Inaugural dissertation

for

obtaining the doctoral degree

of the

Combined Faculty of Mathematics, Engineering and

Natural Sciences

of the

Ruprecht – Karls – Universität

Heidelberg

Presented by

M.Sc. Léa Lecomte

Born in: Alençon, France

Oral examination: 6th of April 2023

Structural studies of DNA-bound Condensin

Referees:
Prof. Dr. Frauke Melchior

Dr. Simone Koehler

Summary of the thesis

Eukaryotic chromosomes undergo massive three-dimensional rearrangements over the course of the cell cycle. Prior to cell division, chromosomes are condensed into rod-shaped structures which ensures the faithful transmission of the genetic material into daughter cells. Emerging evidence from biochemical and chromosome conformation capture experiments suggests that condensin, an evolutionary conserved multi-subunit protein complex, catalyzes chromosome condensation through the extrusion of DNA loops using energy derived from ATP hydrolysis. However, the underlying molecular mechanism remains largely elusive. The objective of my PhD thesis was to gain insights into the structural basis of DNA loop extrusion by condensin.

In this thesis, I report the structure of *Saccharomyces cerevisiae* (*Sc*) condensin bound to ATP and to DNA solved by cryo-electron microscopy (cryo-EM) at resolutions of 3.5 to 3.9 Å. The structure reveals that DNA binds at two distinct condensin modules, which I named "core" and "periphery". At both modules, DNA is entrapped into topological compartments formed by the condensin subunits. In addition, I present a low-resolution structure (~10 Å) of *Sc* condensin core bound to DNA, without ATP. The integration of structural information obtained during my PhD with previously available condensin structures, biochemical experiments and single-molecule data results in a detailed model for condensin-mediated DNA loop extrusion. My work provides a molecular explanation for the motor and anchor functions that are required for condensin to reel DNA into a loop while stably holding onto the DNA. The molecular rearrangements of condensin core upon DNA and ATP binding suggest a "power-stroke" motor mechanism of condensin.

In addition, I present first experiments that investigate the DNA sequence specificity of condensin binding, its association with nucleosomes, and its interaction with the general transcription factor TFIIIC. These new experimental data pave the way to understand how condensin is loaded onto chromatin.

Zusammenfassung der Arbeit

Eukaryotische Chromosomen werden im Laufe des Zellzyklus beachtlich in ihrer dreidimensionalen Struktur umgestaltet. Vor der Zellteilung werden die Chromosomen zu stäbchenförmigen Strukturen kondensiert, um die zuverlässige Weitergabe des genetischen Materials an die Tochterzellen zu gewährleisten. Neue Erkenntnisse aus biochemischen Experimenten und Experimenten zur Erfassung der Chromosomenkonformation deuten darauf hin, dass Condensin – ein evolutionär konservierter Proteinkomplex mit mehreren Untereinheiten – mittels ATP-Hydrolyse die Chromosomenkondensation durch die Extrusion von DNA-Schleifen katalysiert. Der zugrundeliegende molekulare Mechanismus ist jedoch noch weitgehend ungeklärt. Ziel meiner Doktorarbeit war es, Einblicke in die strukturellen Grundlagen der Extrusion von DNA-Schleifen durch Condensin zu gewinnen.

In dieser Arbeit berichte ich über die Struktur von Saccharomyces cerevisiae (Sc) Condensin, das an ATP und an DNA gebunden ist, gelöst durch Cryo-Elektronenmikroskopie (Cryo-EM) mit einer Auflösung von 3,5 bis 3,9 Å. Die Struktur zeigt, dass DNA an zwei unterschiedliche Condensin-Module bindet, die ich "Kern" und "Peripherie" genannt habe. An beiden Modulen wird die DNA in topologische Kompartimente eingeschlossen, die von den Condensin-Untereinheiten gebildet werden. Zusätzlich zur ATP-gebundenen Struktur präsentiere ich eine niedrig aufgelöste Struktur (~10 Å) des an DNA gebundenen Sc Condensin-Kerns ohne ATP. Die Integration der während meiner Promotion gewonnenen strukturellen Informationen, mit bereits verfügbaren Condensin-Strukturen, biochemischen Experimenten und Einzelmoleküldaten führt zu einem detaillierten Modell für die Condensin-vermittelte DNA-Schleifen-Extrusion. Meine Arbeit liefert eine molekulare Erklärung für die Motor- und Ankerfunktionen, die erforderlich sind, damit Condensin die DNA zu einer Schleife aufrollen und dabei stabil an der DNA festhalten kann. Die molekularen Umstrukturierungen des Condensin-Kerns bei DNA- und ATP-Bindung deuten auf einen Condensin-Motor-Mechanismus hin, der "Power-Stroke" genannt wird. Schließlich präsentiere ich erste Experimente, die die DNA-Sequenzspezifität der Condensin-Bindung, seine Assoziation mit Nukleosomen und seine Interaktion mit dem allgemeinen Transkriptionsfaktor TFIIIC untersuchen. Diese neuen experimentellen Daten ebnen den Weg, um zu verstehen, wie Condensin auf Chromatin geladen wird.

Publications

Shaltiel IA, Datta S *, **Lecomte L** *, Hassler M, Kschonsak M, Bravo S, Stober C, Ormanns J, Eustermann S, Haering CH. Science (2022). A hold-and-feed mechanism drives directional DNA loop extrusion by condensin.

Datta S *, **Lecomte L** *, Haering CH. Current opinion in structural biology (2020). Structural insights into DNA loop extrusion by SMC protein complexes.

*: equal contribution

Lee B-G, Merkel F, Allegretti M, Hassler M, Cawood C, **Lecomte L**, O'Reilly FJ, Sinn LR, Gutierrez-Escribano P, Kschonsak M, Bravo S, Nakane T, Rappsilber J, Aragon L, Beck M, Löwe J, Haering CH (2020). Cryo-EM structures of holo condensin reveal a subunit flip-flop mechanism.

Acknowledgments

I would like to thank Christian Haering for giving me the opportunity to work on this fascinating protein complex and for his support throughout my PhD. I thank him for supporting me to stay at EMBL to pursue my cryo-EM project.

I am thankful to Sebastian Eustermann, for his guidance, support and for generously welcoming me in his group. I am grateful to have had the chance to benefit from his expertise in cryo-EM.

Thank you to Prof. Dr. Melchior and Dr. Simone Koehler for accepting to review this thesis. Thank you to my Thesis Advisory Committee (TAC) and my thesis committee members: Dr. Simone Koehler, Prof. Dr. Knop, Prof. Dr. Frauke Melchior, Dr. Christoph Mueller, Dr. Helgo Schmidt, Prof. Dr. Irmi Sinning for their support and guidance.

I thank Markus Hassler for his tremendous contribution to this project through his expertise in model building. A huge thank you to Felix Weiss for answering my despair calls from the microscopes and for his tips. I am very grateful to Wim Hagen and Simon Fromm for their support on the microscopes and for maintaining the machines. It was a privilege for me to learn from the experts. Thank you to Thomas Hoffmann for his patience, for maintaining the cluster and for solving an endless list of processing issues. I thank Fabian Merkel for introducing me to cryo-EM sample preparation and data processing. I thank Sol Bravo for introducing me to protein purification and for maintaining the AKTA. I am grateful to Karine Lapouge for teaching me how to use the mass photometer.

I'd like to thank all my colleagues from the Haering lab and especially Indra Shaltiel and Sumanjit Datta, with whom I collaborated closely. Thank you for all the work you have put into our paper and for working with such a collaborative spirit. Our endless kicker games were an incredible source of joy. I'd also like to thank my colleagues from the Eustermann lab for welcoming me so warmly. An immense thank you goes to Dana Grozavu for sharing with me a minuscule office and for her kindness and support during

the writing of this manuscript. I also thank her for performing binding assay to probe the interaction of condensin with the di-nucleosome library that she generated. Special thank you to Anna Jungblut for proofreading the German summary of my thesis. Thank you to Min Zhang for sharing cryo-EM analysis tips. Thank you to Olga Kolesnikova for managing the lab and for her support and tips. Thank you to Thomas Dahlet for performing the SELEX-seq experiments.

Thank you to Wolfram Seifert Davila for providing the TFIIIC purified sample and for working with me (and sharing my enthusiasm) on the condensin-TFIIIC interaction project.

Thank you to Chris Rickerby for putting together amazing and challenging cross-training workouts every week at EMBL. This helped me to stay on tracks during the ups and downs of my PhD.

Thank you to my friend Kevin, for his trust and for his friendship, for making me laugh and for his wisdom that helped me to put things in perspectives, especially throughout the difficult times. I am extremely grateful to my lovely friend Andrea, she taught me tolerance and determination, and offered me an unshakeable friendship.

I'd like to thank my parents for their support and love. Thank you for teaching me ambition and determination. Writing these words, I am thinking to my grandmother Solange, who wanted to study but was not given the chance, as many women of her generation. I thank her for her care and kindness.

I thank my merveilleuse sister Juliette for her support and for her love. I am extremely grateful to have her by my side.

Last, I would like to thank my lover, partner, best friend Arthur from whom I lived far away during my PhD.

Table of content

List of figures

Introduction

I. Overview of the eukaryotic the cell cycle

DNA contains all information required for the functioning, development and reproduction of all living organisms. The cell cycle describes the process that controls the events that lead to the division of a cell into two daughter cells, which each inherit a complete set of genomic DNA from their mother cell. In eukaryotic cells, which are characterized by a compartmentalization of their genomic material into a nucleus, the cell cycle can be divided in two phases: (i) interphase, which can be further subdivided into G1 (gap 1), S and G2 (gap 2) phases, and (ii) M phase, which consists of mitosis and cytokinesis.
In growing phases G1and G2, cellular components and organelles are produced. During S-phase, DNA is duplicated to create two identical copies of every chromosome, called sister chromatids, that will be segregated during M-phase. All steps of the cell cycle are highly regulated in order to maintain genomic integrity and to ensure the faithful transmission of the entire genetic material into the daughter cells.
Throughout the cell cycle, DNA molecules are continuously rearranged, culminating in the transition from a loose interphase chromatin fiber to compacted chromosomes during mitosis. Given the size of eukaryotic genomes (from ~10 mega bases (Mb) in some fungi up to >100 000 Mb in some animals and plants (Sessions, 2013)) organizing chromosomes during interphase and mitosis represents a major cellular challenge and requires efficient strategies that take care of disentangling and compacting the DNA.

II. Eukaryotic chromatin and chromosome architectures

Nucleosomal packing

The micrometer-sized nucleus of the eukaryotic cell contains about two meters of DNA molecules. In order to fit into the nucleus, DNA is compacted and organized into

chromatin. The basic packing unit of chromatin is the nucleosome. It results from the wrapping of 147 base pairs (bp) of DNA around histone octamers, composed of two copies of H2A/H2B and of H3/H4 dimers (Luger et al., 1997). This first level of organization of chromatin results in the formation of a 11-nm fiber that was first observed *in vitro* on reconstituted nucleosome arrays (Rattner and Hamkalo, 1978) and described as the "beads-on-string" architecture (Olins and Olins, 2003). Whether or not interactions between neighbor nucleosomes on the same DNA are involved in the higher-order ultrastructure of interphase and mitotic chromosomes, has been largely debated in the field of chromosome biology (Maeshima et al., 2010). A 30-nm fiber was reported from studies of *in vitro* reconstituted nucleosome arrays (Huynh et al., 2005; Robinson et al., 2006). Later, Rhodes and co-workers reconstituted nucleosome arrays *in vitro* using chicken erythrocyte linker histone H5 and showed that the chromatin organized into a 30-nm fiber (Robinson and Rhodes, 2006). Two models attempted to describe the 30-nm fiber organization observed *in vitro*: (i) the zig-zag model and (ii) the solenoid model. The adoption of one rather than the other would depend on the inter-nucleosomal DNA linker length (Routh et al., 2008).

On the contrary, data obtained *in vivo* do not support the existence of the 30-nm fiber in most cells, but rather suggest a disordered architecture where chromatin is organized into regions of different densities and accessibility. Pioneer cryo-electron microscopy (cryo-EM) work from Dubochet et *al.* carried out on mammalian mitotic cells did not reveal the presence of defined higher-ordered chromatin structures, such as the 30-nm fiber (Dubochet et al., 1988; McDowall et al., 1986). Later, cryo-EM based approaches imaging the isolated *Xenopus* chromosomes or human cells following fixation and selective DNA staining could also not visualize a 30-nm fiber. In the latter case, the interphase chromatin density appeared variable, arguing against the existence of a stable 30-nm fiber *in vivo*. The discrepancy between *in vitro* reconstituted nucleosomal arrays and the results obtained *in vivo* or using purified chromosomes might result from different

concentration effects: the nucleosomes contained in the *in vitro* generated array were highly diluted compared to nucleosomes in purified chromosomes or in the nuclei of human cells. Thus, in a diluted context, the interactions that occur between neighboring nucleosomes on the same DNA (intra-molecular) are favored compared to interactions between different DNAs (inter-molecular), which are certainly more likely *in vivo*. As a result, chromatin is best described as a dynamically folding fiber, referred to as "polymer melt". This description implies that the chromatin fiber constantly rearranges locally by switching between intra- and inter-molecular nucleosome interactions. (Ou et al., 2017).

Chromosome territories

In the interphase nucleus of most eukaryotes, chromosomes are not randomly entangled but are instead organized into chromosome territories (Bolzer et al., 2005). Fluorescence *in situ* hybridization (FISH) studies revealed that human chromosomes form spherical domains of ~2 micrometers in diameter. The partitioning of chromosomes into distinct territories in the nucleus limits the interactions between different chromosomes.

The nuclear envelope acts as a tethering point for chromosomes: The nuclear lamina, a meshwork of filamentous lamin proteins situated along the inner nuclear membrane, preferentially cross-links to certain chromosomal regions termed lamina-associated domains (LADs). These regions are transcriptionally silent and are generally composed of heterochromatin (see below) (Briand and Collas, 2020). Other chromosomal regions cluster at the nucleolus due to the association of their ribosomal RNA-coding genes (rDNA), which results in the coordinated expression of these genes. In addition to the rRNA-coding genes, additional sequences, present in most chromosomes and most frequently transcriptionally silent, surrounds the nucleolus and create was has been called nucleolar-associated domains (NADs) (van Koningsbruggen et al., 2010).

One factor responsible for the establishment of chromosomal territories in *D. melanogaster* is the protein complex condensin II (see below for an extended description) (Bauer

et al., 2012). By inducing axial compaction of interphase chromosomes, condensin II strongly reduces the inter-chromosomal contacts, which results in chromosome partition. In addition, nucleoli of neural stem cells are over-compacted or "hyper-clustered" when the condensin II is mutated, suggesting a key role of this complex in the establishment of the nucleolus (Nishide and Hirano, 2014).

Chromatin compartments

The compaction of chromatin regulates its accessibility to chromosomal proteins. Relaxed chromatin, or euchromatin, is generally accessible and positively correlates with transcriptional activity. On the contrary, more compacted chromatin, or heterochromatin, is usually transcriptionally silent. Hi-C, a method based on the detection of spatially proximal genomic regions, revealed that euchromatin and heterochromatin in human cells are organized into separate compartments that span a few mega-bases and are called A (transcriptionally active) and B (transcriptionally silent) compartments (Lieberman-Aiden et al., 2009). The formation of so-called "transcription factories'' might assist in the formation of these chromatin compartments (Sutherland and Bickmore, 2009). Clustering of functionally related genes was also reported in yeast (Teichmann and Veitia, 2004) and in mouse cells in the case of globin genes (Osborne et al., 2004). Numerous studies (Chubb et al., 2002; Heun et al., 2001; Vazquez et al., 2001) highlighted the mobility of chromatin, which suggests that transcription factories can assemble dynamically. This finding challenges the view that RNA polymerase II (RNA pol II) would solely be recruited at active promoters but rather suggests that active genes are recruited to RNA pol II-containing compartments to achieve a fast-transcriptional adaptive response.

Topologically associating domains (TADs)

Hi-C and related approaches revealed a suborganization of chromatin inside the compartments into so-called Topological Associating Domains (TADs), which extend over

few hundred kilobases (kb). By definition, Intra-TAD contact frequencies are usually twice as high as inter-TAD contacts (Dekker et al., 2002; Dixon et al., 2012; Hou et al., 2012; Nora et al., 2012; Sexton et al., 2012). Based on these criteria, more than 2,000 TADs have been reported in human and mouse cells, which broadly cover more than 90% of each genome. Interestingly, the positioning of TADs seems to be conserved across mammalian species and among different cell types (Dixon et al., 2012). TAD boundaries are mostly found in the proximity of transcription start sites (TSS) of genes associated with active histone marks (H3K4me3 and H3K36me3) and are enriched in CTCF (CCCTC-binding factor) and cohesin proteins. CTCF associates with more than 75% of the TAD boundaries in mouse cells (Dixon et al., 2012). Depletion of CTCF reduced intra-TAD interactions and increased inter-TAD interactions, which suggests than CTCF plays an insulating role for TAD formation (Zuin et al., 2014). Surprisingly, depletion of cohesin did not affect TADs boundaries in human cells, although it led to a general diminution of chromatin contacts.

TADs also correlate with transcriptional activity. A study carried out in *Drosophila melanogaster* classified TADs into subtypes according to their epigenetic profiles, differentiating active TADs (associated with the euchromatin epigenetic marks H3K4me3 and H3K36me3) from the constitutively repressed TADs (associated with HP1 binding and the H3K9me2 heterochromatin epigenetic mark) and the facultatively repressed TADs (associated with the Polycomb-group proteins (PcG) and the H3K27me3 heterochromatin epigenetic mark) (Sexton et al., 2012). Furthermore, cohesin or CTCF depletion in human cells led to misregulation of distinct groups of genes (Sexton et al., 2012).

Loop domains

TADs can be further subdivided into loop domains. Interestingly, loops of about 100-200 kilobases in length are conserved across numerous cell types and between organisms (Rao et al., 2014). Loops are frequently found at the boundaries of TADs. Loops are delimited

by convergent CTCF binding sites (i.e. the binding motifs are facing each other) in >90% of the cases, as opposed to TAD boundaries, which are enriched in divergent CTCF binding sites (Nanni et al., 2020). Loops frequently bring together enhancers and promoters, providing an explanation for how these distant genomic regions, which can be several kilobases apart, can come into contact in order to recruit transcription factors and activate gene expression. Loops are thought to result from cohesin loop extrusion activity (see section IV).

Mitotic chromosome organization

Prior to division, cells individualize and compact their chromosomes into rod-shaped structures, which ensures the faithful segregation of the genetic material. The compaction of chromosomes is a drastic structural rearrangement of the chromatin fibers, which has remained incompletely understood since the first microscopic images of mitotic chromosomes obtained by W. Flemming during the late 19th century (Flemming, 1882). During this process, most of the architectural compartmentalization of chromatin described in the previous sections is lost (Naumova et al., 2013). Furthermore, the transcriptional activity is greatly reduced in many species.

The individualization of the chromosomes requires the resolution of catenanes generated between sister chromatids as a result of DNA replication torsional stress. Formation of catenanes is avoided by the action of topoisomerases, which can be divided into type I and type II topoisomerases based on their catalytic activities. Type-I topoisomerases cut one of the two strands of double-stranded DNA (dsDNA) in order to resolve torsional stress in the DNA double helix (Promonet et al., 2020; Tuduri et al., 2009). On the contrary, type-II topoisomerases induce double-stranded DNA breaks, creating a gap through which they pass another dsDNA (Holm et al., 1985; Sundin and Varshavsky, 1981; Uemura et al., 1987). Doing so, type-II topoisomerases can resolve the above-mentioned catenanes formed between chromatids. The faithful distribution of chromosomes

to the daughter cells further requires the identification and cohesion of the sister chromatids. Cohesion is generated during DNA replication and is maintained until anaphase, when sister chromatids are separated from each other. Factors responsible for sister chromatid cohesion were discovered by genetic screens carried out with mitotically arrested yeast cells that identified mutants unable to hold together sister chromatids (reviewed in Nasmyth et al., 2000). These studies identified the cohesin protein complex, which is composed of five subunits (Smc1, Smc3, Scc1, Scc3 and Scc2-Scc4). The cohesin-associated factor Pds5 was also identified by these screens, together with Eco1, an enzyme responsible for cohesin acetylation (Rolef Ben-Shahar et al., 2008). These pioneering studies in yeast have been followed up with experiments with cell-free extracts prepared from *Xenopus* oocytes, which showed that cohesin depletion reduced cohesion, and with similar genetics and cell biology experiments carried out in *S. pombe* and *A. thaliana* (Bai et al., 1999; Bhatt et al., 1999; Nasmyth, 2001; Tomonaga et al., 2000). Cohesin is thought to embrace the sister chromatids within a ring structure formed by three of its subunits: Smc1, Smc3 and Scc1 (see the next section: Structural Maintenance of Chromosome complexes). Sister chromatid cohesion is abolished during anaphase as a result of the cleavage of the Scc1 subunit by an enzyme called separase (Uhlmann et al., 1999). The minimal factors required to compact chromosomes were identified using cell-free extracts obtained from *Xenopus* oocytes : the core histones, three histone chaperones (nucleoplasmin, Nap1 and FACT), topoisomerase II and condensin protein complexes (Shintomi et al., 2015). M-phase *Xenopus* egg extracts depleted for condensin complexes did not trigger compaction of sperm chromatin into mitotic chromosomes. In addition, depletion of condensin subunits in human cells led to defects in kinetochore structure and function, inducing chromosomal alignment impairment and mis-segregation (Samoshkin et al., 2009). Vertebrate cells contain two types of condensin complexes termed condensin I and condensin II (Ono et al., 2003). In addition to differences in their subunit composition, the two complexes differ in their cellular localization during the cell cycle:

whereas condensin II can be detected in the nucleus throughout the cell cycle and starts to accumulate on the chromosomes at the beginning of the prophase, condensin I only binds to chromosomes after nuclear envelope breakdown (Hirota et al., 2004). Localization and concentration of chromosome-associated condensin I and condensin II also differ, as shown by fluorescence correlation spectroscopy (FSC)-calibrated imaging and super-resolution microscopy experiments carried out with human cells (Walther et al., 2018). About 35,000 condensin II complexes stably associate with human chromosomes throughout mitosis. Condensin I is five to six times more abundant (~195,000 complexes) and binds to chromosomes in two distinct waves: first during prometaphase and second during early anaphase. Condensin II complexes are confined to the central axis of the chromosome, whereas condensin I complexes are distributed over half the width of the chromatids, which suggests that the two complexes might have different roles in chromosome formation. Depletion experiments indeed suggest that condensin II promotes axial compaction of chromosomes, whereas condensin I further accentuates compression of mitotic chromosomes by triggering their lateral compaction (Green et al., 2012; Maddox et al., 2006; Ono et al., 2003; Yamashita et al., 2011). The compaction of mitotic chromosomes driven by condensins is thought to occur through the formation of chromatin loops, a process known as DNA loop extrusion. I will describe this molecular process in depth in section IV.

Topoisomerase II is also a key factor of mitotic chromosome compaction. It localizes to the mitotic chromosome axis (Earnshaw and Heck, 1985; Gasser et al., 1986; Kireeva et al., 2004; Maeshima and Laemmli, 2003) and its inactivation in *S. pombe* coincides with the formation of unresolved mitotic chromosomes (Uemura et al., 1987). Removal of topoisomerase II alpha (one of the two isomers produced by vertebrate cells) does not perturb early prophase compaction but it instead impairs prometaphase compaction of the chromosomes, leading to exit from mitosis (Nielsen et al., 2020). This suggests that topoisomerase II alpha is a determinant for the maintenance of mitotic chromosomes,

most likely by catalyzing the disentanglement of chromatin loops generated by condensins.

The protein Kif4A was also reported to promote chromatin compaction (Mazumdar et al., 2004; Takahashi et al., 2016), but its exact role in mitotic chromosome formation is still elusive. In addition, mitosis-specific epigenetic histone marks might facilitate the formation of the chromosomal helical scaffold by promoting inter-nucleosomal interactions (Kschonsak and Haering, 2015).

Figure 1.1 summarizes the different levels of organization of the eukaryotic chromosomes during interphase and mitosis, described in the previous sections, and highlights the importance of the cohesin and condensin loop extrusion factors (LEF) to establish and maintain these structures.

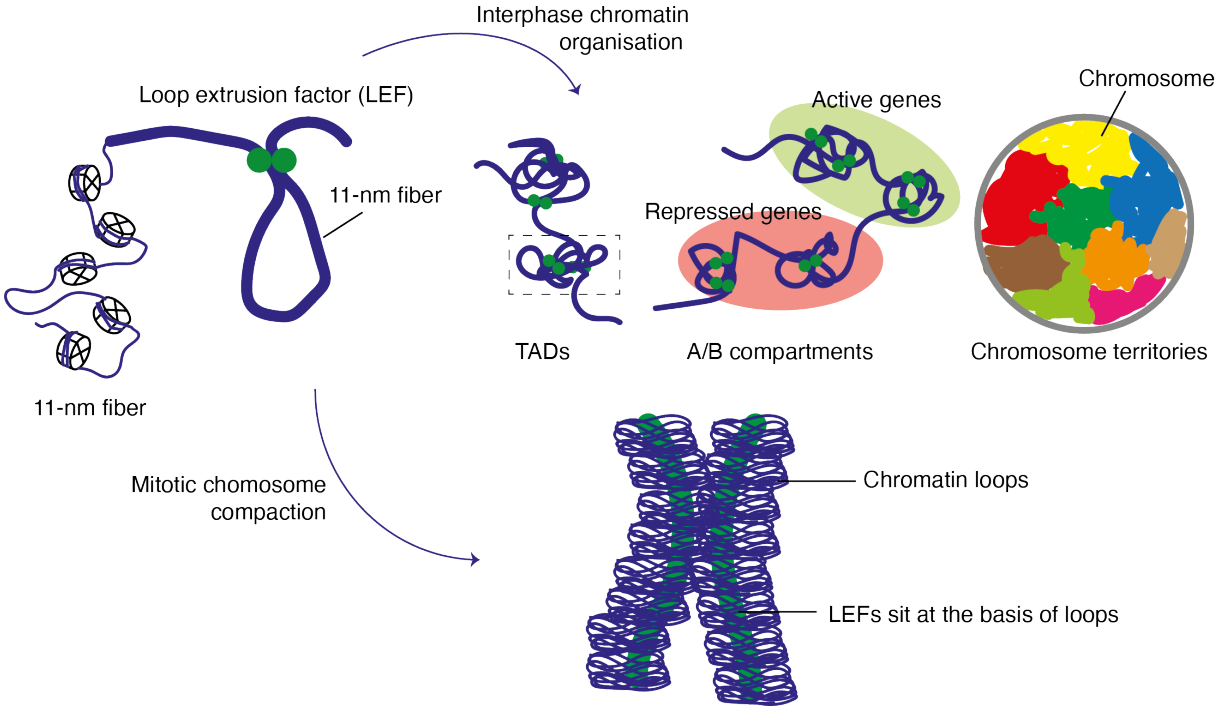


Figure 3.1. Schematic representation of the eukaryotic genome organization. The genome is organized into an 11-nm chromatin fiber. During interphase, the genome folds into additional tri-dimensional elements such as Topological Associated Domains (TADs) under the action of SMC complex loop extruding factors (LEF). Genomic regions are additionally organized into compartments and chromosome territories. During mitosis, LEF compact the chromatin to form rod-like chromosome structures.

III. Prokaryotic chromosome architecture

Organization of the genome by nucleiod-associated proteins (NAPs)

Prokaryotes are unicellular organisms divided into two domains: archaea and bacteria. Prokaryotic cells are devoid of nuclei. Their genomes, usually organized into circular dsDNA chromosomes of several million base pairs in size, are instead organized into so-called nucleoids. These membrane-less entities are structured by DNA-associated proteins and can contain RNAs. Archaeal genomes are packed by specific histones into "nucleosomes" of yet unknown structure (Henneman et al., 2018). Bacterial genomes are deprived of histones and therefore are not primarily packed into nucleosomes. Instead, bacterial genomes are organized by Nucleoid Associated Proteins (NAPs) of various nature that differ between species (for a review, see Hołówka and Zakrzewska-Czerwińska, 2020). Some NAPs are reminiscent of the eukaryotic histones and induce a strong DNA bent or even DNA looping. For example, the histone-like nucleoid structuring protein (H-NS) found in *E. coli* binds to DNA with a preference for AT-rich sequences. Its multimerization induces the formation of nucleoprotein filaments. By multimerizing, H-NS can also promote DNA looping by bridging DNA segments (Dame et al., 2006, 2000). Other proteins called IHF (integration host factor), HU (histone-like protein) or Fis (factor of inversion stimulation) were found in diverse bacteria. These proteins induce the formation of DNA bents of variable degrees, create DNA supercoiling, form protein filaments, and generate stable plectonemes (Dame, 2005; Dame et al., 2011).

Chromosomal interaction domains

The *E. coli* chromosome is divided into macrodomains of about 1 Mb: left arm, right arm, origin (ORI) and terminus (TER) (Valens et al., 2004). Similarly to the TADs found in interphase eukaryotic cells, bacterial chromosomes are folded into interaction domains (CIDs) of tens of Mbp (Dame et al., 2020). Hi-C experiments carried out in

Caulobacter crescentus (C. *crescentus*) highlighted a functional clustering of housekeeping genes that are associated with a high transcription rate (Le and Laub, 2016). Hi-C data obtained from *E. coli* revealed that its chromosome is partitioned differently at its TER compartment compared to the rest of the chromosome (non-TER regions). In the non-TER regions, concerted action of NAPs and the bacterial SMC complex (called MukBEF in this organism) induce the formation of long-range contacts by forming loops of ~800 kb. By contrast, in the TER region, only short-range interactions were reported (Lioy et al., 2018). CID boundaries are often associated with highly expressed genes that are oriented in opposite directions (as ribosomal RNA coding genes) or coincide with NAP binding sites (Le et al., 2013; Marbouty et al., 2015). Like the eukaryotic condensin complexes, MukBEF is thought to structure the chromosome of *E. coli* via loop extrusion.

IV. Structural Maintenance of Chromosomes (SMC) complexes

As discussed in the previous section, the two protein complexes called cohesin and condensin are key regulators of chromatin architecture throughout the life cycle of prokaryotic and eukaryotic cells. These complexes belong to a family called Structural Maintenance of Chromosomes (SMC) complexes, which are involved in a wide range of functions associated with DNA architecture and metabolism. In this section, I briefly recapitulate the roles of the SMC complexes and describe in details their architectures and their molecular mechanisms.

a. The variety of functions of SMC complexes

Structural maintenance of chromosomes (SMC) complexes are conserved from prokaryotes (ScpAB, MukBEF) to eukaryotes (condensin, cohesin, Smc5/6). They are molecular motors that use the energy provided by the ATP hydrolysis to carry out varying

chromosomal regulatory functions. Prokaryotic SMC complexes promote chromosome arm alignment ensuring a faithful DNA segregation. They also structure the genomes by forming CIDs (Danilova et al., 2007; Le et al., 2013; Marbouty et al., 2015).

In eukaryotes, SMC complexes called condensins compact mitotic chromosomes, promote chromosome individualization and thereby guarantee the even distribution of the genetic material into the daughter cells (Paul et al., 2019). A species-specific condensin version called condensin I 'DC' (dosage compensation) down-regulates the expression of genes associated with sex chromosomes in *Caenorhabditis elegans* (Csankovszki et al., 2009). In vertebrates and in budding yeast, condensin is also responsible for tRNA gene clustering and the formation of the nucleolus (Haeusler et al., 2008).

Cohesin complexes form and maintain TADs during interphase. Doing so, they regulate gene expression by bringing enhancer and promoter regions in proximity (Kagey et al., 2010). In addition, cohesin also promotes sister chromatid cohesion (Guacci et al., 1997; Losada et al., 1998; Michaelis et al., 1997).

Eukaryotic cells have an additional SMC complex, called Smc5/6, which has been associated with DNA damage repair (Menolfi et al., 2015).

b. SMC complexes in human diseases

Cohesin is a key player of gene regulation in human cells. Cohesin mutations have therefore been described in diverse human diseases (for review, see Liu and Krantz, 2008). The most prominent example is a developmental disorder called Cornelia de Lange Syndrome, which is associated with diverse phenotypes, such as facial dysmorphia, extremity malformation or intellectual disabilities. Another example of a developmental disorder, this time resulting from cohesin misregulation as a result of mutations in the cohesin regulator Esco2, is Roberts-SC phocomelia syndrome, which has been associated with impaired intellectual and physical development. The ensemble of diseases associated with

cohesin are called cohesinopathies. Cohesin mutations are also frequently found in a wide range of human cancers, with most mutations found in the STAG2 subunit (Hill et al., 2016). Because chromosomal instability is a characteristic of cancer cells, mutations of condensin subunits in these cells are frequent. In particular, condensin II mutations were reported in T-cell lymphoma (Woodward et al., 2016).

c. Loop extrusion as a unifying molecular mechanism for the SMC complexes

Definition and *in vivo* and *in silico* evidence for DNA loop extrusion

A mechanism called DNA loop extrusion, which relies on the progressive formation of a DNA loop, was proposed to explain the variety of biological functions executed by prokaryotic and eukaryotic SMC complexes (Nasmyth, 2001; Riggs et al., 1990).

The visualization by electron microscopy of DNA loops emerging from a central chromosomal scaffold initially led to the idea of DNA looping as a compaction mechanism (Paulson and Laemmli, 1977). Loop extrusion was later supported by theoretical models (Alipour and Marko, 2012) and by chromosome conformation capture data coupled with polymer simulations (Naumova et al., 2013). In the latter study, metaphase chromosomes were predicted to be constituted of loops of about 80-120 kb. Additional polymer simulations demonstrated that DNA loop extrusion can recapitulate mitotic compaction, segregation and disentanglement (Goloborodko et al., 2016).

Later, Hi-C profile analysis of chicken DT40 cells synchronously entering mitosis, coupled with coarse-grained models and polymer simulations, revealed that mitotic chromosomes are constituted of DNA loops that are helically piled up on top of each other, with condensin locating at the basis of the loops (Gibcus et al., 2018).

Symmetric and asymmetric DNA loop extrusion

Recently, DNA loop extrusion was visualized *in vitro* for different eukaryotic SMC complexes – condensin, cohesin and Smc5/6 – from different species (Davidson et al., 2019;

Ganji et al., 2018; Golfier et al., 2020; Kim et al., 2019; Kim et al., 2022). So far, DNA loop extrusion by prokaryotic SMC complexes could not be directly visualized *in vitro*, but their structural similarities with the Smc5/6 complex (see section below) suggest that they probably also possess DNA loop extrusion activity.

One of the single-molecule setups to monitor DNA loop extrusion *in vitro* is based on the attachment of both extremities of a linear DNA molecule to a passivated surface. This surface, contained into a microfluidic chip that can be injected with buffer solutions, is then imaged by TIRF (Total Internal Reflection Fluorescence) microscopy. Fluorescent labeling of protein and DNA allows the visualization of DNA loop extrusion driven by SMC complexes. These *in vitro* experiments revealed that SMC complexes are sufficient to catalyze DNA loop extrusion. The DNA loop extrusion activity of SMC complexes is abolished when their ATP hydrolysis is inhibited, which demonstrates that the DNA loop formation observed in the single-molecule experiments results from the SMC complex motor activity. In these experiments, different SMC complexes showed similar DNA loop extrusion rates and stall forces. For instance, condensin extruded DNA at a velocity of 50 to 1,500 bp/s, depending on the DNA tension, and hydrolyzed about two ATP molecules per second (Ganji et al., 2018). Thus, condensin – and SMC complexes in general – take unconventionally large steps onto DNA in comparison to translocases and helicases. For instance, EcoR124l and DnaB takes steps of 1-2 bp per hydrolyzed ATP (Rajendran et al., 2000; Seidel et al., 2008).

In vitro loop extrusion assays also shed light on the biophysical properties associated with different types of SMC complexes. Cohesin was reported to loop extrude as a monomer (Davidson et al., 2019) or as a dimer (Kim et al., 2019), whereas Smc5/6 was shown to loop extrude exclusively as a dimer (Pradhan et al., 2022). DNA loops extruded by condensin were generated by monomeric complexes (Ganji et al., 2018). Furthermore, SMC complexes formed DNA loops by progressively extruding a DNA segment from either one or both sides. One-sided extruders were defined as asymmetric, whereas two-

sided extruders were defined as symmetric (**Figure 1.2**). Condensin is an asymmetric loop extruder (Ganji et al., 2018). On the contrary, cohesin was reported to extrude DNA symmetrically (Davidson et al., 2019; Kim et al., 2019). However, a recent *in vitro* loop extrusion study of cohesin (Davidson et al., 2022) performed in presence of the TAD boundary factor CTCF revealed that CTCF induces cohesin to extrude DNA asymmetrically by tethering cohesin to a fixed position on the DNA, likely through the interaction of its N-terminus with the SA2 and Scc1 subunits of cohesin (Li et al., 2020). This result explains how cohesin accumulates at TAD boundaries *in vivo.* In the case of the SMC5/6 complex, symmetric extrusion (Pradhan et al., 2022) could result from dimerization of the complex.

Interestingly, asymmetric DNA loop extrusion failed to recapitulate condensin-driven chromosomal compaction in simulations, leaving gaps in-between DNA loops (Banigan and Mirny, 2020). The observation that condensin molecules form "Z-loops" by bypassing each other in *in vitro* DNA loop extrusion assay provides a explanation how asymmetric loop extrusion could close such gaps (Kim et al., 2020).

Obstacles to DNA loop extrusion

Recently, cohesin was shown to be able to bypass nucleosomes, Cas9, or even larger (~200 nm) artificial obstacles bound to DNA *in vitro*. This suggests that cohesin might also be competent to overcome macromolecular assemblies associated with DNA in a physiological context (e.g. RNA polymerases, poly-nucleosomes, or even entire DNA replisomes) (Pradhan et al., 2021). In cells, in addition to these molecular obstacles, DNA loop extruders have to deal with supercoiled DNA that results from enzymatic activities, such as transcription. Recently, condensin's DNA loop extrusion activity was reported to be stimulated by DNA supercoiling *in vitro*, with plectoneme tips being the preferential binding site of the complex (Kim et al., 2022).

Molecular models for DNA loop extrusion

Over the years, several molecular models that attempt to explain the mechanistic basis of DNA loop extrusion by SMC complexes have been proposed, including tethered-inch-worm, DNA-segment-capture and scrunching-type models (reviewed in Datta et al., 2020; Hassler et al., 2018; Higashi and Uhlmann, 2022). All of these models are based on major ATP-dependent molecular rearrangements in SMC complexes and require at least two independent DNA-binding sites. In the case of an asymmetric extruder (i.e. condensin or cohesin in presence of CTCF), the two DNA-binding sites would have to perform DNA motor or DNA anchor functions, respectively. The motor site feeds the growing DNA loop in an ATP-dependent manner, while the anchor site tethers the complex to a stationary DNA site (**Figure 1.2.**). In this context, the ATP-dependent conformational changes have to guarantee the processivity and the directionality of the process.

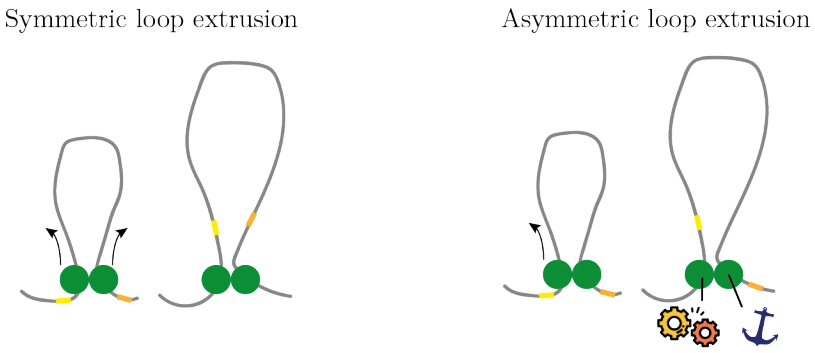


Figure 1.2. Symmetric and asymmetric loop extrusion. A loop extrusion factor with two DNA binding sites (in green) can extrude DNA loop symmetrically (left) or asymmetrically (right). In the case of an asymmetric loop extrusion, the loop extrusion factor must reel-in DNA to form a loop (motor binding site), while stably holding onto the DNA (anchor binding site).

V. Common structural organization of the SMC complexes

The structural similarities between different SMC complexes suggest that they perform their plethora of functions using a common molecular mechanism (**Figure 1.3**). SMC complexes are characterized by two SMC subunits (called ν-Smc and κ-Smc,

according to their interaction with the kleisin) that form either homo- or heterodimers through their 'hinge' dimerization domains. All SMC subunits form intramolecular coiled coils of ~50 nanometers (nm) in length. The SMC dimers sandwich two ATP molecules between their ATPase 'head' domains. These ATP binding sites belong to the ATP-binding cassette (ABC) family, characterized by signature, Walker A (P-loop) and Walker B motifs that are responsible for ATP binding and hydrolysis.

Also shared by all SMC complexes is the presence of a kleisin subunit. These subunits are usually largely unstructured and contact the ν-SMC and κ-SMC via their N- and C-termini, respectively (**Figure 1.3**). The kleisin subunit of MukBEF contains a four-helix bundle at its N-terminus, whereas the kleisin subunits of other SMC complexes contain a helix-turn-helix (HTH) motif. All kleisin subunits contain a winged-helix domain at their C-termini. The binding to different regions of the two subunits of an SMC dimer creates asymmetric complexes, even in the case of Smc subunit homodimers.

Kleisin subunits recruit additional non-SMC subunits called KITE (Kleisin-Interacting Tadem winged-helix Elements) for bacterial SMC and for Smc5/6, and HAWK (HEAT repeat proteins Associated With Kleisin) for cohesin and condensin. Although structurally different, the HAWK and KITE subunits of different SMC complexes share ATPase regulatory and DNA binding functions (Kschonsak et al., 2017; Li et al., 2018; Zabrady et al., 2016).

A

	Human	Budding yeast		Human	Budding yeast
	Condensin			Cohesin	
k-SMC	Smc4	Smc4	k-SMC	Smc1	Smc1
v-SMC	Smc2	Smc2	v-SMC	Smc3	Smc3
g-Kleisin	CAPH	Brn1	a-Kleisin	Rad21	Scc1
HEAT-IA	CAP-D2	Ycs4	HEAT-AI	Nipbl	Scc2
HEAT-IB	CAP-G	Ycg1	HEAT-AII	Pds5A/B	Pds5
b-Kleisin	CAPH2	-	HEAT-B	Stag1/2	Scc3
HEAT-IIA	CAP-D3	-			
HEAT-IIB					

B

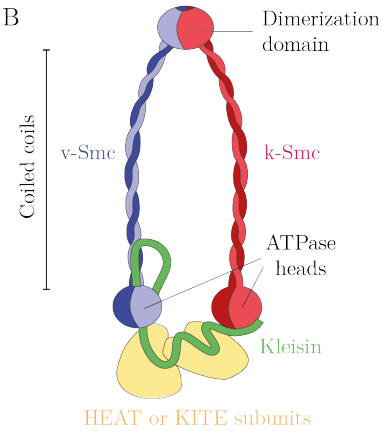


Figure 1.3. Subunit composition of condensin and cohesin and schematic representation of SMC complexes. The subunit composition and nomenclature of condensin and cohesin is given for human and budding yeast (**A**). The schematic organization of SMC complexes is shown (**B**). The complexes are pentamers formed of

two Smc proteins called ν- and κ- (blue and red) according to their interaction with the kleisin (in green). They dimerize at their hinge domain and upon ATP binding to their ATPase head domains. Two HAWK (HEAT-repeat subunits) or KITE subunits are associated with the kleisin.

The ring structure of the SMC complexes condensin and cohesin, created by the Smc and kleisin subunits, was proposed to be the molecular basis of their chromosome compaction and sister chromatid cohesion functions. The "ring model" proposed that condensin and cohesin topologically entrap two DNA strands inside their ring structures (Cuylen et al., 2011; Gruber et al., 2003; Haering et al., 2002). Thus, the embracement of two DNA strands from the same DNA molecule by condensin or cohesin would form a loop, whereas the encirclement of two strands of two different DNA molecules would explain how cohesin can promote sister chromatid cohesion. According to the ring model, entry "gates" are required for the DNA molecule(s) to enter condensin and cohesin rings. One of these gates could result from the dissociation of the kleisin N-terminus from the ν-Smc subunit of condensin and cohesin, driven by ATP binding (Hassler et al., 2019; Muir et al., 2020). Alternative gates could be located at the Smc hinge dimerization interface or at the interface between the C-terminus of the kleisin and the κ-Smc subunit. Opening of any of these interfaces is, however, dispensable for DNA loading and loop extrusion (Davidson et al., 2019; Pradhan et al., 2021; Shaltiel et al., 2022)'. Therefore, the ring model, in which DNA is topologically encircled inside the Smc-kleisin ring of cohesin and condensin, fails to explain DNA loading and loop extrusion. Nevertheless, the ring model might still explain how cohesin mediates sister chromatid cohesion. The exact molecular mechanisms underlying DNA loop extrusion and sister chromatid cohesion still remain to be elucidated.

VI. Chromatin recruitment of the eukaryotic condensin

Condensin has been reported to associate with mitotic chromosomes in all the organisms that have been investigated so far. The chromatin association of condensin I and II in vertebrate cells is differentially regulated at the spatial and temporal level (Hirota et al., 2004). If condensin II localizes into the nucleus in interface and interacts with the chromatin from prophase onwards, condensin I only access compacted chromatin after nuclear envelope breakdown (NEBD) in prometaphase. Only a single condensin complex is present in both fission and budding yeast model organisms, and it associates with chromatin at the beginning of mitosis (Sutani et al., 1999). Based on sequence comparison, the condensin complex found in yeast species is closer related to vertebrate condensin I than condensin II.

Preferential binding of both condensin I and II was mapped at telomeric and centromeric regions, at tRNA and rRNA gene repeats and at highly transcribed genes regions on the genomes of cells of various types (Kim et al., 2013). In particular, nucleosome-depleted regions (NDR) at the transcription start sites (TSS) are enriched in condensin binding, even though transcriptional activity strongly decreases during mitosis (Palozola et al., 2019), suggesting that they constitute preferential loading sites. However, nucleosome depletion cannot be sufficient to promote condensin loading, since many NDRs were reported to be devoid of condensin in fission yeast (Toselli-Mollereau et al., 2016). Other parameters that link condensin loading to transcription, like the formation of R-loops or the creation of transcription-dependent DNA topological stress, were suggested to play a role in condensin recruitment to chromatin (Bernard and Vanoosthuyse, 2015). The RSC complex (Remodels the Structure of Chromatin), a molecular machine that is responsible for the eviction of nucleosomes from chromatin and therefore increases DNA accessibility, was reported to directly interact with cohesin (Muñoz et al., 2020) and to assists condensin chromosomal loading in fission yeast by acting in concert with the

histone acetyltransferase Gcn5 (Toselli-Mollereau et al., 2016). These data suggest that nucleosomes constitute a barrier for condensin recruitment onto chromatin. In contrast, direct interactions between histones and condensin I and II were reported. For example, the C-terminal extremity of HEAT-repeat subunit CAP-D2^{Ycs4} of human condensin I was suggested to contain an interaction domain responsible for histone H1 and H3 recognition (Ball, et al., 2002). The kleisin subunit CAPH of human condensin I was reported to interact with histones H2A and H2A.Z tails (Tada et al., 2011) in a phosphorylation-dependent manner. Moreover, human condensin II HEAT-repeat subunit CAP-D3 was reported to bind to the epigenetic mark H4K20me (Liu et al., 2010). Although these interactions likely facilitate the recruitment and/or the loading of condensin complexes onto chromatin, they are very unlikely to provide the core molecular mechanism that underlies chromosome loading, given that prokaryotic SMC complexes have evolved prior to histones and that eukaryotic condensins preferentially bind to NDR.

Several other chromatin components were suggested to facilitate condensin loading onto chromatin by directly interacting with condensin (reviewed in Piazza et al., 2013). For example, mapping of condensin binding onto the budding yeast genome revealed a condensin enrichment at RNA pol III-transcribed regions (5S rRNA and tRNA-coding genes). General transcription factors of RNA pol III transcription were found to co-localize to these sites (D'Ambrosio et al., 2008). In yeast, tRNA gene clustering is condensin-dependent, and condensin co-immunoprecipitated with RNA pol III core transcription factors TFIIIB and TFIIIC (Haeusler et al., 2018). Furthermore, in human embryonic kidney (HEK) 293T cells, TFIIIC co-immunoprecipitated with condensin II (Yuen et al., 2017).

VII. The architecture of the budding yeast condensin complex

In this section, I focus on the condensin complex from the budding yeast *Saccharomyces cerevisiae* (*Sc)* and on available structural data for the complex. I introduce the motor and anchor functions and nominate potential candidate sites that could bring about these functions.

a. Introduction to the motor and anchor functions of condensin

Condensin is composed of Smc2 and Smc4 subunits, two HEAT-repeat subunits (called Ycg1 and Ycs4 in *Sc*[1]) and a kleisin subunit (Brn1 in *Sc*[1]). As discussed in the previous sections, condensin is a one-side DNA loop extruder, forming loops in an asymmetric fashion, whereas cohesin and Smc5/6 extrude loops symmetrically (Davidson et al., 2019; Ganji et al., 2018; Kim et al., 2019; Pradhan et al., 2022). This behavior can best be explained by the existence of two condensin functional sites: (i) an anchor site that tethers condensin to its initial DNA binding location and (ii) a motor site from which the DNA is progressively looped (**Figure 1.2**).

b. Ycg1-Brn1: a candidate for the DNA anchor site in condensin

The HEAT-repeat subunit Ycg1 is arranged in a solenoid shape (Kschonsak et al., 2017). It interacts with the carboxyterminal part of the Brn1 kleisin subunit. Co-crystal structures of DNA-bound to Ycg1-Brn1 sub-complex revealed that Ycg1 accommodates DNA in its positively charged concave surface, while Brn1 encircles the DNA, forming a so-called 'safety-belt' peptide loop that locks the DNA into place (**Figure 1.4B**). Loosening

[1] For clarity, the *Saccharomyces cerevisiae* (*Sc)* nomenclature will exclusively be used for the rest of the thesis

of the safety belt by mutating DNA-interacting residues caused condensin to slip while it extruded DNA loops in *in vitro* assay (Ganji et al., 2018), which suggests that the condensin module formed by Ycg1 and Brn1 plays a DNA anchor role (Kschonsak et al., 2017). In addition, condensin recruitment to human chromosomes is abolished when closure of the safety belt is impaired by mutations that destabilize the interaction of the "latch" and "buckle" Brn1 interfaces.

c. Candidate sites for the condensin DNA motor function

Holo-complex structures of condensin and other SMC complexes

Cryo-EM structures of the *Sc* condensin holo-complex were obtained in ATP-bound and ATP-free states (Lee et al., 2020) (**Figures 1.4 A, C, D**). A striking unexpected feature of the ATP-free holo-complex is that the coiled coils are aligned in a rod conformation and bend over at a flexure point called "elbow", which had previously been described for the MukBEF complex (Bürmann et al., 2019). To conformations of the ATPase head domains were resolved for the ATP-free state: In one, Ycs4 associates only with Smc4 (pdb: 6YVU), whereas in the other conformation, Ycs4 bridges the Smc2 and Smc4 ATPase head domains (pdb: 6YVV) and keeps them apart (Lee et al., 2020). This study revealed that ATP binding and head engagement lead to coiled-coil opening, with a variety of opening degrees observed among different molecules. The comparison of the structures of condensin ATP-free and ATP-bound states revealed a mutually exclusive interaction of the Ycs4 and Ycg1 HEAT-repeat subunits with the Smc heads (Lee et al., 2020). ATP binding triggers Ycs4 replacement by Ycg1 at the ATPase head interface, referred to as a "flip-flop" mechanism. As a result, only one HEAT-repeat subunit binds to the ATPase head in a given state, while the other subunit is only tethered to the complex through its kleisin interaction.

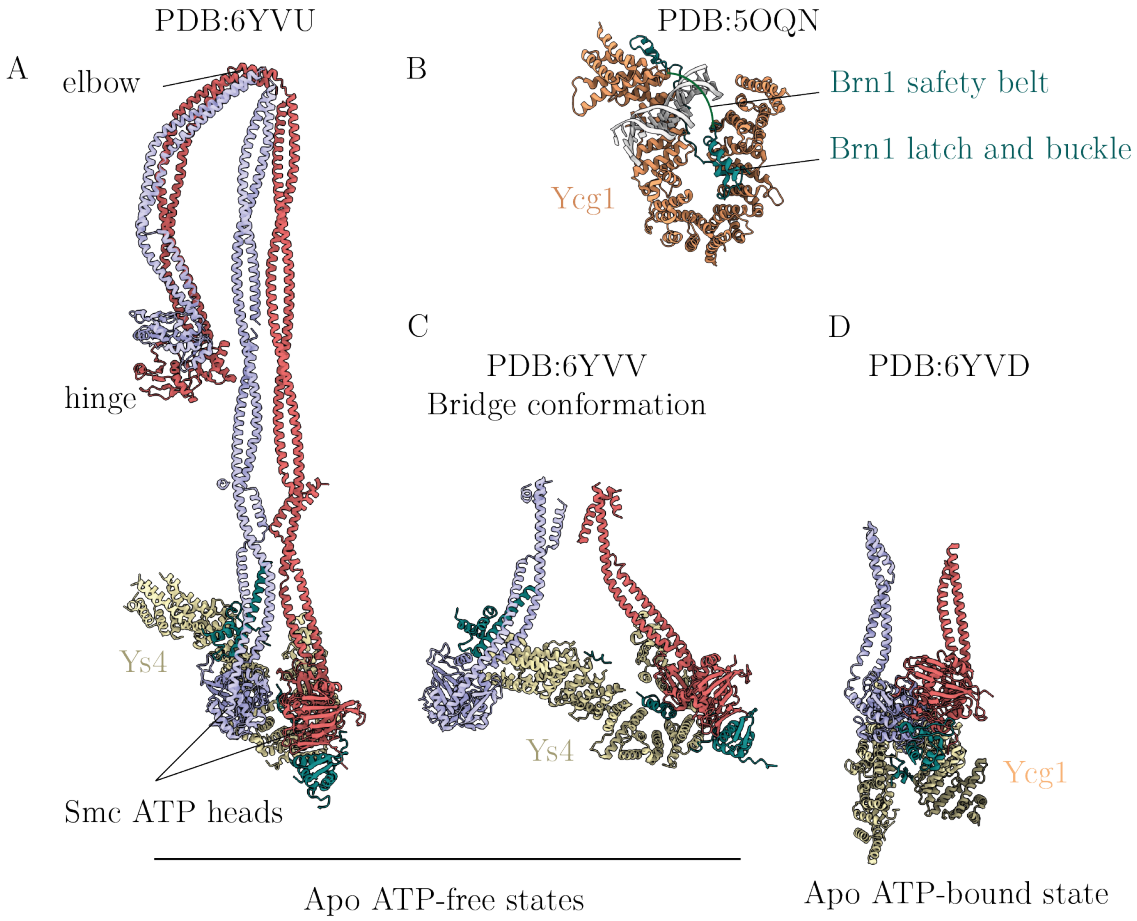


Figure 1.4. Structures of the *Sc* condensin complex. Cryo-EM structures of the *Sc* condensin holo-complex in the apo ATP-free states (**A** and **C**) and in the apo ATP-bound state (**D**). In absence of ATP, *Sc* condensin holo-complex adopts two conformations where the coiled coils align and bent at the elbow: in the first one, the Ycs4 HEAT-repeat subunits binds to the closed ATPase heads (A), whereas it spreads them apart in a second conformation called "bridge" (**C**). When ATP binds, the Smc heads engage and the coiled coils presumably open (D). In that case, the Ycg1-HEAT repeat subunit binds to the ATPase heads.
Crystal structure of the *Sc* condensin subcomplex formed by Ycg1-HEAT repeat, by Brn1 and DNA (**B**). The Brn1 kleisin subunit forms a "safety-belt" peptide loop around the DNA. The safety belt closes at an interface formed by the Brn1 "latch" and "buckle".

Comparison with other SMC complexes

Because the DNA substrate was absent from the holo-complex cryo-EM structures of *Sc* condensin, these structures fell short in explaining the DNA motor function required for DNA loop extrusion. However, structures of other SMC complexes suggest that the condensin 'core' region, composed of the Smc2/4 heads, the Ycs4 HEAT-repeat subunit and the N-terminal extremity of the kleisin subunit, might constitute a DNA binding site with a putative DNA motor function.

Cryo-EM structures of the SMC-related Rad50-Mre11 complex (Rojowska et al., 2014) revealed that DNA binds at positively charged patches found at the top of the dimerized ATPase heads, which are also present in condensin. More recently and during the course of my thesis work, cryo-EM structures were published that revealed the molecular details of the interaction between DNA and the cohesin complex from different species (Collier et al., 2020; Higashi et al., 2020; Shi et al., 2020). In these structures, DNA inserts in-between the open Smc coiled coils and binds on the top of the ATPase heads. DNA also contacts to the N-terminal extremity of the HEAT-repeat subunit Scc2, the homolog of condensin Ycs4. In addition, the N-terminus of the kleisin Scc1, homologous to condensin Brn1, loops around the DNA. Recent DNA-bound structures of MukBEF revealed a similar conformation (Bürmann et al., 2021). These structures thus suggest a shared DNA binding mode among the SMC complexes (see **Discussion**, **Figure 6.3**).

As discussed above, the ATP-bound and ATP-free *Sc* condensin structures resolved in the absence of DNA show that the core module is highly dynamic and that its conformation depends on the nucleotide state. In addition, ATP binding induces massive rearrangements of the long coiled coils, which transition from a folded rod shape to an open and flexible conformation. Interestingly, coiled-coil opening can regulate the accessibility of the positively charged patches on the condensin ATPase heads that have been shown to bind DNA in other SMC complexes. In addition, the opening of the coiled coils can promote access to the inner surface of the hinge domain (**Figure 1.4A**), which had previously been found to bind DNA (Griese et al., 2010). For all these reasons, the core region is a good candidate for a condensin motor site.

VIII. Cryo-Electron Microscopy as a technique of choice to study SMC complexes

In the previous section, I discussed putative functional roles of condensin modules inferred from numerous structural studies. Cryo-EM has proven to be the technique of

choice to study the structure of the large and dynamic SMC complexes. In this section, I introduce this technique and its advantages by giving a brief overview of its evolution over the past few years.

The past ten years have seen the 'resolution revolution' of cryo-EM. It became a widely applicable approach for the structural characterization of macro-molecular complexes. It has the advantage over X-ray protein crystallography that is does not require the formation of protein crystals. Furthermore, it allows the structural determination of large macromolecules (e.g. viral particles) as well as proteins below one hundred kilodaltons. The resolution revolution in the field of cryo-EM results from several improvements and developments over the years: sample preparation techniques, advent of direct electron detectors (DED) in conjunction with other technical improvements of electron microscopes and powerful analysis software.

In the early days of electron microscopy, biological samples were embedded in heavy metals in order to improve their contrast. As a consequence, only low-resolution structures of their envelopes could be obtained. In order to increase the resolution while preserving biological ultrastructure, cryo-genic sample vitrification was developed in the 1980s (Dubochet et al., 1988), with a technique known as cryo-EM. In cryo-EM, the particles or interest are entrapped in a thin vitreous ice layer by fast cooling of the aqueous solution in liquid ethane, which prevents the formation of ice crystals and makes the layer more transparent to the electron beam. In addition to improvements in sample preparation, imaging acquisition techniques were developed to limit the effect of beam induced motion and damage. DED enable the rapid acquisition of individual movie frames, which can be aligned in order to correct for beam-induced motions and filtered to limit the effect of beam-induced damage (Glaeser, 2016; Li et al., 2013). On top of that, advancements in software development allow structural reconstruction by averaging thousands of experimental 2D projections (i.e. the sample particles) in order to

increase the low signal-to-noise ratio (SNR) of macromolecules that are composed of low-atomic number elements. Two of these innovative software packages are RELION (REgularised LIkelihood Optimisation) (Scheres, 2012) and CryoSPARC (Cryo-EM Single Particle Ab-Initio Reconstruction and Classification) (Punjani et al., 2017). RELION uses an expectation maximization approach, whereas CryoSPARC is partially based on a gradient descent algorithm. Both approaches aim to accurately assign projection angles of particles from 2D micrographs in order to obtain a 3D reconstruction. Particles can be classified according to their structural heterogeneity. However, the motions of macromolecular complexes are often continuous, adopting a wide range of conformations. In order to account for this type of conformational heterogeneity which cannot be described by a small number of discrete states, other approaches based on neural network algorithms were recently implemented for cryo-EM image analysis (Zhong et al., 2021). Neural networks were also recently used to improve particle picking from crowed and heterogeneous micrographs (Bepler et al., 2020).

Over the past few years, tremendous improvements in the field of cryo-EM allowed to resolve macromolecular structures at atomic resolutions (Wang, 2022) and to characterize the conformation of heterogenous samples (Jonić, 2016). With the recent development of cryo-electron tomography, which provides then means to resolve the structure of biomolecular complexes in their native cellular context, the coming years promise to be exiting for structural biology.

Aims and scope of this thesis

Accumulating data suggests that the formation of mitotic chromosomes depends on the DNA loop extrusion activity of condensin complexes. However, the molecular details of how condensin binds DNA and extrudes loops are unknown, largely due to a lack of structural information on DNA-bound condensin complexes. The aim of my PhD thesis work was to structurally and biochemically characterize the interaction of condensin with DNA, at different stages of its ATPase cycle, in order to reveal the molecular basis of loop extrusion.

I chose cryo-EM as an approach, as it has proven to be successful to resolve the structures of other DNA-bound SMC complexes. In order to perform structural studies of DNA-bound condensin complexes by cryo-EM, I initially optimized the biochemical reconstitution as well as the cryo-EM sample preparation of the complex (see **Chapter 1**).

Firstly, I aimed to identify the condensin anchor and motor sites that are required for condensin to reel DNA into a loop while stably holding onto the DNA. For this purpose, I solved the structure of the *Saccharomyces cerevisiae* (*Sc*) condensin bound to DNA and to ATP at a resolution of 3.4 to 3.9 Å. This allowed me to identify two independent DNA-bound modules, which I named "core" and "periphery" associated with the motor and anchor functions of condensin, respectively (see **Chapter 2**).

Secondly, I aimed to provide the structural basis of the motor functions of condensin core. To do so, I solved the structure of the ATP-free DNA-bound *Sc* condensin core at a resolution of ~10 Å (see **Chapter 3**). My structural work identified a potential "power-stroke" motor mechanism for condensin. The integration of this structural information with biochemical and single-molecule microscopy data, obtained by Dr. I. A. Shaltiel, led to a detailed loop extrusion model mechanism for condensin (see **Discussion**).

Thirdly, to interrogate the molecular mechanism associated with condensin loading onto the DNA, I obtained a preliminary structure of the *Chaetomnium thermophilum* (*Ct*)

condensin resolved at ~8 Å. It revealed a previously uncharacterized conformation of the complex (see **Chapter 3**).

Finally, I initiated three new collaborative research directions, working closely with Dr. Thomas Dahlet, Dana Grozavu and Wolfram Heinrich Seifert Davila, to interrogate the DNA sequence specificity of condensin binding via a SELEX experiment, to biochemically and structurally characterize binding of condensin to the transcription factor TFIIIC, and to assay condensin binding to a di-nucleosome substrate (see **Chapter 4**). These experiments aim to further understand how condensin functions in the context of chromatin.

Results

Chapter 1

Technical considerations on condensin sample preparations and 3D reconstruction pipelines

I. Protein purification of condensin complexes

a. *S. cerevisiae* (*Sc*) holo-complexes

Condensin holo-complexes were over-expressed in *S. cerevisiae* from galactose-inducible promoters on 2-microns high-copy plasmids and purified as previously described (Terakawa et *al.*, 2017). Briefly, the protocol is divided into three steps: affinity purification via a poly-histidine tag fused to Brn1, affinity purification via a Strep-tag fuse to Smc4, and size exclusion chromatography. The single peak in the size exclusion chromatography profile and SDS-PAGE analysis suggest that the condensin complex obtained via this strategy is pure and stochiometric (**Figure 2.1A**).

b. *C. thermophilum* (*Ct*) holo-complexes, Δ Ycg1 and 'bonsai' complexes

Ct condensin holo-complexes, ΔYcg1 tetrameric complexes and 'bonsai' complexes were generated in Sf21 cells using the Multibac expression system (Sari et al., 2016). The 'bonsai' construct lacks the majority of the Smc coiled coils and the Ycg1 HEAT-repeat subunit, and it contains a shorter version of the Brn1 kleisin (see methods and dedicated results section for details). Briefly, the purification consists of three steps: (i) an affinity purification via a polyhistidine-tag fused to the Brn1 subunit, (ii) an ion exchange chromatography followed by (iii) a size exclusion chromatography. The

chromatograms and the SDS-gels corresponding to the purified complexes are shown in **Figures 2.1.B, C** and **D**.

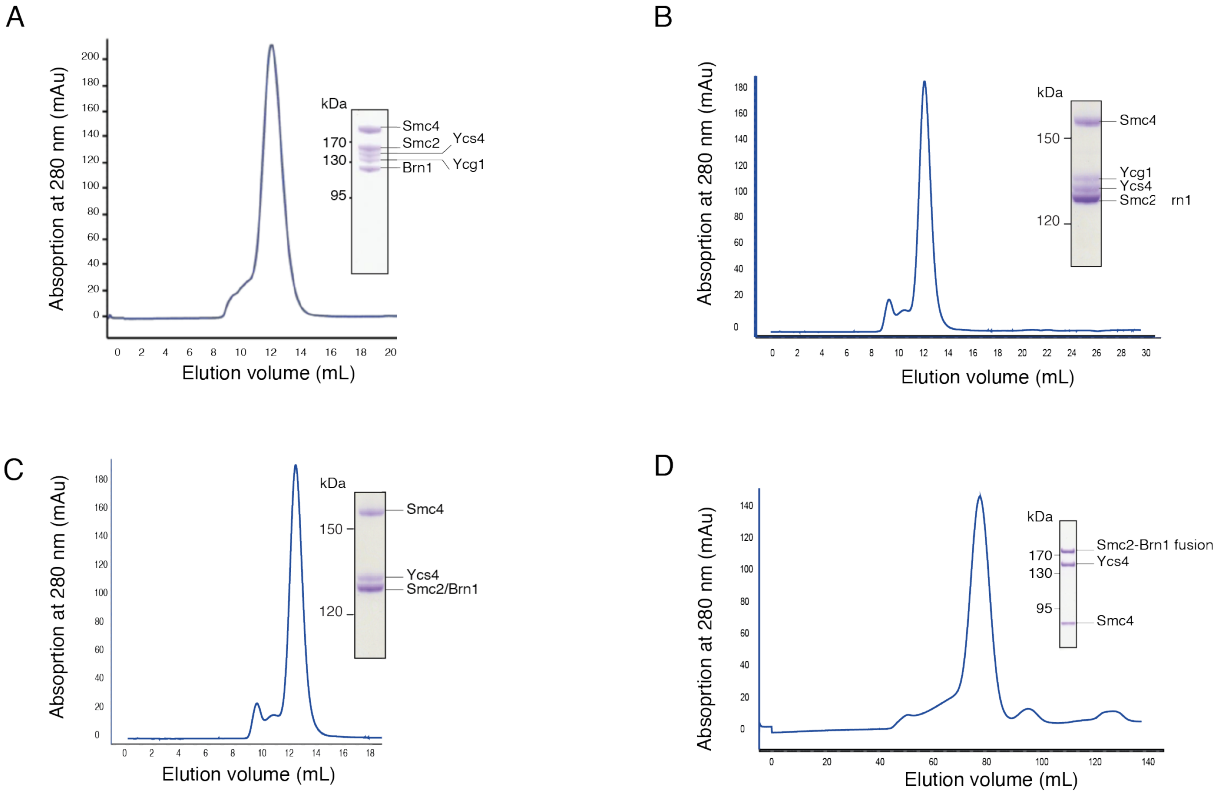


Figure 2.1. Condensin protein complex samples. Gel filtration profiles and corresponding SDS-PAGE gels of the *Sc* holo-complex (**A**), *Ct* holo-complex (**B**), *Ct* ΔYcg1 complex (**C**), *Ct* "bonsai" complex (**D**). The protein constructs were purified by Sol Bravo in the laboratory of Pr. C. H. Haering.

II. DNA-binding activity of condensin complexes

To assess the DNA binding activity of *Sc* condensin holo-complexes using gel filtration, *Sc* condensin complexes (5 mM) were mixed with 75-bp DNA duplexes (2 mM) labelled with a 6-FAM dye (6-carboxyfluorescine) and with ADP (1 mM), $AlCl_3$ (0.5 mM) and NaF (10 mM) where indicated (**Figure 2.2**). ADP-AlF_4 was used to capture the ATP-hydrolysis transition state in which the Smc2-Smc4 heads are engaged (Malär et al., 2021). In similar buffer and concentration, the experiments in presence and absence of ADP-AlF_4 showed the formation of DNA-bound condensin complexes. However, in the presence of ADP-AlF_4, condensin bound DNA more stably as judged by the higher

propension of free-DNA in the absence of ADP-AlF_4. I therefore decided to capture condensin in an engaged ATP-bound state for structural studies of DNA binding (see below).

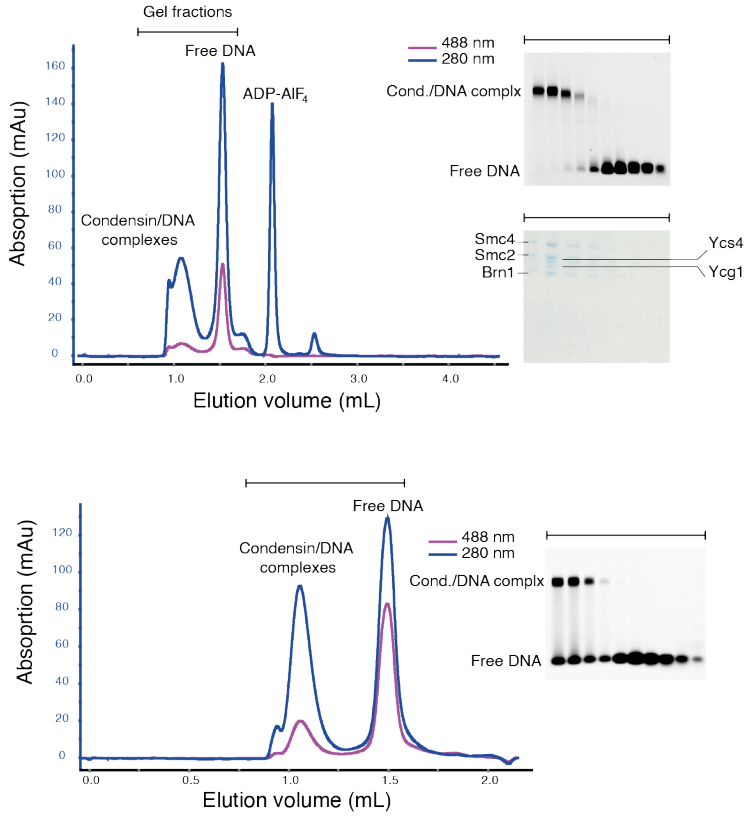


Figure 2.2. Analytical gel filtration experiments testing condensin DNA binding activity. Gel filtration profiles showing the complex formation between Sc condensin and DNA, obtained using the Superose inc. 3.2/300 column, in presence (top) or absence (bottom) of ADP-AlF4. The corresponding agarose gels visualize the free DNA and the condensin-DNA complexes. The chromatogram blue curves correspond to the 280-nm signal. The 488-nm signal, probing the 6-FAM labels of the DNA, is shown in pink. These experiments demonstrate that the Smc head engagement induced by the ADP-AlF4 stabilizes condensin DNA-binding.

In parallel, Dr. Sumanjit Datta performed electrophoretic mobility shift assays (EMSA), to characterize the interaction between *Ct* and *Sc* condensin complexes and DNA (**Figure 2.3**). To this end, 50-bp DNA (5 nM) was mixed with the indicated concentrations of *Ct* or *Sc* condensin holo-complexes, in presence or absence of ATP (1 mM). Interestingly, these experiments demonstrate that the DNA binding activity of *Sc* and *Ct* condensin is only mildly affected by the presence of ATP. The deletion of the Ycg1 HEAT-repeat

subunit had only a minor effect on the DNA-binding activity of *Ct* condensin DNA binding, whereas it strongly diminished the DNA binding activity of *Sc* condensin.

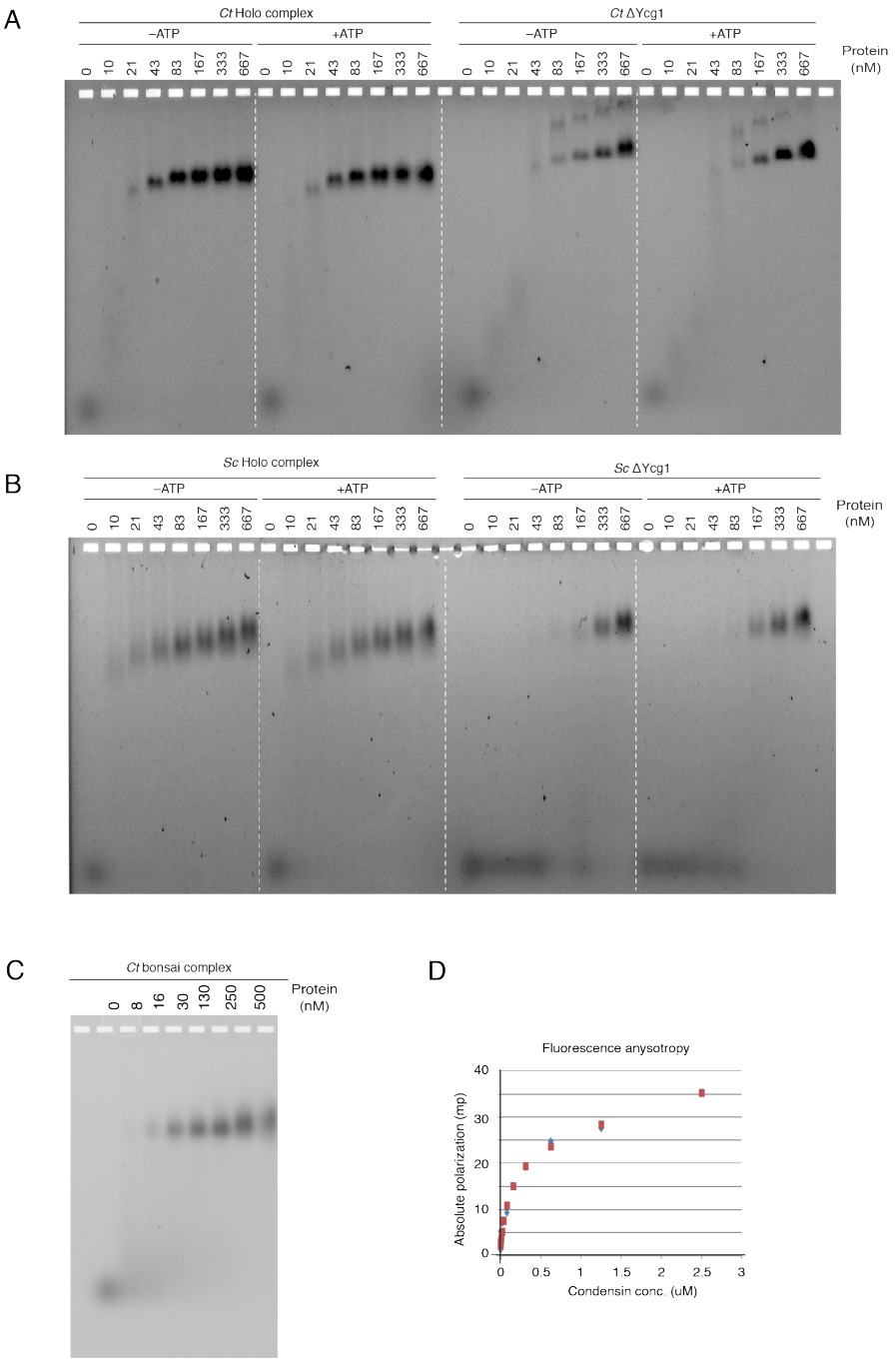


Figure 2.3. Electrophoretic mobility Shift Assay (EMSA) and Fluorescence Anisotropy of condensin complexes. EMSA experiments probing the DNA-binding activity of the *Ct* condensin holo-complex and *Ct* ΔYcg1 complex with and without ATP (**A**) or *Sc* condensin holo-complex and *Sc* ΔYcg1 complex with and without ATP (**B**), using a 6-FAM labelled 50-bp DNA duplex. These EMSA experiments were done by Dr. Sumanjit Data in Pr. C. H. Haering lab. EMSA (**C**) and fluorescence anisotropy (**D**) experiments using a 6-FAM labelled 30-bp DNA duplex and probing the DNA-binding activity of the *Ct* condensin "bonsai"

complex without ATP. These experiments were performed by Dr. Markus Hassler in the laboratory of Pr. C. H. Haering. The panels A and B are adapted from Shaltiel et al., 2022.

To quantify the DNA binding affinity of the *Ct* "bonsai" condensin, fluorescence anisotropy experiments using a 6-FAM labelled 30-bp DNA were performed by Dr. M. Hassler (**Figure 2.3.D.**). The measured K_D of ~130 nM shows that the DNA binding activity of the *Ct* "bonsai" condensin is comparable to that of the holo-complex, despite the deletion of the coiled coils and in the absence of the Ycg1 HEAT-repeat subunit. This experiment suggests that the Smc head domains, together with a minimal version of the kleisin and the Ycs4 HEAT-repeat subunit, are sufficient to support the condensin DNA binding affinity.

III. Cryo-EM sample preparation: screening and optimization

In this section, I describe the key parameters that I identified during the cryo-EM screening procedure in order to obtain high-resolution condensin structures. I screened numerous sample compositions and cryo-grid types, using a Talos Artica 200-kV microscope equipped with a Falcon 3 direct electron detector. The sample preparation conditions for each structure that I describe in this thesis are detailed in the methods or in the result sections.

During the screening process, I assessed data quality by performing two-dimensional (2D) classification (subsequently referred to as "diagnostic 2D classification"). In order to minimize the computational costs, I performed diagnostic 2D classifications using particle sets obtained from a subset of ~1,000 micrographs.

The first challenge in obtaining structures of condensin resided in the sample preparation. In the course of the screening procedure, I identified several key parameters that influenced sample quality, even though individual cryo-grids were not highly reproducible. First, I selected the best cryo-grid support material to maintain condensin integrity. Using copper as a grid support material instead of gold resulted in disintegrated

condensin particles. Second, because condensin preferentially interacted with the grid support, the protein concentration and the wait time before cryo-grid blotting were important parameters in order to condensin to locate into the cryo-grid holes after saturating the grid support. Maintaining low salt concentrations (< 80 mM NaCl) was also essential to stabilize condensin-DNA complexes. I reproducibly observed a non-homogeneous spreading of the sample over the cryo-grid squares, which I attributed to condensin's preferential interaction with the grid support and the low salt concentration. The non-even spreading prevented the formation of the thin layer of vitreous ice required for the acquisition of high-quality data. To tackle this issue, I used grids with resized spacing in-between the holes (2 μm holes spaced by 1 μm, or "R2/1" Quantifoil grids) and applied detergent (Octyl-beta-Glucopyranosid) to some of the samples. As an additional strategy, I used PEGylated gold grids to prevent the particles from interacting with the support (Zhang et al., 2021) in the case of the *Sc* ATP-free DNA-bound condensin structure (see **Chapter 3**).

I also tested different DNA substrate lengths (30, 50, 75, 150 bp). Whereas 50 and 75-bp double stranded DNA were suitable to obtain stable condensin-DNA complexes, the shorter (30-bp DNA) and longer (150-bp DNA) substrates were not. The protein:DNA ratio mattered, especially when I used condensin at a low concentration (0.3 μM). In this case, an excess of DNA resulted in condensin being absent from the cryo-grid holes, which were then only filled with vitrified free DNA. Using gel filtration experiments, I showed that Smc2-Smc4 head engagement stabilizes condensin DNA binding (**Figure 2.2**). Therefore, I used the "double Walker-B EQ" ($Smc4_{E1352Q}$, $Smc2_{E1113Q}$) mutations at both ATPase sites to stabilize the engaged *Sc* condensin complex. The conserved glutamic acid of the Walker-B motif has a key function in the hydrolysis of the ATP gamma-phosphate (Hassler et al., 2019). Its mutation to glutamine strongly diminishes ATP hydrolysis, while preserving ATP binding.

In addition to the *Sc* condensin holo-complex, I used *Ct* condensin deletion mutants. In particular, I used a "bonsai" version of condensin with shorter coiled coils, with the expectation that it might help to decrease the affinity of condensin for the grid support and thereby increase the ice quality, and also facilitate particle alignment during data processing (see below). Because I observed that the *Ct* bonsai construct led to a higher ice quality on a cryo-grid than the *Sc* holo-complex, I decided to also use the *Ct* holo-complex and a *Ct* condensin version lacking the Ycg1 HEAT-repeat subunit for structural cryo-EM studies. The advantages and limitations of these approaches are discussed in the corresponding sections (See **Chapter 3**).

IV. Strategies to analyze structural heterogeneity

The second challenge to obtain condensin 3D structures is the sample heterogeneity. To illustrate this point, I take the example of the DNA- and ATP-bound condensin data set. The same reasoning was employed for all data sets.

a. Picking strategies and particle sorting

Most micrographs showed a dense distribution of heterogeneous particles, which made a manual picking approach very challenging (**Figure 3.1**, **Chapter 2**). Instead, I used an automated approach based on the WARP neural-network algorithm (Tegunov and Cramer, 2019) as a first picking step. This allowed me to filter out possible contaminations and aggregated particles. I performed diagnostic 2D classification to assess particle heterogeneity and pooled particles into group types. The different classes showed diverse particles, views and orientations, but nevertheless allowed me to remove obvious artefacts from the particle sets. In order to obtain a larger set of particles of interest, ultimately required to determine cryo-EM densities at higher resolution, I used two other

picking strategies. I trained the neural network Topaz (Bepler et al., 2020) for picking, using a clearly identified particle set used to generate a 3D *ab initio* model, which I then used for 3D classification. This step drastically increased the number of particles. As an example, in the case of the "core" of the ATP- and DNA-bound condensin data set (see **Chapter 2**), the Topaz particle training set of 91,522 particles led to a picked set of 251,513 particles subsequently used for 3D classifications. I then used a third picking strategy based on 3D cross-correlation. Using a refined 3D model of the core as a template, this approach allowed the specific picking of 1,203,041 particles. These approaches helped me to drastically increase the resolution of the core going from ~ 5-6 Å to 3.5 Å.

b. 3D classification-based strategies

As an additional way to deal with sample heterogeneity, I performed extensive 3D classifications of the particles, with and without performing particle alignment. In the latter case, I kept the Euler angles, determined during a prior refinement step, fixed. This allowed a better sorting of particle heterogeneity and better visualization of detailed molecular rearrangements.

In addition to molecular heterogeneity, the 50-nm long coiled coils of condensin represented a challenge for performing 2D and 3D reconstitutions. Regardless of whether the long coiled-coil regions of condensin align into a rod or adopt an open conformation, they form linear structures that generate strong signals for the alignment algorithms and hence can be a possible source of bias by generating preferentially assigned particle orientations. To circumvent this issue, I performed particle centering and masking to partially exclude the coils while performing 3D classifications. I also attempted to automatically subtract the signal of the flexible coiled coil regions of the ATP-free DNA-bound *Sc* condensin particles.

Chapter 2

DNA- and ATP- bound structure of *Sc* condensin

I. Determination of the DNA- and ATP-bound condensin structure and description of its overall architecture

In order to obtain the structure of the ATP-engaged condensin complex bound to DNA, I applied equimolar ratios of *Sc* condensin double EQ mutant (1 mM, 0.65 mg/mL) and 50 base pair (bp) long double-stranded (ds) DNA onto cryo-grids (mesh 200 Au R2/2, Quantifoil) in the presence of ATP (1 mM). To saturate the grid support, I used a wait time of 120 sec before blotting using the FEI Vitrobot Mark IV. I then acquired a data set of 6,544 micrographs on a Titan Krios 300 kV microscope equipped with a K2 direct electron detector.

As discussed above and as visible on a raw micrograph, the sample displayed a large degree of heterogeneity (**Figure 3.1A**). I sorted particles into four categories: (i) C1 and (ii) C2 (representing core conformations 1 and 2), (iii) Y showing the HEAT-repeat subunits Ycg1 and/or Ycs4 and (iv) classes presumably corresponding to artefacts and/or smaller groups of particles that might correspond to the Smc hinge domain (**Figure 3.1B**). C2 classes were reminiscent of the apo ATP-free condensin 2D classes previously reported (Lee et *al.*, 2020). Therefore, I used C1 and Y classes only to generate 3D *ab initio* models of the core and of the Ycg1 peripheral complex. The resulting analyses are based on a series of 3D classifications with or without performing particle alignment and several rounds of refinements using RELION (Scheres, 2012) and cryoSPARC (Punjani et al., 2017) (**Figures 3.2. and 3.3.,** see Methods for further details).

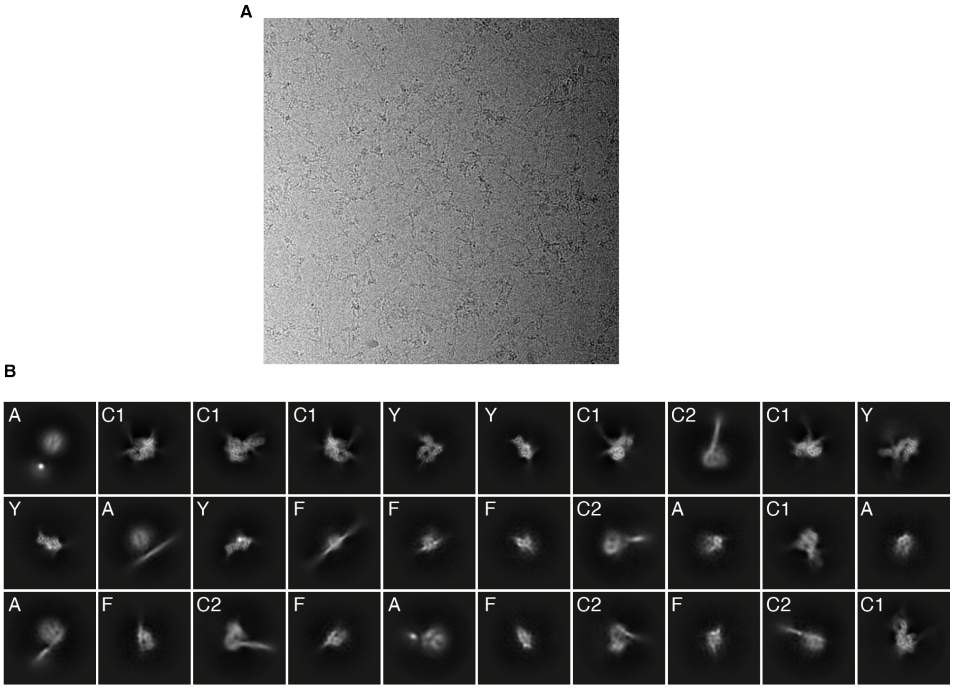


Figure 3.1. ATP- and DNA-bound condensin cryo-EM sample. Example of micrograph showing condensin particles (**A**). Diagnostic 2D classification of the ATP- and DNA-bound condensin cryo-EM data set showing the heterogeneity of the sample (**B**). The particles are classified into different types: C1 for core 1 and C2 for core 2, respectively correspond to the core condensin complexes in the ATP and DNA-bound state or in the ATP-free state. The Y-type of particles correspond to Ycs4 and/or Ycg1 HEAT-repeat subunits, the F-type to protein fragments that possibly show coiled-coils and hinge domain, and the A-type to cryo-EM artefacts. This figure is adapted from Shaltiel et *al.*, 2022.

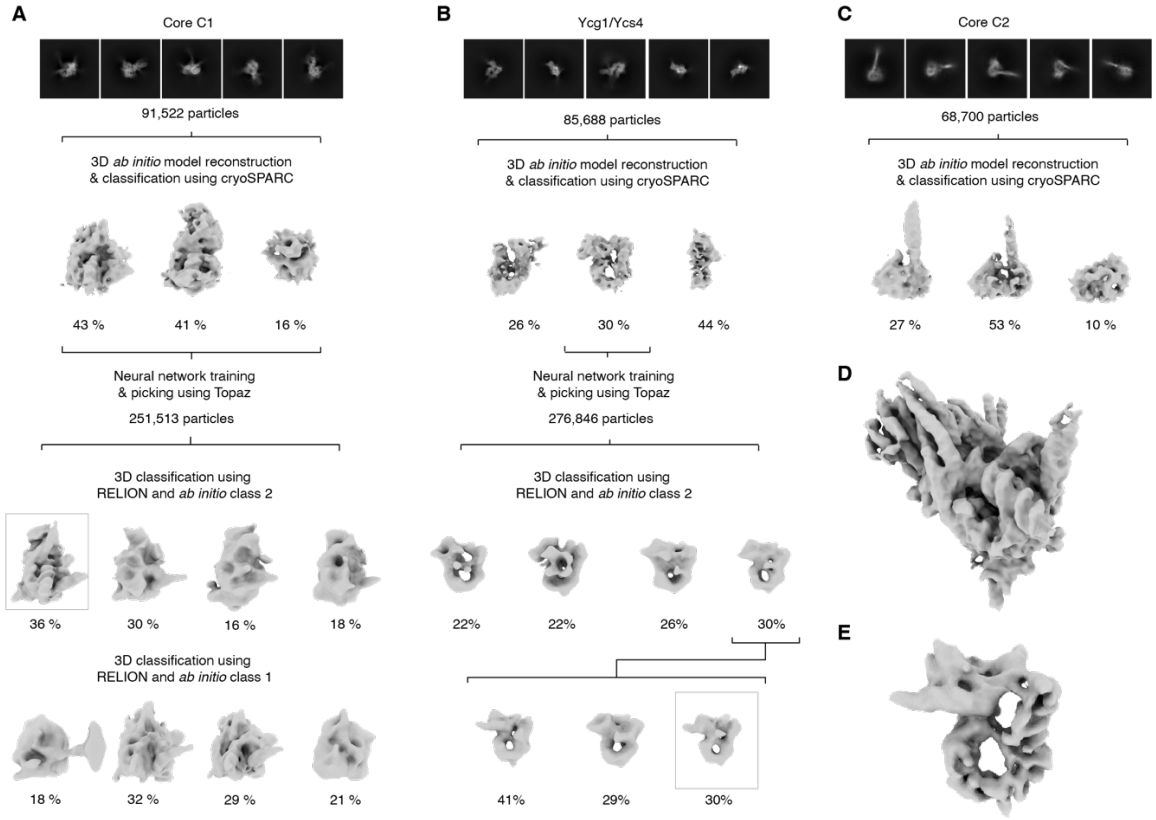


Figure 3.2. Beginning of the analysis pipeline of the ATP- and DNA-bound condensin cryo-EM data set. Calculation of 3D *ab initio* model and subsequent 3D classifications of the particles types corresponding to core 1 (**A**), Ycg1/Ycs4 (**B**) and core 2 (**C**). 3D model of the core 1 condensin complex (**D**) and of the Ycg1 peripheral complex (**E**). Figure adapted from Shaltiel et *al.*, 2022.

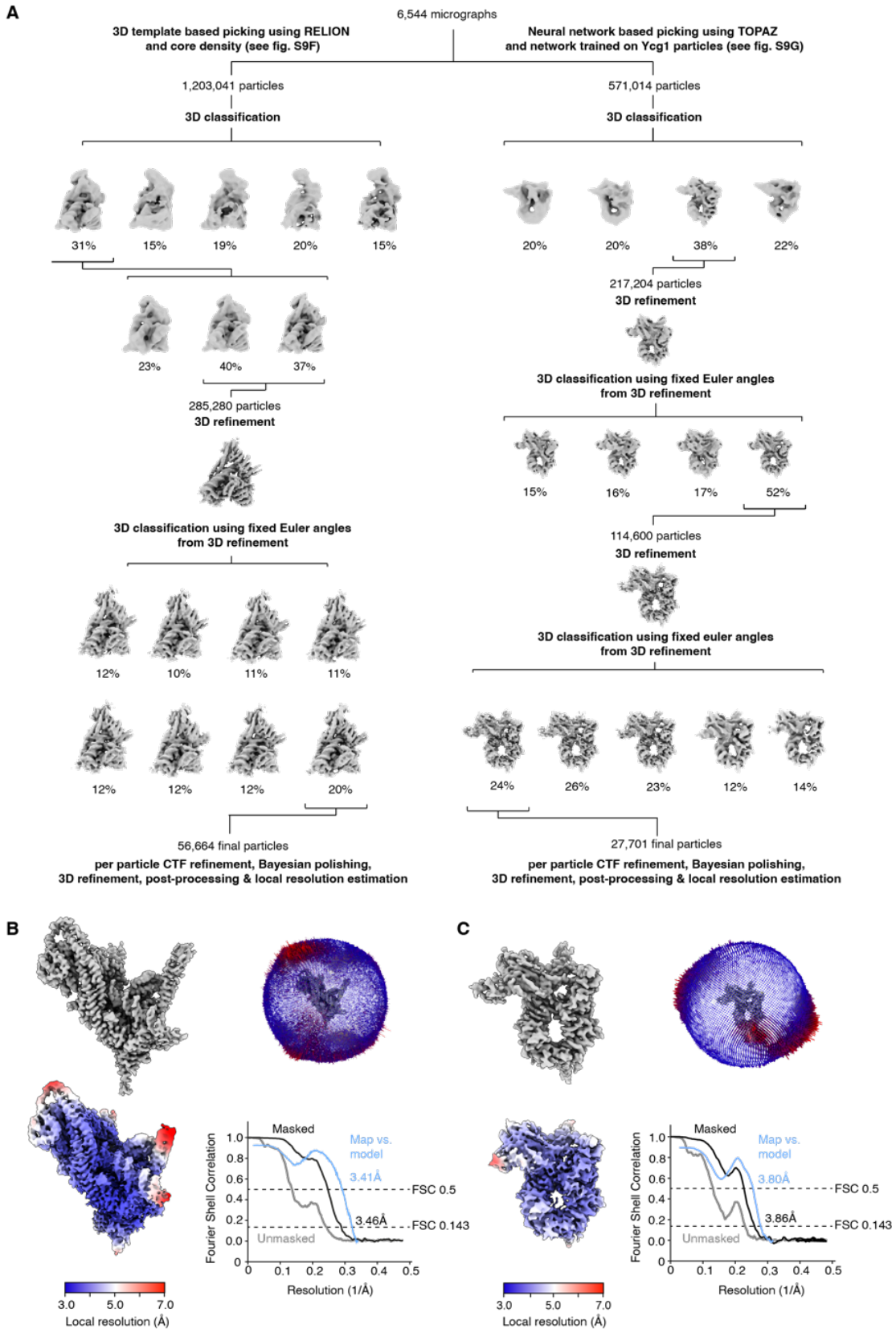


Figure 3.3. Last steps of the analysis pipeline of the ATP- and DNA-bound condensin cryo-EM data set.

Analysis pipeline leading to high resolution maps of the core condensin complex (left) and the Ycg1 peripheral condensin complex (right) (**A**). Final cryo-EM map, plot of the local resolutions, distribution of the particle views and Fourier Shell Correlation (FSC) plot of the core condensin complex (**B**) and the Ycg1 peripheral complex (**C**). Figure adapted from Shaltiel et *al.*, 2022.

The particle classes reveal that condensin is composed of two independent modules that I named 'core' and 'peripheral' complex (**Figure 3.4.A**). I was able to resolve the structures of the core and of the peripheral complex with nominal resolutions of 3.5 Å or 3.9 Å, respectively. Smc2, Smc4, Ycs4 and their respective kleisin interacting fragments constitute the core (**Figure 3.4.B**). DNA is clamped by the Ycs4 N-terminus, while it is flanked in-between the ATPase engaged heads. The N-terminal extremity of Brn1 loops around the DNA. As a result of ATP head engagement, the coiled coils adopt an open conformation. They could not be visualized along their entire length, presumably due to their inherent flexibility. The peripheral complex is formed by the Ycg1 HEAT-repeat subunit and its interacting kleisin fragment. DNA is accommodated in the Ycg1 HEAT hook shape, while it is encircled by the Brn1 kleisin.

Thus, the cryo-EM structures demonstrate that condensin binds DNA at two topological chambers formed by the Brn1 kleisin. The first chamber is created by the first ~200 residues of Brn1 (chamber I). The second chamber is created by the ~130 residues of Brn1 'safety-belt' peptide loop formed along the hook-shaped Ycg1 peripheral complex (chamber II). In addition to forming a topological compartment for DNA entrapment, the Brn1 kleisin threads through the entire condensin complex and tethers the core and the peripheral subcomplexes. Doing so, the kleisin forms an additional topological "intermediate" chamber (chamber IA) (**Figure 3.4A**).

Due to the flexible nature of the Brn1 kleisin, which is predicted to be largely unstructured for the stretches that tether the core to the peripheral subcomplex, I could not visualize the topology of the kleisin between these two subcomplexes. To determine whether DNA co-occupies both DNA binding sites of condensin, a single-molecule tracing approach to position the subcomplexes relatively to each other was performed by Luis Hauptmann in the laboratory of Dr. Sebastian Eustermann. The principle of this

approach consists of mapping the particle coordinates of the core subcomplex relative to the coordinates of the peripheral subcomplex onto the micrographs and of determining whether they adopt a random relative positioning or if they are closer to each other than the estimated random distance. To be considered parts of the same condensin molecule, cores and peripheral complexes have to be within a distance of 30-40 nm (corresponding to fully stretched 65 amino acids separating them) from one another. The analysis did not allow the conclusion that the particles used to generate the 3D model of the core and the peripheral subcomplexes were part of the same condensin molecule. However, this approach was limited by the fact that the mapped particles originated from several rounds of 3D classifications. In this process, the particles were selected based on the resolution of the resulting densities. Therefore, it is likely that the classification procedure caused the limitation of this approach.

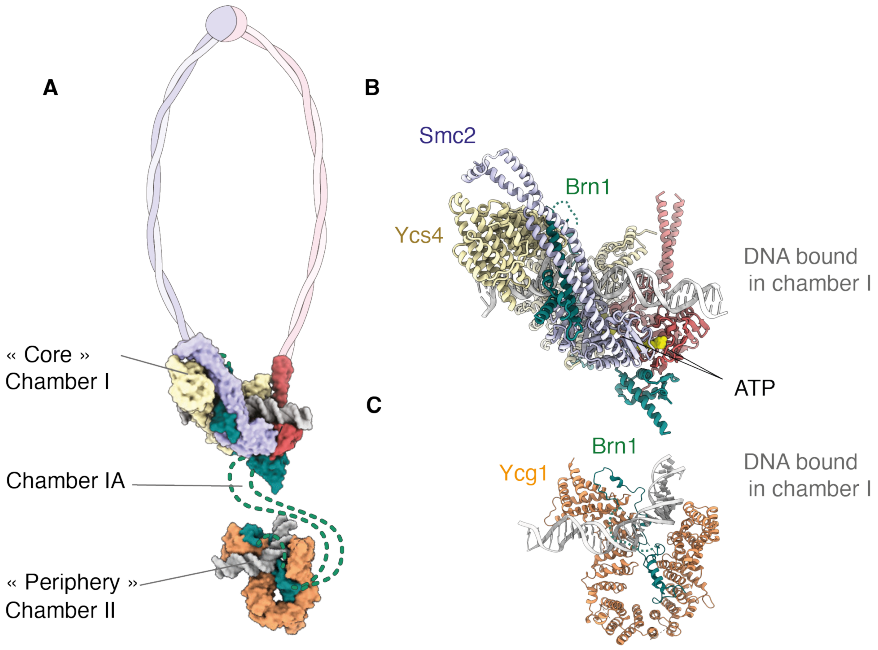


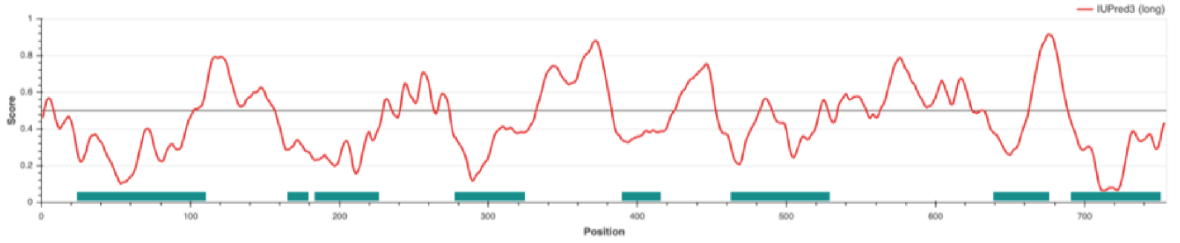
Figure 3.4. Architecture of the ATP- and DNA-bound condensin complex. Atomic model of the ATP- and DNA-bound condensin complex with surface representation (**A**) and cartoon representation (**B**, **C**). The flexible and unresolved Brn1 kleisin fragments are represented by dashed lines. The kleisin chambers I, II and intermediary (IA) are indicated. The unresolved coiled coils are drawn (PDB: 7qen). Figure partially adapted from Shaltiel et *al.*, 2022.

II. Structural analysis of protein interfaces of the DNA- and ATP-bound condensin complex

a. Brn1 conformation and interaction interfaces

The kleisin Brn1 subunit threads through the entire complex (**Figure 3.5**), forming dedicated compartments for DNA while simultaneously interacting with all other subunits. Overall, about 50% of the Brn1 protein sequence was resolved, and 100% of the regions predicted to be structured are visible in the structures (**Figure 3.5A**). This suggests that the non-resolved parts are indeed unstructured and/or too flexible to be resolved. The path and the length of the unstructured DNA-containing chamber regions are indicated in **Figure 3.5B**.

A



B

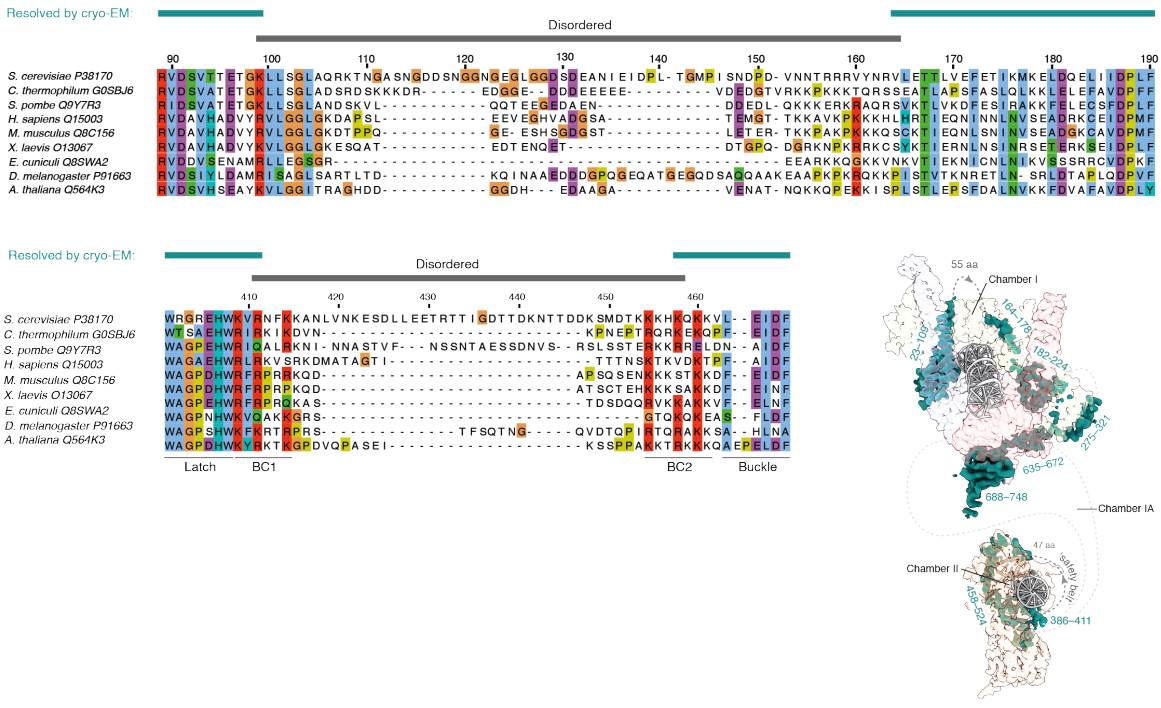


Figure 3.5. Disordered regions of the *Sc* Brn1 kleisin. IUPred secondary structure predictions of Brn1. The part resolved in the ATP- and DNA-bound structure of condensin are shown by the green bars (**A**). Clustal X protein alignment of Brn1 protein sequences across eukaryotic species of chamber I (*top*) and chamber II (*bottom*) regions with resolved and disordered sections respectively indicated by green and grey bars. Cryo-EM density showing the path of the Brn1 kleisin and the architecture of chambers I, II and intermediary (IA)(**B**). Figure partially adapted from Shaltiel et *al.*, 2022.

As observed in other eukaryotic SMC complexes, the N-terminal extremity of the kleisin subunit folds into a Helix-Turn-Helix (HTH) motif and interacts with the Smc2 neck region, formed by the head-proximal region of the coiled coils (**Figure 3.6**). The interaction is partially mediated by $Brn1_{Y86}$, which inserts in between the N- and the C-terminal helices of the coiled coil, in a similar way to what has been observed for the co-structure of the *Chaetomnium thermophilum* (*Ct*) Smc2 neck and Brn1 HTH domains obtained by X-ray crystallography (Hassler et al., 2018). $Brn1_{Y86}$ is highly conserved among eukaryotes (**Figure 3.6A**) and its mutation to phenylalanine partially hinders spore growth in a yeast survival assay, whereas the $Brn1_{Y86A}$ mutation is lethal in this assay (Hassler et al., 2018). In a similar fashion to what was observed for the *Ct* structure, $Brn1_{R89}$ contacts Smc2, which is consistent with the fact that a $Brn1_{R89D}$ mutation induces yeast spore lethality. I identified additional candidate residues responsible for the Smc2-Brn1 interaction, but their impact on the interface maintenance was not yet tested biochemically. For example, $Smc2_{E1016}$ contacts $Brn1_{Y58}$ in the α2 helix of the Brn1 HTH by stacking against the C-terminal extremity of the Smc2 neck region (**Figure 3.6B** and **D**).

The next resolved kleisin fragment, starting 55 residues distal from the HTH, can be divided into three Ycs4-interacting patches (**Figure 3.7**). Patches 1 and 2 (residues $Brn1_{164\text{-}178}$ and $Brn1_{178\text{-}224}$, respectively) form what has been previously identified as a "low-affinity patch" based on biochemical experiments performed with *Ct* Ycs4-Brn1 subcomplexes (Hassler et al., 2018). This interface is mediated by the interactions between hydrophobic residues (e.g., $Brn1_{F188}$ and $Brn1_{F218}$ – $Ycs4_{M1036}$ and $Ycs4_{F1040}$) and by polar

contacts (e.g. Brn1$_{K223}$ and Ycs4$_{D643}$) (**Figure 3.7B and C**). The sequences of these two patches are conserved (**Figure 3.7E**).

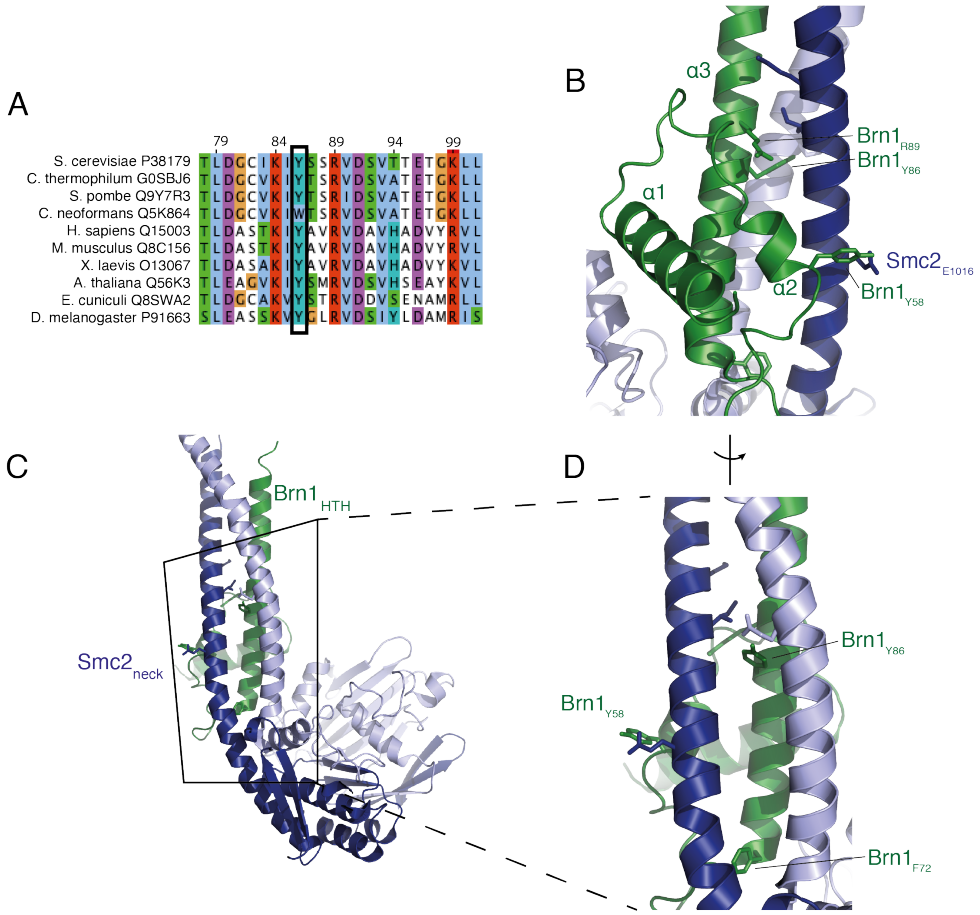


Figure 3.6. Interaction between Brn1 and Smc2 neck region in the context of the *Sc* ATP- and DNA-bound condensin complex. ClustalX alignment of the Brn1 protein sequences across eukaryotic species (**A**). Atomic model (PDB:7QEN) of the Smc2 and Brn1 HTH (helix-turn-helix domain) (**C**) and zoom in on this interface in two different views (**B**, **D**). Residues implicated in the interaction between Brn1 HTH and Smc2 neck region are indicated.

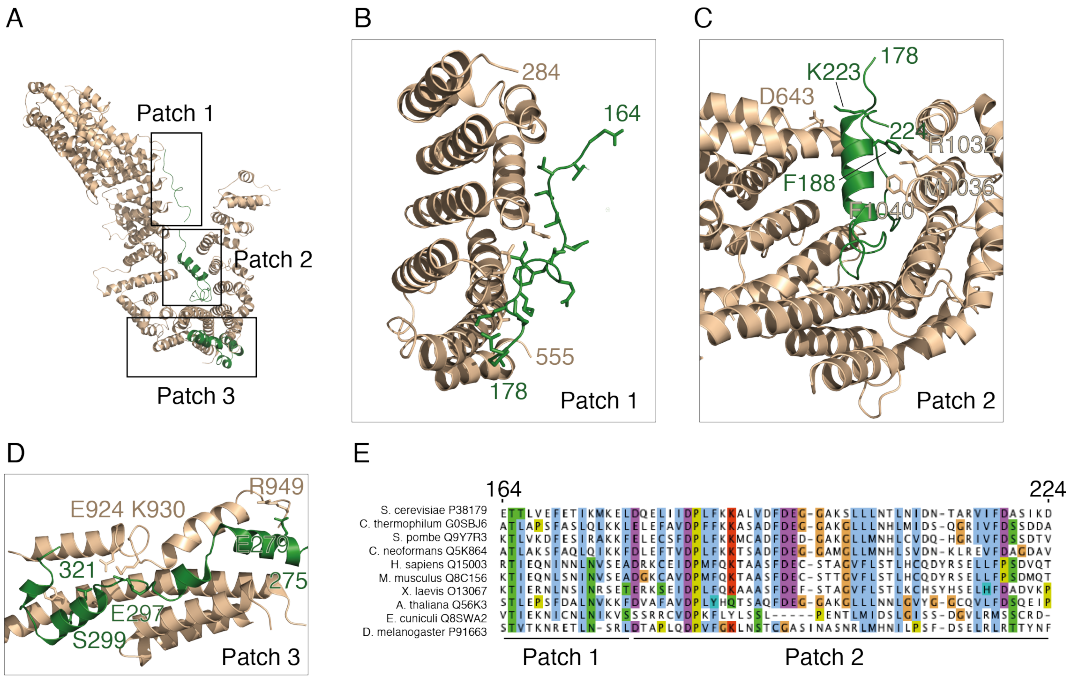


Figure 3.7. Interaction between Ycs4 and Brn1 in the context of the *Sc* ATP- and DNA-bound condensin complex. Atomic model of the Ycs4 and Brn1 interaction patches 1, 2 and 3 (PDB:7QEN) (**A**), with zoom in on the different patches (**B**, **C** and **D**). ClustalX alignment of Brn1 patches 1 and 2 protein sequences across eukaryotic species (**E**).

In comparison to the crystal structure of the *Ct* Ycs4-Brn1 subcomplex (Hassler et *al.*, 2018 – PDB: 6QJ3), my cryo-EM structure reveals the folding of a previously unassigned region of the kleisin. The resolution of the cryo-EM map allows the visualization of the alpha-helical fold of Brn1 residues 275 to 321, to which I refer as Patch 3 (**Figure 3.7D**). The corresponding region in the *Ct* Ycs4-Brn1 crystal could not be resolved with confidence. Despite its low sequence conservation, Patch 3 ($Brn1_{275\text{-}321}$) binds with high affinity to Ycs4 (Hassler et *al.*, 2018). My structure reveals that this interaction is mediated by ionic interactions ($Brn1_{E297}$ – $Ycs4_{K930}$, $Brn1_{E279}$ – $Ycs4_{R949}$) and by hydrogen bound formation ($Brn1_{S299}$ - $Ycs4_{E924}$).

The next two kleisin fragments, which include Brn1 residues 321 to 386 and 524 to 635, tether the Ycg1 peripheral complex to the core and could not be resolved by cryo-EM (**Figure 3.5A**). The structures of the Brn1 fragments that contact Ycg1 (Patch 4, $Brn1_{386\text{-}412}$ and Patch 5, $Brn1_{458\text{-}524}$) are similar to the ones describes in the previously published

crystal structure (Kschonsak et *al.*, 2018). These interaction interfaces are depicted in **Figure 3.8**. The most important Ycg1-interacting residues of the Brn1 kleisin are highlighted and were previously validated by mutagenesis experiments (Kschonsak et *al.*, 2018). Only the so-called "safety belt" region that loops around the DNA could not be visualized, presumably due to it being unstructured (**Figure 3.5A**).

Finally, the Brn1 C-terminal region interacts with the Smc4 head domain through a Winged-helix domain (WH – Brn1 residues 688 to 748), in a similar way to what has been described for the *Ct* Smc4-Brn1 co-crystal structure (Hassler et *al.*, 2018).

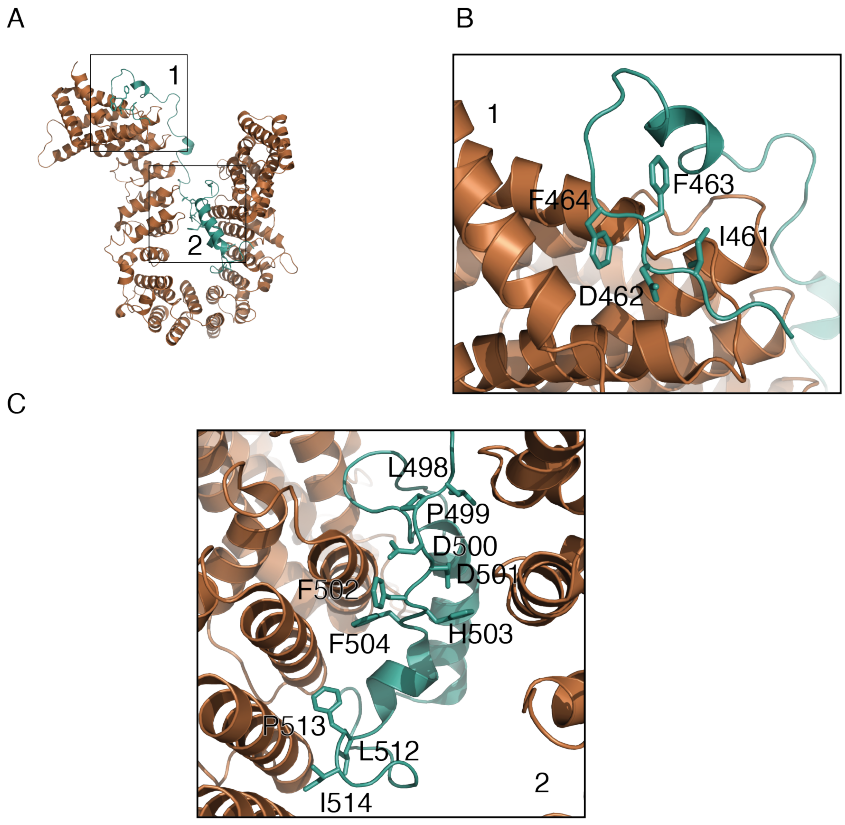


Figure 3.8 Interaction between Ycg1 and Brn1 in the context of the *Sc* ATP- and DNA-bound condensin complex. Atomic model of the Ycg1 and Brn1 interaction patches 1, and 2 (**A**), with zoom in on patch 1 (**B**) and 2 (**C**), and key residues highlighted (**C**).

b. Conformation and interaction interfaces of the Smc subunits

The density for each ATP molecule can unambiguously be distinguished at both Smc active binding sites (**Figure 3.9A**). ATP binding engages the Smc2 and Smc4 heads in an antiparallel orientation, with the signature motif (LSGGQ) of the helical lobe of one

head aligning with the Walker A (P-loop) and the Walker B motifs of the RecA-like lobe of the other head (**Figure 3.9B**).

Due to the inherent flexibility of the 50-nm intramolecular coiled coils of Smc2 and Smc4, their entire length could not be resolved ($Smc2_{241-946}$ and $Smc4_{369-1231}$ are not resolved). As a consequence, the connectivity between the heads and the hinge could not be visualized. Some 2D classes (**Figure 3.1B.**, class "F") strikingly resemble the previously published hinge crystal structure. I could not obtain a high-resolution model from these particles due to their small size.

The structure of the core partially reveals the N-terminal extension of Smc4 that was predicted to be largely unstructured (**Figure 3.9C.**, $Brn1_{125-140}$). The resolution of the obtained electron density map was insufficient to build amino acid side chains, this stretch was therefore built as a poly-alanine model. The sequence of this region did not show any notable conservation among the eukaryotic kingdom, but appeared highly conserved among yeast species (**Figures 3.9D and 3.9E**). Although lengths and amino acid sequences of the eukaryotic Smc4 N-terminal extensions are extremely variable, they contain an important fraction of phosphorylatable residues, such as serines and tyrosines. Intriguingly, the Smc4 N-terminal extensions contain a large proportion of proline residues that can be clustered in conserved patches. *C. thermophilum*, possesses a particularly long extension (120 residues longer than *S. cerevisiae*), with a very high density of prolines (about 20% of the residues, 10% in the case of *S. cerevisiae*).

In the core structure, the resolved $Smc4_{125-140}$ N-terminal stretch contacts Ycs4 in the close proximity of the conserved KG-loop on its posterior surface, which is responsible for the interaction between Ycs4 and Smc4 (Hassler et *al.*, 2018, see the next section). Thus, it is tempting to speculate that the Smc4 N-terminal extension could play a role in regulating the Smc4-Ycs4 interaction. Given the amino acid content of this extension, its function is likely to be regulated post-translationally by phosphorylation of the serine or tyrosine residues or by hydroxylation of the proline residues.

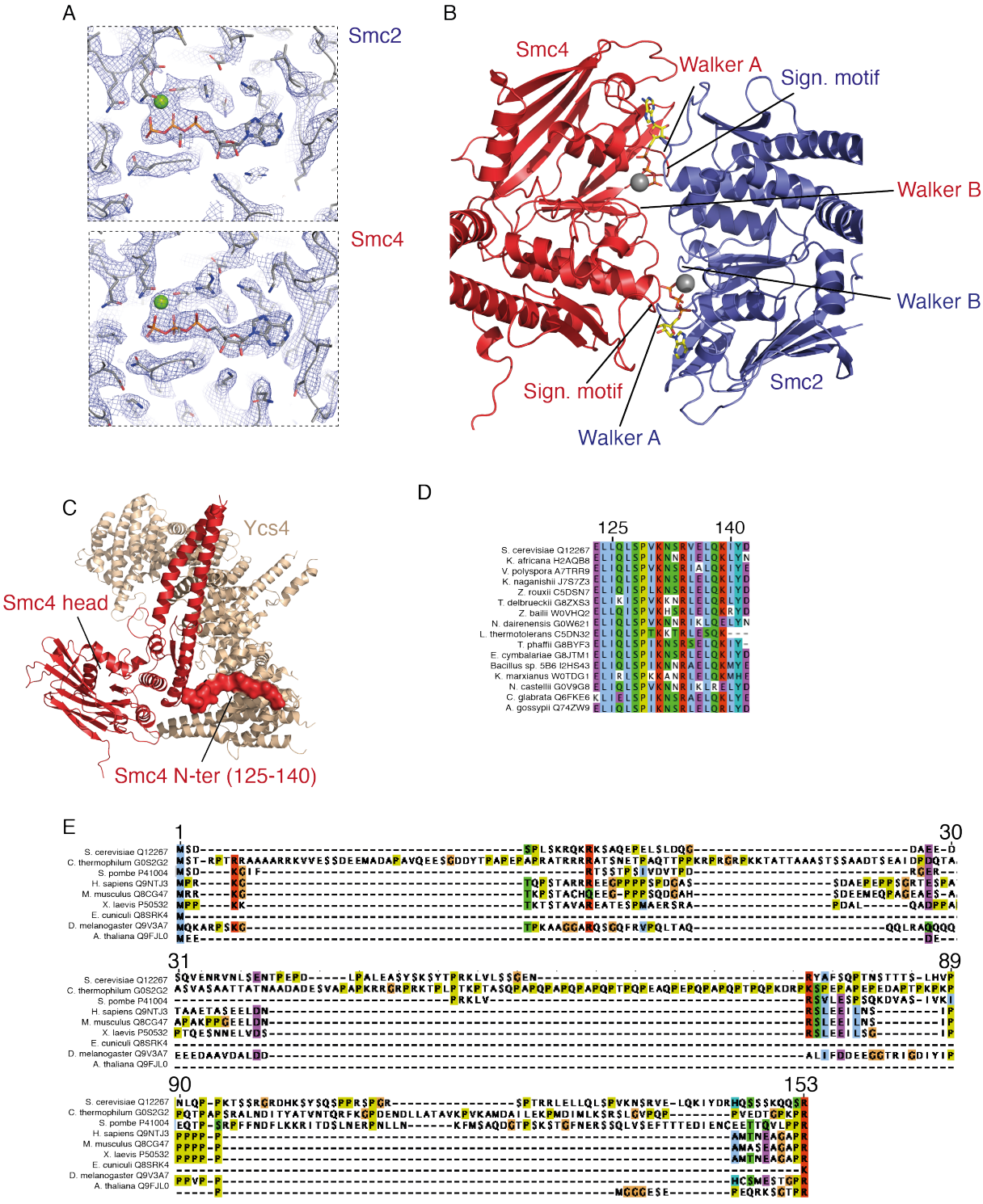


Figure 3.9. Architecture of the condensin ATPase Smc subunits in the context of the *Sc* ATP- and DNA-bound state. Atomic model and electron density of the ATP- and DNA-bound condensin complex showing the ATP bound to the Smc2 (top) and to the Smc4 (bottom) catalytical pockets (**A**). This panel is adapted from Shaltiel et *al.*, 2022. Atomic model of the engaged ATPase domains of Smc2 and Smc4 with Walker A, B and signature motifs together with ATP and Mg^{2+} indicated (**B**). Atomic models of the Smc4 and Ycs4 subunits of the ATP- and DNA-bound condensin complex in cartoon representation, and in surface representation for the N-terminal (N-ter) extension of Smc4. The Smc4 N-ter contacts Ycs4 (**C**). ClustalX alignment of the protein sequences of the Smc4 N-ter region, that is resolved in the *Sc* ATP- and DNA-bound condensin cryo-EM structure (125-140), across yeast species and showing a high degree of

conservation (**D**). ClustalX alignment of the protein sequences of the Smc4 N-ter region (from residue 1 to 153) across eukaryotic species and showing a poor degree of conservation and variable sequence length (**E**).

c. Ycs4 interacts with the Smc2 and Smc4 proteins

Ycs4 binds to both Smc subunits. The first interaction interface is located in-between the Ycs4 N-terminal lobe and the Smc2 neck region (**Figure 3.10A**). This interface seems rather sparse, and it seems to be only mediated by a couple of residues ($Smc2_{E194}$-$Ycs4_{K286}$, $Smc2_{N226}$-$Ycs4_{T186}$, $Smc2_{N988}$-$Ycs4_{R198}$). The interface between Smc4 and Ycs4 observed in my cryo-EM structure is similar to the interface that has been described by a previous crystal structure (Hassler et *al.* 2018) (**Figure 3.10.B.**). The so-called D-loop of Smc4 and the W-loop of Ycs4 mediate the contact between the two proteins. In addition to these regions, the HEAT-repeat helix 15b of Ycs4 contacts the Smc2 ATPase head. This interface will be further discussed in **Chapter 3**.

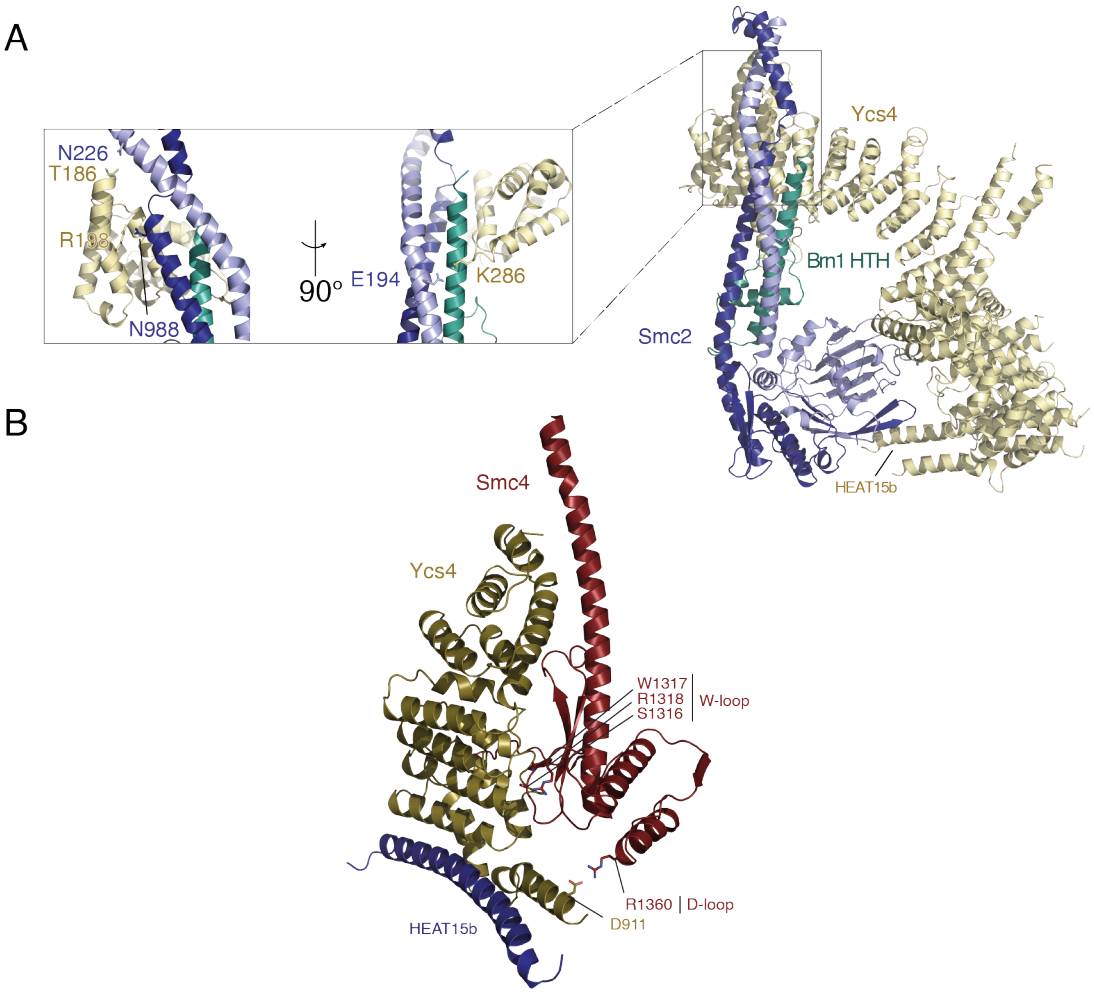


Figure 3.10. Interaction between Ycs4 and the Smc subunits in the context of the *Sc* ATP- and DNA-bound condensin complex. Atomic model of the Smc2, Ycs4 and Brn1 subunits of the ATP- and DNA-bound condensin complex (**A**). Ycs4 N-terminal lobe contacts Smc2 neck region and Brn1 HTH (Helix-Turn-Helix) domain. Important residues are highlighted. Atomic model of the Smc4 and Ycs4 subunits of the *Sc* ATP- and DNA-bound condensin complex showing the interaction interface between the two subunits at the Smc4 W- and D-loop regions. The helix HEAT15b of Ycs4 is highlighted in blue (**B**). Panel B is adapted from Shaltiel et *al.*, 2022.

III. Molecular details of the DNA-binding regions of ATP-bound condensin

a. Overview of topological DNA-binding chambers

DNA is topologically enclosed in two condensin chambers: the core motor chamber and the peripheral anchor chamber (**Figure 3.4.B and C**). In this section, I will focus on describing the molecular interactions between condensin and DNA in both chambers.

At the core motor chamber, DNA is clamped in-between the Smc2-Smc4 coiled coils by the N-terminal extremity of the Ycs4 HEAT-repeat subunit in a similar way as has been observed for related SMC complexes (Bürmann et al., 2021; Collier et al., 2020; Higashi et al., 2020; Shi et al., 2020; Yu et al., 2022). My structure shows 35 bp of the 50-bp DNA double helix present in the sample. At the peripheral anchor chamber, which is formed by Ycg1 and its interacting kleisin fragment, 31 bp out of the 50-bp DNA are resolved. In both cases, the DNA phosphate backbone could be traced with certainty. DNA was, however, not resolved at single base-pair resolution and was hence built as a poly-AT chain. The lack of resolution could be explained by condensin binding DNA in different registers. The condensin-DNA interactions display different modalities: (i) insertion of residues into the minor groove, (ii) insertion of residues into the major groove, and (iii) residue-phosphate backbone interaction. Whereas the latter suggests an unspecific interaction, the two former interactions cannot exclude that some DNA sequences might be preferentially bound by the condensin complex in a cellular context.

b. Molecular details of DNA binding at the core chamber

i. Ycs4

Residues of the Ycs4 N-terminal HEAT-repeat helices 8b, 7b, 5b and 3b are in direct contact with the DNA (**Figure 3.11A**). Positively charged residues ($Ycs4_{K292}$, $Ycs4_{K300}$, $Ycs4_{K377}$, $Ycs4_{K384}$, $Ycs4_{R416}$, $Ycs4_{K420}$) contact the DNA at the phosphate backbone or by inserting into the minor groove (for the case of $Ycs4_{R416}$). The side chain of $Ycs4_{Y373}$ enters the DNA minor groove and is presumably involved in a base-stacking interaction. The presence of an aliphatic residue at this position is a strictly conserved feature of Ycs4 sequences among eukaryotes. Likewise, the majority of the positively charged residues in the vicinity of the DNA are highly conserved among eukaryotes (**Figure 3.11F**) and form a positively charged patch (**Figure 3.11G**). Mutation to glutamate of the DNA-proximal residues $Ycs4_{K292}$, $Ycs4_{K300}$, $Ycs4_{K377}$, $Ycs4_{K384}$, $Ycs4_{R416}$, $Ycs4_{K420}$ impaired or reduced yeast

survival (**Figure 3.11A and B**). In line with these results, double mutant $Ycs4_{K292E/K300E}$ or $Ycs4_{K377E/K384E}$ condensin complexes neither showed DNA-stimulated ATPase nor DNA loop extrusion activities, but maintained a stoichiometry similar to wild-type complexes, as shown by SDS-gel electrophoresis (**Figure 3.11C**, **D** and **E**).

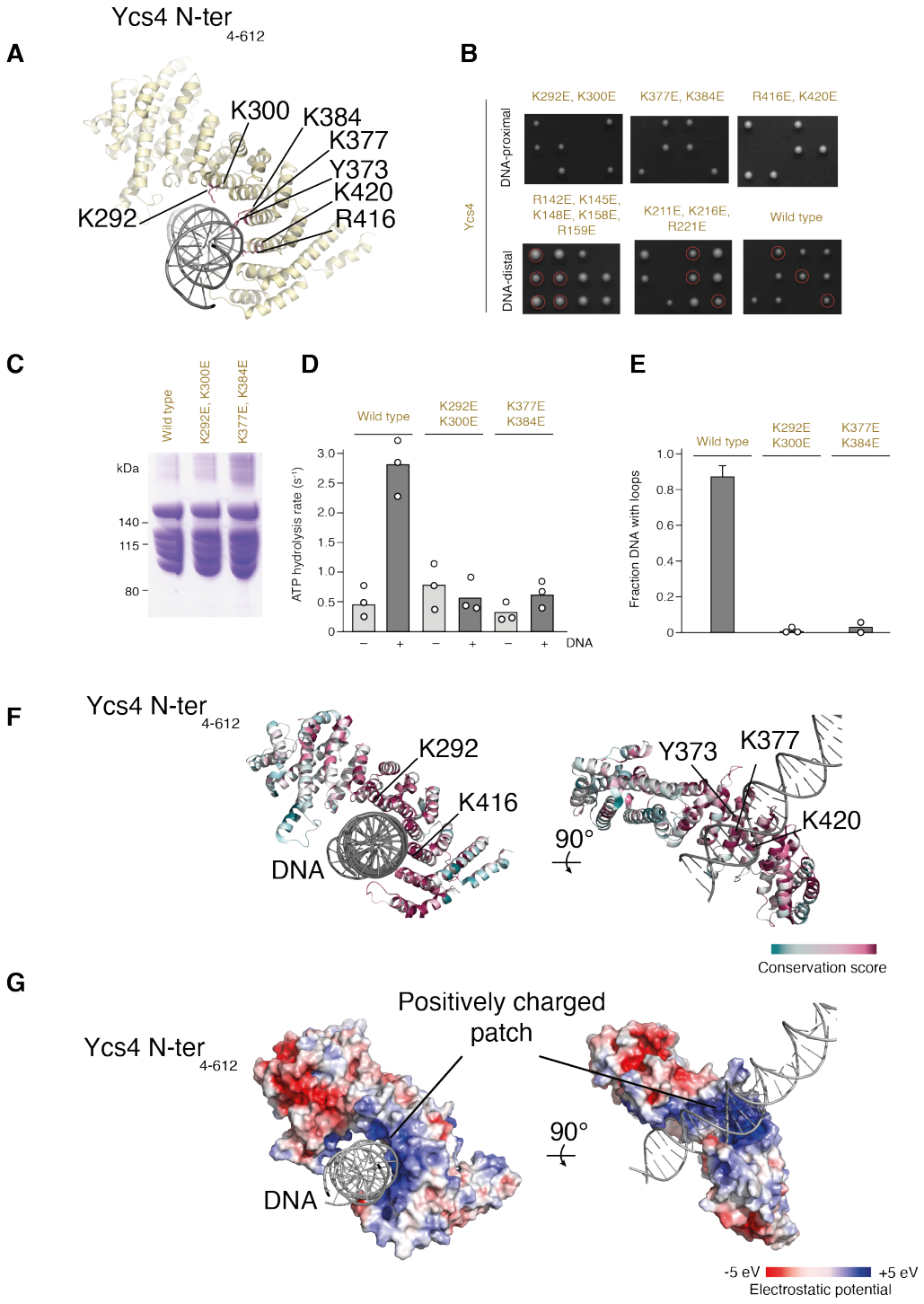


Figure 3.11. DNA binding of the Ycs4 subunit in the context of the *Sc* ATP- and DNA-bound condensin complex. Atomic model of Ycs4 N-terminal lobe contacting the DNA. Important residues are highlighted (**A**). Effect of *Sc* Ycs4 reversal charge mutations on budding yeast survival (**B**). SDS-PAGE showing wild-type or mutated *Sc* condensin

holo-complex stoichiometry (**C**). ATPase and loop extrusion activities of wild-type or mutated *Sc* condensin holo-complexes (**D**, **E**). Experiments corresponding to panels **B**, **C**, **D** and **E** were performed by Dr. Indra Shaltiel in Pr. C. H. Haering lab. Panels **B**, **C**, **D** and **E** are adapted from Shaltiel et *al.* 2022. Conservation plot of the DNA-bound region of Ycs4 showing a high sequence conservation (**F**). Electrostatic potential of the DNA-bound region of Ycs4 showing a positively charged patch responsible for the DNA binding (**G**).

ii. Smc2 and Smc4

The engaged ATPase head domains contain three positively charged patches that contact the DNA (**Figure 3.12A** and **B**). Smc2 binds DNA through the helical and RecA-like lobes of the ATPase head and through the inner side of the coiled-coil neck region. In total, I identified nine basic residues that interact with the DNA phosphate backbone (**Figure 3.12A**). Six of them are strictly conserved among eukaryotes. The DNA phosphate backbone binding mode suggests that Smc2 binds to its DNA substrate with little or no sequence specificity. Charge reversal mutations of $Smc2_{K114}$ and $Smc2_{K121}$ to glutamate impaired and abolished yeast survival, respectively, while the mutation to glutamate of the DNA-distal lysine residue $Smc2_{K76}$ had no effect (**Figure 3.12C**).

Five basic residues are spread across the Smc4 head domain and contact DNA, with four of them being strictly conserved. Among the conserved residues, $Smc4_{R205}$ inserts into the DNA minor groove. $Smc4_{K312}$ and $Smc4_{K317}$ are part of a loop (thereafter referred to as K-loop) that fully inserts into the DNA major groove. The K-loop is strictly conserved (**Figure 3.12D**), which suggests that it is important for condensin function. The insertion of Smc4 residues into the two DNA grooves implies that this subunit might be able to recognize specific DNA sequences, but the resolution of the DNA did not allow me to visualize individual bases and thereby confirm my hypothesis.

iii. Brn1

The core DNA-binding chamber is formed by the N-terminal ~200 residues of Brn1, which includes the Helix-Turn-Helix (HTH) motif (**Figure 3.4**). Simultaneous charge reversal mutations of the N-terminal and DNA-proximal residues $Brn1_{N41}$ and $Brn1_{K42}$ to glutamate render condensin non-functional in budding yeast (**Figure 3.12E**). The flexible and unstructured kleisin region of 55 amino acids that follows the HTH motif to the first

binding site for Ycs4 (residues 110–163) wraps around the DNA to form a closed compartment. Although this peptide stretch could not be visualized in the cryo-EM structure, its it is likely that this region merely serves as a linker. While this connecting linker is 54-residues long in *S. cerevisiae,* it is considerably shorter in other species (**Figure 3.5B**).

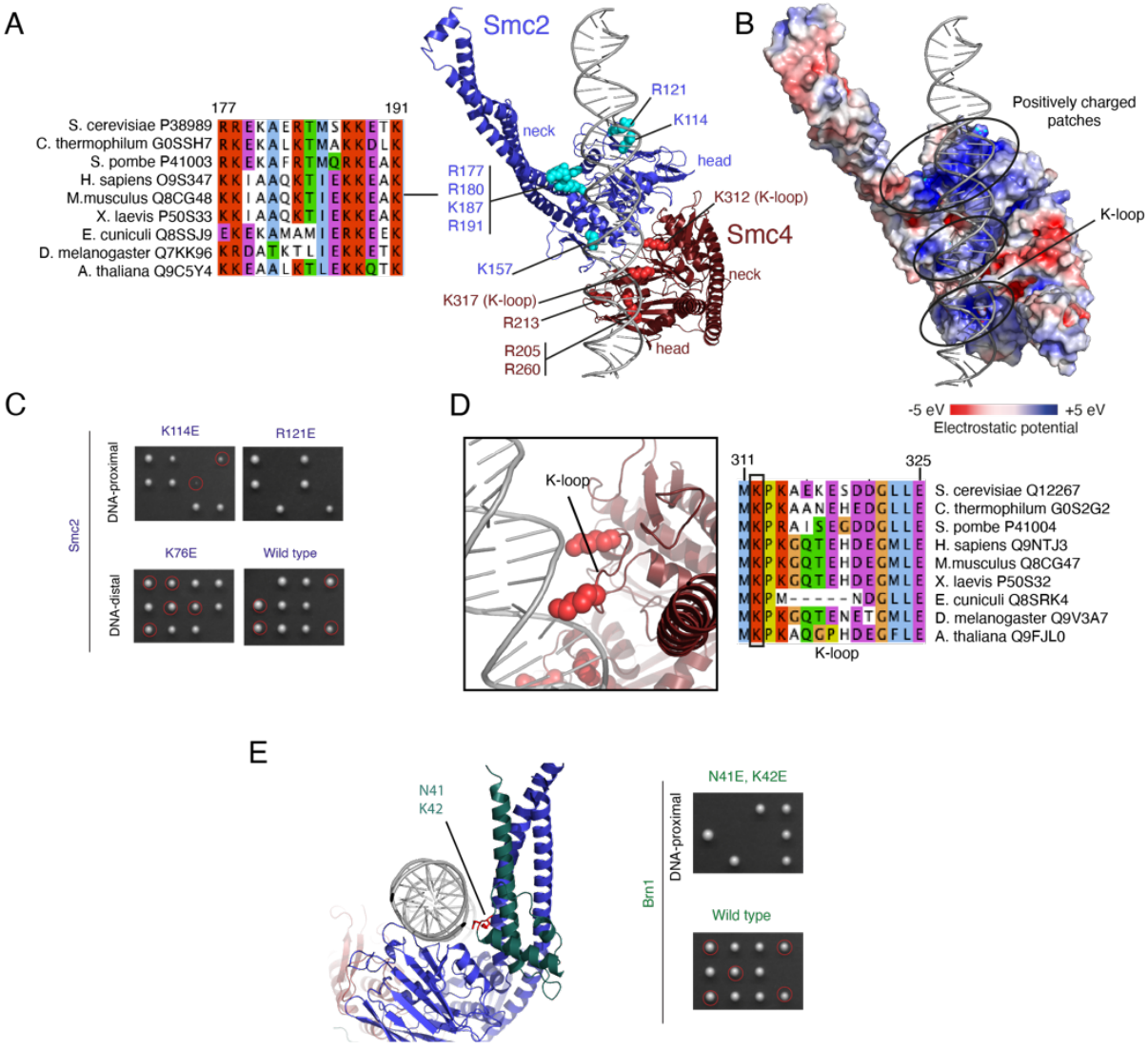


Figure 3.12. DNA binding of the Smc and the Brn1 subunits in the context of the *Sc* ATP- and DNA-bound condensin complex. Atomic model of the Smc subunits of the ATP- and DNA-bound condensin complex with important residues highlighted (**A**, right). ClustalX alignment of the protein sequence corresponding to a Smc2 patch that contacts the DNA, across eukaryotic species (**A**, left). Atomic model of the Smc subunits of the ATP-bound condensin complex bound to the DNA with surface representation plotting the electrostatic potential and showing positively charged patches at the top ATPase head domains, the K-loop is highlighted (**B**). Yeast survival assay showing the effect of reversal charge mutation of Smc2 residues implicated in DNA binding (**C**). ClustalX alignment of protein sequences corresponding to the Smc4 K-loop that contacts the DNA, across different eukaryotic species (right) and zoom-in on this region (left) (**D**). Atomic model of the Smc2 and Brn1 subunits of the ATP- and DNA-bound condensin

complex. Brn1 residues contacting the DNA are highlighted (**E**, left). Yeast survival assay showing the effect of reversal charge mutation of Brn1 residues implicated in DNA binding (**E**, right). Yeast survival assays shown in panels **C** and **E** were performed by Dr. Indra Shaltiel. These panels are partially adapted from Shaltiel et *al.*, 2022.

IV. Molecular details of DNA binding at the peripheral chamber

The interaction of the peripheral subcomplex with DNA is mediated by both, the Ycg1 HEAT-repeat subunit and the Brn1 kleisin subunit (**Figure 3.13**). The DNA is bent in a ~90° angle and trapped in-between the N- and C-terminal lobes of Ycg1. In addition, the 130 residues of the Brn1 kleisin safety belt loop around the DNA. My structure resolves ~65% of the safety belt residues. The missing part is predicted to be unstructured, and neither its length nor its sequence are conserved among eukaryotes. This suggests that its role is restricted to the formation of a topological loop around the DNA without making specific contacts (**Figure 3.5B**).

Ycg1 contains many basic residues that contact the DNA (**Figure 3.13B**). The majority is located at the N- or at the C-terminal lobe of the protein. Six of these residues are highly conserved among eukaryotes (**Figure 3.13A**). Three of these residues ($Ycg1_{R170}$, $Ycg1_{R849}$, $Ycg1_{R895}$) contact the DNA by inserting into the deformed grooves at the bending point. $Ycg1_{R77}$ and $Ycg1_{K70}$, localized at the N-terminal lobe, and $Ycg1_{R253}$, localized at the "body" of the protein, interact with the DNA by contacting the phosphate backbone and inserting into the minor groove, stabilizing one of the DNA extremities.

Seven basic residues of Brn1 contact the DNA. Only two of these residues are conserved among eukaryotes. Interestingly, and similar to what is observed for Ycg1, these residues are located at the bending point of the DNA. $Brn1_{409}$ and $Brn1_{411}$, situated just upstream the unresolved safety belt peptide loop ($Brn1_{413–457}$), inserts into the distorted major groove of the DNA double helix.

Because the DNA could not be resolved at single-base pair resolution, it remains unknown whether these residues form base-specific interactions. Nevertheless, DNA bending might be an indication that the peripheral complex recognizes a specific sequence or

shape of the DNA. In addition, DNA bending could provide the molecular explanation for how the peripheral complex can perform an anchoring function by helping condensin to resist longitudinal pulling forces during DNA loop extrusion.

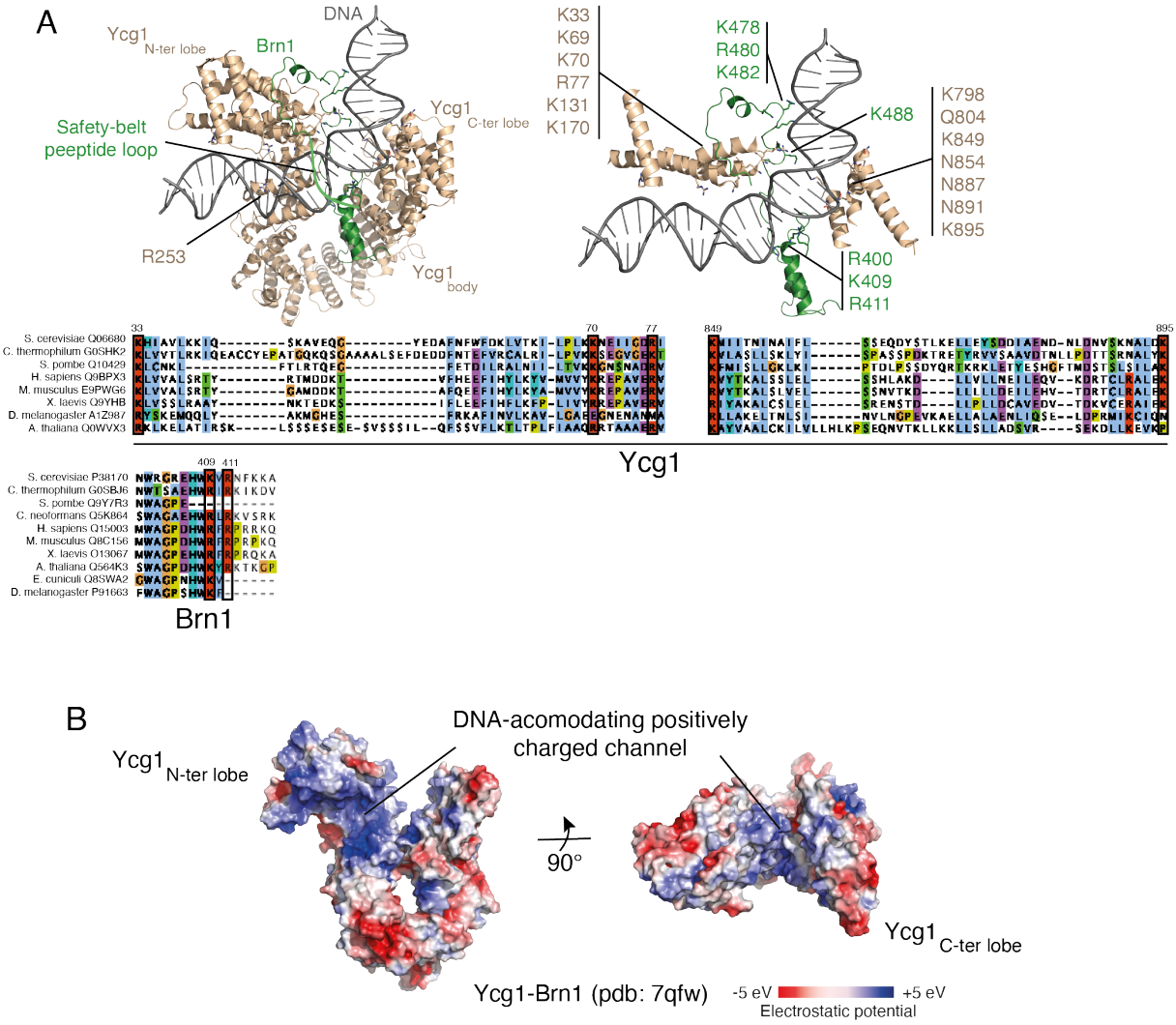


Figure 3.13. DNA binding of the Ycg1 and Brn1 subunits in the context of the *Sc* ATP- and DNA-bound condensin complex (PDB: 7QEN). Atomic model of the Ycg1 and Brn1 subunits of the ATP-bound condensin complex bound to the DNA with important residues highlighted (**A**, top). ClustalX alignment of protein sequences of Ycg1 and Brn1 across eukaryotic species showing the conservation of the DNA binding residues (**A**, bottom). Atomic model of the Ycg1 and Brn1 subunits of the ATP-bound condensin complex bound to the DNA using a surface representation plotting the electrostatic potential (**B**). The model shows a DNA-accommodating positively charged channel on the interior surface of the N-terminal lobe of the Ycg1 subunit, and in between its N- and C-terminal lobes. Brn1 also contributes to the formation of this positively charged channel.

V. Comparison with the ATP-free DNA-free state condensin structure

a. Smc subunit rearrangements

As a result of ATP binding, the Smc2 and Smc4 heads engage and the coiled coils open. A comparison of the ATP- and DNA-bound state of condensin (this work, PDB:7QEN) with the previously resolved structure of the nucleotide-free apo state (Lee et *al.*, 2020, PDB: 6YVU) reveals major molecular rearrangements (**Figure 3.14**): When the Smc4 RecA-like lobes were used as an alignment point, it becomes obvious that Smc4 RecA-like and helical lobes of both states align almost perfectly (**Figure 3.14A**). Taking the Smc2 Rec-A like lobes as an alignment point for comparing the two states reveals that the Q-loop, which contains a highly conserved glutamine residue ($Smc2_{Q147}$), is reoriented between the two states (**Figure 3.14C**). At a strategic position in-between the RecA-like and the helical lobes, the Q-loop could have a mechanistic coupling role by sensing the ATP state and by translating it into an increase of the coiled-coil opening angle by about 25°. This principle is reminiscent of the "coupling helix" observed in the related ABC transporters (Hassler et al., 2019; Wen and Tajkhorshid, 2011).

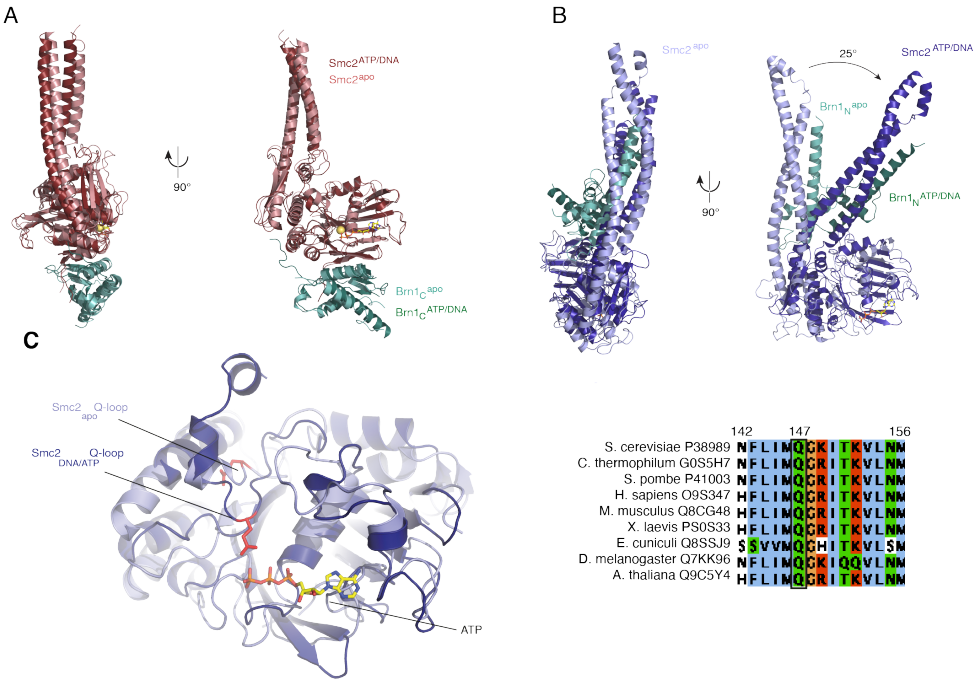


Figure 3.14. Molecular rearrangements of the Smc subunits of *Sc* condensin upon ATP and DNA binding (PDB: 6YVU, Lee et al., 2020; 7QEN, this work). Alignment of atomic models of the Smc4 head domain bound to the C-terminus of Brn1 of the apo ATP-free state (PDB: 6YVU) to the DNA- and ATP-bound state (PDB: 7QEN) of condensin, showing a perfect overlap (**A**). Aligned atomic models of the Smc2 head domain bound to the N-terminus of Brn1 of the apo ATP-free state (PDB: 6YVU) and the DNA- and ATP-bound state (PDB: 7QEN) of condensin, showing a 25° reorientation of the Smc2 coiled coils regarding to the ATPase domain. The ATPase domains show no major rearrangement upon ATP and DNA binding (**B**). Panels **A** and **B** are adapted from Shaltiel et *al.*, 2022. Zoom in onto the Smc2 ATPase domains of the *Sc* apo ATP-free state (PDB: 6YVU) and the *Sc* DNA- and ATP-bound state (PDB: 7QEN) of condensin, showing a reorientation of the Q-loop, siting at the interface between the helical and the RecA-like lobe of Smc2. The Q-loop contacts the ATP in the DNA-bound condensin state (**C**, left). ClustalX alignement of the protein sequence of Smc2 Q-loop across eukaryotic species showing a strict conservation of the glutamine residue of the Q-loop (**C**, right).

b. Ycs4 HEAT-repeat subunit rearrangement

Ycs4 is composed of 21 HEAT repeats. It also contains an additional helical protrusion previously called “proboscis”. A comparison of the core structure with the structure of the Ycs4 subunit obtained in the context of the ATP-free apo state (Lee et *al.*, 2019, PDB: 6YVU) was performed by aligning the RecA-like Smc4 lobes (**Figure 3.15**). Despite the fact that the Ycs4 body rotates by about 30°, the interaction interface between the Ycs4 KG-loop and the Smc4 W-loop is maintained (Hassler et al., 2018, **Figure 3.10B**). Strikingly, the Ycs4 N-terminal lobe bends at a 90° angle (**Figure 3.15**), which induces the formation of a protein interface with the Smc2 neck and the Brn1 HTH (**Figure 3.10A**). The inner surface of the Ycs4 N-terminal lobe is highly positively charged, which explains how it can efficiently clamp onto the DNA molecule (**Figure 3.11F**).

In addition to these large molecular rearrangements, the tilt of the HEAT15b helix towards the bottom of the Smc4 head domain explains how the Ycs4 positioning is compatible with ATP binding at both of Smc active sites. The observation that the Ycs4 HEAT15b helix inserts in between the ATPase heads, directly contacting the nucleotide binding sites in the ATP-free apo structure, suggests that this helix might inhibit ATP binding (**Figure 4.15**, see **Chapter 3**).

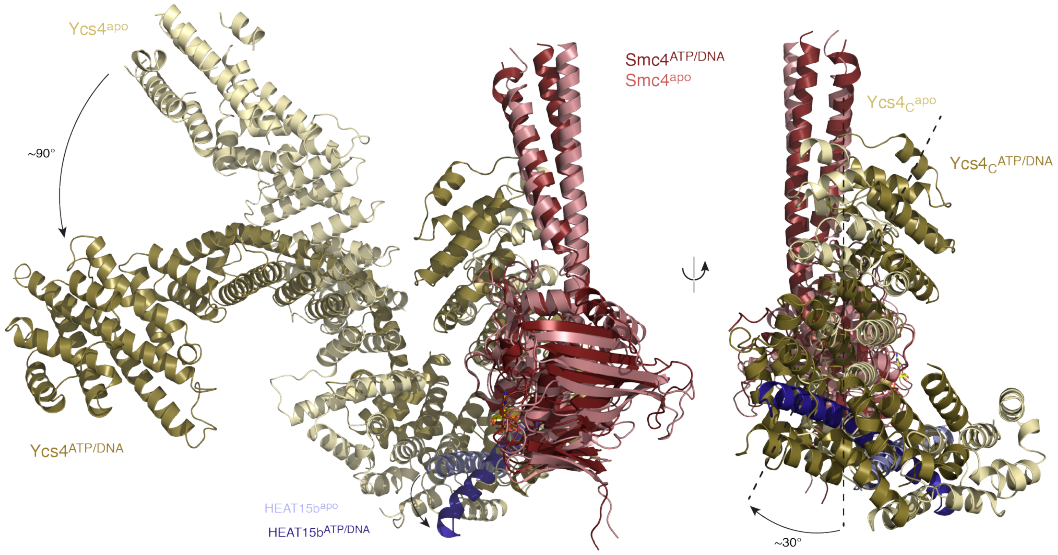


Figure 3.15. Ycs4 molecular rearrangements upon ATP- and DNA-binding to *Sc* condensin. Aligned atomic models of the Smc4 head domain and Ycs4 corresponding to the apo ATP-free state (PDB: 6YVU) and the DNA- and ATP-bound state (PDB: 7QEN) of *Sc* condensin, taking Smc4 RecA-like lobe as a reference point. The N-terminal lobe of Ycs4 bends with a 90° angle upon ATP- and DNA-binding, in addition of a 30° rotation of the all protein. The Ycs4 HEAT 15b, highlighted in blue, is shifted downwards upon ATP- and DNA-binding. This figure is adapted from Shaltiel et *al.*, 2022.

c. Mapping the position of DNA in the ATP-free condensin core

To identify DNA-bound compartments of condensin in the absence of ATP, an immuno-precipitation assay was performed by Dr. I. A. Shaltiel. Condensin was first loaded onto a plasmid in an ATP-dependent manner. After removing ATP with several washing steps, chambers I and II were specifically cross-linked, followed by protein denaturation by SDS at 65 °C (see methods). The formation of SDS-resistant DNA-protein catenanes inside the crosslinked chambers attests for topological DNA entrapment, which should be released by TEV protease cleavage of Brn1 at the cross-linked chamber.

Cysteine pairs for cross-linking chambers I and II were designed to be compatible with both, the ATP-free and the ATP-bound state of the complex (PDB: this work, 7QEN; 6YVU). However, since ATP was extensively washed away before cross-linking, is it likely that the ATP-free condensin state was probed during these experiments. Closure of chamber I, formed by a cross-link between $Smc2_{Q71C}$ and $Brn1_{E197C}$, led to the formation of SDS-resistant DNA-protein catenanes. Similarly, closure of chamber II,

which resulted from $Brn1_{S384C}$ and $Brn1_{S524C}$ cross-linking, retained DNA-protein catenanes. Upon TEV treatment, DNA was released from either of the two chambers (**Figure 3.16A-C**). Thus, in analogy to what I described for the ATP-bound condensin structure, this data suggests that chambers I and II topologically encircle DNA also in the absence of ATP. Based on these data, I docked DNA into chamber I of the ATP-free apo structure of condensin (**Figure 3.16D and E**). Comparison of chambers I of the ATP-free DNA-free and ATP-bound and DNA-bound structures reveals a similar organization. In addition, some of the residues that contact DNA in the ATP-bound structure would also be correctly positioned to do so in the ATP-free apo state.

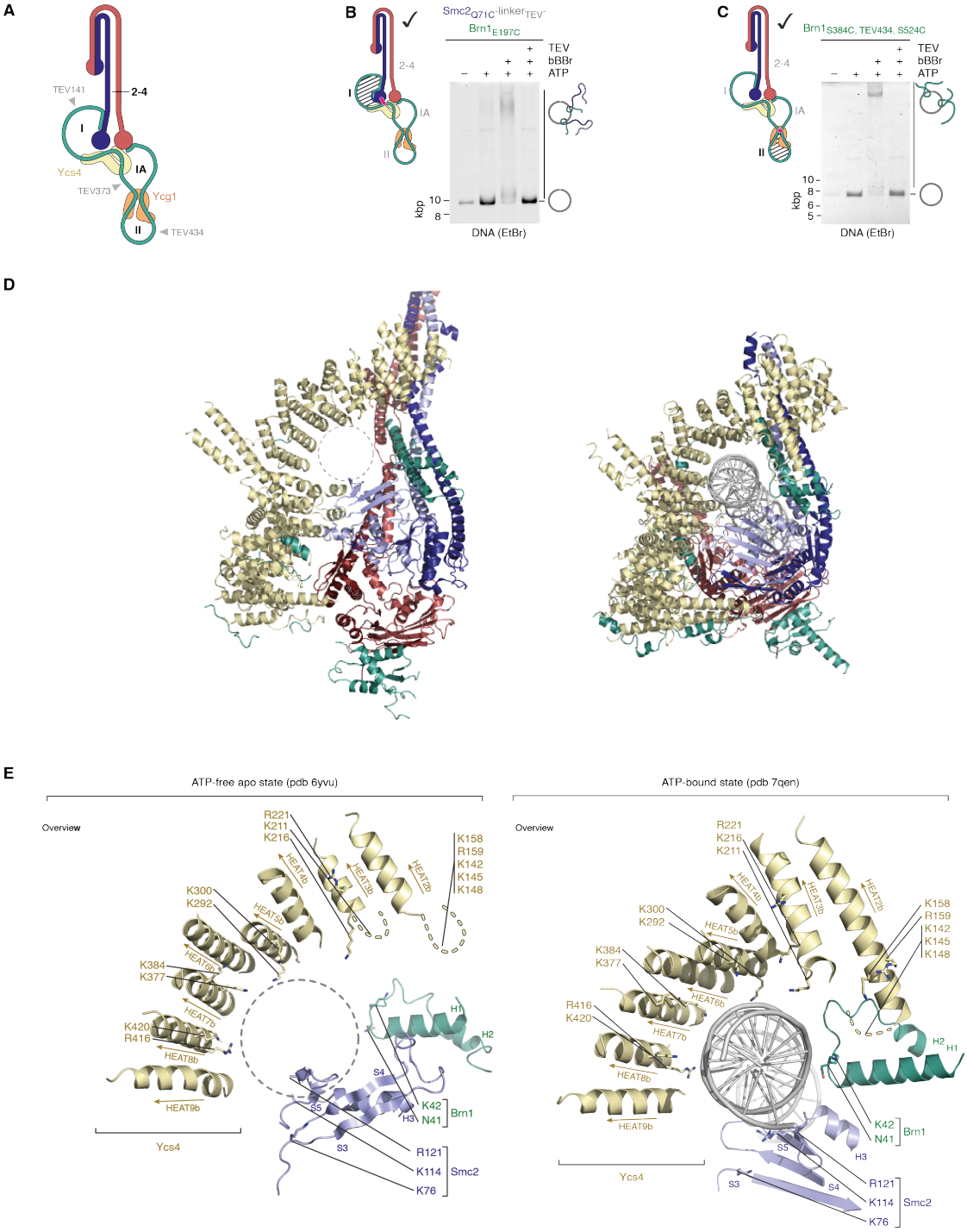


Figure 3.16. DNA sits at the condensin chambers I and II. Schematic representation of condensin complex chambers. The TEV sites, used to cleave the different chambers, are depicted (**A**). Cross-linking experiments showing topological entrapment of DNA into chamber I (**B**) and chamber II (**C**) of condensin. Comparison of chambers I of the ATP-free apo state of *Sc* condensin (left, PDB:6YVU, Lee et al., 2020) to the ATP- and DNA-bound state of *Sc* condensin (right, PDB: 7QEN, this work) showing a similar architecture (**D**). Zoom in on chambers I of the ATP-free apo state of *Sc* condensin (left, PDB:6yvu, Lee et al., 2020) and on the ATP- and DNA-bound state of *Sc* condensin (right, PDB: 7qen, this work) (**E**). The residues contacting the DNA in the ATP- and DNA-bound state of condensin are correctly positioned to do so in the ATP-free apo state of condensin. Figure adapted from Shaltiel et *al.*, 2022. The cross-linking experiments corresponding to panel **B** and **C** were performed by Dr. Indra Shaltiel.

d. A power-stroke motion as a molecular explanation for the condensin motor function

As discussed in the introduction, condensin extrudes DNA in an asymmetric fashion (Ganji et *al.*, 2018). This requires that a single condensin molecule has two independent modules; one needs to translocate DNA (motor module) and the other one needs to stationary hold on to the DNA (anchor module).

In order to identify the motor and the anchor modules, single-molecule DNA loop extrusion assays were performed by Dr. I. A. Shaltiel (**Figure 3.17**). Condensin complexes were labelled with an ATTO647N fluorophore and their loop extrusion activity was monitored after cleaving DNA chambers I or II with TEV protease in a similar way to what I described above (**Figure 3.17A**). Cutting the motor chamber should retain condensin at its anchor site, corresponding to the extrusion start site. On the contrary, cutting the anchor chamber should retain condensin at the translocating end of the loop.

Cutting core chamber I led to loop rupture and the return of condensin to the loop extrusion start position in the majority of the cases (56/65). This indicates that the condensin motor function resides at chamber I (**Figure 3.17C** and **D**). By contrast, cutting the peripheral chamber II induced loop rupture and condensin to be retained at the translocating position in half of the cases (36/76). For the other loop rupture events, condensin either dissociated (8/76) or remained at the anchor position (32/76) (**Figure 3.17C** and **D**). These experiments suggest that the core module performs the condensin motor function, whereas the peripheral module ensures the anchoring of condensin to the DNA.

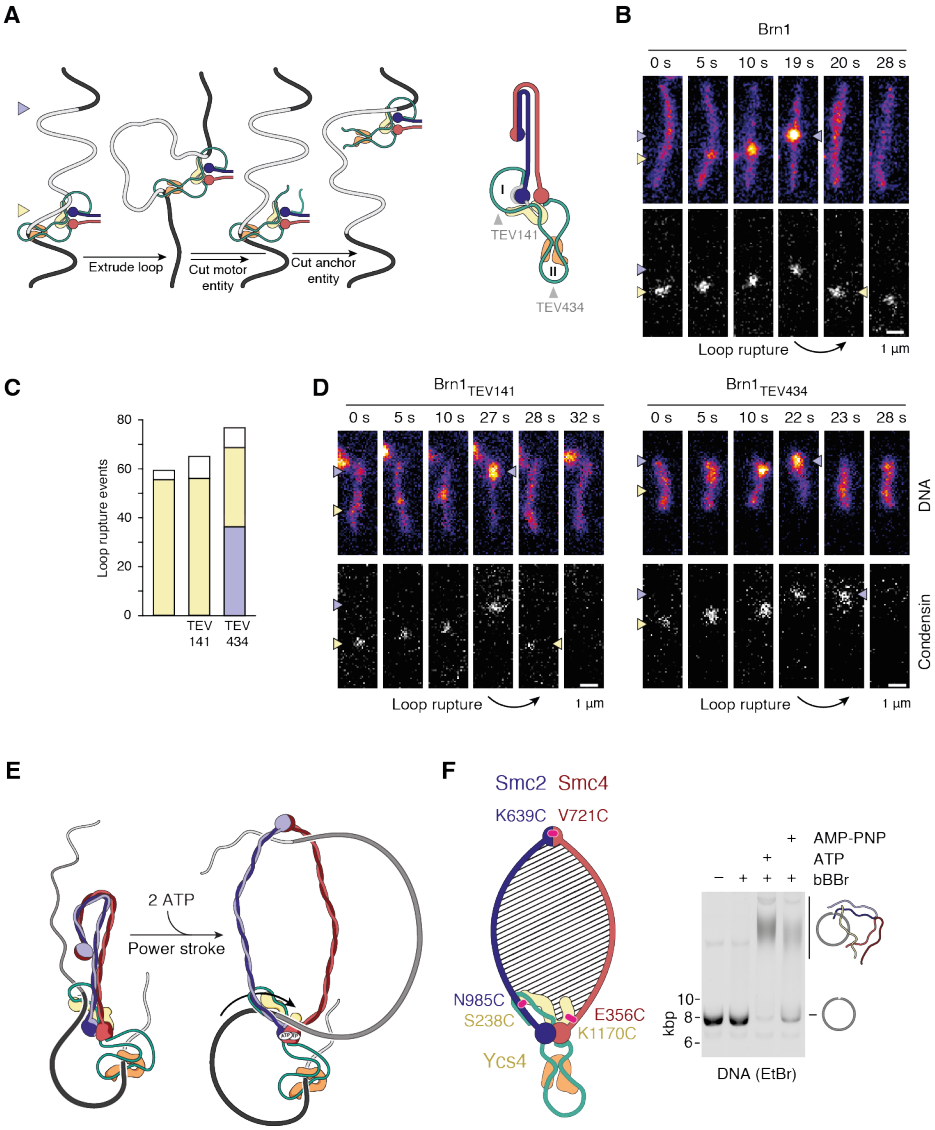


Figure 3.17. Chamber I and II are respectively associated with the motor and anchor functions of condensin. Schematic representation of the loop extrusion experiment based on the TEV cleavage of chambers I and II (**A**, left) and of the condensin chambers I and II (**A**, right). The TEV sites are highlighted. Loop extrusion experiments showing the SitOx orange labelled DNA (top panel) and the fluorescently labelled condensin (bottom panel) without Brn1 cleavage (**B**), or with chambers I and II cleavages (**D**). Statistics associated with the loop extrusion experiments (**C**). Schematic representation of condensin power-stroke showing a loop formed in between the Smc coiled coils as a result of ATP binding (**E**). Cross-linking experiments showing the topological entrapment of DNA into the ATP-dependent compartment formed by the Ycs4 subunit and by the Smc coiled coils. This compartment and the cross-linking sites are schematically represented (**F**). Figure adapted from Shaltiel et *al.*, 2022. The cross-linking and loop extrusion experiments were performed by Dr. Indra Shaltiel.

As I described in the previous sections, DNA is present in chamber I – or motor chamber – in the presence and the absence of ATP (**Figure 3.16**), as attested by my structure and by the cross-linking data. In addition, DNA can be docked into chamber I of the

ATP-free apo condensin structure with high confidence. This docking is in concordance with a low-resolution DNA-bound ATP-free structure described that I describe in **Chapter 3** (see also **Discussion**).

A comparison between the ATP-free DNA-free and the ATP-bound and DNA-bound states suggests a molecular mechanism for the condensin motor function: ATP binding induces the opening of the coiled coils concomitantly with the bending of Ycs4 onto the DNA, which pushes the DNA onto the Smc heads. The heads then make numerous contacts with the DNA (**Figure 3.17E**). Interestingly, in the course of my cryo-EM analysis, I identified an additional structural conformation of the core in which Ycs4 bends to a greater extent onto the DNA and in which Brn1 HTH slips downwards at the Smc2 neck.

Altogether, these rearrangements – which I refer to as "power-stroke" – induce a tilt in the DNA that leads to the formation of a DNA loop enclosed in the coiled-coil lumen. The formation of this loop was verified by the concomitant cross-linking of the hinge interface ($Smc2_{K639C}$-$Smc4_{V721C}$) and the interfaces between Ycs4 and the Smc subunits ($Smc2_{N985C}$-$Ycs4_{S238C}$ and $Smc4_{E356C}$-$Ycs4_{K1170C}$) described by my structure (**Figure 3.17F**). Cross-linking resulted in the retainment of the DNA in the upper coiled-coil compartment after protein denaturation in a strictly ATP-dependent manner. In line with my structural predication, this result demonstrates the formation of a DNA loop in the Smc compartment as a result of ATP-dependent rearrangements.

e. Comparison to the ATP-bound DNA-free state

The coiled coils display various opening angles in the cryo-EM classes of the ATP-bound DNA-free state (PDB:6YVV), ranging from rod-shaped, to mildly open at the head-proximal region, to presumably fully open, even though the later could not be visualized by cryo-EM due to the flexible nature of the coils in this state. Unfortunately, the opening

angle of different particles could not be directly correlated with the nucleotide state, since the medium resolution of the structure (7.6 Å) did not allow the visualization of the nucleotide occupancy at the binding sites. A comparison of the ATP-bound DNA-free and the ATP-bound DNA-bound states by aligning the RecA-like lobes of Smc4 shows a bend of the Smc2 coiled coils of about 67° at a flexure point at the beginning of the coils (**Figure 3.18A**). In both states, the Smc2 signature motif aligns with the Smc4 Walker A and B motifs and the Smc4 signature motif aligns with the Smc2 Walker A and B motifs, as expected from an ATP-engaged state. As a result, the alignment of the RecA-like and helical lobes of each head domain is almost perfect between the two states. The orientation of the Q-loop – the role of which has been discussed in the previous section – could not be assessed with certainty in the ATP-bound DNA-free state. Thus, whether the coiled-coil bend is causally related to DNA binding remains unclear. The overlapping positions of the Ycs4 and Ycg1 subunits that bind to the heads in the DNA-bound and in the DNA-free engaged states, respectively, suggest that binding of these subunits is mutually exclusive (**Figure 3.18B**). A steric clash is similarly observed when comparing the ATP-bound with the ATP-free apo states (Lee et *al.*, 2019).

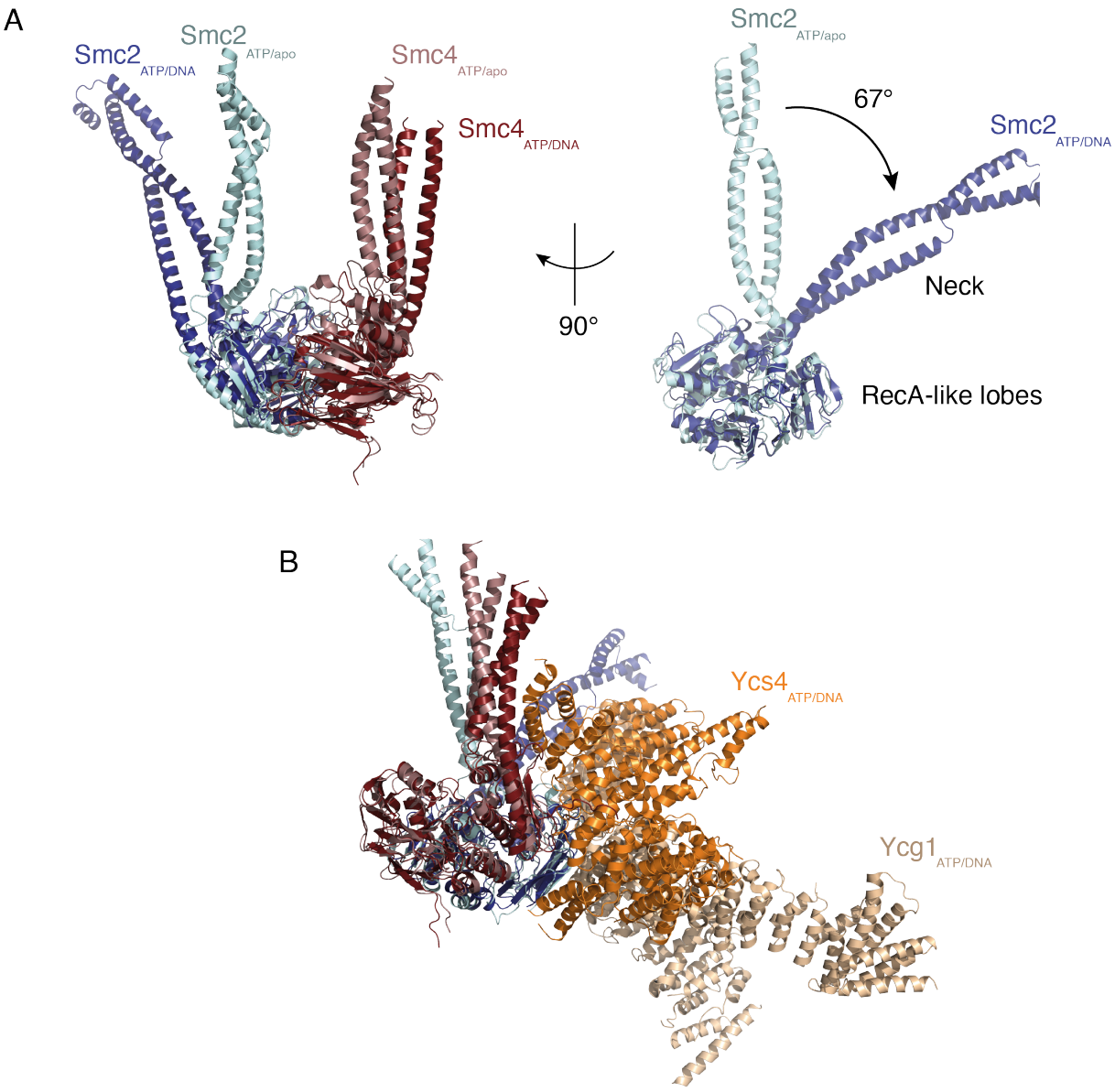


Figure 3.18. Comparison of the ATP-bound apo *Sc* condensin (PDB:6YVV, Lee et al., 2020) and the ATP- and DNA-bound *Sc* condensin (PDB:7QEN, this work). Aligned atomic models of the Smc subunits of the ATP-bound apo state (PDB: 6YVV) and the DNA- and ATP-bound state (PDB: 7QEN) of *Sc* condensin, taking Smc2 RecA-like lobe as a reference point. Upon DNA binding, Smc2 coiled coils undergo an 67° opening whereas the Smc4 protein does not show a massive rearrangement (**A**). The binding of Ycs4 and Ycg1 subunits with the Smc ATPase heads, respectively in the ATP- and DNA-bound and in the apo ATP-free *Sc* condensin states, are mutually exclusive (**B**).

Chapter 3

ATP-free and DNA-bound structures of Condensin

In **Chapter 2**, I proposed a "power-stroke" mechanism by which a DNA loop is actively fed into the Smc coiled-coil compartment. DNA is held within condensin chamber I throughout the power-stroke movement, while ATP-driven coiled coil opening and Ycs4 molecular rearrangements actively tilt the DNA onto a binding platform formed upon Smc head engagement. I described biochemical experiments that support this mechanism. In this chapter, I describe my attempts to visualize another important step in the condensin ATPase cycle by obtaining a condensin structure of the DNA-bound ATP-free state.

I. Attempt to resolve the ATP-free DNA-bound state of *Sc* condensin

a. Sample preparation

In order to obtain a structure of the DNA-bound ATP-free condensin, I initially used *Sc* condensin holo-complex and similar buffer conditions as I had used for the ATP- and DNA-bound complex. I mixed equimolar ratio of condensin and of 75-bp dsDNA (1 µM) on ice and applied the sample onto a R2/2 Ultrafoil cryo-grid. In order to reduce unwanted effects of the complex binding to the grid support, I pre-treated the gold support of the grids with a mono-layer of thiolated PEG (Zhang et al., 2021). I then cleaned the grids by glow discharging before a 72-hour incubation in (11-mercaptoundecyl) hexaethylenglycol (see methods for details). This improved the spreading of the ice over the grid squares. To avoid over-saturation of the holes with particles, I skipped the waiting time between sample application and blotting.

b. Challenges and strategies to resolve the ATP-free DNA-bound structure of *Sc* condensin

I faced several limitations during the collection and the analysis of the data set obtained for the ATP-free DNA-bound condensin complex. **Figure 4.1** summarizes the analysis pipeline. Although PEGylation of the grid led to a better spreading of the sample, the latter appeared too thick in the majority of the squares, and only very few regions were suitable for data acquisition. I could therefore acquire only a small data set of 4,269 micrographs. Second, contrary to what had been the case for the ATP- and DNA-bound data set, application of the Topaz neural network algorithm did not drastically increase the number of picked particles in comparison with the training set, leading to only a relatively small number of particles (33,663 particles) that resulted into a 3D density resolved at only ~10 Å (**Figure 4.1C**). Third, analysis of the angular distributions of the projections revealed a biased orientation (**Figure 4.1C**), that could either result from (i) some missing experimental projections (i.e., the complex would be preferentially oriented on the grid) or (ii) projections being incorrectly assigned during refinement. The latter option could result from the strong signal that emerged from the coiled-coil rods. To improve the quality of the density, I subtracted the signal from the coiled coils by masking them out during 3D classification. This strategy nevertheless failed to improve the density. This is possibly explained by the small size of the 'core' module that lacks the coiled coils, which would not generate enough signal for the alignment of the particles.

c. Analysis of the density and emphasis on the power-stroke mechanism

The low-resolution density resembles the apo ATP-free structure (PDB: 6YVU, Lee et al., 2020). As a result, I manually docked this model into the electron density (**Figure 4.2**).

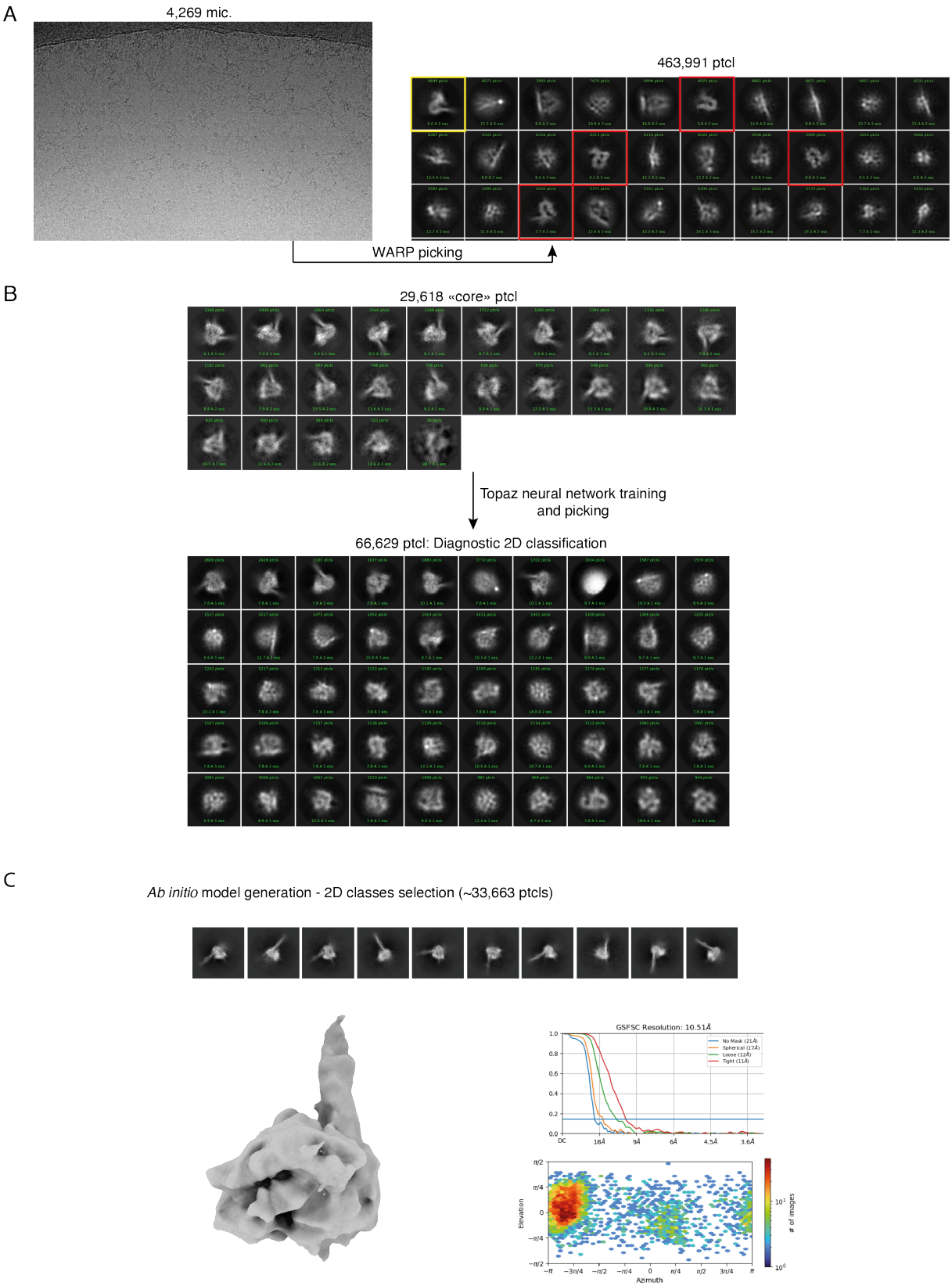


Figure 4.1. Analysis pipeline of the data set corresponding to the DNA-bound ATP-free *Sc* condensin holo-complex. Example of micrograph of the DNA-bound ATP-free *Sc* condensin holocomplex data set (**A**, left) and diagnostic 2D classification of the WARP picked particles (**A**, right). 2D classifications of the particles used as input for Topaz training (**B**, top) and resulting from the Topaz picking (**B**, bottom).

Generation of *ab initio* 3D model corresponding to the DNA-bound ATP-free *Sc* condensin holo-complex (**C**, left). The Fourier Shell Correlation plot and the plot corresponding to the orientations of the particle views are shown (**C**, right).

Interestingly, a region of the density, located bellow the N-terminus of Ycs4, is not accounted for by the structure of the apo complex and could potentially correspond to DNA (**Figure 4.2.B**). Thus, the DNA would be positioned inside the kleisin chamber I. The N-terminal lobe of the Ycs4 subunit seems to bend further than it does in the apo ATP-free state, which could be a consequence of the interaction of Ycs4 with DNA. The region of the electron density that corresponds to the Smc4 head is incomplete and the density that would correspond to the Brn1 Winged-Helix domain is absent (**Figure 4.2B** and **C**). Despite its incompleteness and its low resolution, this density is consistent with the power-stroke mechanism proposed in **Chapter 2**. If the additional density corresponded to DNA, then DNA would be contained in chamber I, in line with the cross-linking data described in **Chapter 2**. The DNA would be clamped by the Ycs4 N-terminal lobe and could then be flipped in-between the open coiled coils of the Smc subunits upon ATP-binding. Similarly to what I show in **Figure 3.16** of **Chapter 2**, most of the residues that make contacts with the DNA in the ATP-bound state would also be correctly positioned to do so in the configuration that is suggested by this low-resolution structure. Nevertheless, a higher resolution cryo-EM structure is required to further support the conclusions regarding the conformation of the DNA-bound ATP-free state of condensin.

II. Attempts to resolve the ATP-free DNA-bound state of *Ct* condensin

a. *Ct* condensin holo-complex and *Ct* condensin ΔYcg1 complex

i. *Ct* holo-complex sample preparation

Because of its thermophilic nature, Ct is an organism of choice for structural studies of DNA-bound protein complexes. As described in Chapter 1, Ct condensin, even in the

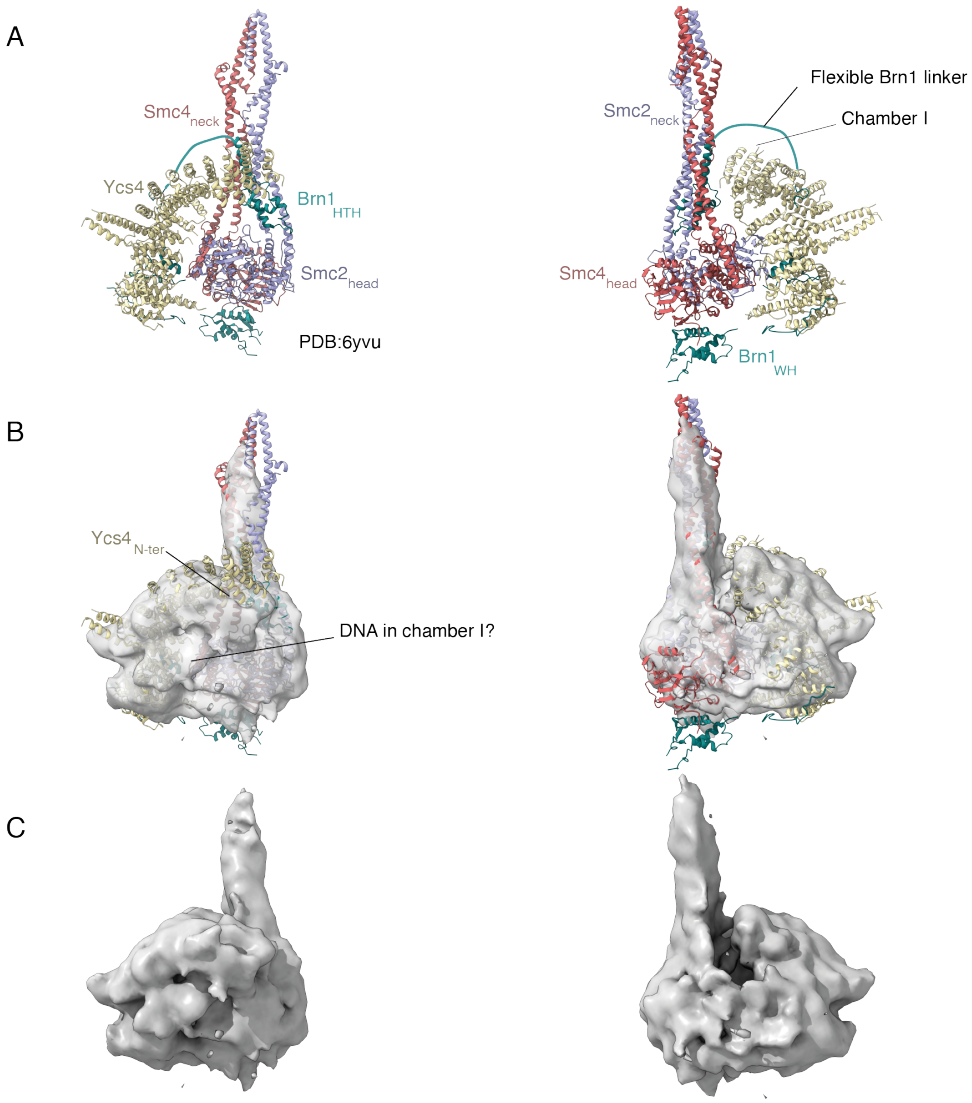


Figure 4.2. Interpretation of the DNA-bound ATP-free *Sc* condensin holo-complex electron density. The atomic model of the apo ATP-free *Sc* condensin holo-complex (PDB: 6YVU) is docked into the electron density of the DNA-bound ATP-free *Sc* condensin holo-complex (**A** and **B**). Unassigned density below the N-terminal lobe of Ycs4 could correspond to DNA, that would fit into chamber I (**A**, **B** and **C**). The unresolved regions of the kleisin are drawn with a green line.

absence of its Ycg1 subunit, maintains its DNA-binding activity. Therefore, I reasoned that this complex might be of use to obtain higher-resolved electron density maps to build the structure of the ATP-free DNA-bound condensin complex. The *Ct* condensin holo-complex could be purified at a higher concentration than the *Sc* condensin holo-complex (14 mM compared to 8 mM). This allowed me to supplement condensin cryo-EM samples with detergent, which can improve the ice quality while protecting the sample from the denaturing air/water interface (Li et al., 2021). Using equimolar

protein:DNA ratios (4 µM), I supplemented the sample with 0.02% (m/v) β-octyl glucoside. This resulted in an improvement of the ice quality and in an ideal particle distribution (**Figure 4.3A**).

ii. Challenges and strategies to resolve the structure of *Ct* condensin holo-complex

I acquired a data set of 8,762 micrographs using a Krios microscope equipped with a K3 electron detector and obtained an initial model of *Ct* condensin holo-complex. The analysis pipeline is briefly described in **Figure 4.3**. Similarly to the case described above for the *Sc* condensin holo-complex, I observed a preferential distribution of the views (**Figure 4.3B**). In this instance, the bias did not seem to result from a preferential positioning of the sample on the cryo-grid, since I could visualize numerous orientations and views of the complexes during analytical 2D classification of the particles used to generate the 3D *ab inito* model (**Figure 4.3C** and **D**). Instead, the anisotropy of the density resulted in information loss at the region that corresponds to the Smc2 head domain. I used different approaches to tackle this problem, based on masking and successive rounds of 3D classifications with or without keeping the Euler angles fixed. In addition, I tried different neural network picking strategies with or without performing preliminary algorithm training to either be the least stringent as possible and pick the majority of the particles, or to strictly select for particles views that appeared rare in the 2D classification process. None of these approaches proved promising. Instead, 2D classes showed 'blurred' protein regions (**Figure 4.3D**), suggesting an inherent flexibility of the complex. These data suggest that the Smc2 head domain undergoes a flexible motion relative to the rest of the core; a motion that has been previously reported in the case of the *Sc* apo ATP-free structure (Lee et al., 2020).

iii. Analysis of the *Ct* condensin holo-complex density

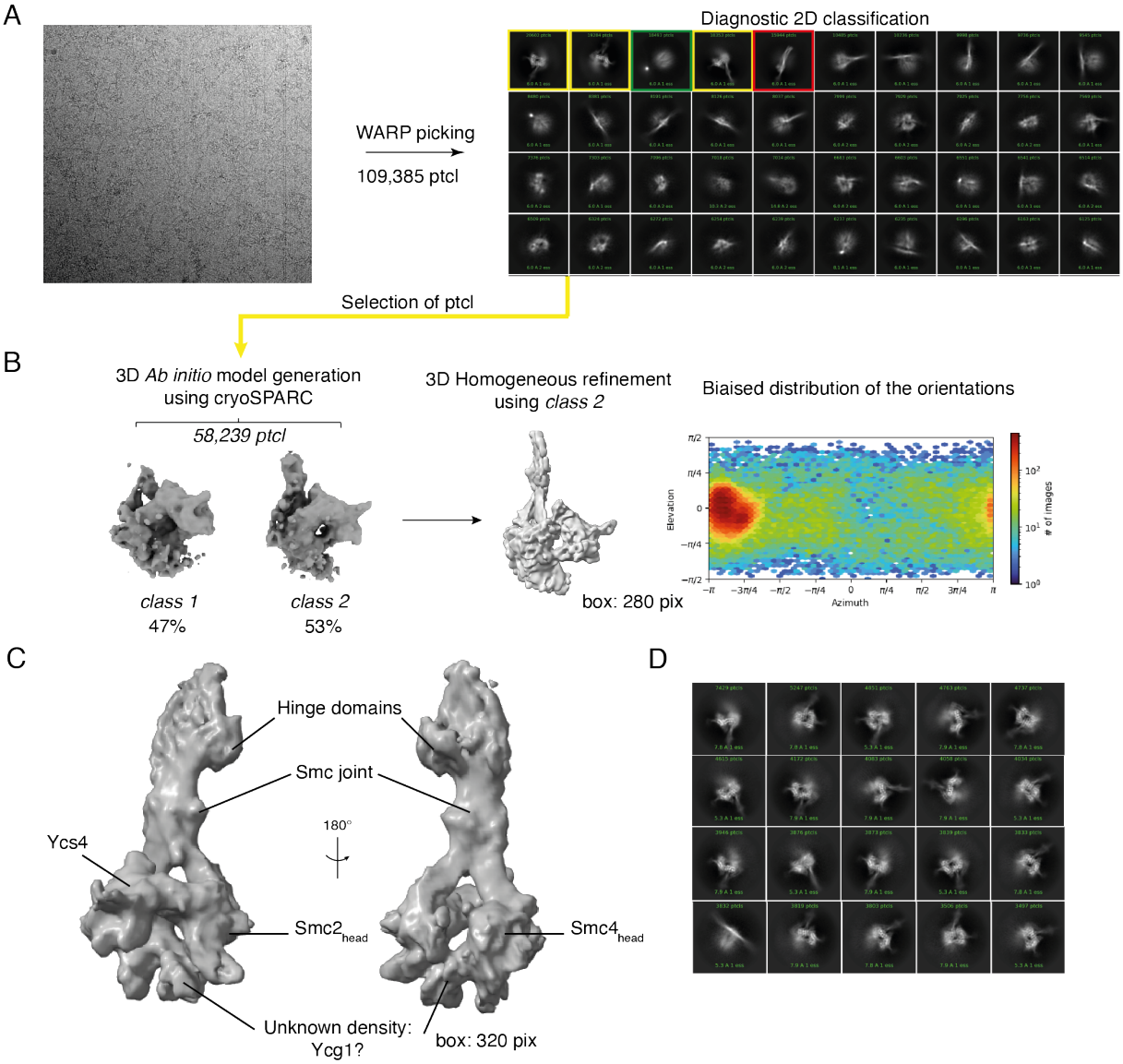


Figure 4.3. Analysis pipeline of ATP-free *Ct* condensin holo-complex data set. Example of micrograph from the ATP-free *Ct* condensin holo-complex data set (**A**, left) and diagnostic 2D classification of the WARP picked particles (**A**, right). Generation of 3D *ab initio* and refined 3D model corresponding to the *Ct* condensin holo-complex (**B**, left). The plot corresponding to the orientations of the particle views shows a biased distribution (**B**, right). Annotated electron density of the ATP-free *Ct* condensin holo-complex (**C**) and 2D classification of the particles used to generate this model (**D**).

The density is composed of ~30,000 particles and reached a resolution of 7-8 Å. Despite its low resolution, the density provided some insights. Firstly, DNA does not seem to be bound to the *Ct* condensin holo-complex, as seen in both 2D classes and by the 3D density. Secondly, the ATPase heads seem to be much further apart than in the apo ATP-free state but closer than in the 'bridged' apo ATP-free *Sc* condensin state (Lee et

al., 2020) (**Figure 4.4.A**, **Figure 1.4** of **Introduction**). The new density therefore corresponds to a previously undescribed condensin conformation, in which the ATPase heads are disengaged and spread apart. Some unidentified density in between the heads could possibly correspond to the Ycg1 HEAT-repeat subunit. In contrast to the data set that corresponds to the ATP- and DNA-bound complex discussed in **Chapter 2**, I did not observe any 2D classes that correspond to the Ycg1 subunit. This could be possibly explained if (i) Ycg1 were bound to the core and not solely flexibly tethered to it through the kleisin or (ii) Ycg1 dissociated from the complex during the protein purification process. To exclude the latter possibility, I used mass photometry (**Figure 4.4B**). The presence of a monodisperse species with a molecular weight in the range of the theoretical value of *Ct* condensin holo-complex (MW: 680 kDa) attested the integrity of the complex. The putative insertion of the Ycg1 subunit in between the Smc2 and Smc4 heads would explain the lower basal ATPase activity of the *Ct* holo-complex in comparison to the *Ct* ΔYcg1 complex (Shaltiel et al., 2022). The absence of bound DNA could suggest that this condensin conformation is not compatible with stable DNA binding. Despite the accumulation of structural knowledge about the condensin holo-complex, the condensin DNA-loading mechanism remains unclear. It is tempting to speculate that this conformation could possibly correspond to a DNA pre-loading state of condensin and could thus be an important step towards the identification a loading mechanism.

iv. *Ct* ΔYcg1 condensin complex sample preparation

As discussed in the previous section, Ycg1 seemed to be positioned in-between the non-engaged Smc heads, locking the complex in a conformation that cannot engage with DNA. I therefore used a *Ct* ΔYcg1 condensin complex to obtain a DNA-bound *Ct* complex in absence of ATP.

Although I used similar sample preparation conditions as the ones used for the *Ct* holo-complex, the *Ct* ΔYcg1 condensin sample was deteriorated during plunge freezing and I only observed broken or unfolded particles after 2D classification (not shown).

Following the hypothesis that the *Ct* holo-complex conformation described in the previous section corresponds to a DNA pre-loading state of condensin, I hypothesized that passage through an ATPase cycle is required for the complex to load onto the DNA.

To this end, I incubated the *Ct* holo-complex with DNA and ATP, before adding EDTA to chelate the Mg^{2+} ions present in the dilution buffer. This approach was equally unsuccessful, as the analysis of the micrographs only revealed 2D classes that correspond to broken/unrecognizable particles (not shown). This suggests that this sample preparation condition was not suitable.

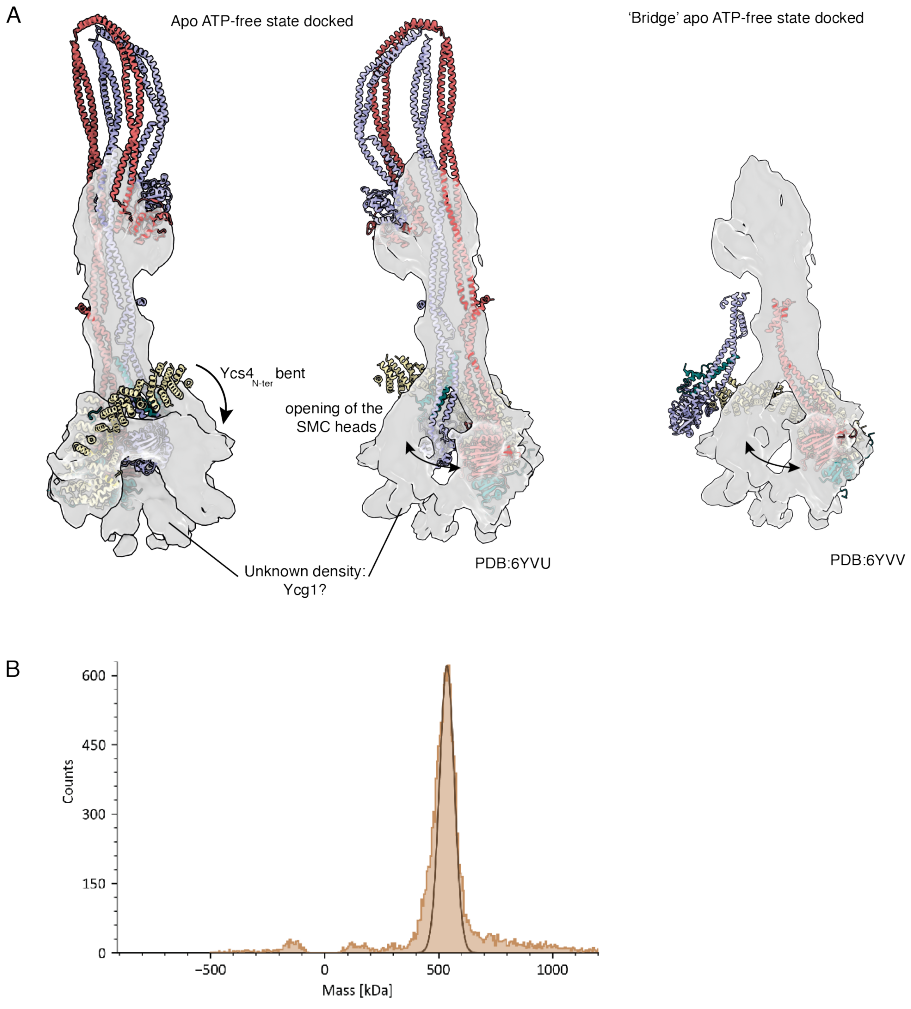


Figure 4.4. Interpretation of the ATP-free *Ct* condensin holo-complex electron density. The atomic models of the apo ATP-free *Sc* condensin holo-complexes (PDB: 6YVU, 6YVV) are docked into the electron

density of the ATP-free *Ct* condensin holo-complex (**A**). None of these models can be fully accommodated by the electron density, suggesting that it corresponds to an unknown condensin conformation. Mass photometer profile of the *Ct* condensin holo-complex attesting the integrity of the complex (**B**).

b. Condensin « bonsai » complex

i. Sample preparation

I next attempted to use a *Ct* condensin "bonsai" protein construct to obtain a structure of the ATP-free DNA-bound condensin complex. The bonsai protein construct lacks the majority of the Smc2 and Smc4 coiled coils and the C- and N-terminal extremities of Smc2 and of Smc4 coils are covalently joined at the "joint" regions. The shorter Smc2 protein (1-257/917-1179) is fused at its C-terminal extremity to a minimal Brn1 kleisin that lacks the Ycg1 interacting region (Δ515-634). The complex also contains a shorter Smc4 (263-510/1328-1542) and the complete Ycs4 subunit. Following the hypothesis that Ycg1 binding to the ATPase Smc heads would prevent the complex to load onto DNA, the Ycg1 subunit is absent from the "bonsai" engineered construct (**Figure 4.5**).

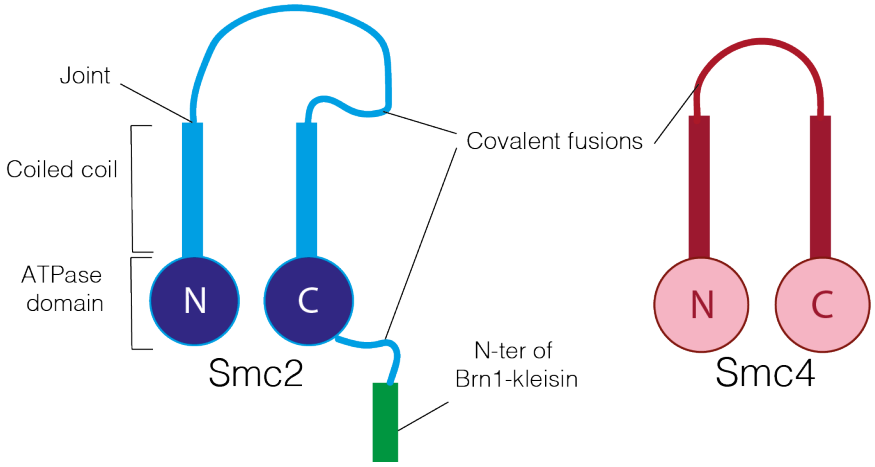


Figure 4.5. Schematic representation of the *Ct* condensin "bonsai" protein construct. The N- and C-terminal extremities of Smc2 (1-257/917-1179) and Smc4 (263-510/1328-1542) are covalently tethered at the "joint" regions. The majority of the coiled coils is deleted. The N-terminus of the kleisin Brn1 is covalently linked to the C-terminal extremity of Smc2. The Brn1 kleisin is deleted from the Ycg1 interacting region (Δ515-634) and from the Ycg1 HEAT-repeat subunit. The full-length Ycs4 HEAT-repeat subunit (not represented on the schematic) is contained in the "bonsai" protein construct. This protein construct was cloned by Dr. M. Hassler.

In order to obtain the DNA-bound ATP-free state of the "bonsai" condensin complex, I incubated it with an equimolar ratio of DNA (1μM), diluted in a similar buffer as the

ATP- and DNA-bound condensin complex (see Methods), and supplemented the sample with 0.02% (m/v) β-octyl glucoside in order to obtain a thin and homogeneous ice layer.

ii. Analysis of the cryo-EM density of the condensin "bonsai"

The analysis pipeline is described in **Figure 4.6**. The final density resulted from 63,339 particles and is resolved at ~6 Å (**Figure 4.6E**). The electron density clearly shows Ysc4 bound to the Smc4 head domain. No DNA is visible, despite being present during the preparation. The Smc2 head is absent from the density, which suggests that it dissociates from the rest of the complex in absence of the full-length coiled coils and in absence of ATP, which induces ATPase head dimerization. Some 2D classes might correspond to the individual Smc2 head, but the particles were too small to be further processed. The *ab initio* 3D model of the bonsai complex was isotropic, in contrast to what I had observed with the previously described *Ct* ATP-free apo structures. This is consistent with the absence of the Smc2 subunit, which seemed to be the source of the anisotropy. The fact that the complex lacks Smc2 also indicates that this region of condensin might be flexible in absence of ATP. The subsequent 3D classifications (not shown) were not discriminant and therefore did not result in an improvement of the electron density resolution.

Given the medium resolution of the density (6 Å), I could fit the crystal structure of the *Ct* Smc4 head (PDB: 6QJ2) and the Alpha fold prediction (Jumper et al., 2021) of the Ycs4 subunit (AF-G0SB82) in to the density. For Ycs4, the prediction was preferred over the available crystal structure, as the latter was obtained from a construct depleted of flexible loop regions (Hassler et al., 2019).

The density reveals an interface composed of the Ycs4$^{\text{KG-loop}}$ and of the Smc2$^{\text{W-loop}}$, similar to the one observed in the apo ATP-free structure (**Figure 3.10 – Chapter 2**). This density reveals an additional contact between the Smc4 catalytic pocket and the HEAT-repeat 15 helices of Ycs4, even though the resolution and the possible flexibility of the

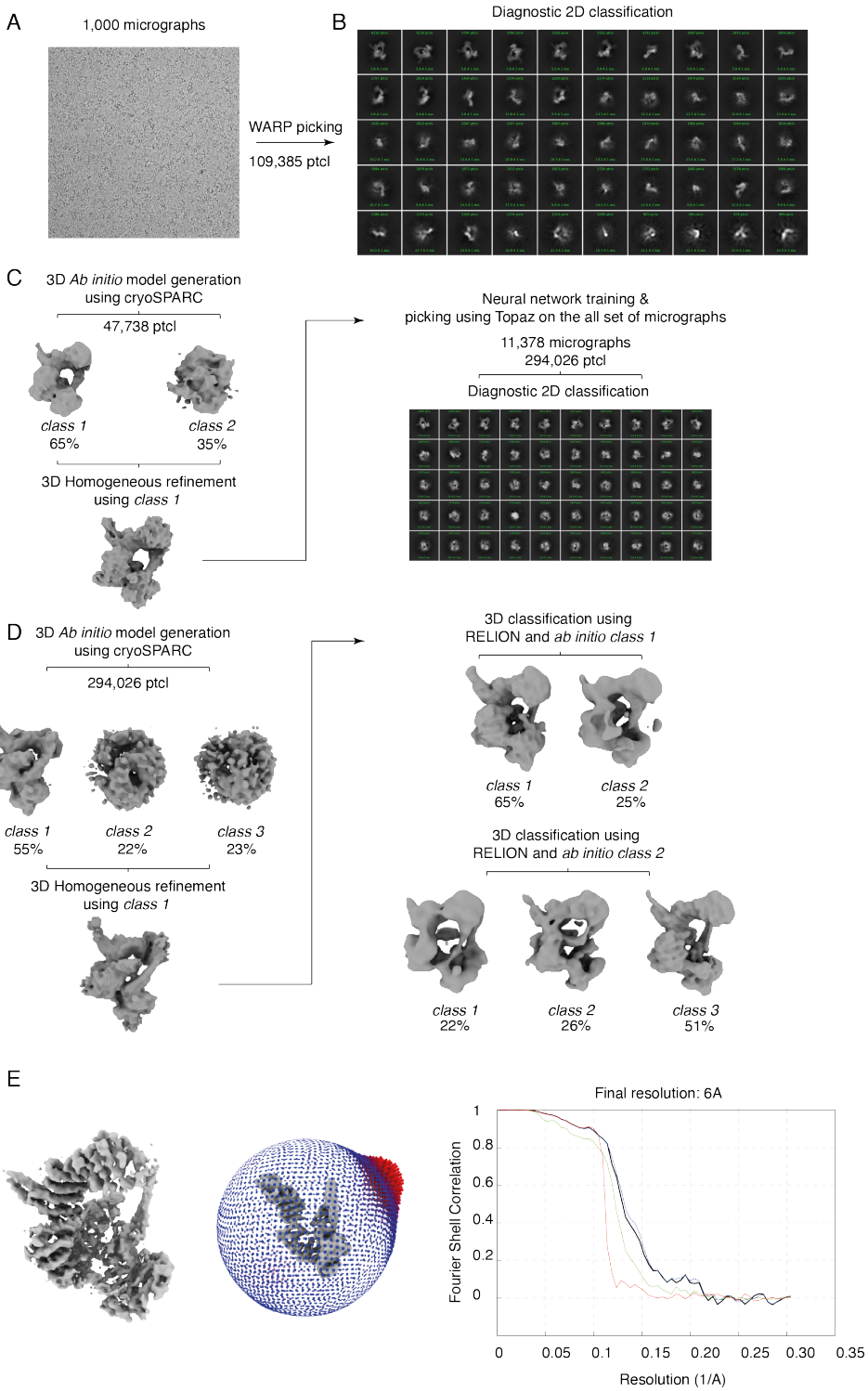


Figure 4.6. Analysis pipeline of the *Ct* condensin "bonsai" data set. Example of micrograph from the *Ct* condensin "bonsai" data set (**A**). Diagnostic 2D classification of the WARP picked particles (**B**). Generation of *ab initio* 3D model of the *Ct* condensin "bonsai" (**C**, left) and topaz neural network training and picking pipelines (**C**, right). Generation of *ab initio* 3D model of the *Ct* condensin "bonsai" using topaz picked particles (**D**, left) and subsequent 3D classification steps (**D**, right). Final density of the *Ct* condensin "bonsai" resolved at 6 angstroms according to the FSC gold standard. Distribution of the orientations of the views are plotted onto the density (**E**).

region did not allow the visualization of the contact at the single-residue level. According to the density, and in accordance with the Alphafold model, helix 15b contacts the Walker A and B motif-proximal regions of Smc4 (**Figure 4.7**). Interestingly, the sequence of helix 15b is conserved among eukaryotes (**Figure 4.7B**). I previously discussed in **Chapter 2** that the HEAT-repeat 15 helices are differentially positioned in the ATP-free apo state in comparison to the ATP- and DNA-bound state. If these helices insert in between the ATPase heads in absence of DNA and ATP, they shift downwards and interact with the inferior surface of the heads when ATP and DNA are bound to condensin (**Figure 3.10** – **Chapter 2**).

To investigate the interaction dynamics of Ycs4 and Smc4, I performed a gel filtration experiment using a *Ct* condensin trimeric version composed of the full-length Smc2 and Smc4 proteins in complex with a minimal version of the Brn1 kleisin ($Brn1_{NC}$) that lacks the HEAT-repeat subunit interaction sites (Δ225-634) and a construct corresponding to *Ct* Ycs4 and its interacting kleisin fragment (225-512) (**Figure 4.8**). I mixed both, the trimeric condensin and the Ycs4-Brn1 subcomplexes, in equimolar ratios (15 µM) before loading them onto a Superose 6 increase 3.2/300 gel filtration column. I analyzed the collected fractions by SDS-PAGE. This experiment showed that Ycs4 binding to the trimeric complex is abolished in the presence of ATP, in accordance with the cryo-EM structure of the ATP-bound apo condensin holo-complex (pdb: 6YVV; Lee et al., 2020). A biochemical and structural study revealed that the disruption of the interaction between $Ycs4^{KG\text{-}loop}$ and $Smc2^{W\text{-}loop}$, as a result of ATP binding to the Smc4 catalytic site, originated from the molecular rearrangement of the Smc4 head domain (Hassler et al., 2018). Thus, the structural insight obtained from the "bonsai" condensin construct adds another layer of complexity to a scenario where ATP and Ycs4 binding to the Smc4 head are tightly interconnected and are mutually exclusive, also because they are both engaged with the catalytic site of Smc4.

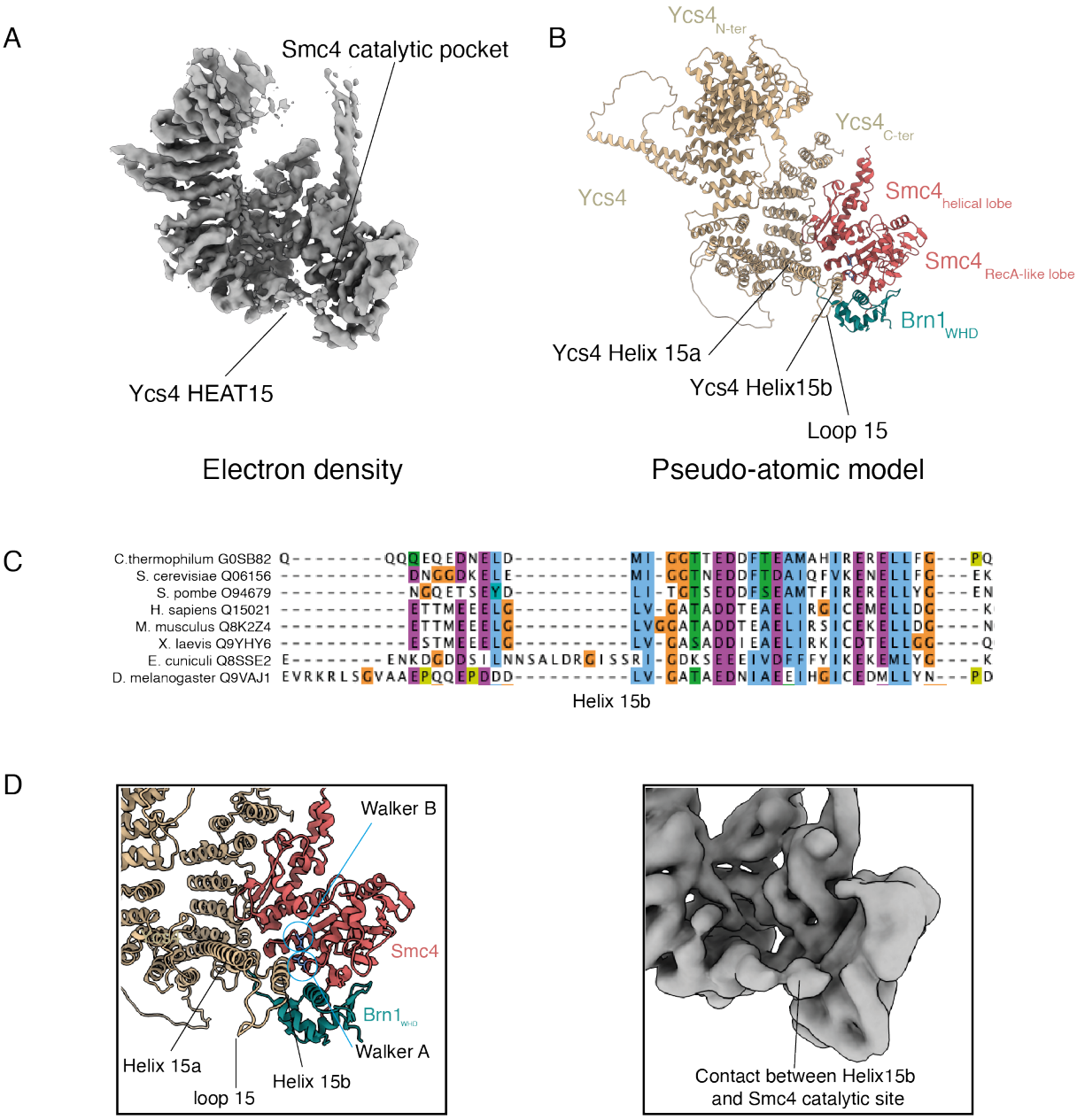


Figure 4.7. Interpretation of the *Ct* condensin "bonsai" electron density. Electron density corresponding to the *Ct* condensin "bonsai" (**A**), the interaction region between the Ycs4 HEAT15 and the catalytical pocket of Smc4 is shown. Pseudo-atomic model of the *Ct* condensin bonsai composed of the AlphaFold simulation (AF-G0SB82) of the *Ct* Ycs4 and the co-crystal structure of Smc4 head domain and Brn1 C-terminus (PDB:6QJ2) (**B**), the interaction region between the Ycs4 HEAT15 and the catalytical pocket of Smc4 is shown. ClustalX alignment of the protein sequences of HEAT15 helix b showing a high degree of conservation across eukaryotic species (**C**). Interaction interface between the Ycs4 HEAT15 and the catalytical pocket of Smc4 is shown on the pseudo-atomic model (**D**, left) and the electron density (**D**, right).

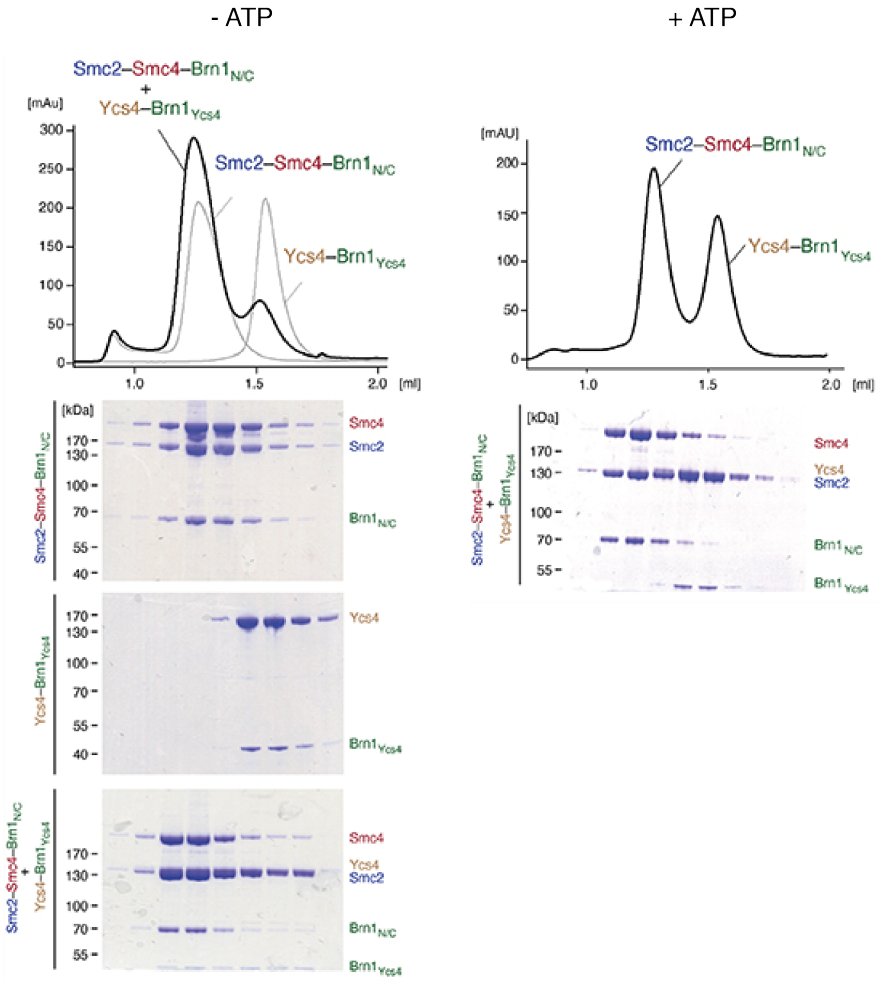


Figure 4.8. Gel filtration experiment probing the *Ct* Ycs4 binding to the *Ct* Smc2-Smc4-Brn1$_{NC}$ trimeric complex. Gel filtration profiles and SDS-PAGE corresponding to the *Ct* Ycs4 and *Ct* Smc2-Smc4-Brn1$_{NC}$ trimeric complex binding experiments in absence (**left**) and in presence of ATP (**right**). *Ct* Ycs4 binding to the *Ct* Smc2-Smc4-Brn1$_{NC}$ trimeric complex is incompatible with ATP binding. This figure is adapted from Lee et al., 2020.

Chapter 4

Towards the understanding of Condensin chromatin association

In the previous chapters, I presented structural and biochemical data characterizing the interaction of condensin with DNA. However, in a physiological context, condensin functions on genomic DNA, which is packaged by histones into chromatin and binds to a myriad of chromatin-associated factors. In this last chapter of the results section, I will discuss preliminary experiments that aim to explore the interaction of condensin with chromatin factors and DNA sequences which might play a functional role in its recruitment. First, I will describe interaction studies with condensin and the general transcription factor TFIIIC and with di-nucleosome substrates. Then, I will describe experiments that aim to identify a DNA sequence specificity of the peripheral module of condensin.

I. Interaction between condensin and TFIIIC

Mapping the binding sites of condensin in the budding yeast genome revealed condensin enrichment at RNA-polymerase (RNA-pol) III transcribed regions (i.e., 5S RNA- and tRNA-coding genes). In addition, the core transcription factor TFIIIC was shown to co-immuno-precipitate with condensin from both, budding yeast and human cells (Haeusler et al., 2018, Yuen et al., 2017). I therefore investigated a potential *in vitro* interaction between *Sc* condensin and *Sc* TFIIIC complexes. The latter were purified by Wolfram Seifert Davila.

a. DNA interacts with individual *Sc* TFIIIC and *Sc* condensin complexes

Sc TFIIIC is a ~548-kDa six-subunit protein complex. The six subunits form τA and τB subcomplexes, which bind to highly conserved sequence motifs located in intra-genic promoter regions of tRNA coding genes called "A box" and "B box", respectively. *Sc* TFIIIC τA is composed of τ131, τ95 and τ55 and flexibly tethered to the τ138 protein of τB, which in addition contains τ91 and τ60 (Male et al., 2015). Purified *Sc* TFIIIC holo-complex was obtained by a three-step purification strategy that included (i) affinity purification via a poly-histidine tag, (ii) affinity purification via a Strep-tag, followed by (iii) ion exchange chromatography (**Figure 5.1.**, see Methods for details).

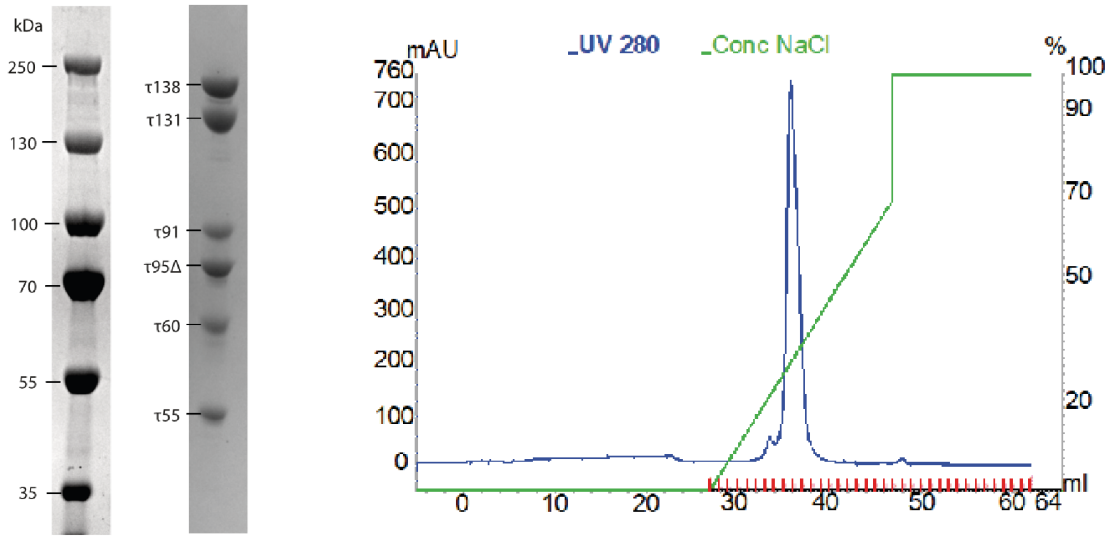


Figure 5.1. Protein purification of the *Sc* TFIIIC complex. SDS-PAGE of purified *Sc* TFIIIC (**left**). The subunits τ138, τ131, τ91, τ95Δ, τ60 and τ55 are purified with a stochiometric ratio. The anion exchange chromatogram (**right**) indicates a monodisperse peak during the last step of the purification indicating that the sample is pure and homogeneous. The purification was performed by Wolfram Seifert Davila in the laboratory of Dr. C. Mueller at EMBL, Heidelberg.

I first asked whether *Sc* TFIIIC forms a complex with DNA. I used mass photometry to assess complex formation between *Sc* TFIIIC and a 120-bp double stranded DNA ligand of the gene that encodes the 37-kDa *Sc* $tRNA^{His}$ and contains the "A box" and "B box" sequence motifs required for τA and τB binding (Stillman and Geiduschek, 1984). Mass photometry has the advantages that it requires very little sample (10 µL of sample at a

concentration of 20-40 nM) and that it does not require any sample labelling (Young et al., 2018). Mass photometry measures molecular weights in a large range (40 kDa – 3,000 kDa). However, it relies on calibration with molecular weight standards, which may limit its accuracy for proteins in complex with nucleic acids (see below).

The mass photometry profile of *Sc* TFIIIC alone measured at a concertation of 40 nM shows a bi-modal distribution, which corresponds to two species (**Figure 5.2A**, pink). The 330-kDa species likely corresponds to the τA subcomplex, whereas the 540-kDa species likely corresponds to the 548 kDa *Sc* TFIIIC complex. I then mixed *Sc* TFIIIC and *Sc* $tRNA^{His}$ DNA in an equimolar ratio at a concentration of 400 nM before diluting the sample 10-fold upon application to the cover slip of the calibrated mass photometer. In addition to the molecular weight (MW) species at ~330 and ~550-kDa that I had detected for *Sc* TFIIIC alone, I observed a ~590-kDa species. The latter presumably corresponds to complexes formed between *Sc* TFIIIC and the 120-bp *Sc* $tRNA^{His}$ DNA (37 kDa)(**Figure 5.2A**, orange).

I then applied the same approach and method to assess the interaction between *Sc* condensin and DNA (**Figure 5.2B**). In the presence of *Sc* $tRNA^{His}$ DNA, I detected a 685-kDa species, which corresponds to a shift in MW of ~60-kDa compared to the signal obtained for *Sc* condensin alone, which has a theoretical MW of 625 kDa. This shift roughly corresponds to the theoretical MW of one or two *Sc* 120-bp $tRNA^{His}$ DNA substrates (37 or 74 kDa, respectively). In addition, I detected a low-intensity signal at a MW of ~1,330 kDa, which likely corresponds to the binding of two *Sc* condensins to *Sc* $tRNA^{His}$ DNA. In fact, the 120-bp *Sc* $tRNA^{His}$ DNA is sufficiently long to accommodate two *Sc* condensin complexes, which require each approximately ~35 bp for interaction at each of their DNA binding sites (see **Chapter 2**).

b. Interaction between *Sc* TFIIIC, *Sc* condensin and DNA

To test whether *Sc* condensin and *Sc* TFIIIC form a complex, I mixed both protein complexes in an equimolar ratio (400 nM) and measured the mass photometer signal at a final concentration of 40 nM. I observed two independent peaks at ~490 kDa and ~550 kDa, respectively, which suggest that the two complexes do not interact under the conditions of this assay (**Figure 5.3A**). Note that the measured MW did not accurately match the theoretical MW of *Sc* TFIIIC and *Sc* condensin (548 and 650 kDa, respectively).

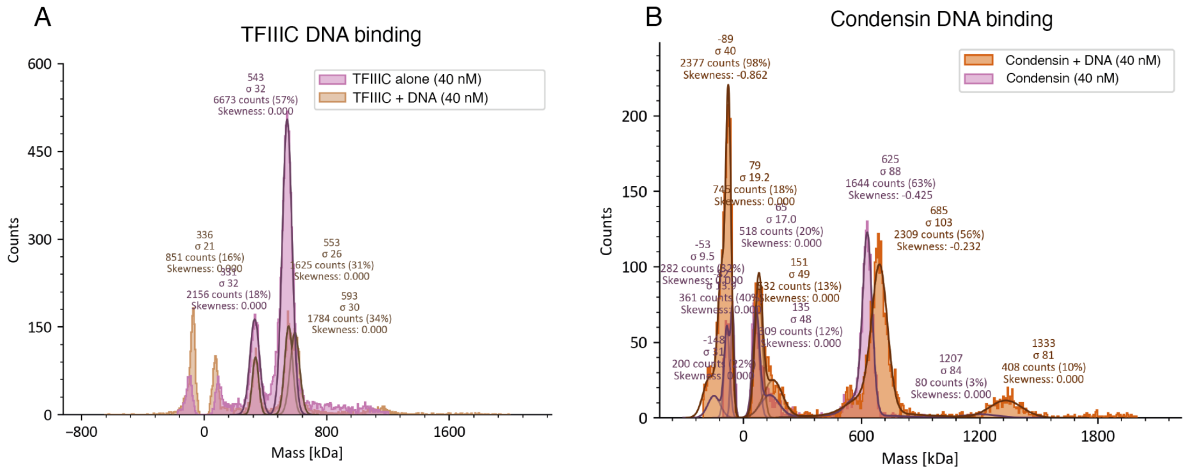


Figure 4.2. Mass photometry profiles probing the DNA-binding activities of *Sc* TFIIIC and *Sc* condensin. In absence of *Sc* tRNA^{His} DNA, *Sc* TFIIIC produces a bi-modal signal. The 330-kDa species likely corresponds to τA, whereas the 540-kDa species likely corresponds to the full *Sc* TFIIIC complex. The binding of *Sc* TFIIIC to tRNA^{His} DNA is attested by the formation of a mono-disperse species at ~590 kDa (**A**). In absence of DNA, *Sc* condensin is measured as a single species in solution (652 kDa). Upon *Sc* tRNA^{His} DNA binding, species are measured at 685 kDa and at ~1,300 kDa. These signals likely correspond to one or two *Sc* condensin bound to tRNA^{His} DNA, respectively (**B**).

Upon *Sc* tRNA^{His} DNA addition to the equimolar mix of *Sc* TFIIIC and *Sc* condensin, I observed the formation of higher molecular weight species at ~1,100, 1,700 and 2,200-kDa (**Figure 5.3B**). In addition, I detected a species at ~550 kDa, which could either correspond to single *Sc* TFIIIC or to *Sc* condensin. The periodicity of ~500-600 kDa between the different peaks is close to the theoretical MWs of both *Sc* TFIIIC or *Sc* condensin. The formation of the high-molecular weight species is DNA-dependent, yet it

was not observed when *Sc* TFIIIC or *Sc* condensin is mixed with DNA in absence of the other protein complex. Although mass photometry was unable to reveal the composition of the different high-molecular weight species, my results suggest that *Sc* TFIIIC, *Sc* condensin and *Sc* $tRNA^{His}$ DNA can assemble into complexes.

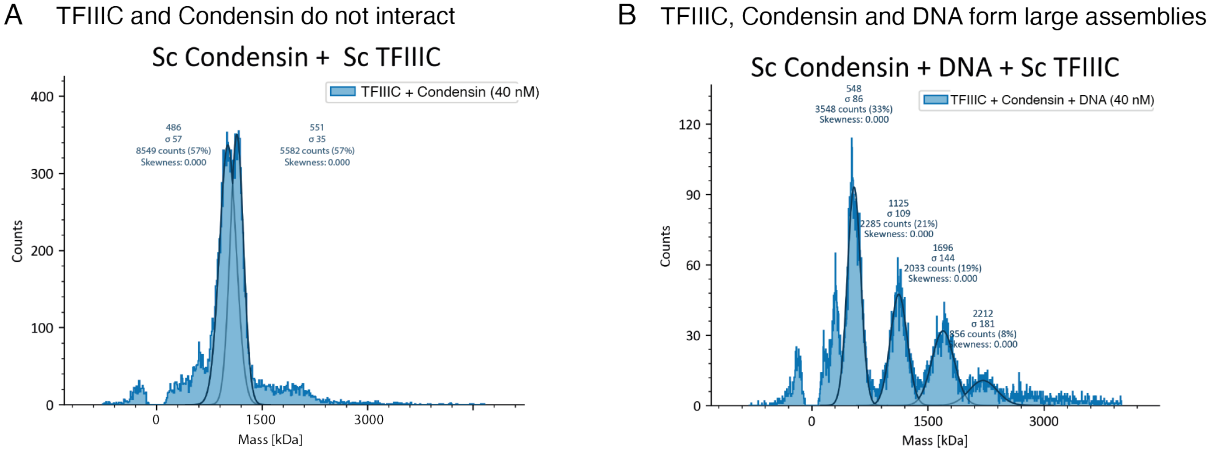


Figure 5.3. Mass photometry profiles probing the interaction between *Sc* TFIIIC and *Sc* condensin. In absence of *Sc* $tRNA^{His}$ DNA (**A**), *Sc* TFIIIC and *Sc* condensin do not interact, as attested by the formation of a bi-modal signal a 486 kDa and 551 kDa, corresponding to *Sc* TFIIIC and *Sc* condensin respectively. Upon addition of DNA (**B**), higher molecular species are formed at 1125, 1696 and 2212 kDa, likely corresponding to *Sc* TFIIIC-condensin-DNA assemblies.

Using the same buffer conditions, I prepared a cryo-EM sample to visualize the *Sc* TFIIIC-condensin-DNA assemblies observed by mass photometry (see Methods). A data set of ~6,000 micrographs was acquired with a Titan Krios 300 kV microscope equipped with a K3 direct electron detector by Wolfram Seifert Davila (**Figure 5.4A**). WARP was used to pick particles from the micrographs, following a similar pipeline as the one that I described for the processing of the data sets in **Chapters 2** and **3**. The diagnostic 2D classification led to the identification of individual classes corresponding to *Sc* condensin (**Figure 5.4B**, orange) and to *Sc* TFIIIC, which was identified either as a monomer or as a dimer (**Figure 5.4B**, blue). This analysis did, however, not provide evidence for the larger *Sc* TFIIIC-condensin-DNA assemblies detected by mass photometry. Optimization of cryo-EM conditions and alternative biochemical approaches will be required to validate and characterize these assemblies.

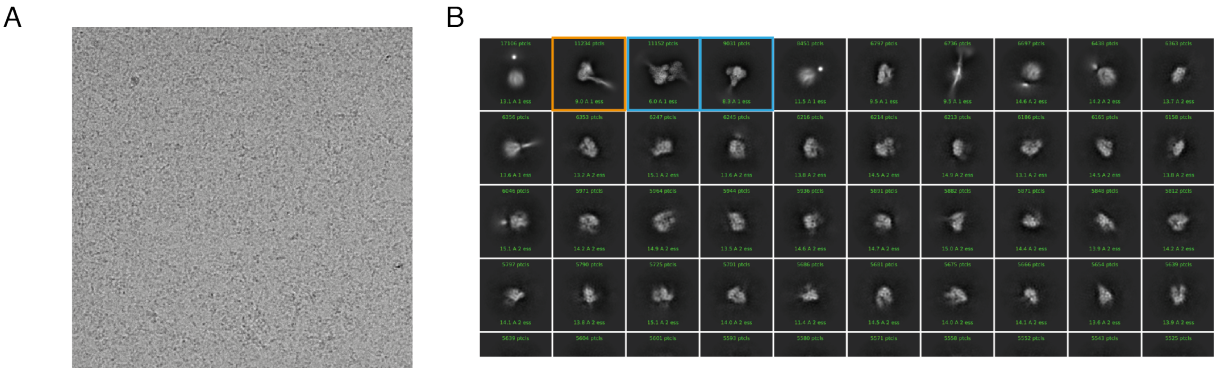


Figure 5.4. Cryo-EM study of a Sc-TFIIIC-condensin-DNA assemblies. Example of micrograph collected from a sample containing equimolar ratio of Sc TFIIIC, Sc condensin and Sc tRNAHis DNA (A). Diagnostic 2D classification showing classes representing Sc condensin in the ATP-free state (orange), and Sc TFIIIC in a monomeric or dimeric form (left and right blue frames) (B). The preliminary analysis of the data set suggests no interaction between Sc TFIIIC and Sc condensin in the tested conditions.

II. Interaction between *Sc* condensin and di-nucleosomes

ChIP-seq studies revealed that condensin is enriched at nucleosome-depleted regions (NDRs) (Piazza et al., 2013). Thus, the formation of inter-nucleosomal free DNA regions might be a key factor for the recruitment of condensin onto the chromatin. Moreover, a direct interactions between histones and condensin had been reported by several studies (Ball, et al., 2002; Liu et al., 2010; Tada et al., 2011).

To investigate the interaction between condensin and nucleosomes, a di-nucleosome substrate was used for a condensin-binding experiment. This substrate, generated by Dana Grozavu, consists of two nucleosomes assembled on the 601 Widom sequence (Lowary and Widom, 1998) spaced by 15 to 60 bp of DNA in 1-bp steps. The varying spacer length allow an estimate of the minimal inter-nucleosomal DNA length required for condensin binding. In addition, a DNA sequence of 30 bp is located on the extremities of the di-nucleosome constructs (**Figure 5.5A**). Complex formation between *Sc* condensin and the di-nucleosome library was assessed by EMSA, in an experiment performed by Dana Grozavu. For this assay a Walker-B mutant version of *Sc* condensin, which binds ATP but hydrolyses it with very low turnover, was mixed with this library at varying concentrations, in the presence of ATP. For the condition corresponding to a partial shift

of the di-nucleosome library (10 nM condensin), shifted and non-shifted di-nucleosome species were extracted (**Figure 5.5B**, **orange**). The shifted band is expected to contain the preferentially bound di-nucleosomes. The corresponding DNA library was isolated, amplified and sequenced. At the time of writing, sequence analysis had not been completed.

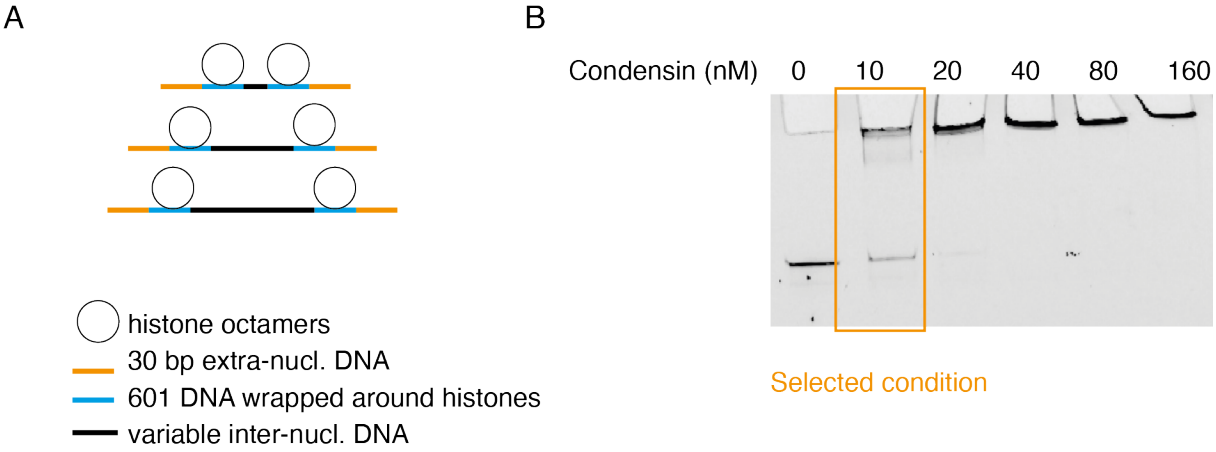


Figure 5.5. Electrophoretic mobility shift assay (EMSA) to probe the binding of *Sc* condensin to a di-nucleosome library. A schematic representation of the library is shown (**A**). Histone octamers were assembled into nucleosomes using the 601 sequence (blue). The variable inter-nucleosomal length is depicted in black and corresponds to a sequence comprised between 15 and 60 bp, with a 1-bp step. A sequence of 30 bp is found at the extremities of the di-nucleosomes (orange). Agarose gel showing the binding of *Sc* condensin to the di-nucleosome library (**B**). The condition corresponding to a concentration of 10 nM of *Sc* condensin was selected to perform the sequence analysis. This experiment was performed by Dana Grozavu in the laboratory of Dr. S. Eustermann.

According to the cryo-EM structure of the condensin core module, I would expect that a minimal length of ~35 bp of DNA should be required to accommodate the condensin binding between the nucleosomes. A crystal structure of the peripheral module revealed that it forms a complex with 18 bp of DNA (Kschonsak et al., 2017). Increasing concentrations of condensin caused a large band shift close to the loading wells (**Figure 5.5B**). This indicates oligomerization or aggregation of the complex. I plan to optimize the experiment by using a di-nucleosome library without flanking extra-nucleosomal DNA, which might be the possible cause the large band shift resulting of the multiple binding of condensin to the di-nucleosome substrates. Moreover, I will use the di-nucleosome substrates for structural studies to uncover potential interactions of condensin with linker

DNA and nucleosome core particles. The analysis of the sequencing results might identify a preferential substrate to perform such a study.

III. DNA sequence specificity of condensin binding

Although condensin has been shown to be enriched at nucleosome-free regions and AT-rich sequences (Piazza et al., 2014), the sequence specificity of condensin binding has not yet been explored in further detail. As I discussed in **Chapter 2**, condensin binds DNA in two topological chambers located at the core and at the peripheral modules. At both chambers, condensin subunits mainly contact the DNA phosphate backbone, which argues in favor of an interaction mode that is independent of the DNA sequence. Nevertheless, the 90° DNA bend observed at the peripheral module is an indication that condensin might recognize specific DNA sequences or shapes.

To probe whether the condensin peripheral module displays any sequence specificity for DNA binding, a SELEX-seq (Systematic Evolution of Ligands by Exponential Enrichment) experiment was performed (Riley et al., 2014). Two rounds of selection were performed by using a complex formed between the *Ct* Ycg1 and its interacting kleisin Brn1 fragment (residues 518-634) (**Figure 5.6A**). This experiment was performed by Thomas Dahlet. As a starting point for this assay, EMSA conditions from Kschonsak et al., 2017 were used. *Ct* Ycg1-Brn1 was mixed with a DNA library containing a central 16-bp segment of randomized DNA sequences, flanked by DNA elements required for sequencing. For the condition where ~10% of the DNA library was shifted, the DNA was extracted from the gel and PCR amplified for a second round of selection (**Figure 5.6B**). Notably, the PCR amplified pool of DNA sequences from the first selection round was shifted, under similar conditions, to a larger extent indicating a successful enrichment of high affinity DNA sequences. The analysis of the sequenced DNA pools was performed as previously described (Riley et al., 2014). Following this method, an affinity score R

(from 0 to 1) was calculated based on the enrichment of the shifted DNA pool after the first and after the second round of selection compared to the initial library.

Several sequence motifs were identified with a high affinity score (R>0.7). Two motifs of 6 bp were identified from 173 and 180 motifs, respectively, with affinity scores of 0.8 and 0.9, respectively (**Figure 5.6C**). Interestingly, both motifs contain two consecutive guanosine residues followed by two or three consecutive thymidine residues ("GGNTTT" for the first selection round and "SGGWTT" for the second selection round, where N can be any nucleotide, S cytidine or guanosine, and W thymidine or adenosine). These results cannot be interpreted in the light of the available structural data since the register of the DNA could not be identified in the cryo-EM structure of the peripheral module. Given the short lengths of these motifs, DNA shape parameters such as the roll, the helix twist or the propeller twist cannot be yet confidently assigned. A deeper analysis of this data will be performed in collaboration with the Bio-IT group at EMBL Heidelberg.

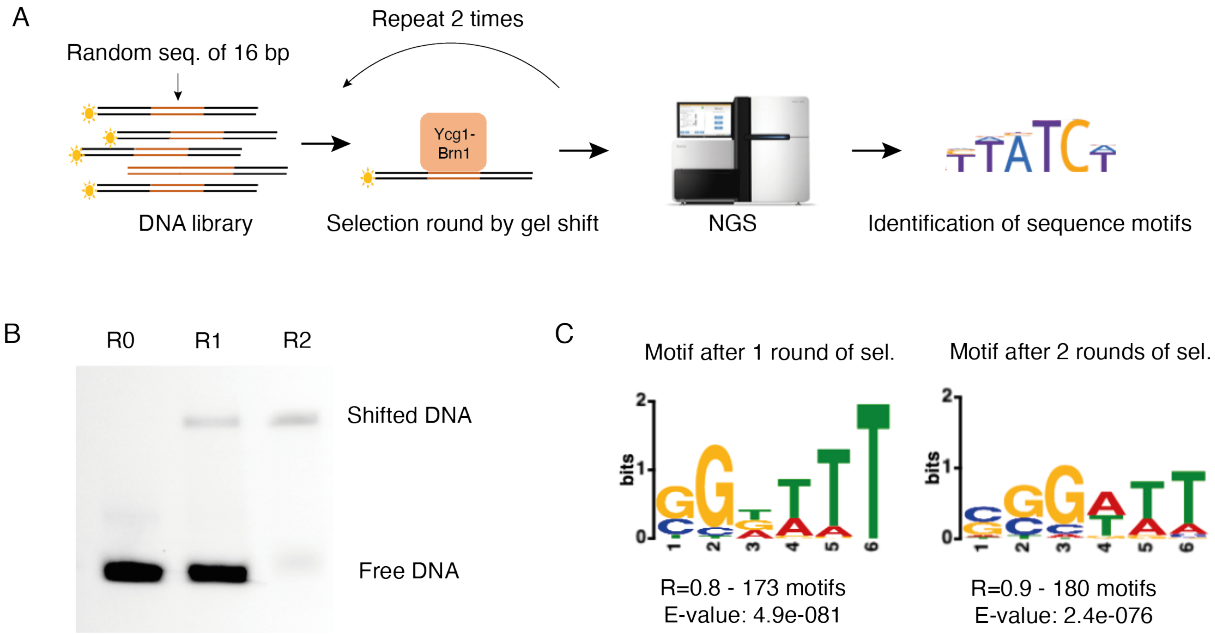


Figure 5.6. SELEX-seq experiment probing the DNA sequence specificity of the *Ct* Ycg1-Brn1 peripheral module. The pipeline of the experiment is shown (**A**). Two EMSA selection rounds are performed by probing the binding of *Ct* Ycg1-Brn1 to a DNA library composed of a random DNA sequence of 16-bp framed by sequencing barcodes. After extraction and PCR amplification of the DNA bound to *Ct* Ycg1-Brn1 at each selection round, ILLUMINA sequencing is performed to identified enriched DNA sequences. EMSA selection rounds are shown in panel **B**. The DNA motifs derived from the two selection rounds are shown in panel **C**. These experiments were performed by Dr. T. Dahlet in the lab of Dr. S. Eustermann.

Discussion

In this thesis, I presented structural and biochemical data characterizing the molecular mechanism of the condensin complex. I focused the DNA-binding activity of condensin at different stages of its reaction cycle in order to understand how the complex extrudes DNA loops.

In **Chapter 1**, I presented a biochemical characterization of condensin-DNA complexes and discussed general considerations regarding how samples of such complexes were optimized for structural studies by cryo-EM.

In **Chapter 2**, I presented cryo-EM structures of two separate modules of ATP-bound condensin in complex with DNA substrates, which I termed "core" and "periphery" subcomplexes. I showed that both create topological chambers for DNA binding. Integration of structural information, cross-linking data and single-molecule loop extrusion data resulted to the proposal of a "power-stroke" mechanism as the driving force for DNA loop extrusion.

In **Chapter 3**, I described low-resolution cryo-EM structures of condensin as a first step towards capturing the ATP-free DNA-bound state of condensin. I will discuss implications from these preliminary data in the following section.

In **Chapter 4**, I presented preliminary data characterizing the interaction of condensin with chromatinized DNA and with the general transcription factor TFIIIC. I also described first experimental steps to probe the DNA-sequence specificity of the condensin peripheral module.

In this concluding section, I will integrate the results of the different parts and assemble them with the structural knowledge available for the SMC complex family in light of their diverse functions through the unique mechanism of DNA loop extrusion. I will present a model for the loop extrusion mechanism of condensin by integrating condensin structures that correspond to different enzymatic states. In addition, I will compare the DNA-binding modes of condensin, cohesin, SMC5/6 and MukBEF SMC complexes to

challenge the power-stroke hypothesis. For this, I will focus on the importance of the molecular rearrangements of the Smc coiled coils and of the HAWK or KITE subunits in the regulation of loop extrusion. Finally, I will discuss exciting open questions in respect to condensin loading onto DNA and chromatin.

I. A DNA loop extrusion model for condensin

a. A hold-and feed model for DNA loop extrusion

Loop extrusion has been proposed as the unifying mechanism underlying SMC complex functions. A loop extrusion model that integrates the structural data obtained during my PhD with protein cross-linking and single-molecule microscopy data obtained by my colleagues Dr. I. A. Shaltiel and Dr. S. Datta, together with previously solved cryo-EM structures of ATP-free condensin states, is described in **Figure 6.1**. The model predicts that the DNA is entrapped inside two topological chambers delimited by the kleisin subunit and that loop extrusion is driven by the engagement of the ATPase heads. Before ATP binds, DNA is entrapped in the core chamber, which is delimited by the first ~200 residues of Brn1, the Smc2 and Smc4 subunits and the Ycs4 HEAT-repeat subunit. At the peripheral chamber, DNA is encircled by the Brn1 "safety-belt" peptide loop and inserts in-between the N- and C-terminal lobe of the Ycg1 HEAT-repeat subunit. Both chambers are required to hold a DNA loop between them (**Figure 6.1**, panel "**loop 1**"). Upon ATP binding, molecular rearrangements at the core create a power-stroke movement: ATPase head engagement and rotation and N-terminal bending of Ycs4 induce tilting of the DNA in-between the open coiled coils. This DNA tilt creates a new DNA loop inside the Smc2 and Smc4 coiled coils (**Figure 6.1**, panel "**loop 2**"). The size of this loop will presumably depend on the tension in the DNA helix, in agreement with

experimental measurements (Ganji et al., 2018), with more slack in the DNA leading to the formation of a larger loop.

Upon ATP hydrolysis, the ATPase heads disengage and the coils collapse, allowing the passage of the DNA segment that had been held into the core chamber. Thereby, the newly-formed loop (**loop 2**) can merge with the initially formed loop (**loop 1**) (**Figure 6.1C**, **"loop merging"**). The top segment of loop 2 is, in contrast, held back by a positively charged patch (**Figure 3.11**, **Chapter 2**) located on the inferior surface of the Ycs4 N-terminal lobe. The transition from the engaged to the bridged states (**Figure 6.1**, panel **B** to panel **C** transition) could reflect the "O-to-B" transitions (open to collapsed coiled coils) observed by AFM (Ryu et al., 2019). Finally, DNA flips back into the core chamber to reset condensin to the starting configuration, now with a longer DNA loop held between the core and the peripheral chambers (**Figure 6.1**, panel **A**). In this model, DNA remains entrapped in the peripheral chamber throughout the loop extrusion cycle.

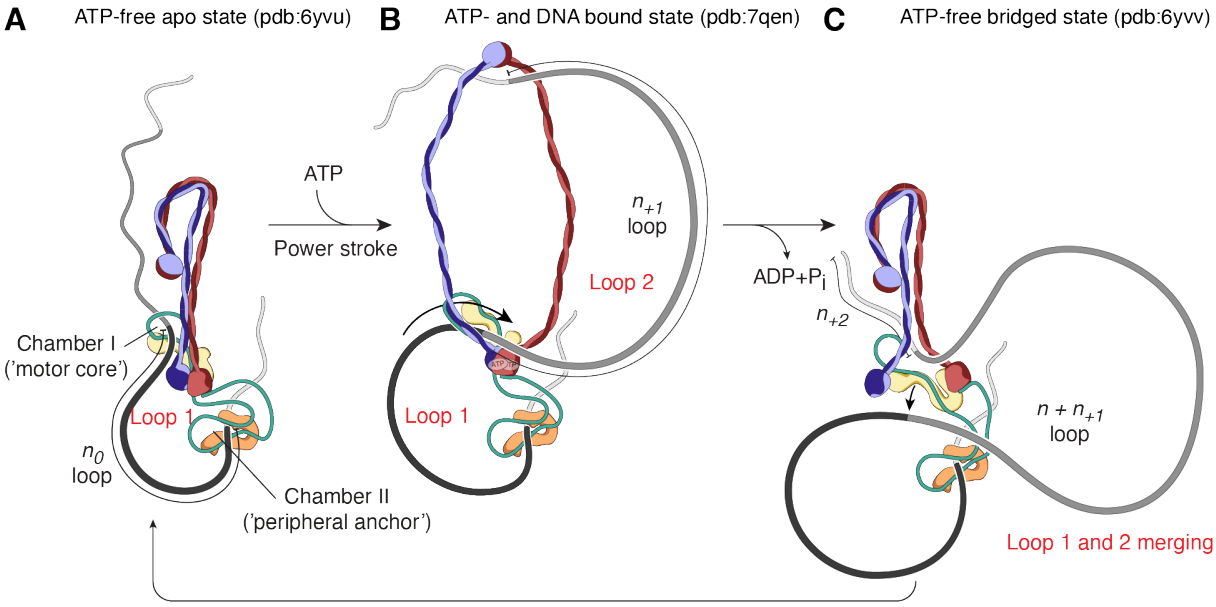


Figure 6.1. The hold-and-feed loop extrusion mechanism proposed for condensin loop extrusion.
DNA is entrapped inside chambers I and II, delimited by the kleisin subunit. Loop extrusion is driven by the engagement of the ATPase heads. Before ATP binds (**A**), DNA is entrapped in the core chamber I, which is delimited by Brn1, Smc2, Smc4 and Ycs4 HEAT-repeat subunit. At the peripheral chamber II, DNA is encircled by the Brn1 "safety-belt" and by the Ycg1 HEAT-repeat subunit. Chambers I and II hold a DNA loop between them ("**loop 1**"). ATP binding (**B**) leads to molecular rearrangements of the core, inducing a power-stroke movement: ATPase head engagement and rotation and N-terminal bending of Ycs4 induce tilting of the DNA in-between the open coiled coils. This DNA tilt creates a new DNA loop

inside the Smc2 and Smc4 coiled coils ("**loop 2**"). Upon ATP hydrolysis (**C**), the ATPase heads disengage and the coils collapse, allowing the passage of the DNA segment that had been held into the core chamber. Thereby, the newly-formed loop (**loop 2**) can merge with the initially formed one (**loop 1**) ("**loop merging**"). The top segment of loop 2 is held back by the Ycs4 N-terminal lobe. Finally, DNA flips back into the core chamber to reset condensin to the starting configuration, now with a longer DNA loop held between the core chamber I and the peripheral chamber II (**A**). In this model, DNA remains entrapped in the peripheral chamber II throughout the loop extrusion cycle.

An essential feature of the hold-and-feed model is that loop extrusion is not based on the topological entrapment of the DNA by the "ring" formed by the Smc and the kleisin subunits, as it was initially proposed (Gruber et al., 2003; Haering et al., 2002). This model is consistent with the finding that a modified cohesin complex, where the Smc and the kleisin subunits were designed as a single protein chain and cross-linked at the hinge, can loop extrude past obstacles that are several orders of magnitude bigger than the "ring" formed by their Smc and kleisin subunits (Pradhan et al., 2021). Instead, our model is based on the topological compartmentalization of the DNA inside separate core and the peripheral chambers. This topological organization guarantees that condensin continuously maintains its contact with DNA while achieving enzymatic processivity. In addition, this topological compartmentalization promotes directional loop extrusion. A DNA compartmentalization strategy is also found among other enzymes with roles in DNA metabolism, such as DNA polymerase sliding clamps or type-II topoisomerases (Beattie and Bell, 2011).

b. The Role of the Smc hinge in DNA loop extrusion

The hinge dimerization domains of the Smc subunits have been found to bind DNA, with a preference for single-stranded over double-stranded DNA (Alt et al., 2017; Griese et al., 2010; Griese and Hopfner, 2011; Piazza et al., 2014). Surprisingly, *B. subtilis* ScpAB SMC complexes are able to support cell division when their hinge domains were replaced by the zinc-hook dimerization domain of Rad50 (Bürmann et al., 2017). Although

structurally different, both, the hinge and the zinc-hook domains, possess basic patches that might have a (non-specific) DNA binding activity (Hopfner et al., 2002). Interestingly, mutation of this basic patch in the cohesin hinge leads to a reduction of the DNA-binding activity and abrogates *in vitro* loop extrusion (Bauer et al., 2021). Therefore, the presence of a DNA binding site at the dimerization extremity of the coiled coils appears to be essential for SMC complex function, whereas the structure of the dimerization domain is not conserved among the SMC family. In the context of the hold-and-feed loop extrusion model described above, the hinge could stabilize the loop formed inside the Smc coiled-coils by interacting with the DNA.

c. Loop extrusion properties of different SMC complexes

The model described in the above section is likely to apply to other SMC complexes: In the case of cohesin, cryo-EM structures highly similar to the ones of the ATP-free apo and ATP- and DNA-bound states of condensin (left and middle panel of **Figure 6.1**) have been reported and describe similar molecular rearrangements of the core module upon ATP binding (Collier et al., 2020; Petela et al., 2021) (see next section, **Figure 6.2**). The tendency of the cohesin $HAWK^{Scc2/Mis4}$ to increase ATPase head spacing in the absence of ATP in AFM studies would suggest that cohesin can also adopt the bridged conformation observed for condensin (**Figure 6.1**, right panel) (Bauer et al., 2021; Higashi et al., 2020, Lee et al., 2020)

However, despite their similar architectures and conformational transitions, condensin and cohesin display different DNA loop extruding properties. Unlike cohesin, condensin extrudes DNA asymmetrically, reeling-in one side of the DNA molecule while stably holding onto the other one (Davidson et al., 2019; Ganji et al., 2018; Golfier et al., 2020; Kim et al., 2019). One-sided loop extrusion requires that one of the DNA-bound chambers anchors condensin onto the DNA.

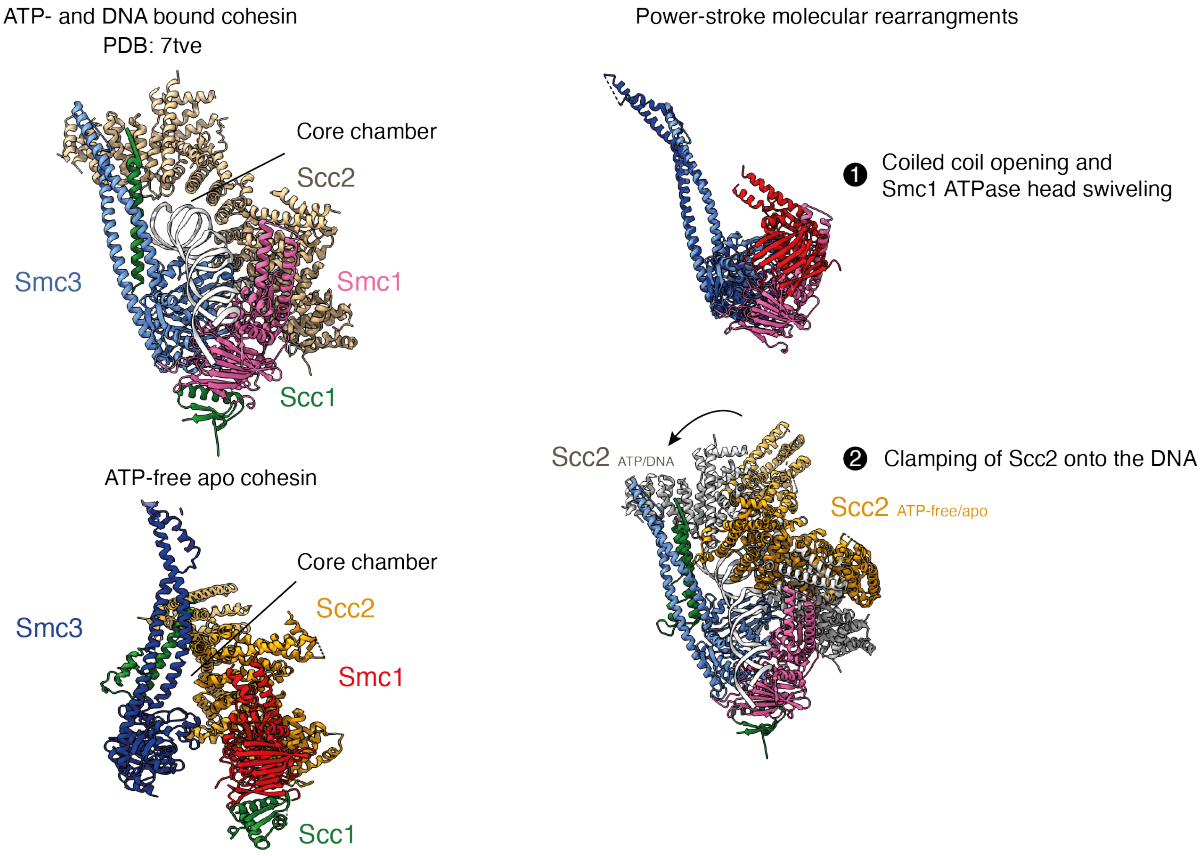


Figure 6.2. Structural comparison of the cryo-EM structures of cohesin in the ATP- and DNA-bound state and in the ATP-free apo state suggesting a power-stroke movement of cohesin.
The ATP-bound cohesin accommodates the DNA in a core chamber that highly resembles to the one described for condensin. The DNA is entrapped by the engaged Smc heads, by the kleisin Scc1 and the HEAT-repeat subunit Scc2 (top, left panel). The core chamber is shown in the ATP-free apo cohesin (bottom, left panel). Superimposition of the Smc subunits of the structures corresponding to the ATP- and DNA-bound state and to the ATP-free apo state of cohesin (**1**). Upon ATP and DNA binding, the coiled-coils open and the Smc1 ATPase head swivels ($Smc1_{apo}$ in red, $Smc1_{ATP/DNA}$ in pink). Superimposition of the Scc2 subunits of the structures corresponding to the ATP- and DNA-bound state and to the ATP-free apo state of cohesin taking the RecA-like lobe of Smc3 as a reference point (**2**), revealing a clamping of Scc2 onto the DNA upon ATP-binding ($Scc2_{ATP/DNA}$ in grey, $Scc2_{ATP\text{-}free/apo}$ in yellow).

The anchoring function of condensin is dependent on the Ycg1 subunit and the Brn1 safety-belt peptide loop of the peripheral module (Kschonsak et al., 2017; Ganji et al., 2018; Saltiel et al., 2021, **Chapter 2**). The 90° bent of the DNA bound to the peripheral module might explain how it can prevent condensin from sliding along DNA during loop extrusion. The cohesin kleisin Scc1 does presumably not have a safety belt that topologically encircles the DNA (Li et al., 2018). Therefore, DNA would not be topologically entrapped at the cohesin peripheral module. Thus, a frequent exchange of the DNA

strands between the core chamber and the peripheral binding site could possibly explain symmetric loop extrusion.

Smc5/6 complexes were reported to extrude DNA loops symmetrically, reeling in DNA from both sides (Pradhan et al., 2022). In the majority of the cases, the extruded DNA loops resulted from the action of Smc5/6 dimers. The same study also revealed that the KITE$^{Nse5/6}$ subunits reduce the loop initiation rate and loop persistence. Although structurally different, KITE$^{Nse5/6}$ subunits could play a similar role as the condensin HAWK^{Ycg1} by downregulating the basal ATPase activity (Shaltiel et al., 2021) or by preventing complex dimerization.

The MukBEF complex has not yet been shown to extrude DNA *in vitro*. Nevertheless, its structural similarity to the Smc5/6 complex would suggest that it has this capacity. As discussed in the next section, cryo-EM structures of the DNA- and ATP-bound Smc5/6 and MukBEF revealed that both complexes also compartmentalize DNA by forming topological chambers delimitated by their engaged ATPase domains and by their kleisin subunits.

In light of this overall comparison, it is likely that the hold-and-feed mechanism of DNA loop extrusion proposed for condensin might apply to the entire family of SMC complexes.

d. The power stroke as a driving force for DNA loop extrusion

In **Chapter 2**, I described the cryo-EM structure of condensin bound to DNA and to ATP. The structure reveals the organization of two DNA-bound condensin modules. One of them, which I call the "core", forms a topological chamber for DNA, which is clamped by the N-terminal lobe of Ycs4 onto the engaged Smc ATPases while it is encircled by Brn1. The single-molecule DNA loop extrusion experiments presented in **Chapter 2** reveal that the core module provides the motor function of condensin. To identify the

molecular mechanism associated with the condensin motor function, I docked DNA into the core chamber of the condensin nucleotide-free apo state, since cross-linking experiments demonstrated DNA entrapment in this chamber in absence of ATP (**Chapter 2**, **Figure 3.16**, PDB: 6YVU). A structural comparison between the nucleotide-free DNA-docked state and the ATP-engaged DNA-bound state shows a tilt of the DNA, which inserts in-between the open coiled coils of the engaged Smc heads upon ATP binding. Concomitantly, Ycs4 subunits rotates and its N-terminus bends over the DNA (**Chapter 2**, **Figures 3.16** and **3.17**). These rearrangements are likely to constitute the loop extrusion "power stroke", where bending of the N-terminus of Ycs4 feeds the DNA in-between the open Smc coiled coils.

Interestingly, structures of other SMC complexes bound to DNA reveal a similar organization of the "core" chamber, with DNA being topologically entrapped by the kleisin, the ATPase subunits and the HAWK or KITE subunits (**Figure 6.3**). Comparisons of the ATP- and DNA-bound and the nucleotide-free apo states of both cohesin and Smc5/6 reveal rearrangements at the core chambers similar to the ones reported for condensin, consisting on the opening of the Smc coiled coils concomitantly with the rotation and bending of the HAWK^{Scc2} or KITE$^{Nse1/3}$ subunits that clamp onto the DNA (**Figure 6.2** and **Figure 6.4**). In the case of MukBEF, two DNA molecules insert in-between the MukB subunit coiled coils. One of them is clamped at the top of the MukB ATPase heads by the KITE^{MukE} dimer and by the kleisin MukF, in a similar manner to what is observed for condensin, cohesin and SMC5/6. The second DNA is held by the MatP dimer at the MukB joint region (**Figure 6.3**) (Bürmann et al., 2021). The conservation of the structural organization of the core among the different SMC complexes could indicate that the power-stroke is the shared motor mechanism among these complexes.

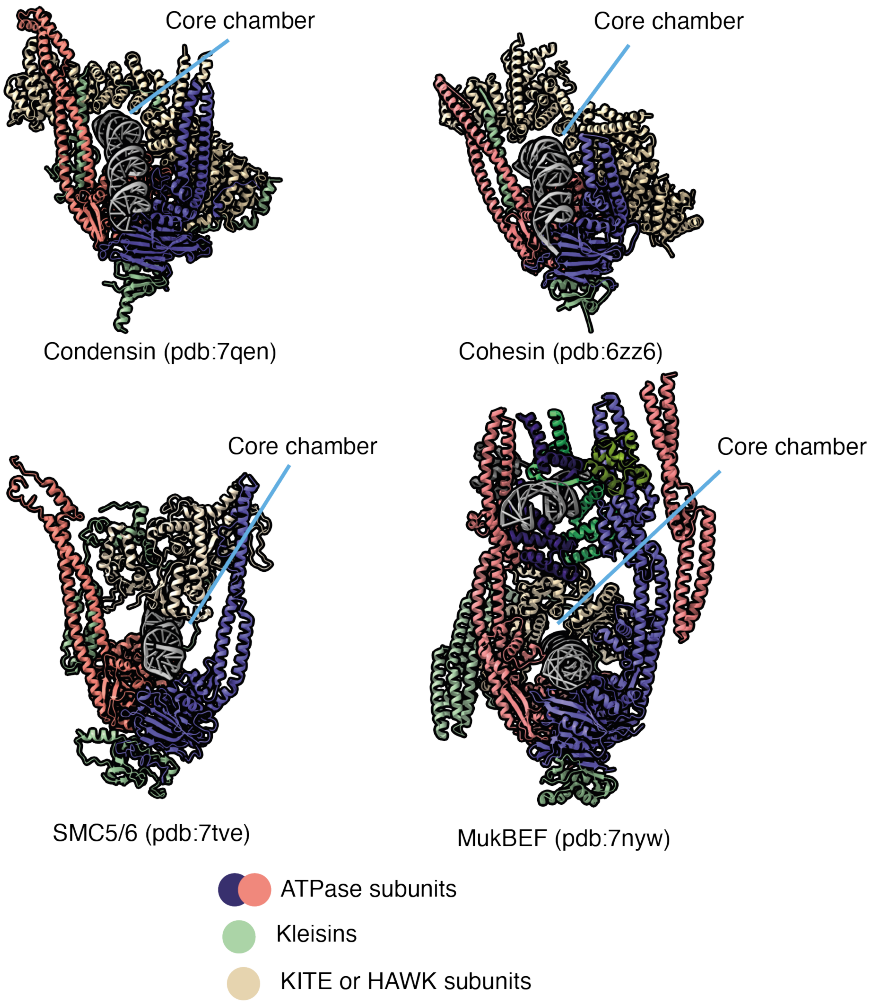


Figure 6.3. Structural comparison of the core chambers of condensin, cohesin, Smc5/6 and MukBEF. The ATPase subunits are shown in purple and pink, the kleisin subunits are shown in green and the HAWK or KITE subunits are shown in beige. In all of the cases, the DNA is bound at the superior surface of the engaged ATPase subunits and encircled by the kleisins. The HAWK or KITE subunits clamp onto the DNA.

In the case of cohesin, the DNA tilt predicted by the power-stroke hypothesis is supported by cross-linking with a 3-kb DNA in presence and absence of ATP (Higashi et al., 2020). In absence of ATP, DNA-protein cross-links were mainly located at the HAWK^{Scc2} N-terminus and at the Smc neck; in presence of ATP, a larger number of cross-links were detected throughout the N-terminal lobe of HAWK^{Scc2} and at the top of the Smc ATPase heads. An additional argument in favor of the power-stroke hypothesis comes from single-molecule experiments, which indicate that the generation of loop extrusion steps

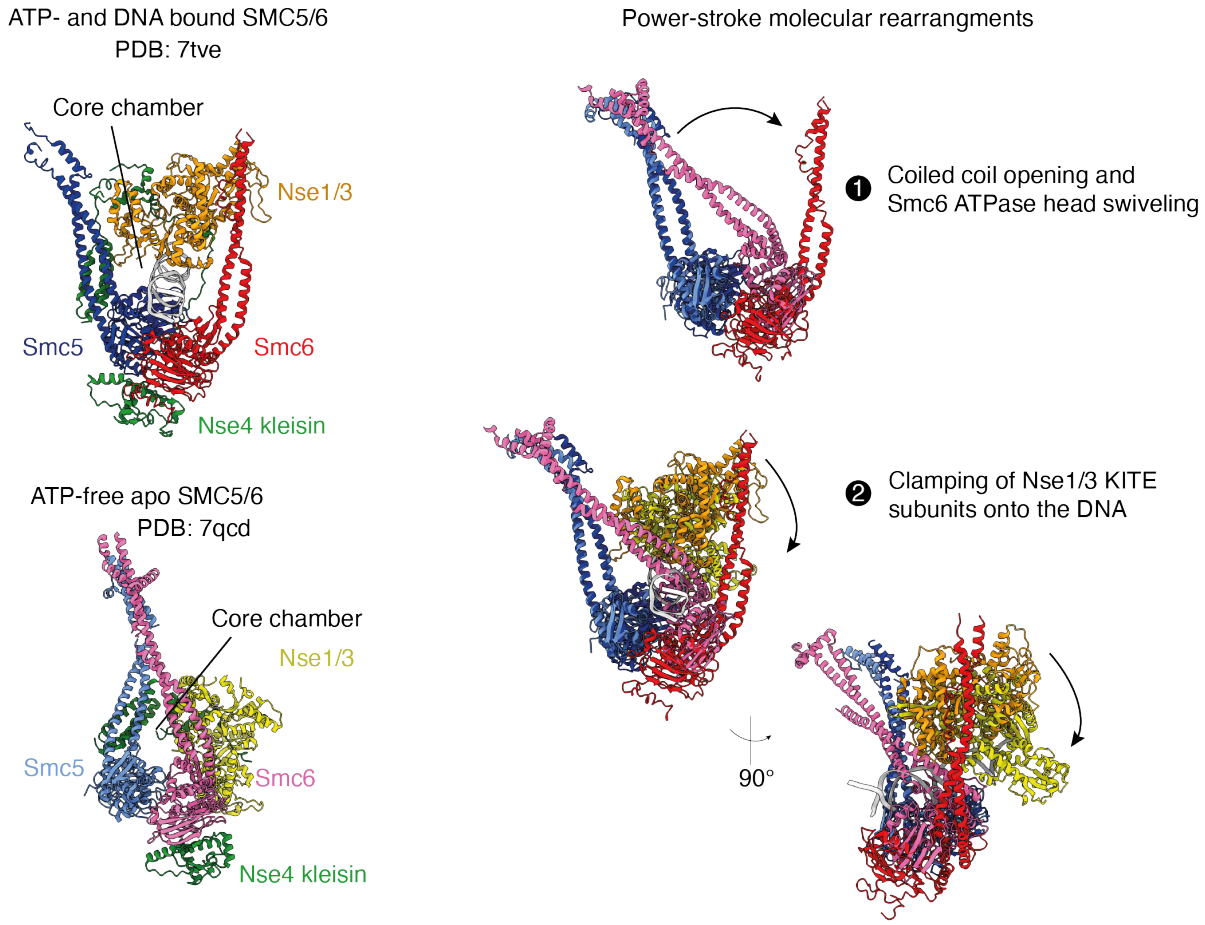


Figure 6.4. Structural comparison of the cryo-EM structures of Smc5/6 in the ATP- and DNA-bound state and in the ATP-free apo state suggesting a power-stroke movement of SMC5/6 core.
The ATP-bound Smc5/6 accommodates the DNA in a core chamber that highly resembles to the one described for condensin and cohesin. The DNA is entrapped by the engaged Smc heads, by the kleisin Nse4 and the KITE subunits Nse1/3 (top, left panel). The core chamber is shown in the ATP-free apo SMC5/6 (bottom, left panel). Superimposition of the Smc subunits of the structures corresponding to the ATP- and DNA-bound state and to the ATP-free apo state of SMC5/6 (**1**). Upon ATP and DNA binding, the coiled-coils open and the Smc6 ATPase head swivels ($Smc6_{apo}$ in pink, $Smc6_{ATP/DNA}$ in red). Superimposition of the Nse1/3 subunits of the structures corresponding to the ATP- and DNA-bound state and to the ATP-free apo state of SMC5/6 taking the RecA-like lobe of Smc5 as a reference point (**2**), revealing a clamping of Nse1/3 onto the DNA upon ATP-binding ($Nse1/3_{ATP/DNA}$ in yellow, $Nse1/3_{ATP\text{-}free/apo}$ in orange).

results from Smc head engagement (Ryu et al., 2019), which depends on ATP binding but not ATP hydrolysis (Bauer et al., 2021). Indeed, condensin Walker-B mutants defective in ATP hydrolysis take a single step onto the DNA in magnetic tweezers experiments (Ryu et al., 2019), which suggests that the loop extrusion step results from ATP binding, not from ATP hydrolysis.

Molecular dynamics simulations of the conformational changes of SMC complexes (Nomidis et al., 2022) based on structural data collected on cohesin, condensin and MukBEF complexes (Bürmann et al., 2021; Collier et al., 2020; Hallett et al., 2021; Higashi et al., 2020; Lee et al., 2020; Shaltiel et al., 2021; Shi et al., 2020) were able to reproduce a power-stroke tilts the DNA into the open Smc coiled coils. The same study proposed that the thermal Brownian motions of the DNA could facilitate the ATP-dependent power-stroke, by inducing the capture of loops by the open Smc coiled coils during loop extrusion.

II. Smc coiled-coils dynamics across SMC complexes

The condensin Smc coiled coils describe a continuous opening motion upon ATP binding (Lee et al., 2020). Cryo-EM structures of cohesin and MukBEF complexes bound to ATP and DNA displayed coiled-coils opening restrained to the head proximal region (Bürmann et al., 2021; Petela et al., 2021). AFM studies suggest, instead, opening of the cohesin and condensin coiled coils along their entire length, and thus independently of ATP binding (Davidson et al., 2019; Eeftens et al., 2016).

The coils contain discontinuities in their helical organization at regions called 'joint' and 'elbow'. The joint forms a flexure point at the head-proximal region and consists of three α-helices that encircle the continuous N-terminal ascending α-helix of the Smc arm (Diebold-Durand et al., 2017). The elbow, present at a conserved position in the Smc arms in cohesin, condensin and MukBEF (Bürmann et al., 2019), results in an asymmetric bent of the coils. Elbow bending has been reported both *in vitro* and *in vivo* (Bürmann et al., 2019; Lee et al., 2020; Petela et al., 2021). As it has been observed in both, ATP-bound and ATP-free states, its relevance is not clear, but it was suggested to promote hinge-HAWK subunit interaction in the case of cohesin (Bauer et al., 2021; Petela et al., 2021). A cohesin AFM study of cohesin revealed that the heads are systematically

disengaged in the case of zipped-up and bent Smc coiled coils (Bauer et al., 2021). It is therefore conceivable that the coiled coil bent stabilizes the disengagement of the ATPase heads. The longer coiled coils in cohesin (580 Å vs 470 Å in condensin) allow the folded hinge to reach up to the joint region and thereby to contact the HAWK subunits (Petela et al., 2021), in contrary to condensin, for which none of these contacts were observed. In the case of MukBEF, which has 580 Å-long coils, the folded hinge reaches down to the head domains and forms an interface between the υ-MukB arm and the κ-MukB joint. This interaction likely stabilizes the coils in a bent conformation, but also prevent the middle domain of the kleisin to interact with κ-MukB to form MukBEF polymers (Bürmann et al., 2021). In summary, the bend of the coiled coils induces the Smc hinge domains of different SMC complexes to either contact the coiled-coil arms, the joint, the HAWK subunits or the ATPase heads.

Elbow folding was not observed for the coiled coils of Smc5/6 or Smc-ScpAB complexes (Soh et al., 2015; Taschner et al., 2021; Yu et al., 2021). This suggests that the presence of an elbow and the bending of the coiled coils result from a divergent evolution of SMC complexes. Molecular dynamics simulations revealed that the rigidity of the elbow and joint regions does presumably not impair the translocation of SMC complexes (Nomidis et al., 2022). The same study even discussed the fact that fully flexible coiled coils could eliminate the need for strong DNA bending and facilitate the extrusion process. In addition, the authors revealed that the reduction of the length of the coiled coils from 50-nm for 40-nm only minimally affected the step size of loop extrusion in the simulations.

In order to fully understand the role of the coiled-coils geometry for SMC complex functions, and particularly for loop extrusion, further studies are required. For instance, performing loop extrusion assays using SMC complexes with elbows locked in a straight or in a bent position could be insightful.

III. Open questions

SMC complexes are conserved from bacteria to eukaryotes. Even though loop extrusion has been demonstrated exclusively for eukaryotic SMC complexes *in vitro*, the structural similarities between prokaryotic and eukaryotic SMC complexes suggests that DNA loop extrusion is a shared property among the SMC family. Loop extrusion is a fascinating molecular mechanism with the ability to explain a plethora of functions associated with SMC complexes, such as chromosome compaction and gene regulation (Davidson and Peters, 2021; Gibcus et al., 2018; Hirano, 2017; van Ruiten and Rowland, 2021; Yatskevich et al., 2019). The accumulation of structural, biochemical, computational and single-molecule imaging data led to a recent progress in our understanding of the molecular mechanism of loop extrusion. Nevertheless, open questions remain about the regulation of SMC complexes and their mechanisms of action in a complex cellular context. In this concluding section, I discuss a few exciting open questions.

a. How do SMC complexes interact with chromatin?

Eukaryotic and prokaryotic genomes are highly organized, with DNA bound to numerous factors to create what has been called "chromatin" (see **Introduction**). Recently, cohesin has been shown to be able to by-pass nucleosomes in an *in vitro* loop extrusion setup (Pradhan et al., 2021). Nevertheless, the DNA used in these experiments only contained a low density of nucleosomes, which might not reflect the physiological condition.

To assay the role of chromatinization on SMC-mediated loop extrusion, it would be of interest to study the interplay between chromatin-accessibility regulatory factors and SMC complexes. For example, a recent study uncovered that TFIIH, a helicase responsible for the formation of the Pol-II pre-initiation complex (Rimel and Taatjes, 2018), is essential for a condensin-dependent formation of mitotic-like chromosomes in *Xenopus*

egg extract, and thus independently of its Pol-II transcriptional activity (Haase et al., 2022). Interestingly, the TFIIH-dependence of condensin for loading onto chromatin was abolished when the nucleosome density was reduced. Another study had revealed that H2A and H2B mutations prevent the lethality caused by the deletion of the cohesin Scc4 subunit in yeast, a subunit that has been implicated in cohesin loading onto chromosomes (Petela et al., 2021). These mutations restored cohesin loading onto DNA. This suggests that nucleosomes restrict cohesin and condensin loading onto chromatin, most likely by reducing DNA accessibility (Piazza et al., 2014). Nevertheless, the interaction of condensin and cohesin with chromatin is poorly understood and constitutes exciting future research directions.

In **Chapter 4**, I presented preliminary attempts to reconstitute complexes between condensin and a di-nucleosome library. In this library, two nucleosomes are spaced by DNA linker lengths varying between 15 and 60 bp in 1-bp steps. The di-nucleosome substrate that corresponds to the minimal inter-nucleosomal DNA length required for condensin loading might soon be identified by sequencing. In addition, single-particle cryo-EM with condensin bound to di-nucleosomes might be able to visualize potential interactions between condensin and histones. In the future, the use of longer arrays of 10 to 15 nucleosomes in a cryo-tomography approach would be interesting to understand how condensin interacts with chromatin. Along the same line, performing loop extrusion assay using nucleosomal DNA with differential nucleosome spacing as a substrate for condensin would be insightful to understand condensin loading and processivity in a chromosomal context.

b. How does DNA sequence, shape and supercoiling affect condensin?

In **Chapter 2**, I presented the cryo-EM structure of the peripheral module, composed of Ycg1 and its interacting Brn1 fragments, bound to DNA. The 90-degrees angle adopted

by the DNA suggests that the peripheral module targets preferential sequences of the genome that are amenable to such a drastic bending.

In **Chapter 4**, I presented preliminary experiments that test the DNA sequence specificity of the binding of the condensin peripheral module. The preliminary analysis identified the DNA motifs "GGNTTT" and "SGGWTT" in the first and second round of selection, respectively. These results are difficult to interpret, since the register of the DNA could not be identified in the peripheral module cryo-EM structure. In addition, given the short length of this motifs, I cannot draw conclusions about DNA shape parameters. Further analysis of this data will be done in collaboration with the Bio-IT unit at EMBL Heidelberg.

Previous studies suggested that R-loops and DNA supercoiling can stimulate the activity of SMC complexes (Bernard and Vanoosthuyse, 2015). Exploring these questions further using biochemical assay such as gel shift or single-molecule loop extrusion experiments in the future would be of great interest.

c. Does condensin interact with chromatin-bound factors?

Unlike for cohesin, no distinct loading complex has been identified for condensin. Condensin was nevertheless found to interact with chromatin factors in pull-down experiments. It is conceivable that these factors facilitate and/or target condensin loading onto chromatin. Among these factors, TFIIIC, a transcription factor associated with RNA polymerase III, was reported to interact with condensin in yeast and human cells (Haeusler et al., 2008; Yuen et al., 2017). In **Chapter 4**, I presented preliminary attempts to investigate this interaction further using mass photometry and cryo-EM. In the near future, alternative biochemical approaches should be used to confirm the interaction suggested by the mass photometry results. In addition, negative-staining EM, which can image particles at low-concentration and independently of parameters that influence the

ice quality in cryo-EM, could be of used to characterize the assemblies formed between TFIIIC, DNA and condensin.

d. How is condensin loaded onto chromatin?

Most of the above-mentioned questions have to do with condensin loading onto DNA. Throughout this thesis, I have shown than condensin's DNA loop extrusion activity is based on the entrapment of DNA in topological chambers delimited by the Brn1 kleisin subunit and regulated by Smc ATPase head engagement. However, how DNA can be initially entrapped into these chambers is unknown. Answering this question will require additional biochemical and structural studies.

In **Chapter 4**, I presented a low-resolution structure of the *Ct* condensin holocomplex that might possibly correspond to a DNA-loading conformation. In this state, the Smc ATPase heads are spread apart but connected by extra density that I hypothesize to correspond to the Ycg1-Brn1 module (see **Chapter 3**, **Figure 4.3**). Interestingly, a conformation that shows the Ycg1-Brn1 module bound to the ATP-engaged heads, in absence of DNA, was reported for condensin (Lee et al., 2020). None of these structures were solved at a sufficient resolution to determine whether the Brn1 "safety-belt" (i.e., the kleisin peptide loop responsible for encircling the DNA) was open or closed. Hypothetically, and assuming that the safety belt opening would be specifically regulated, having the Ycg1-Brn1 module inserted in-between the ATPase heads could transfer the DNA from the safety-belt chamber to the core chamber delimitated by the Smc heads and by the N-terminus of Brn1. Alternatively, the binding of the Ycg1-Brn1 module to the ATPase heads could regulate their catalytical activity, as shown by the higher basal ATPase activity of the *Ct* ΔYcg1 complex when compared to the *Ct* holo-complex (Shaltiel et al., 2022). Alternative loading mechanisms have been suggested, such as the opening of the "gate" formed between the N-terminus of the kleisin and the Smc2 neck

(Murayama and Uhlmann, 2015) or the opening of the hinge in the case of cohesin (Gruber et al., 2006). Nevertheless, the covalent closure of both of these interfaces in cohesin did neither impair DNA loading nor loop extrusion (Davidson et al., 2019), which argues against the hypothesis that opening of the Smc-kleisin ring is required for DNA loading.

Thus, despite accumulating structural data that correspond to different conformations of condensin, the loading of condensin onto DNA is not yet understood. Obtaining higher resolution structures of the *Ct* holo-complex in the above-described state could provide insights into the DNA loading mechanism. An additional approach that could be of interest to resolve the condensin DNA loading mechanism is time-resolved cryo-EM. Several strategies have so far been deveeloped to achieve temporal resolution in cryo-EM (Dandey et al., 2020; Mäeots et al., 2020). These approaches are mostly based on the time-controlled mixing of the reaction components using microfluidics, followed by rapid freezing of the samples after mixing. Capturing intermediate conformations by freezing condensin after adding DNA and/or ATP, in a time-controlled manner, could provide novel insights into condensin loading onto the DNA.

Materials and Methods

I. Software

Software	Source
Cryo-Electron Microscopy	
SerialEM	Mastronade, 2005
RELION 3 and RELION 4	Zivanov et al., 2018
	J. Zivanov et al., 2022
CryoSPARC2	Punjani et al., 2017
MotionCorr2	Zheng et al., 2017
WARP	Tegunov and Cramer, 2019
DeepEMhancer	Sanchez-Garcia et al., 2021
Topaz	Bepler et al., 2019
Data visualisation and analysis	
UCSF ChimeraX	Goddard et al., 2018
Jalview 2.11.2.5	Waterhouse et al., 2009
Consurf	Ashkenazy et al., 2016
Fiji 2.0.0	Schindelin et al., 2012
Coot	Emsley and Cowtan, 2004
Data availability and online resources	
PDB	https://rcsb.org/
EMDB	https://www.ebi.ac.uk/pdbe/emdb/
Thesis writing	
Word 16.16.27	Microsoft
Illustrator 2023	Adobe
Zotero 6.0.5	Zotero.org

II. Cryo-EM data table

The below table is adapted from Shaltiel et *al.*, 2022.

	Sc Condensin Core PDB 7QEN EMDB 13934	Sc Condensin Peripheral PDB 7QFW EMDB 13950
Data collection and Processing (for each dataset)		
Microscope	Titan Krios	Titan Krios
Voltage (keV)	300	300
Camera	GIF-K2 counting	GIF-K2 counting
Magnification	×130000	×130000
Pixel size at detector (Å/pixel)	1.04	1.04
Total electron exposure ($e^-/Å^2$)	40.03	40.03
Exposure rate (e-/pixel/sec)	4.8	4.8
Number of frames collected during exposure	40	40
Defocus range (μm)	-1.0 to -2.0	-1.0 to -2.0
Automation software	SerialEM	SerialEM
Energy filter slit width (if used)	20.0	20.0
Micrographs collected (no.)	6544	6544
Micrographs used (no.)	6544	6544
Total extracted particles (no.)	1203041	571014
For each reconstruction:		
Refined particles (no.)	285280	114600
Final particles (no.)	56664	27701
Point-group or helical symmetry parameters	C1	C1
Estimated error translations (Å) and rotations (°)	0.51, 0.95	0.97, 2.47
Resolution (global, Å)		
FSC 0.5 (unmasked/masked)	7.23/4.00	7.61/4.43
FSC 0.143 (unmasked/masked)	4.16/3.46	4.36/3,86
Resolution range (local, Å)	3.26 – 9.06	3.77 – 7.81
Map sharpening *B* factor ($Å^2$)	-80.6	-117.2
Map sharpening methods *map used for model refinement	RELION local resolution filtering and Rosenthal & Henderson weightening* DeepEMhancer	RELION local resolution filtering and Rosenthal & Henderson weightening DeepEMhancer*
Model composition (for each model)		
Protein	2226	917
Ligands	4	0
RNA/DNA	76	64
Model Refinement (for each model)		
Refinement package	Phenix 1.14-3260	Phenix 1.14-3260
- real or reciprocal space	Real space	Real space
- resolution cutoff	3.2	3.7
Model-Map scores		
Overall correlation coefficients		
CC (mask)	0.82	0.68
CC (peaks)	0.74	0.68
CC (volume)	0.8	0.68
Map versus Model FSC 0.5 (Å)	3.42	3.80
***B* factors ($Å^2$) min/max/mean**		
Protein residues	23.97/160.23/83.62	40.14/103.34/61.41
Ligands	30.02/73.17/55.72	-
RNA/DNA	103.32/289.69/174.04	76.19/199.17/127.77
R.m.s. deviations from ideal values		
Bond lengths (Å)	0.008	0.010
Bond angles (°)	0.897	1.062
Validation (for each model)		
MolProbity score	1.37	1.64
CaBLAM outliers (%)	1.71	2.02
Clashscore	3.96	6.46
Poor rotamers (%)	0.30	0.48
C-beta deviations	0.00	0.00
EMRinger score (if better than 4 Å resolution)	2.23	1.67
Ramachandran plot		
Favored (%)	96.85	95.91
Outliers (%)	0.00	0.11

III. List of plasmids

Plasmid	Source	ID
pET-MCN-Brn1$_{225\text{-}512}$-Ycs4 (N-terminal 6xHIS-tag-TEV-cleavage-site, residues 225-512 of *Ct* Brn1 and residues 3-1222 of untagged *Ct* Ycs4)	Lee et al., 2020	2221
pAceBac1-SMC2-SMC4-StrepII-Brn1-His8 (*Ct* Smc2 and *Ct* Smc4, residues 1-224 and 635-898 of *Ct* Brn1)	Lee et al., 2020	3565
pAceBac1-SMC2-SMC4-StrepII-Brn1-His8-Ycg1-Ycs4 (*Ct* Smc2 and *Ct* Smc4, residues 1-224 and 635-898 of *Ct* Brn1, 24-1006 of untagged *Ct* Ycg1, 3-1222 of untagged *Ct* Ycs4)	Shaltiel et al., 2022	3551
pAceBac1-SMC2-SMC4-StrepII-Brn1-His8-Ycs4 (*Ct* Smc2 and *Ct* Smc4, residues 1-224 and 635-898 of *Ct* Brn1, 3-1222 of untagged *Ct* Ycs4)	Shaltiel et al., 2022	3550

IV. List of yeast strains

The below table is adapted from Shaltiel et *al.*, 2022.

Yeast Genotypes	Source	ID
MAT<u>a</u>, lys2::pGAL1 GAL4::LYS2, pep4::HIS3, bar1::hisG, [2μ pGAL7 SMC4-StrepII$_3$, pGAL10 SMC2, pGAL1 BRN1-His$_{12}$-HA$_3$, TRP1], [2μ pGAL1 YCG1, pGAL10 YCS4, URA3]	Terakawa et al., 2017	C4491
MAT<u>a</u>, lys2::pGAL1 GAL4::LYS2, pep4::HIS3, bar1::hisG, [2μ pGAL7 SMC4(E1352Q)-StrepII$_3$, pGAL10 SMC2(E1113Q), pGAL1 BRN1-His$_{12}$-HA$_3$, TRP1], [2μ pGAL1 YCG1, pGAL10 YCS4, URA3]	Lee et al., 2020	C4520

MAT<u>a</u>, lys2::pGAL1 GAL4::LYS2, pep4::HIS3, bar1::hisG, [2µ pGAL7 SMC4-StrepII$_3$, pGAL10 SMC2, pGAL1 BRN1(13-24ybbR)-His$_{12}$-HA$_3$, TRP1], [2µ pGAL1 YCG1, pGAL10 YCS4, URA3]	Shaltiel et al., 2022	C4982
MAT<u>a</u>, lys2::pGAL1 GAL4::LYS2, pep4::HIS3, bar1::hisG, [2µ pGAL7 SMC4(V721C)-StrepII$_3$, pGAL10 SMC2(K639C), pGAL1 BRN1-His$_{12}$-HA$_3$, TRP1], [2µ pGAL1 YCG1, pGAL10 YCS4, URA3]	Shaltiel et al., 2022	C4723
MAT<u>a</u>, lys2::pGAL1 GAL4::LYS2, pep4::HIS3, bar1::hisG, [2µ pGAL7 SMC4(Q302L)-StrepII$_3$, pGAL10 SMC2(Q147L), pGAL1 BRN1-His$_{12}$-HA$_3$, TRP1], [2µ pGAL1 YCG1, pGAL10 YCS4, URA3]	Terakawa et al., 2017	C4724
MAT<u>a</u>, lys2::pGAL1 GAL4::LYS2, pep4::HIS3, bar1::hisG, [2µ pGAL7 SMC4(S1324R)-StrepII$_3$, pGAL10 SMC2(S1085R), pGAL1 BRN1-His$_{12}$-HA$_3$, TRP1], [2µ pGAL1 YCG1, pGAL10 YCS4, URA3]	Shaltiel et al., 2022	C4726
MAT<u>a</u>, lys2::pGAL1 GAL4::LYS2, pep4::HIS3, bar1::hisG, [2µ pGAL7 SMC4-StrepII$_3$, pGAL10 SMC2, pGAL1 BRN1(13-24ybbR, TEV$_3$141)-His$_{12}$-HA$_3$, TRP1], [2µ pGAL1 YCG1, pGAL10 YCS4, URA3]	Ganji et al., 2018	C5066
MAT<u>a</u>, lys2::pGAL1 GAL4::LYS2, pep4::HIS3, bar1::hisG, [2µ pGAL7 SMC4-StrepII$_3$, pGAL10 SMC2, pGAL1 BRN1(TEV$_3$141)-His$_{12}$-HA$_3$, TRP1], [2µ pGAL1 YCG1, pGAL10 YCS4, URA3]	Hassler et al., 2019	C4896
MAT<u>a</u>, lys2::pGAL1 GAL4::LYS2, pep4::HIS3, bar1::hisG, [2µ pGAL7 SMC4-StrepII$_3$, pGAL7 SMC2-TEV$_3$- BRN1-His$_{12}$-HA$_3$, TRP1], [2µ pGAL1 YCG1, pGAL10 YCS4, URA3]	Shaltiel et al., 2022	C4897
MAT<u>a</u>, lys2::pGAL1 GAL4::LYS2, pep4::HIS3, bar1::hisG, [2µ pGAL1 BRN1-TEV$_3$-SMC4 -His$_{12}$-HA$_3$, pGAL10 SMC2, TRP1], [2µ pGAL1 YCG1, pGAL10 YCS4, URA3]	Shaltiel et al., 2022	C5504
MATα ade2-1, trp1-1, can1-100, leu2-3,112, GAL, psi+, smc2::HIS3, ura3::SMC2-PK$_6$::URA3	Shaltiel et al., 2022	C2003
MATα ade2-1, trp1-1, can1-100, leu2-3,112, GAL, psi+, brn1::HIS3, ura3::BRN1-HA$_6$::URA3	Shaltiel et al., 2022	C2312
MATα ade2-1, trp1-1, can1-100, ura3, GAL, psi+, smc4::natMX, leu2::SMC4-PK$_6$::LEU2	Shaltiel et al., 2022	C3250
MATα, ade2-1, trp1-1, can1-100, GAL, psi+, brn1::natMX, smc2::HIS3, smc4::natMX, leu2::SMC4::LEU2, ura3::SMC2-TEV$_3$-BRN1-Prec$_3$-HA$_6$::URA3	Shaltiel et al., 2022	C3303

MATa, ade2-1, trp1-1, can1-100, GAL, psi+, brn1::natMX, smc2::HIS3, smc4::natMX, leu2::SMC4::LEU2, ura3::SMC2-TEV$_3$-BRN1-Prec$_3$-HA$_6$::URA3	Shaltiel et al., 2022	C3354
MATα ade2-1, trp1-1, can1-100, leu2-3,112, GAL, psi+, brn1::HIS3, smc4::natMX, ura3::BRN1-SMC4-myc$_{18}$::URA3	Shaltiel et al., 2022	C1713
MATα ade2-1, trp1-1, can1-100, leu2-3,112, GAL, psi+, brn1::HIS3, smc4::natMX, ura3::BRN1-SMC4-myc$_{18}$::URA3	Shaltiel et al., 2022	C1714
MATa, scc1::kanMX, smc1::kanMX, smc3::HIS3, leu2::SMC1(G22C, K639C)-myc$_9$::LEU2, ura3::SMC3(E570C)-NoTEV$_3$-SCC1(A547C)-TEV$_3$-HA$_6$::URA3 [2.3 kbp TRP1-ARS1-CEN4]	Shaltiel et al., 2022	C2529
MATa, brn1::natMX, smc2::HIS3, smc4::natMX, leu2::SMC4(V721C, R1417C)::LEU2, ura3::SMC2(K639C)-TEV$_3$-BRN1(K709C)-Prec$_3$-HA$_6$::URA3 [2.3 kbp TRP1-ARS1-CEN4]	Shaltiel et al., 2022	C4137
MAT<u>a</u>, lys2::pGAL1 GAL4::LYS2, pep4::HIS3, bar1::hisG, [2μ pGAL7 SMC4-StrepII$_3$, pGAL10 SMC2, pGAL1 BRN1(13-24ybbR, S384C, TEV434, S524C)-His$_{12}$-HA$_3$, TRP1], [2μ pGAL1 YCG1, pGAL10 YCS4, URA3]	Shaltiel et al., 2022	C5143
MAT<u>a</u>, lys2::pGAL1 GAL4::LYS2, pep4::HIS3, bar1::hisG, [2μ pGAL7 SMC4-StrepII$_3$, pGAL7 SMC2-noTEV$_3$- BRN1(13-24ybbR, TEV141)-His$_{12}$-HA$_3$, TRP1], [2μ pGAL1 YCG1, pGAL10 YCS4, URA3]	Shaltiel et al., 2022	C5503
MAT<u>a</u>, lys2::pGAL1 GAL4::LYS2, pep4::HIS3, bar1::hisG, [2μ pGAL7 SMC4-StrepII$_3$, pGAL10 SMC2, pGAL1 BRN1(13-24ybbR, TEV373, S384C, TEV434, S524C)-His$_{12}$-HA$_3$, TRP1], [2μ pGAL1 YCG1, pGAL10 YCS4, URA3]	Shaltiel et al., 2022	C5157
MAT<u>a</u>, lys2::pGAL1 GAL4::LYS2, pep4::HIS3, bar1::hisG, [2μ pGAL7 SMC4-StrepII$_3$, pGAL10 SMC2, pGAL1 BRN1(13-24ybbR, TEV373, S384C, S524C)-His$_{12}$-HA$_3$, TRP1], [2μ pGAL1 YCG1, pGAL10 YCS4, URA3]	Shaltiel et al., 2022	C5158
MAT<u>a</u>, lys2::pGAL1 GAL4::LYS2, pep4::HIS3, bar1::hisG, [2μ pGAL7 SMC4(V721C, R1417C)-StrepII$_3$, pGAL7 SMC2(K639C)-TEV$_3$-BRN1(13-24ybbR, K709C)-His$_{12}$-HA$_3$, TRP1], [2μ pGAL1 YCG1, pGAL10 YCS4, URA3]	Shaltiel et al., 2022	C5200

MAT<u>a</u>, lys2::pGAL1 GAL4::LYS2, pep4::HIS3, bar1::hisG, [2µ pGAL7 SMC4-StrepII$_3$, pGAL10 SMC2, pGAL1 BRN1(13-24ybbR, N307insC, TEV373, T637C)-His$_{12}$-HA$_3$, TRP1], [2µ pGAL1 YCG1, pGAL10 YCS4, URA3]	Shaltiel et al., 2022	C5360
MAT<u>a</u>, lys2::pGAL1 GAL4::LYS2, pep4::HIS3, bar1::hisG, [2µ pGAL7 SMC4-StrepII$_3$, pGAL10 SMC2, pGAL1 BRN1(13-24ybbR, N307insC, TEV373, S384C, S528C, T637C)-His$_{12}$-HA$_3$, TRP1], [2µ pGAL1 YCG1, pGAL10 YCS4, URA3]	Shaltiel et al., 2022	C5372
MAT<u>a</u>, lys2::pGAL1 GAL4::LYS2, pep4::HIS3, bar1::hisG, [2µ pGAL7 SMC4-StrepII$_3$, pGAL10 SMC2, pGAL1 BRN1(13-24ybbR, N307insC, S384C, TEV434, S524C, T637C)-His$_{12}$-HA$_3$, TRP1], [2µ pGAL1 YCG1, pGAL10 YCS4, URA3]	Shaltiel et al., 2022	C5373
MAT<u>a</u>, lys2::pGAL1 GAL4::LYS2, pep4::HIS3, bar1::hisG, [2µ pGAL7 SMC4-StrepII$_3$, pGAL10 SMC2, pGAL1 BRN1(13-24ybbR, TEV141, S384C, TEV434, S524C)-His$_{12}$-HA$_3$, TRP1], [2µ pGAL1 YCG1, pGAL10 YCS4, URA3]	Shaltiel et al., 2022	C5385
MAT<u>a</u>, lys2::pGAL1 GAL4::LYS2, pep4::HIS3, bar1::hisG, [2µ pGAL7 SMC4-StrepII$_3$, pGAL7 SMC2(Q71C)-TEV$_3$- BRN1(13-24ybbR, E197C)-His$_{12}$-HA$_3$, TRP1], [2µ pGAL1 YCG1, pGAL10 YCS4, URA3]	Shaltiel et al., 2022	C5422
MAT<u>a</u>, lys2::pGAL1 GAL4::LYS2, pep4::HIS3, bar1::hisG, [2µ pGAL7 SMC4-StrepII$_3$, pGAL7 SMC2(T75C)-TEV$_3$- BRN1(13-24ybbR, D196C)-His$_{12}$-HA$_3$, TRP1], [2µ pGAL1 YCG1, pGAL10 YCS4, URA3]	Shaltiel et al., 2022	C5424
MAT<u>a</u>, lys2::pGAL1 GAL4::LYS2, pep4::HIS3, bar1::hisG, [2µ pGAL7 SMC4-StrepII$_3$, pGAL7 SMC2(T75C)-TEV$_3$- BRN1(13-24ybbR, G198C)-His$_{12}$-HA$_3$, TRP1], [2µ pGAL1 YCG1, pGAL10 YCS4, URA3]	Shaltiel et al., 2022	C5425
MAT<u>a</u>, lys2::pGAL1 GAL4::LYS2, pep4::HIS3, bar1::hisG, [2µ pGAL7 SMC4(V721C, R1417C)-StrepII$_3$, pGAL7 SMC2(K639C)-noTEV$_3$-BRN1(13-24ybbR, N307insC, TEV434, T637C, K709C)-His$_{12}$-HA$_3$, TRP1], [2µ pGAL1 YCG1, pGAL10 YCS4, URA3]	Shaltiel et al., 2022	C5428
MAT<u>a</u>, lys2::pGAL1 GAL4::LYS2, pep4::HIS3, bar1::hisG, [2µ pGAL7 SMC4-StrepII$_3$, pGAL10 SMC2, pGAL1 BRN1-His$_{12}$-HA$_3$, TRP1], [2µ pGAL1 YCG1, pGAL10 YCS4(K292E, K300E, URA3]	Shaltiel et al., 2022	C5455

MAT<u>a</u>, lys2::pGAL1 GAL4::LYS2, pep4::HIS3, bar1::hisG, [2µ pGAL7 SMC4-StrepII$_3$, pGAL10 SMC2, pGAL1 BRN1-His$_{12}$-HA$_3$, TRP1], [2µ pGAL1 YCG1, pGAL10 YCS4(K377E, K384E, URA3]	Shaltiel et al., 2022	C5474
MAT<u>a</u>, lys2::pGAL1 GAL4::LYS2, pep4::HIS3, bar1::hisG, [2µ pGAL7 SMC4(N266C, V721C)-ybbR-StrepII$_3$, pGAL10 SMC2(K639C, E1080C), pGAL1 BRN1-His$_{12}$-HA$_3$, TRP1], [2µ pGAL1 YCG1, pGAL10 YCS4, URA3]	Shaltiel et al., 2022	C5490
MAT<u>a</u>, lys2::pGAL1 GAL4::LYS2, pep4::HIS3, bar1::hisG, [2µ pGAL7 SMC4-StrepII$_3$, pGAL10 SMC2, pGAL1 BRN1(13-24ybbR, TEV$_3$141)-His$_{12}$-HA$_3$, TRP1], [2µ pGAL10 YCS4, URA3]	Hassler et al., 2019	C5110
MAT<u>a</u>, lys2::pGAL1 GAL4::LYS2, pep4::HIS3, bar1::hisG, [2µ pGAL7 SMC4(E356C, V721C) -StrepII$_3$, pGAL10 SMC2(K639C, N985C), pGAL1 BRN1-His$_{12}$-HA$_3$, TRP1], [2µ pGAL1 YCG1, pGAL10 YCS4(S238C, K1170C)-ybbR, URA3]	Shaltiel et al., 2022	C5509
MAT<u>a</u>/α ade2-1, trp1-1, can1-100, leu2-3,112, GAL, psi+, ycs4::kanMX6/YCS4, ura3::YCS4-PK$_6$::URA3/ura3	Shaltiel et al., 2022	C5440
MAT<u>a</u>/α ade2-1, trp1-1, can1-100, leu2-3,112, GAL, psi+, ycs4::kanMX6/YCS4, ura3::YCS4(R142E, K145E, K148E, K158E, R159E)-PK$_6$::URA3/ura3	Shaltiel et al., 2022	C5435
MAT<u>a</u>/α ade2-1, trp1-1, can1-100, leu2-3,112, GAL, psi+, ycs4::kanMX6/YCS4, ura3::YCS4(K211E, K216E, R221E)-PK$_6$::URA3/ura3	Shaltiel et al., 2022	C5436
MAT<u>a</u>/α ade2-1, trp1-1, can1-100, leu2-3,112, GAL, psi+, ycs4::kanMX6/YCS4, ura3::ycs4(K292E, K300E)-PK$_6$::URA3/ura3	Shaltiel et al., 2022	C5443
MAT<u>a</u>/α ade2-1, trp1-1, can1-100, leu2-3,112, GAL, psi+, ycs4::kanMX6/YCS4, ura3::ycs4(K377E, K384E)-PK$_6$::URA3/ura3	Shaltiel et al., 2022	C5442
MAT<u>a</u>/α ade2-1, trp1-1, can1-100, leu2-3,112, GAL, psi+, brn1::HIS3/BRN1, ura3::BRN1-PK$_6$::URA3/ura3	Hassler et al., 2019	C4239
MAT<u>a</u>/α ade2-1, trp1-1, can1-100, leu2-3,112, GAL, psi+, brn1::HIS3/BRN1, ura3::brn1(N41E, K42E)-PK$_6$::URA3/ura3	Shaltiel et al., 2022	C5459
MAT<u>a</u>/α ade2-1, trp1-1, can1-100, leu2-3,112, GAL, psi+, brn1::HIS3/BRN1, ura3::brn1(K99E)-PK$_6$::URA3/ura3	Shaltiel et al., 2022	C5460

MATa/α ade2-1, trp1-1, can1-100, leu2-3,112, GAL, psi+, smc2::natMX/SMC2, ura3::SMC2-PK$_6$::URA3/ura3	Shaltiel et al., 2022	C5461
MATa/α ade2-1, trp1-1, can1-100, leu2-3,112, GAL, psi+, smc2::natMX/SMC2, ura3::smc2(R121E)-PK$_6$::URA3/ura3	Shaltiel et al., 2022	C5462
MATa/α ade2-1, trp1-1, can1-100, leu2-3,112, GAL, psi+, smc2::natMX/SMC2, ura3::SMC2(K114E)-PK$_6$::URA3/ura3	Shaltiel et al., 2022	C5463
MATa/α ade2-1, trp1-1, can1-100, leu2-3,112, GAL, psi+, smc2::natMX/SMC2, ura3::SMC2(K76E)-PK$_6$::URA3/ura3	Shaltiel et al., 2022	C5464

V. Protein purification

a. *Saccharomyces cerevisiae* (*Sc*) condensin

As previously described (Terakawa et al., 2017), *Sc* condensin was overexpressed in budding yeast under the control of galactose-inducible promoters from two 2-micron-based high-copy plasmids and subsequently purified in a three steps process. One plasmid contained *pGAL7-SMC4-StrepII$_3$ pGAL1-SMC2 pGAL1-BRN1-His$_{12}$-HA$_3$ URA3*, with or without Walker B mutations (Smc4$_{E1352Q}$, Smc2$_{E1113Q}$). The second plasmid contained *pGAL10-YCS4* and *pGAL1-YCG1 TRP1*. The same plasmid combination was used for the production of fusion or mutated proteins.

Budding yeast cells were grown at 30 ºC in –Trp–Ura media and 2 % (w/v) D-glucose, before transfer to –Trp–Ura media and 2 % (w/v) raffinose for a duration of 6 h. Over-expression was induced by the addition D-galactose to 2 % (w/v) for 16 h. Cell lysis was carried out in a FreezerMill (Spex) in a buffer containing: 200 mM NaCl, 50 mM Tris-HCl pH 7.5, 5 % (v/v) glycerol, 5 mM β-mercaptoethanol, 20 mM imidazole and one tablet of cOmplete EDTA-free protease inhibitor mix per 30 mL of buffer (11873580001, Merck). The lysate was cleared by centrifugation (1h, 50,000xg) and loaded onto a HisTrapp FF column (Cytiva – 5 mL). Proteins were eluted by addition of imidazole to a concentration of 220 mM. After an incubation of elution fractions for

16h in the elution buffer plus 0.01 % (v/v) Tween-20, 1 mM EDTA and 0.2 mM PMSF, and Strep-Tactin Superflow high-capacity resin (2-1208-010, IBA), proteins bound to the Strep-Tactin resin were eluted in a buffer containing 50 mM Tris-HCl pH 7.5, 200 mM NaCl, 5 % (v/v) glycerol, 1 mM dithiothreitol (DTT) and 10 mM desthiobiotin (D1411, Merck). After ultrafiltration, elution fractions were loaded onto a gel filtration column pre-equilibrated in a buffer containing: 50 mM Tris-HCl pH 7.5, 200 mM NaCl, 5 % (v/v) glycerol, 1 mM dithiothreitol (DTT) (Superose 6 increase 10/300 column, Cytiva). Condensin complexes were then stored at -80°C after flash freezing in liquid nitrogen.

b. *Chaetomium thermophilum* (Ct) condensin

The Multibac expression system (Sari et al., 2016) was used to overexpress the *Ct* condensin holocomplex and ΔYcg1 tetrameric complex under the control of the polyhedrin promoter (see table in section II for construct details) in Sf21 insect cells cultured in SF 900 III serum-free medium (12658019, ThermoFisher) at 27 °C. After centrifugation, cells were resuspended in a buffer containing 200 mM NaCl, 50 mM Tris-HCl pH 7.5, 5 % (v/v) glycerol, 5 mM β-mercaptoethanol, 20 mM imidazole and one tablet of cOmplete EDTA-free protease inhibitor mix for 30 mL of buffer (11873580001, Merck). Resuspended cells were sonicated for 45 sec at 4°C with 50% amplitude at "output control" 6, repeated 5 times (SFX-550, Branson). After centrifugation (45,000 × *g* for 1 h), the lysate was incubated with Nickel Sepharose (175268, Cytiva) beads for 3 h, at 4°C. The beads were subsequently washed with 40 to 50 column volumes of a buffer containing 50 mM Tris-HCl pH 7.5, 200 mM NaCl, 5 % (v/v) glycerol, 20 mM imidazole, 5 mM 2-mecaptoethanol. Bound proteins were eluted by addition of 300 mM imidazole to the washing buffer. After exchange to a low salt buffer using a desalting column (HiPrep 26/10 – Cytiva) (25 mM Tris-HCl pH 7.5, 100 mM NaCl, 5 % (v/v) glycerol, 1 mM DTT) using a desalting column, the elution fractions were loaded onto a RESOURCE Q

anion exchange column pre-equilibrated in low-salt buffer (Cytiva). The elution was carried out by increasing the NaCl concentration to 1 M in a linear gradient of 60 mL. Selected fractions were collected and concentrated by ultrafiltration (Vivaspin 30,000 MWCO, Sartorius) before loading onto a gel filtration column (Superose 6 increase 10/300 - Cytiva) pre-equilibrated with a buffer containing 200 mM NaCl, 50 mM TrisHCl pH 7.5, 5% (v/v) glycerol and 1 mM DTT. After analysis of the elution fractions by SDS-PAGE, the selected fractions were pooled and frozen at –80 ºC until use.

c. Expression and purification of *S. cerevisiae* TFIIIC

A six-gene *Sc* TFIIIC expression plasmid was cloned via biGBac and used to generate a recombinant baculovirus for insect cells (Hi5) expression (N-terminal TEV-cleavable his-tag on τ95 and a C-terminal TEV-cleavable TwinStrep tag on τ138). The cells were lysed by sonication using 3 mL buffer/gram of pellet containing 500 mM NaCl, 20 mM imidazole, 50 mM TrisHCl pH 7.5, 4 mM beta-mercapto-ethanol, 2 mM $MgCl_2$, 10% glycerol supplemented with 1 tablet of protease inhibitors (11873580001, Merck) per 20 g of cell pellet, 5 µL benzonase (150 units) and 500 µL DNase I (10mg/mL). The lysate was cleared by centrifugation for 1h at 4°C at 45,000×g. The lysate was loaded onto a Nickel column (5 mL GE Healthcare Nickel NTA FF column) and eluted into a Strep column (5 mL Strep-Tactin XP column (IBA Life Sciences) using a buffer containing 150 mM NaCl, 5 % glycerol, 50 mM TrisHCl pH 7.5, 300 mM Imidazole, 2 mM DTT. After binding of the sample to the Strep-Tactin column, the latter was washed with a buffer containing 150 mM NaCl, 50 mM TrisHCl pH 7.5, 5 mM DDT. Elution from the Strep column was done by addition of biotin to a concentration of 50 mM. The sample was then loaded onto an anion exchange column (MonoQ 5/50 GL column, Thermofisher), pre-equilibrated with a buffer containing 150 mM NaCl, 20 mM HepesKOH pH 7.5, 5 mM DTT. A linear salt gradient increasing the salt concentration up to 1 M NaCl was

used for elution. Fractions were analyzed by SDS-PAGE, pooled, concentrated and frozen at –80 ºC until use.

d. Expression and purification of *Ct* condensin subcomplexes

The *Ct* Ycs4-Brn1 subcomplex was expressed in *E. coli* Rosetta (DE3) pLysSRARE2 cells from a pET-MCN vector under the expression of a T7*lac* promoter (check the table in section II for construct details) for 18 h in 2×YT medium at 37°C. Overexpression was induced by the addition of IPTG to a concentration of 0.2 mM at 18°C. Cells were sonicated at 4°C in a lysis buffer containing 50 mM Tris-HCl pH 7.5, 500 mM NaCl, 20 mM imidazole, 5 mM β-mercaptoethanol, cOm–EDTA (1 tablet/30 mL of buffer). The lysate was cleared by centrifugation at 45,000×g for 30 min and loaded onto Ni-Sepharose beads (GE Healthcare; 1 mL of beads/mL lysate). The beads were washed with lysis buffer for 50 to 100 column volumes (CVs). Proteins were then eluted with lysis buffer supplemented of 300 mM imidazole. The sample was dialyzed for 15 h at 4°C in 2 L dialysis buffer (25 mM Tris-HCl pH 7.5, 300 mM NaCl, 1 mM DTT) using SnakeSkin dialysis tubing, 10,000 MWCO (Thermo Fisher). After dilution of the sample to reach a final salt concentration of 150 mM NaCl, the sample was loaded onto an anion exchange column (6-mL Resource Q anion exchange column (GE Healthcare)). The elution of the proteins was done by a linear salt gradient of 10 CVs. Selected fractions were then loaded onto a gel filtration column (Superdex 200 26/60 column (GE Healthcare)) equilibrated with a buffer containing 25 mM Tris-HCl pH 7.5, 300 mM NaCl, 1 mM DTT. Lastly, the sample was concentrated by centrifugation using a Vivaspin 30,000 MWCO filter (Sartorius).

The *C. thermophilum* Smc2–Smc4–Brn1$_{NC}$ and the *Ct* bonsai condensin subcomplexes were expressed from *Sf21* insect cells using the Multibac system (Sari et al., 2016). The clearance of the lysate, the nickel affinity purification step and the anion exchange step

were performed as described in the previous paragraph. Before loading onto a Resource Q anion exchange column, the sample was desalted using a HiPrep 26/10 desalting column (GE Healthcare) in low-salt buffer (50 mM Tris-HCl pH 7.5, 100 mM NaCl, 1 mM DTT). After anion exchange, the pooled fractions were loaded onto a gel filtration column (Superose 6 10/300 column (GE Healthcare)), equilibrated with a buffer containing 25 mM Tris-HCl pH 7.5, 200 mM NaCl, 1 mM DTT, 5 % glycerol. Selected fractions were pooled and the sample was concentrated as described for the *Ct* Ycs4-Brn1 subcomplex.

VI. EMSA: Electrophoretic mobility Shift Assay

The EMSA experiments using the *Sc* and *Ct* condensin complexes were based on similar protocols. 6-FAM-labelled DNA (5'-6-FAM-attagttact aagatccttc ctctgtagaa gaatgagatt atcggagaca g-6-FAM-3', Merk) were annealed with the antisense strand, apmplified by PCR and purified via gel filtration. The purified DNA substrates were incubated at a concentration of 5 nM with the indicated concentrations of condensin, in presence of 1 mM ATP where indicated. The incubation was performed on ice for 15 min in a buffer containing 50 mM Tris-HCl pH 7.5, 125 mM NaCl, 5 mM $MgCl_2$, 50 mM KCl, 5 % (v/v) glycerol, 1 mM DTT. The protein-DNA mixtures were then loaded onto 0.8 % (w/v) agarose gels in Tris-acetate buffer for 1.5–2 h at 8 $V \cdot cm^{-1}$ at 4 ºC to resolve e DNA-bound and DNA-free protein species. Sample were visualized using a Typhoon FLA9500 (Cytiva) with a 532-nm laser and a 575-nm LP emission filter and the analysis was performed using FIJI (Schindelin et al., 2012).

VII. Analytical Gel filtration

To test the DNA-binding activity of *Sc* condensin, *Sc* condensin complexes were diluted to a concentration of 2 μM in a buffer containing 20 mM HEPESKOH pH 7.5, 125 mM

NaCl, 0.5 mM DTT and 1 mM $MgCl_2$ on ice for 15 min. 1 mM ADP was added to the mixture. After a 10-min incubation time, 0.5 mM $AlCl_3$ and 10 mM NaF were added where indicated. 6-FAM-labelled 75-bp DNA oligos (Merk) were annealed with the anti-sense strand. Briefly, the oligos were mixed at equimolar ratio (25 mM) in a 2× buffer containing: 20 mM Tris-HCl pH 7.5, 100 mM NaCl, 1 mM EDTA. The temperature was then decreased 0.1°C/sec from 90°C to 4°C.

Annealed DNA was added at a concentration of 5 μM to the condensin solution and incubated on ice for 1 h. After centrifugation, the sample was loaded onto a Superose 6 inc. 3.2/300 (Cytiva) gel filtration column. Fractions were analyzed by both, SDS-PAGE and 1% agarose gel electrophoresis. In the latter case, samples were visualized using a Typhoon FLA9500 (Cytiva) scanner with a 532-nm laser and a 575-nm LP emission filter.

VIII. Fluorescent labelling of purified condensin complexes

The synthesis of coenzyme A-coupled dyes was carried out as previously described (Terakawa et al., 2017). Site-specific covalent labelling of the hydroxyl group of the serine residue in the ybbR tags (GTDSLEFIASKLA) in condensin holo-complexes (2–7 μM) was performed in a buffer containing 200 mM NaCl, 50 mM Tris-HCl pH 7.5, 5 % (v/v) glycerol, 1 mM dithiothreitol (DTT), 1.2 μM Sfp phosphopantetheinyl transferase (P9302, New England Biolabs), 10 mM $MgCl_2$ and CoA-dye conjugate reaction mix present in a 5-fold molar excess. After incubation for 16 h at 6 °C, the complexes were loaded onto a gel filtration column (Superose 6 increase 3.2/300 column (Cytiva)), to separate the labelled from the non-labelled proteins. Selected fractions were pooled and concentrated by ultrafiltration (Vivaspin), and stored at -80°C.

IX. *In vitro* DNA loading assay

For the DNA loading assay, plasmid of 6.4 kb that contains TRP1, CEN4, RDN37 and ARS1 sequences from *S. cerevisiae* and 21 tetO repeats was incubated with the enzyme Nb.BbvCI (R0631, New England Biolabs) or EcoRI (R3101, New England Biolabs) in order to generate the nicked or linear substrates, respectively. The substrates were subsequently isolated by phenol:chloroform extraction and ethanol precipitation. On ice, 100 nM condensin was mixed with either 250 or 500 ng of DNA substrate and 0.1 $g.L^{-1}$ BSA in a reaction buffer containing 125 mM NaCl, 50 mM KCl, 50 mM Tris-HCl pH 7.5, 5 mM $MgCl_2$, 5 % (v/v) glycerol and 1 mM DTT. The mix was then supplemented with 1 mM ATP and incubated for 5 min at 25 ºC. Condensin-DNA complexes were immunoprecipitated on ice for 20 min using 12.5 µL magnetic protein A Dynabeads (10002D, ThermoFisher). Prior to the immunoprecipitation, the beads were loaded with 3 µg of 12CA5 anti-HA monoclonal antibodies. After a first washing step using 100 µL of reaction buffer, a second wash of 300 µL was performed at 0.5 M NaCl. To analyze the reaction products, DNA was isolated by ethanol precipitation. To linearize DNA substrates, 1 µL XhoI was added (R0146, New England Biolabs) to 30 µL reaction buffer before the 0.5 M salt washes, followed by an incubation at 25 ºC for 20 min. TEV protease cleavage was performed to cut the condensin chambers by incubation of the immuno-precipitates with 1.5 µg TEV protease at 25 ºC for 25 min, prior to the high-salt washes. To analyze the immune-precipitates, elution was performed at 65 ºC in reaction buffer supplemented with 1% SDS (w/v) for 10 min. The products were subsequently separated using 0.75% agarose gels ran in TAE buffer (0.5 $mg{\cdot}L^{-1}$ EtBr), for 1–4 $V{\cdot}cm^{-1}$ for 16–4h at 4 ºC. To visualize the samples, a Typhoon FLA9500 (Cytiva) scanner equipped with a 532-nm laser and 575-nm LP emission filter was used. FIJI (Schindelin et al., 2012) was used to analyze the images.

X. Dibromobimane (bBBr) cross-linking

For the cross-linking experiments, the condensin complexes were washed with 400 µL buffer (125 mM NaCl, 50 mM KCl, 50 mM sodium phosphate pH 7.5, 5 mM $MgCl_2$, 5 % (v/v) glycerol) before performing the cross-linking reaction by adding 0.1 mM TCEP and 0.2 mM bBBr (34025, Sigma) or DMSO for 5 min. Where indicated, 1 mM ATP or AMP-PNP was added for 5 min before the bBBr treatment, at 25 ºC to probe the nucleotide-bound condensin state. Quenching of the reactions was done by adding DTT to 10 mM. To cut the DNA-containing chambers of condensin, the complexes were treated with 1.5 µg TEV protease at 25 ºC for 25 min before 0.5 M salt washes. 25% of the sample was used for SDS-PAGE analysis and 75% for the analysis on agarose gel electrophoresis. For the analysis of the agarose gel, ethidium bromide was added after imaging protein fluorescence on a Typhoon FLA9500 (Cytiva) scanner. In-gel fluorescence was also used to assess the efficiency of the TEV cleavage.

Before incubation with DNA, samples were dialyzed against DTT-free buffer before cysteine cross-linking the hinge domain of Smc2 and Smc4 (Smc2K639C; Smc4V721C). The mix was then supplemented with 0.2 mM bBBr or DMSO and 0.1 mM TCEP and the quenching of the reaction was performed by adding 10 mM DTT.

XI. Cryo-EM sample preparation of Condensin-DNA complexes

Purified *Sc* condensin holo-complex double Walker B mutant ($Smc4_{E1352Q}$, $Smc2_{E1113Q}$), *Sc* condensin holo-complex, *Ct* condensin holo-complex, *Ct* condensin bonsai or *Ct* condensin ΔYcg1 complexes were diluted to a concentration of 1µM in a buffer containing 50 mM NaCl, 20 mM HEPES-NaOH pH 7.5, 5 mM $MgCl_2$, 1 mM DTT and incubated with 1 mM ATP where indicated for 10 min on ice. Then, an equimolar ratio of 50-bp dsDNA (5'-

GTTGACAGTGTCGCAACCTGCACAGGCAAGCTGCTGAGTCTGGTGTAGAC-3' annealed to the complementary strand) was added to the proteins in order to form DNA-bound condensin complexes.

3 µL of the reaction were then applied onto gold mesh 200 R2/2 or R2/1 cryo-grid (Quantifoil). Prior to sample application, the grids were plasma cleaned for 45 sec in 25 % oxygen and 75 % argon (FISHIONE Instruments, Nanoclean). Plunge freezing was performed at 100 % humidity at a temperature of 8 ºC on a FEI Vitrobot Mark IV (ThermoFisher). In order to saturate the grid support with protein DNA complexes, a wait time of 2 min was used before blotting the cryo-grid with force 3 for 1 s with Whatman blotting paper (Cytiva).

In the case of the *Sc* DNA-bound ATP-free complex, the sample was applied onto PEGylated grids. For this purpose, Ultrafoil Au 2/2 grids (Quantifoil) were used. The grids were first cleaned by glow discharging in residual air for 2 min using Pelco easiGlow. The grids were then incubated with (11-mercaptoundecyl) hexaethylenglycol for 72 h in sealed PCR tubes. The grids were collected manually with a tweezer, rinsed in 100% EtOH and air-dried directly before use. To avoid over-saturation of the holes with protein complexes, no waiting time in-between sample application and grid blotting was used. The sample was blotted with force 3, for 1 sec.

XII. Cryo-EM data acquisition

In the case of the ATP- and DNA-bound *Sc* condensin complex, a Titan Krios electron microscope (FEI) equipped with a K2 direct electron detector was used to record movies (Gatan). Data collection was performed with SerialEM (Takaba et al., 2020). In total, 6,544 micrographs were collected with defocus ranges comprised between 1.0 and 2.2 µm, in counting mode with a total dose of 40 electrons per Å^2 during exposure times of 8 s. The dose was fractionated into 40 frames. A pixel size of 1.04 Å was used, corresponding

to a magnification of 130,000×.For data processing, RELION 3.1.3 (Scheres, 2012) and CryoSPARC 3.2.0 (Punjani et al., 2017) were used.

In the case of the DNA-bound ATP-free *Sc* condensin complexes, 4,269 micrographs were collected as described above with similar parameters, except that the defocus range was set between 1.0-2.0 µm.

In the case of the *Ct* holo-complex data set, 8,762 micrographs were acquired using a Titan Krios microscope (FEI) equipped with a K3 direct electron detector in counting mode. The defocus range was set to 0.8-1.6 µm. The dose was fractionated into 40 frames. A pixel size of 0.882 Å was used, corresponding to a magnification of 105,000×. The same parameters were used to record 11,378 movies for the *Ct* bonsai condensin data set and the 5,458 movies of the *Ct* ΔYcg1 complex.

In all of the cases, the alignment and the dose weighting of the movie frames was performed with the Motioncor2 implementation from RELION, using 5 × 5 patches (Zheng et al., 2017). CTFFIND-4.1 was used to estimate the CTF parameters (Rohou and Grigorieff, 2015).

XIII. Cryo-EM data processing of the ATP- and DNA-bound condensin complex data set, and other condensin data sets

For the first round of particle picking, WARP (version 1.0.9) trained with a BoxNet2Mask model was used in order to filter out aggregates and ethane contamination (Tegunov and Cramer, 2019). The resultant set of particles was 2D classified using CryoSPARC and showed a high degree of particle heterogeneity. In order to deal with the conformational heterogeneity of the particles, the particles were split into groups corresponding to particle types and a combination of *ab initio* reconstruction and 3D classification in CryoSPARC was used. In addition, the neural network-based algorithm Topaz

(Bepler et al., 2020) was used to enrich the pool of particles of interest, before performing 3D classification in RELION.

2D classes were classified as core 1 (C1), Ycg1/Ycs4 (Y), core 2 (C2) or fragments (F). Particles corresponding to C1, Y and C2 were used as for subsequent *ab initio* reconstructions via simultaneous 3D classification. 3D reconstructions of C2 particles were reminiscent of the previously solved nucleotide-free apo state of condensin (Lee et al., 2020), whereas the C1 particles set seemed to correspond to a unknown conformation. Thus, together with the Y particle set, these particles were used for Topaz neural network training and picking. For the training, micrographs were binned with a factor 4 and the conv_127 CNN model was used. The Topaz-picked particles were subsequently 3D classified using RELION, providing 40-Å low-pass-filtered references in order to avoid high-resolution reference bias. For the same reason, references used for performing 3D classification and 3D refinements were filtered the same way. Further analysis was performed on a subset of 251,513 particles to limit the computational time of the calculations. For the C1 particles, the *ab initio* classes 1 and 2 were used as reference volume for 3D classifications. The classification using class 2 as reference resulted in an alternative DNA-binding mode of condensin. This density highly resembled the one resulting from the 3D classification using class 1 as a reference but could not reach high resolution due to the preferred orientations of the particle projections. Similarly, a subset of 276,846 Y particles was 3D classified for two successive rounds using the *ab initio* class 2 as a 3D reference. Resultant C1 and Y densities were subsequently 3D-refined and used to proceed with the analysis.

In the case of the C1 particles, the 3D volume was used as a template for cross-correlation picking in RELION, which resulted in the picking of 1,203,041 particles that underwent two successive rounds of classifications in 3D. Then, a first 3D refinement was performed on a group of 285,280 particles. In order to improve its resolution further, the resulting density was 3D classified using fixed Euler angles. As a result, 66,205 particles were

identified and submitted to two successive rounds of CTF refinement per-particle, followed by a round of Bayesian polishing. Finally, a last 3D refinement step was performed. For the Y particles, the density, identified as the Ycg1-Brn1-DNA bound complex based on its resemblance to a previously published structure (Kschonsak et al., 2017), was used as a template for neural network Topaz picking, using the above-described particles and parameters for the training. This resulted in the picking of 571,014 particles that underwent 3D classification. The most populated class, corresponding to 217,204 particles, was used for 3D refinement. Then, keeping fixed Euler angles, the same set of particles underwent a round of 3D classification that identified 114,600 particles. These particles were used for two successive rounds of CTF refinement per-particle. Finally, one round of Bayesian polishing and a final 3D refinement step were performed.

Both, the estimation of the B-factor and the sharpening, were carried on in RELION, according to the method from Rosenthal and Henderson (Rosenthal and Henderson, 2003). The nominal resolutions of the densities were estimated according to the Fourier shell correlations at 0.143 between the two half-maps refined independently. Estimations and filtering of the local resolutions estimation of 3D volumes was done according to the estimated B-factors and according to the pipelines implemented in RELION.

The processing of the data sets corresponding to other complexes followed the same general pipeline with the specificities discussed and described in the dedicated results sections.

XIV. Model building of the ATP- DNA-bound condensin complex

An atomic model of the core complex was built into a 3.5-Å B-factor-sharpened map. PDB entries 6YVU and 6YVD (Lee et al., 2020) were used as starting points for model building of the individual subunits, using a first step of rigid body fitting with ChimeraX (Pettersen et al., 2021). In order to adjust for the change in conformations of the C-

terminal extremity of Ycs4 ($Ycs4_{C\text{-}ter}$) and of the interface between the Smc2 neck region and the Brn1 HTH, these parts were fitted individually. Using the interactive sharpening tool of Coot (Emsley et al., 2010), the model was adjusted by *de novo* building of the regions corresponding to helices in $Ycs4_{C\text{-}ter}$, to missing loop regions or to Brn1-Ycs4 interactions patches. Due to the insufficient resolution at the DNA base pairs, a poly-A:T double helix was used as a model, and the Geman-McClure restraints module was used to fit this helix into the electron density. AlphaFold (Jumper et al., 2021) predicted models of the regions corresponding to the N-terminal extremity of Ycs4 (AF-Q06156-F1) and to the Brn1-Ycs4 interactions patches (AF-P38170-F1) were used to fit into the poor local resolution regions of the density. To further improve the model, the ISOLDE module from ChimeraX was used (Croll, 2018), using local resolution-filtered and DeepEMhancer (Sanchez-Garcia et al., 2021) processed maps. One automated refinement round was performed in phenix.real_space_refine against the DeepEMhancer-processed map to finalize the model. For the final model, secondary structure restraints were used (alpha, beta, B-form DNA) together with rotamer restraints and Ramachandran restraints.

In the case of the peripheral complex, model building was performed using the 3.9-Å B-factor-sharpened map. Initially, the 5OQN PDB entry was used as a reference. Model building then started by rigid body fitting of Ycg1 residues 7–233 and 234–910 and Brn1 using ChimeraX. Then, coot interactive sharpening procedures were used to improve the model. Geman-McClure restraints in Coot were used to fit the poly-A:T DNA into the electron density. To improve the model, ISOLDE and Coot were subsequently used, using local resolution-filtered and DeepEMhancer-processed maps. Finally, phenix.real_space_refine refinement was performed using secondary structure restraints (alpha, beta, B-form DNA), rotamer restraints and Ramachandran restraints. Since the map processed with DeepEMhancer showed a better fit with the model, it was chosen over the local resolution-filtered map.

XV. DNA Loop extrusion assay

λ-phage DNA (N3011, New England Biolabs) was ligated to 5'-phosphorylated and 3'-biotinylated oligos using T4 ligase. Glass slides and coverslips were prepared as previously described (Ganji et al., 2018). Glass slides were cleaned with 1 M KOH and methanol treatment before activation with 1 % (v/v) (3-aminopropyl)triethoxysilane (440140, Sigma) and 5 % (v/v) acetic acid in methanol. To PEGylate the glass surface, slides were incubated with a 130 g·L^{-1} 1:45 mixture of biotin-PEG-SVA (MW 5,000 Da) and methyl-PEG-SVA (MW 5,000 Da; BIO-SVA_M-SVA, Laysan) in 0.1 M sodium bicarbonate buffer overnight. Slides were subsequently stored at a temperature of –20 ºC under vacuum. Before use, the surfaces were incubated with 25 mM methyl-PEG$_4$-NHS (22341, ThermoFisher) and 0.1 M sodium bicarbonate. The microfluidic device was then assembled and a rinsing solution containing 50 mM Tris-HCl pH 7.5, 20 mM NaCl, 0.2 mM EDTA and 25 µg·mL^{-1} streptavidin (189730, Merck) was injected into the device. The buffers used for experiments carried out with *Ct* condensin were equilibrated to ~25 ºC before protein addition, whereas the buffers used for *Sc* condensin were kept on ice. The imaging device was then abundantly flushed with the rinsing solution before injection of 1–10 pM biotinylated λ-phage DNA in imaging solution containing 50 mM NaCl, 50 mM Tris-HCl pH 7.5, , 2.5 mM $MgCl_2$, 5 % (w/v) d-glucose, 1 mM DTT, 500 nM Sytox orange [S11368; ThermoFisher], 40 µg.mL^{-1} glucose oxidase (G2133, Sigma), 15 µg.mL^{-1} catalase (C1345, Sigma), 2 mM Trolox (238813, Sigma)) during continuous flow (2–8 µL·min^{-1}) using a PHD2000 syringe pump (Harvard Apparatus). To get rid of unbound DNA molecules before injecting condensin, the imaging solution was flushed into the device. About 20 µL of 0.15–2.0 nM condensin in the imaging buffer supplemented with 0.4–1.0 mM ATP were then introduced into the chamber. In order to avoid multiple extrusion events on the same DNA molecule, the condensin concentration was titrated down. For side-flow experiments, a second inlet was inserted and flushed at a high flow

rate (15–25 μL·min^{-1}) and adjusted to the desired angle. When condensin molecules were TEV cleaved to assign the motor and anchor function of condensin chambers, 0.3 μM TEV protease and 0.4 mM ATP were mixed with fluorescently labelled condensin molecules, just prior to imaging.

The microscopy was done using a Leica GSDIM TIRF microscope equipped with a UPlanSApo NA 1.43 160× objective (Leica) equipped with an EMCCD camera (iXon Ultra 897, Andor). The objective was used in a highly inclined and laminated optical sheet mode (HILO). Exposure times of 100-ms exposure were used with a 532-nm laser alone or in alternation with 150-ms exposures with a 643-nm laser. Imaging rates were 2.0 Hz in the case of dual-color imaging, and 7.3 Hz for the experiments were only DNA was visualized by SxO-labelling.

XVI. DNA Loop extrusion data analysis

Custom scripts in FIJI were used to process the DNA loop extrusion images (Schindelin et al., 2012). To assemble kymographs, about 15-21 pixels (100 nm*pixel^{-1}) corresponding to the SxO signal were counted along the DNA axis for every frame. A distance of 5 pixels above and below the DNA was used as a threshold to define background and noise on every frame. For segmentation, a Gaussian-blurred ($\sigma = 1.5$) copy was generated. A threshold of 1.5× noise + background was used to segment the DNA-containing area. To define the tether length of the DNA, the median distance in-between the furthermost pixel exceeding the median pixel intensity inside the DNA-segmented area minus noise, over all time points, was calculated. In order to access the DNA sections that correspond to the loop section, and the ones corresponding to the above and below sections, the per-frame intensity maxima were detected. The 9 pixels around this maximum were subsequently summed up and assigned to DNA sections. The DNA mean intensity outside the loop section was subtracted from the one of the loop sections and assigned to the other sections (above or below), in order to adjust for non-looped DNA

contained inside the looped section. The DNA intensity was converted into DNA length by normalizing the total signal intensity of the DNA to the total DNA length (48,502 bp). In order to obtain the loop extrusion rates of a user-defined time portion, the slope corresponding to the linear fit with the highest R^2 of 50 consecutive data points was used. Due to the peculiar loop extrusion profile of *Ct* condensin, the loop extrusion rates were calculated by manual selection of the 50 consecutive data points. DNA tension was estimated according to the worm-like chain model. For the particular examples of the TEV cleavage experiments, some DNA molecules were left out from the analysis: when (a) multiple condensin were detected on a DNA (~15 %), b) no loop was formed (~30%), c) if the condensin molecule responsible for the loop bleached before loop rupture (~10 %), d) if the condensin responsible for the loop was not fluorescent (~10 %), or e) when a loop formed but did not resolve (~20%). In order to assign the slippage behavior of *Ct* condensin and its frequent change of extrusion directionality, a cut-off of 5 kb for a minimum duration of 10 sec was used.

XVII. TFIIIC-Condensin binding assay and cryo-EM sample preparation

Sc TFIIIC, Condensin and DNA were mixed on ice in equimolar ratios (400 nM) n a buffer containing: 20 mM HEPESKOH pH 7.5, 1 mM DTT, 50 mM NaCl. The DNA used corresponded to the His_tH(GUG)E2 gene including upstream sequences for *in vitro* transcription and TFIIIC-TFIIIB binding (GTATTACTCGAGCCCG-TAATACAACAGTTCTCCATGAAAAGTCGCCATCT**T**
AGTATAGTGGTTAGTACACATCGTTGTGGCCGATGAAACCCT**GGTTCGAT-TCT**AGGAGATGGCATTTT, with the τA and τB binding sequences highlighted in bold). The sample was diluted 10 times in the above-mentioned buffer to a concentration of 40 nM and applied to the sample holder of a Refeyn Two MP mass photometer (Refeyn). In a separate experiment, TFIIIC, condensin and DNA were mixed in 1 µM

equimolar radios and applied onto a cryo-grid R2/2 Au mesh 200 (Ultrafoil). The sample application and blotting parameters were kept identical to the ones described for condensin-DNA complexes described in the above section.

XVIII. SELEX experiments

The SELEX experiment was based on Riley et al 2014, Methods in Molecular Biology, and Slattery et al 2011, Cell. A single-stranded, fluorescein-labelled SELEX library with the sequence (5'-GAGTTCTACAGTCCGACGATCCGC NNNNNNNNNNNNNNNNCCTGAATTCTCGGGTGCCA-3') was ordered from IDT. A dsSELEX library was generated by Klenow extension (Thermofisher) of the 5'-TGGCACCCGAGAATTCCA-3' primer, following the manufacturer's instructions. The dsSELEX library was purified using a PCR purification kit (Qiagen) with an elution in water and a systematic quantification on the Nanodrop spectrophotometer (NanodropONe, ThermoScientific). ThedsSELEX library was incubated at a concentration of 0.2 mM with 0.1 mM *Ct* Ycg1-Brn1 in the following binding buffer:10 mM HEPES-KOH pH 7.5, 125 mM NaCl, 5 mM $MgCl_2$, 1 mM DTT) for 30 min on ice, before running the mix on a NativePAGE 416%, Bis-Tris, 1.0 mm (ThermoFisher). For the negative control experiment, the dsSELEX library was mixed with buffer only. Fluorescein-labelled DNA was imaged on a Typhoon FLA 9500 scanner and the up-shifted band was cut out from the gel. DNA extraction from the polyacrylamide gel was performed by incubating the gel in diffusion buffer (0.5 M ammonium acetate, 10 mM magnesium acetate, 1 mM EDTA, pH 8, 0.1% SDS), mechanic dissociation using a Pasteur pipette and incubation at 50 °C for 20 min. DNA was then extracted using a PCR purification kit (Qiagen) and amplified by PCR using Phusion polymerase (NEB), the primers 5'-TGGCACCCGAGAATTCCA-3'; 5'-GAGTTCTACAGTCCGACGAT-3' and the following PCR conditions: 95 °C for 3 min, 10 cycles (95 °C for 15 s, 55 °C for 15 s, 72 °C for 30 s), and 72

°C for 1 minute. DNA was purified using a PCR purification kit (Qiagen) and used both for Illumina sequencing library preparation and to perform a second round of the binding experiment using the same conditions as above. Illumina libraries were prepared from the initial dsSELEX library (R0 condition) and, the first and the second round of SELEX binding experiments (R1 and R2 condition respectively) with the NEBNext® Ultra™ II DNA Library Prep Kit for Illumina (NEB) following the manufacturer's instructions. The libraries were paired-end sequenced (1×75 bp) on a Nextseq500 platform at the EMBL GeneCore facility. The quality of the 33 317 635 reads (R0), 37 848 357 reads (R1) and 36 099 514 reads (R2) were validated with fastqc (version 0.11.8). Bioinformatic analysis was performed with R (version 4.1.2) using the SELEX package (version 1.26.0, Riley et al 2014, methods in molecular biology). An affinity score cutoff of 0.6, 0.7, 0.8 and 0.9 was applied for subsequent motif discovery. Enriched motifs were determined using the MEME suite version 5.4.1 with the following parameters: classic mode, zero or one occurrence per sequence, minimum width = 4, maximum width = 9, minimum sites each motif have = 30.

XIX. Di-nucleosome library reconstitution

Two Widom 601 sequences, separated by 15-60 bps in 1-bp steps, were inserted into a pEX A2 vector. The DNA sequences containing the two 601 sequences and 30 bp of DNA on one side of one of the 601 sequences were amplified via PCR with a fluorescein-labelled primer and purified via anion exchange chromatography. Di-nucleosomes were reconstituted via salt gradient dialysis (SGD) from 2M NaCl, 25 mM Tris-HCl pH 7.5, 1 mM EDTA to 50 mM NaCl, 25 mM Tris-HCl pH 7.5, 1 mM EDTA. SGD was performed over the course of 16 h at 4 °C using a peristaltic pump. To ensure complete reconstitution of the library, a competitor DNA was added in equimolar amounts relative to the 601 sequence, and *Xenopus* histone octamers were titrated.

XX. Di-nucleosome library binding assay

The fully reconstituted di-nucleosome library was used for binding experiments at a concentration of 7.5 nM in 25 mM HEPESKOH pH 8, 60 mM KCl, 7 % glycerol (v/v), 1 mM DTT. The binding reactions were incubated on ice for 20 minutes and the products were analyzed on a NativePAGE, 4 to 16 %, Bis-Tris, 1 mm mini protein gel, at 100 V, 4 °C. Gels were imaged with a GE Typhoon FLA 9000 scanner. The shifted and unshifted di-nucleosome bands were cut from the conditions in which the di-nucleosome library was partially shifted. The library DNA was extracted using a PCR purification kit (Qiagen), amplified by PCR and sent for Illumina sequencing.

Bibliography

Alt, A., Dang, H.Q., Wells, O.S., Polo, L.M., Smith, M.A., McGregor, G.A., Welte, T., Lehmann, A.R., Pearl, L.H., Murray, J.M., Oliver, A.W., 2017. Specialized interfaces of Smc5/6 control hinge stability and DNA association. Nat. Commun. 8, 14011. https://doi.org/10.1038/ncomms14011

Bai, X., Peirson, B.N., Dong, F., Xue, C., Makaroff, C.A., 1999. Isolation and characterization of SYN1, a RAD21-like gene essential for meiosis in Arabidopsis. Plant Cell 11, 417–430. https://doi.org/10.1105/tpc.11.3.417

Ball, A.R., Schmiesing, J.A., Zhou, C., Gregson, H.C., Okada, Y., Doi, T., Yokomori, K., 2002. Identification of a Chromosome-Targeting Domain in the Human Condensin Subunit CNAP1/hCAP-D2/Eg7. Mol. Cell. Biol. 22, 5769–5781. https://doi.org/10.1128/MCB.22.16.5769-5781.2002

Banigan, E.J., Mirny, L.A., 2020. The interplay between asymmetric and symmetric DNA loop extrusion. eLife 9, e63528. https://doi.org/10.7554/eLife.63528

Bauer, B.W., Davidson, I.F., Canena, D., Wutz, G., Tang, W., Litos, G., Horn, S., Hinterdorfer, P., Peters, J.-M., 2021. Cohesin mediates DNA loop extrusion by a "swing and clamp" mechanism. Cell 184, 5448-5464.e22. https://doi.org/10.1016/j.cell.2021.09.016

Bauer, C.R., Hartl, T.A., Bosco, G., 2012. Condensin II Promotes the Formation of Chromosome Territories by Inducing Axial Compaction of Polyploid Interphase Chromosomes. PLOS Genet. 8, e1002873. https://doi.org/10.1371/journal.pgen.1002873

Beattie, T.R., Bell, S.D., 2011. Molecular machines in archaeal DNA replication. Curr. Opin. Chem. Biol. 15, 614–619. https://doi.org/10.1016/j.cbpa.2011.07.017

Bepler, T., Kelley, K., Noble, A.J., Berger, B., 2020. Topaz-Denoise: general deep denoising models for cryoEM and cryoET. Nat. Commun. 11, 5208. https://doi.org/10.1038/s41467-020-18952-1

Bernard, P., Vanoosthuyse, V., 2015. Does transcription play a role in creating a condensin binding site? Transcription 6, 12–16. https://doi.org/10.1080/21541264.2015.1012980

Bhatt, A.M., Lister, C., Page, T., Fransz, P., Findlay, K., Jones, G.H., Dickinson, H.G., Dean, C., 1999. The DIF1 gene of Arabidopsis is required for meiotic chromosome segregation and belongs to the REC8/RAD21 cohesin gene family. Plant J. Cell Mol. Biol. 19, 463–472. https://doi.org/10.1046/j.1365-313x.1999.00548.x

Bolzer, A., Kreth, G., Solovei, I., Koehler, D., Saracoglu, K., Fauth, C., Müller, S., Eils, R., Cremer, C., Speicher, M.R., Cremer, T., 2005. Three-dimensional maps of all chromosomes in human male fibroblast nuclei and prometaphase rosettes. PLoS Biol. 3, e157. https://doi.org/10.1371/journal.pbio.0030157

Briand, N., Collas, P., 2020. Lamina-associated domains: peripheral matters and internal affairs. Genome Biol. 21, 85. https://doi.org/10.1186/s13059-020-02003-5

Bürmann, F., Basfeld, A., Vazquez Nunez, R., Diebold-Durand, M.-L., Wilhelm, L., Gruber, S., 2017. Tuned SMC Arms Drive Chromosomal Loading of Prokaryotic Condensin. Mol. Cell 65, 861-872.e9. https://doi.org/10.1016/j.molcel.2017.01.026

Bürmann, F., Funke, L.F.H., Chin, J.W., Löwe, J., 2021. Cryo-EM structure of MukBEF reveals DNA loop entrapment at chromosomal unloading sites. Mol. Cell 81, 4891-4906.e8. https://doi.org/10.1016/j.molcel.2021.10.011

Bürmann, F., Lee, B.-G., Than, T., Sinn, L., O'Reilly, F.J., Yatskevich, S., Rappsilber, J., Hu, B., Nasmyth, K., Löwe, J., 2019. A folded conformation of MukBEF and cohesin. Nat. Struct. Mol. Biol. 26, 227–236. https://doi.org/10.1038/s41594-019-0196-z

Chubb, J.R., Boyle, S., Perry, P., Bickmore, W.A., 2002. Chromatin motion is constrained by association with nuclear compartments in human cells. Curr. Biol. CB 12, 439–445. https://doi.org/10.1016/s0960-9822(02)00695-4

Collier, J.E., Lee, B.-G., Roig, M.B., Yatskevich, S., Petela, N.J., Metson, J., Voulgaris, M., Gonzalez Llamazares, A., Löwe, J., Nasmyth, K.A., 2020. Transport of DNA within cohesin involves clamping on top of engaged heads by Scc2 and entrapment within the ring by Scc3. eLife 9, e59560. https://doi.org/10.7554/eLife.59560

Croll, T.I., 2018. ISOLDE: a physically realistic environment for model building into low-resolution electron-density maps. Acta Crystallogr. Sect. Struct. Biol. 74, 519–530. https://doi.org/10.1107/S2059798318002425

Csankovszki, G., Collette, K., Spahl, K., Carey, J., Snyder, M., Petty, E., Patel, U., Tabuchi, T., Liu, H., McLeod, I., Thompson, J., Sarkesik, A., Yates, J., Meyer, B.J., Hagstrom, K., 2009. Three Distinct Condensin Complexes Control C. elegans Chromosome Dynamics. Curr. Biol. CB 19, 9–19. https://doi.org/10.1016/j.cub.2008.12.006

Cuylen, S., Metz, J., Hruby, A., Haering, C.H., 2013. Entrapment of chromosomes by condensin rings prevents their breakage during cytokinesis. Dev. Cell 27, 469–478. https://doi.org/10.1016/j.devcel.2013.10.018

Dame, R.T., 2005. The role of nucleoid-associated proteins in the organization and compaction of bacterial chromatin. Mol. Microbiol. 56, 858–870. https://doi.org/10.1111/j.1365-2958.2005.04598.x

Dame, R.T., Noom, M.C., Wuite, G.J.L., 2006. Bacterial chromatin organization by H-NS protein unravelled using dual DNA manipulation. Nature 444, 387–390. https://doi.org/10.1038/nature05283

Dame, R.T., Rashid, F.-Z.M., Grainger, D.C., 2020. Chromosome organization in bacteria: mechanistic insights into genome structure and function. Nat. Rev. Genet. 21, 227–242. https://doi.org/10.1038/s41576-019-0185-4

Dame, R.T., Tark-Dame, M., Schiessel, H., 2011. A physical approach to segregation and folding of the Caulobacter crescentus genome. Mol. Microbiol. 82, 1311–1315. https://doi.org/10.1111/j.1365-2958.2011.07898.x

Dame, R.T., Wyman, C., Goosen, N., 2000. H-NS mediated compaction of DNA visualised by atomic force microscopy. Nucleic Acids Res. 28, 3504–3510. https://doi.org/10.1093/nar/28.18.3504

Dandey, V.P., Budell, W.C., Wei, H., Bobe, D., Maruthi, K., Kopylov, M., Eng, E.T., Kahn, P.A., Hinshaw, J.E., Kundu, N., Nimigean, C.M., Fan, C., Sukomon, N., Darst, S.A., Saecker, R.M., Chen, J., Malone, B., Potter, C.S., Carragher, B., 2020. Time-resolved cryo-EM using Spotiton. Nat. Methods 17, 897–900. https://doi.org/10.1038/s41592-020-0925-6

Danilova, O., Reyes-Lamothe, R., Pinskaya, M., Sherratt, D., Possoz, C., 2007. MukB colocalizes with the oriC region and is required for organization of the two Escherichia coli chromosome arms into separate cell halves. Mol. Microbiol. 65, 1485–1492. https://doi.org/10.1111/j.1365-2958.2007.05881.x

Datta, S., Lecomte, L., Haering, C.H., 2020. Structural insights into DNA loop extrusion by SMC protein complexes. Curr. Opin. Struct. Biol. 65, 102–109. https://doi.org/10.1016/j.sbi.2020.06.009

Davidson, I.F., Barth, R., Zaczek, M., Torre, J. van der, Tang, W., Nagasaka, K., Janissen, R., Kerssemakers, J., Wutz, G., Dekker, C., Peters, J.-M., 2022. CTCF is a DNA-tension-dependent barrier to cohesin-mediated DNA loop extrusion. https://doi.org/10.1101/2022.09.08.507093

Davidson, I.F., Bauer, B., Goetz, D., Tang, W., Wutz, G., Peters, J.-M., 2019. DNA loop extrusion by human cohesin. Science 366, 1338–1345. https://doi.org/10.1126/science.aaz3418

Davidson, I.F., Peters, J.-M., 2021. Genome folding through loop extrusion by SMC complexes. Nat. Rev. Mol. Cell Biol. 22, 445–464. https://doi.org/10.1038/s41580-021-00349-7

Dekker, J., Rippe, K., Dekker, M., Kleckner, N., 2002. Capturing chromosome conformation. Science 295, 1306–1311. https://doi.org/10.1126/science.1067799

Diebold-Durand, M.-L., Lee, H., Ruiz Avila, L.B., Noh, H., Shin, H.-C., Im, H., Bock, F.P., Bürmann, F., Durand, A., Basfeld, A., Ham, S., Basquin, J., Oh, B.-H., Gruber, S., 2017. Structure of Full-Length SMC and Rearrangements Required for Chromosome Organization. Mol. Cell 67, 334-347.e5. https://doi.org/10.1016/j.molcel.2017.06.010

Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., Ren, B., 2012. Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature 485, 376–380. https://doi.org/10.1038/nature11082

Dubochet, J., Adrian, M., Chang, J.J., Homo, J.C., Lepault, J., McDowall, A.W., Schultz, P., 1988. Cryo-electron microscopy of vitrified specimens. Q. Rev. Biophys. 21, 129–228. https://doi.org/10.1017/s0033583500004297

Earnshaw, W.C., Heck, M.M., 1985. Localization of topoisomerase II in mitotic chromosomes. J. Cell Biol. 100, 1716–1725. https://doi.org/10.1083/jcb.100.5.1716

Eeftens, J.M., Katan, A.J., Kschonsak, M., Hassler, M., de Wilde, L., Dief, E.M., Haering, C.H., Dekker, C., 2016. Condensin Smc2-Smc4 Dimers Are Flexible and Dynamic. Cell Rep. 14, 1813–1818. https://doi.org/10.1016/j.celrep.2016.01.063

Emsley, P., Lohkamp, B., Scott, W.G., Cowtan, K., 2010. Features and development of Coot. Acta Crystallogr. D Biol. Crystallogr. 66, 486–501. https://doi.org/10.1107/S0907444910007493

Ganji, M., Shaltiel, I.A., Bisht, S., Kim, E., Kalichava, A., Haering, C.H., Dekker, C., 2018. Real-time imaging of DNA loop extrusion by condensin. Science eaar7831. https://doi.org/10.1126/science.aar7831

Gasser, S.M., Laroche, T., Falquet, J., Boy de la Tour, E., Laemmli, U.K., 1986. Metaphase chromosome structure. Involvement of topoisomerase II. J. Mol. Biol. 188, 613–629. https://doi.org/10.1016/s0022-2836(86)80010-9

Gibcus, J.H., Samejima, K., Goloborodko, A., Samejima, I., Naumova, N., Nuebler, J., Kanemaki, M.T., Xie, L., Paulson, J.R., Earnshaw, W.C., Mirny, L.A., Dekker, J., 2018. A pathway for mitotic chromosome formation. Science 359, eaao6135. https://doi.org/10.1126/science.aao6135

Glaeser, R.M., 2016. Specimen Behavior in the Electron Beam. Methods Enzymol. 579, 19–50. https://doi.org/10.1016/bs.mie.2016.04.010

Golfier, S., Quail, T., Kimura, H., Brugués, J., 2020. Cohesin and condensin extrude DNA loops in a cell cycle-dependent manner. eLife 9, e53885. https://doi.org/10.7554/eLife.53885

Goloborodko, A., Imakaev, M.V., Marko, J.F., Mirny, L., 2016. Compaction and segregation of sister chromatids via active loop extrusion. eLife 5, e14864. https://doi.org/10.7554/eLife.14864

Green, L.C., Kalitsis, P., Chang, T.M., Cipetic, M., Kim, J.H., Marshall, O., Turnbull, L., Whitchurch, C.B., Vagnarelli, P., Samejima, K., Earnshaw, W.C., Choo, K.H.A., Hudson, D.F., 2012. Contrasting roles of condensin I and condensin II in mitotic chromosome formation. J. Cell Sci. 125, 1591–1604. https://doi.org/10.1242/jcs.097790

Griese, J.J., Hopfner, K.-P., 2011. Structure and DNA-binding activity of the Pyrococcus furiosus SMC protein hinge domain. Proteins 79, 558–568. https://doi.org/10.1002/prot.22903

Griese, J.J., Witte, G., Hopfner, K.-P., 2010. Structure and DNA binding activity of the mouse condensin hinge domain highlight common and diverse features of SMC proteins. Nucleic Acids Res. 38, 3454–3465. https://doi.org/10.1093/nar/gkq038

Gruber, S., Arumugam, P., Katou, Y., Kuglitsch, D., Helmhart, W., Shirahige, K., Nasmyth, K., 2006. Evidence that loading of cohesin onto chromosomes involves opening of its SMC hinge. Cell 127, 523–537. https://doi.org/10.1016/j.cell.2006.08.048

Gruber, S., Haering, C.H., Nasmyth, K., 2003. Chromosomal cohesin forms a ring. Cell 112, 765–777. https://doi.org/10.1016/s0092-8674(03)00162-4

Guacci, V., Koshland, D., Strunnikov, A., 1997. A Direct Link between Sister Chromatid Cohesion and Chromosome Condensation Revealed through the Analysis of MCD1 in S. cerevisiae. Cell 91, 47–57.

Haase, J., Chen, R., Parker, W.M., Bonner, M.K., Jenkins, L.M., Kelly, A.E., 2022. The TFIIH complex is required to establish and maintain mitotic chromosome structure. eLife 11, e75475. https://doi.org/10.7554/eLife.75475

Haering, C.H., Löwe, J., Hochwagen, A., Nasmyth, K., 2002. Molecular architecture of SMC proteins and the yeast cohesin complex. Mol. Cell 9, 773–788. https://doi.org/10.1016/s1097-2765(02)00515-4

Haeusler, R.A., Pratt-Hyatt, M., Good, P.D., Gipson, T.A., Engelke, D.R., 2008. Clustering of yeast tRNA genes is mediated by specific association of condensin with tRNA gene transcription complexes. Genes Dev. 22, 2204–2214. https://doi.org/10.1101/gad.1675908

Hallett, S.T., Harry, I.C., Schellenberger, P., Zhou, L., Cronin, N.B., Baxter, J., Etheridge, T.J., Murray, J.M., Oliver, A.W., 2021. Cryo-EM structure of the Smc5/6 holo-complex. https://doi.org/10.1101/2021.11.25.470006

Hassler, M., Shaltiel, I.A., Haering, C.H., 2018. Towards a Unified Model of SMC Complex Function. Curr. Biol. CB 28, R1266–R1281. https://doi.org/10.1016/j.cub.2018.08.034

Hassler, M., Shaltiel, I.A., Kschonsak, M., Simon, B., Merkel, F., Thärichen, L., Bailey, H.J., Macošek, J., Bravo, S., Metz, J., Hennig, J., Haering, C.H., 2019. Structural Basis of an Asymmetric Condensin ATPase Cycle. Mol. Cell 74, 1175-1188.e9. https://doi.org/10.1016/j.molcel.2019.03.037

Henneman, B., Emmerik, C. van, Ingen, H. van, Dame, R.T., 2018. Structure and function of archaeal histones. PLOS Genet. 14, e1007582. https://doi.org/10.1371/journal.pgen.1007582

Heun, P., Laroche, T., Shimada, K., Furrer, P., Gasser, S.M., 2001. Chromosome dynamics in the yeast interphase nucleus. Science 294, 2181–2186. https://doi.org/10.1126/science.1065366

Higashi, T.L., Eickhoff, P., Simoes, J.S., Locke, J., Nans, A., Flynn, H.R., Snijders, A.P., Papageorgiou, G., O'Reilly, N., Chen, Z.A., O'Reilly, F.J., Rappsilber, J., Costa, A., Uhlmann, F., 2020. A Structure-Based Mechanism for DNA Entry into the Cohesin Ring. bioRxiv 2020.04.21.052944. https://doi.org/10.1101/2020.04.21.052944

Higashi, T.L., Uhlmann, F., 2022. SMC complexes: Lifting the lid on loop extrusion. Curr. Opin. Cell Biol. 74, 13–22. https://doi.org/10.1016/j.ceb.2021.12.003

Hill, V.K., Kim, J.-S., Waldman, T., 2016. Cohesin mutations in human cancer. Biochim. Biophys. Acta BBA - Rev. Cancer 1866, 1–11. https://doi.org/10.1016/j.bbcan.2016.05.002

Hirano, T., 2017. Capturing condensin in chromosomes. Nat. Genet. 49, 1419–1420. https://doi.org/10.1038/ng.3962

Hirota, T., Gerlich, D., Koch, B., Ellenberg, J., Peters, J.-M., 2004. Distinct functions of condensin I and II in mitotic chromosome assembly. J. Cell Sci. 117, 6435–6445. https://doi.org/10.1242/jcs.01604

Holm, C., Goto, T., Wang, J.C., Botstein, D., 1985. DNA topoisomerase II is required at the time of mitosis in yeast. Cell 41, 553–563. https://doi.org/10.1016/S0092-8674(85)80028-3

Hołówka, J., Zakrzewska-Czerwińska, J., 2020. Nucleoid Associated Proteins: The Small Organizers That Help to Cope With Stress. Front. Microbiol. 11, 590. https://doi.org/10.3389/fmicb.2020.00590

Hopfner, K.-P., Craig, L., Moncalian, G., Zinkel, R.A., Usui, T., Owen, B.A.L., Karcher, A., Henderson, B., Bodmer, J.-L., McMurray, C.T., Carney, J.P., Petrini, J.H.J., Tainer, J.A., 2002. The Rad50 zinc-hook is a structure joining Mre11 complexes in DNA recombination and repair. Nature 418, 562–566. https://doi.org/10.1038/nature00922

Hou, C., Li, L., Qin, Z.S., Corces, V.G., 2012. Gene density, transcription, and insulators contribute to the partition of the Drosophila genome into physical domains. Mol. Cell 48, 471–484. https://doi.org/10.1016/j.molcel.2012.08.031

Huynh, V.A.T., Robinson, P.J.J., Rhodes, D., 2005. A method for the in vitro reconstitution of a defined "30 nm" chromatin fibre containing stoichiometric amounts of the linker histone. J. Mol. Biol. 345, 957–968. https://doi.org/10.1016/j.jmb.2004.10.075

Jonić, S., 2016. Cryo-electron Microscopy Analysis of Structurally Heterogeneous Macromolecular Complexes. Comput. Struct. Biotechnol. J. 14, 385–390. https://doi.org/10.1016/j.csbj.2016.10.002

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S.A.A., Ballard, A.J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T., Petersen, S., Reiman, D., Clancy, E., Zielinski, M., Steinegger, M., Pacholska, M., Berghammer, T., Bodenstein, S., Silver, D., Vinyals, O., Senior, A.W., Kavukcuoglu, K., Kohli, P., Hassabis, D., 2021. Highly accurate protein structure prediction with AlphaFold. Nature 596, 583–589. https://doi.org/10.1038/s41586-021-03819-2

Kagey, M.H., Newman, J.J., Bilodeau, S., Zhan, Y., Orlando, D.A., van Berkum, N.L., Ebmeier, C.C., Goossens, J., Rahl, P.B., Levine, S.S., Taatjes, D.J., Dekker, J.,

Young, R.A., 2010. Mediator and cohesin connect gene expression and chromatin architecture. Nature 467, 430–435. https://doi.org/10.1038/nature09380

Kim, E., Gonzalez, A.M., Pradhan, B., van der Torre, J., Dekker, C., 2022. Condensin-driven loop extrusion on supercoiled DNA. Nat. Struct. Mol. Biol. 29, 719–727. https://doi.org/10.1038/s41594-022-00802-x

Kim, E., Kerssemakers, J., Shaltiel, I.A., Haering, C.H., Dekker, C., 2020. DNA-loop extruding condensin complexes can traverse one another. Nature 579, 438–442. https://doi.org/10.1038/s41586-020-2067-5

Kim, J.H., Zhang, T., Wong, N.C., Davidson, N., Maksimovic, J., Oshlack, A., Earnshaw, W.C., Kalitsis, P., Hudson, D.F., 2013. Condensin I associates with structural and gene regulatory regions in vertebrate chromosomes. Nat. Commun. 4, 2537. https://doi.org/10.1038/ncomms3537

Kim, Y., Shi, Z., Zhang, H., Finkelstein, I.J., Yu, H., 2019. Human cohesin compacts DNA by loop extrusion. Science 366, 1345–1349. https://doi.org/10.1126/science.aaz4475

Kireeva, N., Lakonishok, M., Kireev, I., Hirano, T., Belmont, A.S., 2004. Visualization of early chromosome condensation : a hierarchical folding, axial glue model of chromosome structure. J. Cell Biol. 166, 775–785. https://doi.org/10.1083/jcb.200406049

Kschonsak, M., Haering, C.H., 2015. Shaping mitotic chromosomes: From classical concepts to molecular mechanisms. BioEssays News Rev. Mol. Cell. Dev. Biol. 37, 755–766. https://doi.org/10.1002/bies.201500020

Kschonsak, M., Merkel, F., Bisht, S., Metz, J., Rybin, V., Hassler, M., Haering, C.H., 2017. Structural Basis for a Safety-Belt Mechanism That Anchors Condensin to Chromosomes. Cell 171, 588-600.e24. https://doi.org/10.1016/j.cell.2017.09.008

Le, T.B., Laub, M.T., 2016. Transcription rate and transcript length drive formation of chromosomal interaction domain boundaries. EMBO J. 35, 1582–1595. https://doi.org/10.15252/embj.201593561

Le, T.B.K., Imakaev, M.V., Mirny, L.A., Laub, M.T., 2013. High-resolution mapping of the spatial organization of a bacterial chromosome. Science 342, 731–734. https://doi.org/10.1126/science.1242059

Lee, B.-G., Merkel, F., Allegretti, M., Hassler, M., Cawood, C., Lecomte, L., O'Reilly, F.J., Sinn, L.R., Gutierrez-Escribano, P., Kschonsak, M., Bravo, S., Nakane, T., Rappsilber, J., Aragon, L., Beck, M., Löwe, J., Haering, C.H., 2020. Cryo-EM structures of holo condensin reveal a subunit flip-flop mechanism. Nat. Struct. Mol. Biol. https://doi.org/10.1038/s41594-020-0457-x

Li, B., Zhu, D., Shi, H., Zhang, X., 2021. Effect of charge on protein preferred orientation at the air-water interface in cryo-electron microscopy. J. Struct. Biol. 213, 107783. https://doi.org/10.1016/j.jsb.2021.107783

Li, X., Mooney, P., Zheng, S., Booth, C.R., Braunfeld, M.B., Gubbens, S., Agard, D.A., Cheng, Y., 2013. Electron counting and beam-induced motion correction

enable near-atomic-resolution single-particle cryo-EM. Nat. Methods 10, 584–590. https://doi.org/10.1038/nmeth.2472

Li, Y., Haarhuis, J.H.I., Sedeño Cacciatore, Á., Oldenkamp, R., van Ruiten, M.S., Willems, L., Teunissen, H., Muir, K.W., de Wit, E., Rowland, B.D., Panne, D., 2020. The structural basis for cohesin–CTCF-anchored loops. Nature 578, 472–476. https://doi.org/10.1038/s41586-019-1910-z

Li, Y., Muir, K.W., Bowler, M.W., Metz, J., Haering, C.H., Panne, D., 2018. Structural basis for Scc3-dependent cohesin recruitment to chromatin. eLife 7, e38356. https://doi.org/10.7554/eLife.38356

Lieberman-Aiden, E., van Berkum, N.L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O., Sandstrom, R., Bernstein, B., Bender, M.A., Groudine, M., Gnirke, A., Stamatoyannopoulos, J., Mirny, L.A., Lander, E.S., Dekker, J., 2009. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science 326, 289–293. https://doi.org/10.1126/science.1181369

Lioy, V.S., Cournac, A., Marbouty, M., Duigou, S., Mozziconacci, J., Espéli, O., Boccard, F., Koszul, R., 2018. Multiscale Structuring of the E. coli Chromosome by Nucleoid-Associated and Condensin Proteins. Cell 172, 771-783.e18. https://doi.org/10.1016/j.cell.2017.12.027

Liu, J., Krantz, I.D., 2008. Cohesin and Human Disease. Annu. Rev. Genomics Hum. Genet. 9, 303–320. https://doi.org/10.1146/annurev.genom.9.081307.164211

Liu, W., Tanasa, B., Tyurina, O.V., Zhou, T.Y., Gassmann, R., Liu, W.T., Ohgi, K.A., Benner, C., Garcia-Bassets, I., Aggarwal, A.K., Desai, A., Dorrestein, P.C., Glass, C.K., Rosenfeld, M.G., 2010. PHF8 mediates histone H4 lysine 20 demethylation events involved in cell cycle progression. Nature 466, 508–512. https://doi.org/10.1038/nature09272

Losada, A., Hirano, M., Hirano, T., 1998. Identification of Xenopus SMC protein complexes required for sister chromatid cohesion. Genes Dev. 12, 1986–1997. https://doi.org/10.1101/gad.12.13.1986

Lowary, P.T., Widom, J., 1998. New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning11Edited by T. Richmond. J. Mol. Biol. 276, 19–42. https://doi.org/10.1006/jmbi.1997.1494

Luger, K., Mäder, A.W., Richmond, R.K., Sargent, D.F., Richmond, T.J., 1997. Crystal structure of the nucleosome core particle at 2.8 Å resolution. Nature 389, 251–260. https://doi.org/10.1038/38444

Maddox, P.S., Portier, N., Desai, A., Oegema, K., 2006. Molecular analysis of mitotic chromosome condensation using a quantitative time-resolved fluorescence microscopy assay. Proc. Natl. Acad. Sci. 103, 15097–15102. https://doi.org/10.1073/pnas.0606993103

Mäeots, M.-E., Lee, B., Nans, A., Jeong, S.-G., Esfahani, M.M.N., Ding, S., Smith, D.J., Lee, C.-S., Lee, S.S., Peter, M., Enchev, R.I., 2020. Modular microfluidics

enables kinetic insight from time-resolved cryo-EM. Nat. Commun. 11, 3465. https://doi.org/10.1038/s41467-020-17230-4

Maeshima, K., Hihara, S., Eltsov, M., 2010. Chromatin structure: does the 30-nm fibre exist in vivo? Curr. Opin. Cell Biol. 22, 291–297. https://doi.org/10.1016/j.ceb.2010.03.001

Maeshima, K., Laemmli, U.K., 2003. A two-step scaffolding model for mitotic chromosome assembly. Dev. Cell 4, 467–480. https://doi.org/10.1016/s1534-5807(03)00092-3

Malär, A.A., Wili, N., Völker, L.A., Kozlova, M.I., Cadalbert, R., Däpp, A., Weber, M.E., Zehnder, J., Jeschke, G., Eckert, H., Böckmann, A., Klose, D., Mulkidjanian, A.Y., Meier, B.H., Wiegand, T., 2021. Spectroscopic glimpses of the transition state of ATP hydrolysis trapped in a bacterial DnaB helicase. Nat. Commun. 12, 5293. https://doi.org/10.1038/s41467-021-25599-z

Male, G., von Appen, A., Glatt, S., Taylor, N.M.I., Cristovao, M., Groetsch, H., Beck, M., Müller, C.W., 2015. Architecture of TFIIIC and its role in RNA polymerase III pre-initiation complex assembly. Nat. Commun. 6, 7387. https://doi.org/10.1038/ncomms8387

Marbouty, M., Le Gall, A., Cattoni, D.I., Cournac, A., Koh, A., Fiche, J.-B., Mozziconacci, J., Murray, H., Koszul, R., Nollmann, M., 2015. Condensin- and Replication-Mediated Bacterial Chromosome Folding and Origin Condensation Revealed by Hi-C and Super-resolution Imaging. Mol. Cell 59, 588–602. https://doi.org/10.1016/j.molcel.2015.07.020

Mazumdar, M., Sundareshan, S., Misteli, T., 2004. Human chromokinesin KIF4A functions in chromosome condensation and segregation. J. Cell Biol. 166, 613–620. https://doi.org/10.1083/jcb.200401142

McDowall, A.W., Smith, J.M., Dubochet, J., 1986. Cryo-electron microscopy of vitrified chromosomes in situ. EMBO J. 5, 1395–1402. https://doi.org/10.1002/j.1460-2075.1986.tb04373.x

Menolfi, D., Delamarre, A., Lengronne, A., Pasero, P., Branzei, D., 2015. Essential Roles of the Smc5/6 Complex in Replication through Natural Pausing Sites and Endogenous DNA Damage Tolerance. Mol. Cell 60, 835–846. https://doi.org/10.1016/j.molcel.2015.10.023

Michaelis, C., Ciosk, R., Nasmyth, K., 1997. Cohesins: chromosomal proteins that prevent premature separation of sister chromatids. Cell 91, 35–45.

Muir, K.W., Li, Y., Weis, F., Panne, D., 2020. The structure of the cohesin ATPase elucidates the mechanism of SMC–kleisin ring opening. Nat. Struct. Mol. Biol. 27, 233–239. https://doi.org/10.1038/s41594-020-0379-7

Muñoz, S., Passarelli, F., Uhlmann, F., 2020. Conserved roles of chromatin remodellers in cohesin loading onto chromatin. Curr. Genet. 66, 951–956. https://doi.org/10.1007/s00294-020-01075-x

Murayama, Y., Uhlmann, F., 2015. DNA Entry into and Exit out of the Cohesin Ring by an Interlocking Gate Mechanism. Cell 163, 1628–1640. https://doi.org/10.1016/j.cell.2015.11.030

Nanni, L., Ceri, S., Logie, C., 2020. Spatial patterns of CTCF sites define the anatomy of TADs and their boundaries. Genome Biol. 21, 197. https://doi.org/10.1186/s13059-020-02108-x

Nasmyth, K., 2001. Disseminating the genome: joining, resolving, and separating sister chromatids during mitosis and meiosis. Annu. Rev. Genet. 35, 673–745. https://doi.org/10.1146/annurev.genet.35.102401.091334

Nasmyth, K., Peters, J.M., Uhlmann, F., 2000. Splitting the chromosome: cutting the ties that bind sister chromatids. Science 288, 1379–1385. https://doi.org/10.1126/science.288.5470.1379

Naumova, N., Imakaev, M., Fudenberg, G., Zhan, Y., Lajoie, B.R., Mirny, L.A., Dekker, J., 2013. Organization of the mitotic chromosome. Science 342, 948–953. https://doi.org/10.1126/science.1236083

Nielsen, C.F., Zhang, T., Barisic, M., Kalitsis, P., Hudson, D.F., 2020. Topoisomerase IIα is essential for maintenance of mitotic chromosome structure. Proc. Natl. Acad. Sci. U. S. A. 117, 12131–12142. https://doi.org/10.1073/pnas.2001760117

Nishide, K., Hirano, T., 2014. Overlapping and Non-overlapping Functions of Condensins I and II in Neural Stem Cell Divisions. PLOS Genet. 10, e1004847. https://doi.org/10.1371/journal.pgen.1004847

Nomidis, S.K., Carlon, E., Gruber, S., Marko, J.F., 2022. DNA tension-modulated translocation and loop extrusion by SMC complexes revealed by molecular dynamics simulations. Nucleic Acids Res. 50, 4974–4987. https://doi.org/10.1093/nar/gkac268

Nora, E.P., Lajoie, B.R., Schulz, E.G., Giorgetti, L., Okamoto, I., Servant, N., Piolot, T., van Berkum, N.L., Meisig, J., Sedat, J., Gribnau, J., Barillot, E., Blüthgen, N., Dekker, J., Heard, E., 2012. Spatial partitioning of the regulatory landscape of the X-inactivation center. Nature 485, 381–385. https://doi.org/10.1038/nature11049

Olins, D.E., Olins, A.L., 2003. Chromatin history: our view from the bridge. Nat. Rev. Mol. Cell Biol. 4, 809–814. https://doi.org/10.1038/nrm1225

Ono, T., Losada, A., Hirano, M., Myers, M.P., Neuwald, A.F., Hirano, T., 2003. Differential contributions of condensin I and condensin II to mitotic chromosome architecture in vertebrate cells. Cell 115, 109–121.

Osborne, C.S., Chakalova, L., Brown, K.E., Carter, D., Horton, A., Debrand, E., Goyenechea, B., Mitchell, J.A., Lopes, S., Reik, W., Fraser, P., 2004. Active genes dynamically colocalize to shared sites of ongoing transcription. Nat. Genet. 36, 1065–1071. https://doi.org/10.1038/ng1423

Palozola, K.C., Lerner, J., Zaret, K.S., 2019. A changing paradigm of transcriptional memory propagation through mitosis. Nat. Rev. Mol. Cell Biol. 20, 55–64. https://doi.org/10.1038/s41580-018-0077-z

Paul, M.R., Hochwagen, A., Ercan, S., 2019. Condensin action and compaction. Curr. Genet. 65, 407–415. https://doi.org/10.1007/s00294-018-0899-4

Petela, N.J., Gonzalez Llamazares, A., Dixon, S., Hu, B., Lee, B.-G., Metson, J., Seo, H., Ferrer-Harding, A., Voulgaris, M., Gligoris, T., Collier, J., Oh, B.-H., Löwe, J., Nasmyth, K.A., 2021. Folding of cohesin's coiled coil is important for Scc2/4-induced association with chromosomes. eLife 10, e67268. https://doi.org/10.7554/eLife.67268

Pettersen, E.F., Goddard, T.D., Huang, C.C., Meng, E.C., Couch, G.S., Croll, T.I., Morris, J.H., Ferrin, T.E., 2021. UCSF ChimeraX: Structure visualization for researchers, educators, and developers. Protein Sci. Publ. Protein Soc. 30, 70–82. https://doi.org/10.1002/pro.3943

Piazza, I., Haering, C.H., Rutkowska, A., 2013. Condensin: crafting the chromosome landscape. Chromosoma 122, 175–190. https://doi.org/10.1007/s00412-013-0405-1

Piazza, I., Rutkowska, A., Ori, A., Walczak, M., Metz, J., Pelechano, V., Beck, M., Haering, C.H., 2014. Association of condensin with chromosomes depends on DNA binding by its HEAT-repeat subunits. Nat. Struct. Mol. Biol. 21, 560–568. https://doi.org/10.1038/nsmb.2831

Pradhan, B., Barth, R., Kim, E., Davidson, I.F., Bauer, B., Laar, T. van, Yang, W., Ryu, J.-K., Torre, J. van der, Peters, J.-M., Dekker, C., 2021. SMC complexes can traverse physical roadblocks bigger than their ring size. https://doi.org/10.1101/2021.07.15.452501

Pradhan, B., Kanno, T., Igarashi, M.U., Baaske, M.D., Wong, J.S.K., Jeppsson, K., Björkegren, C., Kim, E., 2022. The Smc5/6 complex is a DNA loop extruding motor. https://doi.org/10.1101/2022.05.13.491800

Promonet, A., Padioleau, I., Liu, Y., Sanz, L., Biernacka, A., Schmitz, A.-L., Skrzypczak, M., Sarrazin, A., Mettling, C., Rowicka, M., Ginalski, K., Chedin, F., Chen, C.-L., Lin, Y.-L., Pasero, P., 2020. Topoisomerase 1 prevents replication stress at R-loop-enriched transcription termination sites. Nat. Commun. 11, 3940. https://doi.org/10.1038/s41467-020-17858-2

Punjani, A., Rubinstein, J.L., Fleet, D.J., Brubaker, M.A., 2017. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nat. Methods 14, 290–296. https://doi.org/10.1038/nmeth.4169

Rajendran, S., Jezewska, M.J., Bujalowski, W., 2000. Multiple-step kinetic mechanism of DNA-independent ATP binding and hydrolysis by Escherichia coli replicative helicase DnaB protein: quantitative analysis using the rapid quench-flow method. J. Mol. Biol. 303, 773–795. https://doi.org/10.1006/jmbi.2000.4124

Rao, S.S.P., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S., Aiden, E.L., 2014. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell 159, 1665–1680. https://doi.org/10.1016/j.cell.2014.11.021

Rattner, J.B., Hamkalo, B.A., 1978. Higher order structure in metaphase chromosomes. I. The 250 A fiber. Chromosoma 69, 363–372. https://doi.org/10.1007/BF00332139

Riggs, A.D., Holliday, R., Monk, M., Pugh, J.E., 1990. DNA methylation and late replication probably aid cell memory, and type I DNA reeling could aid chromosome folding and enhancer function. Philos. Trans. R. Soc. Lond. B Biol. Sci. 326, 285–297. https://doi.org/10.1098/rstb.1990.0012

Riley, T.R., Slattery, M., Abe, N., Rastogi, C., Liu, D., Mann, R.S., Bussemaker, H.J., 2014. SELEX-seq: a method for characterizing the complete repertoire of binding site preferences for transcription factor complexes. Methods Mol. Biol. Clifton NJ 1196, 255–278. https://doi.org/10.1007/978-1-4939-1242-1_16

Rimel, J.K., Taatjes, D.J., 2018. The essential and multifunctional TFIIH complex. Protein Sci. 27, 1018–1037. https://doi.org/10.1002/pro.3424

Robinson, P.J.J., Fairall, L., Huynh, V.A.T., Rhodes, D., 2006. EM measurements define the dimensions of the "30-nm" chromatin fiber: Evidence for a compact, interdigitated structure. Proc. Natl. Acad. Sci. 103, 6506–6511. https://doi.org/10.1073/pnas.0601212103

Robinson, P.J.J., Rhodes, D., 2006. Structure of the "30 nm" chromatin fibre: a key role for the linker histone. Curr. Opin. Struct. Biol. 16, 336–343. https://doi.org/10.1016/j.sbi.2006.05.007

Rohou, A., Grigorieff, N., 2015. CTFFIND4: Fast and accurate defocus estimation from electron micrographs. J. Struct. Biol. 192, 216–221. https://doi.org/10.1016/j.jsb.2015.08.008

Rojowska, A., Lammens, K., Seifert, F.U., Direnberger, C., Feldmann, H., Hopfner, K.-P., 2014. Structure of the Rad50 DNA double-strand break repair protein in complex with DNA. EMBO J. 33, 2847–2859. https://doi.org/10.15252/embj.201488889

Rolef Ben-Shahar, T., Heeger, S., Lehane, C., East, P., Flynn, H., Skehel, M., Uhlmann, F., 2008. Eco1-dependent cohesin acetylation during establishment of sister chromatid cohesion. Science 321, 563–566. https://doi.org/10.1126/science.1157774

Rosenthal, P.B., Henderson, R., 2003. Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. J. Mol. Biol. 333, 721–745. https://doi.org/10.1016/j.jmb.2003.07.013

Routh, A., Sandin, S., Rhodes, D., 2008. Nucleosome repeat length and linker histone stoichiometry determine chromatin fiber structure. Proc. Natl. Acad. Sci. 105, 8872–8877. https://doi.org/10.1073/pnas.0802336105

Ryu, J.-K., Katan, A.J., Sluis, E.O. van der, Wisse, T., Groot, R. de, Haering, C., Dekker, C., 2019. AFM images of open and collapsed states of yeast condensin suggest a scrunching model for DNA loop extrusion. https://doi.org/10.1101/2019.12.13.867358

Samoshkin, A., Arnaoutov, A., Jansen, L.E.T., Ouspenski, I., Dye, L., Karpova, T., McNally, J., Dasso, M., Cleveland, D.W., Strunnikov, A., 2009. Human Condensin Function Is Essential for Centromeric Chromatin Assembly and Proper Sister Kinetochore Orientation. PLoS ONE 4, e6831. https://doi.org/10.1371/journal.pone.0006831

Sanchez-Garcia, R., Gomez-Blanco, J., Cuervo, A., Carazo, J.M., Sorzano, C.O.S., Vargas, J., 2021. DeepEMhancer: a deep learning solution for cryo-EM volume post-processing. Commun. Biol. 4, 874. https://doi.org/10.1038/s42003-021-02399-1

Sari, D., Gupta, K., Thimiri Govinda Raj, D.B., Aubert, A., Drncová, P., Garzoni, F., Fitzgerald, D., Berger, I., 2016. The MultiBac Baculovirus/Insect Cell Expression Vector System for Producing Complex Protein Biologics. Adv. Exp. Med. Biol. 896, 199–215. https://doi.org/10.1007/978-3-319-27216-0_13

Scheres, S.H.W., 2012. RELION: Implementation of a Bayesian approach to cryo-EM structure determination. J. Struct. Biol. 180, 519–530. https://doi.org/10.1016/j.jsb.2012.09.006

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.-Y., White, D.J., Hartenstein, V., Eliceiri, K., Tomancak, P., Cardona, A., 2012. Fiji: an open-source platform for biological-image analysis. Nat. Methods 9, 676–682. https://doi.org/10.1038/nmeth.2019

Seidel, R., Bloom, J.G., Dekker, C., Szczelkun, M.D., 2008. Motor step size and ATP coupling efficiency of the dsDNA translocase EcoR124I. EMBO J. 27, 1388–1398. https://doi.org/10.1038/emboj.2008.69

Sessions, S.K., 2013. Genome Size, in: Maloy, S., Hughes, K. (Eds.), Brenner's Encyclopedia of Genetics (Second Edition). Academic Press, San Diego, pp. 301–305. https://doi.org/10.1016/B978-0-12-374984-0.00639-2

Sexton, T., Yaffe, E., Kenigsberg, E., Bantignies, F., Leblanc, B., Hoichman, M., Parrinello, H., Tanay, A., Cavalli, G., 2012. Three-dimensional folding and functional organization principles of the Drosophila genome. Cell 148, 458–472. https://doi.org/10.1016/j.cell.2012.01.010

Shaltiel, I.A., Datta, S., Lecomte, L., Hassler, M., Kschonsak, M., Bravo, S., Stober, C., Eustermann, S., Haering, C.H., 2021. A hold-and-feed mechanism drives directional DNA loop extrusion by condensin. https://doi.org/10.1101/2021.10.29.466147

Shi, Z., Gao, H., Bai, X.-C., Yu, H., 2020. Cryo-EM structure of the human cohesin-NIPBL-DNA complex. Science 368, 1454–1459. https://doi.org/10.1126/science.abb0981

Shintomi, K., Takahashi, T.S., Hirano, T., 2015. Reconstitution of mitotic chromatids with a minimum set of purified factors. Nat. Cell Biol. 17, 1014–1023. https://doi.org/10.1038/ncb3187

Soh, Y.-M., Bürmann, F., Shin, H.-C., Oda, T., Jin, K.S., Toseland, C.P., Kim, C., Lee, H., Kim, S.J., Kong, M.-S., Durand-Diebold, M.-L., Kim, Y.-G., Kim, H.M., Lee, N.K., Sato, M., Oh, B.-H., Gruber, S., 2015. Molecular basis for SMC rod formation and its dissolution upon DNA binding. Mol. Cell 57, 290–303. https://doi.org/10.1016/j.molcel.2014.11.023

Stillman, D. j., Geiduschek, E. p., 1984. Differential binding of a S. cerevisiae RNA polymerase III transcription factor to two promoter segments of a tRNA gene. EMBO J. 3, 847–853. https://doi.org/10.1002/j.1460-2075.1984.tb01895.x

Sundin, O., Varshavsky, A., 1981. Arrest of segregation leads to accumulation of highly intertwined catenated dimers: Dissection of the final stages of SV40 DNA replication. Cell 25, 659–669. https://doi.org/10.1016/0092-8674(81)90173-2

Sutani, T., Yuasa, T., Tomonaga, T., Dohmae, N., Takio, K., Yanagida, M., 1999. Fission yeast condensin complex: essential roles of non-SMC subunits for condensation and Cdc2 phosphorylation of Cut3/SMC4. Genes Dev. 13, 2271–2283.

Sutherland, H., Bickmore, W.A., 2009. Transcription factories: gene expression in unions? Nat. Rev. Genet. 10, 457–466. https://doi.org/10.1038/nrg2592

Tada, K., Susumu, H., Sakuno, T., Watanabe, Y., 2011. Condensin association with histone H2A shapes mitotic chromosomes. Nature 474, 477–483. https://doi.org/10.1038/nature10179

Takaba, K., Maki-Yonekura, S., Yonekura, K., 2020. Collecting large datasets of rotational electron diffraction with ParallEM and SerialEM. J. Struct. Biol. 211, 107549. https://doi.org/10.1016/j.jsb.2020.107549

Takahashi, M., Wakai, T., Hirota, T., 2016. Condensin I-mediated mitotic chromosome assembly requires association with chromokinesin KIF4A. Genes Dev. 30, 1931–1936. https://doi.org/10.1101/gad.282855.116

Taschner, M., Basquin, J., Steigenberger, B., Schäfer, I.B., Soh, Y.-M., Basquin, C., Lorentzen, E., Räschle, M., Scheltema, R.A., Gruber, S., 2021. Nse5/6 inhibits the Smc5/6 ATPase and modulates DNA substrate binding. EMBO J. 40, e107807. https://doi.org/10.15252/embj.2021107807

Tegunov, D., Cramer, P., 2019. Real-time cryo-electron microscopy data preprocessing with Warp. Nat. Methods 16, 1146–1152. https://doi.org/10.1038/s41592-019-0580-y

Teichmann, S.A., Veitia, R.A., 2004. Genes encoding subunits of stable complexes are clustered on the yeast chromosomes: an interpretation from a dosage balance perspective. Genetics 167, 2121–2125. https://doi.org/10.1534/genetics.103.024505

Terakawa, T., Bisht, S., Eeftens, J.M., Dekker, C., Haering, C.H., Greene, E.C., 2017. The condensin complex is a mechanochemical motor that translocates along DNA. Science 358, 672–676. https://doi.org/10.1126/science.aan6516

The Smc5/6 complex is a DNA loop extruding motor [WWW Document], 2022. https://doi.org/10.21203/rs.3.rs-1673223/v1

Tomonaga, T., Nagao, K., Kawasaki, Y., Furuya, K., Murakami, A., Morishita, J., Yuasa, T., Sutani, T., Kearsey, S.E., Uhlmann, F., Nasmyth, K., Yanagida, M., 2000. Characterization of fission yeast cohesin: essential anaphase proteolysis of Rad21 phosphorylated in the S phase. Genes Dev. 14, 2757–2770. https://doi.org/10.1101/gad.832000

Toselli-Mollereau, E., Robellet, X., Fauque, L., Lemaire, S., Schiklenk, C., Klein, C., Hocquet, C., Legros, P., N'Guyen, L., Mouillard, L., Chautard, E., Auboeuf, D., Haering, C.H., Bernard, P., 2016. Nucleosome eviction in mitosis assists condensin loading and chromosome condensation. EMBO J. 35, 1565–1581. https://doi.org/10.15252/embj.201592849

Tuduri, S., Crabbé, L., Conti, C., Tourrière, H., Holtgreve-Grez, H., Jauch, A., Pantesco, V., De Vos, J., Thomas, A., Theillet, C., Pommier, Y., Tazi, J., Coquelle, A., Pasero, P., 2009. Topoisomerase I suppresses genomic instability by preventing interference between replication and transcription. Nat. Cell Biol. 11, 1315–1324. https://doi.org/10.1038/ncb1984

Uemura, T., Ohkura, H., Adachi, Y., Morino, K., Shiozaki, K., Yanagida, M., 1987. DNA topoisomerase II is required for condensation and separation of mitotic chromosomes in S. pombe. Cell 50, 917–925. https://doi.org/10.1016/0092-8674(87)90518-6

Uhlmann, F., Lottspeich, F., Nasmyth, K., 1999. Sister-chromatid separation at anaphase onset is promoted by cleavage of the cohesin subunit Scc1. Nature 400, 37–42. https://doi.org/10.1038/21831

Valens, M., Penaud, S., Rossignol, M., Cornet, F., Boccard, F., 2004. Macrodomain organization of the Escherichia coli chromosome. EMBO J. 23, 4330–4341. https://doi.org/10.1038/sj.emboj.7600434

van Koningsbruggen, S., Gierlinski, M., Schofield, P., Martin, D., Barton, G.J., Ariyurek, Y., den Dunnen, J.T., Lamond, A.I., 2010. High-resolution whole-genome sequencing reveals that specific chromatin domains from most human chromosomes associate with nucleoli. Mol. Biol. Cell 21, 3735–3748. https://doi.org/10.1091/mbc.E10-06-0508

van Ruiten, M.S., Rowland, B.D., 2021. On the choreography of genome folding: A grand pas de deux of cohesin and CTCF. Curr. Opin. Cell Biol. 70, 84–90. https://doi.org/10.1016/j.ceb.2020.12.001

Vazquez, J., Belmont, A.S., Sedat, J.W., 2001. Multiple regimes of constrained chromosome motion are regulated in the interphase Drosophila nucleus. Curr. Biol. CB 11, 1227–1239. https://doi.org/10.1016/s0960-9822(01)00390-6

Walther, N., Hossain, M.J., Politi, A.Z., Koch, B., Kueblbeck, M., Ødegård-Fougner, Ø., Lampe, M., Ellenberg, J., 2018. A quantitative map of human Condensins provides new insights into mitotic chromosome architecture. J. Cell Biol. 217, 2309–2328. https://doi.org/10.1083/jcb.201801048

Wang, H.-W., 2022. A commentary of "Cryo-EM achieves atomic resolution" in 10 remarkable discoveries from 2020 in Nature. Fundam. Res. 2, 349–350. https://doi.org/10.1016/j.fmre.2022.01.014

Wen, P.-C., Tajkhorshid, E., 2011. Conformational Coupling of the Nucleotide-Binding and the Transmembrane Domains in ABC Transporters. Biophys. J. 101, 680–690. https://doi.org/10.1016/j.bpj.2011.06.031

Woodward, J., Taylor, G.C., Soares, D.C., Boyle, S., Sie, D., Read, D., Chathoth, K., Vukovic, M., Tarrats, N., Jamieson, D., Campbell, K.J., Blyth, K., Acosta, J.C., Ylstra, B., Arends, M.J., Kranc, K.R., Jackson, A.P., Bickmore, W.A., Wood, A.J., 2016. Condensin II mutation causes T-cell lymphoma through tissue-specific genome instability. Genes Dev. 30, 2173–2186. https://doi.org/10.1101/gad.284562.116

Yamashita, D., Shintomi, K., Ono, T., Gavvovidis, I., Schindler, D., Neitzel, H., Trimborn, M., Hirano, T., 2011. MCPH1 regulates chromosome condensation and shaping as a composite modulator of condensin II. J. Cell Biol. 194, 841–854. https://doi.org/10.1083/jcb.201106141

Yatskevich, S., Rhodes, J., Nasmyth, K., 2019. Organization of Chromosomal DNA by SMC Complexes. Annu. Rev. Genet. 53, 445–482. https://doi.org/10.1146/annurev-genet-112618-043633

Young, G., Hundt, N., Cole, D., Fineberg, A., Andrecka, J., Tyler, A., Olerinyova, A., Ansari, A., Marklund, E.G., Collier, M.P., Chandler, S.A., Tkachenko, O., Allen, J., Crispin, M., Billington, N., Takagi, Y., Sellers, J.R., Eichmann, C., Selenko, P., Frey, L., Riek, R., Galpin, M.R., Struwe, W.B., Benesch, J.L.P., Kukura, P., 2018. Quantitative mass imaging of single biological macromolecules. Science 360, 423–427. https://doi.org/10.1126/science.aar5839

Yu, Y., Li, S., Ser, Z., Kuang, H., Than, T., Guan, D., Zhao, X., Patel, D.J., 2022. Cryo-EM structure of DNA-bound Smc5/6 reveals DNA clamping enabled by multi-subunit conformational changes. Proc. Natl. Acad. Sci. 119, e2202799119. https://doi.org/10.1073/pnas.2202799119

Yu, Y., Li, S., Ser, Z., Sanyal, T., Choi, K., Wan, B., Kuang, H., Sali, A., Kentsis, A., Patel, D.J., Zhao, X., 2021. Integrative analysis reveals unique structural and functional features of the Smc5/6 complex. Proc. Natl. Acad. Sci. U. S. A. 118, e2026844118. https://doi.org/10.1073/pnas.2026844118

Yuen, K.C., Slaughter, B.D., Gerton, J.L., 2017. Condensin II is anchored by TFIIIC and H3K4me3 in the mammalian genome and supports the expression of active dense gene clusters. Sci. Adv. 3, e1700191. https://doi.org/10.1126/sciadv.1700191

Zabrady, K., Adamus, M., Vondrova, L., Liao, C., Skoupilova, H., Novakova, M., Jurcisinova, L., Alt, A., Oliver, A.W., Lehmann, A.R., Palecek, J.J., 2016. Chromatin association of the SMC5/6 complex is dependent on binding of its NSE3 subunit to DNA. Nucleic Acids Res. 44, 1064–1079. https://doi.org/10.1093/nar/gkv1021

Zhang, Z., Shigematsu, H., Shimizu, T., Ohto, U., 2021. Improving particle quality in cryo-EM analysis using a PEGylation method. Structure 29, 1192-1199.e4. https://doi.org/10.1016/j.str.2021.05.004

Zheng, S.Q., Palovcak, E., Armache, J.-P., Verba, K.A., Cheng, Y., Agard, D.A., 2017. MotionCor2 - anisotropic correction of beam-induced motion for improved cryo-electron microscopy. Nat. Methods 14, 331–332. https://doi.org/10.1038/nmeth.4193

Zhong, E.D., Bepler, T., Berger, B., Davis, J.H., 2021. CryoDRGN: reconstruction of heterogeneous cryo-EM structures using neural networks. Nat. Methods 18, 176–185. https://doi.org/10.1038/s41592-020-01049-4

Zuin, J., Dixon, J.R., van der Reijden, M.I.J.A., Ye, Z., Kolovos, P., Brouwer, R.W.W., van de Corput, M.P.C., van de Werken, H.J.G., Knoch, T.A., van IJcken, W.F.J., Grosveld, F.G., Ren, B., Wendt, K.S., 2014. Cohesin and CTCF differentially affect chromatin architecture and gene expression in human cells. Proc. Natl. Acad. Sci. U. S. A. 111, 996–1001. https://doi.org/10.1073/pnas.1317788111