

DISSERTATION

STRUCTURAL AND FUNCTIONAL INSIGHT INTO KINETOCHORE PROTEIN CENP-
N AND ITS INTERACTION WITH CENP-A NUCLEOSOME

Submitted by

Keda Zhou

Department of Biochemistry and Molecular Biology

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Spring 2018

Doctoral Committee:

Advisor: Karolin Luger

Jennifer Deluca

Susan Bailey

Tingting Yao

Copyright by Keda Zhou 2018

All Rights Reserved

ABSTRACT

STRUCTURAL AND FUNCTIONAL INSIGHT INTO KINETOCHORE PROTEIN CENP-N AND ITS INTERACTION WITH CENP-A NUCLEOSOME

Proper chromosome segregation during mitosis is one of the most important processes to ensure genome integrity. During this process, the microtubules are captured by a multi-unit complex called kinetochore. The kinetochore is assembled specifically at centromere through recognizing nucleosomes containing the histone H3 variant CENP-A. CENP-N and CENP-C are the only two kinetochore proteins that specifically recognize CENP-A nucleosomes. There are about 1 in 25 nucleosomes that contain CENP-A at the centromere. Therefore, how these two proteins ‘ignore’ the abundant H3 nucleosomes to interact selectively with a handful of centromeric CENP-A nucleosomes has important implications for genome stability during cell division. To obtain deep insight into the mechanism behind this, I solved the structure of CENP-A nucleosome in complex with CENP-N by single particle cryo electron microscopy (cryo-EM) at 4 Å. Through charge and space complementarity, the unique “RG” loop on CENP-A is decoded by CENP-N. CENP-N also engages in extensive interactions with a long segment of the distorted nucleosomal DNA double helix. These interactions were validated *in vitro* and *in vivo*. The DNA ends of CENP-A nucleosome which are disordered in the crystal structure are mostly visible in the cryo-EM structure when it is in complex with CENP-N. By micrococcal nuclease digestion assay, the CENP-A nucleosome DNA ends are shown to be less flexible when CENP-N is presented in solution, which is consistent with structural study. Since CENP-N does not interact with

DNA ends directly, the less dynamics on the DNA ends indicate a more stable nucleosome. By quantitative electrophoretic mobility shift assay (EMSA) and electron microscopy, the stabilizing effect of CENP-N on CENP-A nucleosome was confirmed *in vitro*. However, this effect was not significant *in vivo*, which indicates that the CENP-A nucleosome stability *in vivo* is determined by multiple factors. Besides the change on DNA ends of CENP-A nucleosome, the orientation of H4 N-terminal tail is altered due to its interaction with CENP-N, with important implications for the multiple biological processes involving the H4 N-terminal tail, especially with respect to the formation of chromatin higher order structure. The structural and functional studies in this thesis shed light on how CENP-N ensures that the kinetochore assembles specifically at the centromere.

ACKNOWLEDGEMENTS

Six years' pursuit is a long journey. I am very happy to be a "RAM" student in the department of Biochemistry and Molecular Biology. I am so grateful to meet so many kind people at both Colorado state university and University of Colorado. I extremely enjoyed my time working with them.

I am so proud and honored to be the student of Dr. Karolin Luger. She guided me to enter the world of science. I was greatly motivated by her enthusiasm on science. I grew up so much as a scientist under her mentoring. The work ethic that she shared with me will benefit my whole life. Language cannot express my deep gratitude to her. All I can do is to make her proud of me in future.

I want to thank my committee members, Dr. Jennifer Deluca, Dr. Susan Bailey and Dr. Tingting Yao. They gave me a lot of suggestions on my project. Without their strong supports, I cannot make the works here possible.

I am so happy working with the members in Luger's lab. They helped me on every aspect in the lab. I want to thank previous lab member Dr. Mekonnen Lemma Dechassa. He mentored me on my rotation and PhD project. I learnt so much from him. I want to specially thank Elizabeth Kasney and Pamela Dyer. They helped me to make life easier

in the lab. I also want to thank Dr. Francesca Mattioli and Dr. Uma M Muthurajan. They gave me a lot of suggestions and helps on science.

Finally, I would like to thank my parents and my wife. Their love and support encouraged me through my whole journey during past six year and half.

TABLE OF CONTENTS

ABSTRACT.....	ii
ACKNOWLEDGEMENTS	iv
CHAPTER 1 LITERATURE REVIEW.....	1
Nucleosome	1
The structure of the nucleosome	3
The higher order structures of chromatin	6
Histone modifications on histone tails.....	8
Histone variants in nucleosomes	10
The centromeric, CENP-A containing nucleosome	12
The nucleosome containing the H3 variant CENP-A defines the centromere.....	13
Structure and biophysical property of nucleosome containing CENP-A	13
Constitutive Centromere Associated Network (CCAN)	18
The sub-complexes of CCAN	18
Key player at the centromere: CENP-C	20
Dynamic node at the centromere: CENP-N	24
Specific aims	25
CHAPTER 2 THE CYRO-EM STRUCTURE OF CENP-A NUCLEOSOME IN COMPLEX WITH CENP-N.....	28
Overview	28
Introduction	29
Results	34
Crystal structure of CENP-N ¹⁻²³⁵	34
Cryo-EM analysis of the CENP-N:CENP-A nucleosome complex	43
The CENP-N:CENP-A interface	45
Mutational validation of the CENP-N:CENP-ANCP structure.....	46
Identification of a CENP-C region involved in CENP-LN binding	49
CENP-LN binding motif of CENP-C is required for kinetochore recruitment of CENP-N.....	50
Discussion	54
Methods	62
CHAPTER 3 CENP-A nucleosome is stabilized by both CENP-N and CENP-C <i>in vitro</i>	76
Overview	76
Introduction	77

Results	79
CENP-A nucleosome DNA ends are less accessible to micrococcal nuclease when CENP-N interacts with the nucleosome	79
CENP-N increases the stability of the CENP-A nucleosome <i>in vitro</i>	83
CENP-N and CENP-C additively stabilize CENP-A nucleosomes <i>in vitro</i>	88
Loss of CENP-C and CENP-N does not alter CENP-A nucleosome levels in chromatin	91
Loss of CENP-C and/or CENP-N does not alter the salt extraction of CENP-A from centromeric chromatin.....	95
Mutation of the CENP-A L1 loop to reduce CENP-N affinity does not affect CENP-A or CATD chimera maintenance at centromeres.....	98
Discussion	100
Materials and Methods	106
CHAPTER 4 CENP-A nucleosome dynamics and the potential role of CENP-N on higher order structure of centromere	115
Single particle cryo-EM study on CENP-A nucleosome with and without the presence of CENP-C or CENP-N/CENP-C	115
Overview.....	115
Introduction	116
Result	117
The DNA ends in cryo-EM electron density map from the sub-population of stable CENP-A nucleosomes are well organized	117
All stabilized CENP-A nucleosomes adopt same wrapping conformation of DNA ends....	119
Discussion	122
CENP-N ^{NT} induces CENP-A mono-nucleosome oligomerization in vitro	123
Overview.....	123
Introduction	124
Results.....	126
CenpN mediates the oligomerization of CenpA mono-nucleosomes	126
The H4 N-terminal tail is essential for the oligomerization of CENP-A mono-nucleosomes mediated by CENP-N.....	129
Discussion	129
CHAPTER 5 SUMMARY AND FUTURE DIRECTIONS	133
REFERENCES	137
APPENDIX.....	156

CHAPTER 1 LITERATURE REVIEW

Chromosomes are separated equally into two daughter cells during M phase by microtubules, which ensures the integrity of genome. The kinetochore machinery is essential for this process. It consists of more than 100 proteins that connect microtubules to a special region called “centromere” on each chromosome. This “centromere” is defined by nucleosomes containing the histone H3 variant CENP-A. To guarantee proper chromosome segregation, kinetochore proteins need to be specifically recruited to the centromere in a highly defined manner. In humans, the ‘centromere constitutive associate network’ (CCAN), which is formed by 16 inner kinetochore proteins, is responsible for the interaction with the centromere (Przewloka and Glover, 2009). Recognizing CENP-A nucleosomes by CCAN is key for centromeric localization of the kinetochore. Among the 16 CCAN proteins, CENPC and CENPN were identified as the only two proteins in CCAN that directly and preferably interact with CENP-A nucleosome. The goal of this thesis is to understand the molecular basis on how CENP-N recognizes CENP-A nucleosome and the effect of this interaction.

Nucleosome

As genetic material, DNA is compacted and organized into chromatin (or chromosomes, figure 1.1) in all eukaryotes. Under the electron microscope, the chromatin structure

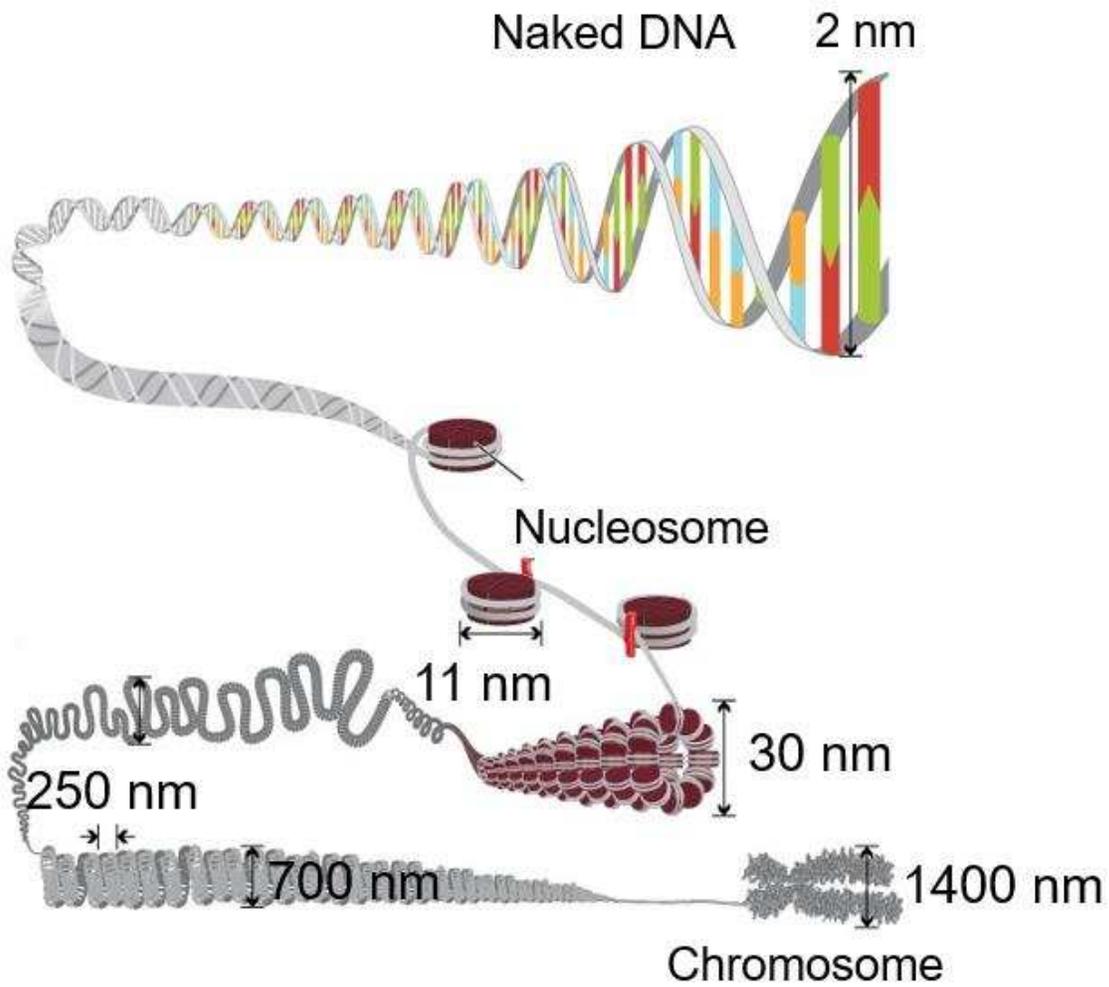


Figure 1.1 The scheme shows how DNA is compacted into chromosome (Annunziato, 2008). The numbers indicate the diameters of the compaction units.

appears like “beads on a string” (Woodcock et al., 1976). This ‘bead’ is the basic unit of chromatin termed “nucleosome” (Kornberg, 1974). In a canonical nucleosome, histones (H3, H4, H2A and H2B) are wrapped and organized by DNA. During past decades, much progress was made regarding the structure and function of nucleosome, allowing us to obtain a deep understanding on how nucleosome, as the primary organizational principle, regulates DNA access during essential biological processes such as transcription, DNA repair and replication under different contexts.

The structure of the nucleosome

A high-resolution (2.8 Å) structure of the nucleosome (Figure 1.2 A) was first obtained by crystallography (Luger et al., 1997). The nucleosome was reconstituted *in vitro* by using palindromic DNA derived from alpha satellite DNA and recombinant histones from *Xenopus laevis*. Histones are wrapped by 147 bp DNA in a left handed supercoil. The main contacts between histones and DNA are through histone-fold domains (Figure 1.2 B) which organize 121 bp DNA. 13 bp DNA at each end binds to H3 α -N (Luger et al., 1997). These interactions between the DNA end and H3 are considered to be important for the stability of the nucleosome (Iwasaki et al., 2013). Inside the histone core of the nucleosome, histones H3 and H4 form a symmetric hetero-tetramer through forming a four helix bundle between two H3 molecules. Two H2A-H2B dimers interact with the (H3-H4)₂ tetramer through multiple interactions including H2B/H4 interactions through four-helix bundles and H2A/(H3-H4) interaction through the H2A “docking domain”

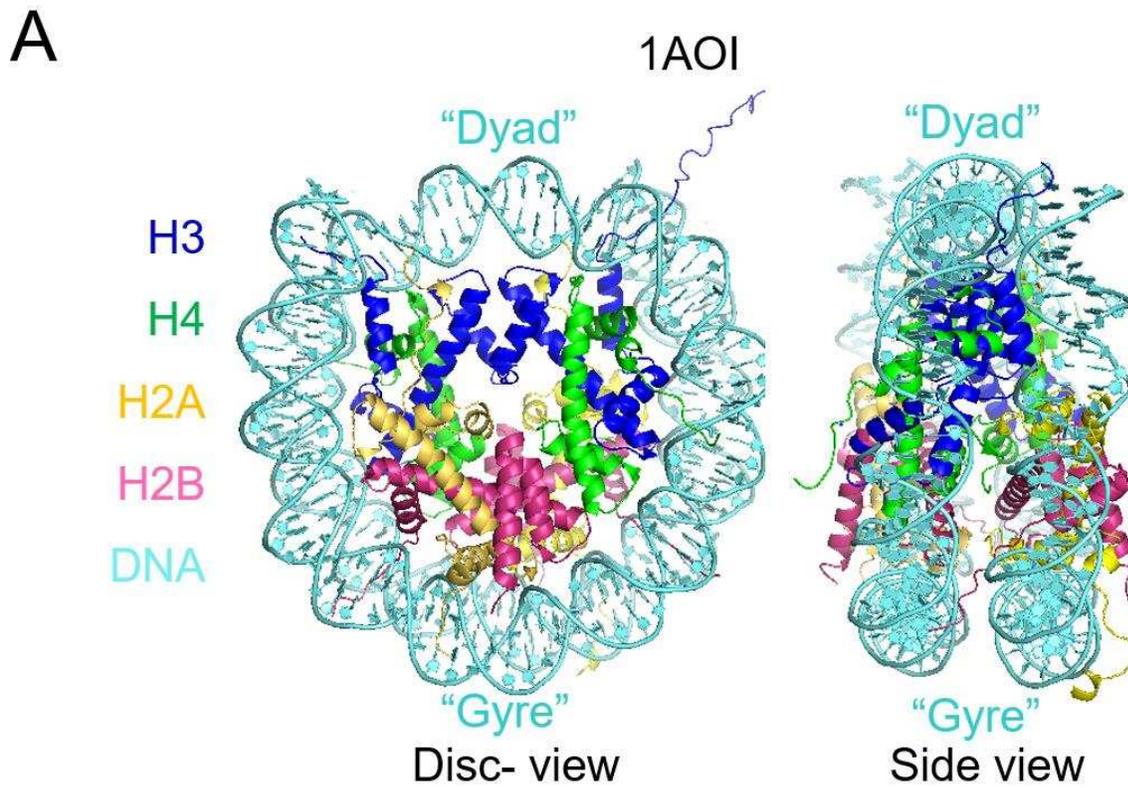


Figure 1.2 A Crystal structure of the nucleosome (1AOI, (Luger et al., 1997)). Each component of the nucleosome is labeled by different color, as indicated. “Dyad” and “Gyre” are used to describe the DNA at different positions as indicated.

B

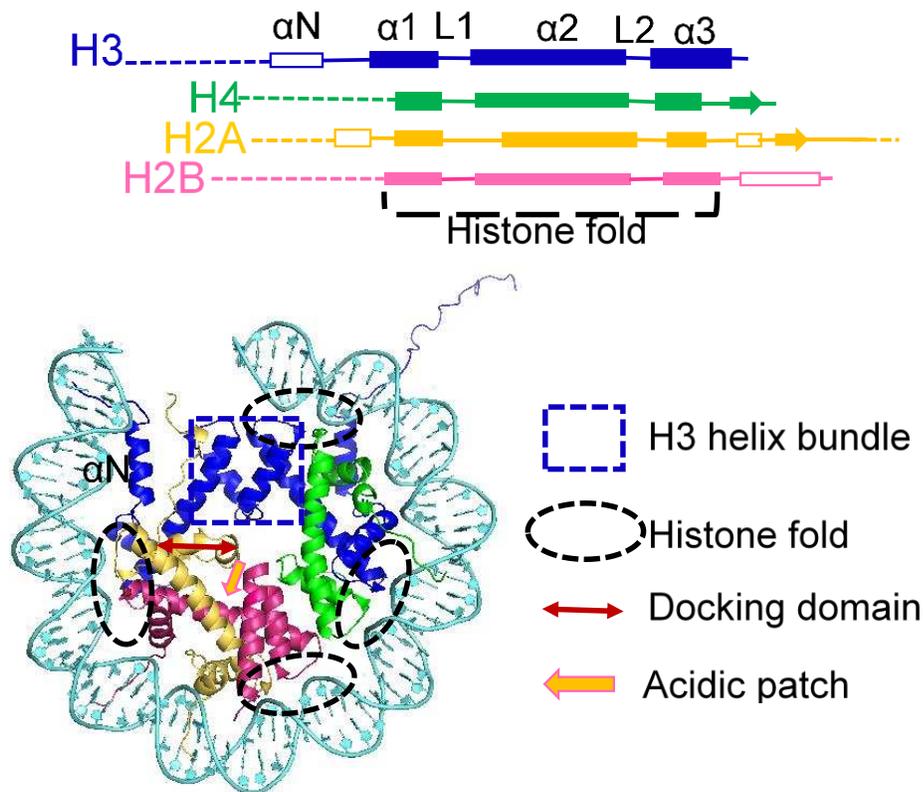


Figure 1.2 B Architecture of the nucleosome (modified based on reference (Luger, 2001)). The upper panel shows an overview of histone secondary structures. Dotted lines show the histone tails; the histone fold region of each histone, comprised of three α -helices (solid boxes), is indicated; the open boxes and arrows represent α -helix and β -strands in histone fold extensions, respectively. The lower panel shows how histones are organized in the nucleosome. Blue dashed box highlights the four-helix bundle formed by two H3; black dashed circles indicate the contacts between histone fold and DNA; red arrow covers the docking domain of H2A; brown arrow points the acidic patch formed by H2A and H2B.

(Figure 1.2 B). Two H2A interact with each other through a very small interface formed by their L1 loops from two sides of the nucleosome. Following studies (Clapier et al., 2008; Tsunaka et al., 2005; White et al., 2001) showed that the nucleosome core particle structures across species (from yeast to human) are almost the same. Recent work on archaeal nucleosome structure revealed that the mechanism of histone-based DNA compaction in archaea is very similar to what is seen in the eukaryotic nucleosome (Mattioli et al., 2017). However, the histones in eukaryotes have evolved into higher complexity. This allowed eukaryotic nucleosomes to develop more structural and functional diversity.

The higher order structures of chromatin

The DNA is organized by nucleosome, which is the first level of DNA compaction (Figure 1.1). The 10-nm fiber, known as ‘beads on the string’ is a structure representative of an extended form for nucleosomes in an array. *In vitro*, the 30 nm fiber is the second level of compaction which is induced by $MgCl_2$ or linker histone H1 (Dorigo et al., 2004). The H4 N-terminal tail plays a key role during $MgCl_2$ induced chromatin compaction (Bradley et al., 2006). The H4 tail interacts with the acidic patch, formed by six H2A and two H2B residues on the nucleosome (Figure 1.2, (Luger et al., 1997)). $MgCl_2$ stabilizes and enhances this interaction *in vitro* (Kalashnikova et al., 2013; Luger et al., 2012). Meanwhile, H1 binds to both DNA “dyad” and DNA linkers. A cryo-EM structure showed that H1 compacts chromatin into a 30 nm fiber with a “Tetra-

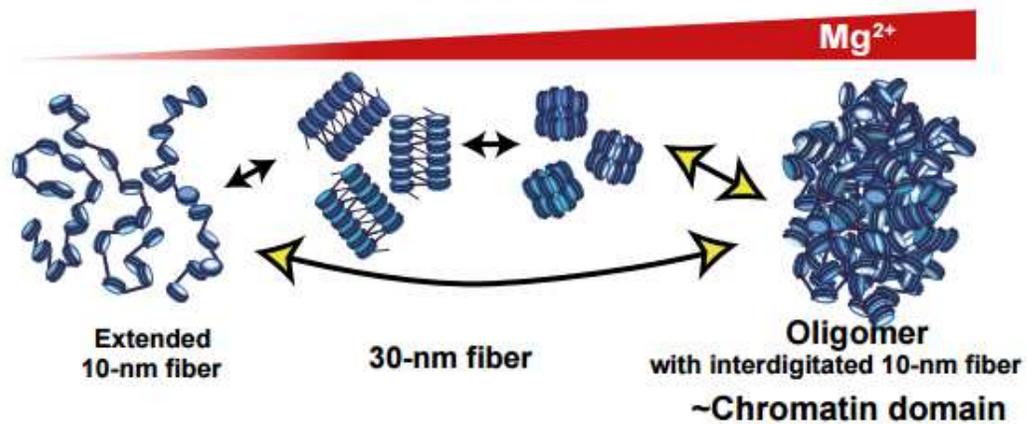


Figure 1.3: Chromatin compaction in solution with different concentrations of $MgCl_2$ (Maeshima et al., 2016)

This model scheme described how nucleosomes in an array compact, induced by different concentration of $MgCl_2$.

nucleosome” as the repeating unit *in vitro*. However, the 30 nm fiber was not observed *in vivo* (Maeshima et al., 2014). “Polymer melt” is an alternative model to describe the chromatin compaction (Hansen, 2012). This model suggests that nucleosomal arrays would directly self-associate into large globular oligomers without folding into 30 nm fiber (Figure 1.3), which can be regulated by linker histones such as H1 (Maeshima et al., 2016). Nevertheless, both H4 tail and linker histone H1 were shown to be very important for chromatin compaction in both models.

Histone modifications on histone tails

In the 1960s, histone acetylation was found to affect gene activity (Allfrey et al., 1964). Later, more types of histone modifications (methylation, phosphorylation, ADP-ribosylation and ubiquitination) were discovered. Most dominant modifications are acetylation and methylation on histone tails. Acetylation and deacetylation of lysine happens on all histones (for example, H3K9, H4K5, H2AK9 and H2BK12) reversibly, and these reactions are catalyzed by histone acetyltransferases (HAT) and histone deacetylases (HDACs) respectively (Bannister and Kouzarides, 2011; Chrun et al., 2017). Acetylation was reported to be necessary for transcription activation and DNA repair (Gong and Miller, 2013; Hebbes et al., 1988). Deacetylation is linked to chromatin condensation and transcription repression (Eberharter and Becker, 2002; Fukuda et al., 2006). Another prevalent modification of histone is methylation. The main targets of methylation are the lysines and arginines on H3 and H4 (for example, H3R2, H3K9,

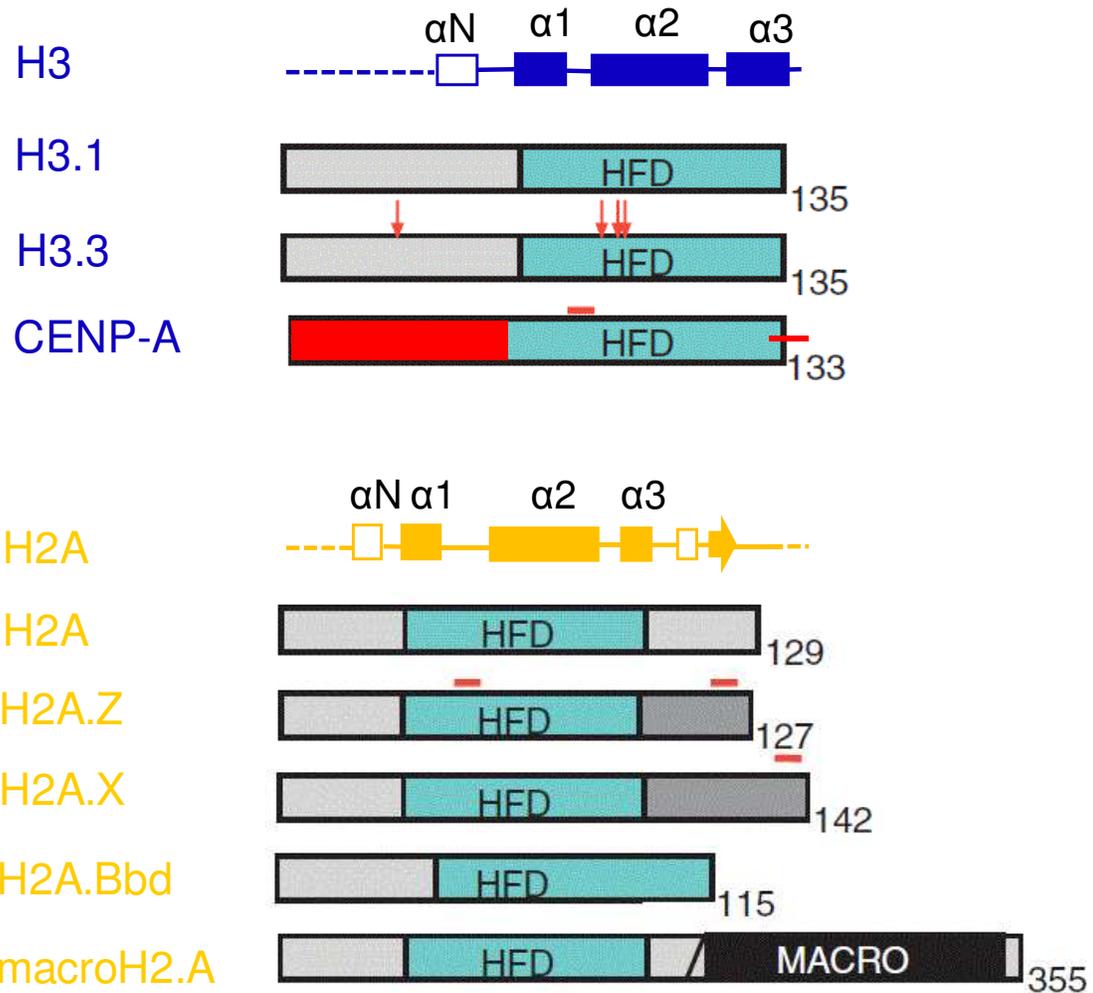


Figure 1.4 Overview of histone variants (Henikoff and Smith, 2015). Histone H3 and H2A variants are compared to their canonical counterpart. The sequence differences are indicated in red (arrow, line and box).

H3K27, H4R3 and H4K20). H3K9 tri-methylation was reported to induce heterochromatin through recruiting heterochromatin protein HP1 (Fischle et al., 2005). These modifications of histones make the nucleosome play key roles in regulating transcription, DNA replication and DNA damage repair.

Histone variants in nucleosomes

Besides histone modification, histone variants act as another epigenetic marker in nucleosome. These variants, only for histones that are self-interacting in the nucleosome (H2A and H3, figure 1.4), differ partially in sequence from the canonical histones. Because of variations in primary sequence, these variants bring unique features to the nucleosome. Their deposition in the genome, and their functions are quite different from their canonical counterparts as well. H2A has four variants including H2A.Z, H2A.X, macroH2A and H2A-Bbd. H2A.Z, a universal variant in Eukaryotes, shares ~60% sequence identity with canonical H2A (Suto et al., 2000). With H2A.Z incorporation, there is a subtle de-stabilization in the “docking domain” between H2A.Z/H2B dimer and (H3/H4)₂ tetramer. The acidic patch on the surface of nucleosome containing H2A.Z is extended, which promotes HP1 α mediated chromatin fiber folding (Fan et al., 2004) and the stimulation of chromatin remodelers such as Imitation Switch (ISWI) (Goldman et al., 2010). The changed L1 loop in H2A.Z also favors homotypic H2A.Z/H2A.Z incorporation (Suto et al., 2000). Due to its unique biophysical properties, the nucleosome containing H2A.Z was reported to regulate

transcription, DNA replication and DNA damage repair (Henikoff and Smith, 2015; Rogakou et al., 2016; Zlatanova and Thakar, 2008).

H2A.X shares almost 95% sequence identity with canonical H2A. H2A.X contains a very unique sequence motif located in the C-terminal tail; 'SQ(E or D)Ø' in which Ø represents a hydrophobic amino acid (Henikoff and Smith, 2015). The serine is phosphorylated in this motif to form "γH2A.X.", an early response to DNA double strand break (Rogakou et al., 1998).

MacroH2A, as its name indicates, has an extra 30 KDa domain (macro domain) at its C-terminus in addition to the N-terminal histone domain. The linker region between the histone domain and macro domain was shown to promote the self-association of nucleosome arrays *in vitro* (Muthurajan et al., 2011). Meanwhile, MacroH2A was found to have a domain which can induce the chromatin compaction *in vitro* (Abbott et al., 2005). More interestingly, the macro domain interacts with Poly (ADP-ribose) polymerase1 (PARP-1) to inhibit its enzymatic activity *in vitro* (Nusinow et al., 2007). This contributes to the silencing of one of two X chromosomes in female mammals.

H2A.Bbd was found to be absent in Barr bodies (Barr body deficient). Unlike MacroH2A, H2A.Bbd is excluded from the inactive X chromosome (Chadwick and Willard, 2001). The nucleosome containing H2A.Bbd is wrapped only by 118 bp DNA due to its unique

“docking domain”. This “relaxed” conformation may be linked to its association with active transcription (Bao et al., 2004; Henikoff and Smith, 2015; Tolstorukov et al., 2012).

H3 has two well-known variants; H3.3 and CENP-A. H3.3 shares almost 97% identity with its canonical counterpart (H3.1), with the only difference being 4 amino acids at the α 2-helix. Despite these seemingly “insignificant” differences, H3.3 shows a totally different deposition pathway (replication independent) from H3.1, which makes H3.3 specifically localize on actively transcribed genes with active transcription (Ahmad and Henikoff, 2002). CENP-A, on the other hand, is very different from H3.1 in its primary sequence. This makes CENP-A act as a very special epigenetic marker at the centromere.

The centromeric, CENP-A containing nucleosome

The centromere is the special region on the chromosome where microtubules attach. There is evidence for a tri-layered structure between the centromere and microtubules apparent under the electron microscope, which is called the kinetochore. The kinetochore (a complicated protein complex) specifically assembles at centromeres through recognizing the centromere marker CENP-A.

The nucleosome containing the H3 variant CENP-A defines the centromere

CENP-A was the first discovered centromeric protein by using autoimmune sera from patients with scleroderma pigmentosa (Earnshaw and Rothfield, 1985). CENP-A was then purified and identified as a histone H3 variant (Palmer et al., 1991) located at centromeres in humans. Although CENP-A and H3 share relatively poor similarity (62%) or identity (51%) in their primary sequences, their secondary structures are almost the same (Sekulic et al., 2010) (Figure 1.5 A). Like H3, CENP-A is incorporated into DNA with other histones (H4, H2A and H2B) to form a nucleosome (Warburton et al., 1997). It has been argued that CENP-A nucleosome can assume a hemisome architecture (Dalal et al., 2007; Malik and Henikoff, 2009). However, the major form of CENP-A nucleosomes is octameric, strongly resembling nucleosomes containing H3.1 (i.e. the canonical nucleosome) (Hasson et al., 2013; Nechemia-Arbely et al., 2017; Tachiwana et al., 2011). Nevertheless, nucleosomes containing CENP-A were found to be essential for kinetochore recruitment (Earnshaw et al., 1989; Foltz et al., 2006; Hooser et al., 2001). Mislocalization of CENP-A promotes the formation of neo-centromeres (Warburton et al., 1997). Therefore, the nucleosome containing CENP-A epigenetically defines the centromere.

Structure and biophysical property of nucleosome containing CENP-A

Comparing the crystal structures of the CENP-A nucleosome (Tachiwana et al., 2011) and the canonical nucleosome (Luger et al., 1997), they are very similar but show three

important differences (Figure 1.5 A): 1) CENP-A possesses a shorter α N helix than H3, so it results in less DNA coordination at the nucleosome entry-exit site. Thus, DNA ends in CENP-A nucleosome are more flexible (Conde et al., 2007; Kato et al., 2013; Roulland et al., 2016) (Figure 1.5 B); and in fact are disordered in the crystal structure (Tachiwana et al., 2011). 2) The CENP-A Targeting Domain (CATD) (Black et al., 2004), which includes the loop1 and helix1 regions of the histone fold domain, contains a two amino acid insertion including residues R80 and G81 (“RG” loop, figure 1.5 C); 3) The C-terminal tail of CENP-A diverges from histone H3 and possesses a “LEEGLG” motif in addition. These three distinct differences make the nucleosome containing CENP-A unique in its biophysical property as well as in its ability to recruit kinetochore proteins.

CENP-A nucleosomes are easier to disassemble than canonical nucleosomes *in vitro* (Arimura et al., 2014; Conde et al., 2007). One hypothesis is that less stably bound DNA ends make CENP-A nucleosomes more prone to dissociation. Rather than organizing 147 bp DNA as in canonical nucleosomes, CENP-A nucleosomes are only wrapped by 121 bp DNA (Tachiwana et al., 2011), as a result of the shorter, more disordered α N helix. Despite of these structural differences, CENP-A nucleosomes appear to be very stable during the cell cycle *in vivo*. The CATD of CENP-A was identified as the sequence element to be responsible for this high degree of stability *in vivo* (Bodor et al., 2013; Tachiwana et al., 2011), but the mechanism is not known. Meanwhile, the CATD and “LEEGLG” motif at C-terminus are very important for the recruitment of inner kinetochore proteins (Carroll et al., 2009, 2010).

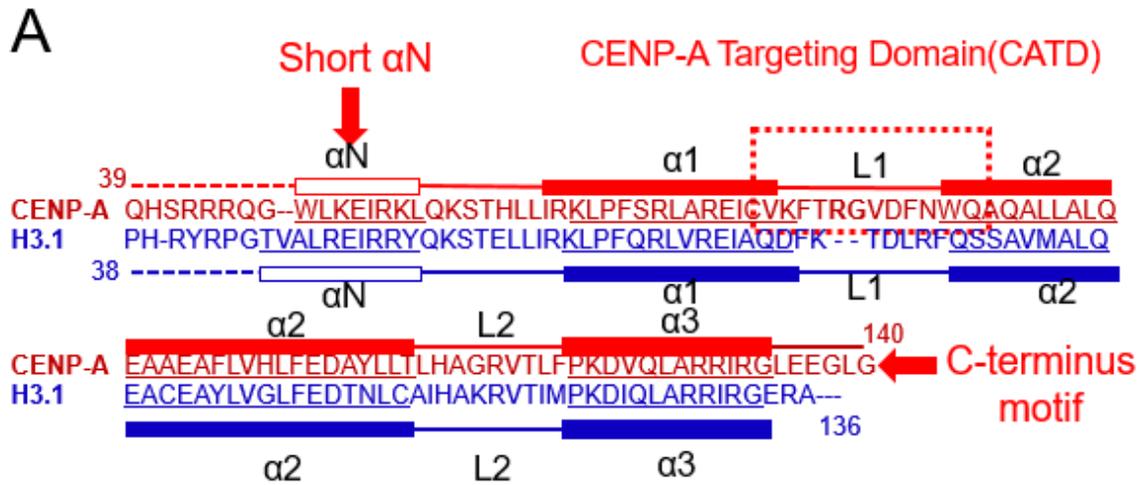


Figure 1.5 A, sequence alignment between histone H3.1 and its variant CENP-A. The N-terminal tail is not included here. The secondary structures are indicated in different colors. The red dotted box shows the CENP-A targeting domain (CATD). Two unique features of CENP-A (short α N and “LEEGLG” motif) are also pointed out by red arrows.

B

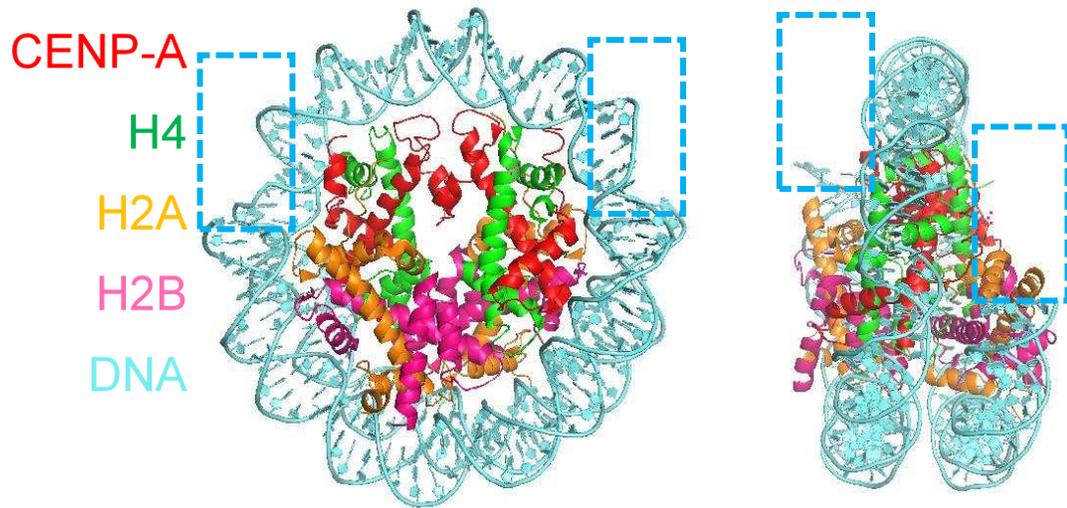


Figure 1.5 B. The nucleosome containing CENP-A has two flexible DNA ends. Each component is labeled in different colors as indicated. The blue boxes highlight the invisible DNA ends in crystal structure (3AN2, (Tachiwana et al., 2011)).

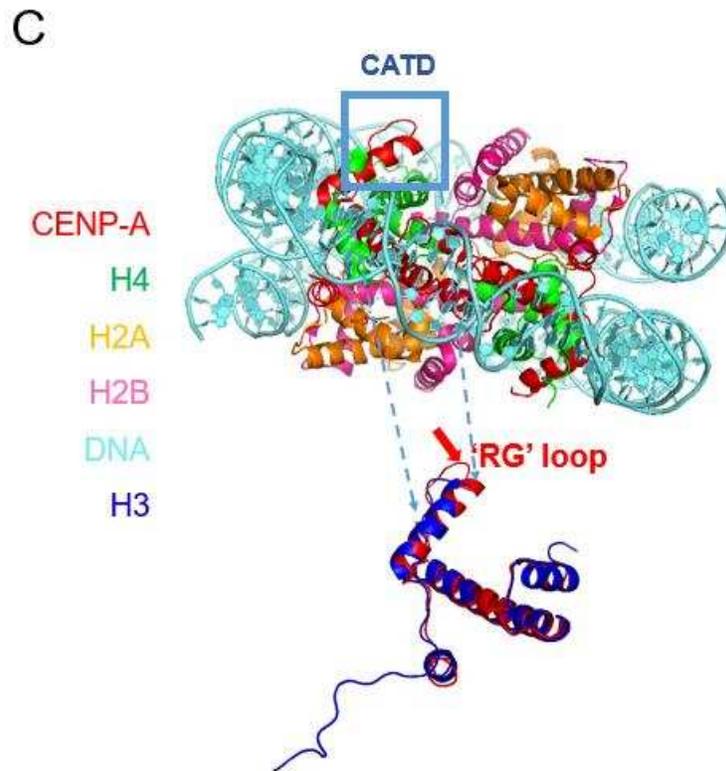


Figure 1.5 C. Structural difference between CENP-A and H3 nucleosome in Loop 1 (L1). Compared to H3, L1 in CENP-A sticks out from nucleosome. There are two extra amino acids (Arginine 80 and Glycine 81) forming a tip in L1. 'RG' represents Arginine 80 and Glycine 81.

Constitutive Centromere Associated Network (CCAN)

The kinetochore is defined as a large protein complex connecting the centromere with microtubules. It is critical for the movement and segregation of chromosomes during mitosis. A tri-layered structure was observed around the centromere by electron microscopy (Brinkley and Stubblefield, 1966; Jokelainen, 1967; Luykx, 1965). There are more than 100 proteins in the kinetochore. The outer plate of the kinetochore connects microtubule tips with a flexible network (Dong et al., 2007; McIntosh et al., 2008; Rev, 2015) formed by the conserved KNL-1/MIS12 complex/Ndc80 complex (KMN) network (Cheeseman et al., 2006). In the inner plate, the Constitutive Centromere Associated network (CCAN) directly and continuously contact with centromere throughout the cell cycle (Pesenti et al., 2016).

The sub-complexes of CCAN

There are 5 sub-complexes in CCAN (Pesenti et al., 2016): CENP-C, the CENP-L/CENP-N complex (NL), the CENP-H/CENP-I/CENP-K/CENP-M complex (HIKM), the CENP-T/CENP-W/CENP-S/CENP-X complex (TWSX), and the CENP-O/CENP-P/CENP-Q/CENP-R/CENP-U complex (OPQRU). These complexes are assembled in a well-controlled manner (Figure 1.6). In the first step, CENP-C and CENP-N directly and independently bind to the nucleosome containing CENP-A (Carroll et al., 2009, 2010), serving as the platform for the kinetochore assembly. CENP-L then bridges CENP-N and CENP-C (Weir et al., 2016). CENP-C recruits CENP-HIKM which is important for

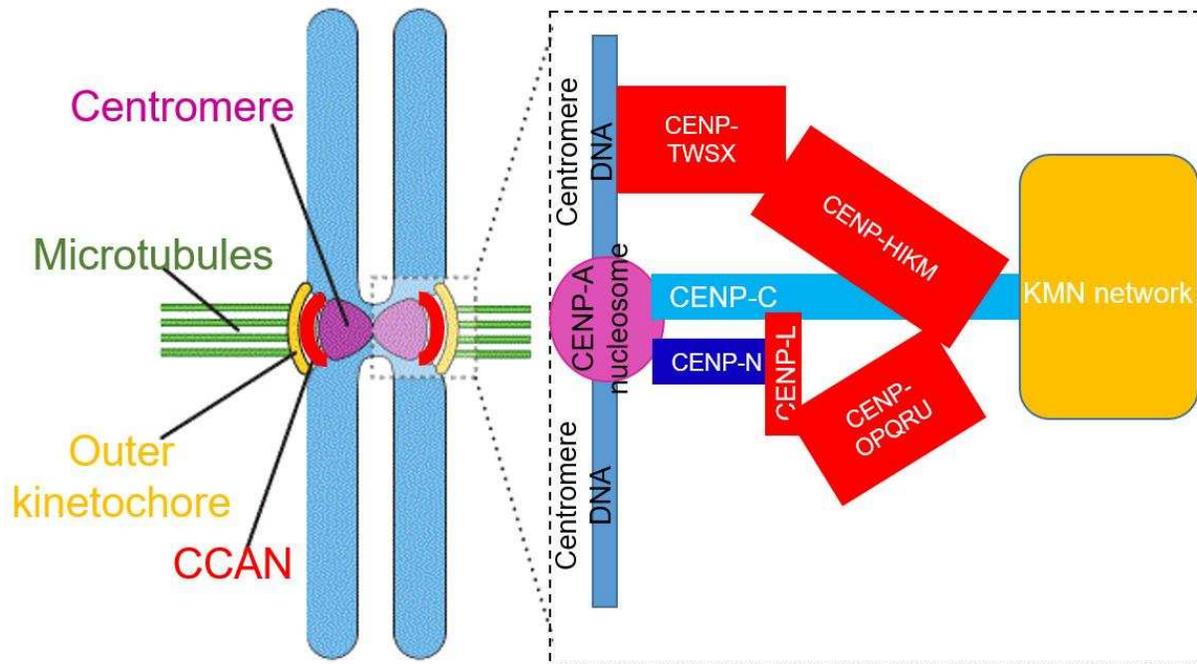


Figure 1.6 Schematic model (modified from the reference (Klare et al., 2015)) for the Centromere Constitutive Associate Network (CCAN). Cyan and blue box represent CENP-C and CENP-N, respectively. Other CCAN components are indicated by red boxes. Centromere is labeled by magenta. Outer kinetochore is in brown. Green sticks represent microtubules.

CENP-TW localization to the centromere through direct interaction (Klare et al., 2015). At the centromere, since CENP-T and CENP-W both contain histone-fold domains to bind centromere DNA (Hori et al., 2008), CENP-TW together with CENP-SX forms a centromeric chromatin structure on DNA (Nishino et al., 2012). CENP-C also directly recruits the KMN network at the outer kinetochore through physical contacts with Mis12 (Screpanti et al., 2011). On the other hand, CENP-N, together with CENP-L, serves as another critical node for the CCAN assembly at centromere (McKinley et al., 2015). Depleting CENP-N causes significant loss of CENP-K, CENP-T and CENP-O-P (McKinley et al., 2015). CENP-L is one key component for the recruitment of CENP-OPQRU (Eskat et al., 2012) which forms a tightly packed complex through multiple pairwise interactions (Eskat et al., 2012). In summary, CENP-C and CENP-N recognize the nucleosome containing CENP-A first. Sequentially, CENP-HIKM, CENP-TWSX and CENP-OPQRU are recruited and maintained to form the CCAN.

Key player at the centromere: CENP-C

CENP-C is a large (943 amino acids in human) and disordered protein (Figure 1.7). CENP-C⁴²⁶⁻⁵³⁷ specifically binds to the nucleosome containing CENP-A through recognizing the CENP-A C-terminal “LEEGLG” motif as well as the acidic patch of H2A-H2B on the nucleosome (Carroll et al., 2010; Kato et al., 2013)(Figure 1.5A). CENP-C¹⁻⁷¹ interacts with Mis12 (KMN subunit) directly (Screpanti et al., 2011), which suggests that CENP-C is a very important linker between the inner and outer kinetochore. CENP-C¹⁸⁹⁻²⁹⁰ binds to CENP-HK (Klare et al., 2015). In addition, CENP-C¹⁶⁶⁻³²⁴ was shown to

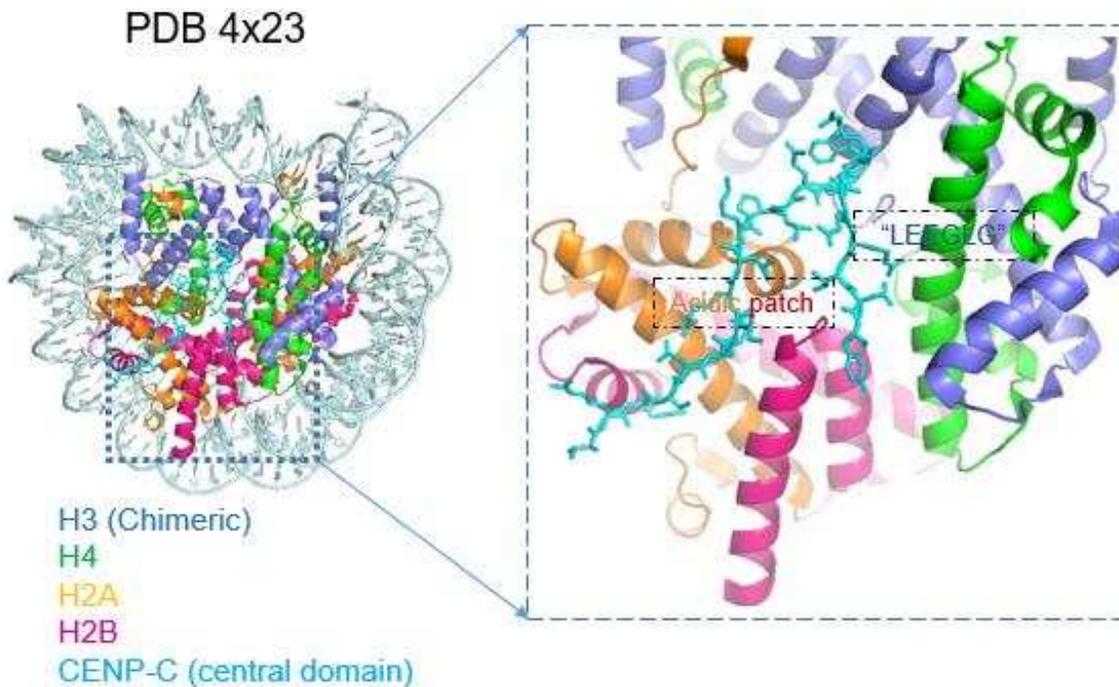
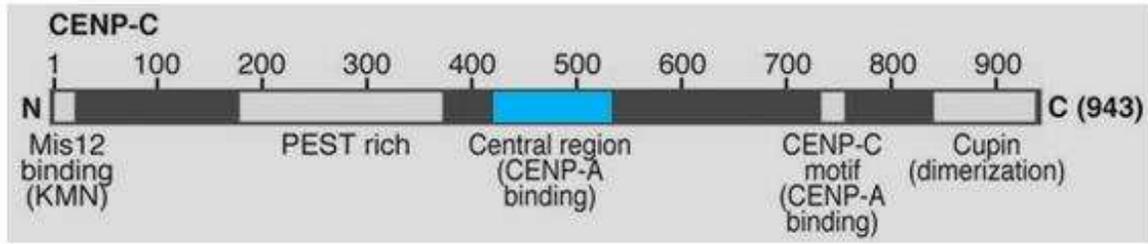


Figure 1.7 CENP-C recognizes CENP-A nucleosome through its central region. The upper panel shows the motifs of human CENP-C (Klare et al., 2015). The cyan box highlights the central region on CENP-C which is responsible for recognizing the CENP-A nucleosome. The lower panel shows the crystal structure (4x23, (Kato et al., 2013)) of CENP-C (central domain, in cyan) in complex with the nucleosome (chimeric H3 with "LEEGLG" motif at C-terminus). Histones and DNA are labeled as indicated.

directly bind to CENP-NL (Nagpal et al., 2015). Through its C-terminal region, CENP-C⁶³⁸⁻⁸¹⁹ (Mif2p homology domain II) binds to alpha satellite DNA *in vivo*; CENP-C⁷⁶⁰⁻⁹⁴³ (Mif2p homology domain III) is essential for its dimerization or oligomerization which could potentially play role in high order (Sugimoto et al., 1997; Trazzi et al., 2009). Structurally, CENP-C serves as a platform for the kinetochore assembly on the nucleosome containing CENP-A. Functionally, CENP-C is essential for the incorporation of CENP-A into centromeres during early G1 phase after the exit of mitosis (Jansen et al., 2007; Moree et al., 2011). The Mis-18 complex (Mis18- α , Mis18- β , Mis18BP1, RbAp48 and RbAp46 in humans) primes or licenses the deposition of CENP-A (Fujita et al., 2007; Hayashi et al., 2004; Stellfox and Foltz, 2014). CENP-C which is associated with 'old' CENP-A nucleosome recruits Mis18BP1 to centromere through direct binding. The dephosphorylated Mis18BP1 binds to the Mis18- α /Mis18- β heterotetramer in late telophase. CENP-A/H4 chaperone Holiday Junction recognition protein (HJURP), which deposits CENP-A onto centromere DNA during G1 phase (Foltz et al., 2010; Hayashi et al., 2004; Nardi et al., 2016), is recruited by (Mis18- α /Mis18- β)₂ (Nardi et al., 2016). Therefore, CENP-C determines the localization of CENP-A at the centromere. Depleting CENP-C inhibits CENP-A chromatin assembly (Moree et al., 2011). CENP-C⁴²⁶⁻⁵³⁷ was also reported to prevent the structural transition of the CENP-A nucleosome through sliding of DNA gyres (Falk et al., 2016). This effect was considered to stabilize CENP-A nucleosomes (Falk et al., 2015). However, no direct evidence was provided *in vitro*.

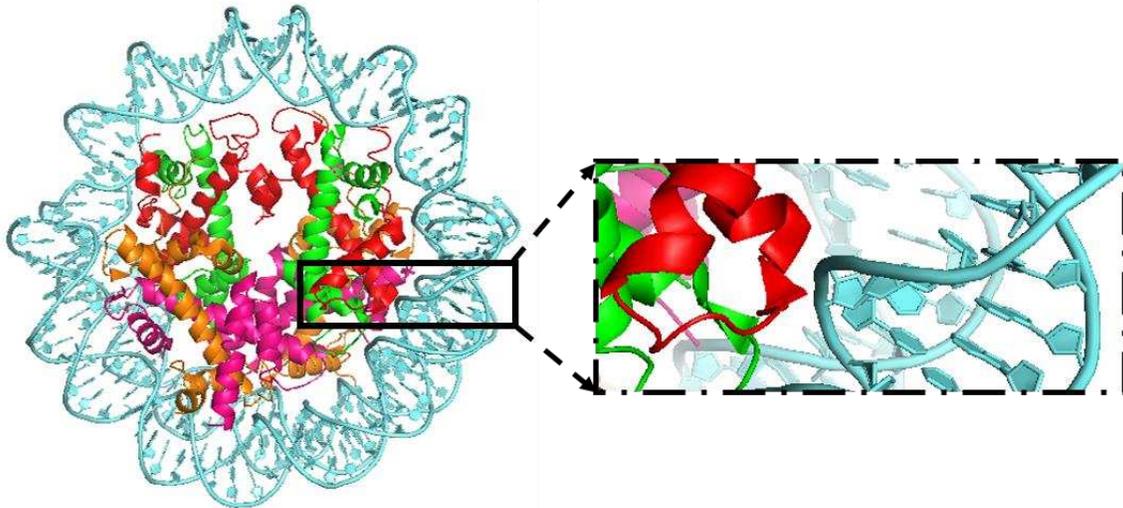
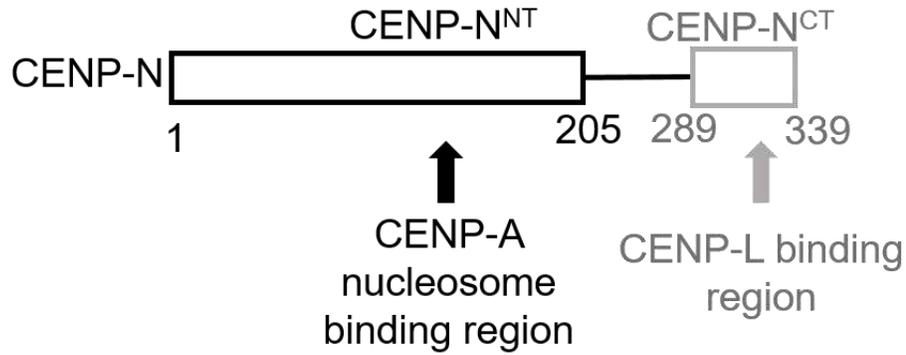


Figure 1.8 CENP-N recognizes CENP-A nucleosome by its N-terminal region. Upper panel is the scheme depicts the domains for different interactions. CENP-N N-terminal region (CENP-N^{NT}) specifically binds to CENP-A nucleosome; CENP-N C-terminal region (CENP-N^{CT}) binds to CENP-L. Lower panel shows the regions on CENP-A nucleosome that CENP-N interacts with. It includes the CATD and the nucleosomal DNA close to CATD (Guo et al., 2017).

Dynamic node at the centromere: CENP-N

As one of the only two proteins which directly recognize a nucleosome containing CENP-A, CENP-N is less well-studied than CENP-C. There are at least five isoforms of CENP-N in human. Isoform1 (full length= 339 amino acids, Figure 1.8) is the only one that was found to associate with both CENP-L and the nucleosome containing CENP-A (Carroll et al., 2009). N-terminal region of isoform1 (1-289) was shown to be responsible on interacting with the nucleosome containing CENP-A and C-terminal region (290-339) directly binds to CENP-L (Carroll et al., 2009). Other isoforms are either shorter (for example 1-204) or are different on C-terminal regions. The isoforms are expressed differently in different tissues (Gerhard et al., 2004; Ota et al., 2004).

The “RG” loop on CATD (CENP-A Targeting Domain) was identified as the key feature to be recognized by the N-terminal region of CENP-N (1-289) (Carroll et al., 2009; Fang et al., 2015). Recently, CENP-N (1-240) was shown to bind the nucleosome containing CENP-A (McKinley et al., 2015). Thus, the nucleosome binding domain is located within the first 240 amino acids. Recent HDX data established that CENP-N simultaneously binds to the CATD and the nucleosomal DNA close to the CATD (Guo et al., 2017). It was proposed that CENP-N fastens CENP-A to the DNA. However, little is known regarding the detailed structural basis. Unlike CENP-C, which is a persistent scaffold at centromere during cell-cycle, CENP-N was shown to be largely absent during Mitosis and G1 phases (Hellwig et al., 2011). By fluorescence resonance energy transfer (FRET), CENP-N was observed to associate with the kinetochore stably in the middle of

S phase, and dissociate again during G2 (Hellwig et al., 2011). It is still unclear how this dynamic behavior is regulated. One report showed that CENP-N would fall off the nucleosome array containing CENP-A once the array formed a more compact structure *in vitro* (Fang et al., 2015). Investigating how CENP-N interacts with CENP-A nucleosome will help us get deeper understanding on the CENP-N dynamics.

The structure and function of CENP-N are still illusive. Structurally, how does CENP-N recognize the CENP-A nucleosome from canonical (H3) nucleosome? Functionally, does CENP-N only serve as platform for CCAN assembly, or would CENP-N affect the nucleosome containing CENP-A structure once they associate with each other? Answering these questions would help us obtain a deeper understanding on how kinetochore specifically assembles at centromere and its following effects.

Specific aims

One of the main goals of this thesis is to completely understand the specific recognition of the nucleosome containing CENP-A by CENP-N at a structural level. Only the N-terminal region of CENP-N is responsible for its binding to the nucleosome containing CENP-A. Thus, I expressed and purified this part of CENP-N from insect cells. I characterized and optimized the binding of CENP-N to CENP-A nucleosome biochemically. By using single particle cryo electron microscopy (cryoEM), I obtained the electron density map for CENP-A nucleosome:CENP-N complex. Through collaboration

with Dr. Andrea Musacchio's group, we solved the structure of CENP-A nucleosome:CENP-N complex. The structure was further validated by mutagenesis *in vitro* and *in vivo*, and the results were published in *elife* (Pentakota et al., 2017). This was the first cryo-EM structure determined in the Luger Lab, and my work helped established the methodology in the lab.

The second goal of this thesis was to evaluate the effect of CENP-N on CENP-A nucleosome upon binding. I proposed that CENP-N stabilizes the CENP-A nucleosome by directly interacting. To test this hypothesis, I performed Electron Mobility Shift Assay (EMSA) and electron microscopy to evaluate the stability of the CENP-A nucleosome in the absence or presence of CENP-N under different *in vitro* conditions. I also tested the stabilizing effect on CENP-A nucleosome when both CENP-C and CENP-N bind to the CENP-A nucleosome. By collaborating with Dr. Aaron Straight's group, we also evaluated the stabilizing effect of either CENP-N or CENP-C on CENP-A nucleosome *in vivo*. By combining *in vitro* and *in vivo* data, we provided deeper insight on how CENP-A nucleosome is maintained at centromere. This work was published in 'Molecular biology of cell' (Cao et al., 2018) The mechanism on how CENP-N stabilizes CENP-A nucleosome was also investigated by cryo-EM.

Finally, the CENP-N was found to be able to oligomerize CENP-A nucleosomes *in vitro*. Analytical ultracentrifugation (AUC) and cryo-EM were used to get deeper insight on its

mechanism. In summary, the works in this thesis provide comprehensive understanding on the structure and function of CENP-N at centromere.

CHAPTER 2 THE CYRO-EM STRUCTURE OF CENP-A NUCLEOSOME IN COMPLEX WITH CENP-N¹

Overview

Centromere protein (CENP) A, a histone H3 variant, is a key epigenetic determinant of chromosome domains known as centromeres. Centromeres nucleate kinetochores, multi-subunit complexes that capture spindle microtubules to promote chromosome segregation during mitosis. Two kinetochore proteins, CENP-C and CENP-N, recognize CENP-A in the context of a rare CENP-A nucleosome. Here, we reveal the structural basis for the exquisite selectivity of CENP-N for centromeres. CENP-N uses charge and space complementarity to decode the L1 loop that is unique to CENP-A. It also engages in extensive interactions with a 15-base pair segment of the distorted nucleosomal DNA double helix, in a position predicted to exclude chromatin remodelling enzymes. Besides CENP-A, stable centromere recruitment of CENP-N requires a coincident interaction with a newly identified binding motif on nucleosome-bound CENP-C. Collectively, our studies clarify how CENP-N and CENP-C decode and stabilize the non-canonical

¹ Pentakota, S.*, Zhou, K.*, Smith, C., Maffini, S., Petrovic, A., Morgan, G.P., Weir, J.R., Vetter, I.R., Musacchio, A., and Luger, K. (2017). Decoding the centromeric nucleosome through CENP-N. *Elife* 6. . I am the co-first (*) author of this paper. My contributions on this work were described as following (as stated in the manuscript): SP and KZ expressed and purified proteins and performed analytical biochemical experiments, including experiments with CENP-N mutants. KZ, with GPM and KL determined the EM structure. In their respective laboratories, KZ and KL, and SP, IRV, AP, and AM, conceived the mutational analysis of CENP-N, which was executed by KZ and SP. AM drafted the paper with substantial contributions from SP, KZ, IRV, and KL.

CENP-A nucleosome to enforce epigenetic centromere specification and kinetochore assembly.

Introduction

Accurate segregation of chromosomes from a mother cell to its two daughters during cell division is a prerequisite for healthy cell physiology and for the transmission of the genetic information across generations (Santaguida and Amon, 2015). Specialized, conserved molecular machinery dedicated to this crucial function has been identified in the majority of eukaryotic organisms studied to date (Drinnenberg et al., 2016; van Hooff et al., 2017). The purpose of this machinery is to generate stable linkages between chromosomes, the carriers of genetic information, and the mitotic spindle, the microtubule-based structure devoted to the segregation of chromosomes into the daughter cells.

In the last two decades, substantial progress in our understanding of the molecular features of the chromosome segregation apparatus has been made. A crucial role in this process is played by centromeres, specialized chromatin domains whose defining mark in almost all known eukaryotes is the enrichment of centromeric protein A (CENP-A, also known as CenH3), which replaces histone H3 in nucleosomes (Fukagawa and Earnshaw, 2014; Musacchio and Desai, 2017). The primary function of centromeres is to provide a platform for the assembly of macromolecular complexes known as

kinetochores, whose task in turn is the physical capture of microtubules of the mitotic spindle. Kinetochores contain approximately 30 core subunits, normally subdivided in centromere-proximal and microtubule-proximal groups. The microtubule-proximal subunits (outer kinetochore), which are directly implicated in microtubule binding, are usually denoted as the KMN assembly, from the name of three sub-complexes, the Knl1, Mis12, and Ndc80 complexes (Musacchio and Desai, 2017). The centromere-proximal subunits (inner kinetochore), which are also organized in sub-complexes, are collectively identified as the constitutive centromere associated network (CCAN) because they appear to reside at centromeres for the entire cell cycle (Cheeseman and Desai, 2008; Foltz et al., 2006; Izuta et al., 2006; Obuse et al., 2004; Okada et al., 2006) (Figure 2.1 A).

The ability of CENP-A to nucleate kinetochores depends on its incorporation into nucleosomes (CENP-A nucleosomes) with histones H2A, H2B, and H4. *In vitro*, these interact specifically and selectively with two CCAN components, CENP-C and CENP-N (Carroll et al., 2010; Carroll et al., 2009; Guo et al., 2017; Guse et al., 2011; Hoffmann et al., 2016; Klare et al., 2015; Nagpal et al., 2015; Samejima et al., 2015; Weir et al., 2016). Binding of these proteins to CENP-A-nucleosomes has been shown to require two regions where the CENP-A sequence diverges significantly from that of histone H3,

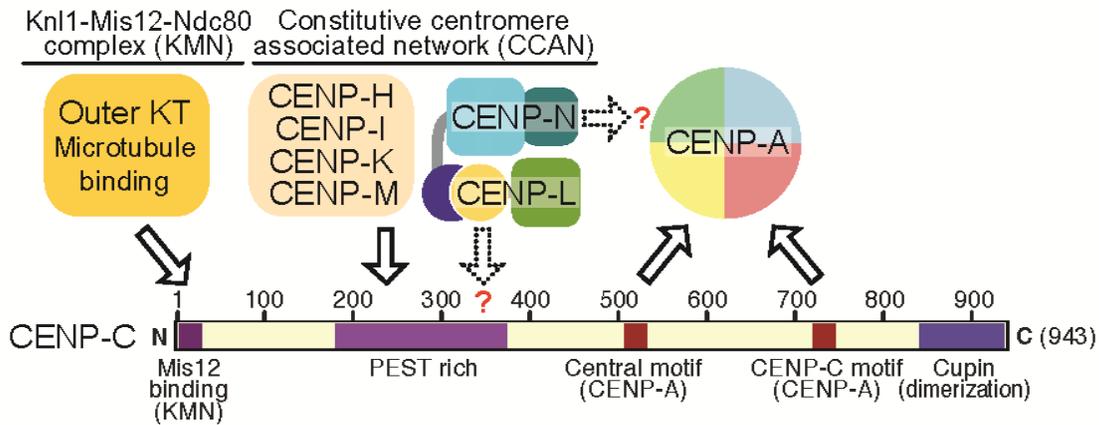
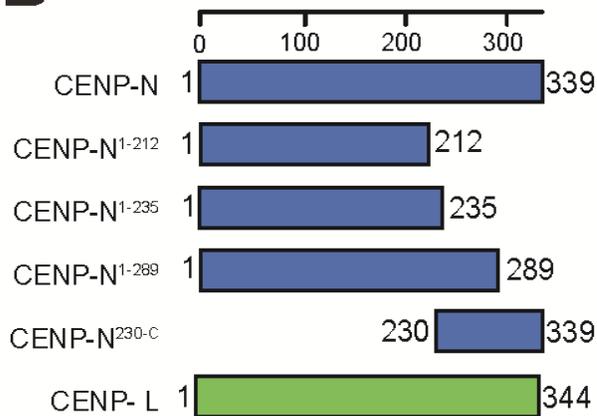
A**B**

Figure 2.1 The interaction of CENP-N with nucleosomes

A. Schematic of crucial CCAN and KMN subunits discussed in the text. The Kn1-Mis12-Ndc80 (KMN) complex is the main microtubule receptor at the kinetochore. Other interactions are discussed in the main text. The question mark indicates that the precise determinants for the recruitment of CENP-LN to CENP-C and for the interaction of CENP-N with the CENP-A nucleosome have not been identified. B. Schematic depicting constructs described in the manuscript.

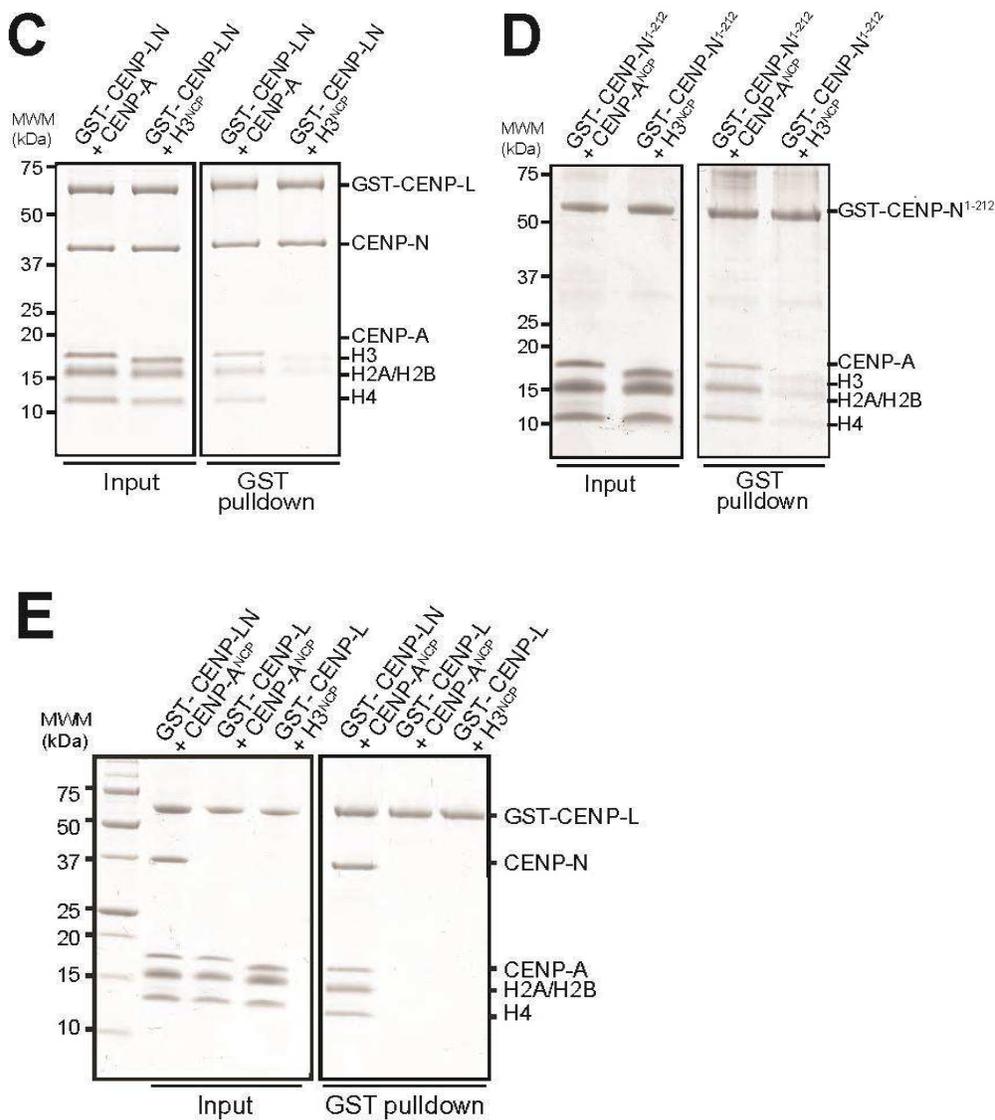


Figure 2.1 The interaction of CENP-N with nucleosomes

C–E. Solid phase binding assays where the indicated GST fusion proteins were immobilized on glutathione-sepharose beads (at a final concentration of 1 mM) and incubated with 3 mM of the indicated nucleosome core particles. After incubation (see Materials and methods), beads were centrifuged, washed, and bound proteins visualized by SDS-PAGE and Coomassie staining.

the L1 loop and the C-terminal tail (Carroll et al., 2009; Fachinetti et al., 2013; Guo et al., 2017; Kato et al., 2013; Logsdon et al., 2015). An evolutionary conserved motif of CENP-C, present in one or two copies in different organisms, is sufficient for recognition of CENP-A *in vitro*. This motif interacts primarily with a solvent-exposed acidic patch on the H2A and H2B subunits of the CENP-A nucleosome and also decodes the divergent C-terminal tail of CENP-A (Guo et al., 2017; Kato et al., 2013). The two copies of this motif in human CENP-C are referred to as the central motif (or domain) and the CENP-C motif (Figure 2.1 A). While at least the central motif has been shown to be required for efficient centromere retention of newly incorporated CENP-A (Guo et al., 2017), neither motif appears to be strictly necessary for centromere localization of CENP-C in human cells (Guo et al., 2017), likely because CENP-C contains binding sites for additional CCAN subunits that can stabilize its centromere localization even in the absence of a direct interaction with CENP-A (Guo et al., 2017; Hinshaw and Harrison, 2013; Klare et al., 2015; McKinley et al., 2015; Nagpal et al., 2015; Weir et al., 2016). The specific succession of binding sites within CENP-C, a protein that secondary structure prediction algorithms identify as being largely intrinsically disordered, has led to suggest that it acts as a blueprint in the establishment of the inner to outer kinetochore axis, with an N-terminal motif involved in stabilizing the outer kinetochore, a middle region involved in stabilizing the inner kinetochore CCAN complex, and a C-terminal region involved in interactions with the centromeric chromatin (Figure 2.1 A) (Gascoigne et al., 2011; Kato et al., 2013; Klare et al., 2015; McKinley et al., 2015; Przewłoka et al., 2011; Screpanti et al., 2011).

CENP-N forms a constitutive complex with CENP-L (designated CENP-LN complex), which in turn interacts with the CENP-HIKM complex and with CENP-C (Guo et al., 2017; Hinshaw and Harrison, 2013; Klare et al., 2015; McKinley et al., 2015; Weir et al., 2016). Binding of CENP-N requires the exposed L1 loop of CENP-A and may also reach into the neighbouring DNA (Carroll et al., 2010; Carroll et al., 2009; Fang et al., 2015; Guo et al., 2017). The structural basis of the interaction of CENP-N with the CENP-A nucleosome, however, has remained elusive. Furthermore, it is unclear whether this interaction is sufficient for the recruitment of CENP-N to the kinetochore, or whether additional interactions with CCAN subunits are also required. Here, we addressed both issues. First, we combined X-ray crystallography and cryo electron microscopy (EM) to gain a high-resolution view of the CENP-N:CENP-A nucleosome complex, and identified and validated the main determinants of this interaction. Second, we defined the determinants of a physical interaction of CENP-LN with CENP-C and demonstrated that kinetochore recruitment of CENP-N requires the coincident presence of CENP-A and CENP-C at kinetochores. Our studies have important implications for kinetochore assembly and epigenetic specification of centromeres.

Results

Crystal structure of CENP-N¹⁻²³⁵

Human CENP-N, a 339-residue protein (Figure 2.1 B), interacts directly with CENP-L (Hinshaw and Harrison, 2013; Weir et al., 2016). When immobilized on solid phase and challenged with CENP-A or H3 nucleosome core particles (NCPs), CENP-LN interacted

Table 2.1

X-ray data collection and refinement statistics

Data collection and processing				
	Native	SeMet 1	SeMet 2	SeMet 1+2
Space group	P4 ₁	P4 ₁	P4 ₁	P4 ₁
Wavelength	0.97793	0.9793	0.9793	0.9793
No. xtals	1	1	1	2
Source	SLS	PETRA	PETRA	PETRA
Detector	Pilatus 6M	Pilatus6M	Pilatus 6M	Pilatus 6M
Mol/AU	2	2	2	2
a,b,c (Å)	87.3 87.3 81.1	88.99 88.99 76.96	89.14 89.14	88.99 88.99 76.96
α, β, γ (°)	90 90 90	90 90 90	90 90 90	90 90 90
Resolution (Å)	87.3-2.74 (2.81-2.74)*	48.7-3.3 (3.9-3.3)	48.8-3.2 (3.3-3.2)	48.7-3.3 (3.4-3.3)
R _{meas}	8.2 (155.1)	17.2 (153.4)	18.8 (173.4)	18.7(167.8)
I/ σ I	17.3 (1.4)	7.5 (1.1)	7.2 (1.0)	10.4 (1.4)
Completeness (%)	99.8 (98.5)	100.0 (100.0)	99.9 (98.8)	100.0(100.0)
Redundancy	9.4 (8.7)	7.1 (7.2)	7.0 (6.3)	14.1 (14.1)
Refinement				Phasing
Resolution (Å)	87.3-2.7			FOM 0.39
No. reflections	17103			BAYES-CC 38.1
R _{work} / R _{free} (%)	21.6/26.1			12 Selenium-sites
No. atoms:				
Protein/ Ligands	3432/6			
Water	10			
aver. B (Å ²)	90.4			
R.m.s. deviations				
Bond lengths (Å)	0.0076	Ramachandran plot: 98.0 % favourable, 0 % outliers		
Bond angles (°)	1.27			

* Values in parentheses are for highest resolution shell

specifically with CENP-A^{NCP} (Figure 2.1 C). As shown previously (Carroll et al., 2009), the CENP-A binding region of the CENP-LN complex lies within the N-terminal region of CENP-N, because a stable fragment encompassing residues 1-212 of human CENP-N (CENP-N¹⁻²¹²) also bound selectively to CENP-A^{NCPs} but not H3^{NCPs} (Figure 2.1 D).

To address the structural features of CENP-N and the basis of its interaction with the CENP-A^{NCP}, we therefore focused our structural analysis on N-terminal constructs of CENP-N (Figure 2.1 B). We obtained well diffracting crystals of the CENP-N¹⁻²³⁵ construct and determined its crystal structure at 2.8 Å resolution (Table 2.1). CENP-N¹⁻²³⁵ consists of two closely juxtaposed domains that interact through an extended interface to form a single structural unit (Figure 2.2 A-B). The first domain (residues 1-77) consists of a 5-helix bundle, whereas the second domain (residues 78-212, cyan in Figure 2.1 A) consists of a 6-stranded anti-parallel β-sheet sandwiched between α-helices (Figure 2.2 C-D). There is no clear density beyond residue ~210, indicating that the structure is disordered after this point. Fold-recognition by DALI (Holm and Rosenstrom, 2010) identified similarity of the first domain to PYRIN domains (PYDs; a superposition is shown in Figure 2.2 Supplement 1A-B). PYDs are ‘death fold’ family domains implicated in protein-protein interactions relevant to inflammation and apoptosis (Ratsimandresy et al., 2013). They have not been previously implicated in interactions with DNA or chromatin.

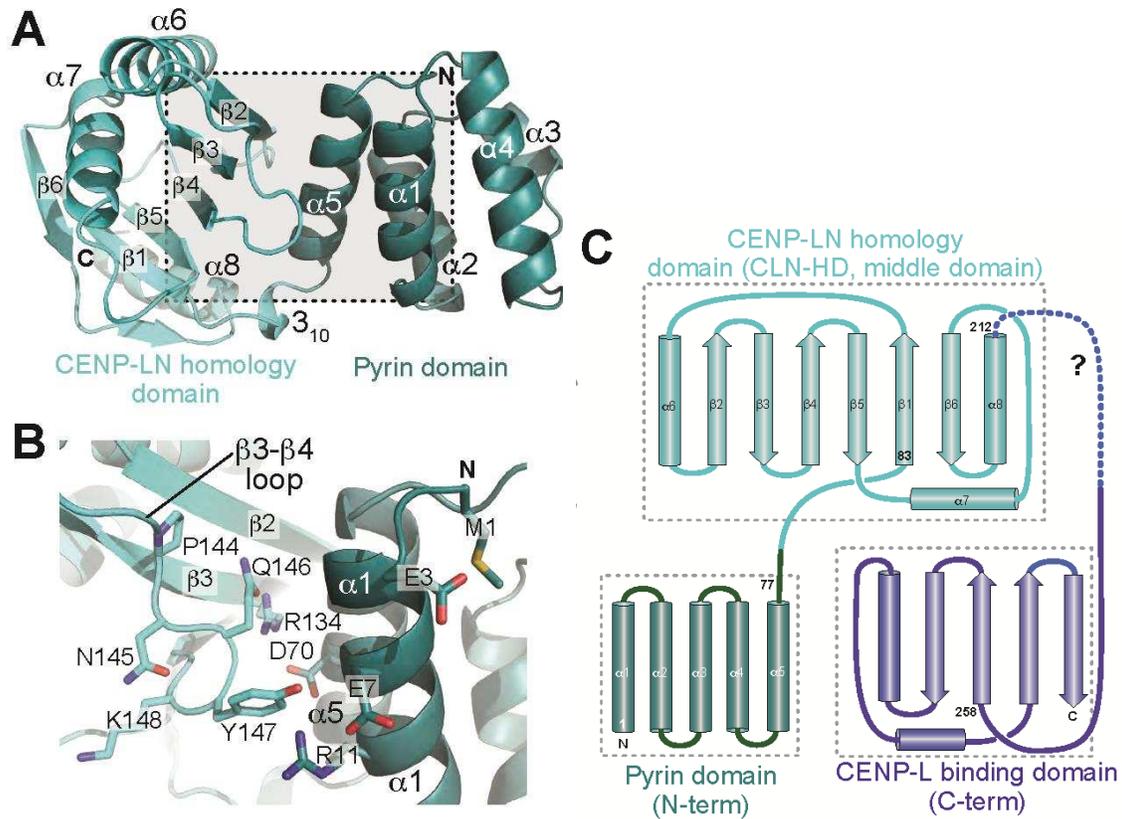


Figure 2.2 Crystal structure of the CENP-A binding region of CENP-N

A. Cartoon model of CENP-N¹⁻²³⁵ with secondary structure and domain organization. B. Close-up of the boxed region in A. C. Topology diagram of CENP-N. The topology of the Pyrin and CLN-HD domains was directly derived from the crystal structure of CENP-N¹⁻²³⁵ reported here. The topology of the CENP-L binding domain was derived from the crystal structure of the Chl4^{CENP-N}:Im13^{CENP-L} yeast homolog (Hinshaw and Harrison, 2013).

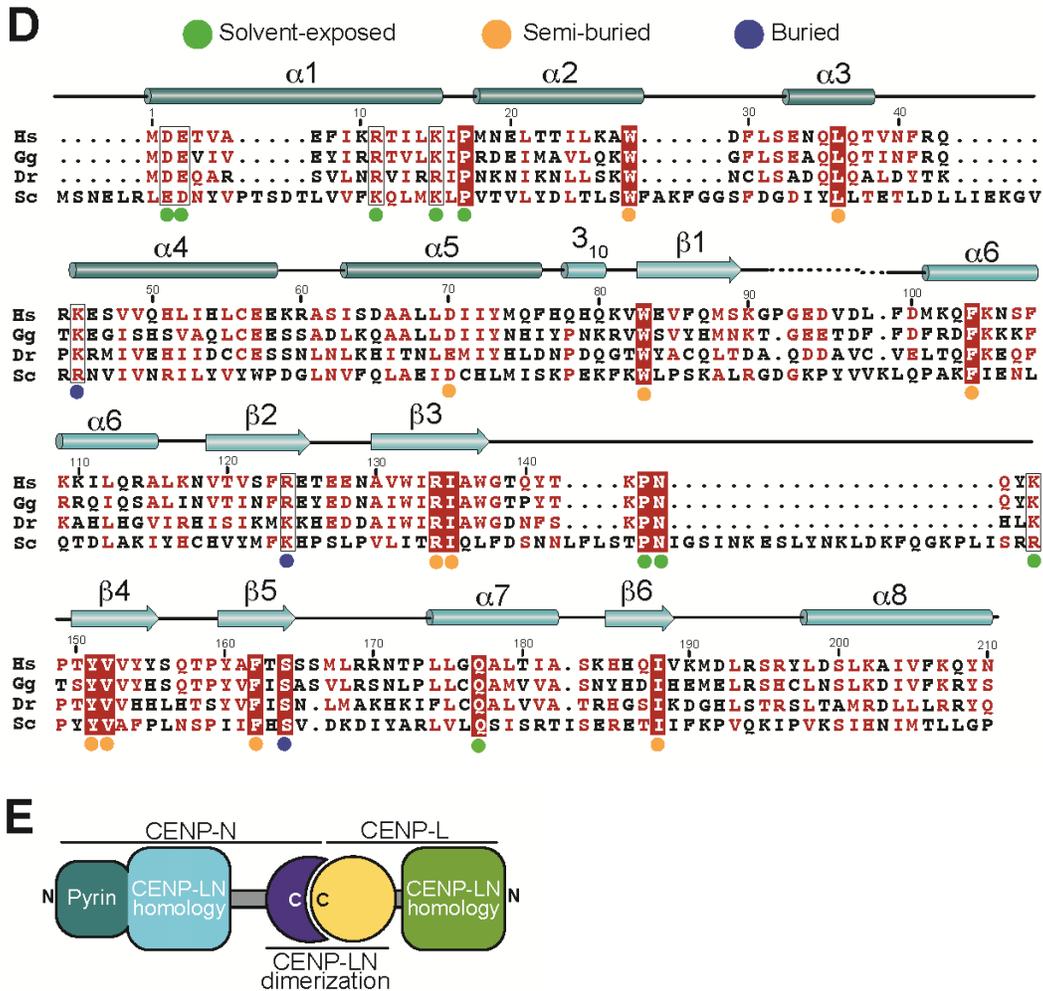


Figure 2.2 Crystal structure of the CENP-A binding region of CENP-N

D. Multiple sequence alignment of CENP-N from the indicated species with secondary structure. Green, blue, and orange dots indicate solvent-exposed, semi-buried, and buried side chains, respectively. Positions with conserved residues are displayed red; positions with conserved side chain charge are boxed. E. Schematic summarizing domain organization of CENP-L, CENP-N, and their dimerization.

Iml3 and Chl4 are fungal orthologs of CENP-L and CENP-N, respectively. We referred to a previously reported crystal structure of the full-length Iml3 protein bound to the C-terminal region of Chl4 (Iml3:Chl4^C, PDB ID 4JE3) (Hinshaw and Harrison, 2013) to deduce the structural organization of the human CENP-LN complex. Iml3 consists of an N-terminal domain (shown in green in Figure 2.2 Supplement 2 A) and a C-terminal domain (the 'insert' domain shown in yellow; the topology of Iml3 is shown in Figure 2.2 Supplement 2 B). Iml3 hetero-dimerizes with Chl4 through a subdomain within the insert domain (Figure 2.2 Supplement 2 B) (Hinshaw and Harrison, 2013). Due to strong sequence similarity of Iml3 and CENP-L throughout their length (not shown), the structure of Iml3 provides an excellent model for the structure of CENP-L. Importantly, although our crystal structure does not encompass the C-terminal region of CENP-N, the sequence of the latter is strongly related to that of Chl4^C (Figure 2.2 Supplement 3 A), which was captured in complex with Iml3 in the Iml3:Chl4^C structure, indicating that they are also structurally related (Figure 2.2 C). Indeed, as already observed (Guo et al., 2017; Hinshaw and Harrison, 2013), the C-terminal region of CENP-N (CENP-N^{230-C}) was sufficient to interact with CENP-L (Figure 2.1 Supplement 1). Thus, the structure of CENP-N¹⁻²³⁵ reported here and that of the Iml3:Chl4^C complex are complementary, and together provide an almost comprehensive view of the CENP-L^{Iml3}:CENP-N^{Chl4} complex (Figure 2.2 Supplement 3 A and Figure 2.2 Supplement 3 B). Besides identifying the N-terminal domain of CENP-N¹⁻²³⁵ as a PYRIN domain, DALI also identified an unanticipated structural homology of the second domain of CENP-N¹⁻²³⁵ with the N-terminal domain of Iml3^{CENP-L} (Figure 2.2 Supplement 3 C). We therefore refer to these domains of CENP-N and CENP-L as CLN-HD (for CENP-L and CENP-N homology

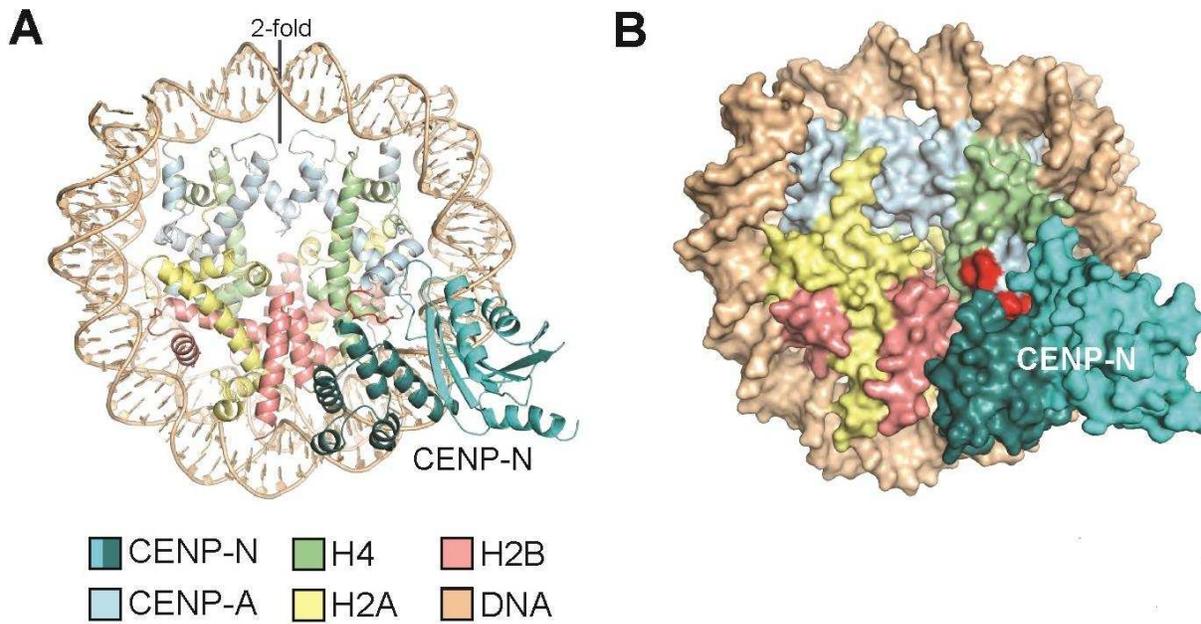


Figure 2.3 The CENP-N:CENP-A^{NCP} complex

A. Cartoon model of the CENP-A^{NCP} with bound CENP-N¹⁻²³⁵, determined by cryo-EM. B. Surface representation of the complex. In A and B, the L1 loop of CENP-A is displayed in red.

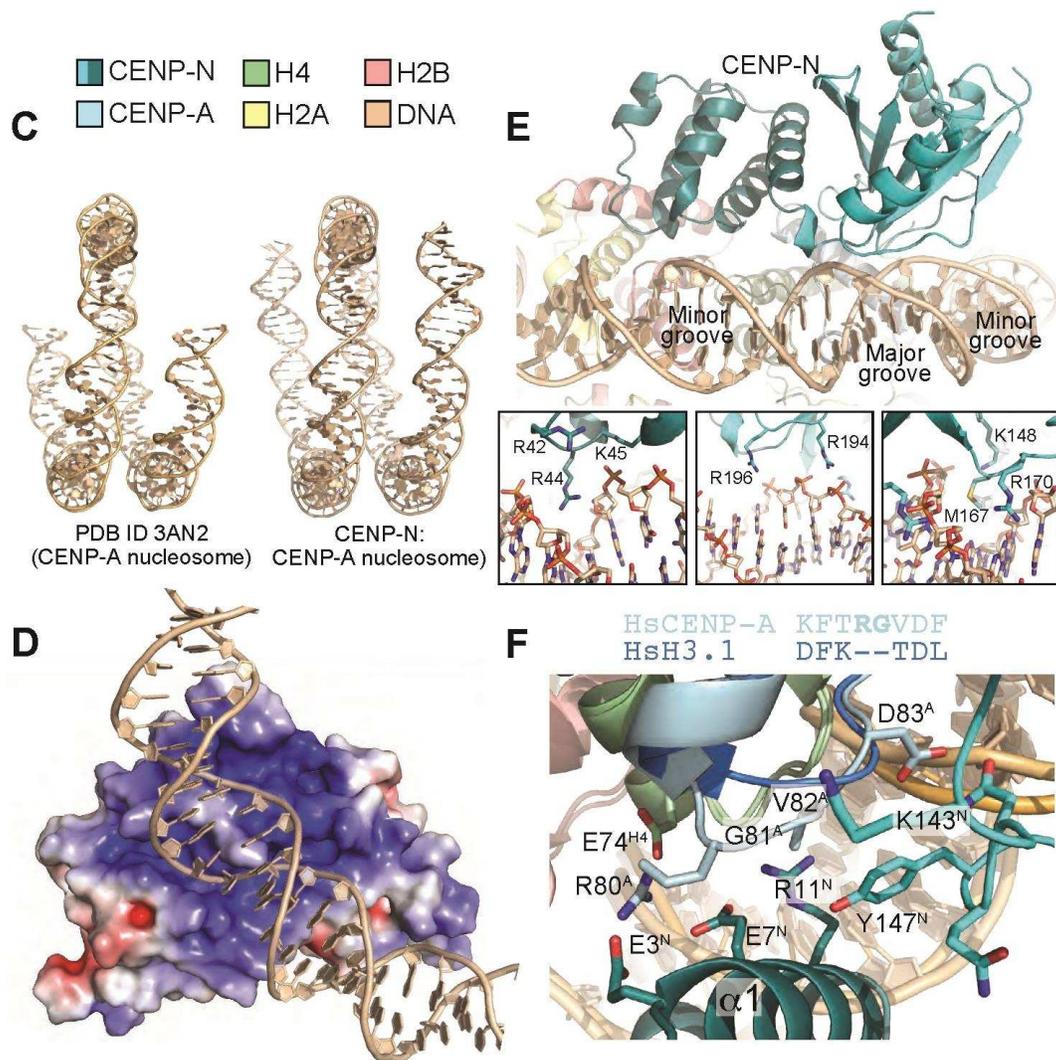


Figure 2.3 The CENP-N:CENP-A^{NCP} complex

C. Comparison of the DNA ends in the crystal structure of the CENP-A nucleosome (Tachiwana et al., 2011) and in the structure of the CENP-A:CENP-N complex. D. Electrostatic potential at the CENP-N DNA binding interface with contour levels $\pm 4 k_B T/e$ (k_B , Boltzmann constant; T, absolute temperature; e, the magnitude of electron charge, calculated with the APBS Pymol plugin). E. Interaction of CENP-N with backbone, minor groove, and major groove of DNA with close-up views of selected interactions. F. Interactions at the CENP-A L1 loop and comparison with superimposed H3.

Table 2.2**EM data collection, processing, and refinement statistics**

Data collection and processing	
Voltage (kV)	300
Magnification	290,000x
Defocus (μm, nominal)	-1.0 to -2.5
Pixel size (\AA)	1.02
Electron dose rate (counts/pixel/s)	10
Total electron dose ($\text{e}^- / \text{\AA}^2$)	80
Exposure time (s)	8
Number of images (collected/processed)	3900/3024
Number of frames per image	40
Initial particle number	1,843,269
Particle number for 3D classification	1,267,674
Final particle for refinement	937,118
Resolution (masked/unmasked) (\AA)	4.0/4.2
Map sharpened b-factor (\AA^2)	-233
Model refinement	
r.m.s. deviation (bonds)	0.005
r.m.s. deviation (angles)	0.97
All-atom clashscore	2.30
Ramachandran plot	
Outliers (%)	0.00
Allowed (%)	4.59
Favored (%)	95.81
CaBLAM analysis:	
Outliers (%)	1.92
Disfavored (%)	6.65
Ca outliers (%)	0.11
Rotamer outliers (%)	0.00

domain). Structural similarities of the CLN-HD suggest that CENP-N and CENP-L are evolutionary related. However, sequence identity of the two domains, even after structural superposition, is minimal, likely explaining why structural similarity had not been predicted (Figure 2.2 Supplement 3 D). CENP-L, or its complex with CENP-N230-C, did not interact with CENP-ANCPs or H3NCPs (Figure 2.1 E and Figure 2.1 Supplement 1). Thus, CENP-L and CENP-N, even if partly structurally related, have clearly distinct functions. In conclusion, the structure of CENP-N contains an N-terminal Pypin domain, a central CLN-HD, and a C-terminal CENP-L dimerization domain, while CENP-L contains an N-terminal CLN-HD, interrupted immediately before the C-terminal helix by an insertion that contains a region required for CENP-N dimerization.

Cryo-EM analysis of the CENP-N:CENP-A nucleosome complex

Using cryo electron microscopy (cryo-EM), we obtained a three-dimensional reconstruction of CENP-N¹⁻²⁸⁹ bound to CENP-A^{NCPs} at ~4.0 Å (Figure 2.3 A-B, Figure 2.3 Supplement 1, and Table 2.2). We built an atomic model of the CENP-N:CENP-A^{NCP} complex by fitting into the EM density high-resolution models of the CENP-A histone core (PDB ID 3AN2) (Tachiwana et al., 2011), combined with DNA derived from a nucleosome reconstituted with the 145-bp 601 DNA sequence (PDB ID 3LZ0; Vasudevan et al., 2010), and the newly determined crystal structure of CENP-N¹⁻²³⁵. Both manual and automatic fitting strategies produced unequivocal fits, allowing the first visualization of the interaction of CENP-N with the CENP-A nucleosome (Figure 2.3 Supplement 2).

The CENP-A nucleosome appears to be stabilized by its interaction with CENP-N (Guo et al., 2017). There is clear density for 139 of the 145 bp of DNA and for the N-terminal helix of CENP-A (Figure 2.3 C and Figure 2.3 Supplement 2 A), two features reported to be largely disordered and thus invisible in the crystal structure of the CENP-A^{NCP} (PDB ID 3AN2) (Tachiwana et al., 2011). CENP-N, whose structure changes very little upon binding to the CENP-A nucleosome, is positioned on top of the L1 loop of CENP-A (also called RG loop for the presence of a conserved arginine-glycine motif at the loop's apex) and contacts approximately 15-bps of the adjacent DNA gyre (Figure 2.3 A). There is clear density only until CENP-N¹⁻²⁸⁹ residue ~210, indicating that the following approximately 80 C-terminal residues (at the opposite end of the nucleosome interaction interface) may be flexible. Of the ~2400 Å² of CENP-A^{NCP} and CENP-N surface area that become buried in the complex, ~1400 Å² are at the CENP-N:DNA interface, where both CENP-N domains form extensive interactions with DNA from bp -21 to -35 relative to the 2-fold axis, or superhelical location [SHL] -2 to -3. There is a marked accumulation of positively charged residues on this DNA binding interface (Figure 2.3 D). Four loops in the CLN-HD straddle the DNA double helix over ~8 bp, and the consecutive 7 bp are bound by the PYRIN domain, which is positioned to insert an arginine (R44) into the compressed minor groove in an arrangement that is reminiscent of the minor groove arginines inserted by the histones (Figure 2.3 E). The highly conserved P17 in the PYRIN domain positions the main chain of CENP-N to latch on to the phosphate backbone of the DNA, with interactions made through the side chains of K15, R42, K45, K81, and R194. There are also likely insertions of CENP-N side chains into two minor grooves (besides R44, also K148, M167, R170) and the intervening

major groove at SHL -3 [R196, see also (Carroll et al., 2009)] (Figure 2.3 E). In agreement with the presence of a large interaction interface with nucleosomal DNA, CENP-N bound more tightly to CENP-A^{NCPs} but retained substantial binding affinity for H3^{NCPs} in electrophoretic mobility shift assays (EMSAs) (Figure 2.3 Supplement 3 A-B). Likely, this residual binding to H3^{NCPs} in the EMSAs, which emerged less clearly in solid phase binding assays (Figure 2.1 C), reflects emphasis on electrostatic interactions under the low salt conditions of the EMSA assays (see Methods), as also discussed in the context of Figure 2.4 Supplement 2.

The CENP-N:CENP-A interface

The substantial interface with DNA explains why CENP-N does not bind CENP-A:H4 tetramers lacking DNA (Carroll et al., 2009). However, while DNA binding clearly contributes to the binding affinity of this interaction, it is unlikely to contribute to the discrimination of CENP-A^{NCPs} from H3^{NCPs}, because CENP-N bound selectively to CENP-A^{NCPs} even when the CENP-A^{NCPs} and H3^{NCPs} contained the same DNA sequence (Figure 2.1 B-C). Conversely, the structure clearly suggests why recognition of the L1 loop is crucial for discrimination (Black et al., 2004; Carroll et al., 2010; Carroll et al., 2009). CENP-N binds the L1 loop through a continuous interface comprising the α 1 helix in CENP-N^{PD} and the β 3- β 4 loop in CENP-N^{CLN-HD}. Several of the infrequent conserved solvent-exposed residues of CENP-N (identified by a green dot in Figure 2.2 D), including E3, E7, R11, K143, P145, N146, and K148 reside in this interface. Y147, which is less conserved, contributes to the stabilization of the relative arrangements of the CENP-N^{PD} and CENP-N^{CLN-HD}, which is largely unchanged in the nucleosome-

bound and free structures of CENP-N. Insertion of the side chain of M1 into the hydrophobic core contributes to the stabilization of the $\alpha 1$ helix. The interaction with the CENP-A L1 loop engages a triad of residues, E3^{CENP-N}, E7^{CENP-N}, and R11^{CENP-N}, whose side chains emerge from the same face of the $\alpha 1$ helix looking towards the L1 loop (Figure 2.3 F).

The CENP-A residues R80^{CENP-A} and G81^{CENP-A} form a 2-residue insertion that is the most conspicuous difference between the L1 loops in CENP-A and H3 (Figure 2.3 F, Figure 2.3 Supplement 4 A-B). The insertion is crucial, because it allows R80^{CENP-A} to form hydrogen bonds with both E3^{CENP-N} and E7^{CENP-N}, while absence of a side chain at G81^{CENP-A} allows the CENP-A loop to insert deeply into a cleft formed between the two CENP-N domains, where the side chain of Y147^{CENP-N} packs tightly against V82^{CENP-A}. In EMSAs, mutation of R80 and G81 to alanine partly ablated the preference of CENP-N for CENP-A^{NCPs} (Figure 2.3 Supplement 3 A-B). The side chain of R11^{CENP-N}, a residue previously shown to be important for the CENP-N:CENP-A^{NCP} interaction (Carroll et al., 2009), on the other hand, is squeezed between the loop 1 region of CENP-A and the loop 2 region of H4, where it may be involved in a double salt bridge with E74^{H4} and E7^{CENP-N} (Figure 2.3 F and Figure 2.3 Supplement 5).

Mutational validation of the CENP-N:CENP-A^{NCP} structure

We generated a collection of single and double alanine point mutants to probe the role of individual CENP-N residues in the interaction with the CENP-A^{NCP}. In pull-down

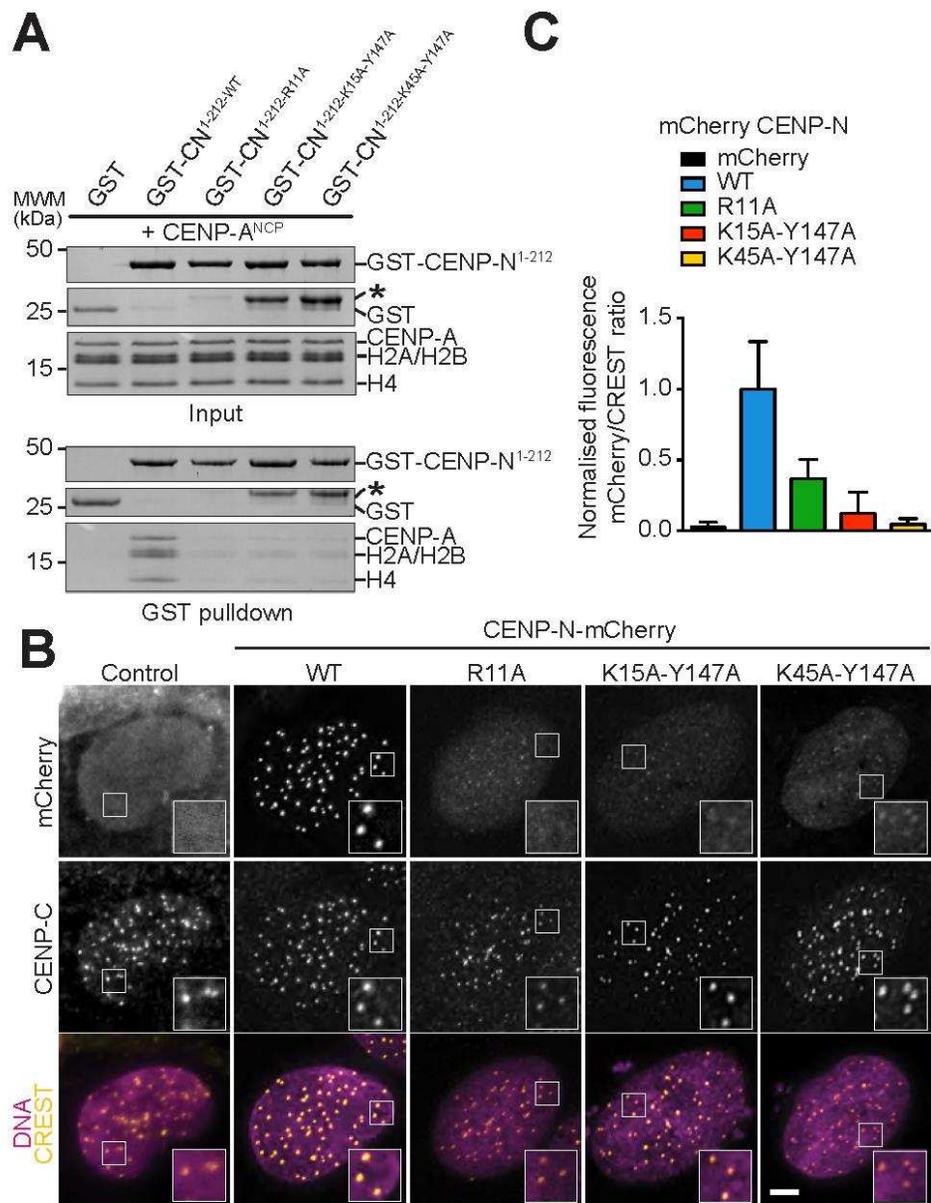


Figure 2.4 Validation of the CENP-N:CENP-A^{NCP} complex

A. *In vitro* binding assay probing the interaction of GST-CENP-N¹⁻²¹² immobilized on solid phase with CENP-A^{NCP}. B. Fluorescence microscopy analysis comparing localization at human kinetochores (U2OS osteosarcoma cells) of a wild type CENP-N-mCherry fluorescent reporter and of its mutant variants. C. Quantification of localization of the mCherry constructs in B normalized to CREST.

assays *in vitro*, we found essentially complete loss of binding with an alanine (A) mutant of R11 (Figure 2.4 A), and substantial reductions of binding with alanine mutants of E7 or Y147, at the CENP-A L1 interface, or of K15 or K45, at the interface with DNA (Figure 2.4 Supplement 1 A). Combining mutations of Y147 with either K15 or K45 almost completely disrupted CENP-A^{NCP} binding (Figure 2.4 A), in line with the idea that recognition of the L1 loop and of the DNA jointly contribute to the binding affinity of CENP-N for the CENP-A nucleosome. CENP-N targeting to centromeres in U2OS cells reflected the observations made *in vitro*, with R11A single mutant and the K15A-Y147A and K45-Y147A double mutants appearing severely impaired in the ability to target centromeres (Figure 2.4 B-C), and other single mutants suffering intermediate effects on binding to centromeres (Figure 2.4 Supplement 1 B).

In competition gel shift assays, CENP-N mutants at the CENP-A L1 loop interface (L1 mutants), including R11A and two double mutants (E3A-E7A and K143A-Y147A), lost the ability of CENP-N^{wt} to discriminate between CENP-A and H3-nucleosomes (Figure 2.4 Supplement 2 A-B). We have shown in Figure 2.4 A that CENP-N^{R11A} does not bind CENP-A^{NCPs} in solid phase binding assays. The residual interaction of this mutant with CENP-A^{NCPs} or H3^{NCPs} in EMSAs likely reflects the extensive binding interface of CENP-N for nucleosomal DNA (whose effects are emphasized under low salt conditions, as already indicated for CENP-N^{wt} in the context of Figure 2.3 Supplement 3). In line with this interpretation, we find that in the EMSAs the L1 mutants of CENP-N bind H3^{NCPs} indistinguishably from CENP-N^{wt}, whereas the same mutants bind to CENP-A^{NCPs} considerably worse than CENP-N^{wt} (Figure 2.4 Supplement 2). Collectively, these

results further emphasize the importance of the L1 loop of CENP-A in selective recognition by CENP-N.

Identification of a CENP-C region involved in CENP-LN binding

As discussed in the Introduction, CENP-C, an intrinsically disordered protein, provides a defined succession of binding sites for different kinetochore components (Klare et al., 2015) (Figure 2.1 A). CENP-C, CENP-LN, and CENP-HIKM, another kinetochore sub-complex located in the vicinity of CENP-A, form a 7-subunit complex (designated CHIKMLN) that binds CENP-A^{NCPs} cooperatively, i.e. with increased binding affinity in comparison to any of the individual subunits or sub-complexes (McKinley et al., 2015; Weir et al., 2016). Within this assembly, CENP-C²⁻⁵⁴⁵ binds the CENP-LN complex *in vitro* [Figure 2.5 Supplement 1 A and (Hinshaw and Harrison, 2013; McKinley et al., 2015; Nagpal et al., 2015; Weir et al., 2016)].

We set out to exploit biochemical reconstitution and our improved structural understanding of the CENP-LN complex to query the importance of this interaction for kinetochore assembly in humans. Trimming of CENP-C²⁻⁵⁴⁵ identified CENP-C²²⁵⁻³⁶⁴ as a minimal CENP-LN interaction domain (Figure 2.5 A), in line with a recent study (Guo et al., 2017). Neither CENP-L nor CENP-N¹⁻²³⁵ bound to CENP-C²⁻⁵⁴⁵ (Figure 2.5 Supplement 1 B-C). However, the CENP-LN^{230-C} dimer bound CENP-C in the absence of nucleosomes (Figure 2.5 B). Thus, CENP-C²²⁵⁻³⁶⁴ binds at or near the CENP-LN dimer interface, possibly also exploiting structural ordering of these regions upon

dimerization. CENP-C²²⁵⁻³⁶⁴ contains a handful of conserved residues, some of which were previously shown to mediate an interaction with the CENP-HIKM complex (Klare et al., 2015) (Figure 2.5 C). We probed an additional conserved linear motif in CENP-C²²⁵⁻³⁶⁴ (residues 302-306) for its potential role in CENP-N binding. A 5-alanine mutant of residues 302-306 (identified as CENP-C^{5A}) failed to interact with CENP-NL, identifying this region of CENP-C as the CENP-LN binding motif (Figure 2.5 D and Figure 2.5 Supplement 2). Importantly, CENP-C^{2-545-5A} did not interact with CENP-LN, but retained binding to CENP-A^{NCPs} and CENP-HIKM (Figure 2.5 E-F). In isothermal titration calorimetry (ITC) experiments, CENP-LN^{230-C} bound CENP-C²²⁵⁻³⁶⁴ with a dissociation constant (K_D) of 1 μ M, but it showed no binding to CENP-C^{225-364-5A} (Figure 2.5 G-H).

CENP-LN binding motif of CENP-C is required for kinetochore recruitment of CENP-N

The availability of a CENP-LN binding mutant of CENP-C gave us an opportunity to ask if the interaction of CENP-LN with CENP-A, besides being necessary, is also sufficient for kinetochore recruitment of CENP-N. For this, we depleted CENP-C by RNAi and replaced it with exogenous wild type (wt) or mutant (5A) copies. Depletion of CENP-C prevented kinetochore localization of CENP-N, showing that nucleosome binding is not sufficient for CENP-N to reach kinetochores at the low cellular concentration of these proteins. Exogenously expressed wild type CENP-C promoted CENP-N recruitment, while CENP-C^{5A} failed to promote it (Figure 2.6 A-B). Thus, the CENP-LN binding site of CENP-C, while not crucial for CENP-C recruitment to kinetochores, is instead crucial for

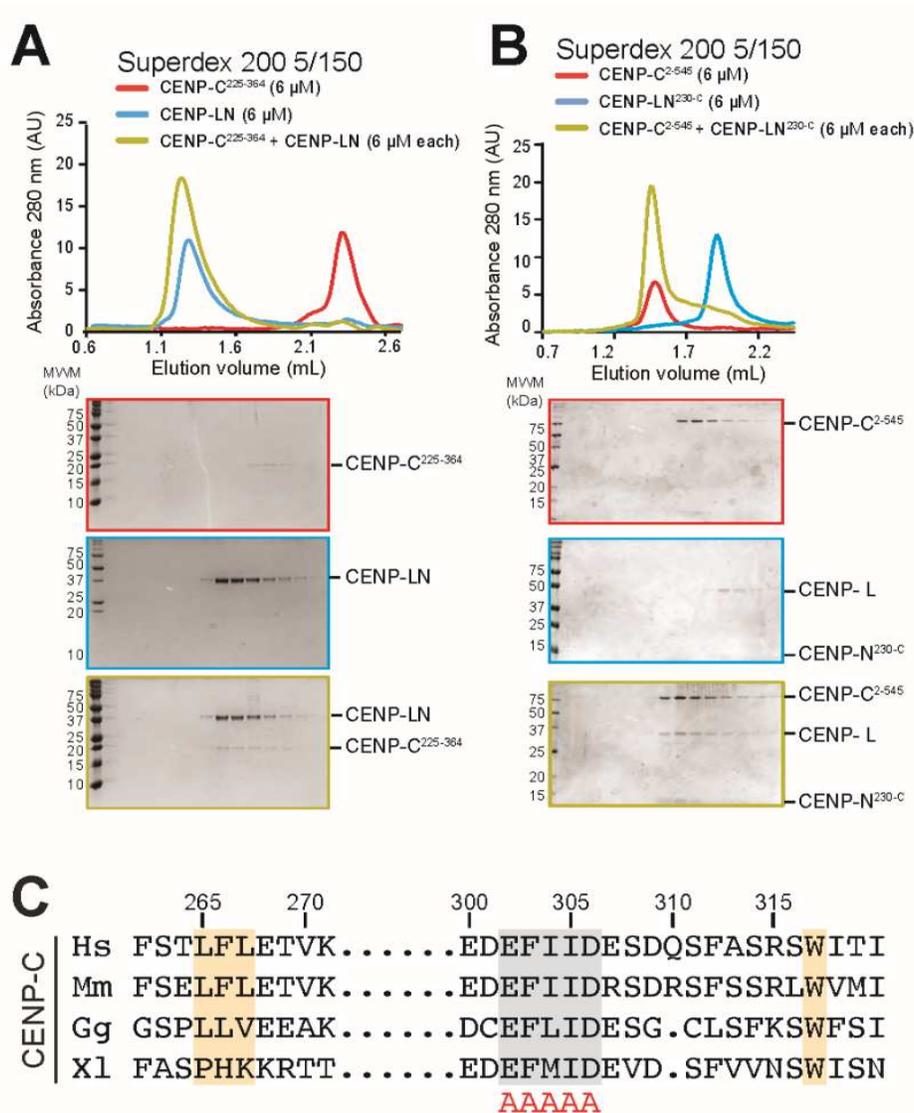


Figure 2.5 Identification of a CENP-N binding site on CENP-C

A) Size exclusion chromatography (SEC) runs of CENP-C²²⁵⁻³⁶⁴, CENP-LN complex, and their combination at the indicated loaded concentrations, identified a binding site for CENP-LN in CENP-C²²⁵⁻³⁶⁴. Elution fractions were separated by SDS-PAGE and visualized by Coomassie staining. B) CENP-LN^{230-C} binds CENP-C²⁻⁵⁴⁵, indicating that the CENP-N N-terminal region is not required for CENP-C binding. C) Sequence of a segment of the PEST-rich domain CENP-C that contains a binding site for the CENP-HIKM complex (Klare et al., 2015) (residues indicated in salmon).

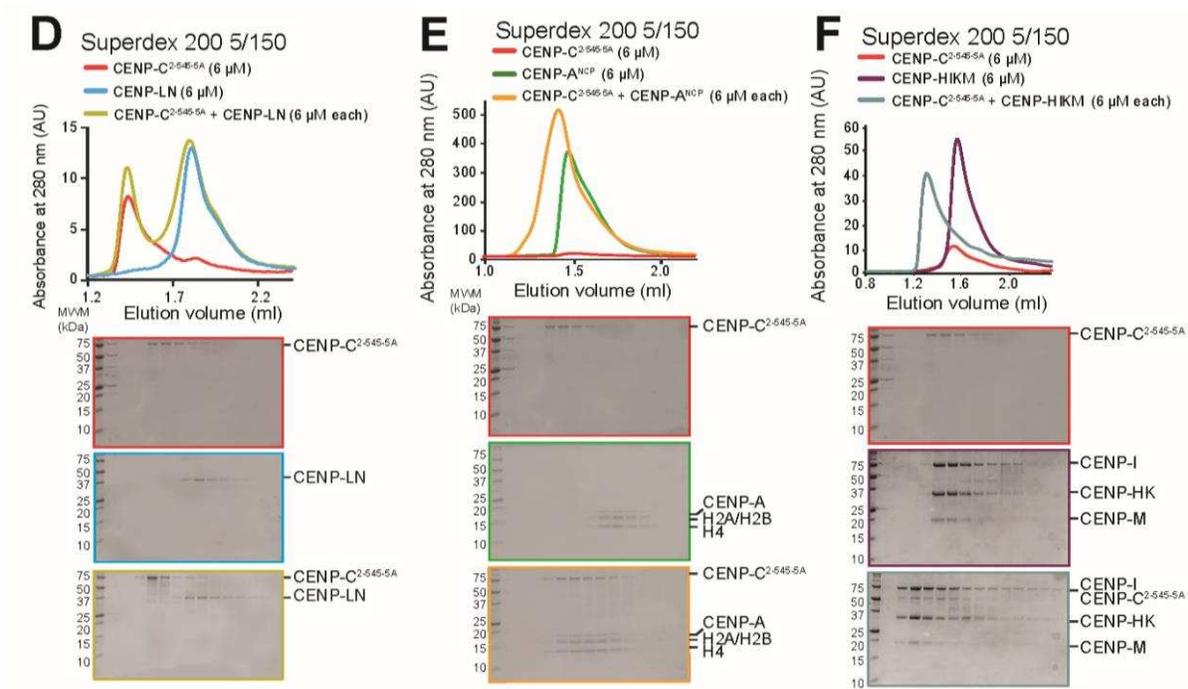


Figure 2.5 Identification of a CENP-N binding site on CENP-C

D. CENP-LN does not bind CENP-C^{2-545-5A}. E. CENP-C^{2-545-5A} retains the ability to bind to CENP-A^{NCP}. F. CENP-C^{2-545-5A} retains the ability to bind to the CENP-HIKM complex.

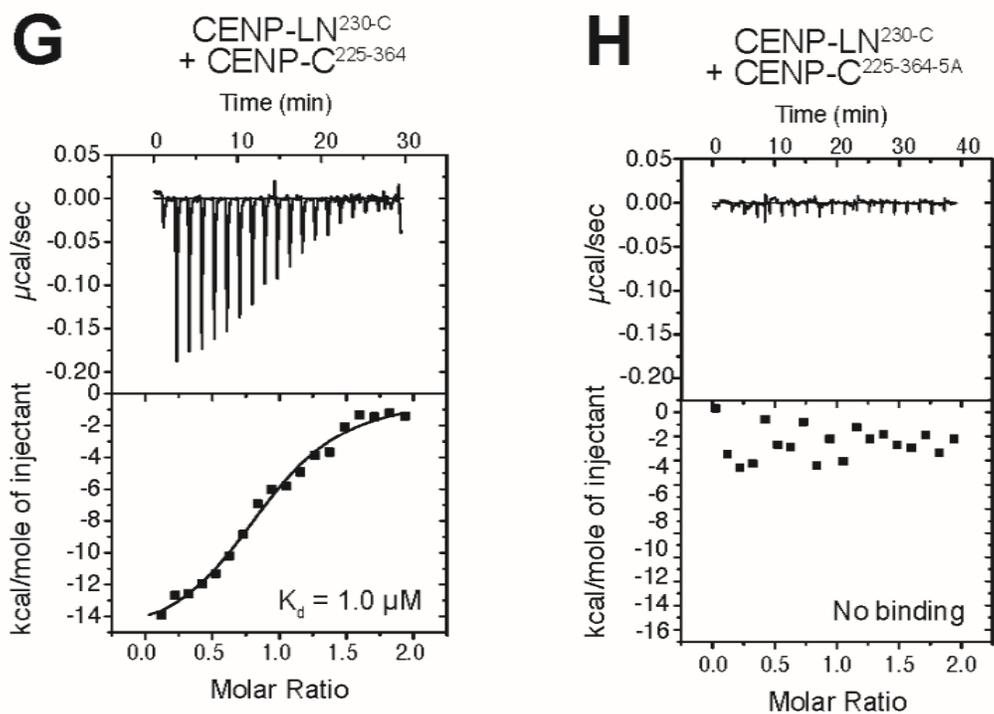


Figure 2.5 Identification of a CENP-N binding site on CENP-C

G. Isothermal titration calorimetry (ITC) experiment quantifying the physical interaction of the CENP-L:CENP-N^{230-C} complex with CENP-C²²⁵⁻³⁶⁴. H. In agreement with the SEC data, CENP-C^{225-364-5A} fails to interact with the CENP-L:CENP-N^{230-C} complex in an ITC experiment.

CENP-N recruitment. Overall, these observations indicate that the CENP-LN complex reads the presence of two features of kinetochores, the presence of CENP-A and the presence of CENP-C, both of which are necessary for its efficient recruitment.

Discussion

The histone H3 variant CENP-A is an essential feature of centromeres and has two main functions. First, it is required for kinetochore assembly through its direct interactions with inner kinetochore subunits that can then seed the assembly of this large macromolecular assembly. Second, it is a landmark that determines the stability of centromere chromatin identity through cell division. Interactions of CENP-C and CENP-N with CENP-A^{NCPs} are the only known direct and specific points of contact of the kinetochore with the centromere and are therefore the crucial effectors through which CENP-A implements its role (Carroll et al., 2010; Carroll et al., 2009; Kato et al., 2013).

While the structural basis of CENP-C binding to CENP-A had been described (Kato et al., 2013), how CENP-N binds CENP-A had remained elusive. Here, we have filled this important gap, as shown schematically in Figure 2.6 D-E. CENP-N nucleosome binding differs from that observed with RCC1 and other nucleosome binders that engage primarily an exposed acidic patch on histones H2A and H2B (Makde et al., 2010). However, in its outline it resembles the interaction of the ATPase domain of SWI2/SNF2 chromatin remodeler with H3^{NCPs} (Farnung et al., 2017; Liu et al., 2017; Narlikar et al.,

2013), with the important difference that SWI2/SNF2 closely approaches H3 without making significant direct contacts with it, whereas CENP-N interacts directly with CENP-A (Figure 2.3 Supplement 6 A-D). There are also similarities with the nucleosome binding mechanism of the bromo-adjacent homology (BAH) domain of Sir3 (PDB ID 3TU4) (Armache et al., 2011), but the latter interacts predominantly with the H4 N-terminal tail, through recognition of K16^{H4}, and with the acidic patch on H2A-H2B, and much less extensively with DNA (Figure 2.3 Supplement 6 E-F). In the CENP-N:CENP-A^{NCP} complex, the normally disordered N-terminal tail of H4 is ordered until R23 and interacts weakly with the CENP-N loop connecting β 3 with β 4 (Figure 2.3 Supplement 2 D). The reported mono-methylation of K20 of H4 in the CENP-A nucleosome (Hori et al., 2014) may further modulate this interaction. In summary, the SWI2/SNF2 and BAH modes of nucleosome binding are predominantly based on interactions with DNA or with histones, respectively, while CENP-N shows a balance of both. The considerable interaction of CENP-N with DNA is a remarkable and unexpected feature of the complex structure.

CENP-C and CENP-N can interact concomitantly with the same CENP-A nucleosome (Carroll et al., 2010), as also confirmed in recent studies (Guo et al., 2017; Weir et al., 2016). The central motif and the CENP-C motif of CENP-C, which confer CENP-A recognition ability *in vitro*, interact through a 'arginine anchor' with the acidic patch of H2A and H2B, and also decode the divergent C-terminal tail of CENP-A (Kato et al., 2013). These determinants of CENP-C binding on CENP-A are located adjacent to, but not overlapping with, the CENP-N binding footprint. Indeed, when modelled on the

CENP-N:CENP-A^{NCP} structure according to the position it adopts in its structure with the nucleosome (PDB ID 4X23) (Kato et al., 2013), the CENP-C motif can be accommodated without steric clashes (Figure 2.6 C,E). Thus, CENP-C and CENP-N interact with CENP-A through complementary interfaces. In the context of a larger CCAN complex, these CENP-A binding motifs cooperate to increase the overall binding affinity for CENP-A (Guo et al., 2017; Weir et al., 2016).

It has been proposed that CENP-N is significantly stabilized upon binding to the CENP-A nucleosome (Guo et al., 2017). Our study did not identify a clear structural basis for this phenomenon, as we failed to identify significant conformational changes in CENP-N in isolation (crystal structure) or in its complex with CENP-A^{NCP}. It has also been proposed that CENP-C reshapes and rigidifies the CENP-A nucleosome and that it modulates the DNA termini to make them match the loose wrap observed at centromeres (Falk et al., 2015). Importantly, these effects of CENP-C binding on the CENP-A nucleosome do not appear to be required for the selective (over H3) interaction of CENP-N, because selectivity for CENP-A was retained in the absence of CENP-C [this study and (Carroll et al., 2009; Weir et al., 2016)]. We also note that the DNA termini appear to be well defined in our structure of the CENP-N:CENP-A^{NCP} complex, contrarily to what was observed in the structure of the isolated CENP-A^{NCP} (PDB ID 3AN2). At present, we cannot definitively conclude whether the stabilization of the termini is due to CENP-N binding to the CENP-A nucleosome, as we have not yet been able to obtain a high-resolution EM structure of the CENP-A nucleosome in isolation for

comparison. It is possible that the cryogenic conditions used for our structural work stabilize a specific conformation of the complex.

In most organisms, centromere identity is not specified by the centromere's DNA sequence, but rather by the enrichment of CENP-A at a defined chromatin domain. *De novo* formation of stably inherited centromeres at previously non-centromeric sites (neo-centromeres) provides clear evidence in favour of this idea. Thus, rather than being genetically (i.e. DNA-sequence) specified, centromeres are epigenetically specified, with the pre-existing enrichment of CENP-A being a necessary condition for continued deposition of new CENP-A at the same site through the generations. There is therefore considerable interest in the molecular mechanisms that promote new CENP-A deposition at centromeres during the cell cycle, and in the mechanisms that promote the stabilization and persistence of CENP-A after its incorporation at centromeres.

Conserved machinery for new CENP-A deposition, including the specialized CENP-A chaperone HJURP (Scm3 in *S. cerevisiae*) and an adaptor complex consisting of the Mis18 and M18BP1 subunits, has been described in recent years (Dunleavy et al., 2009; Foltz et al., 2009; Fujita et al., 2007; Hayashi et al., 2004; Pidoux et al., 2009; Sanchez-Pulido et al., 2009; Williams et al., 2009). Additional machinery, in particular chromatin remodelling enzymes harnessing a source of energy to evict H3, is likely involved in the reaction but has not been univocally identified. This machinery is recruited to centromeres early during the cell-cycle and is believed to promote the replacement of

histone H3 with new CENP-A (Dunleavy et al., 2011; Jansen et al., 2007; Schuh et al., 2007). Likely, the existing CENP-A nucleosome acts as a template in this reaction, as the abundance of CENP-A nucleosomes at a given centromere is, at least in first approximation, constant through subsequent cell divisions (French et al., 2017; Hori et al., 2017; Jansen et al., 2007). This implies that the same number of new CENP-A nucleosomes is incorporated after each cell division as that of originally present CENP-A nucleosomes, suggesting that the deposition machinery targets for H3 eviction and replacement with CENP-A an H3 nucleosome that is likely in close proximity of the CENP-A nucleosome (Musacchio and Desai, 2017). While the mechanistic details of CENP-A deposition remain partly unclear, there is now substantial evidence that recruitment of the CENP-A deposition machinery requires CENP-C and possibly other CCAN factors (Dambacher et al., 2012; Moree et al., 2011; Shono et al., 2015). While CENP-N has not been directly implicated in CENP-A deposition, our observation that CENP-N occupies a region of the nucleosome required for binding by chromatin remodelling enzymes of the SWI2/SNF2 family suggests that CENP-N may protect centromeric nucleosomes from remodelling and eviction, thereby contributing to its stability. Indeed, both CENP-C and CENP-N contribute to the stabilization of newly incorporated CENP-A in centromeric chromatin (Guo et al., 2017). New CENP-N deposition at centromeres occurs in late S phase (Fang et al., 2015; Hellwig et al., 2011), and may trigger a stabilization of centromere organization required for successful kinetochore assembly.

In summary, our analysis of the mechanisms of the interactions of the CENP-NL complex with CENP-A and CENP-C represents a step forward in the molecular dissection of the almost universally conserved functions of CENP-A in eukaryotes, which are required for accurate chromosome segregation and, more generally, for the success of cell division.

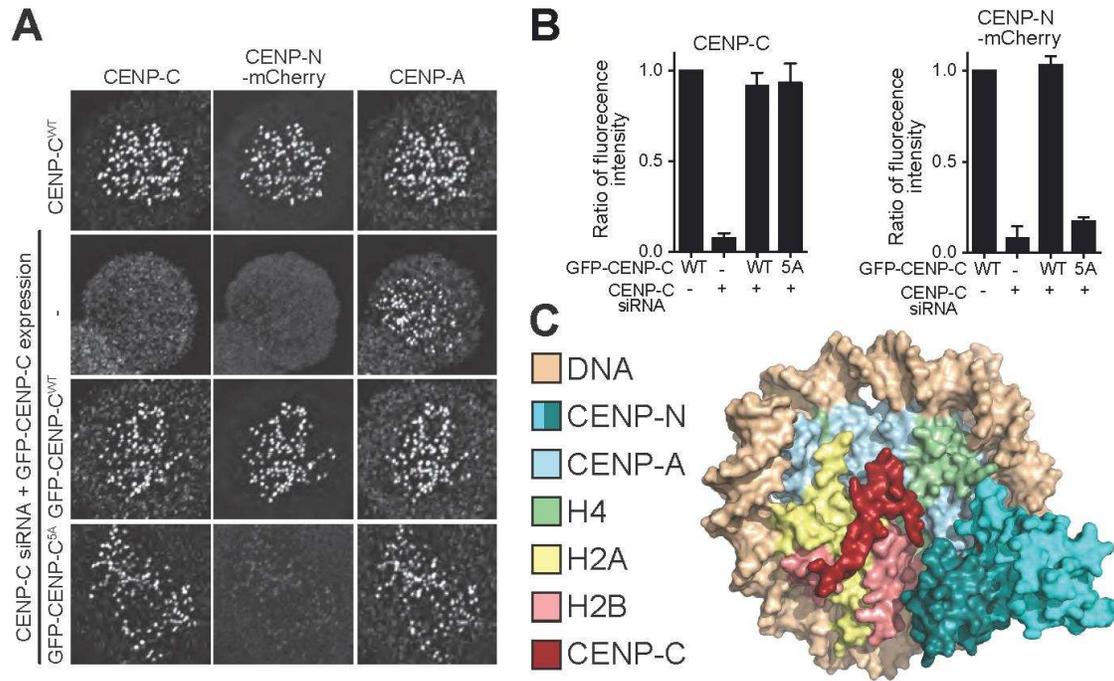


Figure 2.6 Effective CENP-N localization requires CENP-C

A. Fluorescence microscopy analysis comparing kinetochore localization of a wild type CENP-N-mCherry fluorescent reporter in human HeLa FlpIn TRex cells depleted of CENP-C and, where indicated, further expressing wild type GFP-CENP-C or the 5A mutant. B. Quantification of CENP-C (left) and mCherry-CENP-N (right) levels at kinetochores in mitotic cells following the rescue of CENP-C depletion by either GFP-CENP-C^{WT} or the GFP-CENP-C^{5A} mutant. Graphs show kinetochore fluorescence intensity of the indicated protein (antibodies against CENP-C or mCherry) normalized to CENP-C or mCherry-CENP-N kinetochore levels in the absence of RNAi treatment, respectively. Each graph is representative of two independent experiments. C) Surface representation of a composite model built by combining the coordinates of the CENP-C motif (residues 712 to 733 from PDB ID 4X23, describing its interaction with nucleosome) with those of the CENP-N:CENP-A^{NCP} complex.

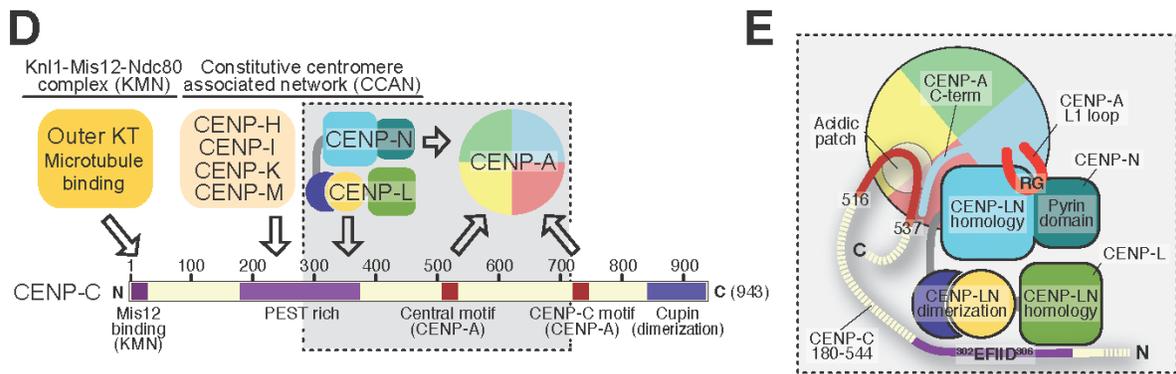


Figure 2.6 Effective CENP-N localization requires CENP-C

D. Schematic of crucial kinetochore interactions, already shown in Figure 2.1 A, but with question marks removed at interactions investigated in the present work. E. The grey box, an enlargement of the box in D, summarizes the details of the interactions reported in this work, as well as previous information on the interaction of CENP-C with the CENP-A^{NCP}.

Methods

Production of recombinant proteins in insect cells GST-CENP-L and its complexes with CENP-N fragments were produced in insect cells. Specifically, the coding sequence expressing 3C cleavable GST-tagged CENP-L was sub-cloned into MCS2 of the MultiBac vector pFL (Bieniossek et al., 2012), where the coding sequences of CENP-N or CENP-N^{230-C} were sub-cloned into MCS1 of pFL for co-expression. All constructs were further transformed into EMBacY cells and subsequently transfected into Sf9 cells in order to produce baculovirus. Baculovirus was amplified in Sf9 cells and used to infect Tnao38 cells. Tnao38 cells expressing GST-CENP-L:CENP-N, GST-CENP-L:CENP-N^{230-C}, or GST-CENP-L virus were cultured for 72 h and isolated (Weir et al., 2016). Briefly, cells were resuspended in lysis buffer containing 50 mM HEPES (pH 7.5), 300 mM NaCl, 10% glycerol, 4 mM 2-mercaptoethanol, 1 mM MgCl₂ pH 7.5 in presence of Benzonase. Resuspended cells were lysed by sonication and centrifuged at 100,000g at 4 °C for 1 h. Cleared lysates were incubated with pre-equilibrated GSH-Sepharose beads (Amintra) at 4°C for 2h. After extensive washing with lysis buffer, the GST fusion proteins were eluted in lysis buffer supplemented with 20 mM reduced glutathione. The eluted proteins were concentrated in a 10 K Da Amicon-Ultra-15 Centrifugal filter either alone or in the presence of GST-tagged 3C protease. Concentrated proteins were subsequently loaded onto a Superdex 200 16/600 column equilibrated in 20 mM HEPES pH 7.5, 2.5% glycerol, 300 mM NaCl and 1 mM TCEP. The corresponding peak fractions were collected and concentrated in a 10 KDa MWCO concentrator and then flash frozen in liquid N₂, and stored at -80 °C until further use.

Full length CENP-N was cloned into a pACEbac1 vector. CENP-N^{NT} constructs (containing the N-terminal constructs 1-289 and 1-238) were generated by introducing a sequence encoding 6His followed by two stop codons at the designated positions. CENP-N^{NT} constructs were expressed in SF21 insect cells (Invitrogen). Pellets from 1L of cell culture (~2 million cells/ml) were lysed in 100 ml lysis buffer (50 mM sodium phosphate pH 7.5, 300 mM NaCl, 20 mM imidazole and 10% glycerol). Nickel NTA beads (1.5 ml) were incubated with the lysed cells overnight at 4°C. Beads were washed with wash buffer (20 mM Tris.Cl pH 7.5, 300 mM NaCl, 30 mM imidazole and 10% glycerol), and eluted with elution buffer (20 mM Tris.Cl pH 7.5, 300 mM NaCl, 150 mM imidazole and 10% glycerol). The concentrated eluate was loaded onto an S200 gel filtration column in buffer containing 20 mM Tris.Cl pH 7.5, 300 mM NaCl, 10% glycerol and 1 mM DTT for the final purification.

Production of recombinant proteins in bacteria Fragments encoding different constructs of CENP-N were sub-cloned from a cDNA into pST50Tr-DHFRHIS for expression of recombinant C-terminally polyhistidine-tagged products (Tan et al., 2005). To produce GST tagged CENP-N, a GST encoding sequence was sub-cloned in pST50Tr-DHFRHIS in frame with the sequence coding for CENP-N¹⁻²¹². Mutant CENP-N¹⁻²¹² constructs were created by site-directed mutagenesis using the QuikChange kit (Agilent Technologies). Escherichia coli (DE3) cells harbouring vectors expressing CENP-N¹⁻²¹²-His or CENP-N¹⁻²³⁵-His were grown in TB media supplemented with 100 µg ampicillin at 37 °C at an OD₆₀₀ of 0.8-1.0. Then the temperature was reduced to 20 °C and protein expression was induced with 0.3 mM IPTG for 16 h. Cells were

harvested at 4600 g for 15 minutes. Bacterial pellets were resuspended in lysis buffer containing 50 mM HEPES (pH 7.5), 300 mM NaCl, 10% glycerol, 2 mM 2-mercaptoethanol, 10 mM MgCl₂ and 5 mM Imidazole. Resuspended cells were lysed by sonication and cleared by centrifugation at 100,000 g at 4°C for 30 minutes. The cleared lysate was incubated with cOmplete™ his tag beads (Roche) and incubated at 4°C for 2 hr. After extensive washing, CENP-N¹⁻²¹²-His or CENP-N¹⁻²³⁵-His were eluted in lysis buffer supplemented with 300 mM Imidazole. A 6 ml ResourceS cation exchange column was pre-equilibrated in 15% buffer B (20 mM HEPES, 1 M NaCl, 10% glycerol, 10 mM MgCl₂ pH 7.5, 1 mM TCEP) and 85% buffer A (20 mM HEPES, 10% glycerol, 10 mM MgCl₂ pH 7.5, 1 mM TCEP). CENP-N¹⁻²¹²-His or CENP-N¹⁻²³⁵-His were diluted with buffer A to reach a final concentration of 150 mM NaCl, loaded onto the ResourceS column and eluted with a linear gradient of buffer B from 150 to 1000 mM NaCl. Fractions containing CENP-N¹⁻²¹² His or CENP-N¹⁻²³⁵ His were concentrated in a 10 kDa MWCO and loaded onto a Superdex200 16/600 column equilibrated in 20 mM HEPES pH 7.5, 2.5% glycerol, 300 mM NaCl and 1 mM TCEP. The corresponding peak fractions were collected and concentrated in a 10 kDa MWCO concentrator and then flash frozen in liquid N₂, and stored at -80 °C until further use. The expression and purification procedure was identical for wild-type or mutants GST-CENP-N sequences, except that the concentration of NaCl in the lysis buffer was raised to 500 mM NaCl (instead of 300 mM).

CENP-C fragments encoding CENP-C²⁻⁵⁴⁵ or CENP-C²²⁵⁻³⁶⁴ were obtained from codon-optimized CENP-C cDNA and subcloned in pGEX-6P-2rbs for 3' fusion to the sequence

encoding GST. Mutant CENP-C constructs were created by site-directed mutagenesis using the QuikChange kit (Agilent Technologies). Expression and purification for all CENP-C constructs and mutants was carried out as described (Klare et al., 2015). All constructs were sequence-verified.

Nucleosome reconstitution Plasmids for the production of *X. laevis* H2A, H2B, H3 and H4 were a gift from D. Rhodes. Plasmids for the production of human CENP-A:H4 histone tetramer were a gift from A.F. Straight. 145-bp DNA (601-Widom) wrapped around CENP-A or H3 octamers was a gift from C.A. Davey. Purification of CENP-A or H3 containing NCPs were performed as previously described (Guse et al., 2011; Guse et al., 2012; Weir et al., 2016). For cryo-EM studies, nucleosomes were reconstituted from recombinant human H2A, H2B, CENP-A, and H4. Nucleosomes were reconstituted using the salt dialysis method (Dyer et al., 2004). '601' DNA (145 bp), H2A:H2B dimer, and CENP-A:H4 (or H3:H4) dimer were mixed at molar ratios of 1 to 2.4 to 2.4 in buffer containing 20 mM Tris.Cl pH 7.5, 2 M NaCl and 1 mM EDTA. Sample was transferred to a dialysis tube (D-tube Dialyzer, EMD Millipore), and dialyzed against RB high (Reconstitution Buffer high: 20 mM Tris.Cl pH7.5, 2 M NaCl, 1 mM EDTA and 1 mM DTT) for 4 hours at 4°C. Gradient dialysis was set up as described (Dyer et al., 2004). A peristaltic pump (flow rate 1.5 ml/minute) was used for replacing RB high with RB low (Reconstitution Buffer low: 20 mM Tris.Cl pH7.5, 0.25 M NaCl, 1 mM EDTA and 1 mM DTT) for 18 hours. Finally, the sample was dialyzed against storage buffer (20 mM Tris.Cl pH7.5, 20 mM NaCl, 1 mM EDTA and 1 mM DTT) for 8 hours.

GST pull-down assays GST pulldown assays were performed as previously described (Klare et al., 2015) with minor modifications. Briefly, pre-blocked GSH Sepharose beads were incubated in pull-down buffer consisting of 10 mM HEPES pH 7.5, 2.5% glycerol, 1 mM TCEP and either 300 mM NaCl for GST-LN-CENP-C assays or 150 mM NaCl for GST-LN-NCPs assays. GST-CENP-LN or GST-CENP-N¹⁻²¹² as baits (at a 1 μ M overall concentration) were incubated with either NCPs or CENP-C as prey at a 3 μ M concentration. The bait was loaded onto 10 μ l pre-blocked beads before adding the prey. The reaction volume was topped up to 40 μ l with buffer and incubated at 4°C for 1 h. Beads were spun down at 500g for 3 min. The supernatant was removed and beads were washed twice with 200 μ l pull-down buffer supplemented with 0.05% Triton-X-100. Supernatant was removed completely, samples boiled in 15 μ l Laemmli sample loading buffer and run on a 14% SDS-PAGE gel. Bands were visualized with Coomassie brilliant blue staining. For pre-blocking of GSH Sepharose beads, 750 μ l of GSH Sepharose beads were washed twice with 1 ml washing buffer (20 mM HEPES pH 7.5, 200 mM NaCl) and incubated in 1 ml blocking buffer (20 mM HEPES pH 7.5, 500 mM NaCl, 500 μ g/ μ l BSA) overnight at 4°C with rotation. Beads were washed five times with 1 ml washing buffer and resuspended in 500 μ l washing buffer to have a 50/50 slurry of beads and buffer.

Analytical SEC analysis Analytical SEC was performed on either Superdex 200 5/150 or Superose 6 5/150 columns in a buffer containing 20 mM HEPES (pH 7.5), 300 mM NaCl, 2.5% glycerol, 2 mM TCEP, pH 7.5 on an ÄKTAmicro system. Proteins were mixed at 6 μ M in a total volume of 50 μ l, incubated for one hour on ice, spun for 15

minutes in a bench-top centrifuge before each injection. All samples were eluted under isocratic conditions at 4°C in SEC buffer (20 mM HEPES pH 7.5, 300 mM NaCl, 2.5% glycerol, 2 mM TCEP) at a flow rate of 0.1 ml/min. Elution of proteins was monitored at 280 nm. Fractions (100 µl) were collected and analyzed by SDS–PAGE and Coomassie blue staining.

CENP-N:CENP-A nucleosome complex preparation Complexes containing CENP-A nucleosome and CENP-N^{NT} (CENP-N:CENP-A nucleosome complex) were generated by mixing CENP-A nucleosome and CENP-N^{NT} at a molar ratio of 1:3, then dialyzed against 20 mM Tris.Cl pH 7.5, 50 mM NaCl, 1 mM EDTA and 1 mM DTT. The complexes were concentrated to approximately 2.5 µM using a Millipore concentrator (MW 50 kD cut-off).

EMSA assays Nucleosomes were mixed with CENP-N^{NT} constructs in buffer containing 20 mM Tris.Cl, pH 7.5, 1 mM EDTA, 150 mM NaCl, 1 mM DTT and 0.1% CHAPS for 30 minutes at room temperature. The mixture was analysed by 5% native PAGE, and the gel was stained with SYBR Gold. The gel was imaged at 488 nm (for SYBR Gold staining or Alexa488 labelled H2B in the nucleosome) and 647 nm (for Atto N 647 labelled H2B in the nucleosome) using a Typhoon imager. Intensity at 647 nm was used for binding curves, with the number of replicates indicated. The intensity of the nucleosome band (not the shifted bands) was quantified, and normalized to the nucleosome sample without CENP-N.

Isothermal titration calorimetry All protein samples were loaded onto a Superdex 200 10/300 column equilibrated in 20 mM HEPES pH 7.5, 2.5% glycerol, 300 mM NaCl and 1 mM TCEP prior to ITC runs. ITC measurements were performed at 25°C on an ITC200 micro calorimeter (GE Healthcare). In each titration, the protein in the cell (at a 5-8 μ M concentration) was titrated with 19 x 2 μ l injections (at 180 sec intervals) of protein ligand (at 50-80 μ M concentration). The injections were continued beyond saturation to allow for determination of heats of ligand dilution. Data were fitted by least-square procedure to a single site binding model using ORIGIN software package (MicroCal).

Mammalian plasmids Plasmids for stable cell lines were generated in pCDNA5/FRT/TO-EGFP-IRES, a modified version of the pCDNA5/FRT/TO vector (Invitrogen, Carlsbad, CA). The starting plasmid for EGFP expression was made by PCR amplifying the EGFP sequence of the pEGFP-C1 plasmid (Takara Bio Inc.) and cloning it into the pCDNA5/FRT/TO vector previously modified to carry an internal ribosomal entry site (IRES) sequence creating the pCDNA5/FRT/TO EGFP-IRES vector (Petrovic et al., 2010). Mammalian expression plasmids used in this study to express N-terminally tagged CENP-C WT or the PEST mutant were derived from the pCDNA5/FRT/TO-EGFP-IRES and used either for genomic integration or transient expression of human CENP-C proteins. A modified version of the pCDNA5/FRT/TO-EGFP-IRES plasmid where EGFP was removed and replaced with mCherry to create the pCDNA5/FRT/TO-mCherry-IRES was used to express all C-terminally tagged CENP-N constructs either via genomic integration or transient expression. To create the

EGFP tagged CENP-C protein and CENP-N tagged with mCherry protein full-length proteins were amplified by PCR from full-length human codon-optimized cDNA synthesized by GeneArt (Life Technologies). Using the Gibson cloning method³⁶ the cDNA sequence was ligated with the PCR amplified pCDNA5/FRT/TO-EGFP-IRES or pCDNA5/FRT/TO-mCherry-IRES to create the final constructs used in this study. Each plasmid was then sequence verified before use.

Cell culture and transfection U2OS cells, a gift from A. Bird (MPI-Dortmund, Germany), were grown in DMEM (PAN Biotech, Aidenbach, Germany) supplemented with 10% FBS (Clontech, part of Takara Bio group, Shiga, Japan), penicillin and streptomycin (GIBCO, Carlsbad, CA), and 2 mM L-glutamine (PAN Biotech). Transient transfections were performed using pCDNA5-FRT-TO plasmids carrying either CENP-N^{wt} or mutants with a C-terminal-mCherry tag transfected into asynchronously growing cells and expressed for 24 h before preparation for immunofluorescence analysis. To analyse whether the CENP-C^{5A} mutant abrogates kinetochore localization of CENP-N, we used Flp-In T-REx HeLa cells (a gift from SS Taylor, University of Manchester, Manchester, England, UK) to generate stable doxycycline-inducible cell lines, which were maintained in DMEM (PAN Biotech, Aidenbach, Germany) with 10% tetracycline-free FBS (Clontech) supplemented with 50 µg/ml Zeocin (Invitrogen) and 2 mM L-glutamine (PAN Biotech). Flp-In T-REx HeLa expression cell lines were generated as previously described (Krenn et al., 2012). FlpIn T-REx HeLa cells stably expressing mCherry-CENP-N were grown on coverslips pre-coated with poly-D-Lysine (Millipore, 15 µg/ml). The day after cells were seeded, expression of mCherry-CENP-N was

induced using 0.2–0.5 µg/ml doxycycline (Sigma, St. Louis, MO). Further, cells were transiently transfected with GFP-CENP-C^{WT} or GFP-CENP-C^{5A} mutant and CENP-C siRNA for 72 h. When required 16 hours before fixation cells were synchronized overnight in 0.33 µM nocodazole (Sigma–Aldrich) to assess mitotic localization. Endogenous CENP-C was depleted by siRNA (target sequence: 5'-GGAUCAUCUCAGAAUAGAA-3' obtained from ThermoFisher), transfected into cells using Lipofectamine RNAiMAX Transfection Reagent (ThermoFisher) as per the manufactures instructions for 72 h.

Immunofluorescence Either U2OS or FlpIn T-REx HeLa cells were fixed using 4% paraformaldehyde in phosphate buffered saline (PBS), permeabilised using 0.25 % Triton X-100 in PBS and blocked in 3 % BSA/PBS. U2OS cells were stained for with the following antibodies: anti-CENP-C [rabbit polyclonal antibody SI410, diluted 1:1000 (Trazzi et al., 2009); and CREST/anti-centromere antibody (Antibodies Inc. 15-234-0001), diluted 1:100]. FlpIn T-REx HeLa cells were stained for GFP (GFP-Boost, Chromotek gba-488, 1:500), mCherry (RFP-Boost, Chromotek, rba-594, 1:500), CENP-C [rabbit polyclonal antibody SI410; 1:1000 (Trazzi et al., 2009)], CREST/anti-centromere antibodies (Antibodies, Inc., Davis, CA, 1:100), or CENP-A (Mouse, Gene Tex GTX13939, 1:500) diluted in blocking buffer for 2-4 h. Donkey anti-human Alexa Fluor 405 and donkey anti-rabbit Alexa Fluor 488, donkey anti-human (Jackson ImmunoResearch Laboratories, Inc., West Grove, PA), as well as donkey anti-mouse (Invitrogen) were used as secondary antibodies. DNA was stained with 0.5 µg/ml DAPI (Serva), and coverslips were mounted with Mowiol mounting media (Calbiochem).

U2OS cells were imaged with a Deltavision Elite System (GE Healthcare, UK) equipped with an IX-71 inverted microscope (Olympus, Japan), a PLAPON 60x/1.42NA objective and a pco.edge sCMOS camera (PCO-TECH Inc., USA). Images were acquired as Z-sections (using the softWoRx software from Deltavision) and converted into maximal intensity projections TIFF files for illustrative purposes while FlpIn T-REx HeLa Cells were imaged at room temperature using a spinning disk confocal device on the 3i Marianas system equipped with an Axio Observer Z1 microscope (Zeiss), a CSU-X1 confocal scanner unit (Yokogawa Electric Corporation, Tokyo, Japan), Plan-Apochromat 63x or 100x/1.4NA Oil Objectives (Zeiss), and Orca Flash 4.0 sCMOS Camera (Hamamatsu). Images were acquired as z-sections at 0.2 μm . Images were converted into maximal intensity projections, exported, and converted into 8-bit. Quantification of kinetochore signals was performed on unmodified 16-bit z-series images using Imaris 7.3.4 32-bit software (Bitplane, Zurich, Switzerland). Measurements were exported in Excel (Microsoft) and graphed with GraphPad Prism 6.0 (GraphPad Software, San Diego California USA). For quantification in Figure 2.4 B-C, after background subtraction, all signals were normalized to CREST and the values obtained were then normalized by the mean of the CENP-N-mCherryWT construct. Quantifications are based on one experiment where a range of 7 to 10 cells and 177 to 393 kinetochores per condition were analysed.

Crystallization of CENP-N¹⁻²³⁵ His and structure determination Initial crystallization hits of CENP-N¹⁻²¹²His6 (His6, hexahistidine tag) or CENP-N¹⁻²³⁵His6 were obtained in sitting drop crystallization experiments at c.a. 6 mg/ml in a 96 well format using a

Mosquito protein crystallization robot (TTP Labtech) at 4°C. Crystals grew in various commercial screens including Qiagen Nextal PEGSII conditions B11 and B12 and Qiagen Nextal PEG conditions H6 and H8 within 24-48 hours, reaching maximum size in 5-7 days. CENP-N¹⁻²³⁵His6 crystals were further optimized in 24-well plates via hanging drop method using a two-dimensional grid screen varying PEG3350 (from 6%-16%) and pH (from 6.6-7.2). Selenomethionine (SeMet) derivatives of CENP-N¹⁻²³⁵His6 were crystallized under similar conditions. Crystals were cryo-cooled in a mother liquor solution containing 20-25% (v/v) glycerol. All data was collected at 100K using a Pilatus 6M detector either at the X10SA beamline at the SLS in Villigen, Switzerland, or at the P11 beamline of PETRA in Hamburg, Germany. All data sets were integrated and scaled using XDS and XSCALE (Kabsch, 2010).

Both native and selenomethionine (SeMet) crystals grew in space group P4₁ with two molecules per asymmetric unit and a relatively similar packing, although the c-axis is approximately 4 Å shorter in the latter. Phasing with PHENIX (Adams et al., 2010) located 12 of the 14 possible SeMet sites (in each monomer a C-terminal Met is disordered, and M167 has very weak anomalous density). Merging of SeMet datasets from two different crystals was essential to improve the anomalous signal. The quality of the phases allowed autobuilding of (mostly) alpha helices into the electron density in spite of the relatively low resolution (3.3 Å). Molecular replacement with PHASER (Collaborative Computational Project, 1994) successfully placed this initial model into the native dataset (conservative resolution at I/sigma=3 of 2.89 Å, data used to 2.74 Å). The sequence was assigned with the help of the anomalous peaks. Refinement with

REFMAC (Collaborative Computational Project, 1994) and PHENIX resulted in a model with very good Ramachandran geometry (98% residues in favoured regions, 0% outliers) and $R_{\text{work}}/R_{\text{free}}$ values of 21.6 and 26.0%, respectively. Data to 2.74 Å were used despite high R factors, since they improved the convergence and quality of the refinement.

Monomers A and B in the asymmetric unit are quite similar except for a minor hinge-bending between the two subdomains of each monomer, and except for loops 137-142 and 164-174. In monomer B, the 137-142 loop is pulled approximately 3 Å away from the remainder of the molecule by a symmetry contact, and in turn pulls on the neighbouring loop causing a 2.6-Å movement of the backbone of K117. This loop has no crystal contacts in monomer A. Similarly, residues 164-174 of monomer B pack against a symmetry related molecule, most likely causing the observed (relative) stabilization of M167 in monomer B and backbone shifts of up to 7.3 Å relative to monomer A. Residues 166-168 have very weak density especially in chain A, corresponding to a weak and multiple anomalous density of SeM167, indicating multiple conformations of this loop. The coordinates of CENP-N have been submitted to the protein data bank (PDB) with code 6EQT.

CryoEM grid preparation and microscopy Quantifoil 2/2 grids (Quantifoil Micro Tools GmbH) were used for the CENP-N^{NT}:CENP-A^{NCP} complex (2.5 μM concentration). The grids were glow discharged (EMIttec) at 40 mA for 20 seconds. 4 μl sample was applied onto the grid before plunge freezing into Ethane, using a Vitrobot (FEI, MARK IV). Blot

time was 4 seconds. All grids were stored in liquid nitrogen before imaging. CENP-N^{NT}:CENP-A^{NCP} complex was imaged at nominal magnification of 29000x on a FEI Titan Krios (300 kV), equipped with a Gatan K2 Summit direct detector. Pixel size was 1.02 Å. The movies were captured in super resolution mode with electron dose rate at 10 electrons per pixel per second for 8 seconds and 0.2 seconds per frame. Defocus range was -1.0 to -2.5 µm.

Single particle analysis of cryoEM images Motioncor2 was used for the alignment of images (motion correction) (Zheng et al., 2017). GCTF was applied for Constant transfer function (CTF) estimation (Zhang, 2016). Images were evaluated manually by inspecting their power spectra. Particles were manually picked for the initial 2D classification (10 class averages) in RELION 2.05 (Fernandez-Leiro and Scheres, 2017). Initial 2D class averages were used for particle auto-picking as described in Relion2.0 tutorials. Particles from auto-picking were extracted and sorted for reference free 2D classification. In 2D classification, 200 class averages were generated. Noisy class averages were discarded. Particles from the retained class averages were used for reference free 3D reconstruction in cryoSPARC (Punjani et al., 2017). The low pass filtered map from the 3D reconstruction in cryoSPARC was used as reference for 3D classification in RELION 2.05 (Fernandez-Leiro and Scheres, 2017). Four 3D classes were created after 3D classification. Particle images from the class at high resolution were used for 3D refinement in Relion 2.1b1. The map was then sharpened in “post-process”, using a mask file created by Relion2.1b1. The local resolution of the map was estimated by Relion (Kucukelbir et al., 2014).

CryoEM structure modelling The original 3D refined map and post-processed map were used for the model fitting and refinement. DNA was taken from a high-resolution crystal structure of a nucleosome with 601 DNA [Protein Data Bank accession code 3LZ0 (Vasudevan et al., 2010)], CENP-A containing histone core was taken from Protein Data Bank accession code 3AN2 (Tachiwana et al., 2011). The crystal structure of CENP-N (amino acids 1-212) was used to fit into the remaining density using UCSF Chimera (Pettersen et al., 2004) (rigid body without allowing flexibility). Based on map density, the model was iteratively modified and locally refined in Coot (Emsley et al., 2010). The final model was subjected to real space refinement in PHENIX (Adams et al., 2010).

CHAPTER 3 CENP-A nucleosome is stabilized by both CENP-N and CENP-C *in vitro*²

Overview

Eukaryotic centromeres are defined by the presence of nucleosomes containing the histone H3 variant, Centromere Protein A (CENP-A). Once incorporated at centromeres, CENP-A nucleosomes are remarkably stable, exhibiting no detectable loss or exchange over many cell cycles. It is currently unclear whether this stability is an intrinsic property of CENP-A containing chromatin or whether it arises from proteins that specifically associate with CENP-A chromatin. Two proteins, CENP-C and CENP-N, are known to bind CENP-A human nucleosomes directly. Here, we test the hypothesis that CENP-C or CENP-N stabilizes CENP-A nucleosomes *in vitro* and in living cells. We show that CENP-N stabilizes CENP-A nucleosomes alone, and additively with CENP-C *in vitro*. However, removal of CENP-C and CENP-N from cells, or mutating CENP-A so that it no longer interacts with CENP-C or CENP-N, had no effect on centromeric CENP-A stability *in vivo*. Thus, the stability of CENP-A nucleosomes in chromatin does not solely arise from its interactions with CENP-C or CENP-N.

² Shengya Cao*, Keda Zhou*, Zhening Zhang, Karolin Luger, Aaron F. Straight, 2018, Constitutive centromere-associated network contacts confer differential stability on CENP-A nucleosomes *in vitro* and in the cell, *molecular biology of the cell*, E17-10_0596. It is a collaboration work with Dr. Aaron Straight's group. I am the co-first (*) author of this paper. My contributions on this work include all *in vitro* assays (Figure 3.1 to Figure 3.3, Supplemental Figure S3.1 and S3.2).

Introduction

During mitosis, vertebrate cells assemble one kinetochore on each chromosome to connect chromosomes to spindle microtubules, monitor chromosome alignment on the spindle, and move chromosomes to poles during anaphase. The assembly site for the kinetochore is the centromere, a specialized chromatin domain that is epigenetically specified by the replacement of histone H3 in nucleosomes with the centromere-specific histone variant CENP-A (McKinley and Cheeseman, 2016). Unlike histones H3.1 and H3.2, CENP-A nucleosome assembly is uncoupled from replication and occurs only after mitotic exit in G1 (Jansen et al., 2007). To maintain centromere identity, CENP-A nucleosomes must remain at centromeres outside of the assembly period (Hoffmann et al., 2016). This occurs because CENP-A nucleosomes are stable once incorporated into chromatin, showing loss only by dilution through replication (Bodor et al., 2013). CENP-A appears to be stably maintained at centromeres over days in dividing cells and for months or even years in post-mitotic cells (Smoak et al., 2016).

Currently, it is unclear whether the stability of CENP-A nucleosomes in chromatin is an intrinsic property of CENP-A containing nucleosomes or results from other factors that bind to and stabilize CENP-A nucleosomes. On a single nucleosome level, purified or reconstituted CENP-A nucleosomes are more prone to disassembly than H3 nucleosomes when challenged with the destabilizing effects of heat, heparin or nucleosome assembly protein 1 (NAP-1) (Arimura et al., 2014; Conde et al., 2007). In dividing cells, CENP-A at centromeres exhibit slower turnover rates than most H3 and

are as stable as the most stable H3 (Bodor et al., 2013). While much of H3 turnover is mediated by active processes like chromatin remodeling and transcription (Henikoff, 2008), it is surprising that these processes do not also act on CENP-A given that there is transcription at active centromeres (McNulty et al., 2017). Therefore, we hypothesize that factors that distinguish between CENP-A and H3 nucleosomes can help stabilize CENP-A nucleosomes either through direct binding or exclusion of these active processes.

Two essential proteins, CENP-C and CENP-N, have been found to preferentially interact with CENP-A nucleosomes at the centromere. Previous studies have suggested that CENP-C stabilizes CENP-A nucleosomes (Falk et al., 2015, 2016). Another study showed that replacing the loop1 and α 2-helix region on H3 with the CATD region of CENP-A, the region that binds CENP-N (Carroll et al., 2009), is sufficient to confer stability to CENP-A nucleosomes *in vivo* (Bodor et al., 2013). In a recent report, it was proposed that centromeres are maintained by CENP-N through fastening of CENP-A to DNA (Guo et al., 2017). We tested the direct effects of CENP-C and CENP-N on CENP-A nucleosomes maintenance *in vitro* and in cells. Here we show that *in vitro*, CENP-N protects CENP-A mono-nucleosomes against the destabilizing effects of salt, dilution, or plunge-freezing on an EM-grid, and this stabilization is further increased by CENP-C. In contrast, rapid *in vivo* degradation of either CENP-C, CENP-N, or both, had no effect on CENP-A stability at centromeres within a single cell cycle. Similarly, mutating residues R80 and G81 in the context of CENP-A, or a chimera of H3 containing just the CENP-A CATD, had no effect on the maintenance of these proteins at centromeres. Our results

indicate that although CENP-C and CENP-N provide added stability to the CENP-A nucleosome *in vitro*, this effect does not account for any additional stability of the CENP-A nucleosome *in vivo*.

Results

CENP-A nucleosome DNA ends are less accessible to micrococcal nuclease when CENP-N interacts with the nucleosome

The CENP-A nucleosome coordinates fewer base pairs of DNA near the DNA entry and exit sites than does the histone H3 nucleosome (Dechassa et al., 2011; Kingston et al., 2011; Tachiwana et al., 2011; Yoda et al., 2000). This reduced wrapping of DNA around the histone core promotes the loss of H2A/H2B and is thought to be one key feature that causes or reflects the instability of centromeric nucleosomes (Voltz et al., 2012). Here, we performed limited micrococcal nuclease digestion (MNase) with both CENP-A nucleosome and H3 nucleosomes, to test whether this affects MNase accessibility. Nucleosomes were reconstituted with a 166 bp DNA fragment derived from the '601' nucleosome positioning sequence identified by Widom and colleagues (Lowary and Widom, 1998) (sequence listed under 'Materials and Methods'). While the majority of DNA in the H3 nucleosome sample remained unaffected by MNase (leaving an intact 166 bp DNA fragment Figure 3.1 A), nearly all DNA in the CENP-A nucleosome was degraded to smaller fragments (157 bp, 140 bp and 126 bp, Figure 3.1 A). This indicates differences in the way in which DNA ends are organized by the histone core, with H3 conferring more stable wrapping than CENP-A, and thus less access for MNase.

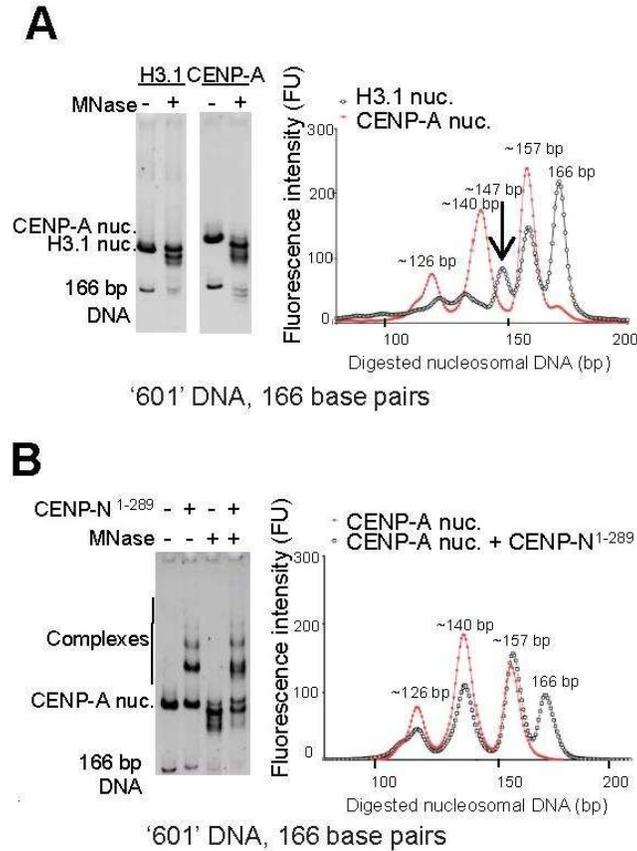


Figure 3.1 CENP-N protects CENP-A nucleosomes from micrococcal nuclease digestion.

A. Comparative MNase digestion analysis of H3.1- and CENP-A nucleosomes, reconstituted with a 166 bp DNA fragment derived from the '601' sequence. After digestion (quenched with 50 mM EDTA), samples were analyzed in a gel shift assay prior to DNA extraction (left panel) and after DNA extraction (right panel). The length of the DNA fragments was determined according to the standard DNA ladder curve obtained from the Bioanalyzer (Agilent).

B. MNase digestion analysis of CENP-A nucleosome (with the same DNA sequence as in A) in the absence and presence of CENP-N¹⁻²⁸⁹. The reactions were performed and analyzed as in A.

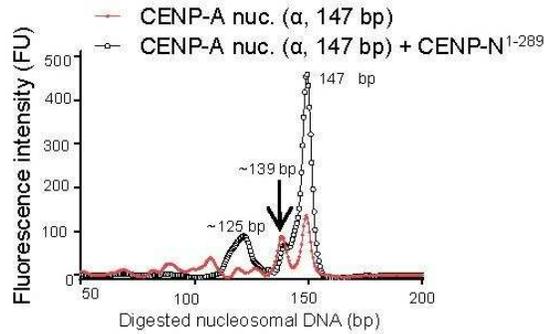
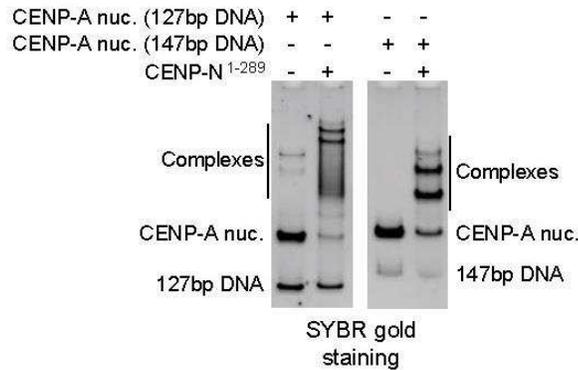
C**D**

Figure 3.1 CENP-N protects CENP-A nucleosomes from micrococcal nuclease digestion.

C. MNase digestion analysis of CENP-A nucleosome (147 bp DNA derived from α satellite DNA) in the absence and presence of CENP-N¹⁻²⁸⁹.

D. The DNA ends of the CENP-A nucleosome are important for the proper orientation of CENP-N¹⁻²⁸⁹ on the CENP-A nucleosome. CENP-A nucleosomes were reconstituted with either 147 bp or 127 bp of 601 nucleosome positioning DNA. CENP-N¹⁻²⁸⁹ was mixed with CENP-A nucleosome at 3:1 ratio. After a 5-minute incubation at 37° C, samples were analyzed by 5% native PAGE.

We then asked whether the nucleosome binding domain (1-289) of CENP-N would confer increased stability on DNA ends when bound to CENP-A nucleosomes, by performing MNase assays in presence of CENP-N. CENP-N¹⁻²⁸⁹ elutes as a monomer from an S200 gel filtration column (Supplemental Figure 3.1 A). Native PAGE suggests that more than one CENP-N¹⁻²⁸⁹ can interact with a CENP-A nucleosome, as expected from the presence of two equivalent binding sites on the nucleosome (Supplemental Figure 3.1B). To determine the effect of CENP-N binding on DNA accessibility, we repeated the MNase treatment with a 1:2 complex of the CENP-A-nucleosome with CENP-N¹⁻²⁸⁹. While the patterns of generated DNA fragments were the same for the CENP-A nucleosome in the presence or absence of CENP-N, more of the 166 bp undigested DNA fragment was extracted from the sample with CENP-N present (Figure 3.1 B). We repeated this with nucleosomes reconstituted with either a 147 bp or 207 bp DNA fragment with the same '601' core sequence, both of which showed similar results (data not shown). To exclude that this effect was a result of the rather unique properties of the '601' sequence, we repeated the experiment with the palindromic DNA derived from α -satellite DNA, which closely resembles the native DNA template for CENP-A at the centromere. Consistent with the results described above, CENP-N conferred increasing resistance towards MNase digestion in CENP-A nucleosomes reconstituted with α -satellite DNA (Figure 3.1 C). Therefore, we conclude that the interaction between CENP-N and CENP-A nucleosome stabilizes the DNA ends in CENP-A nucleosomes, either directly or indirectly, and protects them from digestion by MNase. When we tested a CENP-A nucleosome reconstituted with a 127 bp DNA fragment (where the penultimate 10 bp were removed from the 147 bp "601" DNA fragment used in the other

experiments), these CENP-A nucleosomes formed less defined complexes with CENP-N as judged by native PAGE, where complexes formed a smear rather than two discrete shifted bands (Figure 3.1 D). This implies that the DNA ends are essential for the formation of a well-defined complex between the CENP-A nucleosome and CENP-N.

CENP-N increases the stability of the CENP-A nucleosome *in vitro*

To further explore the effect of CENP-N on CENP-A nucleosome stability, we used EMSA to quantify the amount of stable nucleosome remaining under various destabilizing conditions. First, we incubated nucleosomes in the absence and presence of CENP-N¹⁻²⁸⁹ at increasing ionic strength. CENP-A nucleosomes dissociated in increased ionic strength such that less than half of the input nucleosome was present at 300 and 600 mM salt. This dissociation was reduced in the presence of CENP-N¹⁻²⁸⁹ (Figure 3.2 A), such that 75 % of the nucleosomes remained intact throughout the salt titration series (Figure 2 A, left panel). CENP-N also stabilized CENP-A nucleosomes against dissociation in response to dilution and heat (Figure 3.2 B). Heating CENP-N for 5 minutes at 55 °C was sufficient to denature CENP-N and prevented it from binding to the CENP-A nucleosome, and this served as a control in these experiments. In the presence of CENP-N, 75% of CENP-A nucleosomes remained intact after 2-fold dilution and heat treatment, while only 25-30% of nucleosomes remained intact in the absence

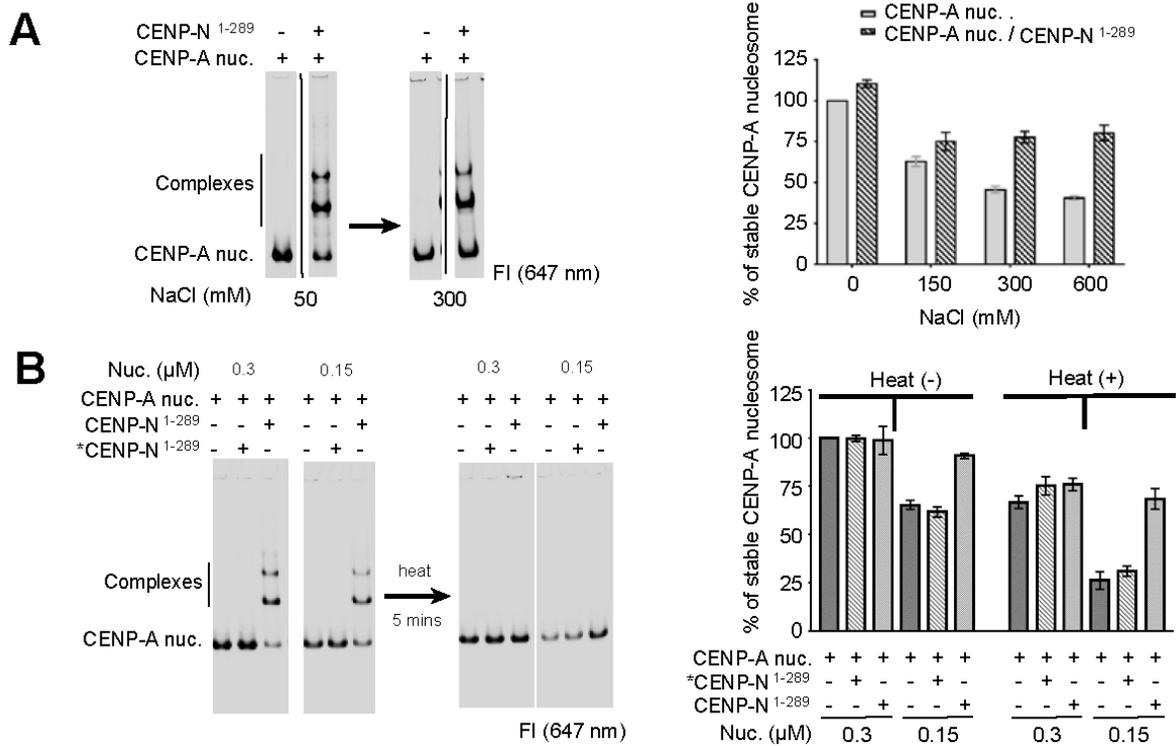


Figure 3.2 CENP-N increases the stability of CENP-A nucleosomes against dissociation *in vitro*.

A. CENP-N¹⁻²⁸⁹ increases the stability of CENP-A nucleosomes at increased ionic strength. 100 nM CENP-A nucleosome was mixed with either buffer or 200 nM CENP-N¹⁻²⁸⁹ at the indicated NaCl concentration. Nucleosome stability was quantified by measuring the intensity of all bands including shifted bands (right panel; normalized as described in methods), n=3. All assays were performed in a final buffer containing 20 mM Tris HCl, pH 7.5, the indicated NaCl concentration, 5% glycerol, 0.5 mM EDTA (no detergent). The ~ 25 % of nucleosomes that are not protected in the presence of CENP-N likely reflects the amount of unbound CENP-A nucleosome (right panel).

B. CENP-N¹⁻²⁸⁹ stabilizes CENP-A nucleosomes against the combined effects of dilution and heat treatment. Left panel: native PAGE. CENP-N¹⁻²⁸⁹ was mixed with CENP-A nucleosomes at a molar ratio of 3:1. As a control, *CENP-N¹⁻²⁸⁹ indicates CENP-N¹⁻²⁸⁹ that was denatured by heating at 55°C for 5 minutes before mixing with CENP-A nucleosomes. The same amount of sample was loaded immediately after treatment. All bands including shifted bands were quantified to determine the percent of remaining nucleosome, and error bars are derived from three independent gels (n=3). Intensity at 647 nm was measured.

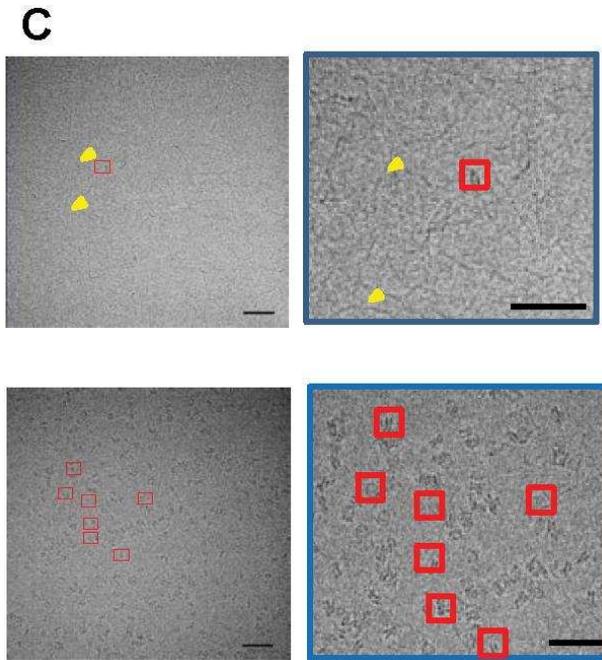


Figure 3.2 CENP-N increases the stability of CENP-A nucleosomes against dissociation *in vitro*.

C. CENP-A mono-nucleosomes on an EM grid are greatly stabilized in the presence of CENP-N¹⁻²⁸⁹. The CENP-A nucleosome sample (2.5 μ M) was mixed with 7.5 μ M CENP-N¹⁻²⁸⁹, and the control was adjusted with buffer. Red boxes indicate nucleosome-shaped particles. Yellow arrows show free DNA. Scale bar=50 nm. The blue box highlights the area from the left micrograph. The number of intact particles were counted and are listed in Table 1.

Table 3.1**Electron microscopy statistics**

Sample	Particle number (per image)	
	Cryo-EM grid	Negative stain
CENP-A nucleosome	< 10	290
Complex AN	700	460
Complex ANC	800	N/A

Particles were picked on images obtained from cryo-EM by using nucleosome like 2D classes as reference in Relion. Particles on grids obtained from negative-stain EM grid were picked and counted by using “dogPicker”, which defined the nucleosome particle by diameter (larger than 10 nm). No negative-stain EM was performed for complex ANC.

of CENP-N, or with CENP-N that had been denatured before addition to CENP-A nucleosomes. Finally, application of CENP-A nucleosomes onto a cryoEM grid (C-Flat gold) followed by plunge-freezing, yielded very few (less than 10 per image) intact nucleosome particles, while the very same preparations, in the presence of a two-to-three fold excess of CENP-N, yielded a large number of well-formed particles (more than 700 nucleosome-like particles per image, Figure 3.2 C and Table 3.1). Negative staining EM was also used to test the stability of the CENP-A nucleosome. Interestingly, the particle sizes varied drastically when no CENP-N was present in solution, while the presence of CENP-N dramatically increased not only the total number but also the quality of particles (Supplemental figure 3.1 C and Table 3.1). Together, this data indicates a direct role for CENP-N in CENP-A nucleosome stabilization *in vitro*. It should be noted that we and others have successfully prepared EM grids with H3 nucleosomes (Chua et al., 2016) and thus the increased fragility of CENP-A containing nucleosomes which is alleviated by CENP-N, is due to CENP-A. A recent report showed that CENP-N fastens the nucleosome by simultaneously binding to the CATD on CENP-A and to nucleosomal DNA around CATD, which could stabilize the CENP-A nucleosome (Guo et al., 2017). This was recently confirmed by cryo-EM studies by us and others (Pentakota et al., 2017; Sagar, Chittori; Jingjun Hong Hayden Saunders, Hanqiao Feng, Rodolfo Ghirlando, Alexander E. Kelly², Yawen Bai², 2017). Our data provided direct evidence on the stabilizing effect of CENP-N on CENP-A nucleosome *in vitro*.

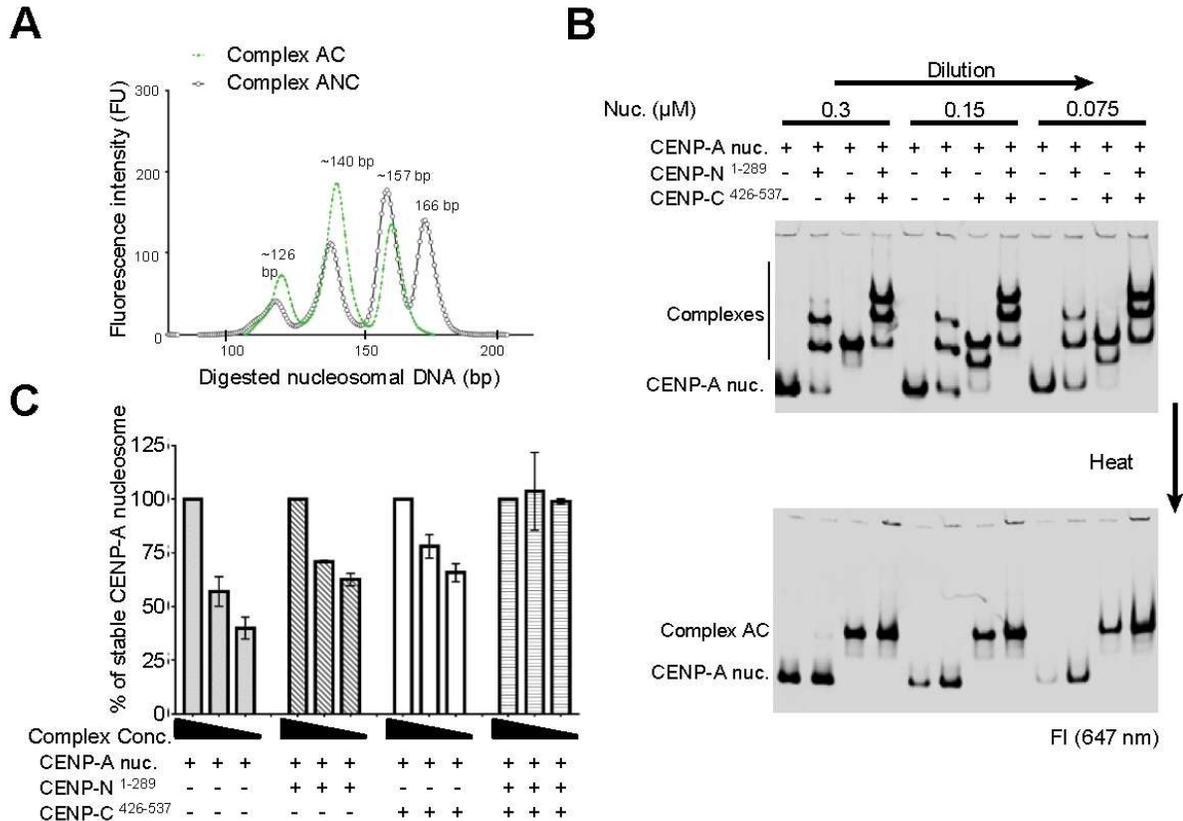


Figure 3.3 the CENP-A nucleosome binding domains of CENP-N and CENP-C have additive effects on stabilizing CENP-A nucleosomes in vitro.

A. MNase digestion analysis for complex AC and complex ANC. CENP-C⁴²⁶⁻⁵³⁷ was mixed with CENP-A nucleosome at a 2:1 ratio to form the complex AC. CENP-N¹⁻²⁸⁹ was mixed with preformed AC complex at 2:1 ratio to form the ANC complex. DNA fragments after digestion were analyzed as in Figure 1A.

B. CENP-N¹⁻²⁸⁹ and CENP-C⁴²⁶⁻⁵³⁷ showed additive effects in stabilizing the CENP-A nucleosome against dilution and heat. CENP-N¹⁻²⁸⁹ or CENP-C⁴²⁶⁻⁵³⁷ was mixed with CENP-A nucleosome at molar ratios of 3:1. The buffer was adjusted to a final 20 mM Tris HCl, pH 7.5, 150 mM NaCl, 2% glycerol, 0.5 mM EDTA. Dilution and heat treatments were done as described in the methods section. Samples were kept at 4°C for 2 hours before analysis by native PAGE.

C. Quantification of the percent of stable nucleosomes under different nucleosome concentrations without heat (from the upper panel in J), n=2. The raw signal (intensity of fluorescence signal at 647 nm) of nucleosome or complex after dilution (150 nM and 75 nM) was normalized to the signal from samples before dilution (at 300 nM NaCl).

D

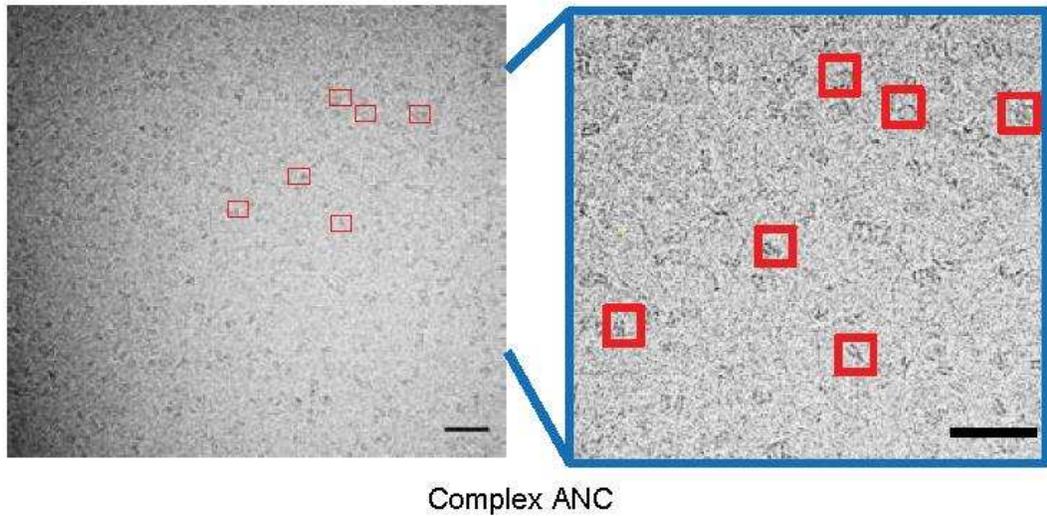


Figure 3.3 the CENP-A nucleosome binding domains of CENP-N and CENP-C have additive effects on stabilizing CENP-A nucleosomes in vitro.

D. The stabilizing effect of CENP-N and CENP-C on CENP-A nucleosome, as demonstrated by cryo-EM. Both CENP-N¹⁻²⁸⁹ and CENP-C⁴²⁶⁻⁵³⁷ were mixed with CENP-A nucleosome at molar ratio 3:1 to form the ANC complex. Buffer was adjusted to the same condition for both samples. Concentration of both samples was 2.5 μM . Red boxes indicate nucleosome-shaped particles. Scale bar=50 nm. Blue box highlights the area from the left micrograph. The number of intact particles were counted and are listed in Table 1.

CENP-N and CENP-C additively stabilize CENP-A nucleosomes *in vitro*

A second CENP-A nucleosome specific binding protein, CENP-C, has also been shown to stabilize CENP-A nucleosomes, and is thought to bind CENP-A nucleosomes through a binding site that differs from that of CENP-N (Carroll et al., 2010). We confirmed that the nucleosome binding domains of CENP-C and CENP-N (CENP-N¹⁻²⁸⁹ and CENP-C⁴²⁶⁻⁵³⁷) form distinct complexes when incubated with CENP-A nucleosomes individually (Supplemental figure 3.2 A). When CENP-N¹⁻²⁸⁹ is added to a pre-formed complex of CENP-C⁴²⁶⁻⁵³⁷ with the CENP-A nucleosome (complex AC) we observe a distinct pattern by EMSA which is consistent with independent binding events. Using pulldown assays (Supplemental figure 3.2 B), we show that indeed CENP-N and CENP-C bind to the same nucleosome (complex ANC), in agreement with previous reports (Carroll et al., 2010; Guo et al., 2017). To test whether simultaneous binding of CENP-C and CENP-N provides added stability to nucleosomes, we repeated the MNase protection experiment with the complex AC and with complex ANC. The complex AC exhibited the same sensitivity towards MNase digestion as CENP-A nucleosomes alone, while the complex ANC exhibited the same level of protection from MNase as the AN complex (Figure 3.3 A). We conclude that CENP-C does not contribute to the protection of DNA ends in CENP-A nucleosomes. Thus, the coordinated binding of CENP-N and CENP-C to CENP-A nucleosomes could further increase the overall stability of the nucleosome. Using EMSA and cryoEM, we found that CENP-C⁴²⁶⁻⁵³⁷ and CENP-N¹⁻²⁸⁹ showed similar stabilizing effects on CENP-A nucleosomes, but the combination of CENP-C⁴²⁶⁻⁵³⁷ and CENP-N¹⁻²⁸⁹ increased the stability of the CENP-A nucleosome more than

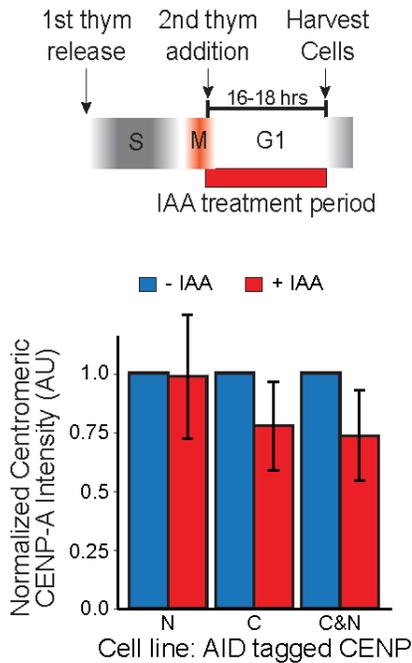
CENP-C or CENP-N alone (Figure 3.3 B, 3.3 C and 3.3 D, Table 3.1). Thus, CENP-N and CENP-C independently protect CENP-A nucleosomes against dissociation *in vitro*.

Loss of CENP-C and CENP-N does not alter CENP-A nucleosome levels in chromatin

We tested whether the stability of CENP-A nucleosomes *in vivo* results from the nucleosome-stabilizing effects of CENP-N and/or CENP-C observed *in vitro*. To this end, we generated cells expressing conditionally degradable CENP-C and/or CENP-N, by fusing the auxin-inducible-degron (AID) tag to the C-terminus of the endogenous CENP-C and CENP-N genes in cells expressing the Fbox protein, Tir1 (Holland et al., 2012; McKinley et al., 2015; Nishimura et al., 2009). We also tagged CENP-N with superfolder GFP (sfGFP), and CENP-C with mRuby2 and a 3xFLAG epitope. We validated that we had modified both alleles of the endogenous genes by PCR or western blotting (Supplemental Figure 3.3 C and 3.3 D). Upon the addition of 1 mM indole-3-acetic acid (IAA), either CENP-N alone or CENP-N and C were degraded to background levels within 30 minutes (Supplemental Figure 3.3).

We validated the specificity of CENP-N degradation by stably introducing mRuby2-3xFLAG-tagged full-length CENP-N, mRuby2-3xFLAG-tagged CENP-N truncations or the mRuby2-3xFLAG tag alone, as transgenes into our AID tagged CENP-N cells. The mRuby2-3xFLAG tag expressed alone was not stable in cells (Supplemental Figure 3.4 A). However, when fused to full length or truncated CENP-N, the fusion proteins were

A Degrade CENP-C and/or CENP-N during G1



B Degrade CENP-C and/or CENP-N during G1 through G2 or S through G2

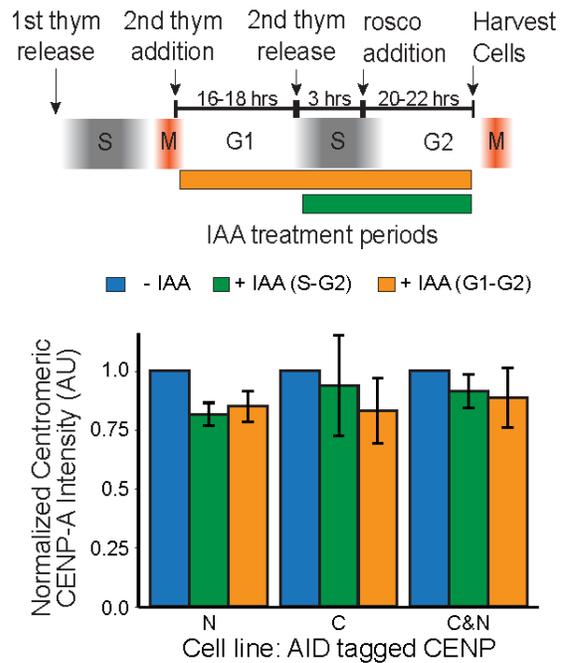


Figure 3.4 CENP-C and CENP-N degradation has no significant effect on centromeric CENP-A maintenance.

A, B. Cell lines containing AID tagged CENP-N, CENP-C or both were treated with IAA to degrade the indicated proteins. Blue bars represent cells not treated with IAA. Cells were maintained in IAA beginning just after mitosis (red bars) and harvested in an early S phase thymidine (thym) arrest (A) or after mitosis (mustard bars) or in early S phase (green bars) and harvested in a G2 phase roscovitine (rosc) arrest (B). Centromeric CENP-A immunofluorescence signal was normalized to the no IAA signal.

Data are presented as mean \pm SEM for three independent replicates. * $p < 0.05$

C Degradate CENP-N during G1 and assay Labeled CENP-A

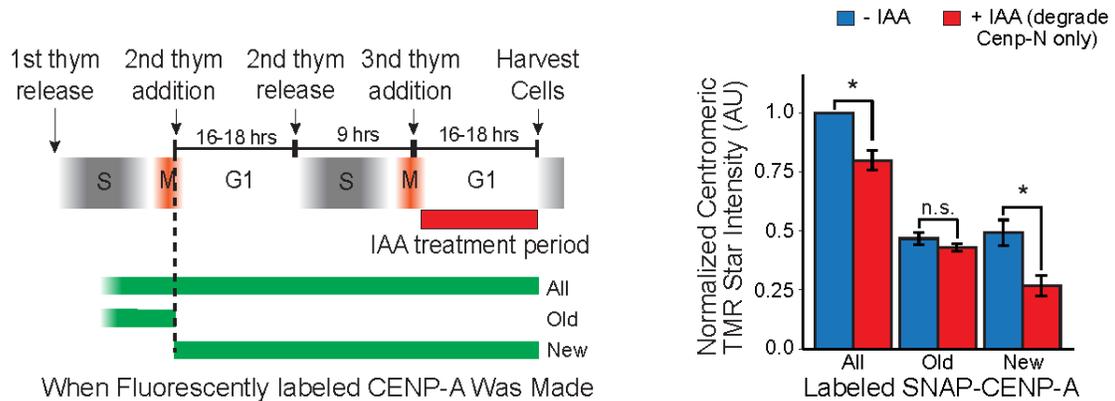


Figure 3.4 CENP-C and CENP-N degradation has no significant effect on centromeric CENP-A maintenance.

C. Degradation of CENP-N inhibits new CENP-A assembly but not preexisting CENP-A in chromatin. The CENP-N AID-sfGFP cell line containing a stably integrated SNAP-tagged CENP-A was either fluorescently labeled or quenched according to the schematic (left panel). Green bars represent the time of synthesis of the fluorescent population of SNAP-tagged CENP-A. IAA was added as in (A). Centromeric TMR star intensity represents the fluorescent population of SNAP-tagged CENP-A.

Data are presented as mean \pm SEM for three independent replicates. * $p < 0.05$

stable and expressed at similar levels (Supplemental Figure 3.4 A). Upon IAA addition to degrade the endogenously tagged CENP-N, cells expressing only the tag or N-terminal truncations of CENP-N showed signs of chromosome segregation defects upon CENP-N loss, indicated by the appearance of micronuclei, as previously described (McKinley et al., 2015). No significant increase in micronuclei was detected in cells expressing the full-length CENP-N, indicating that these defects result specifically from CENP-N degradation and that the full-length CENP-N transgene complements the IAA-induced loss of CENP-N (Supplemental Figure 3.4 B and 3.4 C). We tested whether the loss of CENP-C, CENP-N, or both had an impact on the levels of CENP-A at centromeres by degrading the proteins and measuring the levels of centromere associated CENP-A using an anti-CENP-A antibody. To avoid assaying secondary effects from chromosome segregation defects, we degraded the proteins during drug-induced cell cycle arrests. In all cases, cells were treated with 0.5 mM thymidine to induce arrest at the G1-S boundary. Cells were then released from this arrest and allowed to go through mitosis before the addition of 1 mM IAA. Cells were cultured in the presence of IAA either during G1 ending in a second thymidine arrest at the G1-S boundary (Figure 3.4 A), through a release from the second thymidine arrest into a roscovitine arrest, roughly corresponding to G1-S-G2, or starting from the release from the second thymidine arrest into a roscovitine arrest, roughly corresponding to early S-phase through G2 (Figure 3.4 B). Despite near complete degradation of these two centromere proteins (Supplemental Figure 3.5 A-C), we observed no significant differences in total CENP-A levels as assayed by immunofluorescence (Figure 3.4 A and 3.4 B). To distinguish between the maintenance of CENP-A in chromatin and the

new assembly of CENP-A-nucleosomes, we made a stable cell line that constitutively overexpressed SNAP-tagged CENP-A at about 8x the endogenous level. We then used fluorescent-pulse labeling of the SNAP-tag with TMR Star to measure the levels of preexisting CENP-A in chromatin, newly assembled CENP-A, or total SNAP-tagged CENP-A (Figure 3.4 C). We observed a significant decrease in centromeric TMR signal when we labeled the total SNAP-tagged CENP-A or the newly assembled SNAP-tagged CENP-A but, in contrast to published observations (Guo et al., 2017), we saw no difference in the SNAP-tagged CENP-A previously incorporated into the centromere. We were unable to reproduce a loss of previously incorporated CENP-A in either our cell lines or the published cell lines using the published methods (Supplemental Figure 3.6 and Discussion). Our data demonstrating that CENP-A assembly depends upon CENP-N is consistent with our previous observations (Carroll et al., 2009) but also demonstrates that the loss of CENP-N has no effect on the maintenance of preexisting CENP-A at centromeres.

Loss of CENP-C and/or CENP-N does not alter the salt extraction of CENP-A from centromeric chromatin

Although we saw no change in CENP-A levels in chromatin after degradation of CENP-C and/or CENP-N this does not directly assay nucleosome stability in chromatin. We therefore measured the ease with which centromeric CENP-A could be extracted with salt in the presence or absence of CENP-C and/or CENP-N. The difference in CENP-A nucleosome stability that we see *in vitro* predicts that we would extract CENP-A from chromatin at lower salt concentrations in the absence of CENP-C and/or CENP-N. We

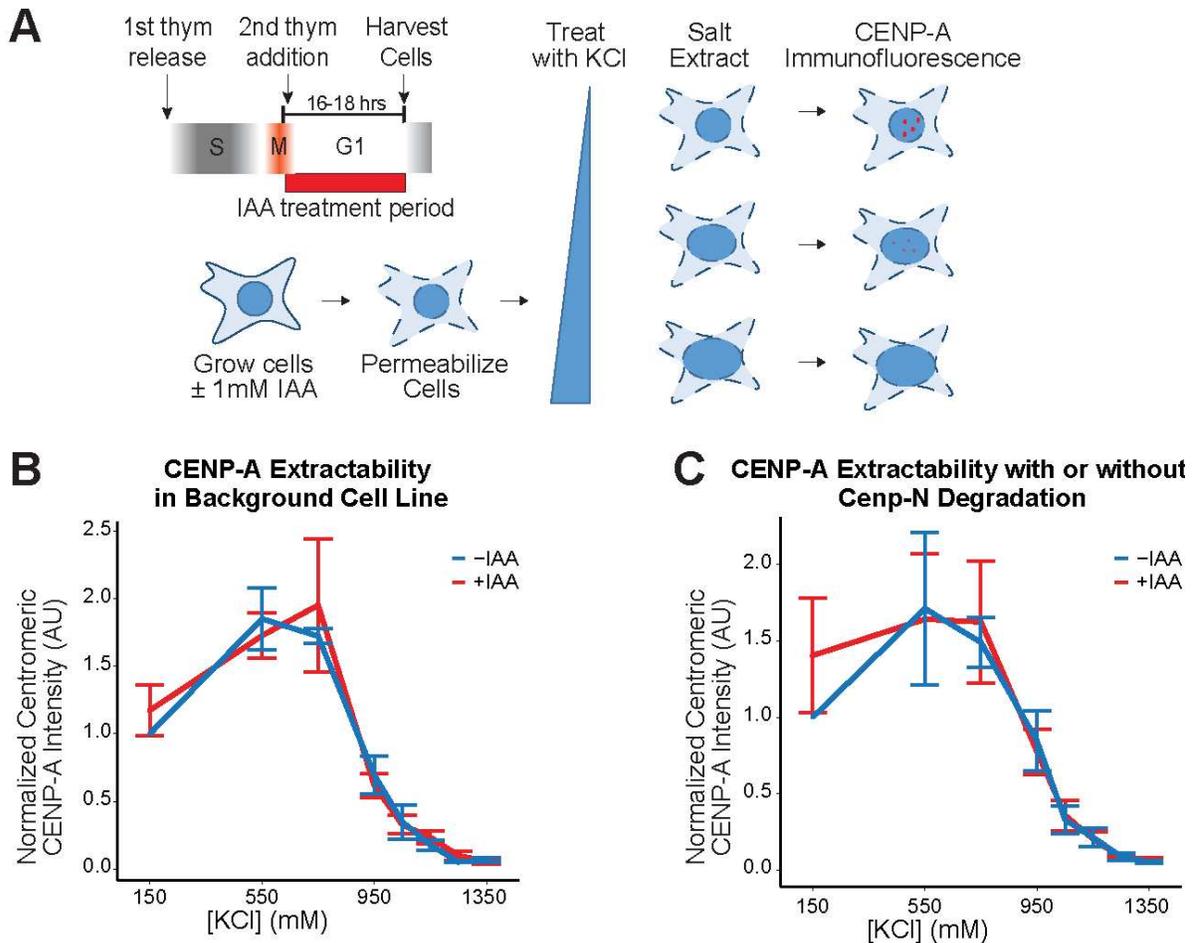


Figure 3.5 CENP-C or CENP-N alone does not affect the salt-extractability of centromeric CENP-A.

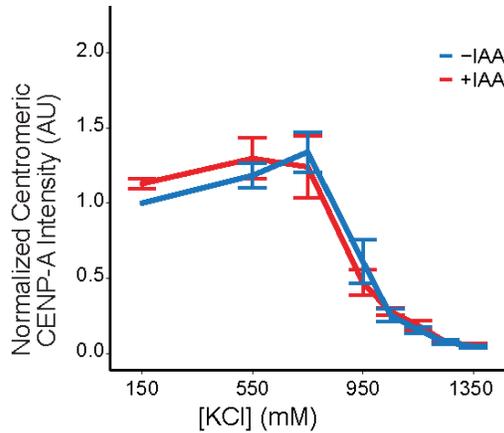
A. Schematic of experimental workflow. Cells were treated with different concentrations of KCl to extract CENP-A from chromatin. Cells were then fixed and stained for CENP-A before imaging.

B. IAA alone has no effect on CENP-A extractability. CENP-A extraction with increasing salt concentration in the parental OsTIR1 expressing cell line with no AID tagged proteins.

C-E. Salt extractability curves for three different cell lines where either CENP-C, CENP-N or both were tagged with AID. Cells were left untreated or treated with IAA to degrade AID-CENP-N (C), AID-CENP-C (D) or both (E). Centromeric CENP-A immunofluorescence signal was normalized to the no IAA, 150 nM KCl signal.

Data are presented as mean \pm SEM for three independent replicates.

D CENP-A Extractability with or without CENP-C Degradation



E CENP-A Extractability with or without CENP-C and N Degradation

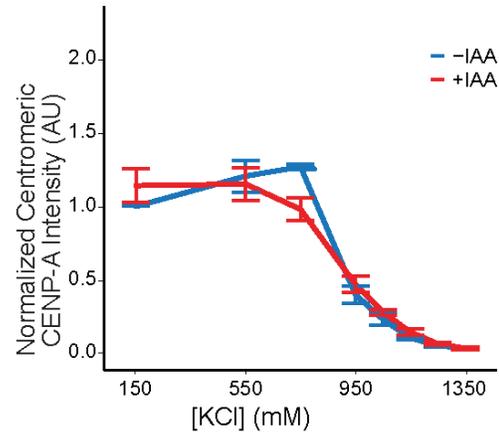


Figure 3.5 CENP-C or CENP-N alone does not affect the salt-extractability of centromeric CENP-A.

C-E. Salt extractability curves for three different cell lines where either CENP-C, CENP-N or both were tagged with AID. Cells were left untreated or treated with IAA to degrade AID-CENP-N (C), AID-CENP-C (D) or both (E). Centromeric CENP-A immunofluorescence signal was normalized to the no IAA, 150 nM KCl signal.

Data are presented as mean \pm SEM for three independent replicates.

permeabilized cells and treated them with increasing concentrations of KCl as previously described (Moree et al., 2011), removed the salt and localized CENP-A using immunofluorescence (Figure 3.5 A). IAA treatment did not have a significant effect on CENP-A salt extraction in cells with no AID tagged proteins (Figure 3.5 B). Degrading either CENP-C or CENP-N for 16-18 hours during G1 also resulted in no difference in salt extraction of CENP-A (Figure 3.5 C and 3.5 D). This lack of difference was not due to loss of CENP-C or CENP-N during the salt extraction because both proteins remained localized at centromeres at all tested KCl concentration without IAA treatment, independent of CENP-A levels (Supplemental Figure 3.5 D-F). When both CENP-C and CENP-N were degraded, we saw a slight effect on CENP-A extractability at 600 mM KCl (p -value = 0.076) but no significant difference at higher or lower salt concentrations (Figure 3.5 E). Our data suggest that the presence of CENP-N and CENP-C at centromeres does not have a strong stabilizing effect on CENP-A nucleosomes *in vivo*.

Mutation of the CENP-A L1 loop to reduce CENP-N affinity does not affect CENP-A or CATD chimera maintenance at centromeres.

CENP-C and CENP-N are both essential proteins making it impossible to test whether long term loss of these factors alters CENP-A nucleosome stability in dividing cells. To circumvent this issue, we engineered stable cell lines that constitutively express four different mutated forms of CENP-A to 4-8x the endogenous level: wild-type CENP-A, a chimeric CENP-A with the CENP-A CATD but the corresponding N and C-terminus from H3.1 or these two species where residues R80 and G81 on CENP-A were mutated to alanine (Figure 3.6 A). The CATD chimera should neither engage CENP-B nor CENP-C

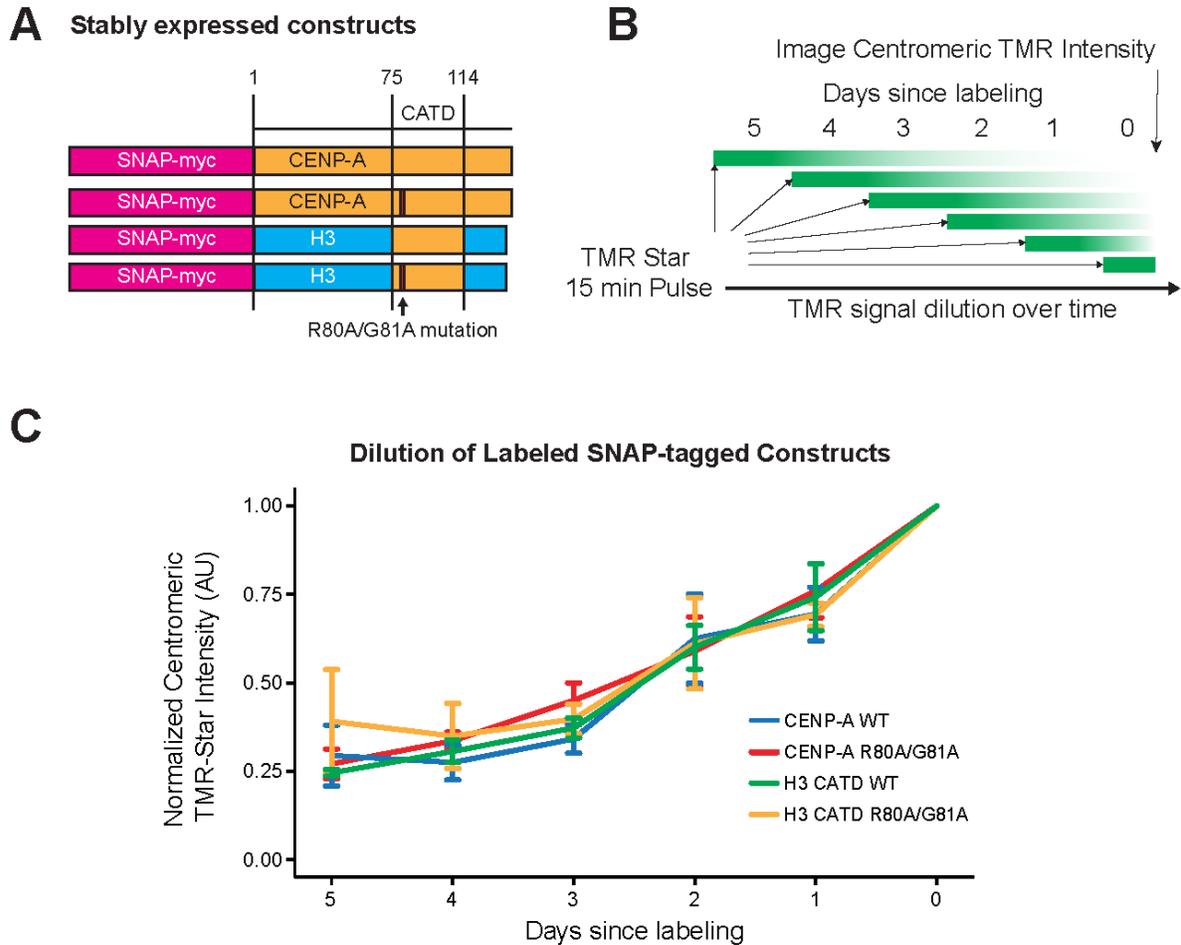


Figure 3.6 R80A G81A mutation does not affect CENP-A or CATD chimera maintenance at centromeres.

A. Schematic of constructs stably integrated into cells for this experiment. Mustard colored bars represent CENP-A sequences. Blue colored bars represent H3.1 sequences.

B. Schematic of labeling scheme used for C). Cells were seeded on coverslips at the same density 5 days before harvest. Cells were labeled with TMR-Star for 15 minutes on different days so that the fluorescent population of CENP-A is diluted to different degrees when cells are harvested at day 0.

C. Centromeric TMR-star intensity was determined by microscopy. Signals across the different cell lines were normalized to the intensity at 0 days since labeling. Intensities were not background subtracted because of variable nuclear background. There are no significant differences between the four curves. Data are presented as mean \pm SEM for three independent replicates.

(Carroll et al., 2010; Fachinetti et al., 2013, 2015; Fujita et al., 2015; Guse et al., 2011) and the CENP-A-R80A,G81A mutation reduces CENP-N's affinity for CENP-A nucleosomes (Fang et al., 2015; Pentakota et al., 2017). These mutations make it possible to compare the long-term maintenance of these four CENP-A mutants as a readout for CENP-A stability when CENP-N and CENP-C binding are perturbed. We measured the CENP-A transgene levels at centromeres by fluorescently labeling cells with a short pulse of SNAP-Cell® TMR-Star each day over the course of five days (Figure 3.6 B). The kinetics of loss were equivalent between all four species until day five after pulse labeling when we could no longer reliably detect a centromeric signal (Figure 3.6 C). Our results recapitulated published data showing that the CATD chimera was maintained similarly to normal CENP-A at centromeres (Bodor et al., 2013). Our findings are consistent with our conclusion that CENP-C and CENP-N binding to CENP-A nucleosomes does not have a strong stabilizing effect on CENP-A in nuclear chromatin.

Discussion

CENP-A nucleosomes were previously shown to be more persistent in chromatin than most H3 nucleosomes (Bodor et al., 2013). However, what precisely causes stabilization of these nucleosomes at the molecular level is unknown. We hypothesized that direct interactors of the CENP-A nucleosome, CENP-C and CENP-N, could stabilize the CENP-A nucleosome resulting in more stable retention. To test this

hypothesis, we examined the functions of CENP-C and CENP-N in stabilizing CENP-A mono-nucleosomes *in vitro* and CENP-A chromatin at centromeres.

We observed that CENP-N (additively with CENP-C) stabilizes the penultimate 10 bp of DNA in a single CENP-A nucleosome, thereby protecting it from disassembly caused by dilution, increased ionic strength, or adsorption onto various EM grids. A recent report pointed out that CENP-N could stabilize the CENP-A nucleosome by fastening CENP-A and DNA (Guo et al., 2017). This was shown very recently by cryo-EM studies of a CENP-A nucleosome in complex with CENP-N (Pentakota et al., 2017; Sagar, Chittori; Jingjun Hong Hayden Saunders, Hanqiao Feng, Rodolfo Ghirlando, Alexander E. Kelly², Yawen Bai², 2017). While no direct interaction with DNA ends was observed in these structures, CENP-N has a large (15 bp) DNA interaction interface near the CENP-A L1 loop, with which it also interacts extensively. Possibly, breathing of DNA ends is prevented by the contacts between CENP-N and nucleosomal DNA further into the particle. Our *in vitro* data is consistent with the structures in that CENP-A mononucleosomes form much more stable particles in the presence of CENP-C and/or CENP-N.

In contrast, in cells, we do not observe a significant change in centromeric CENP-A levels in the absence of CENP-C, CENP-N, or both. Our results are inconsistent with a recently published report (Guo et al., 2017) that shows a loss of CENP-A in chromatin when CENP-N is removed. In attempts to resolve this difference, we obtained the

CENP-N-AID-GFP cell line with SNAP-tagged CENP-A from Guo et al. We directly compared their cell line with the ones we used in this study and replicated their experimental protocol. Despite this, we found no significant difference between the IAA-treated and control cells in any of the cell lines (Supplemental Figure 3.6 E). Notably, when the data was analyzed as in Guo et al., the cell line obtained used in that study exhibited the largest difference, a $23\% \pm 20\%$ decrease upon treatment with IAA-treated, when both centromeric TMR-Star signal (Supplemental Figure 3.6 E, top) and changes in cell number (Supplemental Figure 3.6 E, middle) were factored in. However, this observed difference (p-value 0.3485) might be largely due to non-specific, IAA-dependent changes in cell number (Supplemental Figure 3.6 E, middle) as a cell line expressing no AID-tagged proteins showed a $16 \pm 19\%$ decrease when treated with IAA, when changes in cell number were considered. Without cell number normalization, this value would be $3 \pm 8\%$. Though there are modest differences in SNAP-CENP-A and TIR1 expression levels between the cell lines (Supplemental Figure 3.6 B), these differences cannot account for the lack of destabilization of CENP-A nucleosomes in our hands. Thus, while we observe a stabilization of the CENP-A nucleosome by CENP-C and CENP-N *in vitro*, removing CENP-C and CENP-N *in vivo* does not result in significant destabilization of CENP-A in chromatin. We found it surprising that we did not observe any loss of retention of CENP-A nucleosomes without CENP-C and/or CENP-N in cells given the stabilizing effect of these proteins *in vitro*. We propose three possible explanations for why the loss of CENP-C and CENP-N has relatively little effect on CENP-A nucleosome stability *in vivo*, despite the demonstrated stabilizing effect on the CENP-A nucleosome by CENP-C and CENP-N *in vitro*. One possibility is that CENP-A

nucleosomes interact with other factors at the centromere that promote their stability. For example, CENP-B interacts with the N-terminal tail of CENP-A and also binds to DNA, and thus may compensate for the loss of CENP-C or CENP-N. Currently only CENP-N, CENP-C and CENP-B have been identified as centromere specific CENP-A nucleosome binding proteins. It was recently shown that M18BP1/KNL2 can also bind directly to CENP-A nucleosomes in frogs (French et al., 2017), birds (Hori et al., 2017) and plants (Sandmann et al., 2017), suggesting that other factors may also exist in human cells that bind and stabilize CENP-A nucleosomes.

A second interpretation of our findings is that despite the requirement for CENP-N and CENP-C in centromere assembly, the interaction with CENP-A may be less critical once the CCAN has properly assembled. We found that overexpression of full-length Ruby-tagged CENP-N displaced GFP-tagged CENP-N expressed at endogenous levels, while the overexpression of truncated versions of CENP-N that lacked the CENP-L interaction domain did not displace GFP-tagged CENP-N, despite being able to localize at centromeres (Supplemental Figure 3.4 D and 3.4 E). The fact that their localization does not displace endogenous CENP-N suggests two possible models: 1) that the number of CENP-A binding sites exceeds the number of full length CENP-N molecules at the centromere; or 2) that once CENP-N localizes to the centromere by interacting with CENP-A, it no longer requires CENP-A interaction to remain at the centromere and instead depends on its interactions with other members of the CCAN. This second model is supported by our cell-based salt wash assays in which CENP-A levels decreased with increasing salt yet leaving CENP-C and CENP-N localization to the

centromere unaffected (Supplemental Figure 3.5 D-F). A recent study demonstrated that complete removal of CENP-A from the centromere did not result in a significant decrease in CENP-N that was already bound at the centromere (Hoffmann et al., 2016). A portion of centromeric CENP-C is also resistant to either degradation or replicative dilution of CENP-A (Fachinetti et al., 2013; Hoffmann et al., 2016). If the second model were correct, we cannot assume that CENP-C or CENP-N constitutively bind CENP-A directly, and in the absence of such an interaction we would expect no difference in CENP-A stability without CENP-C or CENP-N.

Third, the environment surrounding CENP-A nucleosomes differ substantially *in vitro* and in cells. *In vitro*, we observed changes in the stability of CENP-A nucleosomes in response to CENP-N binding at intermediate salt concentrations. While these measurements represent real differences in the stability of the CENP-A nucleosome, this stability change may not translate to the cellular environment where there is substantial crowding and regulated levels of monovalent and divalent cations. *In vitro*, a substantial stability difference with and without CENP-N was observed only above 150 mM NaCl. Similarly, in the dilution assay, the presence of CENP-N did not result in additional stabilization at 300 nM nucleosome concentration, while there was a reproducible 3-fold effect at 150 nM. While it is difficult to assess the local concentration of CENP-A in the cell, we observed that CENP-A appears as a diffraction limited spot in microscopy. Using the estimated number of 200 CENP-A nucleosomes per centromere (Bodor et al., 2014), and assuming that the centromere occupies a sphere with a 300 nm diffraction limited radius, then the concentration of CENP-A

nucleosomes would be approximately 3 μM , and thus beyond the 'critical concentration'. The actual concentration of all nucleosomes if we include H3 nucleosomes would be even higher.

Since we did not observe a change in the retention of CENP-A nucleosomes in cells, the nature of this retention remains elusive. Rather than relying on CCAN interactions, CENP-A nucleosomes could be locked in place by interaction with other nearby nucleosomes brought into proximity by the 3-dimensional organization of the chromatin fiber. Several other factors such as post-translational modifications, RNA interactions or yet to be discovered CENP-A interacting proteins might also contribute to CENP-A nucleosome stability. It may be that all nucleosomes meet the basic level of stability to demonstrate this retention in cells but CENP-A nucleosomes may be protected from destabilizing processes that displace nucleosomes from chromatin. Several studies demonstrate that a fraction of H3 nucleosomes persist on chromatin in cycling cells (Kimura and Cook, 2001; Radman-Livaja et al., 2011).

Also, while we focused on rapidly dividing cells, Smoak et al. recently showed that there is long-term retention of CENP-A at centromeres in mouse oocytes for greater than a year with no detectable assembly (Smoak et al., 2016). However, this may not be unique as H3 is one of the longest-lived proteins in post-mitotic tissues (Commerford et al., 1982; Toyama et al., 2013; Waterborg, 1993). Unfortunately, experiments in cultured cells can only access much shorter timescales, and thus the stabilization of CENP-A by

CENP-C and CENP-N that we see *in vitro* may be relevant when observed over the lifetime of an oocyte or egg.

Materials and Methods

Protein purification Cenp-N¹⁻²⁸⁹ with a (His)₆ tag at the C-terminus was constructed in plasmid pACEbac1 and expressed in Sf21 insect cells. Infected cells were collected and suspended in lysis buffer (50 mM sodium phosphate buffer, pH7.5, 300 mM NaCl, 10% glycerol). Cells were sonicated in presence of Benzonase nuclease (2 µl per 40 ml). Nickel beads, equilibrated with lysis buffer, were added and the mixture was incubated overnight at 4°C. Beads were washed and protein was eluted with a 40- to 500 mM imidazole buffer in lysis buffer. Size exclusion chromatography over a GE S200 column was performed in 20 mM Tris HCl, pH7.5, 300 mM NaCl, 10% glycerol, 1 mM DTT, and 0.5 mM EDTA. CENP-A-H4 was co-expressed from a bicistronic plasmid in *E.Coli*, and purified as described (Guse et al., 2012).

Nucleosome reconstitution DNA, (CENP-A - H4)₂ tetramer, and H2A-H2B were mixed at a 1 to 1.1 to 2.2 molar ratio in buffer containing 20 mM Tris HCl, pH 7.5, 2 M NaCl, 1 mM DTT and 1 mM EDTA. Reconstitutions were either performed by step-wise dilution or by dialysis against TEN buffer (20 mM Tris HCl, pH 7.5, 1 mM EDTA, 10 mM NaCl and 1 mM DTT) as described (Muthurajan et al., 2016). Some preparations of

nucleosome were also reconstituted with atto N 647 dye labeled H2B (T112C) (D'Arcy et al., 2013).

The following DNA templates were used in this study:

Widom '601' sequence DNA (147 bp; (Lowary and Widom, 1998)):

```
ATCTGAGAATCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCTCTAGCACCG  
CTTAAACGCACGTACGCGCTGTCCCCCGCGTTTTAACCGCCAAGGGGATTACTCC  
CTAGTCTCCAGGCACGTGTCAGATATATACATCCGAT
```

Widom '601' sequence DNA (127 bp):

```
CCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCTCTAGCACCGCTTAAACGCA  
CGTACGCGCTGTCCCCCGCGTTTTAACCGCCAAGGGGATTACTCCCTAGTCTCCA  
GGCACGTGTCAGATATA
```

Widom '601' sequence DNA (166 bp):

```
ATCCCTATACGCGGCCGCCCTGGAGAATCCCGGTGCCGAGGCCGCTCAATTGGTC  
GTAGACAGCTCTAGCACCGCTTAAACGCACGTACGCGCTGTCCCCCGCGTTTTAA  
CCGCCAAGGGGATTACTCCCTAGTCTCCAGGCACGTGTCAGATATATACATCCGAT
```

DNA derived from α satellite DNA (147 bp):

```
ATCAATATCCACCTGCAGATACTACCAAAGTGTATTTGGAACTGCTCCATCAAAA  
GGCATGTTTCAGCTGGATT  
CCAGCTGAACATGCCTTTTGATGGAGCAGTTTCAAATACACTTTTGGTAGTATCTG  
CAGGTGGATATTGAT
```

Binding assay Nucleosomes or DNA were mixed with the indicated amount of Cenp-N¹⁻²⁸⁹ or CENP-C⁴²⁶⁻⁵³⁷ in buffer (20 mM Tris HCl, pH 7.5, 1 mM EDTA, 50 -150 mM NaCl and 1 mM DTT) for 30 minutes at room temperature, then analyzed by 5% native PAGE. The gel image was captured at 488 nm (for SYBR Gold staining) or 647 nm (for labeled 18 bp DNA) using a Typhoon imager.

Micrococcal nuclease digestion Nucleosome (166 bp DNA, 300 nM, 20 μ l) was incubated with 1U/100 ng MNase in 20 mM Tris HCl, pH 7.5, 1 mM EDTA, 50 mM NaCl and 1 mM DTT at 37°C for the 5 minutes. Reactions were quenched by addition of 50 mM EDTA. Native PAGE was used to analyze the digestion without DNA extraction. In parallel, DNA from the treated sample was extracted with the Minelute DNA extraction kit, followed by quantification on a bioanalyzer (Agilent). The length of DNA fragment was calculated from the standard DNA ladder. For CENP-A nucleosome containing 147 bp α satellite DNA, 2U/100 ng MNase were added into reaction. The reaction was quenched after 10 minutes' incubation at 37°C under same buffer condition above.

Nucleosome stability test assay (gel based) Assembled nucleosomes were incubated in buffer (20 mM Tris HCl, pH 7.5, 1 mM EDTA, and 1 mM DTT) containing the indicated concentration of NaCl. Alternatively, nucleosome samples were diluted with buffer at room temperature or 55°C. Nucleosome stability was defined by dividing the intensity of the remaining nucleosome intensity by that of the input nucleosome on

native PAGE. All raw signals were detected using a Typhoon at 647 nm (H2B in nucleosomes was labeled on residue T122C with atto N 647).

Cryo-EM grid preparation and microscopy for testing stability of nucleosome

Fresh preparations of nucleosomes were divided into two equal aliquots. One was mixed with a 2-3 fold excess of either CENP-N¹⁻²⁸⁹ or CENP-C⁴²⁶⁻⁵³⁷/CENP-N¹⁻²⁸⁹, while the other was mixed with the same amount of buffer to adjust buffer components and sample concentration. 4 μ l of each sample (final concentration of 2.5 μ M) was loaded on to C-Flat(Au) grid, and blotted for 4s in a FEI Vitrobot before plunge freezing in liquid ethane. Samples were imaged at magnification 62Kx by a FEI F30 electron microscope with a Gatan K2 Summit direct detector device. MotionCor2 (Zheng et al., 2017) were used to align 45 frames for each image. Relion2.1 (Scheres, 2016) was used to pick the nucleosome like particles. Particles were extracted and quantified in Relion.

Negative grid preparation and microscopy for testing nucleosome stability

Samples were prepared as above. Final concentration was ~ 30 nM. 4 μ l of each sample was loaded onto holey carbon grids for 30s. Grids were rinsed twice with water. 1% Uranyl Acetate was used to fix and stain the particle. Images were acquired on a FEI F20 microscope equipped with an CCD detector. Particles were picked and quantified by using dogpick in Appion(Lander et al., 2009). Nucleosome particles were defined by the diameter around 10-16 nm.

Cell Culture All cell lines were cultured in RPMI-1640 medium supplemented with 10% fetal bovine serum, 100 U/ml penicillin/ 0.1 mg/ml streptomycin, 2 g/L sodium bicarbonate, and 2 µg/mL puromycin. Degradation during G1 was achieved by treating with 0.5 mM thymidine for 16-18 hours, release into 17 µM deoxycytidine for 3 hours, normal media for 6 hours followed by a second 0.5 mM thymidine arrest and the addition of 1 mM IAA. Degradation during G1 and S phases were done by adding 1 mM IAA during an additional 3 hours release into 17 µM deoxycytidine followed by 20-22 hours in 0.1 mM roscovitine. All drugs were added directly to the media. Expression of transgenes were induced using 1 µg/ml doxycycline.

Generation of Cell Lines CRISPR-Cas9 based genome engineering was used to tag CENP-N with AID-sfGFP at the C-terminus (guide RNA target sequence: GTGCATGTGCAATATCAAGA, 500 base pair homology arms flanking the stop codon were used in a pUC18 backbone donor construct) in osTir1 Flp-In TRex-DLD1 cells (a gift from the Don W. Cleveland Lab, UCSD). CENP-C was tagged with AID-mRuby2-3xFlag at the C-terminus (guide RNA target sequence: GAATGAGTAGACATATTAATC, 1292 base pair and 1251 base pair homology arms before and after the stop codon respectively in a pUC18 backbone donor construct) in the CENP-N-AID-sfGFP background to generate double AID cell lines. Cells were co-transfected with 0.5 µg of plasmid expressing Cas9-GFP (pX458-Addgene 48138) (for CENP-C tagging) or Cas9-mCherry (modified pX458 with GFP replaced with mCherry, a gift from Rajat Rohatgi) (for CENP-N tagging) and guide RNA targeting sequence and 0.5 µg of homology arm donor constructs using Promega Fugene HD in a 6well plate well. Cells were sorted

using a Sony SH800Z Cell sorter for Cas9 expression 2 days post transfection. Sorted cells were outgrown to confluency and then sorted for single cells that expressed the endogenous fusion tags and counter selected for Cas9 expression. CENP-C-AID-YFP cell lines were a gift from the Dani Fachinetti and Don Cleveland, UCSD. CENP-N rescue constructs and SNAP-CENP-A were induced using the FRT/Flp-mediated recombination of pcDNA5/FRT/TO and pEF5/FRT based vectors respectively and selected with 100 $\mu\text{g/ml}$ Hygromycin B. CENP-A R80A/G81A mutants were generating with the following oligonucleotides: ATTTACCGCTGCCGTGGACTTCAACTGG, AGTCCACGGCAGCGGTAAATTTACGCAGATTTC. Full plasmid sequences are available from the Stanford Digital Repository (<https://purl.stanford.edu/gz478gq3828>). Cloning was done using DH5alpha strains of E. coli bacteria. Cells were tested for mycoplasma contamination at the start of cell line generation by PCR and conditions were monitored by the absence of cytoplasmic DNA staining by microscopy.

CENP-A labeling with TMR-Star Cell labeling or blocking was done by incubating cells with 2 μM SNAP-Cell[®] TMR Star (New England BioLabs) or 6.7 μM SNAP-Cell[®] Block (New England Biolabs) diluted with media for 15 minutes, washed with PBS + 1 mM MgCl_2 and CaCl_2 , incubated in media for 1 hour, washed with PBS + 1 mM MgCl_2 and CaCl_2 and put back in media until ready to harvest.

Immunofluorescence, Microscopy For all but the salt extraction experiments, cells were grown on glass coverslips and washed with PBS + 1 mM MgCl_2 and CaCl_2 ,

permeabilized with PBKCl (139.7 mM KCl, 11.8 mM KH₂PO₄) with 0.5% Triton X-100 for 5 min and fixed in PBKCl/0.5% Triton X-100/3.7% formaldehyde for 10 min, blocked in antibody dilution buffer (20 mM Tris HCl, pH 7.4, 150 mM NaCl, 0.1% Triton X-100, 2% BSA, and 0.1% sodium azide). For salt extraction experiments, cells were permeabilized for 10 min followed by treatment with salt solutions (80 mM K-Pipes, pH 6.8, 1 mM MgCl₂, and 1 mM EGTA, 30% glycerol and 0.5% Triton X-100) with additional KCl added as indicated for 15 min. Fixation was done following staining. Proteins were detected with the following primary antibodies:

<u>Antibody</u>	<u>Dilution or Final Concentration</u>
α-centromere derived from human CREST patient serum (Antibodies, Inc. 15-234-0001)	1:100
Custom Rabbit α-CENP-A	2 µg/ml
α-Flag M2 (Sigma, F3165)	2 µg/ml for CENP-C-mRuby2-3xFlag 5 µg/ml for CENP-N-mRuby2-3xFlag truncations
llama nanobody α-GFP directly conjugated with Alexa Fluor 488	1 µg/ml for salt wash experiments

Primary antibodies were detected by either 568 or 647 directly conjugated goat secondary antibodies (Molecular Probes). DNA was stained with 10 µg/ml Hoechst

33258. Those not stained with the nanobody were imaged using endogenous GFP or YFP fluorescence. Images were acquired on a DeltaVision Core deconvolution microscope equipped with a CoolSnap HQ CCD camera (Photometrics). Z-sections were acquired at 0.2 μm steps using a 60x, 1.4 NA objective. Image analysis was done using centromere finder as in (Moree et al., 2011) (<http://cjfuller.github.io/imageanalysisistools/>). Average background signal was subtracted from centromeric signal. All quantification of microscopy experiments covered three independent experiments with at least 30 cells per condition per experiment.

Immunoblotting Cells were harvested using 0.25% Trypsin-EDTA and quenched with media. Cells were spun down and washed with PBS + 1 mM MgCl_2 and CaCl_2 and washed twice with PBKCl + 0.5% Triton X-100 with 1 mM PMSF and 1 mM Benzamidinium HCl, spinning down at 10,000 x g for 5 min at 4°C between washes. Cells were resuspended in 100 μL denaturing protein lysis buffer (20 mM Tris HCl pH 7.4, 15 mM NaCl, 10 mM EDTA, 0.5% Igepal, 0.1% Triton X-100, 0.1% SDS, 0.1% NaDeoxycholate) and sonicated before being separated by SDS-PAGE and transferred onto polyvinylidene fluoride membrane (Bio-Rad Laboratories). Samples were transferred in CAPS transfer buffer (10 mM 3-(cyclohexylamino)-1-propanesulfonic acid, pH 11.3, 0.1% SDS, and 20% methanol) for 3 hours at 4 °C. Rabbit anti-GFP antibodies (custom), rabbit anti-CENP-C antibodies (custom), mouse anti-Flag M2 antibodies (F1804, Sigma-Aldrich) were used at 5 $\mu\text{g}/\text{ml}$. mouse anti-RNA PolIII (ab5408, Abcam) was used at 0.25 $\mu\text{g}/\text{ml}$. These were detected using donkey anti-rabbit

conjugated to IRDye800 used at 1:10,000 (LiCor) or goat anti-mouse directly conjugated to Alexa Fluor 647 (Invitrogen).

CHAPTER 4 CENP-A nucleosome dynamics and the potential role of CENP-N on higher order structure of centromere

This chapter consists of two unpublished works: 1. The single particle cryo-EM analysis of CENP-A nucleosome in different contexts; 2. The potential role of CENP-N on the higher order structure of chromatin containing CENP-A.

Single particle cryo-EM study on CENP-A nucleosome with and without the presence of CENP-C or CENP-N/CENP-C

Overview

In the crystal structure, the DNA ends of the CENP-A nucleosome are too disordered to be observed. However, they are visible and well organized in the cryo-EM structure of CENP-A nucleosome in complex with CENP-N. To explore the mechanism accounting for this difference, cryo-EM was performed on the CENP-A nucleosomes in different contexts. These studies reveal that the DNA ends coming from stable CENP-A nucleosomes are all well-organized, even when there is no stabilizing factor such as CENP-N and CENP-C in solution. However, both CENP-N and CENP-C increase the number of stable particles, which leads to more CENP-A nucleosomes with tightly wrapped DNA ends. Therefore, the less flexible DNA ends in CENP-A nucleosome we observed in 'Chapter 3' is the consequence of the stabilizing effect from CENP-N, as also described in 'Chapter 2'.

Introduction

'DNA breathing' is used to describe the DNA dynamics in nucleosomes. 'DNA breathing' is defined as the wrapping/unwrapping of DNA at the entry/exit points of the nucleosome (Poirier et al., 2008; Polach and Widom, 1995; Tims et al., 2011). This was recently visualized by cryo-EM of canonical, H3 nucleosomes (Bilokapic et al., 2017). There are four clear conformations of nucleosome observed from cryo-EM structures, each with DNA ends in different states of unwrapping. The dominant particles in the whole population (~ 90%) have DNA ends tightly wrapped which is consistent with the crystal structure of nucleosome. The remaining 10% nucleosomes have partially or completely unwrapped DNA ends. These unwrapped DNA ends induce the instability of nucleosome (Bilokapic et al., 2017). For CENP-A nucleosome, the DNA ends were shown to be disordered in the crystal structure, and thus in the majority of the particles (Tachiwana et al., 2011). This indicates that 'DNA breathing' of CENP-A nucleosome might be more pronounced, and that nucleosomes with partially unwrapped DNA ends are the dominant population in CENP-A nucleosome sample. This was validated by MNase digestion data showed in 'Chapter 3' as well as the data from previous study in our lab (Dechassa et al., 2011) . However, no data was reported on whether there is a small proportion of CENP-A nucleosomes with well-wrapped DNA ends. In 'Chapter 2', the DNA ends of CENP-A nucleosome when in complex with CENP-N are completely wrapped, although less tightly (unwrapped angle $\sim 7^\circ$) compared to in the H3 nucleosome (Figure 4.1). In 'Chapter 3', CENP-N¹⁻²⁸⁹ was shown to decrease the accessibility of MNase to the DNA ends. Both data suggested that CENP-N facilitates the DNA wrapping in CENP-A nucleosome. Since there is no direct interaction between

CENP-N¹⁻²⁸⁹ and the DNA ends based on the cryo-EM structure, one explanation is that CENP-N¹⁻²⁸⁹ could constrain the CENP-A nucleosome “breathing” by simultaneously binding to CENP-A and DNA as suggested by our cryo-EM structure, and by hydrogen/deuterium exchange coupled to mass spectrometry (Guo et al., 2017). In this way, there would be a certain proportion of CENP-A nucleosome with wrapped DNA as observed in the cryo-EM structure of CENP-A nucleosome:CENP-N1-289. To test this hypothesis, cryo-EM was performed on CENP-A nucleosome alone. In addition, samples containing CENP-A nucleosome with CENP-C426-537 or CENP-N1-289/CENP-C426-537 were also imaged.

Result

The DNA ends in cryo-EM electron density map from the sub-population of stable CENP-A nucleosomes are well organized

To explore how the DNA ends of CENP-A nucleosomes are organized when there is no CENP-N in solution, a sample containing only CENP-A nucleosome was plunge-frozen. Since most CENP-A nucleosomes fell apart on the grid when the sample concentration was low, as shown in ‘Chapter 3’, the final concentration of sample for the data collection here was ~ 10 μ M, to increase the number of intact particles. Single particle analysis was performed as described in ‘Chapter 2’. The electron density map was obtained at approximately 10 Å resolution (Figure 4.2 A). Compared to the crystal structure, it is obvious that the missing DNA ends are visible in this density map. Since most nucleosomes fell apart during grid preparation, the intact particles, which were

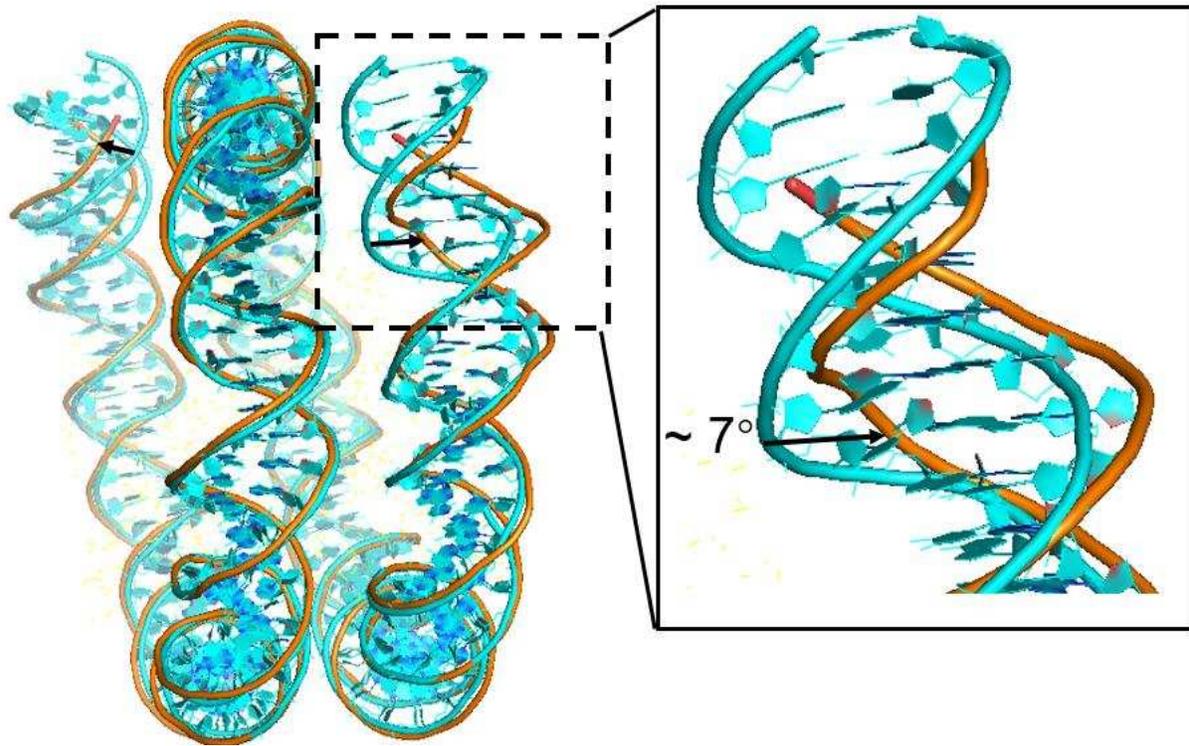


Figure 4.1 DNA ends in H3 nucleosomes wrap slightly tighter than the ones in CENP-A nucleosome with CENP-N

DNA (cyan) from H3 nucleosome (PDB: 3LZ0, (Vasudevan et al., 2010)) was superimposed with the DNA (brown) from cryo-EM structure of CENP-A nucleosome in complex with CENP-N (PDB: 6CW0, (Pentakota et al.)). The arrow highlights the difference on the DNA ends. There is 7° shift between two DNA ends.

less than 10% of the input (Figure 4.3), represent the most stable nucleosomes in the sample. Therefore, the well-organized DNA ends in the electron density map reflect the stable part of whole CENP-A nucleosome population.

All stabilized CENP-A nucleosomes adopt same wrapping conformation of DNA ends

CENP-A nucleosome was shown to be stabilized by CENP-C⁴²⁶⁻⁵³⁷ or CENP-N¹⁻²⁸⁹/CENP-C⁴²⁶⁻⁵³⁷ in 'Chapter 3'. To test whether all of these stabilized nucleosomes have similar levels of DNA wrapping at the ends, cryo-EM was performed on the CENP-A nucleosome sample containing CENP-C⁴²⁶⁻⁵³⁷, or CENP-N¹⁻²⁸⁹/CENP-C⁴²⁶⁻⁵³⁷. The electron density maps from 3D reconstruction showed that the wrapping of DNA ends in these CENP-A nucleosomes are identical to what is observed in CENP-A nucleosomes or CENP-A nucleosomes in complex with CENP-N¹⁻²⁸⁹ (Figure 4.2). Surprisingly, no electron density for either CENP-C⁴²⁶⁻⁵³⁷ or CENP-N¹⁻²⁸⁹ was observed in the map. Cryo-EM can create an unfavorable environment for macromolecular complexes. For example, for a high-quality grid, the solution needs to form a very thin layer on the grid which results in a large water-air interface. The particles could collide with the interface, which would de-stabilize or even denature the complex (Glaeser and Han, 2017). Therefore, both CENP-C⁴²⁶⁻⁵³⁷ and CENP-N¹⁻²⁸⁹ could fall apart before the solution is frozen. This could explain why the majority of particles for reconstructing CENP-A nucleosome: CENP-N¹⁻²⁸⁹ structure only contain one CENP-N (Chapter 2). Meanwhile, since CENP-C⁴²⁶⁻⁵³⁷ is a short disordered peptide, it is also possible that it is not visible

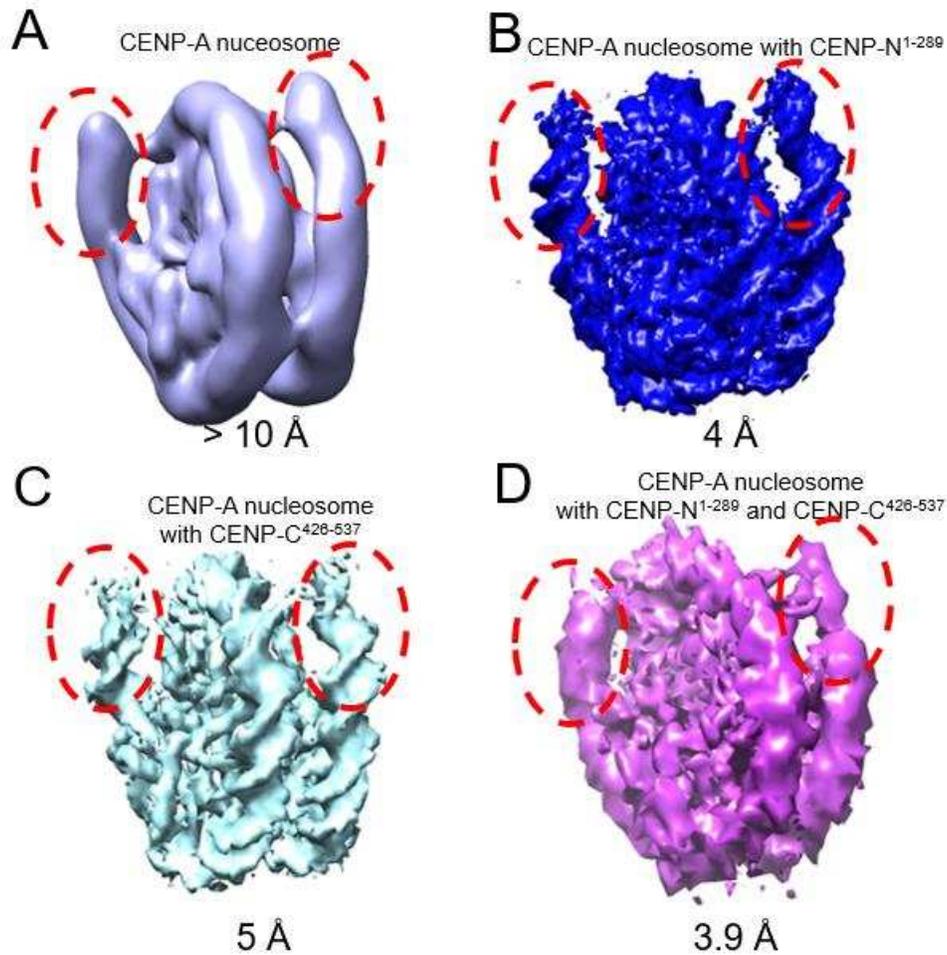


Figure 4.2 Cryo-EM electron density maps of CENP-A nucleosome in different contexts

A. Electron density map of CENP-A nucleosome when there is no other binding protein in solution. The resolution was estimated $> 10 \text{ \AA}$ by Relion 2.1. B. Electron density map of CENP-A nucleosome when CENP-N¹⁻²⁸⁹ is in solution. The resolution was estimated $\sim 4 \text{ \AA}$ by Relion 2.1. C. Electron density map of CENP-A nucleosome when CENP-C⁴²⁶⁻⁵³⁷ is in solution. The resolution was estimated to be around $\sim 5 \text{ \AA}$ by Relion 2.1. D. Electron density map of CENP-A nucleosome when both CENP-N¹⁻²⁸⁹ and CENP-C⁴²⁶⁻⁵³⁷ in solution. The resolution was estimated $\sim 3.9 \text{ \AA}$ by Relion 2.1. The red circles highlight the DNA ends which are missing in crystal structure of CENP-A nucleosome.

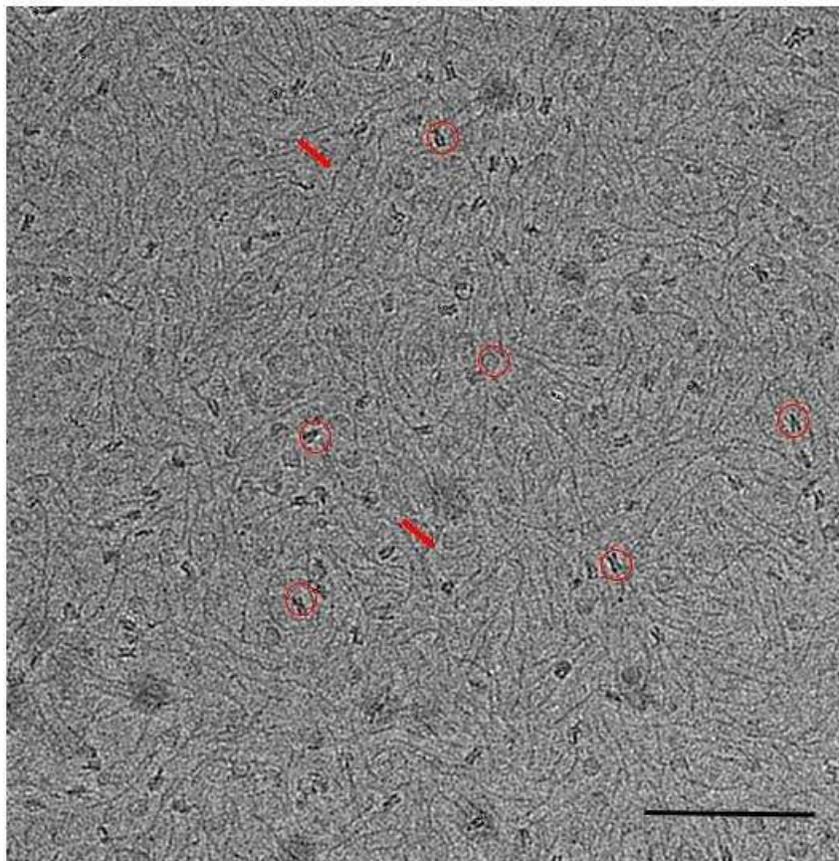


Figure 4.3 ~ 10% CENP-A nucleosome ‘survival’ on cryo-EM grid after plunge freeze

The image represents the data collected for sample only contained CENP-A nucleosome. Red circles highlight the typical “nucleosome-like” particles. The red arrows point out the DNA from the particles that fell apart. The scale bar= 100 nm. Sample concentration was ~ 10 μ M.

due to the low resolution of the map. In summary, my results indicate that the “visible” DNA ends in cryo-EM structure of CENP-A nucleosome represent the most stable state of CENP-A nucleosome.

Discussion

Flexible DNA ends is one of the most significant features unique to the CENP-A nucleosome crystal structure (Tachiwana et al., 2011). However, in the cryo-EM structure of the CENP-A nucleosome in complex with CENP-N¹⁻²⁸⁹ (Chapter 3), the nucleosome DNA ends are organized very well which was also showed by other groups (Chittori et al., 2017; Tian et al., 2018). Previous data from another group showed that the unwrapping states of DNA are dominant for CENP-A nucleosome with 197 bp 601 DNA in solution (Roulland et al., 2016), however, there is a small population of CENP-A nucleosome with DNA ends wrapped. Here, my data showed that the CENP-A nucleosome (147 bp 601 DNA) with well-wrapped DNA ends also exists in solution in absence of interacting proteins, and represents the most stable species of whole population. Recent cryoEM study on H3 nucleosome breathing showed that the histone octamer needs to rearrange subtly to adapt the DNA unwrapping (Bilokapic et al., 2017). Once the DNA unwrapping happens, the α 1-L1- α 2 on H3 moves away from the DNA (SHL 2). If CENP-A nucleosomes shares similar mechanisms for DNA breathing as H3 nucleosome, the movement of CENP-A would be prevented by CENP-N which simultaneously binds to CENP-A “RG” loop and the DNA at SHL 2. Therefore, CENP-N promotes the CENP-A nucleosome breath towards the wrapped state by enhancing the interaction between CENP-A and DNA.

DNA gyres on CENP-A nucleosome slide laterally with respect to each other in solution. CENP-C⁴²⁶⁻⁵³⁷ prevents this DNA gyre sliding in CENP-A nucleosome (Falk et al., 2016). CENP-C⁴²⁶⁻⁵³⁷ binds to the H2A/H2B acidic patch and CENP-A C-terminal 'LEEGLE' motif, which limited the movement of between H2A/H2B and CENP-A. Therefore, CENP-C⁴²⁶⁻⁵³⁷ affects DNA sliding indirectly, and thus the effect of CENP-C⁴²⁶⁻⁵³⁷ on DNA end wrapping could be much more moderate than CENP-N. This is reflected by my cryoEM data that the intact nucleosome particles in the presence of CENP-C, which have well-wrapped DNA ends in electron density map, were not as numerous as those with CENP-N on the cryo-EM grid under same condition (data not shown). MNase digestion result in Chapter 3 showed no significant protection on CENP-A nucleosome DNA ends from CENP-C⁴²⁶⁻⁵³⁷ which also supports the conclusion.

CENP-N^{NT} induces CENP-A mono-nucleosome oligomerization in vitro

Overview

At the centromere, CENP-A nucleosomes are sporadically distributed among canonical nucleosomes. It is not clear how CENP-A nucleosomes, surrounding canonical nucleosomes and kinetochore proteins arrange in the context of higher order structure. By using Sedimentation Velocity Analytical Ultracentrifugation (SV-AUC) and cryo-EM, we show that CenpN^{NT} mediates the oligomerization of CenpA mono-nucleosomes. The H4 tail is also critically involved in this oligomerization. This result indicates that CENP-N could play a role in organizing CENP-A at centromere.

Introduction

Recent data showed that the disorganized or more 'open' DNA ends displayed by CenpA nucleosomes precludes the binding of linker histone H1 in vitro and in vivo (Roulland et al., 2016). Because H1 is implicated in stabilizing the higher order structure of chromatin, it is unknown whether or how chromatin containing CENP-A assumes a compact configuration. One hypothesis is that the kinetochore proteins, especially those in the Constitutive Centromere-Associated Network (CCAN), establish and maintain the local chromatin structure. CENP-C has a dimerization domain in its C-terminal region. The dimerization or oligomerization of CENP-C was suggested to contribute the organization of centromeric chromatin (Sugimoto et al., 1997; Trazzi et al., 2009). However, whether CENP-N plays a role in chromatin structure was not shown before. Based on the cryo-EM structure of CENP-A nucleosome:CENP-N^{NT} complex, the H4 N-terminal tail, which is disordered in most crystal structures, interacts with CENP-N (Figure 4.4) in the loop connecting $\beta 3$ and $\beta 4$. Since the H4 N-tail is involved in chromatin compaction (Bradley et al., 2006; Dorigo et al., 2004), it would be interesting to evaluate whether CENP-N can induce CENP-A chromatin compaction through directing the H4 N-terminal tail. Here, both Analytical Ultracentrifugation (AUC) and cryo-EM were used to test this hypothesis.

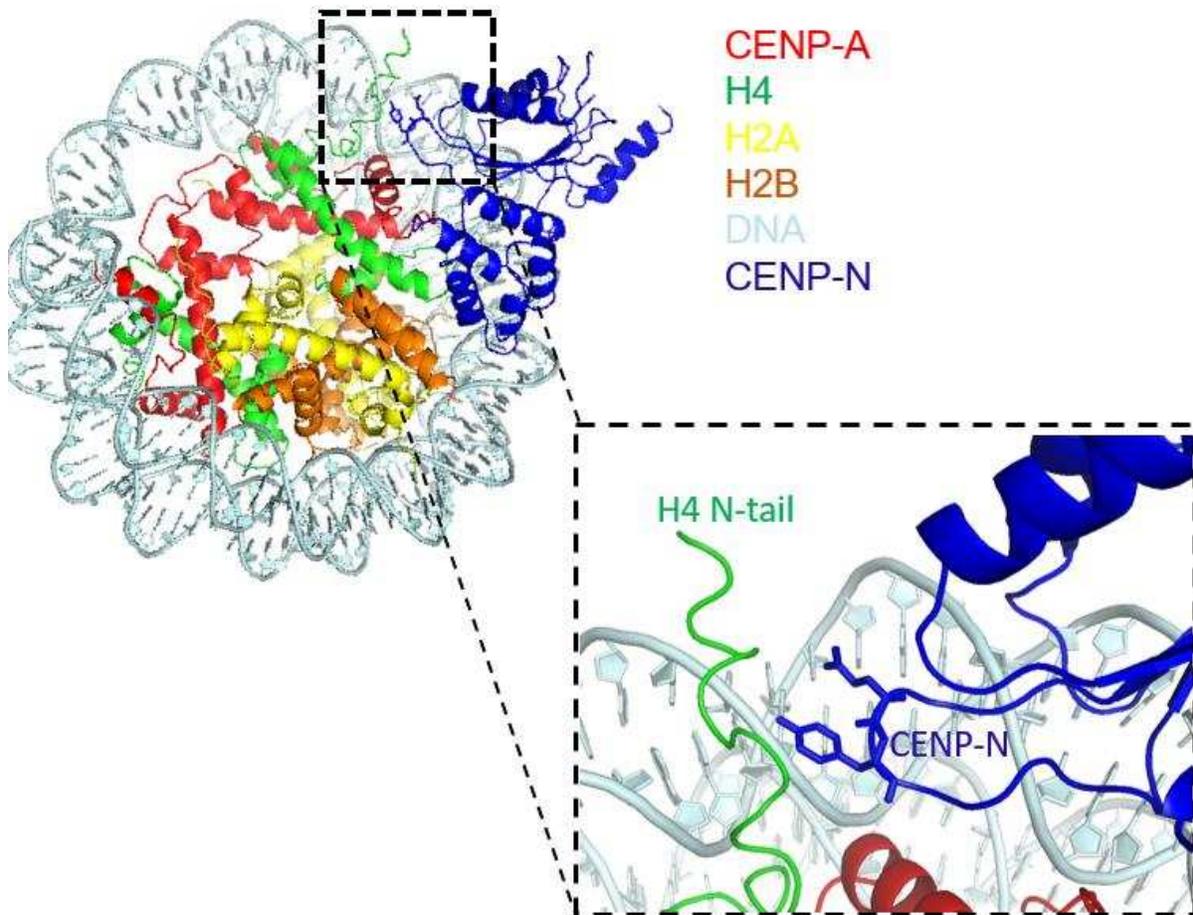


Figure 4.4 CENP-N interacts with H4 N-terminal tail.

The cryo-EM structure indicates that there is interaction between H4 N-tail and the loop between $\beta 3$ and $\beta 4$ on CENP-N (Pentakota et al.). The tail was continually built based on the electron density map (EMD-7326). Different components were labeled by different colors as indicated. The lower panel is the highlighted part of up panel.

Results

CenpN mediates the oligomerization of CenpA mono-nucleosomes

First, SV-AUC was used to evaluate the effect of CenpN¹⁻²⁸⁹ on the association behavior of CenpA mono-nucleosomes. The sedimentation coefficient ($S_{20,w}$ value) of CenpA mono-nucleosomes is around 10.5 S. However, the $S_{20,w}$ value of CenpA nucleosomes in the presence of CenpN increases to larger than 20 S (Figure 4.5), which is close to the s value of the CENP-A tri-nucleosome (three nucleosomes on one DNA template, connected by 30 bp linker DNA). Meanwhile, the sloped distribution of $g(s)$ indicated either the composition or conformation of the complex in solution was very heterogeneous. The s value increases further when the concentration of the salt decreases or the concentration of complex increases. Since CENP-N¹⁻²⁸⁹ molecular weight is estimated around 30 Kd, the total size of complex for one CENP-A nucleosome:CENP-N¹⁻²⁸⁹ is much smaller than two nucleosomes. Therefore, CENP-A nucleosomes together with CENP-N probably are oligomerized. Cryo-EM was used to further validate this. Excitingly, the majority of the CenpA nucleosomes in complex with CenpN are arranged in “ring stacks” on the grid (Figure 4.6 A). No stacks were observed when there was no CENP-N in solution. 2D classification analysis was performed for these particles in the “ring stack”. To understand how nucleosomes were oligomerized, stacks of two nucleosomes were defined as the basic unit for the analysis. Each pair of CENP-A nucleosome particles was picked and extracted for the 2D average calculation, as described in Chapter 2. The classes for the stacks clearly showed that there is a least one CENP-N between two nucleosomes (Figure 4.6 B).

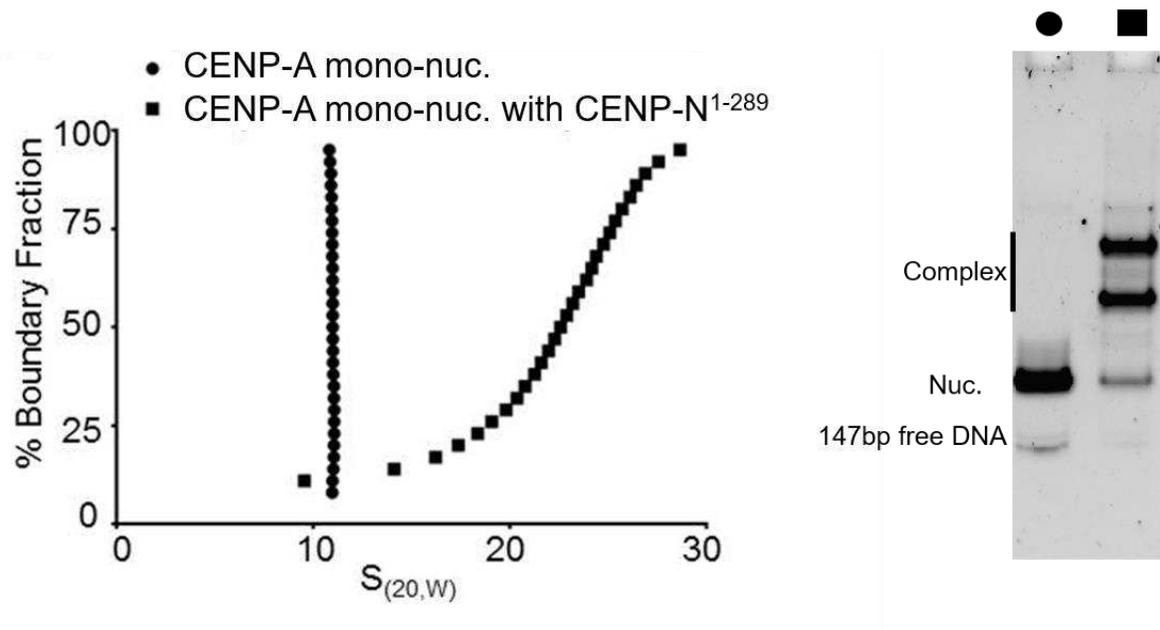


Figure 4.5 CenpN mediates oligomerization of individual CenpA nucleosomes (mono-nucleosome) in solution

Sedimentation coefficient distribution plots (left panel) for CenpA mono-nucleosome and its complex with CenpN¹⁻²⁸⁹ from sedimentation velocity AUC. The $S_{20,w}$ for CENP-A nucleosome is 10.5. The average $S_{20,w}$ of the CENP-A nucleosome in complex with CENP-N¹⁻²⁸⁹ is 25. EMSA (right panel) showed the samples that were analyzed in AUC. "Nuc." is the abbreviation for nucleosome. "Complex" represents CENP-A nucleosome in complex with CENP-N¹⁻²⁸⁹.

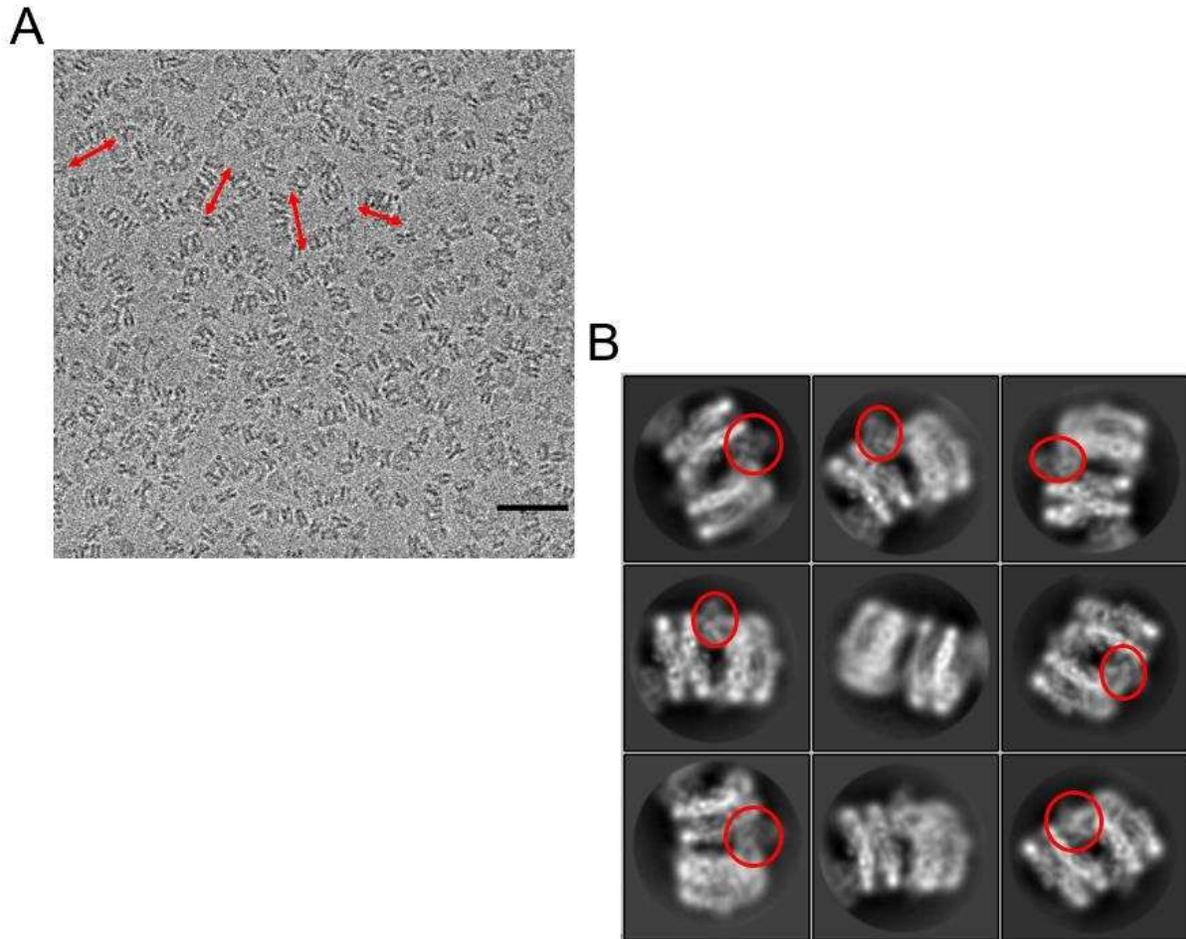


Figure 4.6 CENP-N mediates oligomerization of individual CENP-A nucleosomes (mononucleosome) showed by cryo-EM.

A. Raw image collected for CENP-A nucleosome in complex with CENP-N¹⁻²⁸⁹. The red arrows highlight the “ring stacks” on the grid. The scale bar= 50 nm. B. 2D classification for the basic repeats in the “ring stacks”. The red circles point out the density of CENP-N.

Since the orientation of these stacks is limited to be side-view, it is very challenging to solve the structure.

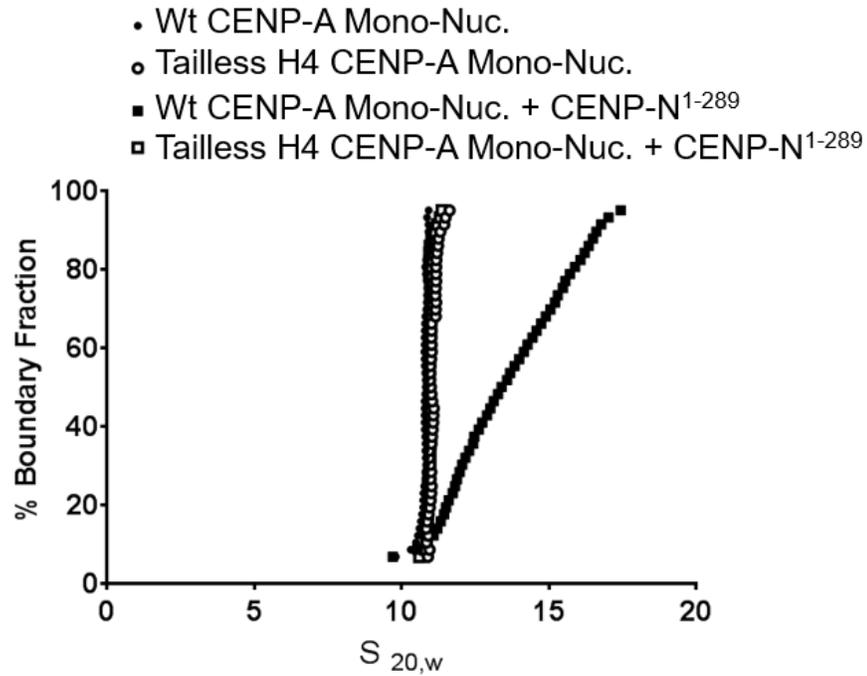
The H4 N-terminal tail is essential for the oligomerization of CENP-A mono-nucleosomes mediated by CENP-N

To test whether this CENP-N mediated oligomerization requires the H4 N-terminal tail, CENP-A nucleosomes were reconstituted with H4 without the N-terminal tail (2-20 deleted, H4 Δ 19). There was no change in the binding of CENP-N to this CENP-A nucleosome with H4 Δ 19, as expected from the only spurious interactions observed by CryoEM. However, unlike CENP-A nucleosome reconstituted with wild type H4, no oligomerization was observed in AUC (Figure 4.7). The average $S_{20,w}$ value for this complex was very close to the $S_{20,w}$ value (10.5 S) of nucleosome. Therefore, the H4 N-terminal tail is essential for CENP-A nucleosome oligomerization shown above.

Discussion

The arrangement of nucleosomes to form higher order structure at the centromere is largely unknown. The interspersion of CENP-A nucleosomes as well as the presence of numerous kinetochore proteins would potentially make the centromere structure very distinct from other regions of the chromosome. Both CENP-N and CENP-C bind to the CENP-A nucleosome surface which is the main target region for chromatin compaction. Therefore, how the CENP-A nucleosome interacts with surrounding canonical nucleosomes or even another CENP-A nucleosome at a distance are intriguing. Here,

A



Tailless H4 = N-terminal a.a. 2-20 deleted

Figure 4.7 H4 N-terminal tail is essential for CENP-A nucleosome oligomerization induced by CENP-N¹⁻²⁸⁹

A. Sedimentation coefficient distribution plots for CenpA mono-nucleosome with wild type H4 or tailless H4 and its complex with CenpN¹⁻²⁸⁹.

B

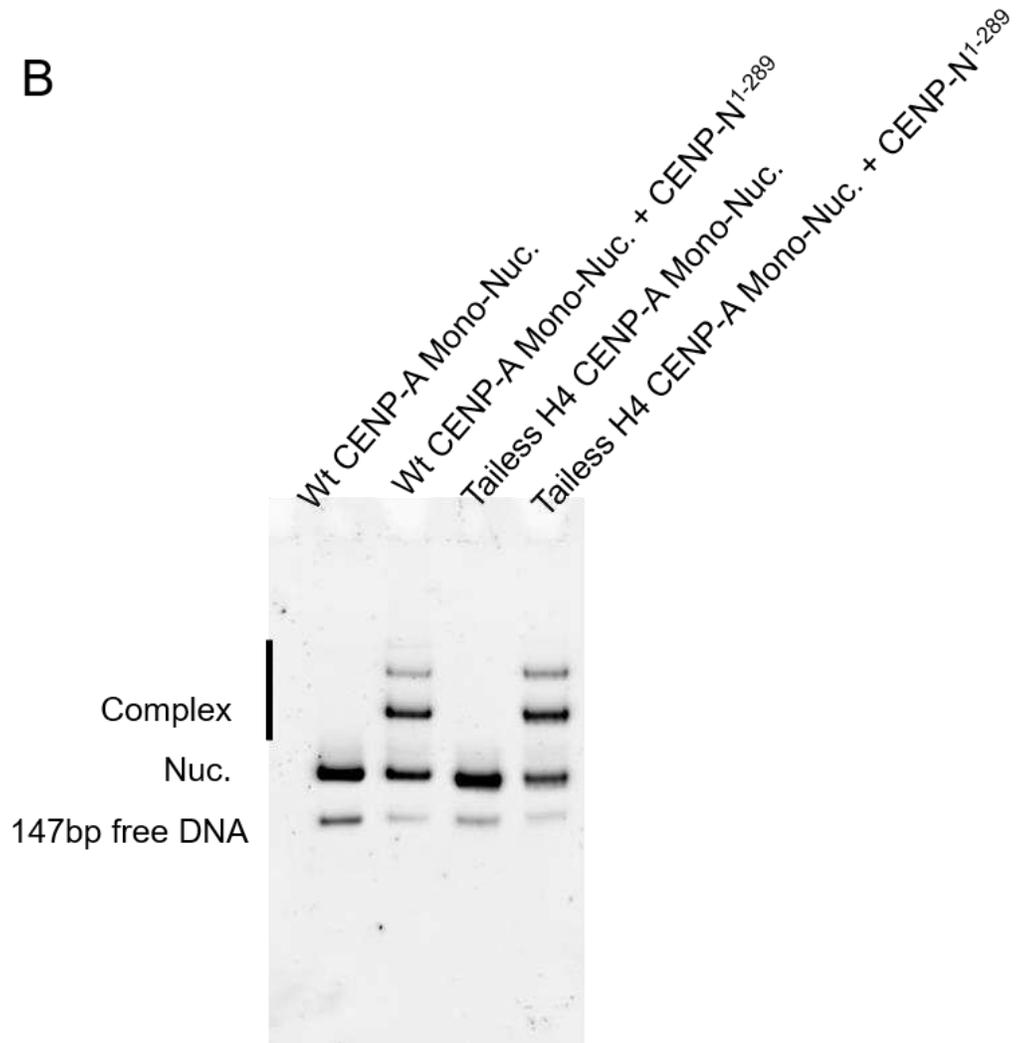


Figure 4.7 The H4 tail is essential for the CENP-A nucleosome oligomerization induced by CENP-N¹⁻²⁸⁹

B. EMSA was performed to show the samples that were analyzed in AUC. “Nuc.” is the abbreviation for nucleosome. “Complex” represents CENP-A nucleosome in complex with CENP-N¹⁻²⁸⁹.

CENP-N^{NT} was shown to be able to oligomerize the CENP-A mono-nucleosomes *in vitro* through H4 N-tail. This provides one potential mechanism for centromeric chromatin compaction or organization. Whether CenpN performs a similar function in the context of a nucleosomal array where nucleosomes are connected by linker DNA is currently being investigated. It is still not clear how CENP-N would perform this function on the level of chromatin which contains both CENP-A nucleosome and H3 nucleosome. CENP-N could induce the H4 tail on CENP-A nucleosome to bind to the neighboring H3 nucleosome. It also could cluster the CENP-A nucleosomes which disperse among H3 nucleosomes in distance. These possibilities need to be investigated in the future.

CHAPTER 5 SUMMARY AND FUTURE DIRECTIONS

For a long time period, the structural basis on how CENP-N recognizes CENP-A nucleosome has been elusive. My main achievement it was to solve the structure of CENP-A nucleosome:CENP-N complex by single particle cryoEM at atomic level (4.0 Å). This provides the first comprehensive understanding on the mechanism of specific binding between CENP-N and CENP-A nucleosome. Consistent with previous biochemical work from other groups (Carroll et al., 2009; Guo et al., 2017), the structure confirms that “RG” loop or L1 loop on CENP-A is recognized by CENP-N. The atomic level insights decipher how CENP-A nucleosome is “decoded” by CENP-N. In addition, the structure uncovers the details of the large interface between CENP-N and ~ 15 bp nucleosomal DNA adjacent to “RG” loop. This contact between CENP-N and nucleosomal DNA could further stabilize the CENP-A nucleosome:CENP-N complex. It also explains why there is substantial binding between CENP-N and H3 nucleosome in low ion strength buffer.

To evaluate the effect of CENP-N on CENP-A nucleosome through binding, the cryo-EM structure of CENP-A nucleosome in complex with CENP-N was compared to the crystal structure of CENP-A nucleosome (3AN2, (Tachiwana et al., 2011)). No significant change was found on CENP-A nucleosome except in two structural features. First, the missing DNA ends of the CENP-A nucleosome in the published crystal structure are mostly visible in cryo-EM structure. By Micrococcal nuclease digestion assay, the DNA ends of CENP-A nucleosome was shown to be less flexible in the

presence of CENP-N. This result demonstrates why the DNA ends of CENP-A nucleosome are visible in cryo-EM structure of the complex. Since there is no direct interaction between CENP-N and the DNA ends in the CENP-A nucleosome, based on the cryo-EM structure, our hypothesis is that CENP-N constrains the nucleosome breathing by simultaneously binding to both CENP-A and DNA. In this way, the CENP-A nucleosome in the presence of CENP-N would be more stable. The stability comparison between CENP-A nucleosome in the absence and presence of CENP-N was analyzed and quantified by EMSA under different “harsh” conditions such as high salt, low concentration and high temperature. Negative staining EM and cryo-EM were also performed to test the stabilizing effect of CENP-N on CENP-A nucleosome. These results show that CENP-N significantly increases the stability of CENP-A nucleosome *in vitro*. In addition, CENP-C, which was proposed to stabilize CENP-A nucleosome, together with CENP-N were shown to additively increase the stability of CENP-A nucleosome. Through additional cryo-EM analysis, I also obtained the electron maps for CENP-A nucleosome with and without CENP-C and CENP-N. These data further supported the hypothesis that CENP-N stabilizes CENP-A nucleosome by constraining “nucleosome breath”. However, no significant stabilizing effect of either CENP-N or CENP-C was observed *in vivo* from our collaborator, contradicting published results by others. This implies that the CENP-A nucleosome is maintained at the centromere by multiple factors. The stabilizing effects of CENP-N and CENP-C on CENP-A nucleosomes are redundant, at least in the short-term.

The second change in the CENP-A nucleosome is that the H4 N terminal tail is more ordered due to its direct interaction with CENP-N. My structural studies show that CENP-N simultaneously occupies nucleosomal DNA around SHL -2 to -3 and binds to H4 N-terminal tail. This mode of interaction resembles the interactions between the ATPase domain of SWI2/SNF2 chromatin remodeler and the H3 nucleosome. CENP-N potentially shields the CENP-A nucleosome from SWI2/SNF2 remodeling. On the other hand, since the H4 N-tail is important for chromatin higher order structure, it will be very interesting to test whether CENP-N brings any effect on chromatin level. My data showed that CENP-N can induce CENP-A nucleosome oligomerization involving H4 tail *in vitro*. It will be exciting to explore the mechanism behind this in future.

Although CENP-N can induce oligomerization of CENP-A mono-nucleosomes *in vitro*, it is still unclear whether this effect persists *in vivo*. There are only 1 in 25 nucleosomes containing CENP-A at centromere. Can CENP-N bridge two CENP-A nucleosomes that are non-adjacent in 'beads-on-a-string' chromatin? Or does CENP-N induce contact between CENP-A nucleosome and its neighboring H3 nucleosomes? It will be necessary to reconstitute the nucleosome arrays with incorporation of CENP-A at designated positions. By using multiple biophysical approaches such as cryo-EM and atomic force microscopy, we can directly observe whether and how CENP-N affects higher order structure of centromeric chromatin. Furthermore, CENP-N associates with other CCAN proteins at the centromere. It is necessary to evaluate the effect of CENP-N on chromatin higher order structure in the context of CCAN. Some key components of CCAN (CENP-CHIKMLN) were reported to form a very stable complex with CENP-A

nucleosome *in vitro* (Weir et al., 2016). CENP-B was also reported to interact with DNA (CENP-B box), CENP-A nucleosome and CENP-C. The next goal on structure study is to solve the structure of this CENP-BCHIKMLN:CENP-A nucleosome mega complex by both cryo-EM (whole complex map) and crystallography (individual component with unknown structure). This will help us to understand how kinetochore proteins are organized around CENP-A nucleosome at atomic level.

REFERENCES

- Abbott, D.W., Chadwick, B.P., Thambirajah, A.A., and Ausió, J. (2005). Beyond the Xi. MacroH2A chromatin distribution and post-translational modification in an avian system. *J. Biol. Chem.* *280*, 16437–16445.
- Ahmad, K., and Henikoff, S. (2002). The Histone Variant H3.3 Marks Active Chromatin by Replication-Independent Nucleosome Assembly ment are not clear. A study in *Tetrahymena* concluded that no protein difference between histone H3 variants was required for replacement histone deposition and. *Mol. Cell* *9*, 1191–1200.
- Allfrey, V.G., Faulkner, R., and Mirsky, A.E. (1964). Acetylation and Methylation of Histones and Their Possible Role in the Regulation of Rna Synthesis. *Proc. Natl. Acad. Sci. U. S. A.* *51*, 786–794.
- Annunziato, A. (2008). DNA Packaging: Nucleosomes and Chromatin. *Nat. Educ.* *1*, 26.
- Arimura, Y., Shirayama, K., Horikoshi, N., Fujita, R., Taguchi, H., Kurumizaka, H., Kagawa, W., and Fukagawa, T. (2014). Crystal structure and stable property of the cancer-associated heterotypic nucleosome containing CENP-A and H3.3. 1–7.
- Bannister, A.J., and Kouzarides, T. (2011). Regulation of chromatin by histone modifications. *Cell Res.* *21*, 381–395.
- Bao, Y., Konesky, K., Park, Y.J., Rosu, S., Dyer, P.N., Rangasamy, D., Tremethick, D.J., Laybourn, P.J., and Luger, K. (2004). Nucleosomes containing the histone variant H2A.Bbd organize only 118 base pairs of DNA. *EMBO J.* *23*, 3314–3324.

Bilokapic, S., Strauss, M., and Halic, M. (2017). Histone octamer rearranges to adapt to DNA unwrapping. *Nat. Struct. Mol. Biol.* *25*, 1–10.

Black, B.E., Foltz, D.R., Chakravarthy, S., Luger, K., Woods, V.L., and Cleveland, D.W. (2004). Structural determinants for generating centromeric chromatin. *Nature* *430*, 578–582.

Bodor, D.L., Valente, L.P., Mata, J.F., Black, B.E., and Jansen, L.E.T. (2013). Assembly in G1 phase and long-term stability are unique intrinsic features of CENP-A nucleosomes. *Mol. Biol. Cell* *24*, 923–932.

Bodor, D.L., Mata, J.F., Sergeev, M., David, A.F., Salimian, K.J., Panchenko, T., Cleveland, D.W., Black, B.E., Shah, J. V, and Jansen, L.E. (2014). The quantitative architecture of centromeric chromatin. *Elife* *3*, e02137.

Bradley, R.S., Hughes, M.K., Crowley, T.J., Baum, S.K., Kim, K.Y., Hyde, W.T., Briffa, K.R., Barnett, T.P., Tett, S.F.B., Bradley, R.S., et al. (2006). Histone H4-K16 Acetylation. *16*, 844–848.

Brinkley, B.R., and Stubblefield, E. (1966). The fine structure of the kinetochore of a mammalian cell in vitro. *Chromosoma* *19*, 28–43.

Cao, S., Zhou, K., Zhang, Z., Luger, K., and Straight, A.F. (2018). Constitutive centromere-associated network contacts confer differential stability on CENP-A nucleosomes in vitro and in the cell. *Mol. Biol. Cell* 1–28.

Carroll, C.W., Silva, M.C.C., Godek, K.M., Jansen, L.E.T., and Straight, A.F. (2009). Centromere assembly requires the direct recognition of CENP-A nucleosomes by

CENP-N. *Nat. Cell Biol.* *11*, 896–902.

Carroll, C.W., Milks, K.J., and Straight, A.F. (2010). Dual recognition of CENP-A nucleosomes is required for centromere assembly. *J. Cell Biol.* *189*, 1143–1155.

Chadwick, B.P., and Willard, H.F. (2001). 2001 - Chadwick, Willard - A novel chromatin protein, distantly related to histone H2A, is largely excluded from the inactive X chromosome.pdf. *152*, 375–384.

Cheeseman, I.M., Chappie, J.S., Wilson-Kubalek, E.M., and Desai, A. (2006). The Conserved KMN Network Constitutes the Core Microtubule-Binding Site of the Kinetochore. *Cell* *127*, 983–997.

Chittori, S., Hong, J., Saunders, H., Feng, H., Ghirlando, R., Kelly, A.E., Bai, Y., and Subramaniam, S. (2017). Structural mechanisms of centromeric nucleosome recognition by the kinetochore protein CENP-N. *Science* (80-.). *2781*.

Chrun, E.S., Modolo, F., and Daniel, F.I. (2017). Histone modifications: A review about the presence of this epigenetic phenomenon in carcinogenesis. *Pathol. Res. Pract.* *213*, 1329–1339.

Chua, E.Y.D., Vogirala, V.K., Inian, O., Wong, A.S.W., Nordenskiöld, L., Plitzko, J.M., Danev, R., and Sandin, S. (2016). 3.9 Å structure of the nucleosome core particle determined by phase-plate cryo-EM. *Nucleic Acids Res.* *44*, 8013–8019.

Clapier, C.R., Chakravarthy, S., Petosa, C., Fernández-Tornero, C., Luger, K., and Müller, C.W. (2008). Structure of the *Drosophila* nucleosome core particle highlights evolutionary constraints on the H2A-H2B histone dimer. *Proteins Struct. Funct. Genet.*

71, 1–7.

Commerford, S.L., Carsten, A.L., and Cronkite, E.P. (1982). Histone turnover within nonproliferating cells. *Proc. Natl. Acad. Sci.* 79, 1163–1165.

Conde, N., Black, B.E., Sivolob, A., Filipski, J., Cleveland, D.W., Prunell, A., Jacques, I., and Umr, M. (2007). CENP-A-containing Nucleosomes: Easier Disassembly versus Exclusive Centromeric Localization. 555–573.

D’Arcy, S., Martin, K., Panchenko, T., Chen, X., Bergeron, S., Stargell, L., Black, B., and Luger, K. (2013). Chaperone Nap1 Shields Histone Surfaces Used in a Nucleosome and Can Put H2A-H2B in an Unconventional Tetrameric Form. *Mol. Cell* 51, 662–677.

Dalal, Y., Wang, H., Lindsay, S., and Henikoff, S. (2007). Tetrameric structure of centromeric nucleosomes in interphase *Drosophila* cells. *PLoS Biol.* 5, 1798–1809.

Dechassa, M.L., Wyns, K., Li, M., Hall, M.A., and Wang, M.D. (2011). Structure and Scm3-mediated assembly of budding yeast centromeric nucleosomes. *Nat. Commun.* 2, 310–313.

Dong, Y., Vanden Beldt, K.J., Meng, X., Khodjakov, A., and McEwen, B.F. (2007). The outer plate in vertebrate kinetochores is a flexible network with multiple microtubule interactions. *Nat. Cell Biol.* 9, 516–522.

Dorigo, B., Schalch, T., Kulangara, A., Duda, S., Schroeder, R.R., and Richmond, T.J. (2004). Nucleosome arrays reveal the two-start organization of the chromatin fiber. *Science* (80-.). 306, 1571–1573.

Earnshaw, W.C., and Rothfield, N. (1985). Identification of a family of human

centromere proteins using autoimmune sera from patients with scleroderma. *Chromosoma* 91, 313–321.

Earnshaw, W.C., Ratrie, H., and Stetten, G. (1989). Visualization of centromere proteins CENP-B and CENP-C on a stable dicentric chromosome in cytological spreads. *Chromosoma* 98, 1–12.

Eberharter, A., and Becker, P.B. (2002). Histone acetylation: A switch between repressive and permissive chromatin. Second in review on chromatin dynamics. *EMBO Rep.* 3, 224–229.

Eskat, A., Deng, W., Hofmeister, A., Rudolphi, S., Emmerth, S., Hellwig, D., Ulbricht, T., Döring, V., Bancroft, J.M., McAinsh, A.D., et al. (2012). Step-Wise Assembly, Maturation and Dynamic Behavior of the Human CENP-P/O/R/Q/U Kinetochores Sub-Complex. *PLoS One* 7.

Fachinetti, D., Folco, H.D., Nechemia-Arbely, Y., Valente, L.P., Nguyen, K., Wong, A.J., Zhu, Q., Holland, A.J., Desai, A., Jansen, L.E.T., et al. (2013). A two-step mechanism for epigenetic specification of centromere identity and function. *Nat. Cell Biol.* 15, 1056–1066.

Fachinetti, D., Han, J.S., McMahon, M.A., Ly, P., Abdullah, A., Wong, A.J., and Cleveland, D.W. (2015). DNA Sequence-Specific Binding of CENP-B Enhances the Fidelity of Human Centromere Function. *Dev. Cell* 33, 314–327.

Falk, S.J., Guo, L.Y., Sekulic, N., Smoak, E.M., Mani, T., Logsdon, G.A., Gupta, K., Jansen, L.E.T., Van Duyne, G.D., Vinogradov, S.A., et al. (2015). CENP-C reshapes and stabilizes CENP-A nucleosomes at the centromere. *Science* 348, 699–703.

Falk, S.J., Lee, J., Sekulic, N., Sennett, M. a, Lee, T.-H., and Black, B.E. (2016). CENP-C directs a structural transition of CENP-A nucleosomes mainly through sliding of DNA gyres. *Nat. Struct. Mol. Biol.* *23*, 204–208.

Fan, J.Y., Rangasamy, D., Luger, K., and Tremethick, D.J. (2004). H2A.Z alters the nucleosome surface to promote HP1 α -mediated chromatin fiber folding. *Mol. Cell* *16*, 655–661.

Fang, J., Liu, Y., Wei, Y., Deng, W., Yu, Z., Huang, L., Teng, Y., Yao, T., You, Q., Ruan, H., et al. (2015). Structural transitions of centromeric chromatin regulate the cell cycle-dependent recruitment of CENP-N. 1058–1073.

Fischle, W., Boo, S.T., Dormann, H.L., Ueberheide, B.M., Garcia, B.A., Shabanowitz, J., Hunt, D.F., Funabiki, H., and Allis, C.D. (2005). Regulation of HP1-chromatin binding by histone H3 methylation and phosphorylation. *Nature* *438*, 1116–1122.

Foltz, D.R., Jansen, L.E.T., Black, B.E., Bailey, A.O., Yates, J.R., and Cleveland, D.W. (2006). The human CENP-A centromeric nucleosome-associated complex. *Nat. Cell Biol.* *8*, 458–469.

Foltz, D.R., Jansen, L.E.T., Bailey, A.O., Iij, J.R.Y., Emily, A., Wood, S., Black, B.E., and Cleveland, D.W. (2010). NIH Public Access. *137*, 472–484.

French, B.T., Westhorpe, F.G., Limouse, C., and Straight, A.F. (2017). *Xenopus laevis* M18BP1 Directly Binds Existing CENP-A Nucleosomes to Promote Centromeric Chromatin Assembly. *Dev. Cell* *42*, 190–199.e10.

Fujita, R., Otake, K., Arimura, Y., Horikoshi, N., Miya, Y., Shiga, T., Osakabe, A.,

Tachiwana, H., Ohzeki, J.I., Larionov, V., et al. (2015). Stable complex formation of CENP-B with the CENP-A nucleosome. *Nucleic Acids Res.* *43*, 4909–4922.

Fujita, Y., Hayashi, T., Kiyomitsu, T., Toyoda, Y., Kokubu, A., Obuse, C., and Yanagida, M. (2007). Priming of Centromere for CENP-A Recruitment by Human hMis18 α , hMis18 β , and M18BP1. *Dev. Cell* *12*, 17–30.

Fukuda, H., Sano, N., Muto, S., and Horikoshi, M. (2006). Simple histone acetylation plays a complex role in the regulation of gene expression. *Briefings Funct. Genomics Proteomics* *5*, 190–208.

Gerhard, D.S., Wagner, L., Feingold, E.A., Shenmen, C.M., Grouse, L.H., Schuler, G., Klein, S.L., Old, S., Rasooly, R., Good, P., et al. (2004). The status, quality, and expansion of the NIH full-length cDNA project: The Mammalian Gene Collection (MGC). *Genome Res.* *14*, 2121–2127.

Glaeser, R.M., and Han, B.-G. (2017). Opinion: hazards faced by macromolecules when confined to thin aqueous films. *Biophys. Reports* *3*, 1–7.

Goldman, J.A., Garlick, J.D., and Kingston, R.E. (2010). Chromatin remodeling by imitation switch (ISWI) class ATP-dependent remodelers is stimulated by histone variant H2A.Z. *J. Biol. Chem.* *285*, 4645–4651.

Gong, F., and Miller, K.M. (2013). Mammalian DNA repair: HATs and HDACs make their mark through histone acetylation. *Mutat. Res. - Fundam. Mol. Mech. Mutagen.* *750*, 23–30.

Guo, L.Y., Allu, P.K., Zandarashvili, L., McKinley, K.L., Sekulic, N., Dawicki-McKenna,

J.M., Fachinetti, D., Jamiolkowski, R.M., Cleveland, D.W., Cheeseman, I.M., et al. (2017). Centromere maintenance is achieved by fastening CENP-A to DNA and directing an arginine anchor-dependent nucleosome structural transition. *Nat. Commun. Press* 8, 1–15.

Guse, A., Carroll, C.W., Moree, B., Fuller, C.J., and Straight, A.F. (2011). In vitro centromere and kinetochore assembly on defined chromatin templates. *Nature* 477, 354–358.

Guse, A., Fuller, C.J., and Straight, A.F. (2012). A cell-free system for functional centromere and kinetochore assembly. *Nat. Protoc.* 7, 1847–1869.

Hansen, J.C. (2012). Human mitotic chromosome structure: What happened to the 30-nm fibre? *EMBO J.* 31, 1621–1623.

Hasson, D., Panchenko, T., Salimian, K.J., Salman, M.U., Sekulic, N., Alonso, A., Warburton, P.E., and Black, B.E. (2013). The octamer is the major form of CENP-A nucleosomes at human centromeres. *Nat. Struct. Mol. Biol.* 20, 687–695.

Hayashi, T., Fujita, Y., Iwasaki, O., Adachi, Y., Takahashi, K., and Yanagida, M. (2004). Mis16 and Mis18 are required for CENP-A loading and histone deacetylation at centromeres. *Cell* 118, 715–729.

Hebbes, T.R., Thorne, A.W., and Crane-Robinson, C. (1988). A direct link between core histone acetylation and transcriptionally active chromatin. *EMBO J.* 7, 1395–1402.

Hellwig, D., Emmerth, S., Ulbricht, T., Döring, V., Hoischen, C., Martin, R., Samora, C.P., McAinsh, A.D., Carroll, C.W., Straight, A.F., et al. (2011). Dynamics of CENP-N

kinetochore binding during the cell cycle. *J. Cell Sci.* *124*, 3871–3883.

Henikoff, S. (2008). Nucleosome destabilization in the epigenetic regulation of gene expression. *Nat. Rev. Genet.* *9*, 15–26.

Henikoff, S., and Smith, M.M. (2015). Histone variants and epigenetics. *Cold Spring Harb. Perspect. Biol.* *7*, 1–25.

Hoffmann, S., Dumont, M., Barra, V., Ly, P., Nechemia-Arbely, Y., McMahon, M.A., Hervé, S., Cleveland, D.W., and Fachinetti, D. (2016). CENP-A Is Dispensable for Mitotic Centromere Function after Initial Centromere/Kinetochore Assembly. *Cell Rep.* *17*, 2394–2404.

Holland, A.J., Fachinetti, D., Han, J.S., and Cleveland, D.W. (2012). Inducible, reversible system for the rapid and complete degradation of proteins in mammalian cells. *Proc. Natl. Acad. Sci. U. S. A.* *109*, E3350-7.

Hooser, A.A. Van, Ouspenski, I.I., Gregson, H.C., Starr, D.A., Yen, T.J., Goldberg, M.L., Yokomori, K., Earnshaw, W.C., Sullivan, K.F., and Brinkley, B.R. (2001). Specification of kinetochore-forming chromatin by the histone H3 variant CENP-A. *J. Cell Sci.* *114*, 3529–3542.

Hori, T., Amano, M., Suzuki, A., Backer, C.B., Welburn, J.P., Dong, Y., McEwen, B.F., Shang, W.H., Suzuki, E., Okawa, K., et al. (2008). CCAN Makes Multiple Contacts with Centromeric DNA to Provide Distinct Pathways to the Outer Kinetochore. *Cell* *135*, 1039–1052.

Hori, T., Shang, W.-H., Hara, M., Ariyoshi, M., Arimura, Y., Fujita, R., Kurumizaka, H.,

and Fukagawa, T. (2017). Association of M18BP1/KNL2 with CENP-A Nucleosome Is Essential for Centromere Formation in Non-mammalian Vertebrates. *Dev. Cell* *42*, 181–189.e3.

Iwasaki, W., Miya, Y., Horikoshi, N., Osakabe, A., Taguchi, H., Tachiwana, H., Shibata, T., Kagawa, W., and Kurumizaka, H. (2013). Contribution of histone N-terminal tails to the structure and stability of nucleosomes. *FEBS Open Bio* *3*, 363–369.

Jansen, L.E.T., Black, B.E., Foltz, D.R., and Cleveland, D.W. (2007). Propagation of centromeric chromatin requires exit from mitosis. *J. Cell Biol.* *176*, 795–805.

Jokelainen, P.T. (1967). The ultrastructure and spatial organization of the metaphase kinetochore in mitotic rat cells. *J. Ultrastruct. Res.* *19*, 19–44.

Kalashnikova, A.A., Porter-Goff, M.E., Muthurajan, U.M., Luger, K., and Hansen, J.C. (2013). The role of the nucleosome acidic patch in modulating higher order chromatin structure. *J. R. Soc. Interface* *10*, 20121022–20121022.

Kato, H., Jiang, J., Zhou, B.-R., Rozendaal, M., Feng, H., Ghirlando, R., Xiao, T.S., Straight, A.F., and Bai, Y. (2013). A conserved mechanism for centromeric nucleosome recognition by centromere protein CENP-C. *Science* *340*, 1110–1113.

Kimura, H., and Cook, P.R. (2001). Kinetics of Core Histones in Living Human Cells: Little Exchange of H3 and H4 and Some Rapid Exchange of H2b. *J. Cell Biol.* *153*, 1341–1354.

Kingston, I.J., Yung, J.S.Y., and Singleton, M.R. (2011). Biophysical characterization of the centromere-specific nucleosome from budding yeast. *J. Biol. Chem.* *286*, 4021–

4026.

Klare, K., Weir, J.R., Basilico, F., Zimniak, T., Massimiliano, L., Ludwigs, N., Herzog, F., and Musacchio, A. (2015). CENP-C is a blueprint for constitutive centromere-associated network assembly within human kinetochores. *J. Cell Biol.* *210*, 11–22.

Kornberg, R.D. (1974). Chromatin Structure: A Repeating Unit of Histones and DNA. *Science* (80-). *184*, 868–871.

Lander, G.C., Stagg, S.M., Voss, N.R., Cheng, A., Fellmann, D., Pulokas, J., Yoshioka, C., Irving, C., Mulder, A., Lau, P.W., et al. (2009). Appion: An integrated, database-driven pipeline to facilitate EM image processing. *J. Struct. Biol.* *166*, 95–102.

Lowary, P., and Widom, J. (1998). New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. *J. Mol. Biol.* *276*, 19–42.

Luger, K. (2001). Nucleosomes: Structure and Function. *Encycl. Life Sci.* 1–8.

Luger, K., Mäder, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* *389*, 251–260.

Luger, K., Dechassa, M.L., and Tremethick, D.J. (2012). New insights into nucleosome and chromatin structure: An ordered state or a disordered affair? *Nat. Rev. Mol. Cell Biol.* *13*, 436–447.

Luykx, P. (1965). The structure of the kinetochore in meiosis and mitosis in *Urechis* eggs. *Exp. Cell Res.* *39*, 643–657.

Maeshima, K., Imai, R., Tamura, S., and Nozaki, T. (2014). Chromatin as dynamic 10-nm fibers. *Chromosoma* 123, 225–237.

Maeshima, K., Rogge, R., Tamura, S., Joti, Y., Hikima, T., Szerlong, H., Krause, C., Herman, J., Seidel, E., DeLuca, J., et al. (2016). Nucleosomal arrays self-assemble into supramolecular globular structures lacking 30-nm fibers. *EMBO J.* 35, 1115–1132.

Malik, H.S., and Henikoff, S. (2009). Major Evolutionary Transitions in Centromere Complexity. *Cell* 138, 1067–1082.

Mattioli, F., Bhattacharyya, S., Dyer, P.N., White, A.E., Sandman, K., Burkhart, B.W., Byrne, K.R., Lee, T., Ahn, N.G., Santangelo, T.J., et al. (2017). Structure of histone-based chromatin in Archaea. *Science* (80-.). 357, 609–612.

McIntosh, J.R., Grishchuk, E.L., Morphew, M.K., Efremov, A.K., Zhudenzov, K., Volkov, V.A., Cheeseman, I.M., Desai, A., Mastronarde, D.N., and Ataulakhanov, F.I. (2008). Fibrils Connect Microtubule Tips with Kinetochores: A Mechanism to Couple Tubulin Dynamics to Chromosome Motion. *Cell* 135, 322–333.

McKinley, K.L., and Cheeseman, I.M. (2016). The molecular basis for centromere identity and function. *Nat Rev Mol Cell Biol* 17, 16–29.

McKinley, K.L., Sekulic, N., Guo, L.Y., Tsinman, T., Black, B.E., and Cheeseman, I.M. (2015). The CENP-L-N Complex Forms a Critical Node in an Integrated Meshwork of Interactions at the Centromere-Kinetochores Interface. *Mol. Cell* 60, 886–898.

McNulty, S.M., Sullivan, L.L., and Sullivan, B.A. (2017). Human Centromeres Produce Chromosome-Specific and Array-Specific Alpha Satellite Transcripts that Are

Complexed with CENP-A and CENP-C. *Dev. Cell* 42, 226–240.e6.

Moree, B., Meyer, C.B., Fuller, C.J., and Straight, A.F. (2011). CENP-C recruits M18BP1 to centromeres to promote CENP-A chromatin assembly. *J. Cell Biol.* 194, 855–871.

Muthurajan, U., Mattioli, F., Bergeron, S., Zhou, K., Gu, Y., Chakravarthy, S., Dyer, P., Irving, T., and Luger, K. (2016). In Vitro Chromatin Assembly: Strategies and Quality Control. In *Methods in Enzymology*, pp. 3–41.

Muthurajan, U.M., McBryant, S.J., Lu, X., Hansen, J.C., and Luger, K. (2011). The linker region of macroH2A promotes self-association of nucleosomal arrays. *J. Biol. Chem.* 286, 23852–23864.

Nagpal, H., Hori, T., Furukawa, A., Sugase, K., Osakabe, A., Kurumizaka, H., and Fukagawa, T. (2015). Dynamic changes in CCAN organization through CENP-C during cell-cycle progression. *Mol. Biol. Cell* 26, 3768–3776.

Nardi, I.K., Zasadzińska, E., Stellfox, M.E., Knippler, C.M., and Foltz, D.R. (2016). Licensing of Centromeric Chromatin Assembly through the Mis18 α -Mis18 β Heterotetramer. *Mol. Cell* 61, 774–787.

Nechemia-Arbely, Y., Fachinetti, D., Miga, K.H., Sekulic, N., Soni, G. V., Kim, D.H., Wong, A.K., Lee, A.Y., Nguyen, K., Dekker, C., et al. (2017). Human centromeric CENP-A chromatin is a homotypic, octameric nucleosome at all cell cycle points. *J. Cell Biol.* 216, 607–621.

Nishimura, K., Fukagawa, T., Takisawa, H., Kakimoto, T., and Kanemaki, M. (2009). An

auxin-based degron system for the rapid depletion of proteins in nonplant cells. *Nat. Methods* 6, 917–923.

Nishino, T., Takeuchi, K., Gascoigne, K.E., Suzuki, A., Hori, T., Oyama, T., Morikawa, K., Cheeseman, I.M., and Fukagawa, T. (2012). CENP-T-W-S-X forms a unique centromeric chromatin structure with a histone-like fold. *Cell* 148, 487–501.

Nusinow, D.A., Hernández-Muñoz, I., Fazzio, T.G., Shah, G.M., Kraus, W.L., and Panning, B. (2007). Poly(ADP-ribose) polymerase 1 is inhibited by a histone H2A variant, macroH2A, and contributes to silencing of the inactive X chromosome. *J. Biol. Chem.* 282, 12851–12859.

Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R., Wakamatsu, A., Hayashi, K., Sato, H., Nagai, K., et al. (2004). Complete sequencing and characterization of 21,243 full-length human cDNAs. *Nat. Genet.* 36, 40–45.

Palmer, D.K., O'Day, K., Trong, H.L., Charbonneau, H., and Margolis, R.L. (1991). Purification of the centromere-specific protein CENP-A and demonstration that it is a distinctive histone. *Proc. Natl. Acad. Sci. U. S. A.* 88, 3734–3738.

Pentakota, S., Zhou, K., Smith, C., and Maffini, S. Decoding the centromeric nucleosome through CENP-N.

Pentakota, S., Zhou, K., Smith, C., Maffini, S., Petrovic, A., Morgan, G.P., Weir, J.R., Vetter, I.R., Musacchio, A., and Luger, K. (2017). Decoding the centromeric nucleosome through CENP-N. *Elife* 6.

Pesenti, M.E., Weir, J.R., and Musacchio, A. (2016). Progress in the structural and

functional characterization of kinetochores. *Curr. Opin. Struct. Biol.* *37*, 152–163.

Poirier, M.G., Bussiek, M., Langowski, J., and Widom, J. (2008). Spontaneous Access to DNA Target Sites in Folded Chromatin Fibers. *J. Mol. Biol.* *379*, 772–786.

Polach, K.J., and Widom, J. (1995). Mechanism of protein access to specific DNA sequences in chromatin: A dynamic equilibrium model for gene regulation. *J. Mol. Biol.* *254*, 130–149.

Przewloka, M.R., and Glover, D.M. (2009). The kinetochore and the centromere: a working long distance relationship. *Annu. Rev. Genet.* *43*, 439–465.

Radman-Livaja, M., Verzijlbergen, K.F., Weiner, A., van Welsem, T., Friedman, N., Rando, O.J., and van Leeuwen, F. (2011). Patterns and mechanisms of Ancestral Histone protein inheritance in Budding yeast. *PLoS Biol.* *9*.

Rev, O. (2015). HHS Public Access. *14*, 871–882.

Rogakou, E., Martin, O.A., Redon, C., Pilch, D., Rogakou, E., Sedelnikova, O., Newrock, K., and Bonner, W. (2016). Histone H2A variants H2AX and H2AZ . *Curr Histone H2A variants H2AX and H2AZ. Curr. Opin. Genet. Dev.* 162–169.

Rogakou, E.P., Pilch, D.R., Orr, A.H., Ivanova, V.S., and Bonner, W.M. (1998). Double-stranded Breaks Induce Histone H2AX phosphorylation on Serine 139. *J. Biol. Chem.* *273*, 5858–5868.

Roulland, Y., Ouararhni, K., Naidenov, M., Ramos, L., Shuaib, M., Syed, S.H., Lone, I.N., Boopathi, R., Fontaine, E., Papai, G., et al. (2016). The Flexible Ends of CENP-A Nucleosome Are Required for Mitotic Fidelity. *Mol. Cell* *63*, 674–685.

Sagar, Chittori; Jingjun Hong Hayden Saunders, Hanqiao Feng, Rodolfo Ghirlando, Alexander E. Kelly², Yawen Bai², S.S. (2017). Structural mechanisms of centromeric nucleosome recognition by the kinetochore protein CENP-N. *Science* (80-.).

Sandmann, M., Talbert, P., Demidov, D., Kuhlmann, M., Rutten, T., Conrad, U., and Lermontova, I. (2017). Targeting of *A. thaliana* KNL2 to centromeres depends on the conserved CENPC-k motif in its C-terminus. *Plant Cell* tpc.00720.2016.

Scheres, S.H.W. (2016). Processing of Structurally Heterogeneous Cryo-EM Data in RELION. *Methods Enzymol.*

Screpanti, E., De Antoni, A., Alushin, G.M., Petrovic, A., Melis, T., Nogales, E., and Musacchio, A. (2011). Direct binding of Cenp-C to the Mis12 complex joins the inner and outer kinetochore. *Curr. Biol.* *21*, 391–398.

Sekulic, N., Bassett, E. a, Rogers, D.J., and Black, B.E. (2010). The structure of (CENP-A–H4)₂ reveals physical features that mark centromeres. *Nature* *467*, 347–351.

Smook, E.M., Stein, P., Schultz, R.M., Lampson, M.A., and Black, B.E. (2016). Long-Term Retention of CENP-A Nucleosomes in Mammalian Oocytes Underpins Transgenerational Inheritance of Centromere Identity. *Curr. Biol.* *26*, 1110–1116.

Stellfox, M.E., and Foltz, D.R. (2014). *NIH Public Access.* *70*, 387–406.

Sugimoto, K., Kuriyama, K., Shibata, A., and Himeno, M. (1997). Characterization of internal DNA-binding and C-terminal dimerization domains of human centromere/kinetochore autoantigen CENP-C in vitro: role of DNA-binding and self-associating activities in kinetochore organization. *Chromosom. Res.* *5*, 132–141.

Suto, R.K., Clarkson, M.J., Tremethick, D.J., and Luger, K. (2000). Crystal structure of a nucleosome core particle containing the variant histone H2A.Z. *Nat. Struct. Biol.* 7, 1121–1124.

Tachiwana, H., Kagawa, W., Shiga, T., Osakabe, A., Miya, Y., Saito, K., Hayashi-Takanaka, Y., Oda, T., Sato, M., Park, S.-Y., et al. (2011). Crystal structure of the human centromeric nucleosome containing CENP-A. *Nature* 476, 232–235.

Tian, T., Li, X., Liu, Y., Wang, C., Liu, X., Bi, G., Zhang, X., Yao, X., Zhou, Z.H., and Zang, J. (2018). Molecular basis for CENP-N recognition of CENP-A nucleosome on the human kinetochore. *Cell Res.* 1–5.

Tims, H.S., Gurunathan, K., Levitus, M., and Widom, J. (2011). Dynamics of nucleosome invasion by DNA binding proteins. *J. Mol. Biol.* 411, 430–448.

Tolstorukov, M.Y., Goldman, J.A., Gilbert, C., Ogryzko, V., Kingston, R.E., and Park, P.J. (2012). Histone Variant H2A.Bbd Is Associated with Active Transcription and mRNA Processing in Human Cells. *Mol. Cell* 47, 596–607.

Toyama, B.H., Savas, J.N., Park, S.K., Harris, M.S., Ingolia, N.T., Iii, J.R.Y., and Hetzer, M.W. (2013). Identification of Long-Lived Proteins Reveals Exceptional Stability of Essential Cellular Structures. *Cell* 154, 971–982.

Trazzi, S., Perini, G., Bernardoni, R., Zoli, M., Reese, J.C., Musacchio, A., and Della Valle, G. (2009). The C-terminal domain of CENP-C displays multiple and critical functions for mammalian centromere formation. *PLoS One* 4, 14–17.

Tsunaka, Y., Kajimura, N., Tate, S.I., and Morikawa, K. (2005). Alteration of the

nucleosomal DNA path in the crystal structure of a human nucleosome core particle. *Nucleic Acids Res.* *33*, 3424–3434.

Vasudevan, D., Chua, E.Y.D., and Davey, C.A. (2010). Crystal Structures of Nucleosome Core Particles Containing the “601” Strong Positioning Sequence. *J. Mol. Biol.* *403*, 1–10.

Voltz, K., Trylska, J., Calimet, N., Smith, J.C., and Langowski, J. (2012). Unwrapping of nucleosomal DNA ends: a multiscale molecular dynamics study. *Biophys. J.* *102*, 849–858.

Warburton, P.E., Cooke, C.A., Bourassa, S., Vafa, O., Sullivan, B.A., Stetten, G., Gimelli, G., Warburton, D., Tyler-Smith, C., Sullivan, K.F., et al. (1997). Immunolocalization of CENP-A suggests a distinct nucleosome structure at the inner kinetochore plate of active centromeres. *Curr. Biol.* *7*, 901–904.

Waterborg, J.H. (1993). Histone synthesis and turnover in alfalfa. Fast loss of highly acetylated replacement histone variant H3.2. *J. Biol. Chem.* *268*, 4912–4917.

Weir, J.R., Faesen, A.C., Klare, K., Petrovic, A., Basilico, F., Fischböck, J., Pentakota, S., Keller, J., Pesenti, M.E., Pan, D., et al. (2016). Insights from biochemical reconstitution into the architecture of human kinetochores. *Nature* *537*, 249–253.

White, C.L., Suto, R.K., and Luger, K. (2001). Structure of the yeast nucleosome core particle reveals fundamental changes in internucleosome interactions. *EMBO J.* *20*, 5207–5218.

Woodcock, C.L.F., Safer, J.P., and Stanchfield, J.E. (1976). Structural repeating units in

chromatin. I. Evidence for their general occurrence. *Exp. Cell Res.* 97, 101–110.

Yoda, K., Ando, S., Morishita, S., Houmura, K., Hashimoto, K., and Takeyasu, K. (2000). Human centromere protein A (CENP-A) can replace histone H3 in nucleosome reconstitution in vitro. 2–7.

Zheng, S.Q., Palovcak, E., Armache, J.-P., Verba, K.A., Cheng, Y., and Agard, D.A. (2017). MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* 14, 331–332.

Zlatanova, J., and Thakar, A. (2008). H2A.Z: View from the Top. *Structure* 16, 166–179.

APPENDIX

SUPPLEMENTARY FIGURES

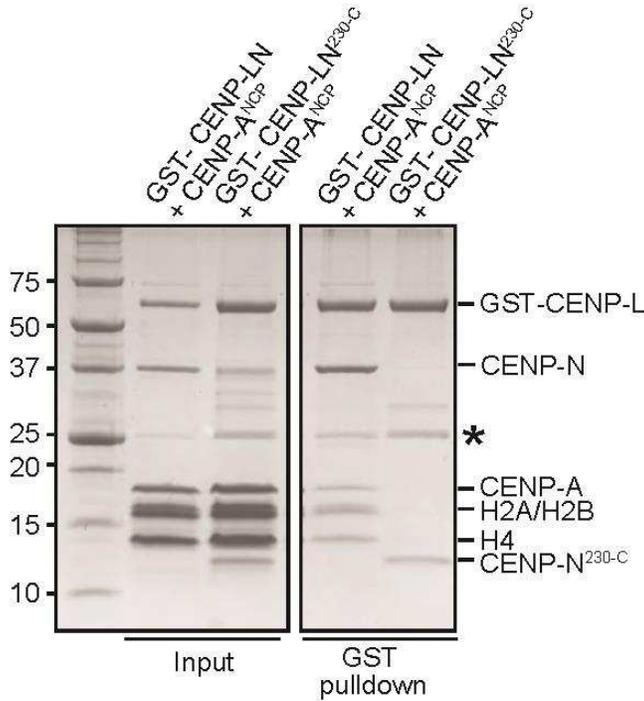


Figure 2.1 Supplement 1

Further interactions of CENP-NL

Solid phase binding assay demonstrating that CENP-LN^{230-C} does not interact with CENP-A nucleosome. The indicated GST fusion proteins were immobilized on glutathione-sepharose beads (at a final concentration of 1 μ M) and incubated with 3 μ M of the indicated nucleosome core particles. After incubation (see Methods), beads were pelleted, washed, and bound proteins visualized by SDS-PAGE and Coomassie staining.

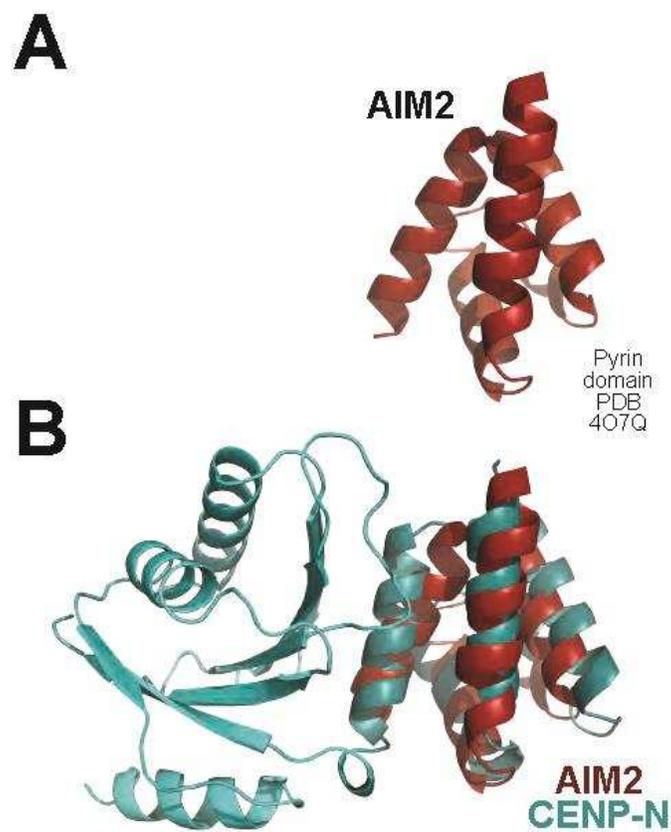


Figure 2.2 Supplement 1

Structural relatedness of CENP-N N-terminal domain with PYRINs

A. Cartoon model of the “death fold” Pyrin domain of AIM2 (Lu et al., 2014). B. Structural superposition of the AIM2 Pyrin domain (red) with the CENP-N Pyrin domain.

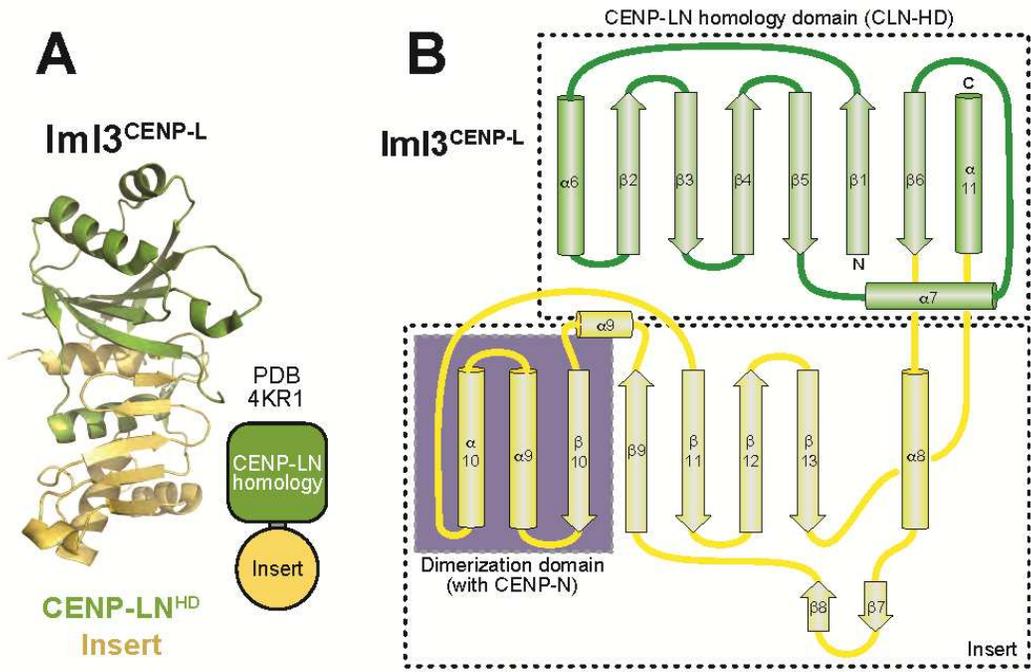


Figure 2.2 Supplement 2
 Topology of Iml3 and CENP-L

A. Overall domain organization of Iml3^{CENP-L} with cartoon model illustrating the CLN-HD in green and the CENP-L insert in yellow. The latter includes the Ch14^{CENP-N}-binding domain, which is conserved in humans, as shown in panel E (Hinshaw and Harrison, 2013). B. Topology diagram of CENP-L, based on the structure of the Iml3^{CENP-L} yeast homolog (Hinshaw and Harrison, 2013) and its comparison with the crystal structure of CENP-N after identification of structural homology by program DALI.

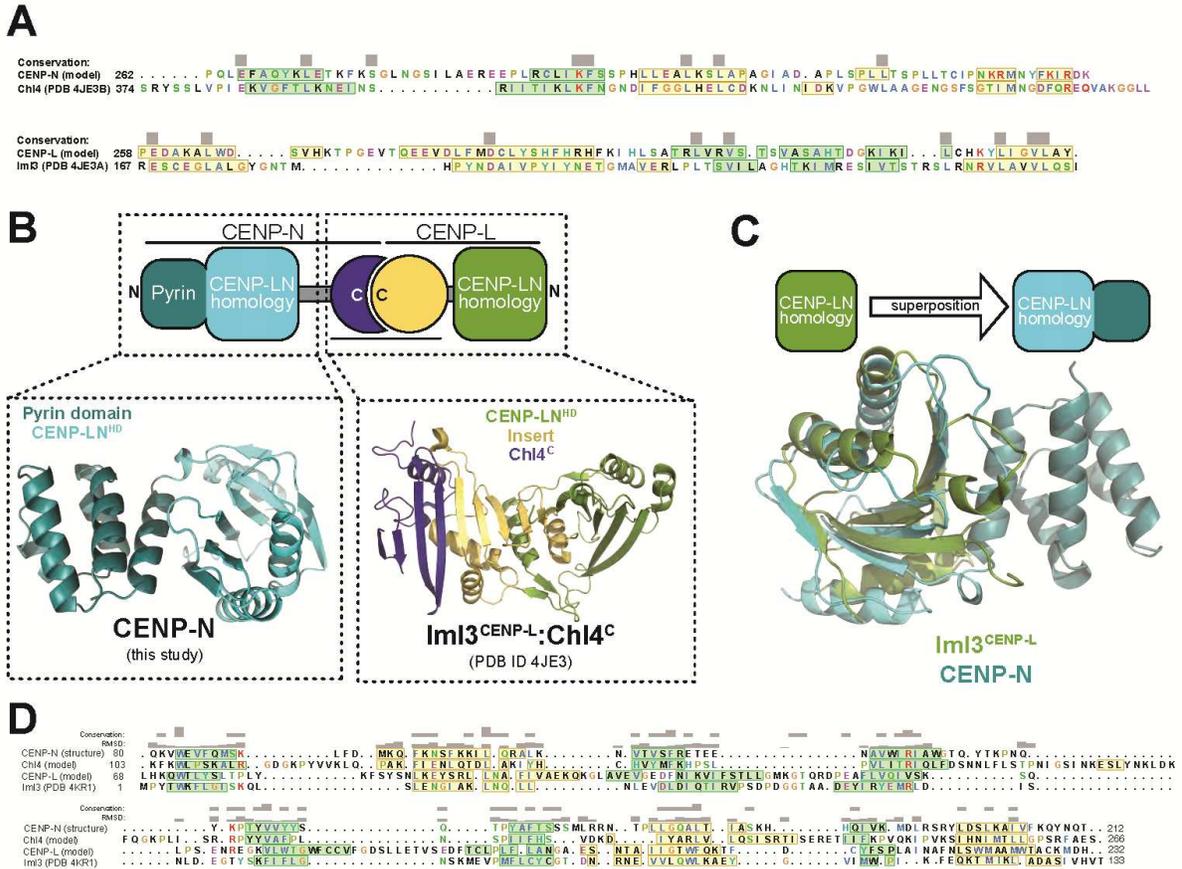


Figure 2.2 Supplement 3

Overall organization of the CENP-LN complex

A) Structure-based sequence alignment generated with program Chimera of the C-terminal region of CENP-N (model) and the C-terminal region of Chl4^{CENP-N} (structure) with β -strands shown in green and α -helices in yellow (secondary structure predictions in case of the model), and of the C-termini of CENP-L (model) and Im13 (PDB structure 4KR1) with the same colour scheme. B) Schematics of the overall organization of the CENP-LN complex, and complementarity of the CENP-N¹⁻²¹² and 4JE3 structures. C) Structural superposition of the CLN-HD of Im13^{CENP-L} (green) on the corresponding middle domain of CENP-N¹⁻²³⁵. D) Structure-based sequence alignment generated with program Chimera of CENP-N, CENP-L, and their yeast homologs Im13^{CENP-L} and Chl4^{CENP-N}, with β -strands shown in green and α -helices in yellow, including the two structures (CENP-N and Im13) and sequences and predicted secondary structures for the two models (Chl4 and CENP-L).

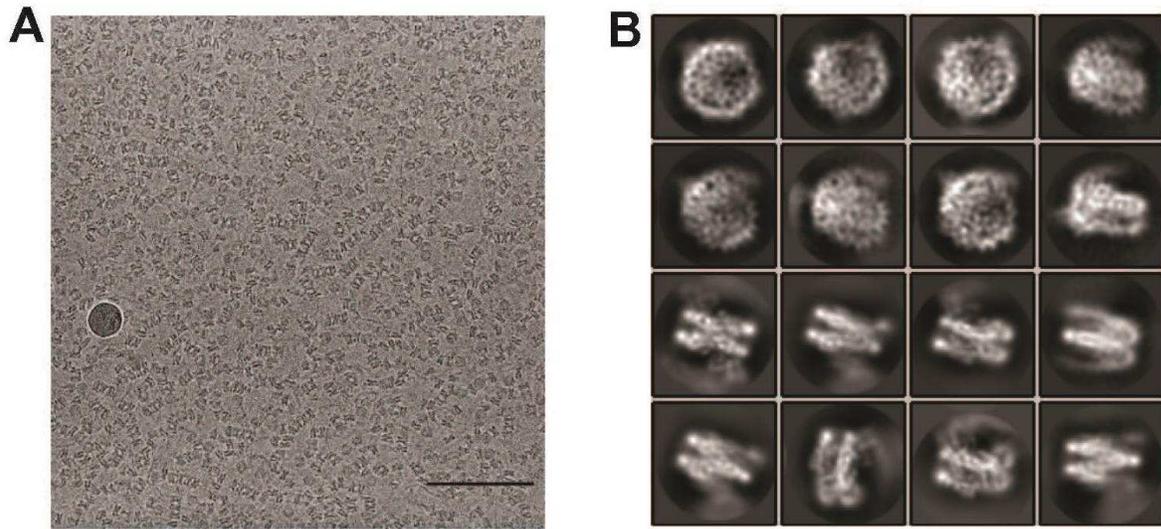


Figure 2.3 Supplement 1

Additional EM data and analysis

A Collected micrograph after CTF correction at $-2.5 \mu\text{m}$ defocus (nominal). Scale bar = 100 nm. B. Representative 2D class averages from RELION (Scheres, 2012) 2D classification.

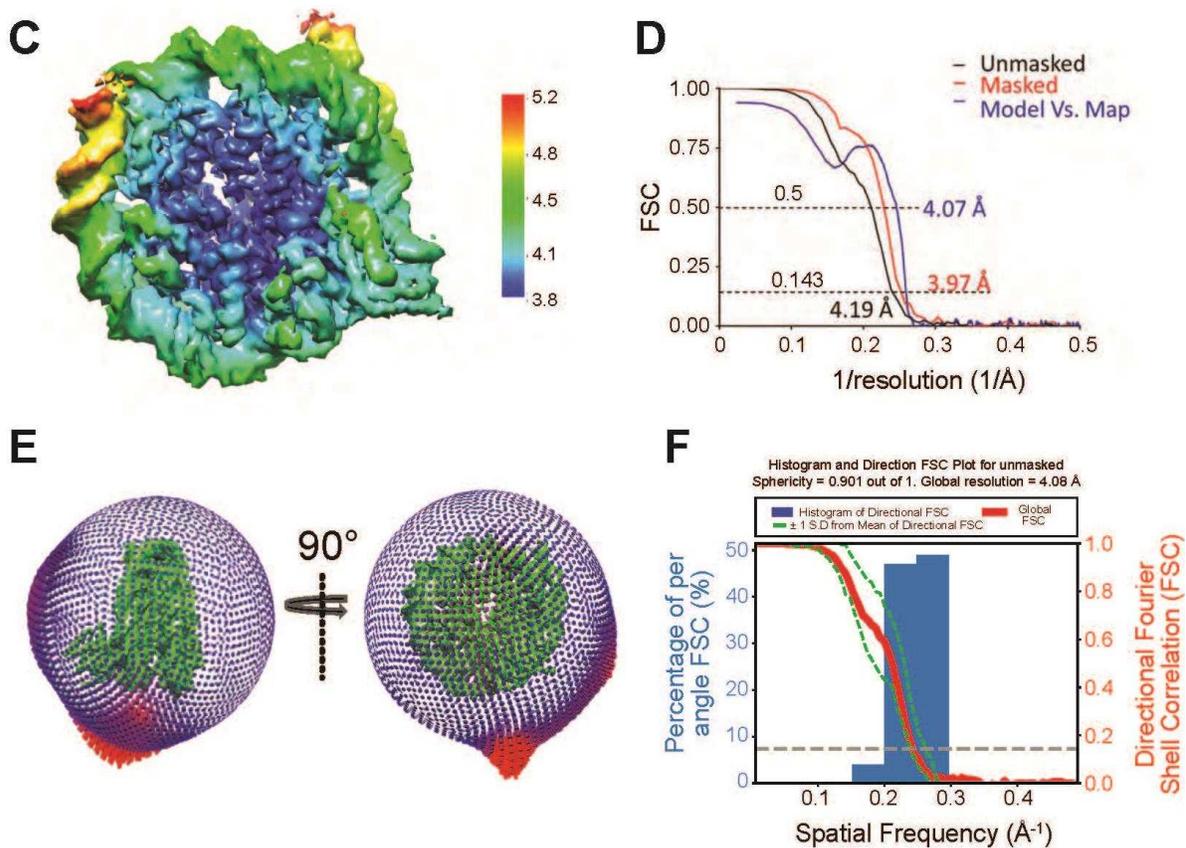


Figure 2.3 Supplement 1

Additional EM data and analysis

C. Fourier Shell Correction (FSC) curve for the maps. D. Estimated local resolution for the CENP-N¹⁻²⁸⁹:CENP-A^{NCP} by RELION. The unit for colour scale is Å. E. Euler angular distribution of the 3D reconstruction. F. Histogram and directional FSC plot for unmasked map. Sphericity = 0.901 out of 1, global resolution = 4.08 Å.

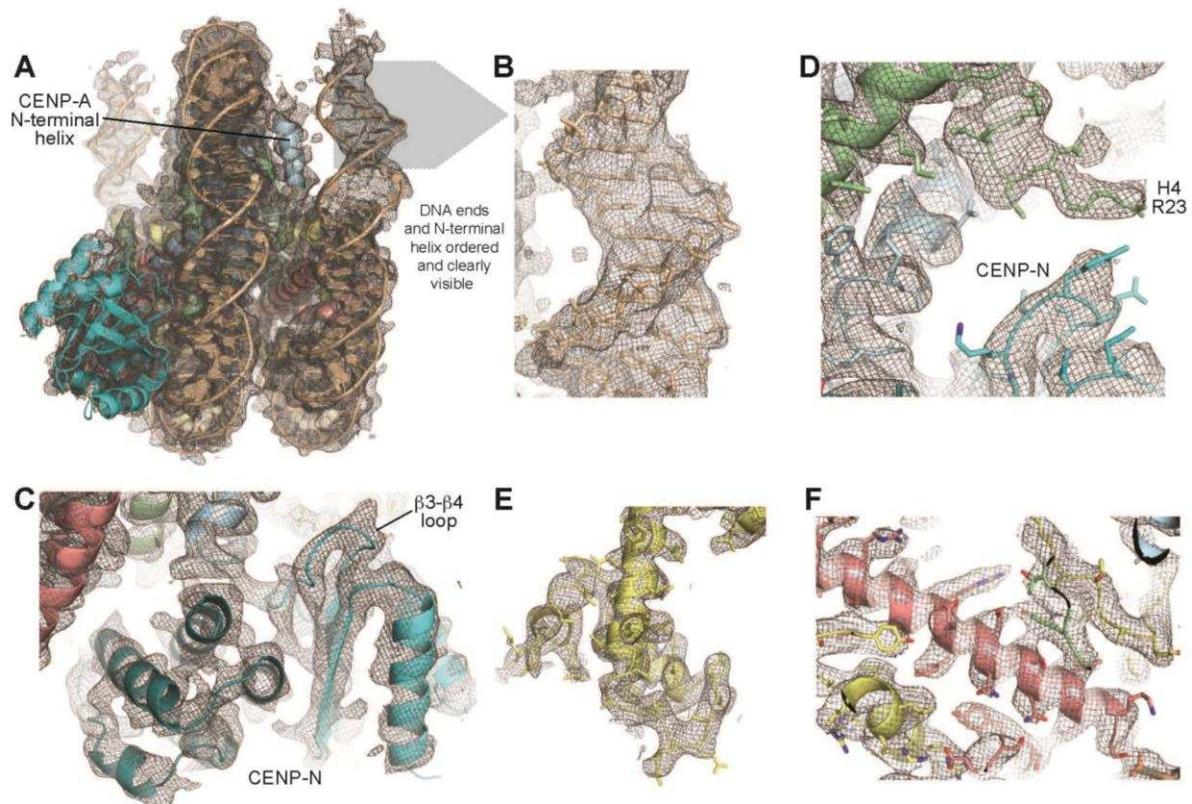


Figure 2.3 Supplement 2

EM maps

A. Representative areas of the map of the CENP-N¹⁻²⁸⁹:CENP-A^{NCP} complex contoured at 3σ shows very clear density for the DNA ends and for the N-terminal helix of CENP-A, starting from residue G46. The grey arrow points to a zoomed-in region close to the DNA end. B. The zoomed-in region. C. Overall fit of the CENP-N¹⁻²³⁵ crystallographic model in the EM density. D. The H4 N-terminal region is clearly ordered starting at residue R23, possibly due to packing against the $\beta 3$ - $\beta 3$ loop of CENP-N. E. Illustrative example of model fit in the map around H2A. F. Illustrative example of model fit in the map around H2B.

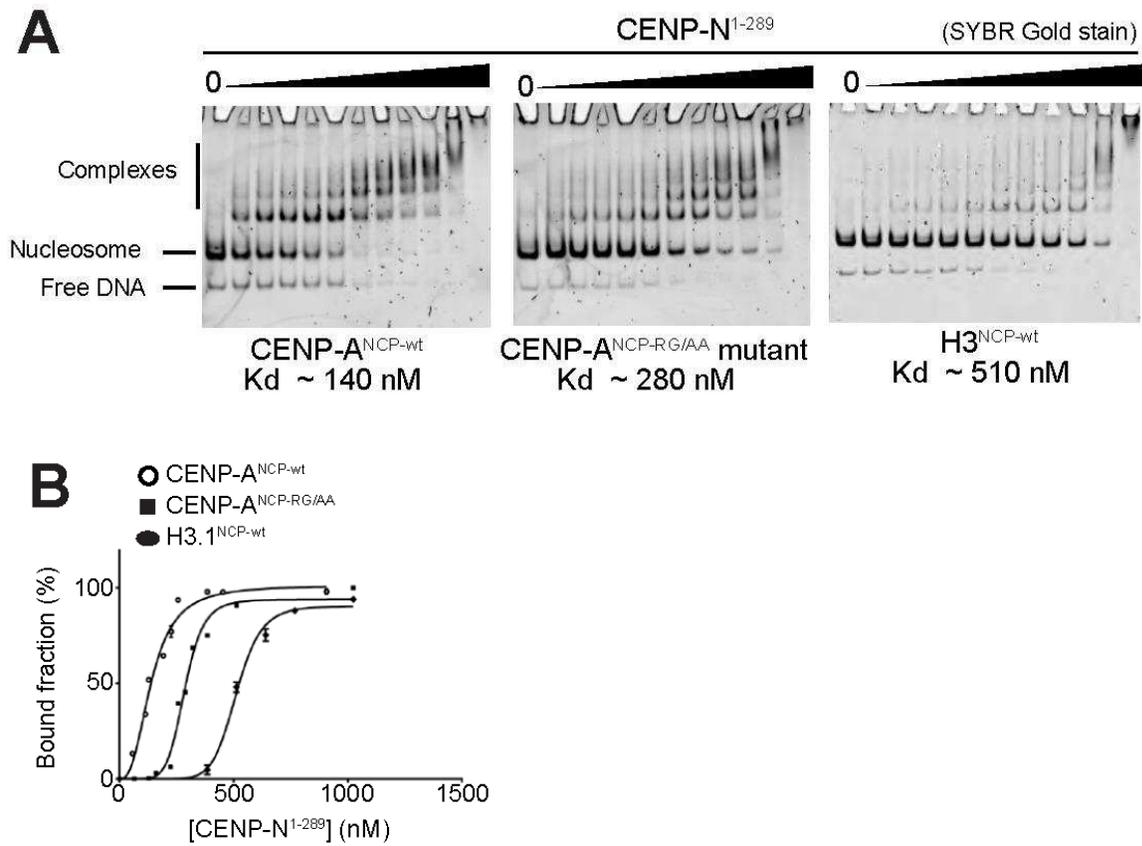


Figure 2.3 Supplement 3

EMSA assays

A. EMSAs (electrophoretic mobility shift assays) with the indicated NCPs and CENP-N¹⁻²⁸⁹.
 B. Quantification of data in A. Data were fitted in Graphpad using the specific binding model with Hill coefficient fitting.

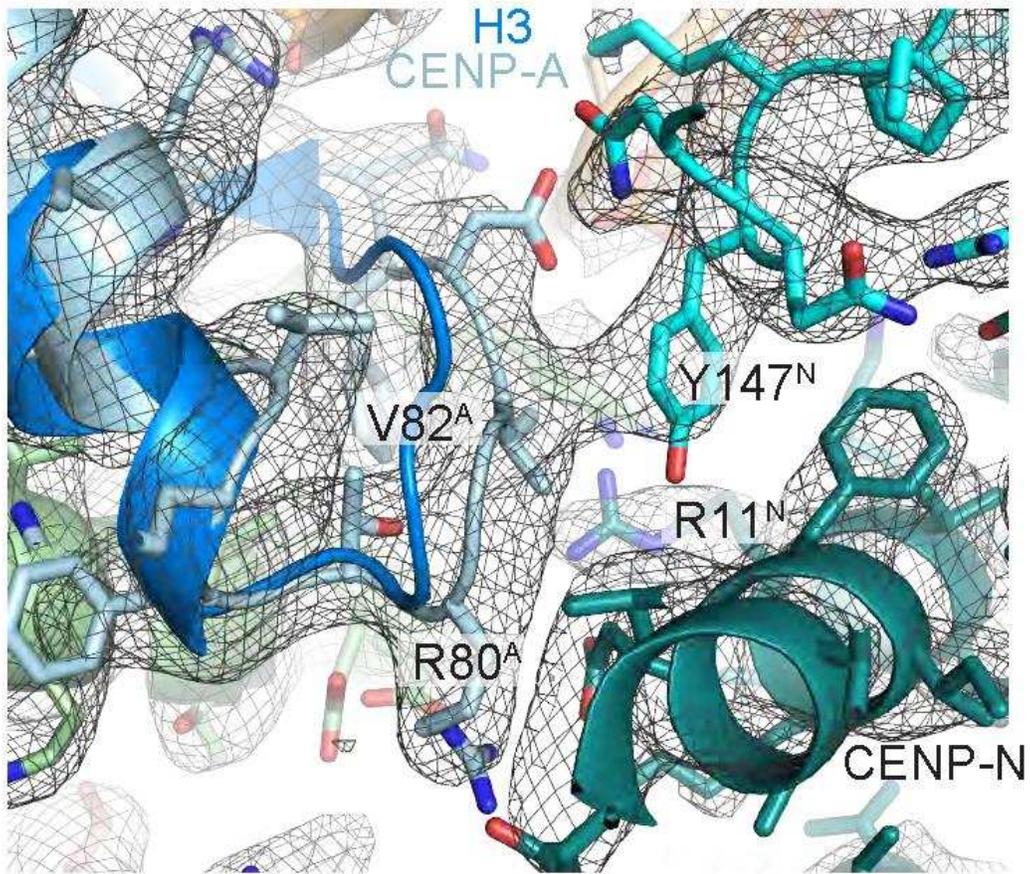


Figure 2.3 Supplement 5

Comparison of CENP-A and H3 and interface with CENP-N

Close-up view of the CENP-A:CENP-N interface with EM density map. The superposition with H3 (dark blue) illustrates that H3 would be unable to form productive contacts with CENP-N. V82 of CENP-A packs against the side chain of Y147 of CENP-N.

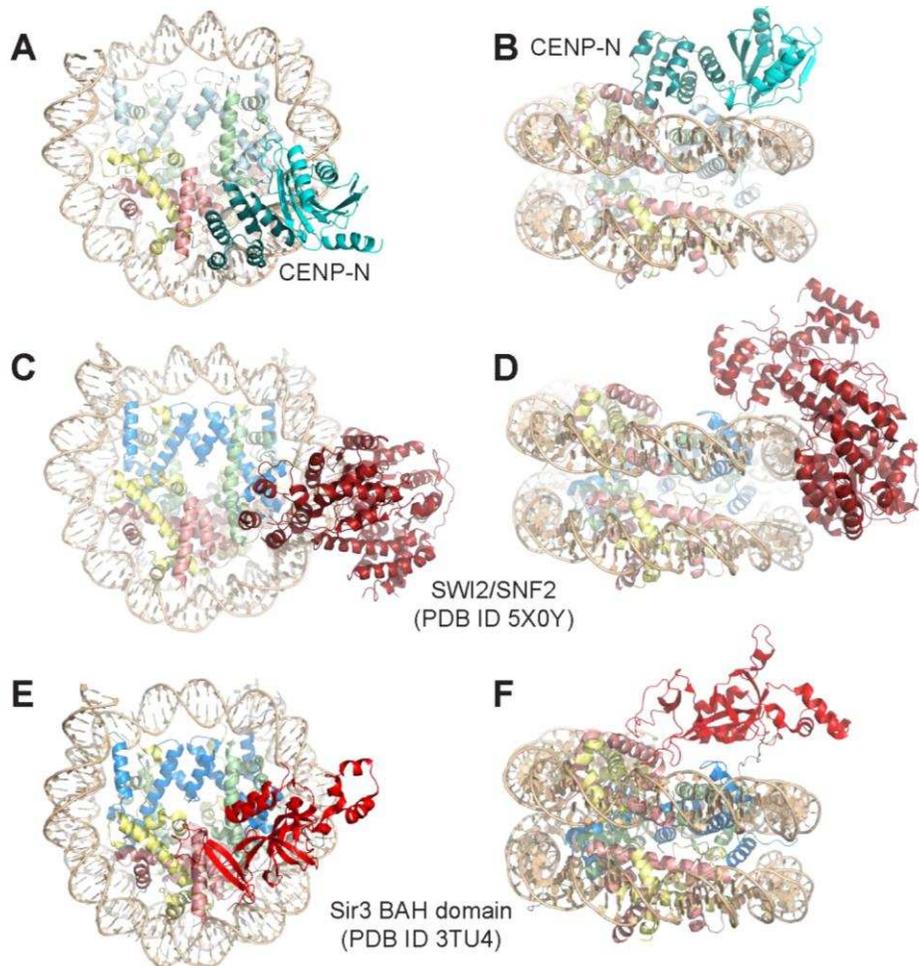


Figure 2.3 Supplement 6

Comparison of nucleosome binding modes

A-B) Two views in different orientation of the structure of CENP-N¹⁻²⁸⁹:CENP-A^{NCP} complex, with CENP-A in light blue and CENP-N in teal and cyan. C-D) Two views in different orientations of the SWI2/SNF2:H3^{NCP} complex (Liu et al., 2017) illustrating similarities in the interactions with the nucleosome. The contacts of SWI2/SNF2 with the proteinaceous part of the H3^{NCP} particle are more limited than for CENP-N. E-F) Two views in different orientations of the BAH:H3^{NCP} complex (Armache et al., 2011). The contacts of the BAH domain with DNA are very limited, while there are extensive interactions with the H2A-H2B acidic patch and with the H4 N-terminal tail and with H3.

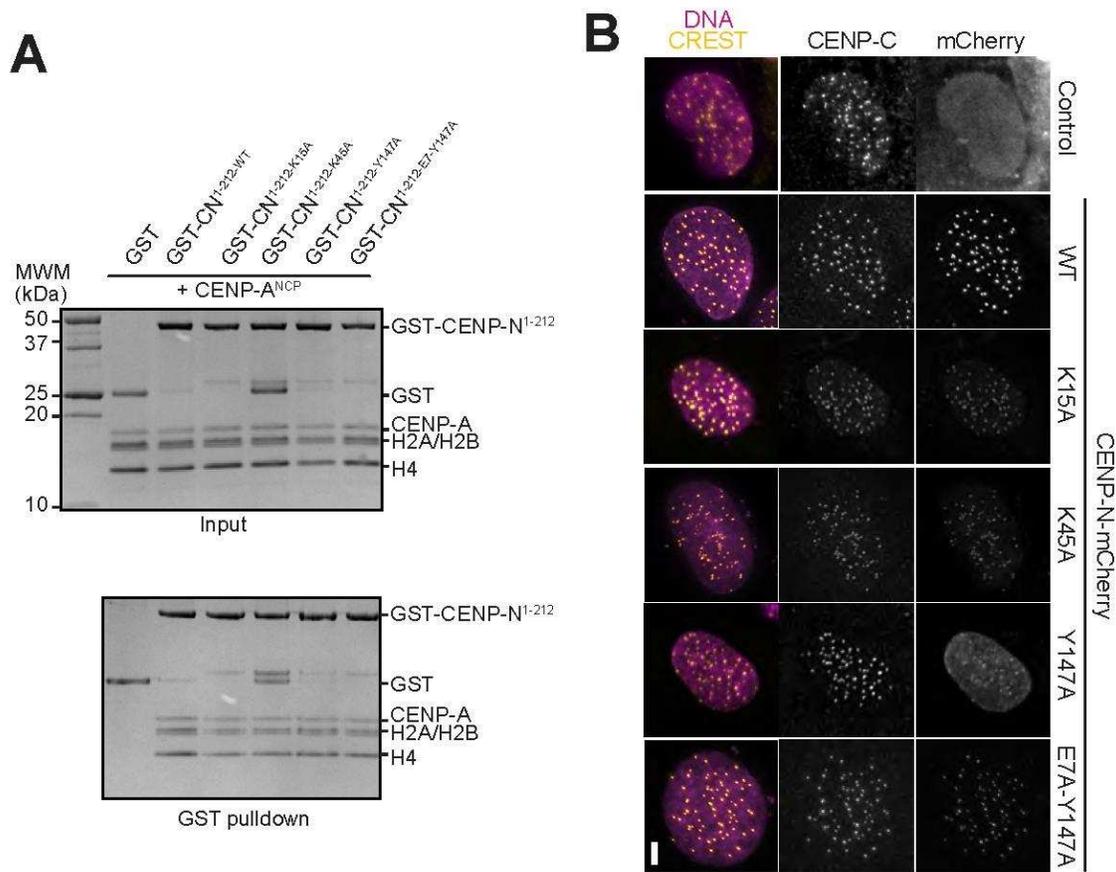


Figure 2.4 Supplement 1

Characterization of CENP-N mutants in solid phase and cell assays

A. *In vitro* binding assay probing the interaction of wild type and mutant GST-CENP-N¹⁻²¹² immobilized on solid phase with CENP-A^{NCP}. B. Fluorescence microscopy analysis comparing kinetochore localization of a wild type CENP-N-mCherry fluorescent reporter and of the indicated mutant variants in U2OS cells.

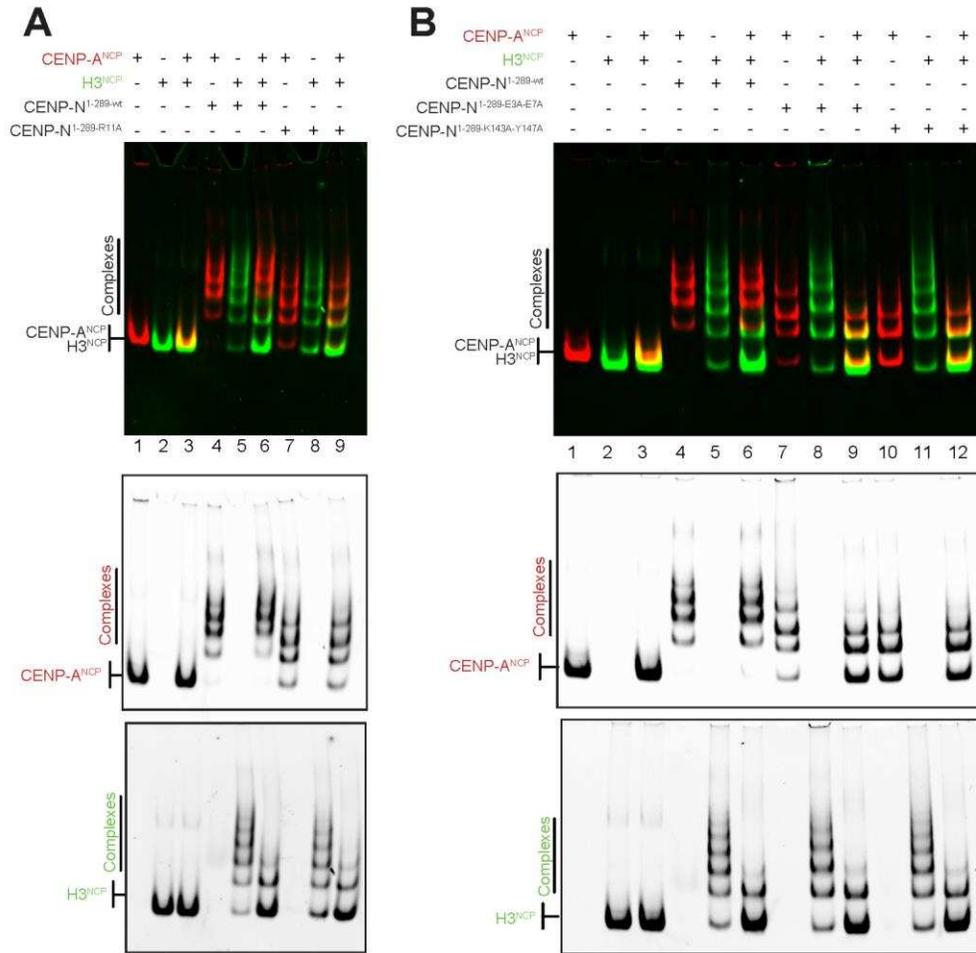


Figure 2.4 Supplement 2

Characterization of CENP-N mutants in competition gel shift assays

A-B) Gel shift assays with differentially labelled CENP-A^{NCPs} (red) and H3^{NCPs} (green) at 125 nM in presence of the indicated CENP-N wild type and mutant species (1 μ M). The two individual channels are shown below. The nature of all slow-migrating species formed upon binding of CENP-N and its mutants to nucleosomes is unknown, but we suspect they represent higher-order oligomerization. Note that mutants affecting the CENP-A L1 loop-binding interface of CENP-N affect CENP-A^{NCPs} binding more than H3^{NCPs} binding, likely because the latter is solely based on the intact DNA-binding interface of CENP-N (see text for additional details).

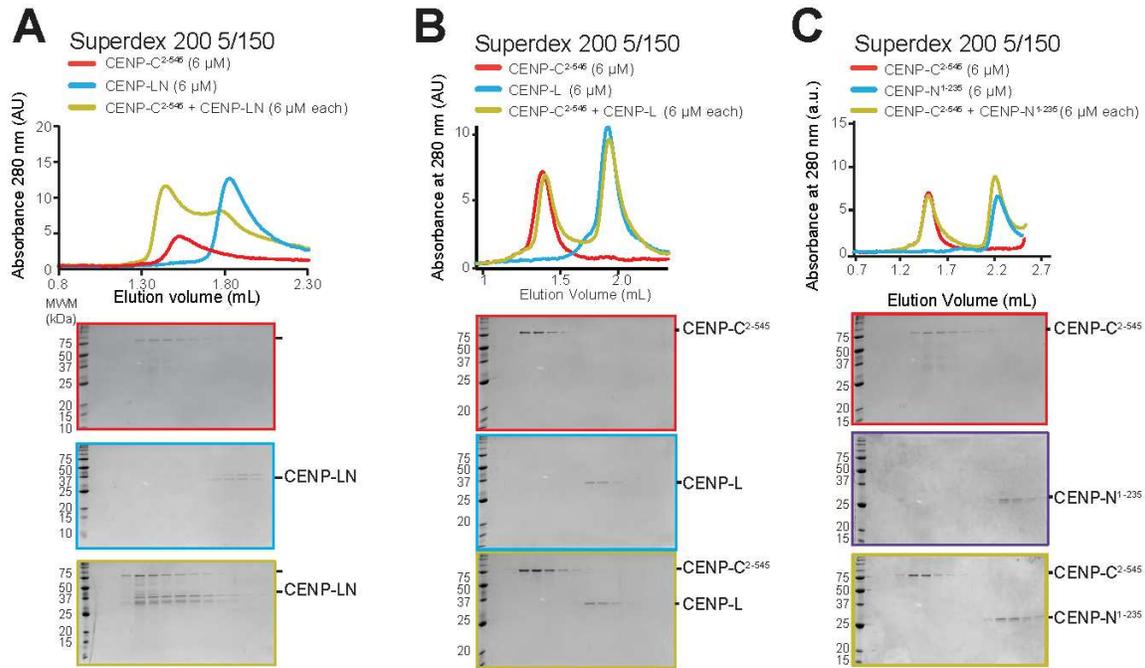


Figure 2.5 Supplement 1

Additional size-exclusion chromatography experiments

A. SEC experiment demonstrating an interaction of CENP-C²⁻⁵⁴⁵ with CENP-LN complex, as shown previously (Weir et al., 2016). B. SEC experiment demonstrating that CENP-L does not bind CENP-C²⁻⁵⁴⁵. C) SEC experiment demonstrating that CENP-N¹⁻²³⁵ does not bind CENP-C²⁻⁵⁴⁵.

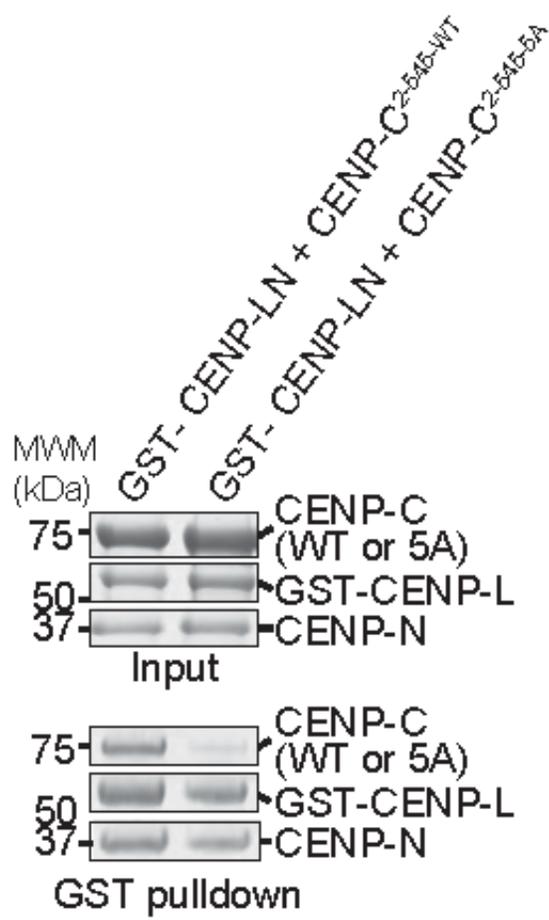
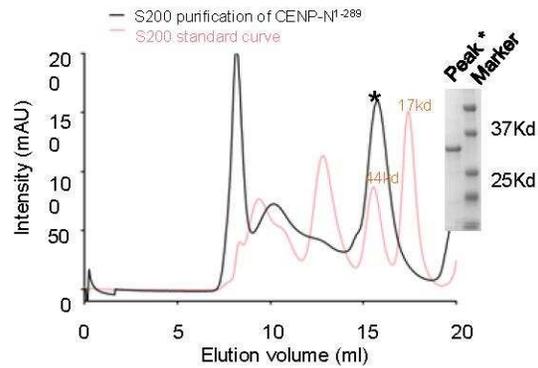
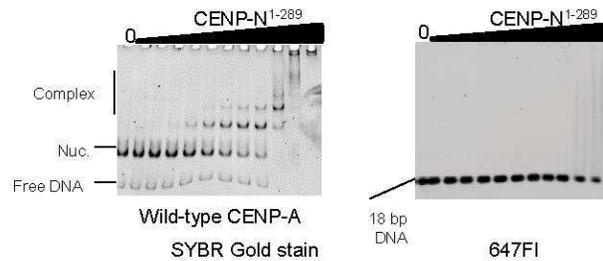
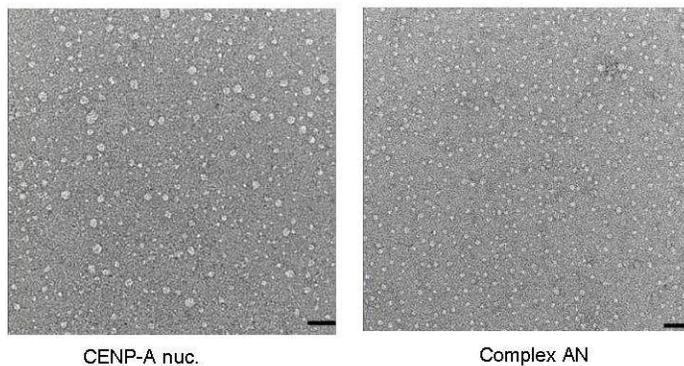


Figure 2.5 Supplement 2

Solid phase binding assays with CENP-C and CENP-C mutant

Solid phase binding assay with immobilized GST-CENP-LN complex and soluble CENP-C²⁻⁵⁴⁵ or CENP-C^{2-545-5A} mutant.

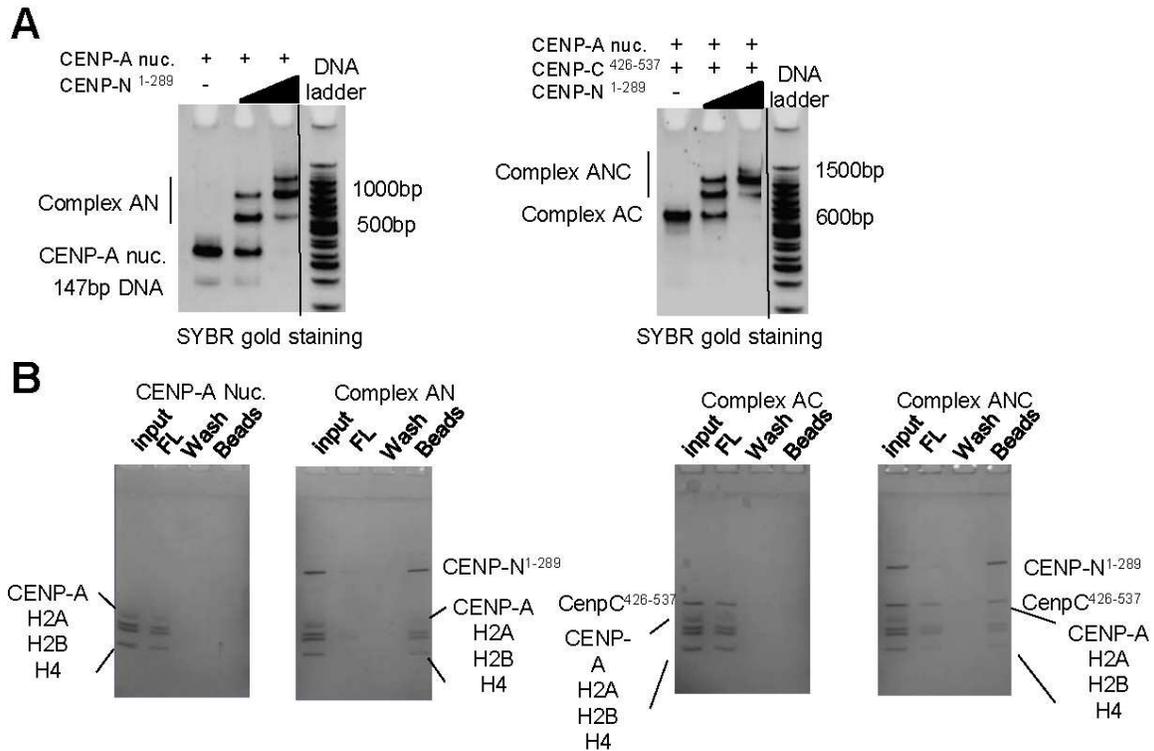
A**B****C**

Supplemental Figure 3.1 CENP-N specifically binds to CENP-A nucleosome in vitro.

A. CENP-N¹⁻²⁸⁹ was purified as a monomer. S200 (superdex 200, GL 10 300) elution profile for CENP-N¹⁻²⁸⁹. The red curve depicts standard proteins for gel filtration (Bio-Rad); the molecular weights are indicated. Inset: SDS-PAGE analysis of protein from the indicated peak.

B. Gel shift assays for the binding of CENP-N¹⁻²⁸⁹ to different substrates as indicated. Both CENP-A nucleosome (10 nM) and DNA (10 nM) were titrated with the same amounts of CENP-N¹⁻²⁸⁹ (from 0 to 1024 nM).

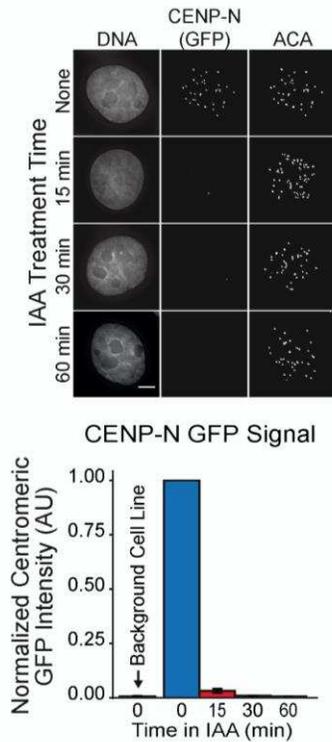
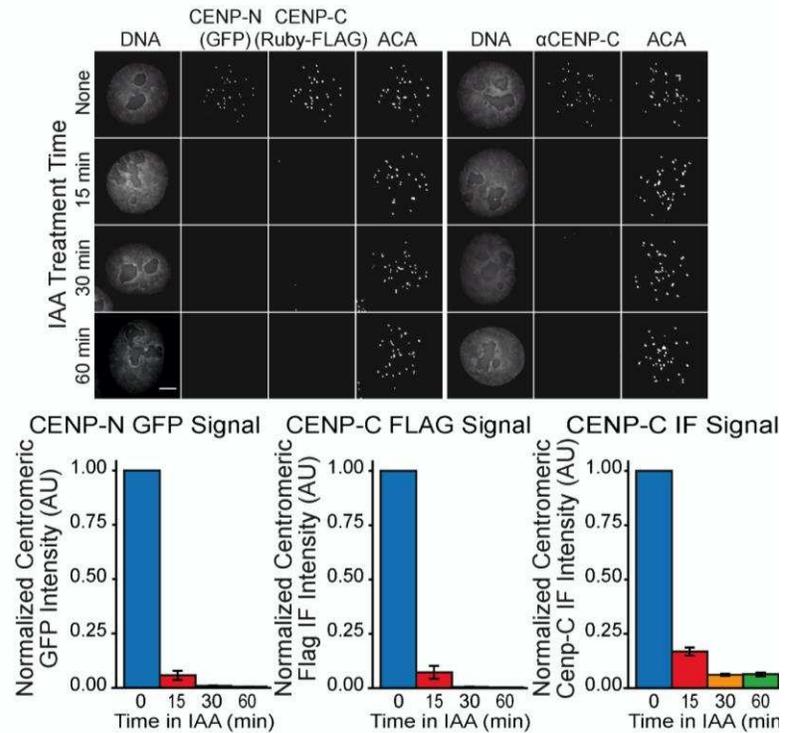
C. Stability of CENP-A nucleosome in the absence and presence of CENP-N¹⁻²⁸⁹, shown by negative staining EM. CENP-N¹⁻²⁸⁹ was mixed with CENP-A nucleosome at molar ratio of 3:1, and the same CENP-A nucleosomes without CENP-N are shown as controls. Concentration of both samples was ~ 30 nM. Red square indicates area magnified to the right. Scale bar=50 nm. The number of intact particles were counted and are listed in Table 1.



Supplemental Figure 3.2 CENP-N¹⁻²⁸⁹ and CENP-C⁴²⁶⁻⁵³⁷ simultaneously bind to the same nucleosome.

A. Native PAGE of 'AN' complex (CENP-A nucleosome with CENP-N¹⁻²⁸⁹) and 'ANC' complex (CENP-A nucleosome with CENP-N¹⁻²⁸⁹ and CENP-C⁴²⁶⁻⁵³⁷). Left panel: 0-, 2-, and 3-fold molar excess of CENP-N¹⁻²⁸⁹ over CENP-A nucleosome. Similarly, a 0-, 2-, and 3-fold excess of CENP-N¹⁻²⁸⁹ was added to 'AC' complex (where the ratio of CENP-C⁴²⁶⁻⁵³⁷ to CENP-A nucleosome is 2:1).

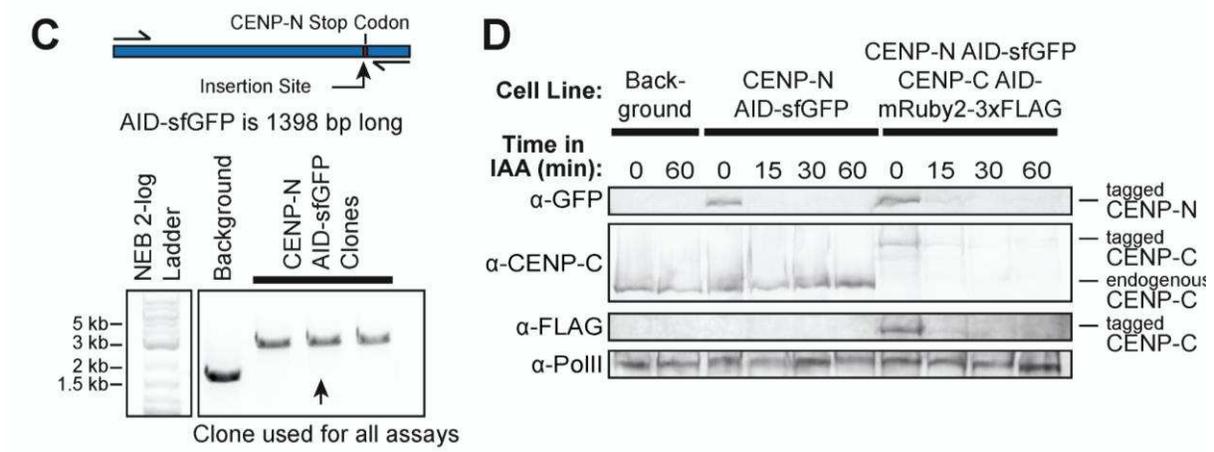
B. Nickel beads pull down assay for 'AN' complex and 'ANC' complex. CENP-N¹⁻²⁸⁹ has a His-tag at the C-terminus. Nickel beads were incubated with the samples at room temperature for 10 minutes. The samples were mixed every 2 minutes. Sample input, flow through (FL), buffer wash, and beads (bound sample) were analyzed by 15% SDS-PAGE.

A CENP-N AID-sfGFP Cell Line**B** CENP-N AID-sfGFP, CENP-C AID-mRuby2-3xFlag Cell Line

Supplemental Figure 3.3 AID tagged proteins are degraded within 30 min following IAA addition

A. Representative images of the CENP-N AID-sfGFP cell line response to 1mM IAA treatment over time. ACA is the human anti-centromere antibody serum used to mark centromeres. Quantification of the centromeric GFP signal is shown below the images. Background was determined from the background Tir1 DLD1 cell line that does not express GFP.

B. Representative images of the CENP-N AID-sfGFP, CENP-C AID-mRuby2-3xFLAG cell line response to 1mM IAA over time. The images in the second and third panels show stained signal assaying the tags. The images in the sixth panels are separate coverslips stained for CENP-C using an antibody. Quantification of these signals are shown below the images. Data are presented as mean \pm SEM for three independent replicates. Scale bar represents 5 μ m.

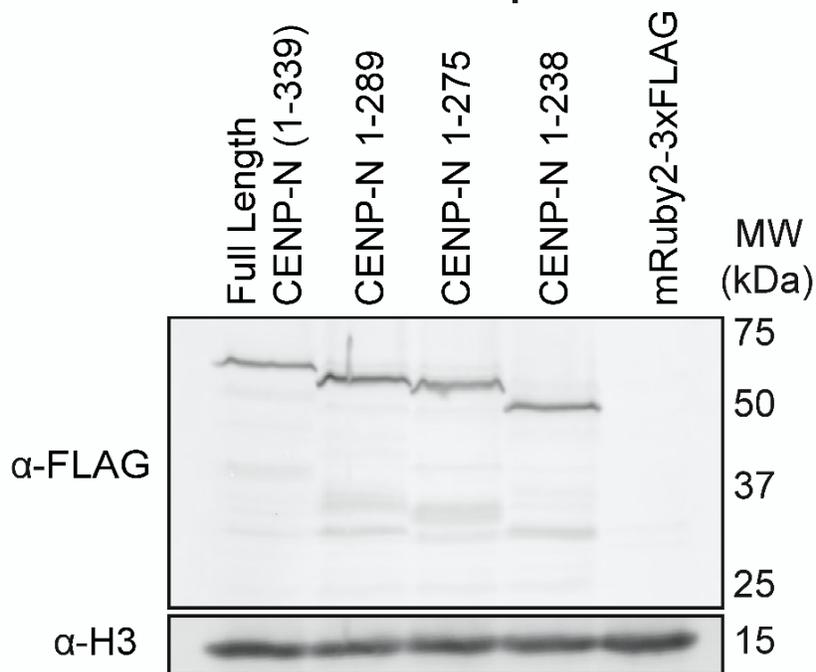


Supplemental Figure 3.3 AID tagged proteins are degraded within 30 min following IAA addition

C. Genomic PCR showing homozygous modification of the endogenous CENP-N locus

D. Representative Western blot showing degradation of tagged species with IAA treatment over time.

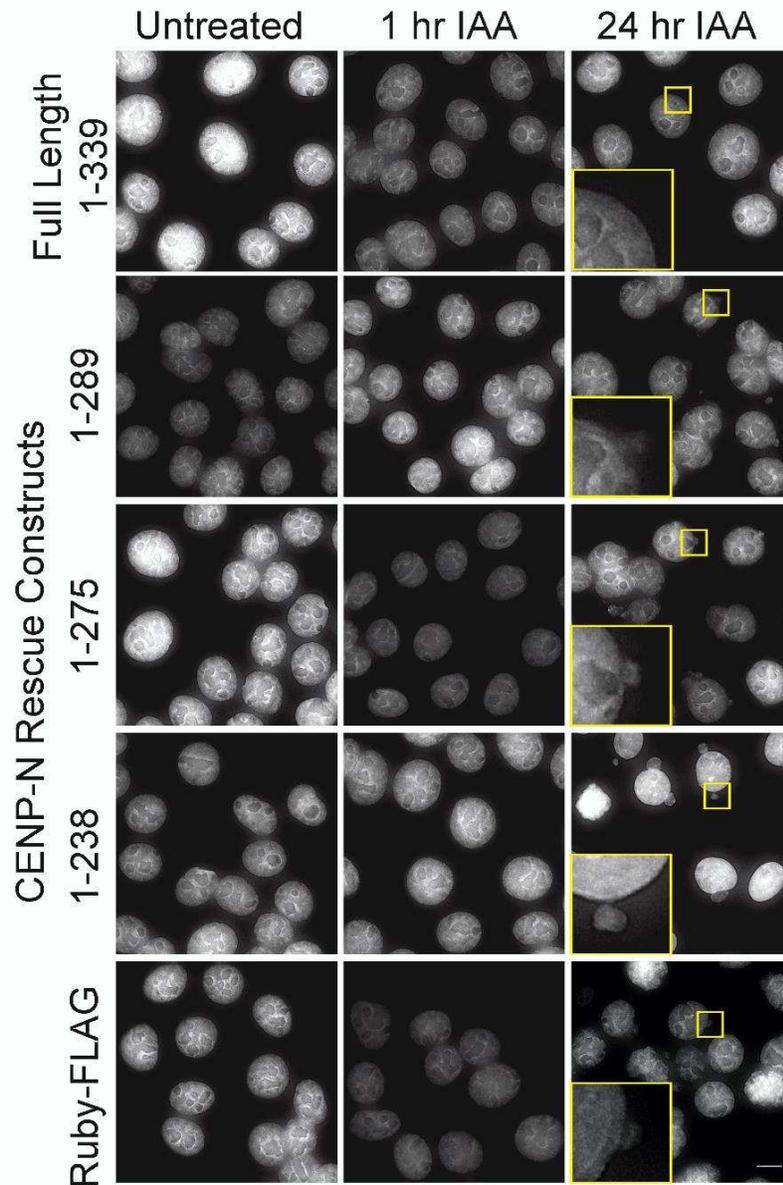
A CENP-N mRuby2-3xFLAG tagged
rescue construct expression



Supplemental Figure 3.4 CENP-N degradation causes micronuclei formation.

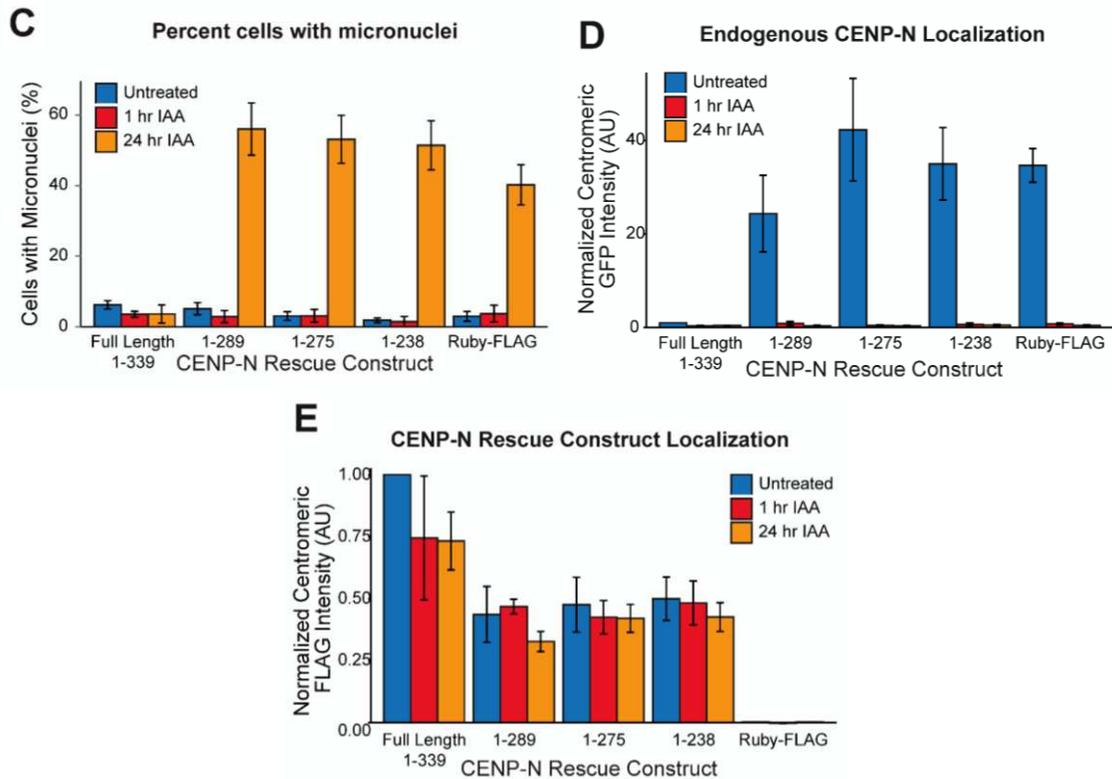
A. Full length or CENP-N truncations rescue constructs were stably integrated into the CENP-N AID-sfGFP background. Cells stably expressing these constructs were Western blotted with anti-Flag antibody to assay expression levels.

B Degrade endogenous CENP-N,
assay rescue by constructs



Supplemental Figure 3.4 CENP-N degradation causes micronuclei formation.

B. Representative images of cells stained with Hoechst 33258 to visualize DNA are shown for cells not treated with IAA or treated for either 1 hour or 24 hours. The inserts highlight the presence or absence of micronuclei.



Supplemental Figure 3.4 CENP-N degradation causes micronuclei formation.

C. Quantification of the percentage of cells containing micronuclei after complementation with different transgenes.

D. Quantification of the levels of endogenous GFP-CENP-N at centromeres in cells expressing the different mRuby2-3XFLAG-tagged transgenes.

E. Quantification of transgene localization to centromeres measured through localization of the centromeric FLAG signal.

Data are presented as mean \pm SEM for three independent replicates. Scale bar represents 10 μ m.

LIST OF PUBLICATIONS

Satyakrishna Pentakota *, **Keda Zhou** *, Charlotte Smith, Stefano Maffini, Arsen Petrovic, Garry P. Morgan, John R. Weir, Ingrid R. Vetter, Andrea Musacchio & Karolin Luger, 2017, Decoding the centromeric nucleosome through CENP-N, *elife*. 6: e33442.

Shengya Cao*, **Keda Zhou***, Zhening Zhang, Karolin Luger, Aaron F. Straight, 2018, Constitutive centromere-associated network contacts confer differential stability on CENP-A nucleosomes in vitro and in the cell, *molecular biology of the cell*, E17-10_0596.

Yelagandula, R., Hume Stroud, Sarah Holec, **Keda Zhou**, Suhua Feng, Xuehua Zhong, Uma M. Muthurajan, Xin Nie, Tomokazu Kawashima, Martin Groth, Karolin Luger, Steven E. Jacobsen, and Frédéric Berger, (2014) The histone variant H2A.W defines heterochromatin and promotes chromatin condensation in *Arabidopsis*. *Cell*.158, 98–109.

Uma Muthurajan, Francesca Mattioli, Serge Bergeron, **Keda Zhou**, Yajie Gu, Srinivas Chakravarthy, Pamela Dyer, Thomas Irving, and Karolin Luger, (2016), In vitro Chromatin Assembly - Strategies and Quality Control. *Methods Enzymol*.573: 3–41.