universität wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„Binding of cohesin to nucleosome-bound DNA“

verfasst von / submitted by

Tabea Gruber, BSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien, 2022/ Vienna, 2022

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	UA 066 834
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Masterstudium Molekulare Biologie
Betreut von / Supervisor:	Dr. Jan-Michael Peters

Abstract

Cohesin, a ring-shaped DNA-binding ATPase, is a member of the SMC protein family that was originally discovered for its importance in sister chromatid cohesion. In recent years cohesin's function in genome folding, where it forms DNA loops and TADs, has been discovered. To fold the genome in interphase cells cohesin connects distal DNA sequences in *cis* through a process called loop extrusion, which was observed *in vitro*. How cohesin can perform this function *in vivo* is still unknown. One of the questions that remain is how cohesin performs loop extrusion not on free but decorated DNA, including nucleosomes. Nucleosomes fold and compact DNA so that large genomes fit into the cell nucleus. Nucleosome-bound DNA is not accessible to most proteins and therefore might be an obstacle for cohesin's function on DNA.

In this study binding of cohesin and NIPBL to nucleosome-bound DNA or free linker-DNA is observed using EMSAs, mass photometry and HS-AFM. To do this, recombinant proteins are expressed in Sf9 cells and purified using tandem affinity purification. Furthermore, two types of nucleosomes are reconstituted using Sf9 histones and 147 bp of DNA, containing no protruding DNA, or 202 bp of DNA, resulting in free linker-DNA protruding from the nucleosome.

The results of this study indicate that cohesin needs free accessible DNA to perform its function on chromatin. Furthermore, NIPBL seems to be able to interact with nucleosome-bound DNA but whether it can load cohesin onto nucleosome-bound DNA and if it can perform this function *in vivo* remains unanswered.

Zusammenfassung

Cohesin ist eine ringförmige DNA-bindende ATPase und gehört zur SMC Protein Familie. Ursprünglich wurde Cohesin für seine Funktion im Zusammenhalt der Schwesterchromatiden entdeckt, doch in den letzten Jahren wurde Cohesin's Fähigkeit in Genomfaltung erforscht. Um das Genom in „Loops" und „TADs" zu falten muss Cohesin entfernte DNA-Sequenzen in *cis* über einen Vorgang verbinden, der sich „Loop Extrusion" nennt. „Loop Extrusion" wurde *in vitro* beobachtet, aber wie dieser Vorgang *in vivo* funktioniert ist weiterhin unklar, da in lebenden Zellen DNA nicht nackt vorliegt, sondern von Proteinen, hauptsächlich Nukleosomen, bedeckt ist. Nukleosomen kompaktieren DNA damit das Genom in den Zellkern passt. Nukleosom-gebundene DNA ist für die meisten Proteine nicht zugänglich und kann daher für Cohesin's Funktion an DNA ein Hindernis sein.

In dieser Arbeit wird die Bindung von Cohesin und seinem HAWK-Protein NIPBL an Nukleosom-gebundene DNA und freie „linker-DNA" beobachtet. Dafür werden EMSAs, Massenphotometrie und HS-AFM Experimente verwendet. Rekombinante Proteine werden in Sf9 Zellen exprimiert und über „Tandem Affinity Purification" aufgereinigt. Zwei verschiedene Typen von Nukleosomen werden mit Hilfe von Sf9 Histonen rekonstituiert: Nukleosomen die 147 bp DNA enthalten und keine freie DNA haben, und Nukleosomen die 202 bp DNA und daher freie hervorstehende DNA enthalten.

Die Ergebnisse dieser Studie deuten darauf hin, dass Cohesin freie DNA braucht, um seine Funktion in lebenden Zellen auszuüben. Zusätzlich scheint es, als könnte NIPBL an Nukleosom-gebundene DNA binden, aber ob es somit auch Cohesin an Nukleosom-gebundene DNA laden kann, und ob es diese Funktion auch *in vivo* ausführen kann, bleibt unklar und erfordert weitere Forschung.

Acknowledgments

First of all, I would like to thank my supervisor Jan-Michael Peters for giving me the opportunity to work in his lab and gaining so much experience through this project.

Most importantly, I would like to thank Sabrina, Gabi and Iain. Thank you, Sabrina, for not only helping and supervising me in this project and answering all my questions, but also for being a friend. Thank you, Gabi, for always offering a helping hand and for saving experiments in the last minute. Thank you, Iain, for always giving suggestions on my project and for being a great bench-neighbour.

I also want to say a big “thank you” to the whole Peters Lab. Each and every single one of you made me feel welcome from my very first day, thank you for the laughs and for the fun I had even on not so funny days.

Lastly, I would like to thank all of my friends and my family. Thank you for always cheering me up, for motivating me when I felt down and for believing in me. A very special “thank you” goes to my parents, who never doubt me and always support me in everything I do

Contents

1. Introduction

1.1. The cell cycle

In eukaryotes the cell cycle is a highly complex process that can be divided into interphase and mitotic (M-)Phase. During mitosis, which is the shortest phase of the cell cycle, the cell divides and its replicated chromosomes are segregated equally into two daughter cells. Here, chromosomes are condensed and appear as rod-shaped sister chromatids. Interphase can be subdivided into G1, S and G2-Phase. While the cell grows and produces proteins as well as cytoplasmic organelles in all three stages, DNA replication only occurs in S-Phase [1][2]. In interphase the genome is not present as long linear DNA molecules, but DNA is packaged and folded in a highly regulated process, which contributes to dynamic genome organization [3].

1.2. 3D genome organization

Eukaryotic genomes are large, especially compared to the nucleus they are packed in to. The human genome for example contains ~3.3 billion base pairs of DNA resulting in double stranded DNA (dsDNA) of approximately 2 meters [4], that has to fit into a cell nucleus of ~10 µm in diameter [5]. To achieve this, DNA is hierarchically folded at multiple levels (Fig. 1). On the smallest scale, dsDNA of 2 nm in diameter is wrapped around histone octamers (see Chapter 1.6.) to assemble 10 nm nucleosome chromatin fibers [6], thereby compacting DNA. On an intermediate scale chromatin is structured into loops and Topologically

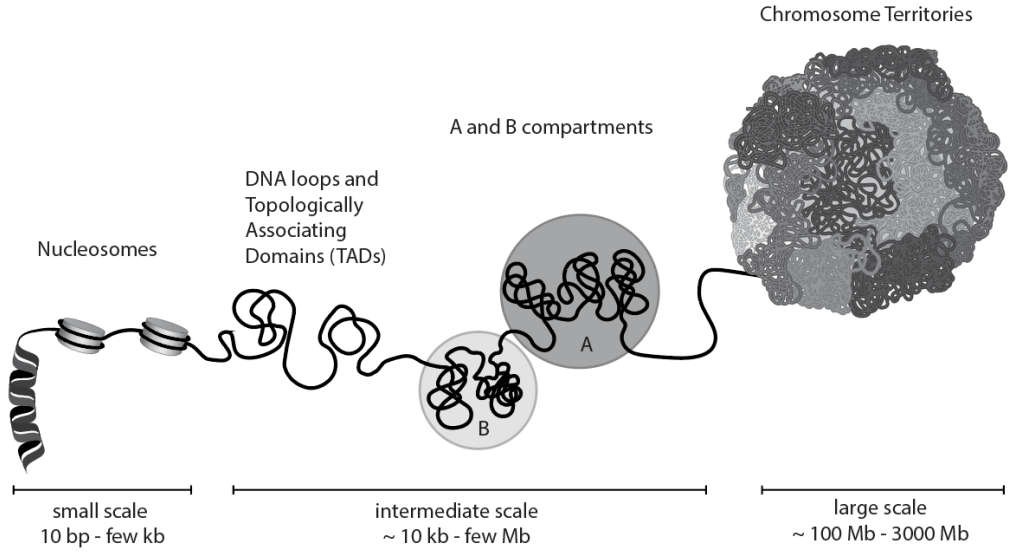

Figure 1 | **3D organization of the genome.** DNA is wrapped around nucleosomes and compacted. Furthermore, loops and TADs are formed by chromatin loops. A and B compartments form active and inactive regions at an intermediate scale of genome organization and at the highest level the interphase nucleus is organized in chromosome territories.

Associated Domains (TADs), which are formed by long-range *cis*-interactions [7]. TADs are self-interacting genomic regions that span hundreds of kilobase pairs (kb) of DNA and are limited by TAD boundaries [8][3]. The formation of chromatin loops promotes interaction of regions within TADs, while simultaneously promoting TAD insulation. The CCCTC binding-factor (CTCF) and the ring-shaped ATPase protein cohesin typically define TAD boundaries and were also found to be enriched at boundaries of loop domains [3][7].

At the next level of genome organization, over distances covering megabase pairs (Mb) of DNA, chromatin is assembled into "A and B compartments". "A compartments" contain transcriptionally active regions that interact with other transcriptionally active regions and usually reside in the interior nuclear space. These regions are marked by chromatin accessibility and active histone modifications and appear as euchromatin in electron micrographs. "B compartments" contain heterochromatin and preferentially reside at the nuclear lamina or the nucleolus. [3][7]

In interphase cells, chromosomes are not randomly distributed but are found in so-called chromosome territories, which display the highest level of genome organization. Within a chromosome territory transcriptionally active "A compartments" are usually positioned at territory boarders and can interact with active regions of other chromosome territories [3].

Overall, 3D genome organization is not only important for gene expression but also for DNA replication, repair, chromosome compaction and segregation. TADs and loops have been found to regulate gene expression by bringing regulatory elements, for example enhancers, silencers and insulators, in close proximity to promoters and active genes, which are often separated by Kb or even Mb of DNA [9]. One of the key factors that is essential for proper DNA folding is the ring-shaped SMC protein complex cohesin [10].

1.3. SMC protein family

Structural maintenance of chromosomes (SMC) proteins are DNA-binding proteins that are conserved in all kingdoms of life and play crucial roles in structural and functional organization of chromosomes [11]. SMC proteins share a unique tripartite ring structure containing two SMC proteins and a kleisin subunit [7]. Three different SMC proteins are known to exist in eukaryotes: condensin is important for condensation of chromosomes [12], cohesin establishes sister chromatid cohesion [13] and SMC5/6 plays a key role in DNA repair [14][15]. In prokaryotes the SMC complexes MukBEF and BsSMC are responsible for chromosome

organization [16][17] (Fig. 2). This study will focus on the SMC protein cohesin and its function in genome organization.

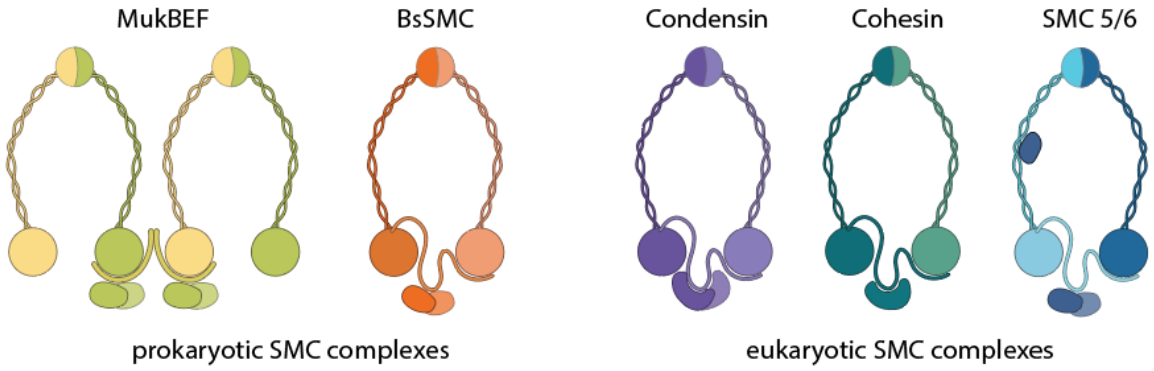

Figure 2 | **Prokaryotic and Eukaryotic SMC complexes**. The prokaryotic SMC proteins MukBEF and BsSMC are pictured on the left. The eukaryotic SMC complexes condensin, cohesin and SMC5/6 are shown on the right.

1.4. Cohesin

1.4.1. Architecture

The ring-shaped ATPase cohesin was first identified for its key role in topologically entrapping sister chromatids within its ring structure and establishing cohesion until the segregation of chromatids in mitosis. The tripartite ring of human cohesin contains two SMC subunits, SMC1 and SMC3, the kleisin RAD21 (SCC1 in yeast), and additional subunits of the HAWK (HEAT (huntingtin, elongation factor 3, A subunit and TOR) repeat proteins associated with kleisins) protein family (Fig. 3). Both SMC subunits consist of a 50 nm-long antiparallel coiled coil domain which stably dimerize at one end to form the so-called hinge domain. On the other side of the coiled coil domain SMC proteins form ATPase head domains that can bind and hydrolyze ATP. RAD21 connects the ATPase heads and thereby closes the ring structure. Additionally, cohesin can interact with different HAWK proteins, namely STAG1/2, NIPBL and PDS5A/B, depending on the type of cohesin complex that is being formed. [18][7]

1.4.2. Function

1.4.2.1. Sister chromatid cohesion

Cohesive Cohesin is crucial for stable sister chromatid cohesion during DNA replication and maintenance of cohesion until the segregation of sister chromatids in mitosis or meiosis [19]. Cohesin establishes the *trans*-connection between sister chromatids by topologically entrapping DNA molecules of each chromatid in its ring structure [20]. During sister chromatid cohesion, the kleisin subunit RAD21 interacts with one of the PDS5A or B isoforms (Fig. 3a). To maintain sister chromatid cohesion two lysine residues (K105 and K106) in the SMC3 protein must be acetylated. Additionally, in higher eukaryotes the cohesin subunit PDS5A/B interacts with the protein sororin, which protects cohesin from chromatin release by WAPL (wings apart-like protein) [7].

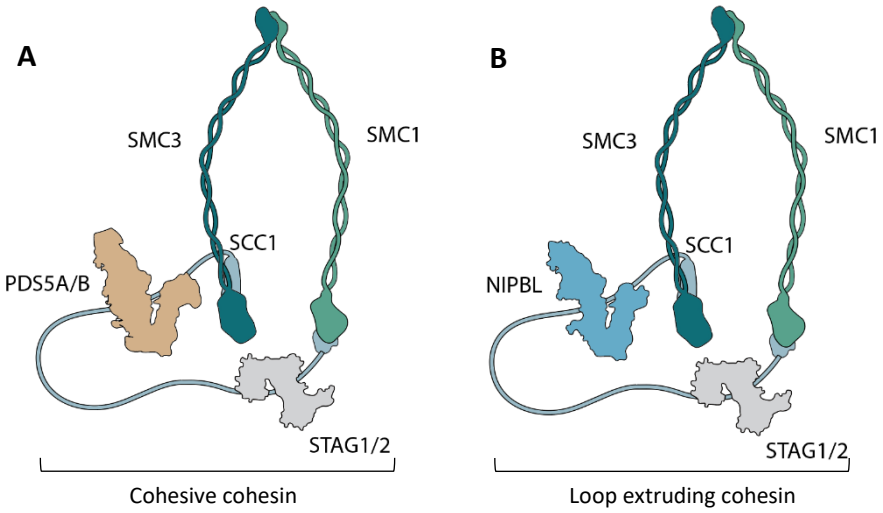

Figure 3 | **Architecture of SMC complexes. A** | Cohesive cohesin. The coiled coils of SMC3 and SMC1 are connected over the ATPase heads by the kleisin protein SCC1/RAD21. Additionally, SCC1/RAD21 associates with the HAWK proteins STAG1/2 and PDS5A/B. This complex cannot extrude loops but might be involved in sister chromatid cohesion. **B** | Loop extruding cohesin. The kleisin SCC1/RAD21 interacts with the HAWK protein NIPBL instead of PDS5A/B as well as STAG1/2. This complex represents the loop extruding form of cohesin.

1.4.2.2. Loop Extrusion in Interphase

In recent years the SMC proteins condensin and cohesin were found to fold chromatin into loops by a process called loop extrusion [7]. This process allows them to regulate DNA compaction, DNA repair and genome stability. The loop extrusion hypothesis has first been proposed in 1990 and states that a loop extruding factor binds to DNA and starts actively extruding a small loop and, by continuously reeling in DNA, enhances loop size over time (Fig. 4) [21]. The conserved SMC complexes cohesin and condensin were discovered to be

good candidates for such a loop extrusion complex since they not only bind to DNA to fulfill their crucial roles as stated above, but also contain ATPase activity which would be necessary for the motor function of an active loop extrusion complex [22]. Budding yeast condensin was shown to translocate along DNA depending on its ATPase activity [23] and recently condensin's ability to extrude DNA loops was observed for the first time *in vitro* [24].

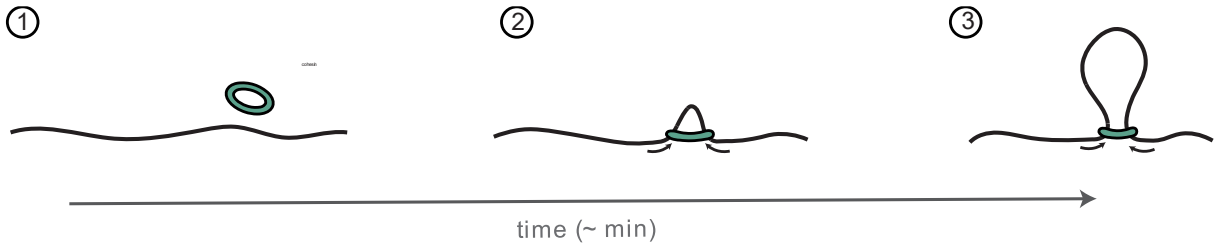

Figure 4 | **The loop extrusion hypothesis. 1**| A loop extrusion complex, pictured here in green, associates with DNA. **2**| DNA is reeled into a loop simultaneously from both sides, thereby creating a small loop. **3**| Over time the loop grows bigger as more DNA is being reeled in.

Additionally, cohesin has been proposed to act as a loop extruding motor in interphase cells. Since it was found to colocalize with the zinc finger protein CTCF at boundaries of loop domains, it was proposed that cohesin would form a small loop and continue to reel DNA into the loop until it encounters convergently oriented CCCTC-binding factor (CTCF) sites bound by CTCF proteins, stopping loop extrusion [7]. Recent single molecule studies have shown that cohesin can indeed actively extrude DNA loops *in vitro* [10][25].

The cohesin complex that connects distant DNA sequences in *cis* by performing loop extrusion interacts with the HAWK protein NIPBL (nipped-b-like protein) (SCC2 in yeast) instead of PDS5A/B (Fig. 3b) [7]. Experimental data showed that cohesin can only extrude loops in the presence of NIPBL *in vitro* [7][25]. When cells are depleted of NIPBL-MAU2 cohesin could barely be detected on chromatin, suggesting that NIPBL-MAU2 may act as cohesin DNA-loader [26]. Additionally, NIPBL is viewed as regulatory subunit since it can activate cohesin's ATPase activity, thereby transforming cohesin into an active holoenzyme [10][25].

Overall, loop extrusion is a conserved process that performs important functions in gene expression, regulation and recombination. How exactly cohesin folds DNA into loops and whether this is done in a topological or pseudo-/non-topological matter is not yet known, although different models including large conformational changes of the SMC complexes were proposed and some experimental data suggested that loop extrusion is performed in a non-topological or pseudo-topological manner. [10][27][7]

1.4.3. Models of Loop Extrusion

Several models of loop extrusion have been proposed and will be discussed briefly in this chapter. All of these models have limitations and further experiments need to be performed to fully understand the mechanisms of loop extrusion [7].

The “walking model” proposes that SMC complexes walk or inchworm along DNA similar to the cytoskeleton motor protein kinesin, which moves along microtubules [28]. Both SMC proteins and kinesin contain a dimer of coiled coils and ATPase domains that could drive conformational changes in an ATP-dependent manner, leading to unidirectional walking [29][7]. Loop extrusion could occur if SMC complexes bind and walk along DNA while simultaneously maintaining contact with a distal DNA sequence via the hinge domain [30][7].

The “pumping model” suggests that DNA may be passed between two distinct chambers that are formed by the SMC subunits and the kleisin in an ATP-dependent manner. The SMC proteins change conformation when the heads engage in the presence of ATP, resulting in opening of the aligned coiled coils to form a ring-like structure. DNA may bind along the upper surface of engaged head domains. ATP hydrolysis leads to disengagement of the heads and the coiled coils will close to again form a rod-shape and the DNA could be pushed into a chamber that is now formed between the disengaged head domains and the kleisin subunit. [31][7]

Another model, called the “scrunching model”, could explain the observation that SMC coiled coils can bend, bringing the hinge and the heads in close proximity. Folding of coiled coils could change the conformation of the hinge and alter its affinity for DNA, resulting in DNA being passed from the hinge to the heads when bending occurs. The conformational changes could be linked to ATP binding and hydrolysis and lead to translocation of DNA. [32][33][7]

In a recent study, that proposes a swing and clamp mechanism for cohesin mediated loop extrusion, NIPBL plays a key role [34]. In brief, the model suggests that the SMC coiled coils spontaneously bend and bring the hinge, which contains a high-affinity DNA binding site, into close proximity with the SMC3 ATPase head, leading to a 50 nm movement. In this state, ATP is absent, NIPBL is bound with its “nose” to the hinge and the heads are disengaged. When ATP binds to the heads DNA is passed from the hinge to the SMC3 head and NIPBL will leave the hinge and clamp onto DNA. Next, the ATPase heads would engage and force the coiled coils to stretch. Subsequent ATP hydrolysis would lead to disengagement of the heads and the starting conformation would be restored. Characterization of the DNA binding sites showed that the SMC3 head and NIPBL DNA binding sites are part of the DNA clamp in the engaged

state and are necessary for ATPase activity, while DNA binding sites in SMC1 head, hinge and STAG1 are not necessary for ATPase activity but still crucial for loop extrusion activity, hence they may play an important role for translocation of DNA into loops. [34]

1.5. NIPBL-MAU2

NIPBL is a hooked shape HAWK protein of 316 kDa that can form heterodimers with its binding partner MAU2. NIPBL was shown to be crucial for cohesin's loop extrusion activity. Furthermore, the heterodimer is thought to load cohesin onto DNA and stimulate its ATPase activity, but how exactly cohesin can bind to DNA and how NIPBL would load cohesin onto DNA is not yet known. [7]

Understanding NIPBL's role in DNA binding and loop extrusion will be important to further understand its role in the human developmental disease Cornelia de Lange Syndrome (CdLS), in which most cases are caused by NIPBL mutations [35]. Furthermore, the "swing and clamp" model, that was described above, offers an explanation on how cohesin could pass by nucleosomes [34], but further investigation will be necessary to fully understand how loop extrusion can act on chromatin and if NIPBL can load cohesin onto nucleosome-bound DNA.

1.6. Nucleosomes

In cells DNA is not present as a linear naked molecule, but is covered by proteins, mainly nucleosomes, that compact the large DNA molecules, which are in some species even meters long, to fit into the ~10 µm nucleus [5]. Compaction of DNA by nucleosomes is the first step in the hierarchical folding process in 3D genome organization as explained above. Nucleosome cores are the fundamental repeating units of chromatin. They consist of a histone octamer around which ~147 bp of DNA wrap approximately 1.67 times forming 10 nm nucleosome chromatin fibers. Under low salt conditions these chromatin fibers appear as "beads-on-a-string" structure [36][37][38]. The histone octamer comprises two copies of each core histone H2A, H2B, H3 and H4, which form H2A-H2B dimers and H3-H4 tetramers in the absence of DNA (Fig. 5) [39][40]. Each core histone consists of a structured part and an 25-40 residue long unstructured N-terminal tail that is accessible for posttranslational modifications [41]. The assembly of the nucleosome core is supported by histone chaperones [42]. Histones H1 and H5 are linker histones that bind DNA entry and exit sites and allow the formation of higher-order structures like the 30 nm chromatin fiber [37], thereby further compacting DNA.

Individual nucleosomes are connected by 20-90 bp long linker DNA [38]. While nucleosomes do not recognize a specific DNA sequence, they show a higher affinity to the so-called 147 bp long 601 widom sequence, which wraps approximately 1.67 times around the histone octamer without any linker DNA protruding from the nucleosome-core [43].

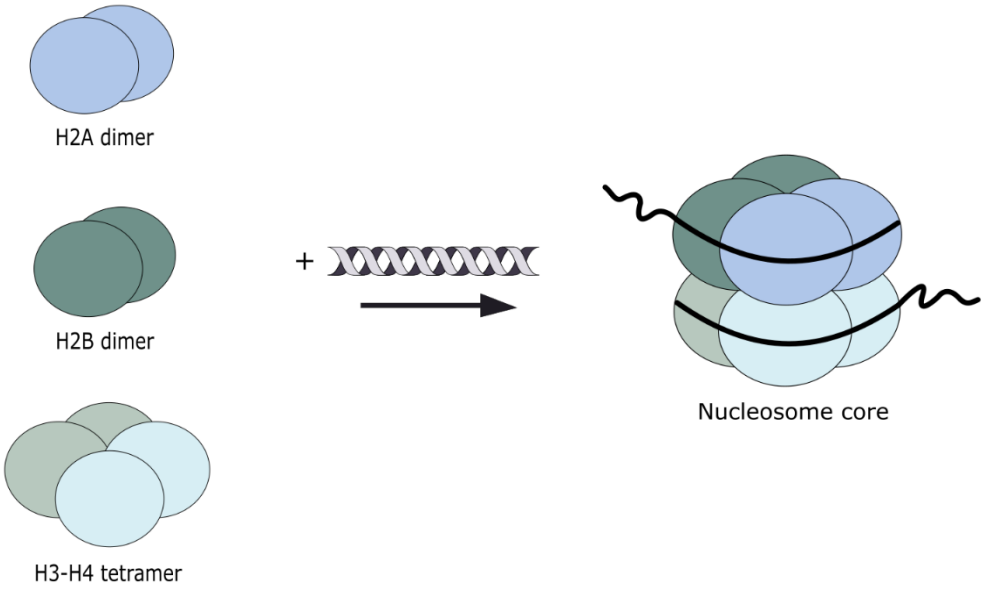

Figure 5 | **Assembly of nucleosome cores.** The core histones consist as H2A dimer, H2B dimer and H3-H4 tetramer in the absence of DNA. When DNA is added the nucleosome core is assembled with the help of histone chaperones. 147 bp of DNA wrap approximately 1.67 times around the histone octamer. Linker DNA protrudes from the nucleosome and connects individual nucleosomes.

Nucleosomes are regulated by ATP-dependent chromatin remodeler complexes that can facilitate processes like nucleosome sliding, loading and eviction from DNA [38]. Furthermore, nucleosomes can be regulated by post-translational modifications of the N-terminal tails of histones, histone variants, DNA methylation and the presence of magnesium ions [37][44]. Acetylation of lysine residues in histone tails for example is important for transcriptional activity: As the basic charge of the residues is neutralized the interaction of the histones with DNA is weakend, which makes the DNA more accessible. Histone acetylation and deacetylation is performed by Histone Acetyltransferases (HATs) and Histone Deacetylases (HDACs). Their regulatory activity is not only important for transcriptional activity but also plays a crucial role in replication, higher-order chromatin structure and interaction of other proteins with nucleosomes [41].

While SMC complexes can act on naked DNA and perform loop extrusion *in vitro*, *in vivo* DNA is covered by nucleosomes and therefore less accessible. How SMC complexes act on chromatin and if they can directly interact with nucleosomes remains unknown. It has been

suggested that the SMC complex condensin may be able to bypass nucleosomes *in vitro* and that nucleosomes are incorporated into loops without being evicted or displaced [45]. Furthermore, it has been shown that cohesin's function on chromatin may require chromatin remodeler complexes like RSC (Remodels the structure of chromatin) and FACT (Facilitates chromatin transcription) in yeast [46][47][48]. How cohesin acts on chromatin in higher eukaryotes and whether the cohesin loader NIPBL-MAU2 can load cohesin onto nucleosome-bound DNA is still unclear.

1.7. Methods for Protein-DNA and Protein-Nucleosome Binding Observation

In this study protein-DNA and protein-nucleosome interactions were observed using different biochemical and biophysical assays that will be explained in this chapter.

1.7.1. Electrophoretic Mobility Shift Assays

Electrophoretic Mobility Shift Assays (EMSAs) are well established assays that are used to study binding of proteins to nucleic acids. The principle of this assay is based on the differences in electrophoretic mobility of protein-DNA (or RNA) complexes compared to free DNA in a gel matrix [49].

Advantages of EMSAs are that they are a simple and rapid method for observation of protein-DNA interaction. Binding conditions can be easily adjusted and when using radioisotope-labeled nucleic acids the method is very sensitive. Additionally, labelling methods like fluorescence or chemiluminescence can be used. Furthermore, nucleic acids of different length and structure, as well as proteins of different sizes can be analysed. Disadvantages of EMSAs are, that there is no information about the position where the protein binds to the nucleic acid, that samples are not at a chemical equilibrium during the electrophoresis and that molecular weights cannot be measured and therefore proteins that are present identified directly [49].

1.7.2. Mass Photometry

While the interaction between biomolecules underlies most biological processes studying those interactions still has many limitations and difficulties. Recently, mass photometry (MP) has been introduced, which allows to observe molecules and their assembly into complexes at low concentrations. MP uses the principle of light scattering to measure the mass of individual proteins and complexes in solution. The molecules bind non-specifically to a glass surface and thereby displace buffer solution, causing light to be scattered where the molecule binds at the interface of glass and buffer. The interference between scattered and reflected light is measured

subsequently and, since the scattering signal is proportional to the molecular mass, is used to measure the mass of the molecule with an accuracy of 2% (Fig. 6) [50].

The main advantages of MP are that it is a sensitive, rapid and accurate method that does not require labelling or immobilization of the single molecules. Its current main limitation is that measurements are only possible at sub-micromolar concentrations, thus only binding affinities in the sub-micromolar range can be observed [50].

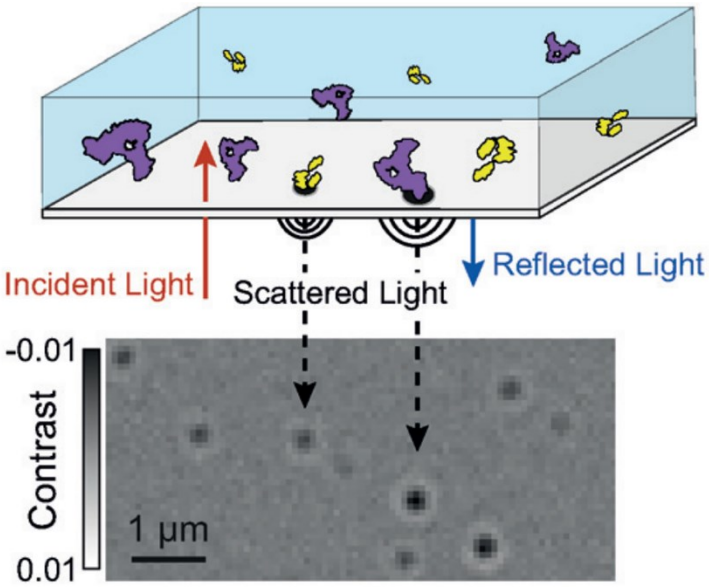

Figure 6 | **Principle of Mass Photometry**. Single proteins bind non-specifically to the glass surface and displace buffer solution, therby creating a change in refractive index. The scattered light that is reflected from the proteins is detected and the molecular mass of each individual protein is measured (from *Soltermann et al. 2020).*

1.7.3. HS-AFM

High-Speed Atomic Force Microscopy (HS-AFM) allows to directly image single molecules under physiological buffer conditions without the need for labelling, staining or fixation. It allows the simultaneous observation of both structural and dynamical behavior of proteins. The proteins are recorded at high spatial resolution, thereby giving insights into their mechanistic function. In HS-AFM a micro-cantilever containing a sharp tip at its free end scans over a sample surface. The cantilever tip is excited and oscillates close to its resonance frequency while tapping over the sample-surface and its contact with the sample results in changes of the oscillation amplitude. An optical beam deflection detector contains a laser beam that reflects back from the cantilever and illuminates a position-sensitive photodetector, thereby detecting changes in the cantilever's deflection pattern. While the cantilever scans the sample in XY-direction it will oscillate and be held constant by movement of the sample-stage in Z-direction through feedback control. The feedback signal is used to acquire a topographic height image of the sample surface. (Fig. 7) [51]

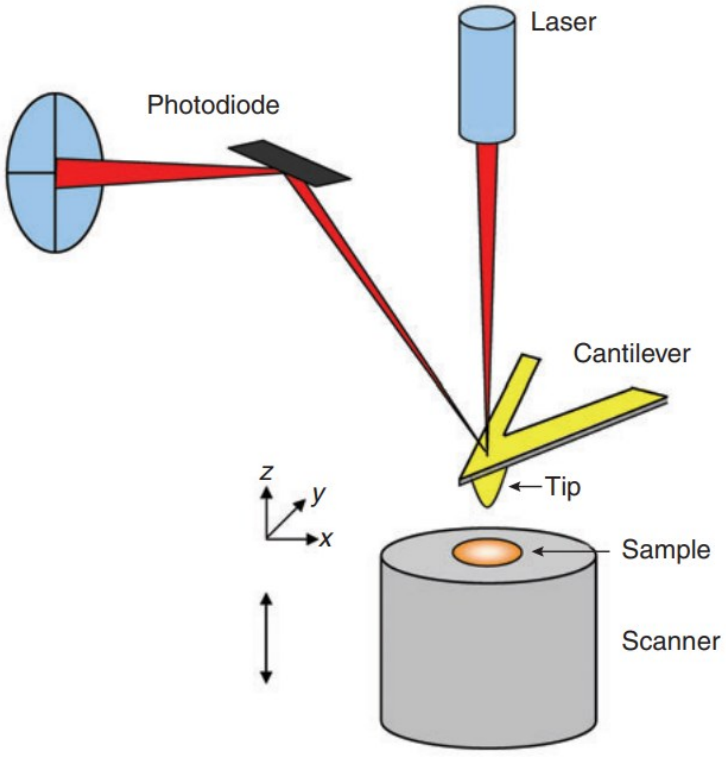

Figure 7 | **Principle of HS-AFM.** A cantilever tip scans over a sample surface in XY-direction while oscillating in Z-direction. A laser beam reflects back from the cantilever and illuminates a position-sensitive photodetector. The feedback signal is used to acquire a topographic height image of the sample (from *Hinterdorfer, P. & Dufrêne, Y. F., 2006).*

1.8. Aim of the Study

The SMC complex cohesin is known for its function in sister chromatid cohesion and since recently for its ability to perform loop extrusion *in vitro* [10][25]. How cohesin can perform loop extrusion *in vivo,* where DNA is compacted by nucleosomes and therefore not freely accessible, remains unknown.

This study aims to understand if cohesin can bind nucleosome-bound DNA or if it needs free linker DNA protruding from the nucleosome core to associate and interact with chromatin. To try and answer this question histones will first be purified from Sf9 insect cell nuclei. Additionally, recombinant histones from *xenopus laevis* will be expressed in *E.coli* cells and purified. 147 bp and 202 bp of DNA will be amplified using a plasmid containing the 601 widom nucleosome-positioning sequence. Nucleosomes will be reconstituted by a salt gradient dialysis and characterized using EMSAs, MP and HS-AFM.

Recombinant proteins will be expressed in Sf9 insect cells and purified by tandem-affinity purification. Binding assays, namely EMSAs and MP will be performed using 147 bp nucleosome-bound DNA, hence no free linker DNA should be present, as well as 202 bp nucleosome-bound DNA, from which approximately 27 bp of DNA should protrude from each side of the nucleosome core as free linker DNA (Fig. 8). Additionally, binding of the proteins to nucleosomes will be compared to binding to naked 147 bp DNA and 202 bp DNA fragments. HS-AFM will be used to characterize nucleosomes and observe interaction of cohesin with linker DNA or nucleosome-bound DNA.

Overall, this work aims to provide insights in cohesin's function on chromatin and answer the question if cohesin can directly interact with nucleosome-bound DNA, if it needs free linker DNA in-between nucleosomes to perform its functions on chromatin or if the cohesin-loader NIPBL may be able to interact with nucleosome-bound DNA and load cohesin onto chromatin.

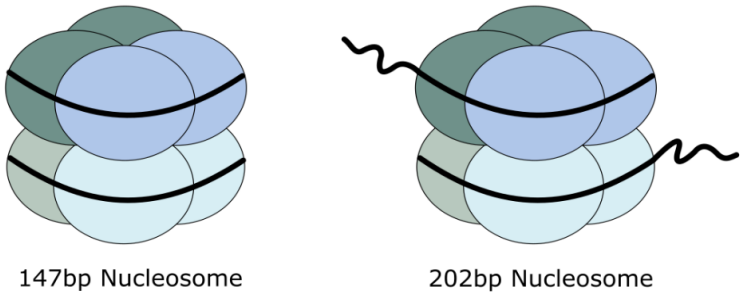

Figure 8 | **147 bp Nucleosome vs. 202 bp Nucleosome.** Nucleosome Reconstitutions using the 147 bp 601 widom sequence result in the formation of nucleosomes that do not contain any linker DNA. When the 202 bp DNA is used linker DNA of approximately 27 bp will protrude from each side of the nucleosome and cohesin should be able to interact with this accessible DNA.

2. Materials & Methods

2.1. General Buffers, Solutions and Media

All general Buffers, Solutions and Media are listed in Table 1. Reagents and Enzymes were ordered from New England Biolabs (NEB), Thermofisher or Sigma-Aldrich, unless stated otherwise. All protein purification steps were performed at 4°C.

Table 1 | **Buffers and Media**

Name	Composition
1M Tris	121.1g Trizima Base dissolved in Sterile Mono Q water and pH is adjusted with 35% HCl.
1x TE Buffer	Mix 200mL 100x TE pH 8 buffer with 1800 ml Mono Q Water and autoclave
Luria Bertani (LB) Medium	1200g LB Broth EZMixTM Powder (Sigma, Art.Nr.: L7658, 1kg) resuspended in 30L Mono Q Water, filled up to 60L, homogenized for 10 min and autoclave
0.5M EDTA pH 8.0	0.5mol/L EDTA dissolved in Sterile Mono Q water. Add NaOH crystals to clear the turbid solution and to increase the pH. Adjust to pH 8 with 10M NaOH and autoclave.
Grace's Insect Medium	112.5g Lactalbumin Hydrolysate, 112.5g Yeastolate, 12g of NaHCO3 and 30 packages of medium powder dissovled in 15L Millipore water. Fill up with Millipore Water to 30L and homogenize for at least 10 min. Adjust the pH with 10M NaOH to pH 6.18. Filter medium, final pH = 6.20.
Sf9 Histone Lysis Buffer	20mM HEPES pH 7.5, 0.25M Sucrose, 3mM $MgCl_2$, 0.5% (v/v) Nonidet P-40, 3mM 2-mercaptoethanol, 0.4mM PMSF
Sf9 Histone Buffer B	20mM HEPES pH 7.5, 3mM $MgCl_2$, 0.2mM EGTA, 3mM 2-mercaptoethanol, 0.4mM PMSF
Sf9 Histone HAP Buffer	50mM sodium phosphate pH 6.8, 0.6M NaCl, 1mM 2-mercaptoethanol, 0.5mM PMSF
SOC medium	100g Bacto Tryptone, 25g Bacto Yeast Extract, 2,92g NaCl, 0,93g KCl, 10,165g MgCl2x6H2O, 12,324g MgSO4x7H2O, 18,015g D-Glucose Anhydrous dissolved in 4L Mono Q Water, pH set to 7,0, filled up to 5L with water

2xTY medium	10g Yeast Extract, 16g Tryptone and 5g NaCl dissolved in 500ml of Sterile MonoQ water and filled up to 1L, pH is adjusted to pH 7.4 with 4M NaOH.
Recombinant Histone Lysis Buffer	20mM Tris pH 8.0, 2M NaCl, 1mM PMSF, 0.5mM TCEP
Ni-Buffer A	20 mM Tris pH 8.0, 2M NaCl, 0.5mM TCEP
Ni-Buffer A/30mM Imidazole	20 mM Tris pH 8.0, 2M NaCl, 0.5mM TCEP, 30 mM Imidazole
Ni-Buffer A/500mM Imidazole	20 mM Tris pH 8.0, 2M NaCl, 0.5mM TCEP, 500 mM Imidazole
RB Buffer High	2M NaCl, 10mM Tris pH 7.5, 1mM EDTA, 1mM DTT
RB Buffer Low	0.25M NaCl, 10mM Tris pH 7.5, 1mM EDTA, 1mM DTT
RP Lysis Buffer	50mM Sodium Phosphate, 500mM NaCl, 5% Glycerol, 10mM Imidazole, 0.1% Tween 20, 3mM β-Mercaptoethanol, 1mM PMSF, 1mM Benzamidin
RP Wash Buffer	50mM Sodium Phosphate, 500mM NaCl, 5% Glycerol, 15mM Imidazole, 0.1% Tween 20, 1mM PMSF, 1mM Benzamidin
RP Ni-Elution Buffer	50mM Sodium Phosphate, 150mM NaCl, 5% Glycerol, 300mM Imidazole
RP FLAG Buffer	50mM Sodium Phosphate, 100mM NaCl, 5% Glycerol, 50mM Imidazole
10x Coh Buffer	200mM Hepes pH 7.6, 500mM NaCl, 50% Glycerol, 20mM MgCl2, 10mM DTT
MP Buffer	20mM Hepes pH 7.6, 50mM NaCl, 2mM MgCl2, 5% Glycerol, 1mM DTT

2.2. Purification of Sf-9 Histones

3L Sf9 insect cells were grown at 27°C in Grace's Insect Medium supplemented with 10% FBS-HI, 1% L-Glutamine, 1% Pluronic F-68 and 1% Penicillin-Streptomycin to 3 mio cells/ml. Subsequently cells were pelleted at 4000rpm for 15 min and washed with 1xPBS. Pellets were resuspended in 20x pellet volume (~500 ml) lysis buffer, transferred to a Dounce homogenizer and lysed with 15 strokes. The lysate was centrifuged at 3000 x g for 15 min. Resuspension and pelleting were repeated twice with lysis buffer and once with buffer B, supplemented with cOmplete EDTA-free tablets, at 3000 x g for 5 min to wash away membranes and obtain a clean nuclei pellet. Next, the nuclei pellet was resuspended in 2x pellet

volume (~20 ml) buffer B, and one total volume (~24 ml) of buffer B/0.6M KCl/10% glycerol was added dropwise while gently stirring at 4°C. The solution was stirred for another 10 min to remove loosely bound proteins from the chromatin. Pelleting at 17500 x g for 30 min yielded the nuclear pellet.

The nuclear pellet was transferred to a Dounce homogenizer in 50 mL Hydroxylapatite (=HAP) buffer, resuspended with 1-2 strokes and sonicated at 20% amplitude for 5 min, 2 s on and 2 s off in a total volume of 100 ml HAP buffer. 10 ml Bio-Gel HTP Hydroxyapatite powder (Bio-Rad #1300420) were equilibrated in HAP buffer by centrifuging twice for 5 min at 1900rpm. The HTP powder was added to the nuclear lysate and rotated for 30 min at 4°C. After centrifugation at 1400rpm for 5 min the pellet was washed four times with 50 ml HAP buffer containing 0.6M NaCl to remove histone H1. The histones were eluted from the DNA by adding 10 ml HAP buffer containing 2.5M NaCl, resuspending the pellet and centrifuging at 1400rpm for 5 min. This step was repeated 10 times to collect 10 fractions. All fractions were loaded onto a 16% Tris-Glycine SDS-PAGE gel and run for 30 min at 230V. The fractions with highest yield were concentrated using Vivaspin 10kDa concentrators and flash frozen in liquid nitrogen.

2.3. Expression and purification of recombinant Histones

2.3.1. Expression of pET29a-YS14

The polycistronic plasmid pET29a-YS14 [52], containing the sequences for histones H3, H4, H2A and H2B from *Xenopus laevis*, was transformed into BL21(DE3)pLysS cells by adding 1 µl plasmid DNA to 50 µl cells followed by an 15 min incubation on ice. Heat shock was performed at 42°C for 30 s and cells were kept another 5 min on ice afterwards. Subsequently, cells were incubated for 1 h in 1 ml SOC Medium at 37°C, plated on a LB agar plate containing 50 µg/ml kanamycin and 25 µg/ml chloramphenicol and grown at 37°C overnight. Colonies were picked, inoculated in 2x 20 ml 2xTY medium and grown for 4 h. These starter cultures were grown in 2x 2 L 2xTY until an OD_{600} of 0.4. Expression of histones was induced by adding 0.4 mM IPTG and cultures were shaken overnight at 37°C. Cells were harvested by centrifugation at 4000 x g for 10 min and pellets were flash frozen until further use.

2.3.2. Nickel-affinity chromatography

Cell pellets were thawed and resuspendend in 100 ml lysis buffer, supplemented with cOmplete EDTA-free tablets, and 8U/ml DNase I, followed by sonication at 50% amplitude for 7 min, 2 seconds on, 2 seconds off. The lysate was centrifuged for 45 min at 18000rpm.

To perform fast protein liquid chromatography (FPLC) a BioRad NGC chromatography system was used. A 5 ml HisTrap Column was equilibrated with 2x 10 ml H_2O, followed by 2x column volumes (CV) Ni-Buffer A, 10x CV Ni-Buffer A/30mM Imidazole and finally 10x CV Ni-Buffer A. After the centrifugation the lysate was adjusted to a final concentration of 30 mM imidazole and loaded onto the column. Next, the bound proteins were eluted by increasing the imidazole concentration gradually from 30 mM to 500 mM over 23.5 CV (=117.5 ml). 2 ml fractions were collected and analysed on a 16% Tris-Glycine SDS-PAGE gel that was run at 230V for 30 min. Stoichiometric fractions were concentrated to 600 µl using a Vivaspin 30 kDa centrifugal concentrator. Subsequently, histones were digested overnight using 5U/100 µg Thrombin to remove the hexahistidine (His_6) tag from histones H2A and H4.

2.3.3. Size exclusion chromatography

A Superdex 200 10/300GL Column was equilibrated using 2x CV H_2O and 2x CV Ni-Buffer A. The thrombin digested proteins were loaded onto the column and 0.5 mL fractions were collected. Fractions were analysed by 16% Tris-Glycine SDS-PAGE gel. Stoichiometric fractions were concentrated in a 10 kDa concentrator.

2.4. Amplification and Purification of 147bp and 202bp DNA

To generate 147 bp and 202 bp DNA fragments a pUC19 plasmid containing the 601 widom sequence was used. Primers were designed to amplify either 147 bp long widom, or a 202 bp DNA sequence with additional base pairs flanking the 601 sequence on both sides. A gradient polymerase-chain-reaction (PCR) was performed using the respective primers to find optimal annealing temperatures and products were analysed on a 1% agarose gel. Using the optimized conditions, a 96-well plate PCR was performed to amplify 147 bp or 202 bp DNA sequences (Table 2). Products were run over a 1% agarose gel to confirm correct amplification.
PCR products were purified by PEG precipitation. One volume of 2.5 M NaCl, 20% PEG6000 was added and samples vortexed thoroughly. Samples were centrifuged at 14500 rpm for 15 min, pellets washed with cold 70% ethanol and centrifuged for 3 min at 14500rpm. The supernatant was discarded, and samples were left to dry for 15-30 min. Dried pellets were resuspended in 500 µl buffer (10mM HEPES pH 7.6, 50mM NaCl, 0,5mM EDTA).

Table 2 | Left: **PCR Mastermix,** right: **protocol for amplification of DNA**

Material	Volume	Steps	Temperature		Time
5x HF Buffer	1 ml	Initial Denaturation	98°C		30s
10mM dNTPs	100 µl	Denaturation	98°C	35x	10s
Plasmid DNA	50 µl	Annealing	50°C		10s
Forward Primer	250 µl	Extension	72°C		15s
Reverse Primer	250 µl	Final Extension	72°C		10 min
DMSO	150 µl				
Phusion Polymerase	50 µl	Hold	4°C		Infinite
Sterile H_2O	3.15 ml				

2.5. Nucleosome Reconstitutions

Nucleosomes were reconstituted via salt gradient dialysis using a peristaltic pump. DNA and histones were mixed in a molar ratio of 1:1 at a high salt concentration (2M NaCl). Additionally, 1,4-Dithiothreit (DTT) was added to a final concentration of 10 mM. The reaction was left on ice for 30 min, placed in 400 ml RB buffer high and dialyzed against 1600 ml RB buffer low using Slide-A-Lyzer MINI Dialysis Device, 20K MWCO. The peristaltic pump was set to a flow rate of 0.7 ml/min, reducing the salt concentration gradually from 2M NaCl to 0.25M NaCl over 36 h, which allows histones and DNA to assemble into nucleosomes spontaneously. After dialysis the DNA concentration was measured using a NanoDrop 1000 Spectrophotometer and nucleosomes were analysed by electrophoretic mobility shift assay (EMSA, see chapter 2.12.).

2.6. Cloning of ΔN-NIPBL2A

To produce a NIPBL DNA binding mutant, which was previously identified and characterized by a postdoc in the lab, primers were designed to mutate Arg_{1867} and Lys_{1870} to Alanine [34]. The template plasmid that was used contained the sequence for N-terminal truncated NIPBL, a Polyhistidine-Tag and FLAG-tag for tandem-affinity purification. An "around-the-horn PCR" (Table 3) was performed with respective primers to introduce the mutations and products were analysed by 1% agarose gel. Template DNA was digested with DpnI for one hour at 37°C and the PCR product purified using the MiniPEx kit (in house DNA cleaning kit). The PCR product was eluted in 26 µl TE buffer and ligated by adding 1 µl PNK4, 1 µl T4 ligase and 3 µl T4 ligase buffer at 37°C for 1 h. Ligated plasmids were transformed into XL10 gold cells and sequenced to confirm successful insertion of the mutations.

Table 3 | Left: **PCR mastermix,** right: **protocol for around-the-horn PCR**

Materials	Volume	Steps	Temperature		Time
5x HF Buffer	50 µl	Initial Denaturation	98°C		30s
10mM dNTPs	5 µl	Denaturation	98°C	35x	10s
Plasmid DNA	5 µl	Annealing	Gradient 50°C-60°C		10s
Forward Primer	12.5 µl	Extension	72°C		7 min
Reverse Primer	12.5 µl	Final Extension	72°C		10 min
DMSO	7.5 µl				
Phusion Polymerase	2.5 µl	Hold	4°C		Infinite
sterile H_2O	155 µl				

2.7. Transformation of plasmids into competent *E. coli* cells

For amplification of respective plasmids transformation into chemically competent DH5α or, for higher efficiencies, XL10 gold cells was performed via heat-shock.
1 or 5 µl of purified plasmid were mixed with 50 µL of competent cells and kept on ice for 10-30 min. Heat-shock was performed at 42°C for 30 s and followed by an incubation on ice for 5-10 min. 1 ml of SOC medium was added and cells were incubated at 37°C for 1 h while shaking at 800 x g. Next, cells were plated on LB plates, supplemented with respective antibiotics, and incubated at 37°C overnight. Colonies were picked, inoculated in 5 ml LB medium containing respective antibiotics and grown overnight at 37°C, 800 x g. For Plasmid Isolation the MiniPEx kit was used.

2.8. Agarose gel electrophoresis

Agarose powder was dissolved in 1xTAE Buffer to a final concentration of 1% (w/v), heated and poured into a gel casting tray where the gel was left to polymerize. SYBR Safe DNA gel stain was added to the gel while it was still liquid. Samples were mixed with 4x DNA loading dye and the gel was run in 1xTAE Buffer at 130V for 30-50 min, depending on the DNA fragment size.

2.9. Sf9 Insect Cell Culture

Sf9 insect cells are cultured at 27°C, 100 rpm, in Grace's Insect Medium supplemented with 10% FBS-HI, 1% L-Glutamine, 1% Pluronic F-68 and 1% Penicillin-Streptomycin. Cell cultures are maintained at a concentration of 1 mio cells/ml for amplification and 1.3 mio cells/ml for infection with recombinant Baculovirus.

2.10. Baculovirus production

To generate recombinant baculoviruses pLIB and pBIG vectors containing NIPBL or cohesin constructs under a polyhedrin promotor for their expression in Sf9 cells were transformed into DH10 Multibac cells to generate bacmid DNA [53]. 3 µl of plasmid DNA are mixed with 50 µl DH10MultiBac cells and kept on ice for 1 min. The cells are transferred into a pre-cooled electroporation cuvette and electroporated at 1800V. After 4 h incubation in 1 ml SOC medium at 37°C, 800 x g, cells are plated onto blue-white selection LB-agar plates supplemented with 50 ng/µl X-gal, 10 ng/µl Tetracycline, 50 ng/µl Kanamycin and 7 ng/µl Gentamycin and incubated at 37°C overnight. A white colony is picked and inoculated in 3 ml LB medium supplemented with antibiotics as mentioned above at 37°C overnight. The cells are centrifuged at 5000rpm for 10 min and resuspended in 300 µl resuspension buffer (MiniPEx kit) and lysed with 300 µl lysis buffer (MiniPEx kit) before 300 µl neutralization buffer (MiniPEx kit) are added. After 10 min centrifugation at 14500rpm, 700 µl of the lysate are mixed with 800 µl isopropanol, inverted 20x and subsequently incubated at -20°C for 20 minr. The sample is pelleted for 1 min at 14500rpm, washed twice with 70% Ethanol, air-dried for ~1 h and finally resuspended in 40 µl sterile H_2O.

For the transfection into Sf9 cells, 1 mio cells are seeded in 3 ml Sf-900 II SFM into a 6-well plate and allowed to attach to the well by incubating at 27°C for 2 h. 10 µl plasmid DNA are mixed with 200 µl Sf-900 II SFM, and 10 µl FuGENE 6 transfection reagent (Promega #E2691) with 100 µl Sf-900 II SFM. Both reactions are incubated at room temperature (RT) for 5 min. The two reactions are mixed and kept at RT for additional 20 min. 150 µl of the reaction is added dropwise to the cells and incubated for ~72 h at 27°C. Afterwards the supernatant which contains recombinant baculovirus (V0) is collected.
To amplify the virus a V1 recombinant baculovirus is produced by adding 3 ml of V0 to 100 ml Sf9 cells and incubating for ~48 h at 27°C.

2.11. Recombinant Protein Expression and Purification

2.11.1. Recombinant Protein Expression

To express recombinant proteins 3 L of Sf9 cells are adjusted to 1.3 mio cells/ml and infected with 1% (v/v) of respective recombinant baculovirus and incubated at 27°C for ~48 h. At ~70% infection rate, cells are harvested at 4000rpm for 15 min. Cell pellets are resuspended in 100 ml 1xPBS and centrifuged for 5 min at 4000rpm. Pellets are washed with 1xPBS, snap frozen in liquid nitrogen and stored at -80°C.

2.11.2. Recombinant Protein Purification

The cell pellets are thawed in a water bath, transferred to a Dounce Homogenizer with 150 ml Recombinant Protein (RP) lysis buffer and lysed with 25 strokes. Lysed cells are centrifuged at 18500rpm for 35 min. In the meantime, 5 ml of Toyopearls are equilibrated by washing one time with 10 ml H_2O and two times with 20 ml RP lysis buffer. The cleared lysate is added to the equilibrated beads and incubated for 2-3 h, rotating. Beads were collected by centrifugation at 1200rpm for 5 min, washed three times with RP wash buffer and transferred to a glass column. The proteins are eluted by adding 5 x 5 ml of RP Ni-Elution buffer. The eluate was added to 5 ml equilibrated ANTI-FLAG beads and incubated for 3 h while rotating. RP FLAG elution buffer was prepared by mixing 0.5 mg/ml 3xFLAG peptide with RP FLAG buffer, rotating for 30 min at RT and incubating on ice until further use. After 3 h the beads were collected, transferred to a glass column and washed 3 x with 25 ml RP FLAG buffer. The beads were incubated with 5 ml of RP FLAG elution buffer for 5 min while stirring occasionally. The eluate was collected and another 4 x 5 ml of elution Buffer were added for a final elution volume of 25 ml. The eluted proteins were concentrated to ~750 µl in a Vivaspin 100 kDa centrifugal concentrator and analysed on 6% Tris-Glycine SDS-PAGE gel.

2.12. Electrophoretic Mobility Shift Assays (EMSAs)

Binding of proteins to DNA or nucleosomes was analysed using Electrophoretic Mobility Shift Assays (EMSAs). 4.5% Acrylamid-Bisacrylamid gels (Table 4) were poured into gel casts and left to solidify for 20 min. Samples were prepared using 5 µl master mix (MM) (Table 4) and 5 µl protein dilution. If no protein was used 5 µl of RP FLAG buffer were added. Reactions were incubated at RT for 10-20 min and run in 0.5x TBE buffer at 100V for 50 min. Gels were stained with SYBR gold (1:10 000) for 10 min and destained in sterile H_2O for 5 min.

Table 4 | Left: **Reaction mix for 4.5% EMSAs,** right: **Mastermix for EMSA samples**

Material	Final concentration	Materials MM	Final Concentration
40% AB	4.5%	10x Coh Buffer	1x
10x TBE	0.5x TBE	30% Sucrose	3.75%
10% APS	0.1%	DNA/Nucleosome	10nM
TEMED	0.05%	ATP	2.5mM
		H_2O	rest

2.13. Quantification of EMSAs

The intensity of nucleosome bands was measured using ImageJ. The band with highest intensity, corresponding to the sample containing no protein, was set to 100% and from all following bands the corresponding percent were calculated. This would represent the percentage of unbound nucleosome fraction. Calculating $100 - x\%$, with x being the amount of unbound nucleosome calculated before, will yield the fraction of bound nucleosome in %.

2.14. Mass Photometry

2.14.1. Glass Slide Preparation

High Precision Cover Glasses 1.5H (Marienfeld #0107222) were prepared by sequential sonications in 2% Hellmanex for 10 min, sterile H_2O for 5 min, 100% isopropanol for 5 min and sterile H_2O for 5 min.
Coverslips were dried under an air stream and silicon buffer gaskets (Grace Bio-Labs CultureWell gaskets) were attached to the glass surface.

2.14.2. APTES modified glass slides

APTES modified slides were cleaned as described above. After drying, the slides were cleaned in a plasma cleaner with oxygen for 8 min. The slides were incubated in 2% APTES in acetone for 1 min while gently shaking. The slides were washed in acetone and incubated at 110 °C for 1h. Finally, the slides were sonicated for 10 min in 100% isopropanol and 5 min in H_2O before drying under an air stream. [54]

2.14.3. Measurements

All measurements were performed on the Refeyn Mass Photometer. Samples containing only one protein were diluted to 200-250 nM in MP buffer. Samples containing mixtures of protein and nucleosome or protein and DNA were diluted to varying concentrations, mixed and incubated at RT for 10 min. 2 uL were further diluted in a gasket filled with 18 uL MP buffer. 6000 frames were recorded using Refeyn AcquireMP and analysed using Refeyn DiscoverMP. For contrast-to-mass calibration a native protein marker was pre-diluted 1:50 in MP buffer followed by a 1:10 dilution for the measurement.

2.15. HS-AFM

All HS-AFM measurements were performed in collaboration with a PhD Student from the Peters Lab using the RIBM HS-AFM NanoExplore (NEX).

The cantilever holder was cleaned by sequential sonication in 2% SDS, MQ, isopropanol, MQ for 10 min and dried subsequently. A 1.5 mm diameter mica disc was placed on a glass rod and fixated with two-phase glue. The mica disc was left to dry overnight. The mica disk was cleaved using tape and modified using 0.01% poly-l-lysine that was incubated on the mica for 5 min before washing with 50mM NaOAc and RB buffer low.
Samples were diluted to ~0.5 ng and incubated on the mica for 5 min before washing with RB buffer low. For measurements a micro cantilever from Olympus (BL-AC10DS-A2) was used and analysed using the Kodec software.

3. Results

3.1. Purification and characterization of human cohesin and human NIPBL

Recombinant human cohesin tetramer and NIPBL were expressed in Sf9 insect cells using baculoviruses that have been generated in the Peters lab before. Proteins were purified using a two-step His_6-3x FLAG tandem affinity purification and analysed on 6% SDS-PAGE gels together with a BSA-standard to determine the respective protein concentration. Following this approach resulted in pure recombinant proteins (Fig. 9B).

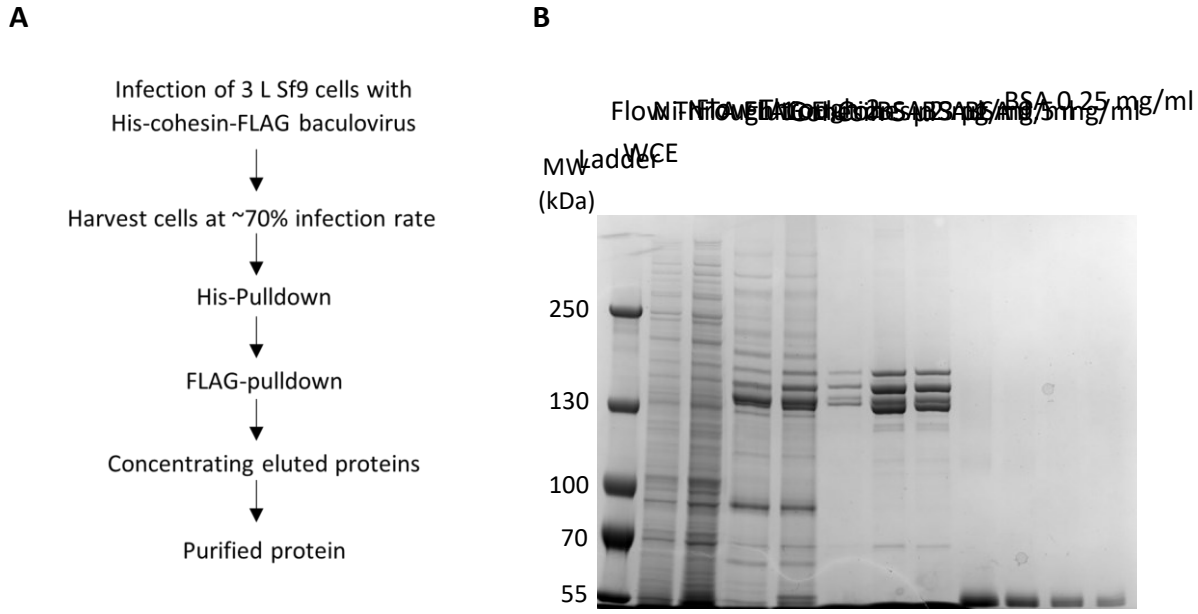

Figure 9 | **Tandem purification of tetrameric cohesin. A** | Schematic overview of the purification steps. **B** | Representative SDS-PAGE of a cohesin purification. WCE= whole cell extract, Flow through 1 sample is taken after incubation with Ni-NTA beads, Flow Through 2 sample is taken after incubation with ANTI-FLAG beads. 10 µl have been loaded after the FLAG elution, 5 µl and 3 µl have been loaded after concentrating the purified protein and quantified using the BSA standard.

3.2. Histone purification

3.2.1. Purification and characterization of Sf9 histones

Sf9 core histones were purified from 3 L insect cells using detergent and physical shear force, yielding the nuclear pellet. DNA from the nuclear extract was bound to hydroxylapatite resin and core histones were eluted using high salt buffer (Fig. 10A). This resulted in relatively pure and stoichiometric fractions of purified histones (Fig. 10B). Fractions with highest yield were pooled and concentrated.

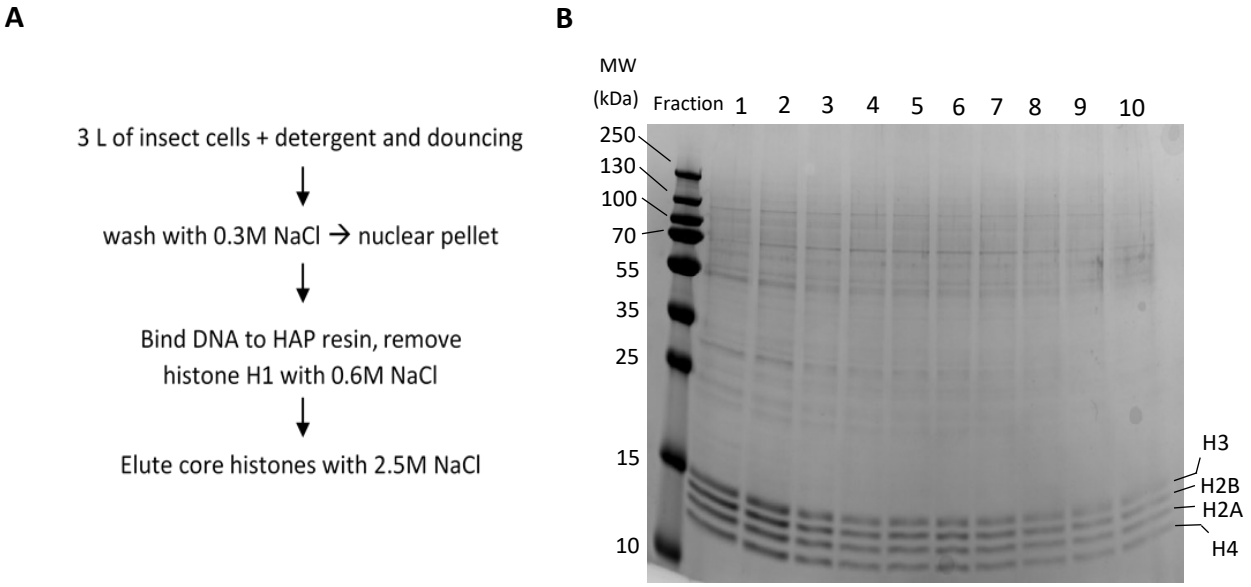

Figure 10 | **Sf9 Histone Purification**. **A** | Schematic overview of Sf9 histone purification steps. **B** | Fractions 1-10 of purified Sf-9 Histones analysed by 16% Tris-Glycine SDS-PAGE. The four bands correspond to histones H3, H2B, H2A and H4.

3.2.2. Purification and characterization of *Xenopus laevis* recombinant histones

Since histones purified from Sf9 cell chromatin may be affected by histone modifications, histones from *Xenopus laevis* were recombinantly expressed and purified from *E.coli* additionally (Fig. 11A). In a first step, histones were purified using Ni-affinity chromatography. Fractions showing the highest stoichiometry were pooled, digested with thrombin and purified in a second step by size-exclusion chromatography (Fig. 11B)

Compared to histone purification from insect cells, this method was more time-consuming and gave with ~ 0.19 mg/ml a 10x lower yield.

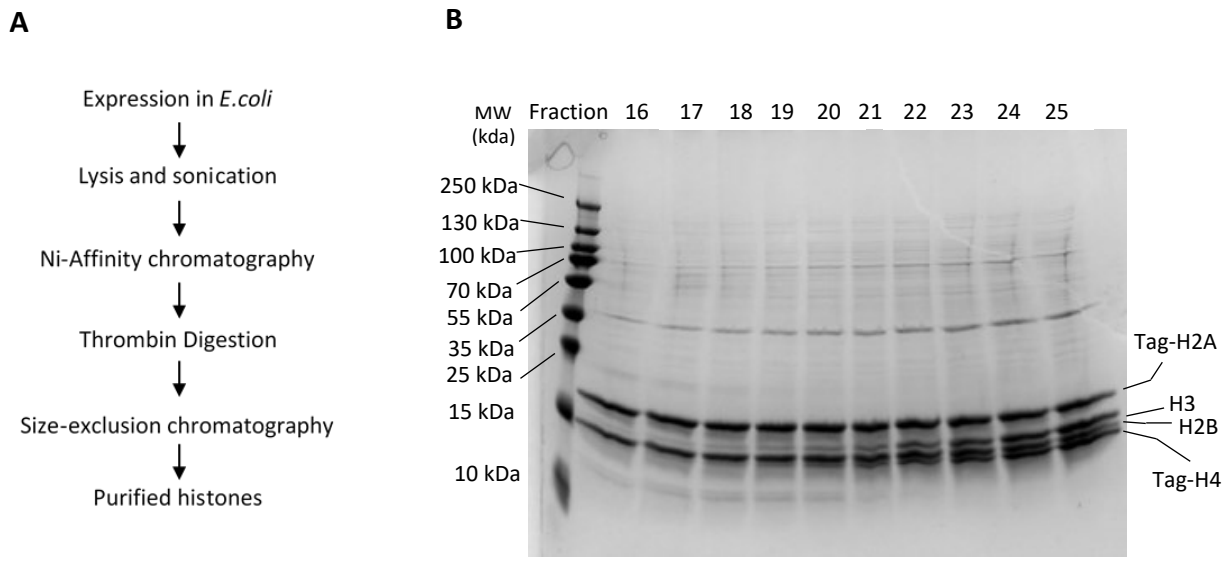

Figure 11 | **Purification of recombinant histones. A** | Schematic overview of recombinant histone purification. **B** | 16% SDS-PAGE gel showing fractions of the four core histones after Ni-affinity chromatography.

3.2.3. Quality control of Sf9 and recombinant histones

Nucleosomes were reconstituted using Sf9 or recombinant histones and the 147 bp long 601 widom sequence and analysed using 4.5% polyacrylamide EMSAs and MP (Fig. 12). MP showed the expected nucleosome size of 180 kDa (Fig. 12B), EMSA running profiles were comparable for both nucleosome reconstitutions (Fig. 12A).

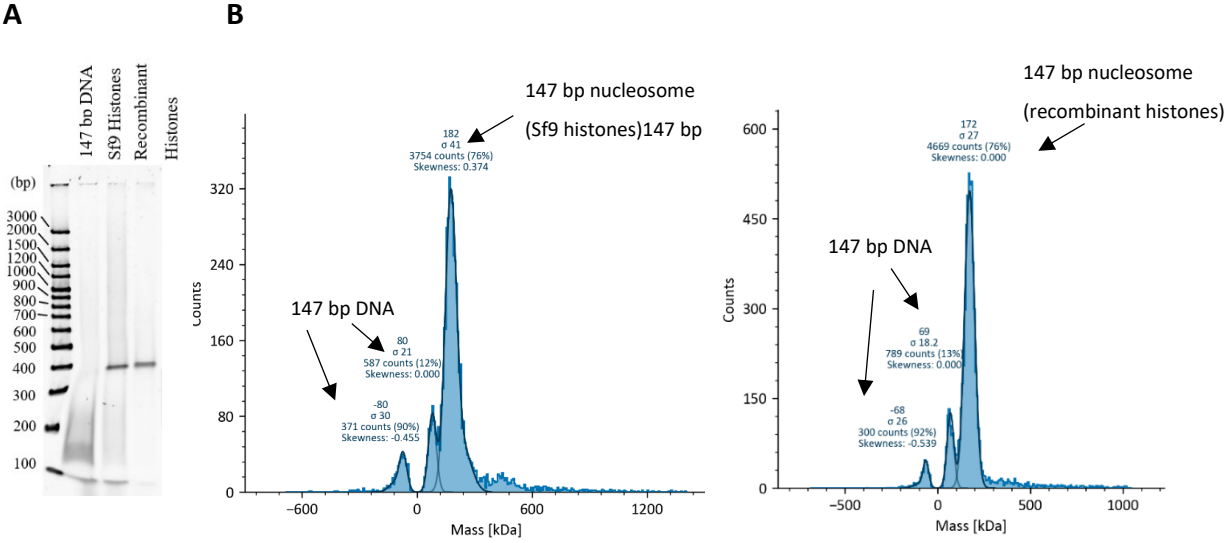

Figure 12 | **Comparing Sf9 and recombinant histone reconstitutions. A** | EMSAs comparing 147 bp nucleosomes containing Sf9 histones (left) and 147 bp nucleosomes containing recombinant histones (right). U = unbound nucleosome, B = bound nucleosome. **B** | Mass Photometry of 147 bp nucleosomes containing Sf9 Histones (left) and recombinant histones (right). Peaks at the negative axis correspond to unbinding events.

3.3. Cohesin binding to nucleosomes analysed by EMSAs

In this study polyacrylamide EMSAs were used to analyse the binding of cohesin complexes and NIPBL to nucleosomes and DNA. The difference in size and mobility of the nucleosome compared to free DNA results in an upward shift of the nucleosome-bound DNA band. When the nucleosome or free DNA is bound to a protein the upward shift in the gel will become even more prominent (Fig. 13A).

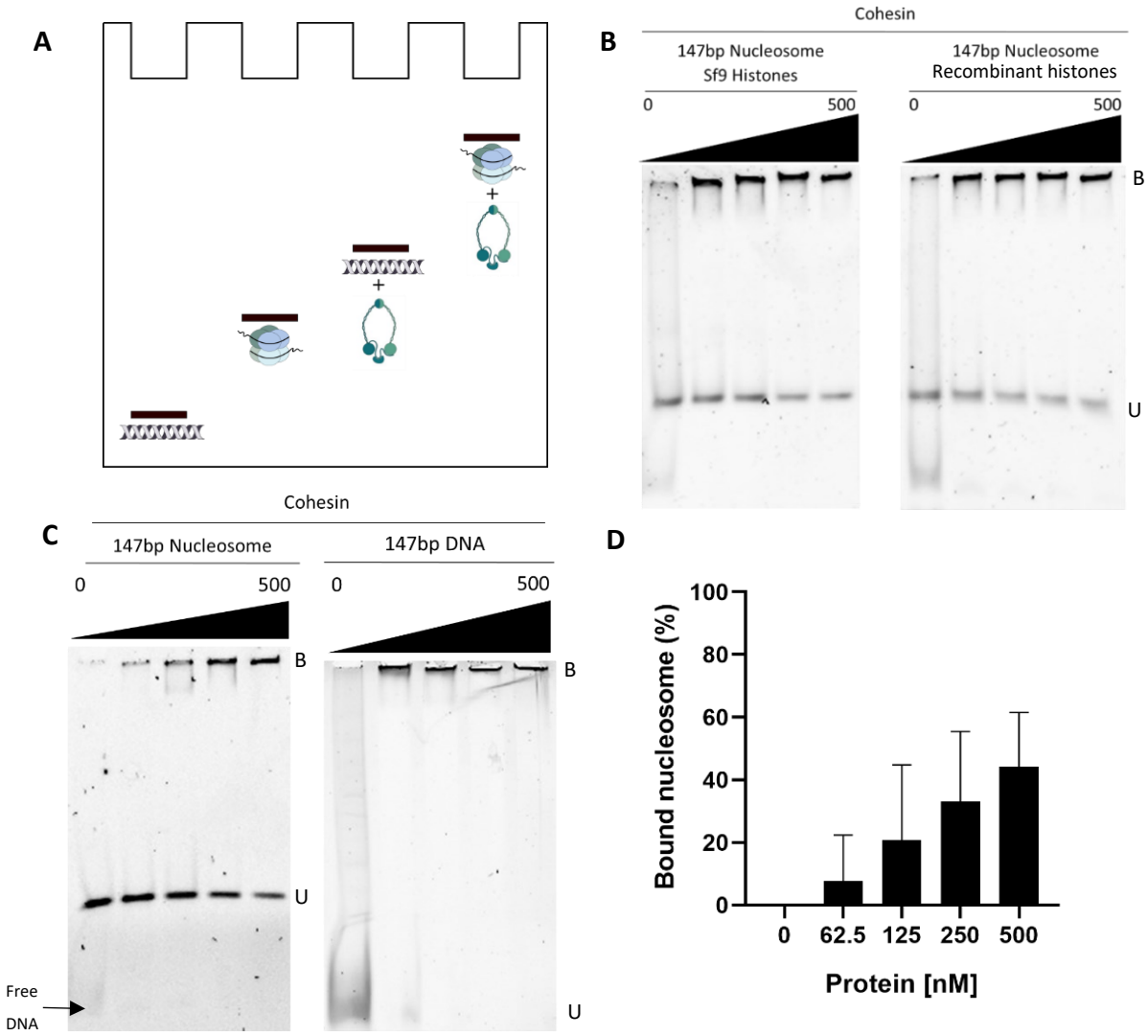

Figure 13 | **Principle of EMSAs. A** | The free DNA has the highest electrophoretic mobility and therefore traverses through the gel the fastest. When the DNA is wrapped around a nucleosome the band is shifted upwards. Similarly, binding of free DNA to cohesin results in a large upward shift and when nucleosome-bound DNA associates with cohesin the shift that can be observed is the largest. **B** | EMSA comparing binding of cohesin to 147 bp nucleosomes that were reconstituted using Sf9 histones (left) and recombinant histones (right). U=unbound nucleosome or unbound DNA, B=bound **C** | Representative EMSAs showing binding of cohesin to 147 bp nucleosome (left) and to 147 bp DNA (right). Cohesin was titrated to nucleosomes and DNA up to a concentration of 500 nM. U=unbound nucleosome or unbound DNA, B=bound **D** | Averaged data showing increase of bound 147 bp nucleosome plotted against the amount of protein that was added. Data represents 12 independent experiments.

3.3.1. Binding of cohesin WT to the 147bp Nucleosome

147 bp nucleosomes and free 147 bp DNA were incubated for 10-15 min with different concentrations of cohesin tetramer in 10x Cohesin Buffer and 2.5 mM ATP and subsequently analysed by EMSAs. Note that both cohesin and NIPBL did not enter the gels but were stuck in the wells. Therefore, binding can only be observed by the disappearance or intensity reduction of the nucleosome bound and free DNA band. An increase in fluorescence signal in the well can also be observed, indicating binding of nucleosome or DNA to the protein, but this proved to be difficult to quantify due to the presence of unbound DNA that did not get incorporated into nucleosomes during the reconstitution (Fig. 13C). This free DNA will bind to proteins with high affinity and thereby increase the fluorescence signal in the well.

Since binding of cohesin to 147 bp nucleosomes containing Sf9 histones or recombinant histones showed no significant difference (Fig. 13B) all further experiments will be performed using Sf9 histones, as the method for Sf9 histone purification is more rapid, simple and results in higher yields.

As expected from literature cohesin showed high affinity to 147 bp free DNA already at low cohesin concentration of 62,5 nM [25][45] (Fig. 13C). In contrast, binding of cohesin to 147 bp nucleosomes was only detected at high concentrations of 250-500 nM (Fig. 13C,D). This suggests that cohesin has low affinities to nucleosome-bound DNA, or that nucleosome-bound DNA is less accessible (Fig. 13D).

3.3.2. Binding of the cohesin EQ mutant to the 147 bp Nucleosome

It is known from cryo-EM structures that DNA is clamped on cohesin's ATPase heads when they are engaged [18]. Based on this, 'forcing' cohesin to stay in its clamped state may allow more stable binding to DNA. Therefore, the ATPase deficient mutant SMC1A-E1157Q/SMC3-E1144Q (EQ), which can bind but not hydrolyse ATP was analysed for its binding to nucleosomes and compared to wildtype (WT) cohesin.

The EQ mutant was titrated to the 147 bp nucleosome in the presence of 2,5 mM ATP and binding efficiencies were compared to WT cohesin (Fig. 14A,B). At lower concentrations the binding efficiencies of WT cohesin and EQ mutant appear similar, but at higher concentrations the EQ mutant seems to bind nucleosomes more efficiently than WT cohesin. This becomes more evident when plotting the bound nucleosome fractions against the final concentration of

cohesin WT or EQ mutant (Fig. 14A) and suggests that the clamped state may help cohesin to stay on nucleosomal DNA.

As a control binding of the EQ mutant to the 147 bp nucleosome was analysed in the absence of ATP, where the binding should be comparable to WT cohesin. Additionally, the binding efficiency was compared to the EQ mutant in the presence of ATP (Fig. 14C,D). Indeed, while the EQ mutant binds to nucleosome-bound DNA in the absence of ATP similarly to WT cohesin, it can bind to DNA already at lower concentrations when ATP is present. When cohesin WT was added to the 147 bp nucleosome without ATP no significant difference could be seen compared to reactions in the presence of ATP.

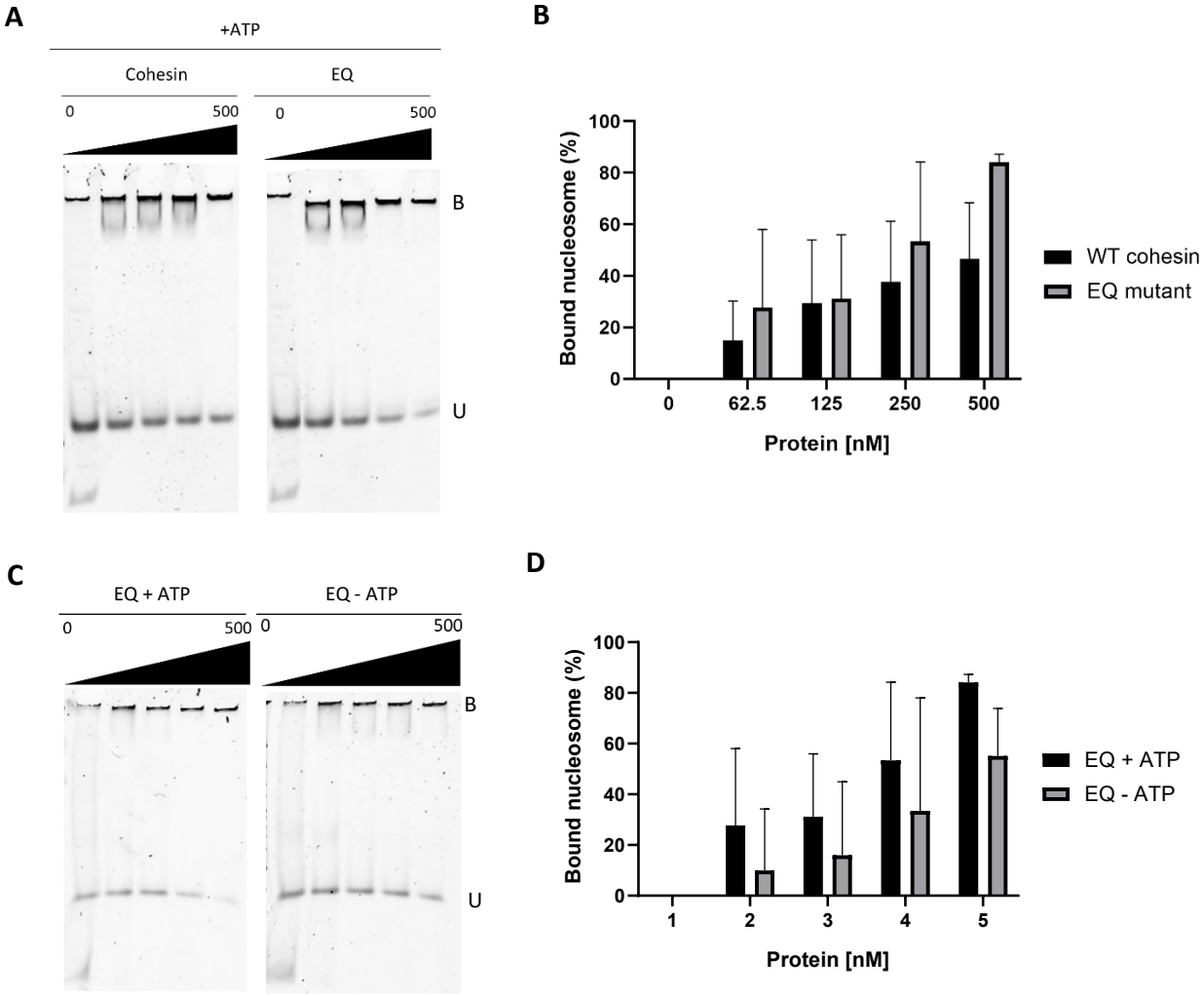

Figure 14 | **Binding of Cohesin WT and EQ mutant to the 147 bp nucleosome. A** | Representative EMSAs comparing binding of 147 bp nucleosome to cohesin WT and EQ mutant. Increasing concentrations of cohesin WT or EQ mutant were incubated for 10-15 min in the presence of ATP and separated by native PAGE. U = unbound nucleosome, B = bound nucleosome. **B** | Binding of cohesin WT and EQ mutant to the 147 bp nucleosome in the presence of ATP of five independent experiments. **C** | Representative EMSAs showing titration of EQ mutant with and without ATP to the 147 bp nucleosome. U = unbound nucleosome, B = bound nucleosome. **D** | Binding of the EQ mutant to the 147 bp nucleosome in the presence and absence of ATP from five independent experiments

Together, these results indicate that cohesin WT can bind to nucleosome-bound DNA with low affinity. Locking cohesin into a head-engaged state increases its ability to bind to nucleosomes in the presence of ATP, suggesting that the clamped-state might stabilize cohesin's interaction with nucleosome-bound DNA.

3.3.3. Binding of cohesin WT to the 202 bp Nucleosome

As cohesin could only bind to nucleosome-bound DNA at high concentrations compared to free DNA, a nucleosome containing 202 bp DNA was reconstituted. Performing EMSAs with the 202bp nucleosome and cohesin will help to ensure that EMSAs are a suitable method to observe nucleosomes and binding of cohesin, because free DNA of approximately 28 bp should protrude from both sides of the nucleosome and be freely accessible to cohesin. To test this EMSAs were performed using the same conditions as described before.

When titrating cohesin to the 202 bp nucleosome, binding can be observed already at a concentration of 62,5 nM while almost all nucleosomes seem to be bound to cohesin at a final protein concentration of 500 nM (Fig. 15). Therefore, the free linker DNA protruding from the 202 bp nucleosome allows cohesin to bind at lower concentrations compared to the 147 bp nucleosome (Fig. 15A), which was confirmed in five independent experiments (Fig. 15B).

While nucleosome-bound DNA seems to be less accessible for cohesin, it can strongly interact with free linker-DNA protruding from the nucleosome core.

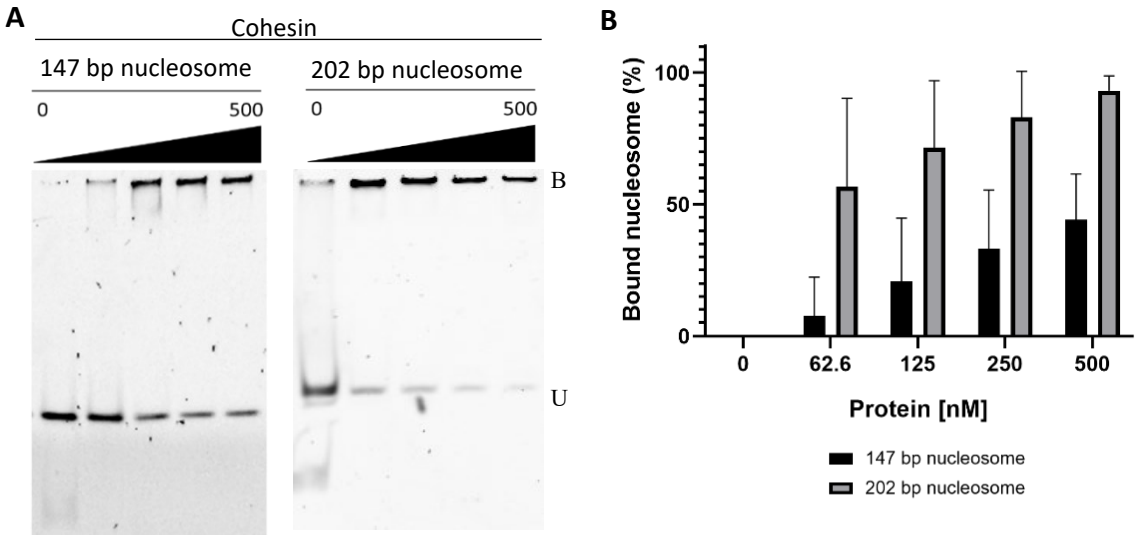

Figure 15 | **Binding of Cohesin to 202 bp nucleosomes. A** | Representative EMSAs comparing binding of cohesin to 147 bp nucleosomes and 202 bp nucleosomes. **B** | Comparison of cohesin's binding efficiency to 147 bp nucleosomes and 202 bp nucleosomes. Data represents five independent experiments., U = unbound nucleosome, B = bound nucleosome

3.3.4. Binding of NIPBL to 147 bp and 202 bp Nucleosomes

NIPBL is important for loading of cohesin onto DNA and for DNA to be clamped on top of cohesin's ATPase heads [34][18]. Therefore, it could play an important role for cohesin's association with nucleosomes.

To test this hypothesis, EMSAs were performed with the same conditions described above. NIPBL strongly interacted with the 202 bp nucleosome, possibly with the free-linker DNA (Fig. 16A). Surprisingly, NIPBL could also interact with 147 bp nucleosomes at low protein concentrations. In fact, when comparing binding of NIPBL to 147 bp and 202 bp nucleosomes similar binding efficiencies are observed (Fig. 16B).

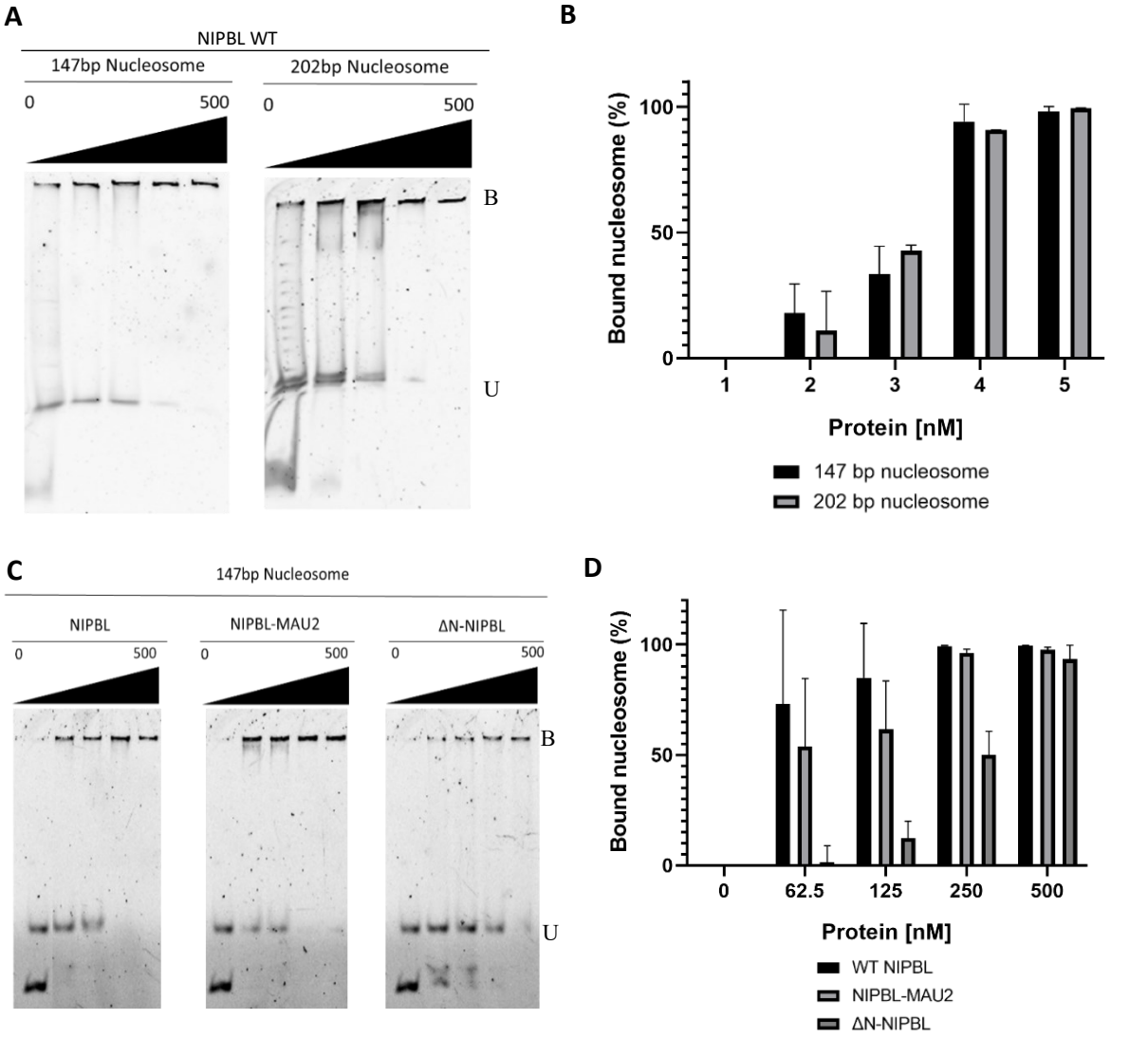

Figure 16 | **Binding of NIPBL to 147 bp and 202 bp nucleosomes**. **A** | Representative EMSAs showing binding of NIPBL WT to 147 bp and 202 bp nucleosomes. U = unbound nucleosome, B = bound nucleosome **B** | Comparison of NIPBL WT binding to 147 bp nucleosomes and 202 bp nucleosomes. The proteins were incubated 10-15min and separated by native PAGE in three independent experiments. **C** | Representative EMSAs showing binding of NIPBL, NIPBL-MAU2 and ΔN-NIPBL variants to the 147 bp nucleosome. U = unbound nucleosome, B = bound nucleosome. **D** | Comparison of NIPBL WT, NIPBL-MAU2 and ΔN-NIPBL binding to 147 bp nucleosomes from three independent experiments.

Because NIPBL forms heterodimers with its binding partner MAU2 to load cohesin onto chromatin *in vivo* [55], and because MAU2 may be important for opening chromatin or to stabilize NIPBL [55][56], binding efficiencies of NIPBL and NIPBL-MAU2 were compared (Fig. 16C). MAU2 doesn't seem to improve binding of NIPBL to nucleosomes and appears to be even slightly reduced compared to NIPBL WT at low concentrations (Fig. 16D), maybe due to steric hindrance when MAU2 is present. This indicates that MAU2 may be dispensable for binding of NIPBL to nucleosomal DNA.

An N-terminal truncated version of NIPBL (ΔN-NIPBL) has been shown to be able to bind to free DNA and to be sufficient for clamping DNA onto cohesin [18][34]. To test if this mutant is also able to bind to DNA wrapped around nucleosomes, it was purified, titrated to nucleosomes and binding observed by EMSAs (Fig. 16C). This revealed that the N-terminal truncated version of NIPBL seems to have lower binding affinities compared to WT NIPBL indicating that the N-terminal region of NIPBL plays a key role in binding to nucleosomes (Fig. 16D).

To address if NIPBL interacts with nucleosome-bound DNA or with the histones themselves a DNA-binding mutant, the NIPBL mutant ΔN-NIPBL2A, was purified, its binding to 147 bp and 202 bp nucleosomes analysed via EMSAs and compared to binding to naked DNA of the respective length (Fig. 17). ΔN-NIPBL2A carries two point mutations that have been shown to completely abolish binding of NIPBL to DNA [34].

Similar to naked DNA, ΔN-NIPBL2A was also deficient in binding to 147 bp nucleosomes, indicating that NIPBL interacts with nucleosome-bound DNA and not with histones directly (Fig. 17A) [34]. To ensure that the reduction in binding is indeed caused by the mutation of the DNA binding sites in NIPBL and not the N-terminal truncation, binding of ΔN-NIPBL WT and ΔN-NIPBL2A to 147 bp nucleosomes were compared (Fig. 17B). This confirmed again that the N-terminal domain of NIPBL plays a key role in nucleosome binding, however nucleosome binding is completely abrogated for the DNA-binding mutant ΔN-NIPBL2A. In line with this, NIPBL2A is also unable to bind to 202 bp nucleosomes or 202 bp DNA (Fig. 17C).

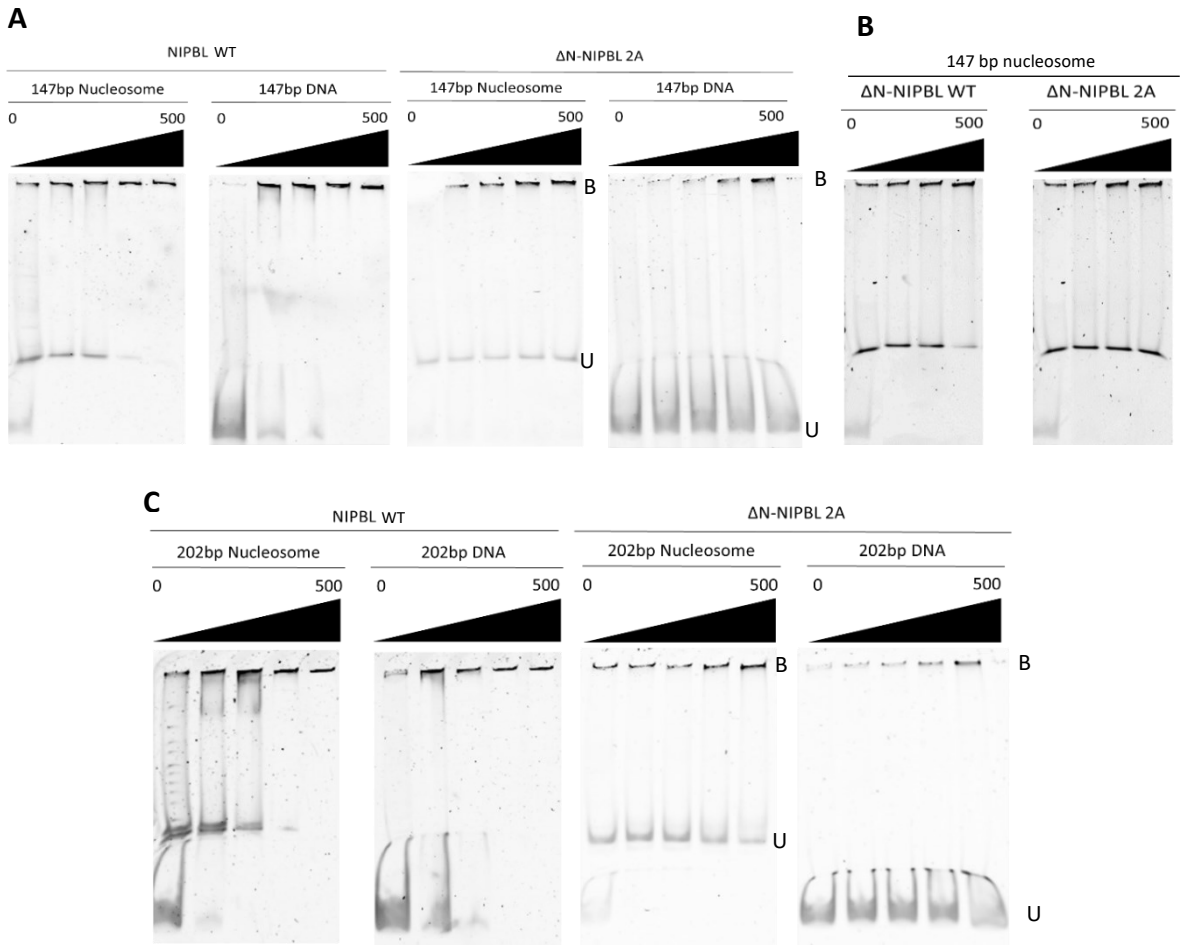

Figure 17 | **Binding of ΔN-NIPBL2A to 147 bp and 202 bp nucleosomes**. **A** | Left: Binding of NIPBL WT to 147 bp nucleosomes and 147 bp DNA. Right: Binding of ΔN-NIPBL2A to 147 bp nucleosomes and 147 bp DNA. **B** | Comparing binding of ΔN-NIPBL WT and ΔN-NIPBL2A to 147 bp nucleosomes. **C** | Left: Binding of NIPBL WT to 202 bp nucleosomes and DNA. Right: Binding of ΔN-NIPBL2A to 202 bp nucleosomes and DNA. In each case NIPBL was titrated from 0 to 500 nM to the nucleosomes, U = unbound nucleosome/DNA, B = bound nucleosome/DNA.

Taken together, these results indicate that NIPBL may play an important role in interaction with nucleosome-bound DNA. However, it is not clear if NIPBL loads cohesin onto the nucleosomes, if it holds cohesin on nucleosome-bound DNA or if its association with nucleosomes has a completely different function.

3.3.5. Binding of Cohesin and NIPBL to 147 bp Nucleosomes

To address if NIPBL can load or stabilize cohesin on nucleosome-bound DNA the two proteins were added to 147 bp nucleosomes in a ratio of 2:1 (NIPBL:cohesin) and EMSAs were performed (Fig. 18).

There is no evident difference between binding of cohesin and NIPBL compared to NIPBL alone. As the proteins do not run through the gel but are stuck in the well it is not possible to

distinguish between complexes of only NIPBL and 147 bp nucleosome or complexes that contain cohesin, NIPBL and 147 bp nucleosome (Fig. 18A).

Additionally, in EMSAs there is no difference in binding of cohesin to the 147 bp nucleosome in the presence of the N-terminal truncated version, ΔN-NIPBL, or ΔN-NIPBL2A (Fig. 18B) compared to binding of NIPBL alone to the 147 bp nucleosomes.

Overall, EMSAs cannot answer the question if NIPBL forms a complex with cohesin to load or stabilize it on nucleosome-bound DNA, or if it binds to the 147 bp nucleosome alone, without incorporating cohesin in the complex.

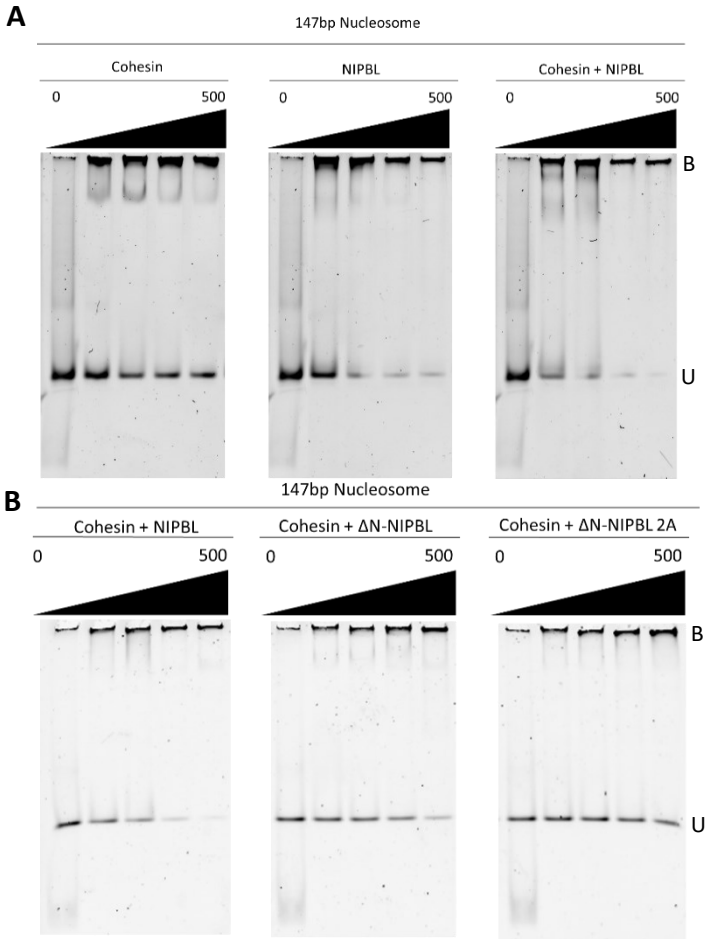

Figure 18 | **Binding of Cohesin and NIPBL to 147 bp nucleosomes. A** | EMSAs comparing the binding of cohesin only, NIPBL only and cohesin + NIPBL to 147 bp nucleosomes. **B** | EMSAs performed with cohesin and different variants of NIPBL. Cohesin was added in concentrations ranging from 32 – 250 nM and NIPBL was titrated from 62,5 to 500 nM, U = unbound nucleosome, B = bound nucleosome.

3.4. Mass Photometry

In this study MP was used as a quality control for protein purifications and to observe binding of proteins to DNA and nucleosome-bound DNA. MP measurements were performed on unmodified glass-slides unless stated otherwise.

3.4.1. Mass Photometry analysis of single components

MP analysis of tetrameric cohesin showed a broad peak at approximately 550 kDa, which fits well with the expected size of 540 kDa (Fig. 19A). Another peak at ~ 140 kDa appears in all measurements of recombinant proteins and is not yet identified, although a possible candidate could be anti-FLAG-tag antibody that was pulled down during the purification process. Peaks on the negative axis correspond to unbinding events.

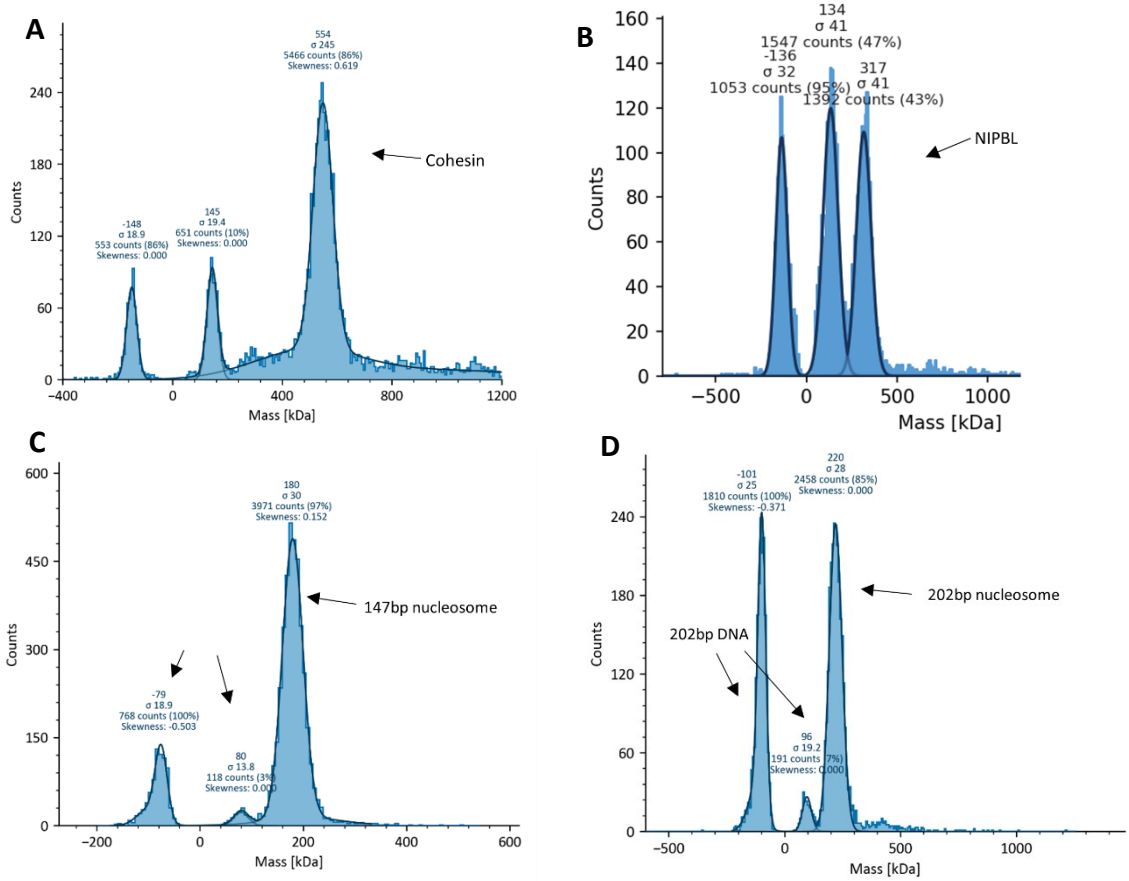

Figure 19 | **Mass Photometry of cohesin, NIPBL and nucleosomes. A** | MP analysis of cohesin added to a final concentration of 20 nM. Unidentified peaks can be observed at 145 and -145 kDa. **B** | MP analysis of NIPBL at a final concentration of 20 nM. High levels of unidentified peak at a mass of 134 kDa were recorded. **C** | MP analysis of 147 bp nucleosomes, measurements were performed at a final concentration of 20 nM. **D** | MP analysis of 202 bp nucleosomes, measured at a final concentration of 20 nM. Arrows indicate expected mass size.

The expected mass for NIPBL is ~ 350 kDa, which could be confirmed in MP with a measured mass of 317 kDa (Fig. 19B). The unidentified peak at ~140 kDa was observed in all NIPBL measurement.

MP measurements of the 147 bp nucleosome and 202 bp nucleosome result in peaks representing the assembled nucleosome cores at masses of 180 kDa and 220 kDa respectively (Fig. 19C,D). This data fits well with the expected size for assembled nucleosome cores, considering that histone octamers have a size of approximately 110 kDa. Additionally, a peak corresponding to free 147 bp and 202 bp DNA which did not get integrated into nucleosomes is visible at 80 kDa and 96 kDa. Free DNA peaks also appear on the negative axis as DNA does not bind well to the unmodified glass slides, hence many unbinding events occur.

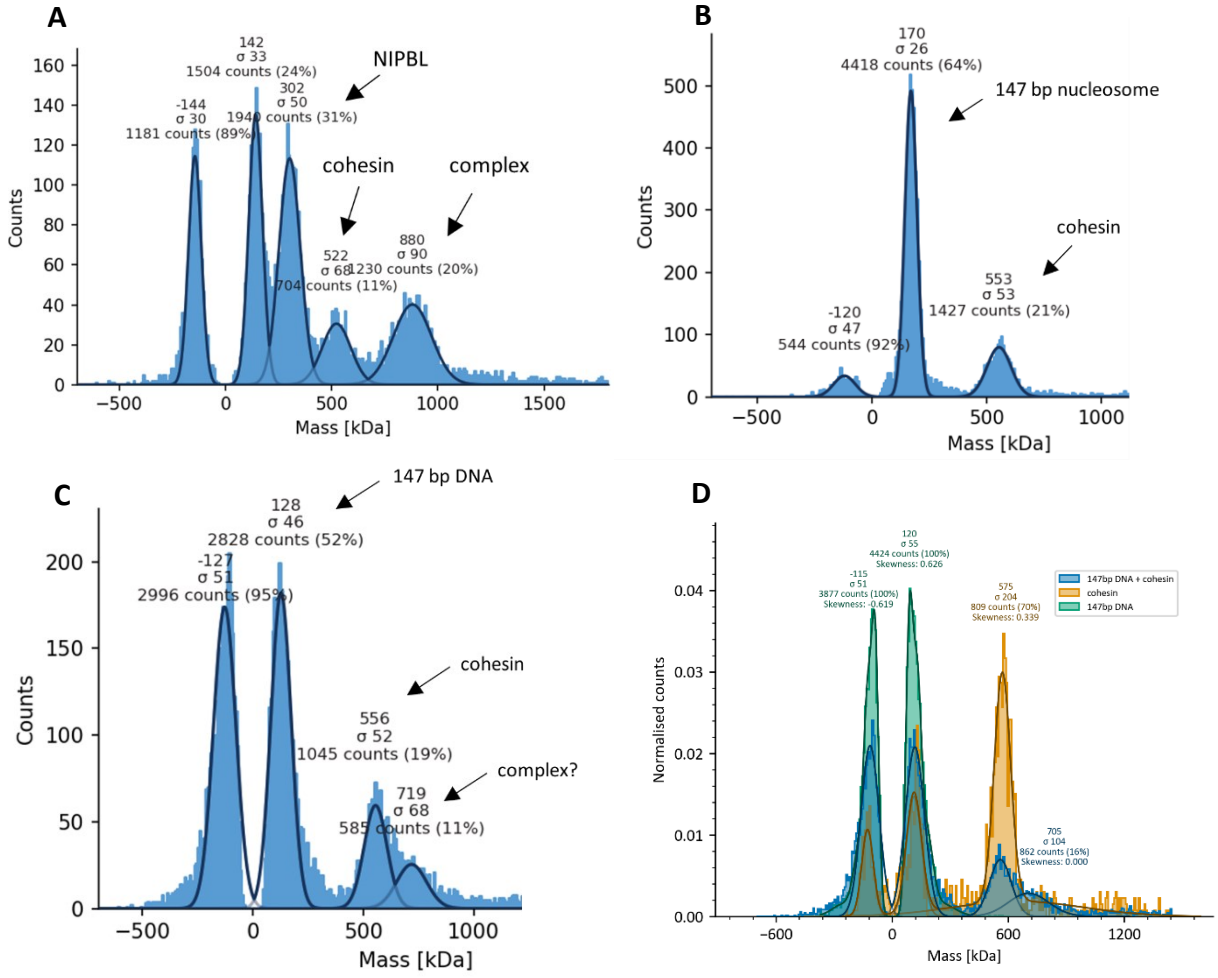

Figure 20 | **Mass Photometry of cohesin-NIPBL and cohesin-147 bp nucleosome/DNA complexes. A** | MP analysis of complex formation between cohesin and NIPBL. NIPBL was added to cohesin in a twofold molar excess. **B** | MP analysis of cohesin and 147 bp nucleosome samples. Nucleosomes were added to cohesin in a twofold molar excess. **C** | MP analysis of complex formation between 147 bp DNA and cohesin. DNA was added to cohesin in a fivefold molar excess. **D** | Merged data of measurements analysing the binding of cohesin to 147 bp DNA. Arrows indicate expected masses.

3.4.2. Mass Photometry analysis of cohesin-NIPBL-nucleosome complexes

Since EMSAs were not suitable to answer if cohesin and NIPBL form a complex on nucleosomes together, MP analysis was used as an alternative approach. First, complex formation of NIPBL and cohesin was analysed by MP to control that NIPBL and cohesin alone can interact with each other. When NIPBL was added to cohesin in a ratio of 2:1 a corresponding complex peak with a mass of 880 kDa could be observed, which fits well with the expected size of 890 kDa and shows that MP is suitable to detect the formation of pentameric cohesin complexes (Fig. 20A).

In a next step, binding of cohesin to 147 bp nucleosomes in the presence or absence of NIPBL was measured using the MP (Fig. 20, Fig. 21). In the absence of NIPBL the peak for nucleosomes is visible at 170 kDa and the peak for cohesin appears at 553 kDa (Fig. 20B). However, no further peak indicating a complex can be observed. When observing binding of cohesin to 147 bp DNA (Fig. 20C) a small peak appears at 719 kDa, which is also visible when

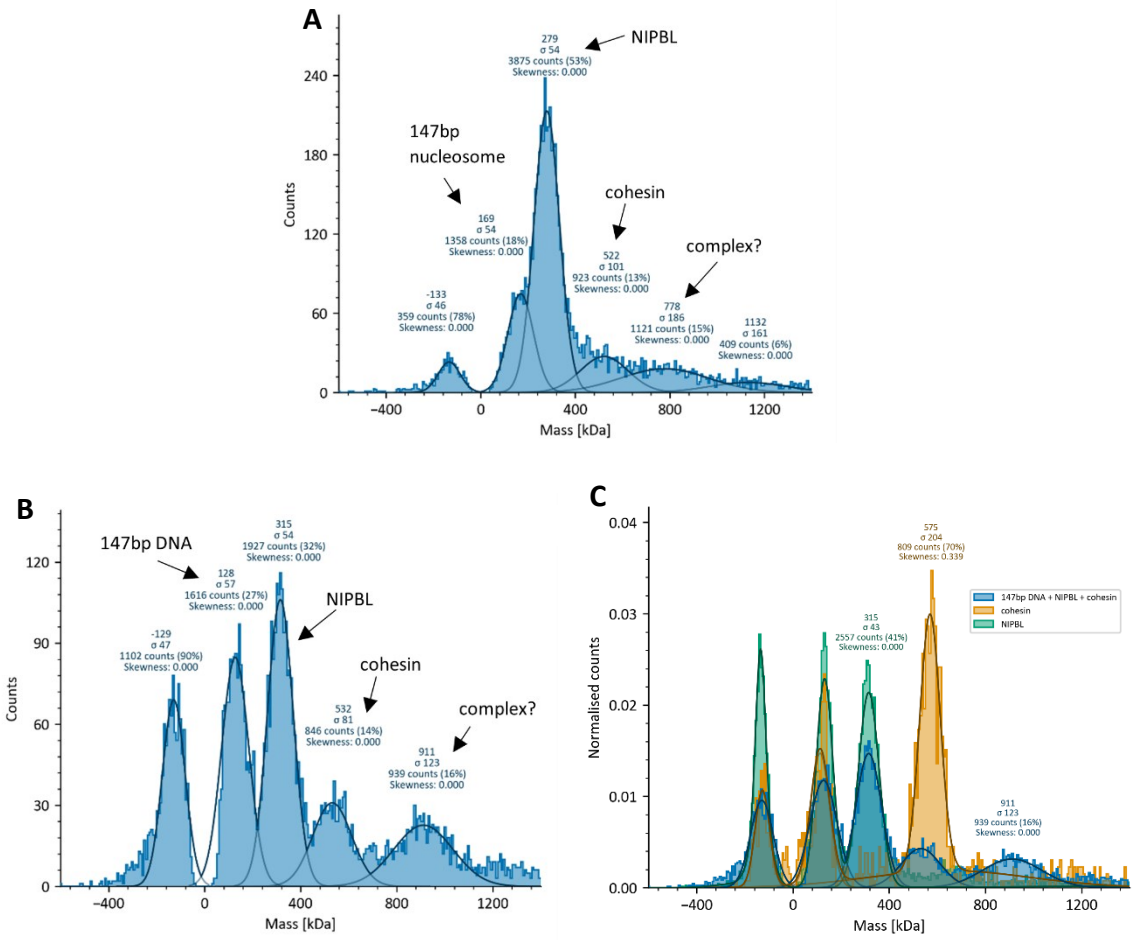

Figure 21 | **MP of cohesin-NIPBL-nucleosome complex formation. A** | MP analysis of cohesin-NIPBL-147 bp nucleosome complex formation. Proteins were added to nucleosomes in a 1.5-fold (cohesin) or twofold (NIPBL) molar excess. **B** | MP analysis of cohesin-NIPBL-DNA complex formation. DNA was added in a fourfold-excess over cohesin, cohesin to NIPBL ratio was 1:2. **C** | Merged data of cohesin-NIPBL-DNA complex formation. Arrows indicate expected mass size.

merging single measurements of cohesin only, DNA only and the complex measurement (Fig. 20D). However, the expected size of a complex between cohesin and 147 bp DNA is with ~ 640 kDa smaller than the observed peak and it cannot be excluded that this peak that makes up only 11% of the sample appears due to background noise.

When NIPBL was added to 147 bp nucleosomes and cohesin no complex formation could be observed (Fig. 21A). The additional peak at 782 kDa does not fit with the expected complex size of ~ 1080 kDa. This peak may represent a small fraction of NIPBL and cohesin complex formation without incorporating the 147 bp nucleosome. When DNA was added to NIPBL and cohesin a small peak appeared at 912 kDa (Fig. 21B), which could fit with the expected complex size of 990 kDa. To confirm whether this peak represents a complex of DNA, cohesin and NIPBL the single measurements were merged (Fig. 21C). Here, it is visible that while the peak only seems to appear when all three components are present, the peak only represents 16% of the sample, and it is not clear if it appears due to background noise

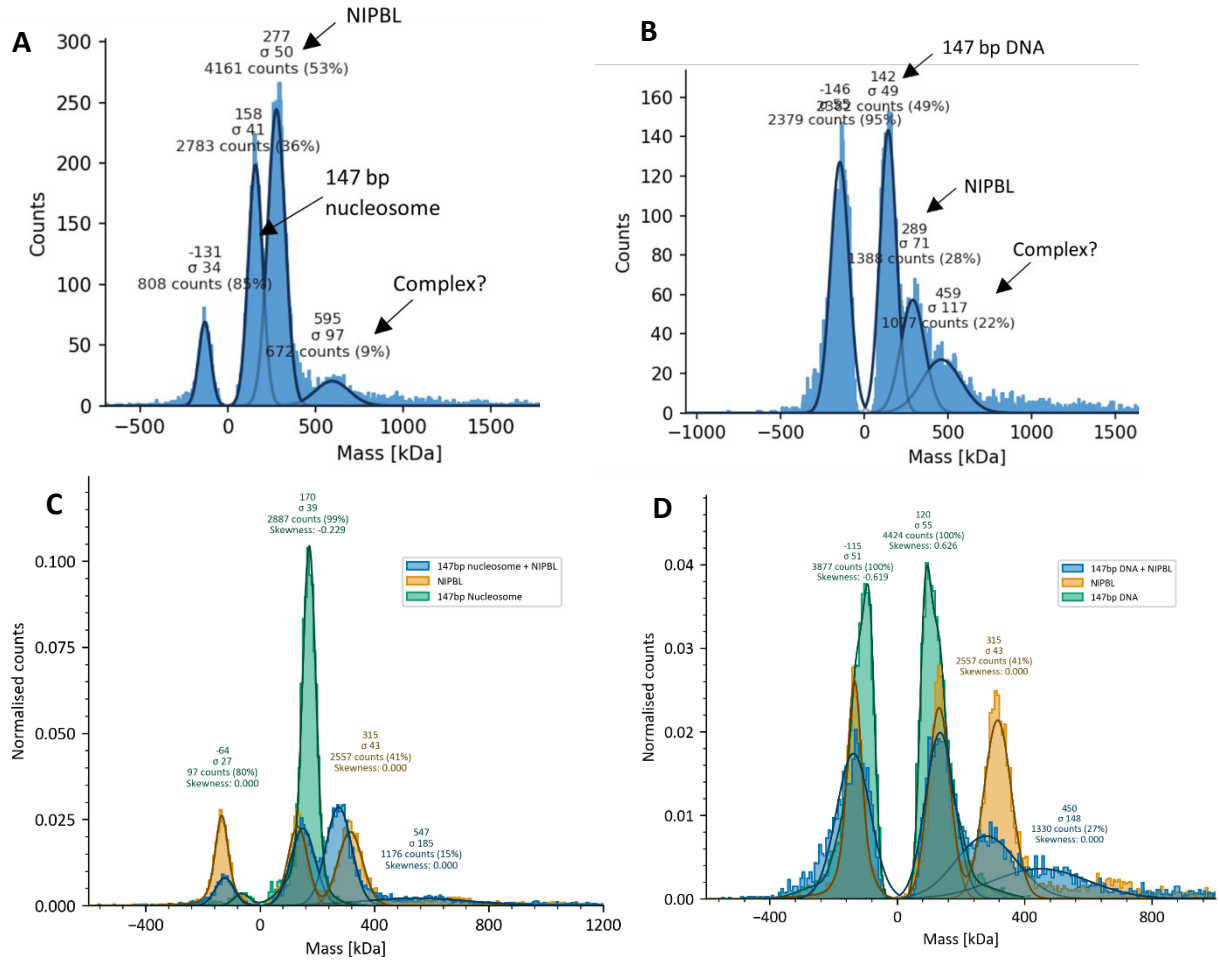

Figure 22 | **Mass Photometry of NIPBL and 147 bp nucleosomes/DNA. A** | MP analysis of 147 bp nucleosomes and NIPBL. NIPBL was added in a twofold molar excess over nucleosomes. **B** | MP analysis of NIPBL and 147 bp DNA. DNA was added in a fivefold molar excess. **C** | Merged MP measurements of 147 bp nucleosomes, NIPBL and the mixed reaction. **D** | Merged Data of 147 bp DNA and NIPBL measurements. Arrows indicate expected mass size.

Interaction of NIPBL alone with 147 bp nucleosomes could also not be observed with MP (Fig. 22A). Although a small peak appeared at 595 kDa, and this peak would fit approximately with the expected size of a NIPBL-nucleosome complex of 550 kDa, when merging the single measurements, it appears that this peak may rather be background noise from the NIPBL sample (Fig. 22C). Similarly, no clear complex peak could be observed when analysing complex formation of NIPBL and 147 bp DNA (Fig. 22B,D). As complexes of NIPBL and nucleosomes, as well as DNA, could clearly be observed with EMSAs, it may be that MP is not ideal for analysing NIPBL-cohesin-nucleosome complex formation.

To confirm this, MP was used to analyse complex formation of 202 bp nucleosomes and cohesin (Fig. 23A), since strong binding of the two molecules was observed in EMSAs. Measuring cohesin binding to 202 bp nucleosomes resulted in two major peaks: 529 kDa corresponding to cohesin and 202 kDa for nucleosomes. Free DNA is visible at -105 kDa. An additional peak with a low abundance appears at 840 kDa. The expected complex size of cohesin and 202 bp nucleosomes would be ~ 750 kDa, thus this peak is too large and is probably caused by background noise. Similarly, when measuring cohesin and 202 bp DNA (Fig. 23B) no complex formation can be observed. These results stand in contrast to the expectation that cohesin would be able to bind free DNA and protruding linker-DNA of nucleosomes and to results previously observed with EMSAs.

Taken together, MP cannot reproduce the results seen previously with EMSAs. However, since binding of cohesin cannot be observed to free DNA as well as to 202 bp nucleosomes, to which cohesin should be able to bind, no conclusion can be made about whether cohesin interacts

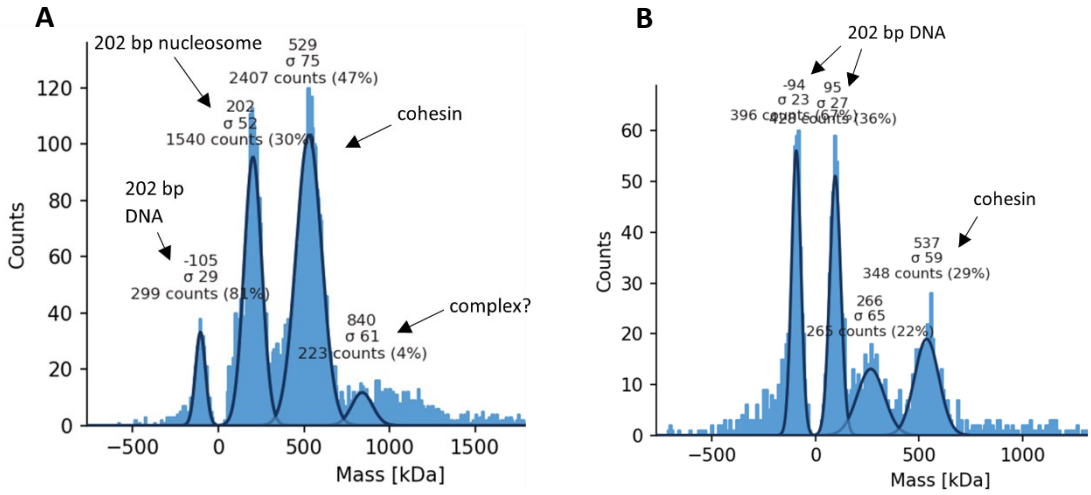

Figure 23| **Mass Photometry of cohesin and 202bp nucleosomes/DNA. A** | MP analysis of 202 bp nucleosome and cohesin. Nucleosome was added in a twofold molar excess over cohesin. **B** | MP analysis of 202 bp DNA and cohesin. DNA was added in a fivefold molar excess. Arrows indicate expected mass size.

with nucleosome-bound DNA. Additionally, because MP is limited in the range of protein concentration and binding affinity it may not be the ideal method to observe binding of cohesin and NIPBL to nucleosome-bound DNA, especially considering that in EMSAs, proteins were added up to 50-fold in excess to observe binding.

3.4.3. Mass Photometry using APTES modified glass slides

An additional problem appears when using MP to measure DNA, as DNA does not bind well to unmodified glass slides. Therefore, glass slides were modified using APTES and binding affinities were observed on modified slides. However, while the unbinding events of DNA decreased, still no complex formation of cohesin or NIPBL to nucleosomes and DNA could be observed, as shown here at the example 147 bp nucleosome/DNA and cohesin (Fig. 24).

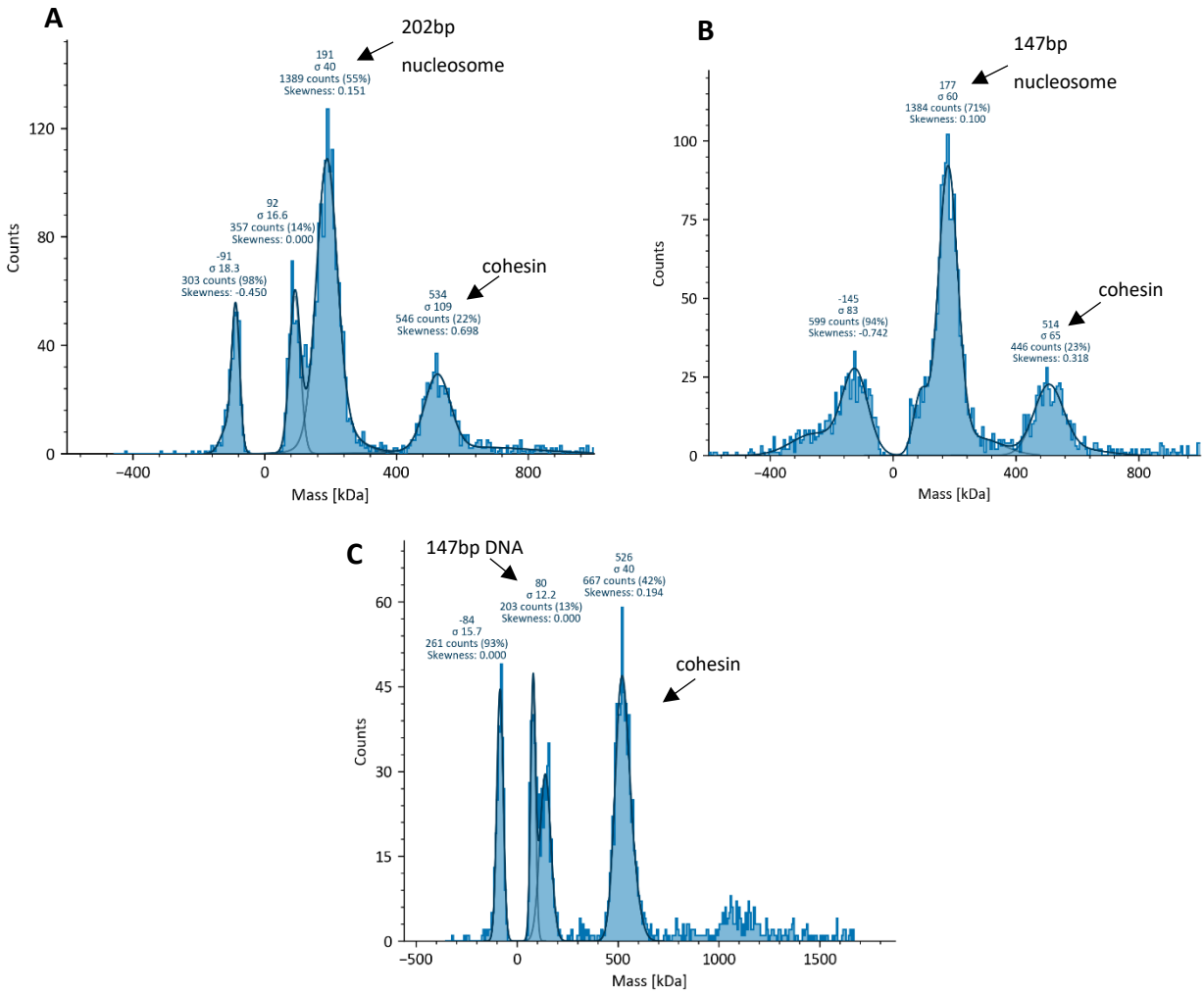

Figure 24 | **Mass Photometry on APTES modified slides. A** | MP analysis of 202 bp nucleosome and cohesin, cohesin was added in a twofold molar excess. **B** | MP analysis of 147 bp nucleosome and cohesin complex formation, cohesin was added in a twofold molar excess. **C** | MP of 147 bp DNA and cohesin, cohesin was added in a twofold molar excess. Arrows indicate expected mass size.

3.5. HS-AFM

In this study HS-AFM was used to observe assembled nucleosomes and their interaction with cohesin.

HS-AFM images of 147 bp nucleosomes revealed homogenously assembled nucleosomes and no visible DNA was protruding from the nucleosome core (Fig. 25A). Sometimes free DNA was visible in the sample. Contrary, free linker-DNA is clearly protruding from 202 bp nucleosomes (Fig. 25B).

When cohesin was added to nucleosomes four round structures could sometimes be observed, two of which would be cohesin's ATPase heads, one could represent STAG1, and one may represent a bound nucleosome (Fig. 25C). However, no conclusion can be made due to the low resolution and similar size of cohesin's ATPase heads, STAG1 and nucleosomes. As NIPBL has a similar phenotype in HS-AFM images as the ATPase heads and nucleosomes, it would also not be distinguishable and was therefore not added to the reaction.

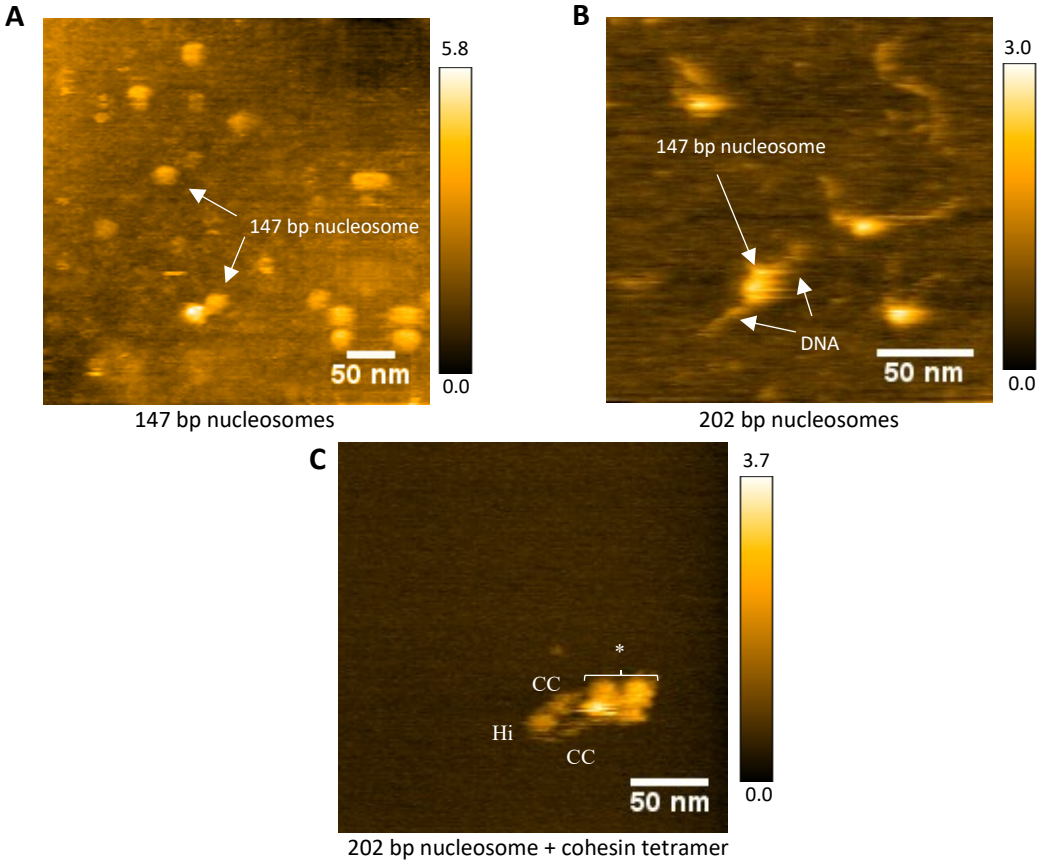

Figure 25 | **HS-AFM of 147 bp and 202 bp nucleosomes. A** | HS-AFM image of 147 bp nucleosomes on poly-L-lysine modified mica. No DNA can be seen protruding from the nucleosome core. **B** | HS-AFM image of 202 bp nucleosome. Free linker-DNA protrudes from the nucleosome core on both sides. **C** | HS-AFM image of 202 bp nucleosome and cohesin tetramer. Measurements with the HS-AFM software Kodec resulted in widths of both cohesin ATPase heads and nucleosome cores of ~18 nm. Hi= Hinge, CC= coiled coils, *= in this area cohesin's ATPase heads are bound to STAG1. Additionally, because four round structures can be seen it could be that a nucleosome is bound as well. The scale bar indicates recorded sample height in nm.

4. Discussion

In this study the binding of cohesin and its HAWK protein NIPBL to nucleosome-bound DNA was investigated. Cohesin, a DNA binding ATPase, belongs to the SMC protein family and plays important roles in sister chromatid cohesion and genome organization. To fold DNA into loops and TADs, and thereby perform functions in gene expression, regulation and recombination, cohesin connects DNA in *cis* through a process called loop extrusion. Cohesin mediated loop extrusion has been observed *in vitro* but how cohesin performs this function *in vivo* remains unknown. The aim of this study was to understand if cohesin can directly bind to and interact with nucleosome-bound DNA, since *in vivo* DNA doesn't appear naked but is covered and compacted by proteins, mainly with nucleosomes. Therefore, cohesin must either be able to act on nucleosome-bound DNA or nucleosomes need to be evicted or displaced before cohesin can perform its function. To address this, recombinant cohesin and NIPBL proteins were expressed in Sf9 cells and purified. Sf9 histones were used to reconstitute two types of nucleosomes containing either 147 bp of DNA, with no DNA protruding from the nucleosome core, or 202 bp of DNA, with 27 bp of DNA protruding from the nucleosome core as linker-DNA.

EMSAs showed that while cohesin has low affinity for nucleosome-bound DNA the binding efficiency increases strongly when free linker-DNA is present. These results are in line with previously published data showing that the SMC protein condensin binds to nucleosomal DNA with low affinity and the binding is drastically increased when free DNA is protruding from the nucleosome [45].

The EQ mutant seemed to be able to associate with nucleosome-bound DNA in the presence of ATP more stably, possibly because it is locked in the clamped state [18][34]. This indicates that cohesin can indeed interact with nucleosome-bound DNA with low efficiency. However, it is unlikely that this low efficiency allows cohesin to bind to and perform its functions on chromatin *in vivo*. Rather, cohesin's affinity to protruding free DNA indicates that nucleosomes may need to be evicted or displaced from DNA, for example by chromatin remodeller complexes, as was shown before in yeast [46][47][48]. Another possibility of how cohesin may extrude loops in the presence of nucleosomes is by "stepping over" the nucleosomes, as it was suggested for condensin [45]. Additionally, cohesin may need other unidentified factors to allow for sufficient binding.

EMSAs performed with NIPBL and nucleosomes indicated that NIPBL can bind to free linker-DNA and, surprisingly, to nucleosome-bound DNA with similar efficiencies. Its binding partner MAU2 seemed to be dispensable for nucleosomal DNA binding. However, studies in yeast suggested that the MAU2 homolog Scc4 is necessary for Scc2 recruitment to chromatin *in vivo* [56]. EMSAs also indicated that the N-terminal domain of NIPBL may be crucial for effective binding to nucleosomes, as the ΔN-NIPBL variant had reduced binding efficiencies. This data may fit to previously published observations that the C-terminal domain of NIPBL yeast homolog Scc2 is sufficient to catalyse loading of cohesin onto naked DNA *in vitro* but not onto chromosomes *in vivo* [47][56]. Additionally, EMSAs performed with the DNA-binding mutant ΔN-NIPBL2A suggested that NIPBL does indeed interact with DNA and not with core histones directly. However, EMSAs could not answer the question if NIPBL and cohesin bind 147 bp nucleosomes in a complex and if NIPBL can load cohesin onto nucleosome-bound DNA, because proteins did not travel through the gel. Therefore, it is not clear what exactly NIPBL's function in nucleosome binding is and further experiments need to be performed to answer these open questions.

Next, mass photometry measurements were performed to control the quality of purified proteins and assembled nucleosomes as well as to measure binding of cohesin and NIPBL to nucleosomes. No binding of cohesin to 147 bp and 202 bp nucleosomes could be detected and similarly, cohesin did not form complexes with free DNA in MP measurements. Additionally, MP was used to analyse complex formation of NIPBL and nucleosomes/DNA. While sometimes small peaks of the expected complex size could be detected, when merging single measurements of NIPBL and nucleosomes, it was not clear if the peak can be attributed to actual complex formation or simply to background noise. In single cases small peaks corresponding to the expected size of NIPBL-DNA complexes could be observed, but again merged data suggested that this could possibly be background noise.

One possible problem in MP measurements could have been that DNA did not bind well to unmodified glass-slides [54]. Therefore, glass-slides were modified with APTES to charge the glass surface positively, but still no complex formation could be observed. Additionally, nucleosomes bound well to unmodified glass-slides, thus it is unlikely that surface characteristics of the glass slide were a problem for detecting complex formation between cohesin/NIPBL and nucleosomes.

SMC complexes are known for their ability to bind naked DNA [25][45], which was confirmed by EMSAs performed in this study. However, since it was not possible to detect complexes of cohesin and free DNA using MP, no conclusion can be made about cohesin's binding affinity to nucleosome-bound DNA and free linker-DNA from MP measurements.

A reason why complex formation couldn't be detected, might be that MP is limited in the range of protein concentration that can be used and binding-affinity that can be observed [50]. While in EMSAs proteins were added up to a 50-fold excess over nucleosomes/DNA, in MP adding proteins in such an excess would lead to too many counts and unusable data. Therefore, proteins or nucleosomes/DNA were only added in a 1- or 2-fold excess. For strong binding-affinities this is sufficient, as could be seen by the complex formation of NIPBL and cohesin, where NIPBL was added in a ratio of 2:1. Overall, MP may not be the ideal method to study binding of cohesin and NIPBL to nucleosomes, as the binding-affinity of these complexes may be too low.

Lastly, HS-AFM was used to visualize assembled nucleosomes and to try and study their interaction with cohesin. HS-AFM showed homogenously assembled 147 bp nucleosomes with no protruding DNA. Furthermore, when imaging 202 bp nucleosomes it was clearly visible that DNA was protruding from both sides of the nucleosome-core, suggesting that the nucleosome assembles with high efficiency on the 601 widom sequence that was placed in the middle of the 202 bp DNA. However, trying to visualize cohesin's interaction with nucleosomes proved to be difficult as nucleosomes appeared to be similar in size to cohesin's ATPase heads. When measuring the approximate width of nucleosomes and cohesin's ATPase heads it confirmed that both molecules have a width of ~ 18 nm. Because of this and because of the limited resolution that was achieved in HS-AFM it was not possible to distinguish between cohesin's heads and nucleosomes and no conclusions could be made.

Taken together, the results of this study suggest that cohesin may need to have access to free DNA to be able to perform its function on chromatin. Additionally, an interesting observation was made in NIPBL's ability to interact with nucleosome-bound DNA when using EMSAs, but if NIPBL can also bind to nucleosomes *in vivo* and what exactly the underlying function is needs to be investigated in future experiments. Furthermore, these results should be replicated with different methods as EMSAs have many limitations, starting with the fact that complex characterization is not possible, and MP and HS-AFM seem to be less than optimal methods for studying the interactions of cohesin and NIPBL with nucleosomes. Experiments that could

be performed in the future include pull-down assays or density gradients to observe if cohesin and NIPBL form a complex with nucleosomes. Additionally, one could vary the length of free linker-DNA to study how long protruding DNA needs to be for cohesin to bind efficiently.

5. References

[1] Cooper GM. The Cell: A Molecular Approach. 2nd edition. Sunderland (MA): Sinauer Associates; 2000. The Eukaryotic Cell Cycle. Available from: https://www.ncbi.nlm.nih.gov/books/NBK9876/

[2] Campbell Neil A., Reece Jane B., Markl Jürgen. Biology. 6th edition. Pearson Studium.

[3] Zheng, H. & Xie, W. (2019). The role of 3D genome organization in development and cell differentiation. Nature reviews Molecular Cell Biology, 20, 535-550. https://doi.org/10.1038/s41580-019-0132-4

https://doi.org/10.1038/s41580-019-0132-4

[4] Piovesan, A., Pelleri, M. C., Antonaros, F., Strippoli, P., Caracausi, M., & Vitale, L. (2019). On the length, weight and GC content of the human genome. BMC research notes, 12(1), 106. https://doi.org/10.1186/s13104-019-4137-z

[5] Lammerding J. (2011). Mechanics of the nucleus. Comprehensive Physiology, 1(2), 783–807. https://doi.org/10.1002/cphy.c100038

[6] Kornberg, R.D. (1974). Chromatin structure: a repeating unit of histones and DNA. Science, 184(4139), 868–871. https://doi.org/10.1126/science.184.4139.868

[7] Davidson, I. F., & Peters, J. M. (2021). Genome folding through loop extrusion by SMC complexes. Nature reviews. Molecular cell biology, 22(7), 445–464.

[8] Rao, S. S. et al. (2014). A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell, 159(7), 1665–1680. https://doi.org/10.1016/j.cell.2014.11.021

[9] Jung, I. et al. (2019). A compendium of promoter-centered long-range chromatin interactions in the human genome. Nature Genetics, 51(10), 1442–1449. https://doi.org/10.1038/s41588-019-0494-8

[10] Davidson, I. F. et al. (2019). DNA loop extrusion by human cohesin. Science, 366(6471), 1338–1345. https://doi.org/10.1126/science.aaz3418

[11] Uhlmann F. (2016). SMC complexes: from DNA to chromosomes. Nature reviews. Molecular cell biology, 17(7), 399–412. https://doi.org/10.1038/nrm.2016.30

[12] Hirano T. (2016). Condensin-Based Chromosome Organization from Bacteria to Vertebrates. Cell, 164(5), 847–857. https://doi.org/10.1016/j.cell.2016.01.033

[13] Losada A. (2014). Cohesin in cancer: chromosome segregation and beyond. Nature reviews. Cancer, 14(6), 389–393. https://doi.org/10.1038/nrc3743

[14] Taylor, E. M. et al. (2001). Characterization of a novel human SMC heterodimer homologous to the Schizosaccharomyces pombe Rad18/Spr18 complex. Molecular biology of the cell, 12(6), 1583–1594. https://doi.org/10.1091/mbc.12.6.1583

[15] Hirano T. (2005). SMC proteins and chromosome mechanics: from bacteria to humans. Philosophical transactions of the Royal Society of London. Series B, Biological sciences, 360(1455), 507–514. https://doi.org/10.1098/rstb.2004.1606

[16] Knust, T., & Graumann, P. L. (2020). Bacillus subtilis Smc condenses chromosomes in a heterologous cell system, which is down-regulated by ScpAB. BMC research notes, 13(1), 524. https://doi.org/10.1186/s13104-020-05344-3

[17] Hiraga S. (2000). Dynamic localization of bacterial and plasmid chromosomes. Annual review of genetics, 34, 21–59. https://doi.org/10.1146/annurev.genet.34.1.21

[18] Shi, Z. et al. (2020). Cryo-EM structure of the human cohesin-NIPBL-DNA complex. Science, 368(6498), 1454–1459. https://doi.org/10.1126/science.abb0981

[19] Michaelis, C. et al. (1997). Cohesins: chromosomal proteins that prevent premature separation of sister chromatids. Cell, 91(1), 35-45.https://doi.org/10.1016/s0092-8674(01)80007-6

[20] Hearing, C.H. et al. (2008). The cohesin ring concatenates sister DNA molecules. Nature, 454(7202), 297-301. https://doi.org/10.1038/nature07098

[21] Riggs, A. D. (1990). DNA methylation and late replication probably aid cell memory, and type I DNA reeling could aid chromosome folding and enhancer function. Philosophical transactions of the Royal Society of London. Series B, Biological sciences, 326(1235), 285–297. https://doi.org/10.1098/rstb.1990.0012t

[22] Hirano T. (2006). At the heart of the chromosome: SMC proteins in action. Nature reviews. Molecular cell biology, 7(5), 311–322. https://doi.org/10.1038/nrm1909

[23] Terakawa, T. et al. (2017). The condensin complex is a mechanochemical motor that translocates along DNA. Science, 358(6363), 672–676. https://doi.org/10.1126/science.aan6516

[24] Ganji, M. et al. (2018). Real-time imaging of DNA loop extrusion by condensin. Science, 360(6384), 102–105. https://doi.org/10.1126/science.aar7831

[25] Kim, Y. et al. (2019). Human cohesin compacts DNA by loop extrusion. Science, 366(6471), 1345–1349. https://doi.org/10.1126/science.aaz4475

[26] Ciosk, R. et al. (2000). Cohesin's binding to chromosomes depends on a separate complex consisting of Scc2 and Scc4 proteins. Molecular Cell, 5(2), 243-254. https://doi.org/10.1016/s1097-2765(00)80420-7

[27] Srinivasan, M. et al. (2018). The Cohesin Ring Uses Its Hinge to Organize DNA Using Non-topological as well as Topological Mechanisms. Cell, 173(6), 1508–1519.e18. https://doi.org/10.1016/j.cell.2018.04.015

[28] Fudenberg, G. et al. (2017). Emerging Evidence of Chromosome Folding by Loop Extrusion. Cold Spring Harbor symposia on quantitative biology, 82, 45–55. https://doi.org/10.1101/sqb.2017.82.034710

[29] Sweeney, H. L., & Holzbaur, E. (2018). Motor Proteins. Cold Spring Harbor perspectives in biology, 10(5), a021931. https://doi.org/10.1101/cshperspect.a021931

[30] Nichols, M. H., & Corces, V. G. (2018). A tethered-inchworm model of SMC DNA translocation. Nature structural & molecular biology, 25(10), 906–910. https://doi.org/10.1038/s41594-018-0135-4

[31] Marko, J. F., De Los Rios, P., Barducci, A., & Gruber, S. (2019). DNA-segment-capture model for loop extrusion by structural maintenance of chromosome (SMC) protein complexes. Nucleic acids research, 47(13), 6956–6972. https://doi.org/10.1093/nar/gkz497

[32] 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. Nature structural & molecular biology, 26(3), 227–236. https://doi.org/10.1038/s41594-019-0196-z

[33] Lee, B. G. et al. (2020). Cryo-EM structures of holo condensin reveal a subunit flip-flop mechanism. Nature structural & molecular biology, 27(8), 743–751. https://doi.org/10.1038/s41594-020-0457-x

[34] Bauer, B. W. et al. (2021). Cohesin mediates DNA loop extrusion by a "swing and clamp" mechanism. Cell, 184(21), 5448–5464.e22. https://doi.org/10.1016/j.cell.2021.09.016

[35] Krantz, I. D. et al. (2004). Cornelia de Lange syndrome is caused by mutations in NIPBL, the human homolog of Drosophila melanogaster Nipped-B. Nature genetics, 36(6), 631–635. https://doi.org/10.1038/ng1364

[36] Davey C.A. et al. (2002). Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 a resolution. Journal of molecular biology, 319(5), 1097-1113. https://doi.org/10.1016/S0022-2836(02)00386-8

[37] Li, G., & Reinberg, D. (2011). Chromatin higher-order structures and gene regulation. Current opinion in genetics & development, 21(2), 175–186. https://doi.org/10.1016/j.gde.2011.01.022

[38] Szerlong, H. J., & Hansen, J. C. (2011). Nucleosome distribution and linker DNA: connecting nuclear function to dynamic chromatin structure. Biochemistry and cell biology, 89(1), 24–34. https://doi.org/10.1139/O10-139

[39] Luger, K. et al. (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution. Nature, 389(6648), 251–260. https://doi.org/10.1038/38444

[40] Chung, S. Y., Hill, W. E., & Doty, P. (1978). Characterization of the histone core complex. Proceedings of the National Academy of Sciences of the United States of America, 75(4), 1680–1684. https://doi.org/10.1073/pnas.75.4.1680

[41] Grant P. A. (2001). A tale of histone modifications. Genome biology, 2(4), REVIEWS0003. https://doi.org/10.1186/gb-2001-2-4-reviews0003

[42] Smith, S. & Stillman, B. (1991). Stepwise assembly of chromain during DNA replication in vitro. The EMBO Journal, 10(4), 971–980.

[43] Lowary, P. T., & Widom, J. (1998). New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. Journal of molecular biology, 276(1), 19–42. https://doi.org/10.1006/jmbi.1997.1494

[44] Grigoryev S. A. (2012). Nucleosome spacing and chromatin higher-order folding. Nucleus, 3(6), 493–499. https://doi.org/10.4161/nucl.22168

[45] Kong, M. et al. (2020). Human Condensin I and II Drive Extensive ATP-Dependent Compaction of Nucleosome-Bound DNA. Molecular cell, 79(1), 99–114.e9. https://doi.org/10.1016/j.molcel.2020.04.026

[46] Huang, J., Hsu, J. M., & Laurent, B. C. (2004). The RSC nucleosome-remodeling complex is required for Cohesin's association with chromosome arms. Molecular cell, 13(5), 739–750. https://doi.org/10.1016/s1097-2765(04)00103-0

[47] Muñoz, S. et al. (2019). A Role for Chromatin Remodeling in Cohesin Loading onto Chromosomes. Molecular cell, 74(4), 664–673.e5. https://doi.org/10.1016/j.molcel.2019.02.027

[48] Garcia-Luis, J. et al. (2019). FACT mediates cohesin function on chromatin. Nature structural & molecular biology, 26(10), 970–979. https://doi.org/10.1038/s41594-019-0307-x

[49] Hellman, L. M., & Fried, M. G. (2007). Electrophoretic mobility shift assay (EMSA) for detecting protein-nucleic acid interactions. Nature protocols, 2(8), 1849–1861. https://doi.org/10.1038/nprot.2007.249

[50] Soltermann, F. et al. (2020). Quantifying Protein-Protein Interactions by Molecular Counting with Mass Photometry. Angewandte Chemie, 59(27), 10774–10779. https://doi.org/10.1002/anie.202001578

[51] Ando T. (2019). High-speed atomic force microscopy. Current opinion in chemical biology, 51, 105–112. https://doi.org/10.1016/j.cbpa.2019.05.010

[52] Shim, Y. et al. (2012). Polycistronic coexpression and nondenaturing purification of histone octamers. Analytical biochemistry, 427(2), 190–192. https://doi.org/10.1016/j.ab.2012.05.006

[53] Weissmann, F. et al. (2016). biGBac enables rapid gene assembly for the expression of large multisubunit protein complexes. Proceedings of the National Academy of Sciences of the United States of America, 113(19), E2564–E2569. https://doi.org/10.1073/pnas.1604935113

[54] Li, Y., Struwe, W. B., & Kukura, P. (2020). Single molecule mass photometry of nucleic acids. Nucleic acids research, 48(17), e97. https://doi.org/10.1093/nar/gkaa632

[55] Parenti, I. et al. (2020). MAU2 and NIPBL Variants Impair the Heterodimerization of the Cohesin Loader Subunits and Cause Cornelia de Lange Syndrome. Cell reports, 31(7), 107647. https://doi.org/10.1016/j.celrep.2020.107647

[56] Chao, W. C. et al. (2015). Structural Studies Reveal the Functional Modularity of the Scc2-Scc4 Cohesin Loader. Cell reports, 12(5), 719–725. https://doi.org/10.1016/j.celrep.2015.06.071

6. List of tables

7. List of Figures