglas Koshland. Cohesin architecture and clustering in vivo. *eLife*, 10:1–49, February 2021.

[281] Tatsuya Hirano. Condensin-Based Chromosome Organization from Bacteria to Vertebrates. *Cell*, 164(5):847–857, 2016.

[282] Iain F. Davidson and Jan Michael Peters. Genome folding through loop extrusion by SMC complexes. *Nature Reviews Molecular Cell Biology*, 22(7):445–464, March 2021.

[283] Markus Hassler, Indra A. Shaltiel, and Christian H. Haering. Towards a Unified Model of SMC Complex Function. *Current Biology*, 28(21):R1266–R1281, November 2018.

[284] Marjon S. van Ruiten and Benjamin D. Rowland. SMC Complexes: Universal DNA Looping Machines with Distinct Regulators. *Trends in Genetics*, 34(6):477–487, June 2018.

[285] Jingchuan Sun, Yi Shi, Roxana E. Georgescu, Zuanning Yuan, Brian T. Chait, Huilin Li, and Michael E. O'Donnell. The architecture of a eukaryotic replisome. *Nature Structural and Molecular Biology*, 22(12):976–982, December 2015.

[286] James D Watson. *Molecular Biology of the Gene*. Pearson Education India, 2004.

[287] Dmitri Ivanov and Kim Nasmyth. A topological interaction between cohesin rings and a circular minichromosome. *Cell*, 122(6):849–860, September 2005.

[288] Frank Bürmann, Louise F.H. Funke, Jason W. Chin, and Jan Löwe. Cryo-EM structure of MukBEF reveals DNA loop entrapment at chromosomal unloading sites. *Molecular Cell*, 81(23):4891–4906.e8, 2021.

[289] Michael H. Nichols and Victor G. Corces. A tethered-inchworm model of SMC DNA translocation. *Nature Structural and Molecular Biology*, 25(10):906–910, September 2018.

[290] Larissa Wilhelm, Frank Bürmann, Anita Minnen, Ho Chul Shin, Christopher P. Toseland, Byung Ha Oh, and Stephan Gruber. SMC condensin entraps chromosomal DNA by an ATP hydrolysis dependent loading mechanism in Bacillus subtilis. *eLife*, 4(MAY), May 2015.

[291] Madhusudhan Srinivasan, Johanna C. Scheinost, Naomi J. Petela, Thomas G. Gligoris, Maria Wissler, Sugako Ogushi, James E. Collier, Menelaos Voulgaris, Alexander Kurze, Kok Lung Chan, Bin Hu, Vincenzo Costanzo, and Kim A. Nasmyth. The Cohesin Ring Uses Its Hinge to Organize DNA Using Non-topological as well as Topological Mechanisms. *Cell*, 173(6):1508–1519.e18, 2018.

[292] Larry E. Tetone, Larry J. Friedman, Melisa L. Osborne, Harini Ravi, Scotty Kyzer, Sarah K. Stumper, Rachel A. Mooney, Robert Landick, and Jeff Gelles. Dynamics of GreB-RNA polymerase interaction allow a proofreading accessory protein to patrol for transcription complexes needing rescue. *Proceedings of the National Academy of Sciences*, 114(7), February 2017.

[293] Humberto Sánchez, Kaley McCluskey, Theo Van Laar, Edo Van Veen, Filip M. Asscher, Belén Solano, John F. X. Diffley, and Nynke H. Dekker. DNA replication origins retain mobile licensing proteins. *Nature Communications*, 12(1):1908, March 2021.

[294] Isabel J. Kingston, Jasmine S.Y. Yung, and Martin R. Singleton. Biophysical Characterization of the Centromere-specific Nucleosome from Budding Yeast. *Journal of Biological Chemistry*, 286(5):4021–4026, February 2011.

[295] Dhananjaya Nayak, Michael Voss, Tricia Windgassen, Rachel Anne Mooney, and Robert Landick. Cys-Pair Reporters Detect a Constrained Trigger Loop in a Paused RNA Polymerase. *Molecular Cell*, 50(6):882–893, June 2013.

[296] Miloš Tišma, Maria Panoukidou, Hammam Antar, Young-Min Soh, Roman Barth, Biswajit Pradhan, Anders Barth, Jaco Van Der Torre, Davide Michieletto, Stephan Gruber, and Cees Dekker. ParB proteins can bypass DNA-bound roadblocks via dimer-dimer recruitment. *Science Advances*, 8(26):eabn3299, July 2022.

[297] Makio Tokunaga, Naoko Imamoto, and Kumiko Sakata-Sogawa. Highly inclined thin illumination enables clear single-molecule imaging in cells. *Nature Methods*, 5(2):159–161, 2008.

[298] Alexander Krull, Tim Oliver Buchholz, and Florian Jug. Noise2void-Learning denoising from single noisy images. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2019-June:2124–2132, 2019.

[299] Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J. van der Walt, Matthew Brett, Joshua Wilson, K. Jarrod Millman, Nikolay Mayorov, Andrew R.J. Nelson, Eric Jones, Robert Kern, Eric Larson, C. J. Carey, İlhan Polat, Yu Feng, Eric W. Moore, Jake VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E. A. Quintero, Charles R. Harris, Anne M. Archibald, Antônio H. Ribeiro, Fabian Pedregosa, Paul van Mulbregt, Aditya Vijaykumar, Alessandro Pietro Bardelli, Alex Rothberg, Andreas Hilboll, Andreas Kloeckner, Anthony Scopatz, Antony Lee, Ariel Rokem, C. Nathan Woods, Chad Fulton, Charles Masson, Christian Häggström, Clark Fitzgerald, David A. Nicholson, David R. Hagen, Dmitrii V. Pasechnik, Emanuele Olivetti, Eric Martin, Eric Wieser, Fabrice Silva, Felix Lenders, Florian Wilhelm, G. Young, Gavin A. Price, Gert Ludwig Ingold, Gregory E. Allen, Gregory R. Lee, Hervé Audren, Irvin Probst, Jörg P. Dietrich, Jacob Silterra, James T. Webber, Janko Slavič, Joel Nothman, Johannes Buchner, Johannes Kulick,

Johannes L. Schönberger, José Vinícius de Miranda Cardoso, Joscha Reimer, Joseph Harrington, Juan Luis Cano Rodríguez, Juan Nunez-Iglesias, Justin Kuczynski, Kevin Tritz, Martin Thoma, Matthew Newville, Matthias Kümmerer, Maximilian Bolingbroke, Michael Tartre, Mikhail Pak, Nathaniel J. Smith, Nikolai Nowaczyk, Nikolay Shebanov, Oleksandr Pavlyk, Per A. Brodtkorb, Perry Lee, Robert T. McGibbon, Roman Feldbauer, Sam Lewis, Sam Tygier, Scott Sievert, Sebastiano Vigna, Stefan Peterson, Surhud More, Tadeusz Pudlik, Takuya Oshima, Thomas J. Pingel, Thomas P. Robitaille, Thomas Spura, Thouis R. Jones, Tim Cera, Tim Leslie, Tiziano Zito, Tom Krauss, Utkarsh Upadhyay, Yaroslav O. Halchenko, and Yoshiki Vázquez-Baeza. SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nature Methods 2020 17:3*, 17(3):261–272, February 2020.

[300] Ryota Sakata, Kyoma Niwa, Diego Ugarte La Torre, Chenyang Gu, Eri Tahara, Shoji Takada, and Tomoko Nishiyama. Opening of cohesin's SMC ring is essential for timely DNA replication and DNA loop formation. *Cell Reports*, 35(4):108999, April 2021.

[301] Jonas N. Pedersen, Rodolphe Marie, Anders Kristensen, and Henrik Flyvbjerg. How to determine local stretching and tension in a flow-stretched DNA molecule. *Physical Review E*, 93(4):042405, April 2016.

[302] Mark J. Schnitzer, Koen Visscher, and Steven M. Block. Force production by single kinesin motors. *Nature Cell Biology*, 2(10):718–723, October 2000.

[303] Ya Guo, Quan Xu, Daniele Canzio, Jia Shou, Jinhuan Li, David U. Gorkin, Inkyung Jung, Haiyang Wu, Yanan Zhai, Yuanxiao Tang, Yichao Lu, Yonghu Wu, Zhilian Jia, Wei Li, Michael Q. Zhang, Bing Ren, Adrian R. Krainer, Tom Maniatis, and Qiang Wu. CRISPR Inversion of CTCF Sites Alters Genome Topology and Enhancer/Promoter Function. *Cell*, 162(4):900–910, 2015.

[304] Anders S. Hansen, Claudia Cattoglio, Xavier Darzacq, and Robert Tjian. Recent evidence that TADs and chromatin loops are dynamic structures. *Nucleus*, 9(1):20–32, December 2018.

[305] Ireneusz Litwin, Ewa Pilarczyk, and Robert Wysocki. The Emerging Role of Cohesin in the DNA Damage Response. *Genes*, 9(12):581, November 2018.

[306] J. Phipps, A. Nasim, and D.R. Miller. Recovery, Repair, and Mutagenesis in Schizosaccharomyces pombe. In *Advances in Genetics*, volume 23, pages 1–72. Elsevier, 1985.

[307] Luis Aragón. The Smc5/6 Complex: New and Old Functions of the Enigmatic Long-Distance Relative. *Annual Review of Genetics*, 52(1):89–107, November 2018.

[308] W.Hayes McDonald, Yelena Pavlova, John R. Yates, and Michael N. Boddy. Novel Essential DNA Repair Proteins Nse1 and Nse2 Are Subunits of the Fission Yeast Smc5-Smc6 Complex. *Journal of Biological Chemistry*, 278(46):45460–45467, November 2003.

[309] Michael Taschner and Stephan Gruber. DNA segment capture by Smc5/6 holocomplexes. *Nature Structural & Molecular Biology*, 30(5):619–628, May 2023.

[310] James Rhodes, Davide Mazza, Kim Nasmyth, and Stephan Uphoff. Scc2/Nipbl hops between chromosomal cohesin rings after loading. *eLife*, 6:e30000, September 2017.

[311] John Jumper, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger, Kathryn Tunyasuvunakool, Russ Bates, Augustin Žídek, Anna Potapenko, Alex Bridgland, Clemens Meyer, Simon A. A. Kohl, Andrew J. Ballard, Andrew Cowie, Bernardino Romera-Paredes, Stanislav Nikolov, Rishub Jain, Jonas Adler, Trevor Back, Stig Petersen,

David Reiman, Ellen Clancy, Michal Zielinski, Martin Steinegger, Michalina Pacholska, Tamas Berghammer, Sebastian Bodenstein, David Silver, Oriol Vinyals, Andrew W. Senior, Koray Kavukcuoglu, Pushmeet Kohli, and Demis Hassabis. Highly accurate protein structure prediction with AlphaFold. *Nature*, 596(7873):583–589, August 2021.

[312] Claudia Cattoglio, Iryna Pustova, Nike Walther, Jaclyn J Ho, Merle Hantsche-Grininger, Carla J Inouye, M Julius Hossain, Gina M Dailey, Jan Ellenberg, Xavier Darzacq, Robert Tjian, and Anders S Hansen. Determining cellular CTCF and cohesin abundances to constrain 3D genome models. *eLife*, 8, June 2019.

[313] Brandon Ho, Anastasia Baryshnikova, and Grant W. Brown. Unification of Protein Abundance Datasets Yields a Quantitative Saccharomyces cerevisiae Proteome. *Cell Systems*, 6(2):192–205.e3, February 2018.

[314] Johann Holzmann, Antonio Z Politi, Kota Nagasaka, Merle Hantsche-Grininger, Nike Walther, Birgit Koch, Johannes Fuchs, Gerhard Dürnberger, Wen Tang, Rene Ladurner, Roman R Stocsits, Georg A Busslinger, Béla Novák, Karl Mechtler, Iain Finley Davidson, Jan Ellenberg, and Jan-Michael Peters. Absolute quantification of cohesin, CTCF and their regulators in human cells. *eLife*, 8:e46269, June 2019.

[315] Qingyao Huang, Damian Szklarczyk, Mingcong Wang, Milan Simonovic, and Christian Von Mering. PaxDb 5.0: Curated Protein Quantification Data Suggests Adaptive Proteome Changes in Yeasts. *Molecular & Cellular Proteomics*, 22(10):100640, October 2023.

[316] Bin Hu, Takehiko Itoh, Ajay Mishra, Yuki Katoh, Kok-Lung Chan, William Upcher, Camilla Godlee, Maurici B. Roig, Katsuhiko Shirahige, and Kim Nasmyth. ATP Hydrolysis Is Required for Relocating Cohesin from Sites Occupied by Its Scc2/4 Loading Complex. *Current Biology*, 21(1):12–24, January 2011.

[317] Geoffrey Fudenberg, Nezar Abdennur, Maxim Imakaev, Anton Goloborodko, and Leonid A. Mirny. Emerging Evidence of Chromosome Folding by Loop Extrusion. *Cold Spring Harbor Symposia on Quantitative Biology*, 82:45–55, 2017.

[318] Atreya Dey, Guang Shi, Ryota Takaki, and D. Thirumalai. Structural changes in chromosomes driven by multiple condensin motors during mitosis. *Cell Reports*, 42(4):112348, April 2023.

[319] Aleksandra Galitsyna, Sergey V. Ulianov, Nikolai S. Bykov, Marina Veil, Meijiang Gao, Kristina Perevoschikova, Mikhail Gelfand, Sergey V. Razin, Leonid Mirny, and Daria Onichtchouk. Extrusion fountains are hallmarks of chromosome organization emerging upon zygotic genome activation. Preprint, Molecular Biology, July 2023.

[320] Bolaji N. Isiaka, Jennifer I. Semple, Anja Haemmerli, Saurabh Thapliyal, Klement Stojanovski, Moushumi Das, Nick Gilbert, Dominique A. Glauser, Benjamin Towbin, Daniel Jost, and Peter Meister. Cohesin forms fountains at active enhancers in *C. elegans*. Preprint, Genomics, July 2023.

[321] Ning Qing Liu, Mikhail Magnitov, Marijne Schijns, Tom Van Schaik, Robin H. Van Der Weide, Hans Teunissen, Bas Van Steensel, and Elzo De Wit. Rapid depletion of CTCF and cohesin proteins reveals dynamic features of chromosome architecture. Preprint, Molecular Biology, August 2021.

[322] Ya Guo, Ediem Al-Jibury, Rosalba Garcia-Millan, Konstantinos Ntagiantas, James W.D. King, Alex J. Nash, Niels Galjart, Boris Lenhard, Daniel Rueckert, Amanda G. Fisher, Gunnar Pruessner, and Matthias Merkenschlager. Chromatin jets define the properties of cohesin-driven in vivo loop extrusion. *Molecular Cell*, 82(20):3769–3780.e5, October 2022.

[323] Eugene Kim, Alejandro Martin Gonzalez, Biswajit Pradhan, Jaco van der Torre, and Cees Dekker. Condensin-driven loop extrusion on supercoiled DNA. *Nature Structural & Molecular Biology*, July 2022.

[324] Sang Ryul Park, Jesse Hauver, Yunxiang Zhang, Andrey Revyakin, Robert A. Coleman, Robert Tjian, Steven Chu, and Alexandros Pertsinidis. A Single-Molecule Surface-Based Platform to Detect the Assembly and Function of the Human RNA Polymerase II Transcription Machinery. *Structure*, 28(12):1337–1343.e4, December 2020.

[325] Kristian Jeppsson, Biswajit Pradhan, Takashi Sutani, Toyonori Sakata, Miki Umeda Igarashi, Davide Giorgio Berta, Takaharu Kanno, Ryuichiro Nakato, Katsuhiko Shirahige, Eugene Kim, and Camilla Björkegren. Loop-extruding Smc5/6 organizes transcription-induced positive DNA supercoils. *Molecular Cell*, page S1097276524000066, January 2024.

[326] Peter Kapusta. Absolute diffusion coefficients: Compilation of reference data for FCS calibration., 2010.

[327] Milot Mirdita, Konstantin Schütze, Yoshitaka Moriwaki, Lim Heo, Sergey Ovchinnikov, and Martin Steinegger. ColabFold: Making protein folding accessible to all. *Nature Methods*, 19(6):679–682, June 2022.

[328] Thomas D. Goddard, Conrad C. Huang, Elaine C. Meng, Eric F. Pettersen, Gregory S. Couch, John H. Morris, and Thomas E. Ferrin. UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Science*, 27(1):14–25, January 2018.

[329] John G. Kemeny and J. Laurie Snell. *Finite Markov Chains*. Van Nostrand, Princeton, N.J., 1960.

[330] T. W. Anderson and Leo A. Goodman. Statistical Inference about Markov Chains. *The Annals of Mathematical Statistics*, 28(1):89–110, March 1957.

[331] Susan H Harvey, Michael Je Krien, and Matthew J O'Connell. Structural maintenance of chromosomes (SMC) proteins, a family of conserved ATPases. *Genome Biology*, 3(2):reviews3003.1, 2002.

[332] Paul Batty, Christoph Ch Langer, Zsuzsanna Takács, Wen Tang, Claudia Blaukopf, Jan-Michael Peters, and Daniel W Gerlich. Cohesin-mediated DNA loop extrusion resolves sister chromatids in G2 phase. *The EMBO Journal*, 42(16):e113475, August 2023.

[333] Kirsten A. Hagstrom, Victor F. Holmes, Nicholas R. Cozzarelli, and Barbara J. Meyer. *C. Elegans* condensin promotes mitotic chromosome architecture, centromere organization, and sister chromatid segregation during mitosis and meiosis. *Genes & Development*, 16(6):729–742, March 2002.

[334] Kristina Krassovsky, Rajarshi P. Ghosh, and Barbara J. Meyer. Genome-wide profiling reveals functional interplay of DNA sequence composition, transcriptional activity, and nucleosome positioning in driving DNA supercoiling and helix destabilization in *C. elegans*. *Genome Research*, 31(7):1187–1202, July 2021.

[335] Hana Dohnalová, Mona Seifert, Eva Matoušková, Flávia S. Papini, Jan Lipfert, David Dulin, and Filip Lankaš. Temperature-dependent twist of double-stranded RNA probed by magnetic tweezers experiments and molecular dynamics simulations. Preprint, Biophysics, June 2023.

[336] Francesco Mosconi, Jean François Allemand, David Bensimon, and Vincent Croquette. Measurement of the Torque on a Single Stretched and Twisted DNA Using Magnetic Tweezers. *Physical Review Letters*, 102(7):078301, February 2009.

[337] T. R. Strick, J.-F. Allemand, D. Bensimon, A. Bensimon, and V. Croquette. The Elasticity of a Single Supercoiled DNA Molecule. *Science*, 271(5257):1835–1837, March 1996.

[338] Rifka Vlijm, Alireza Mashaghi, Stéphanie Bernard, Mauro Modesti, and Cees Dekker. Experimental phase diagram of negatively supercoiled DNA measured by magnetic tweezers and fluorescence. *Nanoscale*, 7(7):3205–3216, 2015.

[339] Eugen Ostrofet, Flávia Stal Papini, and David Dulin. Correction-free force calibration for magnetic tweezers experiments. *Scientific Reports*, 8(1):15920, October 2018.

[340] Scott Forth, Christopher Deufel, Maxim Y. Sheinin, Bryan Daniels, James P. Sethna, and Michelle D. Wang. Abrupt Buckling Transition Observed during the Plectoneme Formation of Individual DNA Molecules. *Physical Review Letters*, 100(14):148301, April 2008.

[341] Timothée Lionnet, Alexandre Dawid, Sarah Bigot, François-Xavier Barre, Omar A. Saleh, François Heslot, Jean-François Allemand, David Bensimon, and Vincent Croquette. DNA mechanics as a tool to probe helicase and translocase activity. *Nucleic Acids Research*, 34(15):4232–4244, September 2006.

[342] Michael Taschner, Jérôme Basquin, Barbara Steigenberger, Ingmar B Schäfer, Young-Min Soh, Claire Basquin, Esben Lorentzen, Markus Räschle, Richard A Scheltema, and Stephan Gruber. Nse5/6 inhibits the Smc5/6 ATPase and modulates DNA substrate binding. *The EMBO Journal*, 40(15):e107807, August 2021.

[343] Karla Kirkegaard and James C. Wang. Bacterial DNA topoisomerase I can relax positively supercoiled DNA containing a single-stranded loop. *Journal of Molecular Biology*, 185(3):625–637, October 1985.

[344] Kemin Tan, Qingxuan Zhou, Bokun Cheng, Zhongtao Zhang, Andrzej Joachimiak, and Yuk-Ching Tse-Dinh. Structural basis for suppression of hypernegative DNA supercoiling by *E. coli* topoisomerase I. *Nucleic Acids Research*, 43(22):11031–11046, December 2015.

[345] He Meng, Johan Bosman, Thijn van der Heijden, and John van Noort. Coexistence of Twisted, Plectonemic, and Melted DNA in Small Topological Domains. *Biophysical Journal*, 106(5):1174–1181, March 2014.

[346] J. Baxter, N. Sen, V. López Martínez, M. E. Monturus De Carandini, J. B. Schvartzman, J. F.X. Diffley, and L. Aragón. Positive supercoiling of mitotic DNA drives decatenation by topoisomerase II in eukaryotes. *Science*, 331(6022):1328–1332, March 2011.

[347] Janet C. Lindow, Robert A. Britton, and Alan D. Grossman. Structural Maintenance of Chromosomes Protein of *Bacillus subtilis* Affects Supercoiling In Vivo. *Journal of Bacteriology*, 184(19):5317–5322, October 2002.

[348] James A. Sawitzke and Stuart Austin. Suppression of chromosome segregation defects of *Escherichia coli muk* mutants by mutations in topoisomerase I. *Proceedings of the National Academy of Sciences*, 97(4):1671–1676, February 2000.

[349] Michael T. Hons, Pim J. Huis In 'T Veld, Jan Kaesler, Pascaline Rombaut, Alexander Schleiffer, Franz Herzog, Holger Stark, and Jan-Michael Peters. Topology and structure of an engineered human cohesin complex bound to Pds5B. *Nature Communications*, 7(1):12523, August 2016.

[350] Stephen T Hallett, Pascale Schellenberger, Lihong Zhou, Fabienne Beuron, Ed Morris, Johanne M Murray, and Antony W Oliver. Nse5/6 is a negative regulator of the ATPase activity of the Smc5/6 complex. *Nucleic Acids Research*, 49(8):4534–4549, May 2021.

[351] Li Hsin Chang, Sourav Ghosh, and Daan Noordermeer. TADs and Their Borders: Free Movement or Building a Wall? *Journal of Molecular Biology*, 432(3):643–652, December 2020.

[352] Dusan Racko, Fabrizio Benedetti, Julien Dorier, and Andrzej Stasiak. Are TADs

supercoiled? *Nucleic Acids Research*, 47(2):521–532, 2019.

[353] Catherine Naughton, Nicolaos Avlonitis, Samuel Corless, James G Prendergast, Ioulia K Mati, Paul P Eijk, Scott L Cockroft, Mark Bradley, Bauke Ylstra, and Nick Gilbert. Transcription forms and remodels supercoiling domains unfolding large-scale chromatin structures. *Nature Structural & Molecular Biology*, 20(3):387–395, March 2013.

[354] Dusan Racko, Fabrizio Benedetti, Julien Dorier, and Andrzej Stasiak. Transcription-induced supercoiling as the driving force of chromatin loop extrusion during formation of TADs in interphase chromosomes. *Nucleic Acids Research*, 46(4):1648–1660, 2018.

[355] Fabrizio Benedetti, Julien Dorier, Yannis Burnier, and Andrzej Stasiak. Models that include supercoiling of topological domains reproduce several known features of interphase chromosomes. *Nucleic Acids Research*, 42(5):2848–2855, March 2014.

[356] Fabrizio Benedetti, Julien Dorier, and Andrzej Stasiak. Effects of supercoiling on enhancer–promoter contacts. *Nucleic Acids Research*, 42(16):10425–10432, September 2014.

[357] Aurélie Diman, Gaël Panis, Cédric Castrogiovanni, Julien Prados, Bastien Baechler, and Michel Strubin. Human Smc5/6 recognises transcription-generated positive DNA supercoils. Preprint, Molecular Biology, May 2023.

[358] Daniel G Gibson, Lei Young, Ray-Yuan Chuang, J Craig Venter, Clyde A Hutchison, and Hamilton O Smith. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods*, 6(5):343–345, May 2009.

[359] Florian Weissmann, Georg Petzold, Ryan VanderLinden, Pim J. Huis In 'T Veld, Nicholas G. Brown, Fabienne Lampert, Stefan Westermann, Holger Stark, Brenda A. Schulman, and Jan-Michael Peters. biGBac enables rapid gene assembly for the expression of large multisubunit protein complexes. *Proceedings of the National Academy of Sciences*, 113(19), May 2016.

[360] Florian Weissmann and Jan-Michael Peters. Expressing Multi-subunit Complexes Using biGBac. In Joseph A. Marsh, editor, *Protein Complex Assembly*, volume 1764, pages 329–343. Springer New York, New York, NY, 2018.

[361] Richard Janissen, Bojk A. Berghuis, David Dulin, Max Wink, Theo van Laar, and Nynke H. Dekker. Invincible DNA tethers: Covalent DNA anchoring for enhanced temporal and force stability in magnetic tweezers experiments. *Nucleic Acids Research*, 42(18):e137–e137, October 2014.

[362] Jan Lipfert, Jacob W J Kerssemakers, Tessa Jager, and Nynke H Dekker. Magnetic torque tweezers: Measuring torsional stiffness in DNA and RecA-DNA filaments. *Nature Methods*, 7(12):977–980, December 2010.

[363] Richard Janissen, Behrouz Eslami-Mossallam, Irina Artsimovitch, Martin Depken, and Nynke H. Dekker. High-throughput single-molecule experiments reveal heterogeneity, state switching, and three interconnected pause states in transcription. *Cell Reports*, 39(4):110749, April 2022.

[364] J. P. Cnossen, D. Dulin, and N. H. Dekker. An optimized software framework for real-time, high-throughput tracking of spherical beads. *Review of Scientific Instruments*, 85(10):103712, October 2014.

[365] Naomi J. Petela, Andres Gonzalez Llamazares, Sarah Dixon, Bin Hu, Byung Gil Lee, Jean Metson, Heekyo Seo, Antonio Ferrer-Harding, Menelaos Voulgaris, Thomas Gligoris, James Collier, Byung Ha Oh, Jan Löwe, and Kim A. Nasmyth. Folding of cohesin's coiled coil is important for scc2/4-induced association with chromosomes. *eLife*, 10:1–32, 2021.

[366] Wenxuan Xu, David Dunlap, and Laura Finzi. Energetics of twisted DNA topologies. *Biophysical Journal*, 120(16):3242–3252, August 2021.

[367] Ryota Takaki, Atreya Dey, Guang Shi, and D. Thirumalai. Theory and simulations of condensin mediated loop extrusion in DNA. *Nature Communications*, 12(1):5865, October 2021.

[368] Yeonee Seol, Hongliang Zhang, Yves Pommier, and Keir C. Neuman. A kinetic clutch governs religation by type IB topoisomerases and determines camptothecin sensitivity. *Proceedings of the National Academy of Sciences*, 109(40):16125–16130, October 2012.

[369] Ksenia Terekhova, Kathryn H. Gunn, John F. Marko, and Alfonso Mondragón. Bacterial topoisomerase I and topoisomerase III relax supercoiled DNA via distinct pathways. *Nucleic Acids Research*, 40(20):10432–10440, November 2012.

[370] E Goldstein and K Drlica. Regulation of bacterial DNA supercoiling: Plasmid linking numbers vary with growth temperature. *Proceedings of the National Academy of Sciences*, 81(13):4046–4050, July 1984.

[371] D H Lee and R F Schleif. In vivo DNA loops in araCBAD: Size limits and helical repeat. *Proceedings of the National Academy of Sciences*, 86(2):476–480, January 1989.

[372] Daniel Klaue and Ralf Seidel. Torsional Stiffness of Single Superparamagnetic Microspheres in an External Magnetic Field. *Physical Review Letters*, 102(2):028302, January 2009.

[373] Luuk Loeff, Jacob W.J. Kerssemakers, Chirlmin Joo, and Cees Dekker. AutoStepfinder: A fast and automated step detection method for single-molecule analysis. *Patterns*, 2(5):100256, May 2021.

[374] Johannes Buheitel and Olaf Stemmann. Prophase pathway-dependent removal of cohesin from human chromosomes requires opening of the Smc3–Scc1 gate. *The EMBO Journal*, 32(5):666–676, January 2013.

[375] Stephan Gruber, Prakash Arumugam, Yuki Katou, Daria Kuglitsch, Wolfgang Helmhart, Katsuhiko Shirahige, and Kim Nasmyth. Evidence that Loading of Cohesin Onto Chromosomes Involves Opening of Its SMC Hinge. *Cell*, 127(3):523–537, November 2006.

[376] Gamze Ö. Çamdere, Kristian K. Carlborg, and Douglas Koshland. Intermediate step of cohesin's ATPase cycle allows cohesin to entrap DNA. *Proceedings of the National Academy of Sciences*, 115(39):9732–9737, September 2018.

[377] Masashi Minamino, Torahiko L Higashi, Céline Bouchoux, and Frank Uhlmann. Topological in vitro loading of the budding yeast cohesin ring onto DNA. *Life Science Alliance*, 1(5):e201800143, October 2018.

[378] Markus Hassler, Indra A. Shaltiel, Marc Kschonsak, Bernd Simon, Fabian Merkel, Lena Thärichen, Henry J. Bailey, Jakub Macošek, Sol Bravo, Jutta Metz, Janosch Hennig, and Christian H. Haering. Structural Basis of an Asymmetric Condensin ATPase Cycle. *Molecular Cell*, 74(6):1175–1188.e9, 2019.

[379] Zhaoqing Ba, Jiangman Lou, Adam Yongxin Ye, Hai-Qiang Dai, Edward W. Dring, Sherry G. Lin, Suvi Jain, Nia Kyritsis, Kyong-Rim Kieffer-Kwon, Rafael Casellas, and Frederick W. Alt. CTCF orchestrates long-range cohesin-driven V(D)J recombinational scanning. *Nature*, 586(7828):305–310, October 2020.

[380] Louisa Hill, Anja Ebert, Markus Jaritz, Gordana Wutz, Kota Nagasaka, Hiromi Tagoh, Daniela Kostanova-Poliakova, Karina Schindler, Qiong Sun, Peter Bönelt, Maria Fischer, Jan-Michael Peters, and Meinrad Busslinger. Wapl repression by Pax5 promotes V gene recombination by

Igh loop extrusion. *Nature*, 584(7819):142–147, August 2020.

[381] Teresa Olbrich, Maria Vega-Sendino, Desiree Tillo, Wei Wu, Nicholas Zolnerowich, Raphael Pavani, Andy D. Tran, Catherine N. Domingo, Mariajose Franco, Marta Markiewicz-Potoczny, Gianluca Pegoraro, Peter C. FitzGerald, Michael J. Kruhlak, Eros Lazzerini-Denchi, Elphege P. Nora, André Nussenzweig, and Sergio Ruiz. CTCF is a barrier for 2C-like reprogramming. *Nature Communications*, 12(1):4856, August 2021.

[382] Julie Segueni and Daan Noordermeer. CTCF: A misguided jack-of-all-trades in cancer cells. *Computational and Structural Biotechnology Journal*, 20:2685–2698, 2022.

[383] Mayilaadumveettil Nishana, Caryn Ha, Javier Rodriguez-Hernaez, Ali Ranjbaran, Erica Chio, Elphege P. Nora, Sana B. Badri, Andreas Kloetgen, Benoit G. Bruneau, Aristotelis Tsirigos, and Jane A. Skok. Defining the relative and combined contribution of CTCF and CTCFL to genomic regulation. *Genome Biology*, 21(1):108, December 2020.

[384] Michael Mitter, Catherina Gasser, Zsuzsanna Takacs, Christoph C. H. Langer, Wen Tang, Gregor Jessberger, Charlie T. Beales, Eva Neuner, Stefan L. Ameres, Jan-Michael Peters, Anton Goloborodko, Ronald Micura, and Daniel W. Gerlich. Conformation of sister chromatids in the replicated human genome. *Nature*, 586(7827):139–144, October 2020.

[385] Adam C. Bell and Gary Felsenfeld. Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. *Nature*, 405(6785):482–485, May 2000.

[386] Adam C Bell, Adam G West, and Gary Felsenfeld. The Protein CTCF Is Required for the Enhancer Blocking Activity of Vertebrate Insulators. *Cell*, 98(3):387–396, August 1999.

[387] Harsha Agarwal, Matthias Reisser, Celina Wortmann, and J. Christof M. Gebhardt. Direct Observation of Cell-Cycle-Dependent Interactions between CTCF and Chromatin. *Biophysical Journal*, 112(10):2051–2055, May 2017.

[388] Anders S Hansen, Iryna Pustova, Claudia Cattoglio, Robert Tjian, and Xavier Darzacq. CTCF and cohesin regulate chromatin loop stability with distinct dynamics. *eLife*, 6, 2017.

[389] Anders S. Hansen, Assaf Amitai, Claudia Cattoglio, Robert Tjian, and Xavier Darzacq. Guided nuclear exploration increases CTCF target search efficiency. *Nature Chemical Biology*, 16(3):257–266, 2020.

[390] Kyong-Rim Kieffer-Kwon, Keisuke Nimura, Suhas S.P. Rao, Jianliang Xu, Seolkyoung Jung, Aleksandra Pekowska, Marei Dose, Evan Stevens, Ewy Mathe, Peng Dong, Su-Chen Huang, Maria Aurelia Ricci, Laura Baranello, Ying Zheng, Francesco Tomassoni Ardori, Wolfgang Resch, Diana Stavreva, Steevenson Nelson, Michael McAndrew, Adriel Casellas, Elizabeth Finn, Charles Gregory, Brian Glenn St. Hilaire, Steven M. Johnson, Wendy Dubois, Maria Pia Cosma, Eric Batchelor, David Levens, Robert D. Phair, Tom Misteli, Lino Tessarollo, Gordon Hager, Melike Lakadamyali, Zhe Liu, Monique Floer, Hari Shroff, Erez Lieberman Aiden, and Rafael Casellas. Myc Regulates Chromatin Decompaction and Nuclear Architecture during B Cell Activation. *Molecular Cell*, 67(4):566–578.e10, August 2017.

[391] Hirotaka Nakahashi, Kyong-Rim Kieffer Kwon, Wolfgang Resch, Laura Vian, Marei Dose, Diana Stavreva, Ofir Hakim, Nathanael Pruett, Steevenson Nelson, Arito Yamane, Jason Qian, Wendy Dubois, Scott Welsh, Robert D. Phair, B. Franklin Pugh, Victor Lobanenkov, Gordon L. Hager, and Rafael Casellas. A Genome-wide Map of CTCF Multivalency Redefines the CTCF Code. *Cell Reports*, 3(5):1678–1689, May 2013.

[392] Widia Soochit, Frank Sleutels, Gregoire Stik, Marek Bartkuhn, Sreya Basu, Silvia C. Hernandez, Sarra Merzouk, Enrique Vidal, Ruben Boers, Joachim Boers, Michael Van Der Reijden, Bart Geverts, Wiggert A. Van Cappellen, Mirjam Van Den Hout, Zeliha Ozgur, Wilfred F. J. Van IJcken, Joost Gribnau, Rainer Renkawitz, Thomas Graf, Adriaan Houtsmuller, Frank Grosveld, Ralph Stadhouders, and Niels Galjart. CTCF chromatin residence time controls three-dimensional genome organization, gene expression and DNA methylation in pluripotent cells. *Nature Cell Biology*, 23(8):881–893, August 2021.

[393] Hongshan Zhang, Zhubing Shi, Edward J. Banigan, Yoori Kim, Hongtao Yu, Xiaochen Bai, and Ilya J. Finkelstein. CTCF and R-loops are boundaries of cohesin-mediated DNA looping. *Molecular Cell*, 83(16):2856–2871.e8, August 2023.

[394] Selena R. Martinez and Jj L. Miranda. CTCF terminal segments are unstructured. *Protein Science*, 19(5):1110–1116, May 2010.

[395] Gordana Wutz, Rene Ladurner, Brian Glenn St Hilaire, Roman R Stocsits, Kota Nagasaka, Benoit Pignard, Adrian Sanborn, Wen Tang, Csilla Várnai, Miroslav P Ivanov, Stefan Schoenfelder, Petra Van Der Lelij, Xingfan Huang, Gerhard Dürnberger, Elisabeth Roitinger, Karl Mechtler, Iain Finley Davidson, Peter Fraser, Erez Lieberman-Aiden, and Jan-Michael Peters. ESCO1 and CTCF enable formation of long chromatin loops by protecting cohesinSTAG1 from WAPL. *eLife*, 9:e52091, February 2020.

[396] R. N. Plasschaert, S. Vigneau, I. Tempera, R. Gupta, J. Maksimoska, L. Everett, R. Davuluri, R. Mamorstein, P. M. Lieberman, D. Schultz, S. Hannenhalli, and M. S. Bartolomei. CTCF binding site sequence differences are associated with unique regulatory and functional trends during embryonic stem cell differentiation. *Nucleic Acids Research*, 42(2):774–789, January 2014.

[397] F. Ann Ran, Patrick D. Hsu, Chie-Yu Lin, Jonathan S. Gootenberg, Silvana Konermann, Alexandro E. Trevino, David A. Scott, Azusa Inoue, Shogo Matoba, Yi Zhang, and Feng Zhang. Double Nicking by RNA-Guided CRISPR Cas9 for Enhanced Genome Editing Specificity. *Cell*, 154(6):1380–1389, September 2013.

[398] Jan Lipfert, Xiaomin Hao, and Nynke H. Dekker. Quantitative Modeling and Optimization of Magnetic Tweezers. *Biophysical Journal*, 96(12):5040–5049, June 2009.

[399] Iwijn De Vlaminck, Marijn T.J. van Loenhout, Ludovit Zweifel, Johan den Blanken, Koen Hooning, Susanne Hage, Jacob Kerssemakers, and Cees Dekker. Mechanism of Homology Recognition in DNA Recombination from Dual-Molecule Experiments. *Molecular Cell*, 46(5):616–624, June 2012.

[400] Daniel B. Allan, Thomas Caswell, Nathan C. Keim, Casper M. van der Wel, and Ruben W. Verweij. Soft-matter/trackpy: Trackpy v0.5.0. April 2021.

[401] Andreas S. Biebricher, Iddo Heller, Roel F. H. Roijmans, Tjalle P. Hoekstra, Erwin J. G. Peterman, and Gijs J. L. Wuite. The impact of DNA intercalators on DNA and DNA-processing enzymes elucidated through force-dependent binding kinetics. *Nature Communications*, 6(1):7304, November 2015.

[402] C. Bouchiat, M. D. Wang, J. F. Allemand, T. Strick, S. M. Block, and Vincent Croquette. Estimating the persistence length of a worm-like chain molecule from force-extension measurements. *Biophysical Journal*, 76(1 I):409–413, 1999.

[403] S.H. Chung and R.A. Kennedy. Forward-backward non-linear filtering technique for extracting small biological signals from noise. *Journal of Neuroscience Methods*, 40(1):71–86, November 1991.

[404] Sri Rama Koti Ainavarapu, Jasna Brujić, Hector H. Huang, Arun P. Wiita, Hui Lu, Lewyn Li, Kirstin A. Walther, Mariano Carrion-Vazquez, Hongbin Li, and Julio M. Fernandez. Contour Length and Refolding Rate of a Small Protein Controlled by Engineered Disulfide Bonds. *Biophysical Journal*, 92(1):225–233, January 2007.

[405] Werner Kuhn. Über die Gestalt fadenförmiger Moleküle in Lösungen. *Kolloid-Zeitschrift*, 68(1):2–15, February 1934.

[406] Anders S. Hansen, Tsung Han S. Hsieh, Claudia Cattoglio, Iryna Pustova, Ricardo Saldaña-Meyer, Danny Reinberg, Xavier Darzacq, and Robert Tjian. Distinct Classes of Chromatin Loops Revealed by Deletion of an RNA-Binding Region in CTCF. *Molecular Cell*, 76(3):395–411.e13, 2019.

[407] Kodai Hara, Ge Zheng, Qianhui Qu, Hong Liu, Zhuqing Ouyang, Zhe Chen, Diana R Tomchick, and Hongtao Yu. Structure of cohesin subcomplex pinpoints direct shugoshin-Wapl antagonism in centromeric cohesion. *Nature Structural & Molecular Biology*, 21(10):864–870, October 2014.

[408] Ola Orgil, Avi Matityahu, Thomas Eng, Vincent Guacci, Douglas Koshland, and Itay Onn. A Conserved Domain in the Scc3 Subunit of Cohesin Mediates the Interaction with Both Mcd1 and the Cohesin Loader Complex. *PLOS Genetics*, 11(3):e1005036, March 2015.

[409] Nieves Pezzi, Ignacio Prieto, Leonor Kremer, Luis A. Perez Jurado, Carmen Valero, Jesús Del Mazo, Carlos Martínez-A., and Jose L. Barbero. *STAG3* , a novel gene encoding a protein involved in meiotic chromosome pairing and location of *STAG3* -related genes flanking the Williams-Beuren syndrome deletion. *The FASEB Journal*, 14(3):581–592, March 2000.

[410] Maurici B. Roig, Jan Löwe, Kok-Lung Chan, Frédéric Beckouët, Jean Metson, and Kim Nasmyth. Structure and function of cohesin's Scc3/SA regulatory subunit. *FEBS Letters*, 588(20):3692–3702, October 2014.

[411] Anders S. Hansen. CTCF as a boundary factor for cohesin-mediated loop extrusion: Evidence for a multi-step mechanism. *Nucleus*, 11(1):132–148, January 2020.

[412] Ricardo Saldaña-Meyer, Javier Rodriguez-Hernaez, Thelma Escobar, Mayilaadumveettil Nishana, Karina Jácome-López, Elphege P. Nora, Benoit G. Bruneau, Aristotelis Tsirigos, Mayra Furlan-Magaril, Jane Skok, and Danny Reinberg. RNA Interactions Are Essential for CTCF-Mediated Genome Organization. *Molecular Cell*, 76(3):412–422.e5, November 2019.

[413] Masataka Yamauchi, Giovanni Bruno Brandani, Tsuyoshi Terakawa, and Shoji Takada. SMC complex unidirectionally translocates DNA by coupling segment capture with an asymmetric kleisin path, May 2024.

[414] Tiaojiang Xiao, Julie Wallace, and Gary Felsenfeld. 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. *Molecular and Cellular Biology*, 31(11):2174–2183, June 2011.

[415] Jiao Li, Kaimeng Huang, Gongcheng Hu, Isaac A. Babarinde, Yaoyi Li, Xiaotao Dong, Yu-Sheng Chen, Liping Shang, Wenjing Guo, Junwei Wang, Zhaoming Chen, Andrew P. Hutchins, Yun-Gui Yang, and Hongjie Yao. An alternative CTCF isoform antagonizes canonical CTCF occupancy and changes chromatin architecture to promote apoptosis. *Nature Communications*, 10(1):1535, April 2019.

[416] Carine Berezin, Fabian Glaser, Josef Rosenberg, Inbal Paz, Tal Pupko, Piero Fariselli, Rita Casadio, and Nir Ben-Tal. ConSeq: The identification of functionally and structurally important residues in protein sequences. *Bioinformatics*, 20(8):1322–1324, May 2004.

[417] Barak Yariv, Elon Yariv, Amit Kessel, Gal Masrati, Adi Ben Chorin, Eric Martz, Itay Mayrose, Tal Pupko, and Nir Ben-Tal. Using evolutionary data to make sense of macromolecules with a "face-lifted" ConSurf. *Protein Science*, 32(3):e4582, March 2023.

[418] Richard Evans, Michael O'Neill, Alexander Pritzel, Natasha Antropova, Andrew Senior, Tim Green, Augustin Žídek, Russ Bates, Sam Blackwell, Jason Yim, Olaf Ronneberger, Sebastian Bodenstein, Michal Zielinski, Alex Bridgland, Anna Potapenko, Andrew Cowie, Kathryn Tunyasuvunakool, Rishub Jain, Ellen Clancy, Pushmeet Kohli, John Jumper, and Demis Hassabis. Protein complex prediction with AlphaFold-Multimer, October 2021.

[419] Brett Robison, Vincent Guacci, and Douglas Koshland. A role for the Smc3 hinge domain in the maintenance of sister chromatid cohesion. *Molecular Biology of the Cell*, 29(3):339–355, February 2018.

[420] D. A. Smith. A Quantitative Method for the Detection of Edges in Noisy Time-Series. *Philosophical Transactions: Biological Sciences*, 353(1378):1969–1981, 1998.

[421] Tobias Jachowski. Tobiasjj/stepfinder, March 2023.

[422] Charles Truong, Laurent Oudre, and Nicolas Vayatis. Selective review of offline change point detection methods. *Signal Processing*, 167:107299, February 2020.

[423] Schrödinger, LLC. The PyMOL molecular graphics system, version 1.8. November 2015.

[424] Izabella Krystkowiak and Norman E. Davey. SLiMSearch: A framework for proteome-wide discovery and annotation of functional modules in intrinsically disordered regions. *Nucleic Acids Research*, 45(W1):W464–W469, July 2017.

[425] Peter Jehl, Jean Manguy, Denis C. Shields, Desmond G. Higgins, and Norman E. Davey. ProViz—a web-based visualization tool to investigate the functional and evolutionary features of protein sequences. *Nucleic Acids Research*, 44(W1):W11–W15, July 2016.

[426] Selena R. Martinez and Jj L. Miranda. CTCF terminal segments are unstructured. *Protein Science*, 19(5):1110–1116, May 2010.

[427] Xindan Wang, Tung B.K. Le, Bryan R. Lajoie, Job Dekker, Michael T. Laub, and David Z. Rudner. Condensin promotes the juxtaposition of DNA flanking its loading site in *Bacillus subtilis*. *Genes & Development*, 29(15):1661–1675, August 2015.

[428] Jian Li, Jiangqing Dong, Weitao Wang, Daqi Yu, Xinyu Fan, Yan Chit Hui, Clare S.K. Lee, Wai Hei Lam, Nathan Alary, Yang Yang, Yingyi Zhang, Qian Zhao, Chun-Long Chen, Bik-Kwoon Tye, Shangyu Dang, and Yuanliang Zhai. The human pre-replication complex is an open complex. *Cell*, 186(1):98–111.e21, January 2023.

[429] Jorge A. Fallas, George Ueda, William Sheffler, Vanessa Nguyen, Dan E. McNamara, Banumathi Sankaran, Jose Henrique Pereira, Fabio Parmeggiani, T. J. Brunette, Duilio Cascio, Todd R. Yeates, Peter Zwart, and David Baker. Computational design of self-assembling cyclic protein homooligomers. *Nature Chemistry*, 9(4):353–360, April 2017.

[430] Scott E. Boyken, Zibo Chen, Benjamin Groves, Robert A. Langan, Gustav Oberdorfer, Alex Ford, Jason M. Gilmore, Chunfu Xu, Frank DiMaio, Jose Henrique Pereira, Banumathi Sankaran, Georg Seelig, Peter H. Zwart, and David Baker. De novo design of protein homooligomers with modular hydrogen-bond network–mediated specificity. *Science*, 352(6286):680–687, May 2016.

[431] Cameron J. Glasscock, Robert Pecoraro, Ryan McHugh, Lindsey A. Doyle, Wei Chen, Olivier Boivin, Beau Lonnquist,

Emily Na, Yuliya Politanska, Hugh K. Haddox, David Cox, Christoffer Norn, Brian Coventry, Inna Goreshnik, Dionne Vafeados, Gyu Rie Lee, Raluca Gordan, Barry L. Stoddard, Frank DiMaio, and David Baker. Computational design of sequence-specific DNA-binding proteins. Preprint, Synthetic Biology, September 2023.

[432] Simon Blouin, Timothy D. Craggs, Daniel A. Lafontaine, and J. Carlos Penedo. Functional Studies of DNA-Protein Interactions Using FRET Techniques. In Benoît Leblanc and Tom Moss, editors, *DNA-Protein Interactions*, volume 543, pages 475–502. Humana Press, Totowa, NJ, 2009.

[433] Sungchul Hohng, Chirlmin Joo, and Taekjip Ha. Single-Molecule Three-Color FRET. *Biophysical Journal*, 87(2):1328–1337, August 2004.

[434] Charlotte Cockram, Agnès Thierry, Aurore Gorlas, Roxane Lestini, and Romain Koszul. Euryarchaeal genomes are folded into SMC-dependent loops and domains, but lack transcription-mediated compartmentalization. *Molecular Cell*, 81(3):459–472.e10, February 2021.

[435] Aleem Syed and John A. Tainer. The MRE11–RAD50–NBS1 Complex Conducts the Orchestration of Damage Signaling and Outcomes to Stress in DNA Replication and Repair. *Annual Review of Biochemistry*, 87(1):263–294, June 2018.

[436] Naomichi Takemata, Rachel Y. Samson, and Stephen D. Bell. Physical and Functional Compartmentalization of Archaeal Chromosomes. *Cell*, 179(1):165–179.e18, September 2019.

[437] N. J. Sargentini and K. C. Smith. Characterization of an Escherichia coli mutant (radB101) sensitive to gamma and uv radiation, and methyl methanesulfonate. *Radiation Research*, 93(3):461–478, March 1983.

[438] Robert G. Lloyd, Steven M. Picksley, and Catherine Prescott. Inducible expression of a gene specific to the RecF pathway for recombination in Escherichia coli K12. *Molecular and General Genetics MGG*, 190(1):162–167, April 1983.

[439] Amar Deep, Yajie Gu, Yong-Qi Gao, Kaori M. Ego, Mark A. Herzik, Huilin Zhou, and Kevin D. Corbett. The SMC-family Wadjet complex protects bacteria from plasmid transformation by recognition and cleavage of closed-circular DNA. *Molecular Cell*, 82(21):4145–4159.e7, November 2022.

[440] Hon Wing Liu, Florian Roisné-Hamelin, Bertrand Beckert, Yan Li, Alexander Myasnikov, and Stephan Gruber. DNA-measuring Wadjet SMC ATPases restrict smaller circular plasmids by DNA cleavage. *Molecular Cell*, 82(24):4727–4740.e6, December 2022.

[441] Karl-Peter Hopfner, Lisa Craig, Gabriel Moncalian, Robert A. Zinkel, Takehiko Usui, Barbara A. L. Owen, Annette Karcher, Brendan Henderson, Jean-Luc Bodmer, Cynthia T. McMurray, James P. Carney, John H. J. Petrini, and John A. Tainer. The Rad50 zinc-hook is a structure joining Mre11 complexes in DNA recombination and repair. *Nature*, 418(6897):562–566, August 2002.

[442] Fernando Moreno-Herrero, Martijn De Jager, Nynke H. Dekker, Roland Kanaar, Claire Wyman, and Cees Dekker. Mesoscale conformational changes in the DNA-repair complex Rad50/Mre11/Nbs1 upon binding DNA. *Nature*, 437(7057):440–443, September 2005.

[443] Young-Min Soh, Frank Bürmann, Ho-Chul Shin, Takashi Oda, Kyeong Sik Jin, Christopher P. Toseland, Cheolhee Kim, Hansol Lee, Soo Jin Kim, Min-Seok Kong, Marie-Laure Durand-Diebold, Yeon-Gil Kim, Ho Min Kim, Nam Ki Lee, Mamoru Sato, Byung-Ha Oh, and Stephan Gruber. Molecular Basis for SMC Rod Forma-

tion and Its Dissolution upon DNA Binding. *Molecular Cell*, 57(2):290–303, January 2015.

[444] Hiroki Koide, Noriyuki Kodera, Shveta Bisht, Shoji Takada, and Tsuyoshi Terakawa. Modeling of DNA binding to the condensin hinge domain using molecular dynamics simulations guided by atomic force microscopy. *PLOS Computational Biology*, 17(7):e1009265, July 2021.

[445] Michiko Hirano and Tatsuya Hirano. Opening Closed Arms: Long-Distance Activation of SMC ATPase by Hinge-DNA Interactions. *Molecular Cell*, 21(2):175–186, January 2006.

[446] Allen Chiu, Ekaterina Revenkova, and Rolf Jessberger. DNA Interaction and Dimerization of Eukaryotic SMC Hinge Domains. *Journal of Biological Chemistry*, 279(25):26233–26242, June 2004.

[447] Hisashi Tatebe, Chew Theng Lim, Hiroki Konno, Kazuhiro Shiozaki, Akira Shinohara, Takayuki Uchihashi, and Asako Furukohri. Rad50 zinc hook functions as a constitutive dimerization module interchangeable with SMC hinge. *Nature Communications*, 11(1):370, January 2020.

[448] Naomichi Takemata and Stephen D. Bell. Multi-scale architecture of archaeal chromosomes. *Molecular Cell*, 81(3):473–487.e6, February 2021.

[449] Li Liu, Kaiyuan Han, Huimin Sun, Lu Han, Dong Gao, Qilemuge Xi, Lirong Zhang, and Hao Lin. A comprehensive review of bioinformatics tools for chromatin loop calling. *Briefings in Bioinformatics*, 24(2):bbad072, March 2023.

[450] Xingya Xu and Mitsuhiro Yanagida. Suppressor screening reveals common kleisin–hinge interaction in condensin and cohesin, but different modes of regulation. *Proceedings of the National Academy of Sciences*, 116(22):10889–10898, May 2019.

[451] Jung-Sik Kim, Xiaoyuan He, Jie Liu, Zhijun Duan, Taeyeon Kim, Julia Gerard, Brian Kim, Manoj M. Pillai, William S. Lane, William S. Noble, Bogdan Budnik, and Todd Waldman. Systematic proteomics of endogenous human cohesin reveals an interaction with diverse splicing factors and RNA-binding proteins required for mitotic progression. *Journal of Biological Chemistry*, 294(22):8760–8772, May 2019.

[452] Nan Wu and Hongtao Yu. The Smc complexes in DNA damage response. *Cell & Bioscience*, 2(1):5, 2012.

[453] Jianhua Song, Andrea Lafont, Jingrong Chen, Frank M. Wu, Katsuhiko Shirahige, and Susannah Rankin. Cohesin Acetylation Promotes Sister Chromatid Cohesion Only in Association with the Replication Machinery. *Journal of Biological Chemistry*, 287(41):34325–34336, October 2012.

[454] Benjamin D. Rowland, Maurici B. Roig, Tatsuya Nishino, Alexander Kurze, Pelin Uluocak, Ajay Mishra, Frédéric Beckouët, Philippa Underwood, Jean Metson, Richard Imre, Karl Mechtler, Vittorio L. Katis, and Kim Nasmyth. Building Sister Chromatid Cohesion: Smc3 Acetylation Counteracts an Antiestablishment Activity. *Molecular Cell*, 33(6):763–774, March 2009.

[455] Marjon S. Van Ruiten, Démi Van Gent, Ángela Sedeño Cacciatore, Astrid Fauster, Laureen Willems, Maarten L. Hekkelman, Liesbeth Hoekman, Maarten Altelaar, Judith H. I. Haarhuis, Thijn R. Brummelkamp, Elzo De Wit, and Benjamin D. Rowland. The cohesin acetylation cycle controls chromatin loop length through a PDS5A brake mechanism. *Nature Structural & Molecular Biology*, 29(6):586–591, June 2022.

[456] Amit K. Singh, Qingrong Chen, Cu Nguyen, Daoud Meerzaman, and Dinah S. Singer. Cohesin regulates alternative splicing. *Science Advances*, 9(9):eade3876, March 2023.

[457] Dingquan Yu, Grzegorz Chojnowski, Maria Rosenthal, and Jan Kosinski. AlphaPulldown—a python package for protein–protein interaction screens using AlphaFold-Multimer. *Bioinformatics*, 39(1):btac749, January 2023.

[458] Nicolas L. Jean, Trevor J. Rutherford, and Jan Löwe. FtsK in motion reveals its mechanism for double-stranded DNA translocation. *Proceedings of the National Academy of Sciences*, 117(25):14202–14208, June 2020.

[459] Peixuan Guo, Zhengyi Zhao, Jeannie Haak, Shaoying Wang, Dong Wu, Bing Meng, and Tao Weitao. Common mechanisms of DNA translocation motors in bacteria and viruses using one-way revolution mechanism without rotation. *Biotechnology Advances*, 32(4):853–872, July 2014.

[460] Estelle Crozat and Ian Grainge. FtsK DNA Translocase: The Fast Motor That Knows Where It's Going. *ChemBioChem*, 11(16):2232–2243, November 2010.

[461] Paul J. Pease, Oren Levy, Gregory J. Cost, Jeff Gore, Jerod L. Ptacin, David Sherratt, Carlos Bustamante, and Nicholas R. Cozzarelli. Sequence-Directed DNA Translocation by Purified FtsK. *Science*, 307(5709):586–590, January 2005.

[462] Hong-Yun Wang, Tim Elston, Alexander Mogilner, and George Oster. Force Generation in RNA Polymerase. *Biophysical Journal*, 74(3):1186–1202, March 1998.

[463] James E. Graham, David J. Sherratt, and Mark D. Szczelkun. Sequence-specific assembly of FtsK hexamers establishes directional translocation on DNA. *Proceedings of the National Academy of Sciences*, 107(47):20263–20268, November 2010.

[464] Elio A. Abbondanzieri, William J. Greenleaf, Joshua W. Shaevitz, Robert Landick, and Steven M. Block. Direct observation of base-pair stepping by RNA polymerase. *Nature*, 438(7067):460–465, November 2005.

[465] V. Epshtein. Transcription through the roadblocks: The role of RNA polymerase cooperation. *The EMBO Journal*, 22(18):4719–4727, September 2003.

[466] Ja Yil Lee, Ilya J. Finkelstein, Lidia K. Arciszewska, David J. Sherratt, and Eric C. Greene. Single-Molecule Imaging of FtsK Translocation Reveals Mechanistic Features of Protein-Protein Collisions on DNA. *Molecular Cell*, 54(5):832–843, June 2014.

[467] Kathleen A. Marquis, Briana M. Burton, Marcelo Nollmann, Jerod L. Ptacin, Carlos Bustamante, Sigal Ben-Yehuda, and David Z. Rudner. SpoIIIE strips proteins off the DNA during chromosome translocation. *Genes & Development*, 22(13):1786–1795, July 2008.

Acknowledgments

"The dwarf sees farther than the giant, when he has the giant's shoulder to mount on."

These words of Samuel Taylor Coleridge in *The Friend* resonate deeply with me. Undoubtedly, the journey to becoming a scientist is a collaborative venture that transcends the boundaries of science alone. Along this path, I have been hoisted by many giants. Being immersed in such an incredible community of family, friends, and phenomenal scientists is a fortune beyond what could have been wished for.
Thank you for all you have contributed in one of innumerable ways!

Roman
Delft, May 2024

Curriculum Vitæ

Roman Barth

1995/08/09	Born in Bruchsal, Germany
2013-2016	**Bachelor of Science in Physics** Humboldt-Universität zu Berlin, Berlin, Germany *Thesis:* Implementation and testing of 405 nm laser diodes as excitation source for the detection of singlet oxygen luminescence using fiber optics. **Supervisor**: Prof. Dr. rer. habil. Beate Röder
2015-2016	**Working student** Fraunhofer Institute for Telecommunications, Berlin, Germany Wafer processing for integrated optical elements
2015-2016	**Research internship** Laboratoire de Biologie Moléculaire Eucaryote (CBI; CNRS), Toulouse, France Quantitative mapping of chromatin dynamics in living cells **Supervisor**: Prof. dr. Kerstin Bystricky
2017-2019	**Master of Science in Applied Physics** Delft University of Technology, Delft, The Netherlands Track Bionanoscience *Thesis:* Genome-sized DNA compaction by condensin **Supervisor**: Prof. dr. Cees Dekker *Cum laude distinction*
2019	**Industry internship** F. Hoffmann-La Roche AG, Basel, Switzerland Oligonucleotide encapsulation in lipid nanoparticles for brain delivery
2020-2024	**PhD in Single Molecule Biophysics** *Thesis:* Exploring SMC motor proteins and how to stop them **Promoters**: Prof. dr. Cees Dekker & Prof. dr. A. Marileen Dogterom

List of Publications

* denotes equal contribution

Included in this thesis.

R. Barth*, R. Janissen*, L. Muras*, J. v.d. Torre, E. v.d. Sluis, A. v.d. Berg, I. F. Davidson, J.-M. Peters, C. Dekker: Two CTCF motifs impede cohesin-mediated DNA loop extrusion, *in preparation*

I. F. Davidson, **R. Barth**, S. Horn, R. Janissen, K. Nagasaka, G. Wutz, R. R. Stocsits, B. Bauer, C. Dekker, J.-M. Peters: Cohesin supercoils DNA during loop extrusion, bioRxiv, 2024

R. Janissen*, **R. Barth***, I. F. Davidson, M. Taschner, S. Gruber, J.-M. Peters, C. Dekker: All eukaryotic SMC proteins induce a twist of -0.6 at each DNA-loop-extrusion step, bioRxiv, 2024

R. Barth, I. F. Davidson, J. van der Torre, M. Taschner, S. Gruber, J.-M. Peters, C. Dekker: SMC motor proteins extrude DNA asymmetrically and contain a directional switch, bioRxiv, 2023

B. T. Analikwu*, A. Deshayes*, J. van der Torre, T. Guérin, A. J. Katan, C. Béneut, **R. Barth**, J. Phipps, V. Scolari, X. Veaute, C. Barrington, D. Busso, F. Uhlmann, K. Dubrana, S. Mattarocci, C. Dekker, S. Marcand: Telomere protein arrays stall DNA loop extrusion by condensin, bioRxiv, 2023

R. Janissen, **R. Barth**, M. Polinder, J. van der Torre, C. Dekker: Single-molecule visualization of twin-supercoiled domains generated during transcription, Nucleic Acids Research, 2023

M. Tišma, R. Janissen, H. Antar, A. Martin-Gonzalez, **R. Barth**, T. Beekman, J. van der Torre, D. Michieletto, S. Gruber, C. Dekker: Dynamic ParB–DNA interactions initiate and maintain a partition condensate for bacterial chromosome segregation, Nucleic Acids Research, 2023

N. de Franceschi, **R. Barth**, S. Meindlhumer, A. Fragasso, C. Dekker: Dynamin A as a one-component division machinery for synthetic cells, Nature Nanotechnology, 2023

Roman Barth*, Biswajit Pradhan*, E. Kim, I. Davidson, J. van der Torre, J.-M. Peters, C. Dekker: Testing pseudotopological and nontopological models for SMC-driven DNA loop extrusion against roadblock-traversal experiments, Scientific Reports, 2023

E. Kim*, **R. Barth***, C. Dekker: Looping the genome with SMC complexes, Annual Rev. of Biochemistry, 2023

I. Davidson*, **R. Barth***, M. Zaczek*, J. van der Torre, W. Tang, K. Nagasako, R. Janissen, J. Kerssemakers, G. Wutz, C. Dekker, J.-M. Peters: CTCF is a DNA-tension-dependent barrier to cohesin-mediated loop extrusion, Nature, 2023

R. Barth, H. A. Shaban: Spatially coherent diffusion of human RNA Pol II depends on transcriptional state rather than chromatin motion, Nucleus, 2022

B. Pradhan, **R. Barth**, E. Kim, I.F. Davidson, B. Bauer, T. van Laar, W. Yang, J.-K. Ryu, J. van der Torre, J.-M. Peters, C. Dekker: SMC complexes can traverse physical roadblocks bigger than their ring size, Cell Reports, 2022

M. Tišma, M. Panoukidou, H. Antar, Y. M. Soh, **R. Barth**, B. Pradhan, A. Barth, J. van der Torre, D. Michieletto, S. Gruber, C. Dekker: ParB proteins can bypass DNA-bound roadblocks via dimer-dimer recruitment, Science Advances, 2022

C. Valades-Cruz*, **R. Barth***, M. Abdellah*, H. A. Shaban: Genome-wide analysis of the dynamic and biophysical properties of chromatin and nuclear proteins in living cells with Hi-D, Nature Protocols, 2024, *accepted*

H. A. Shaban, **R. Barth**, K. Bystricky: Navigating the crowd: visualizing coordination between genome dynamics, structure, and transcription, Genome Biology, 2020

R. Barth, K. Bystricky, H. A. Shaban: Coupling chromatin structure and dynamics by live super-resolution imaging, Science Advances, 2020

H. A. Shaban, **R. Barth**, L. Recoules, K. Bystricky: Hi-D: nanoscale mapping of nuclear dynamics in single living cells, Genome Biology, 2020

R. Barth, G. Fourel, H. A. Shaban: Dynamics as a cause for the nanoscale organization of the genome, Nucleus, 2020

H. A. Shaban*, **R. Barth***, K. Bystricky: Formation of correlated chromatin domains at nanoscale dynamic resolution during transcription, Nucleic Acids Research, 2018

