Insight into the role of CENP-N in kinetochore structure and function

Inaugural-Dissertation

Zur

Erlangung des Doktorgrades

Dr. rer. nat.

der Fakultät für Biologie

an der

Universität Duisburg-Essen

Vorgelegt von

Satya Krishna Pentakota

aus, Visakhapatnam, Indien

durchgeführt am

Max Planck Institut für molekulare Physiologie

Abteilung für mechanistische Zellbiologie

December 2018

Die der vorliegenden Arbeit zugrunde liegenden Experimente wurden am Max Planck Institut für molekulare Physiologie in der Abteilung für mechanistische Zellbiologie durchgeführt.

1. Gutachter: Prof. Dr. Andrea Musacchio

2. Gutachter: Prof. Dr. Stefan Westermann

3. Gutachter: Prof. Dr. Elsa Sanchez-Garcia

Vorsitzender des Prüfungsausschusses: Prof. Dr. Stefan Westermann

Tag der mündlichen Prüfung: 05-02-2019.

In the context of this doctoral work, the following articles were published:

- 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.

- John R. Weir*, Alex C. Faesen*, Kerstin Klare*, Arsen Petrovic, Federica Basilico, Josef Fiscböck, Satyakrishna Pentakota, Jenny Keller, Marion E. Pesenti, Dongqing Pan, Doro Vogt, Sabine Wohlgemuth, Franz Herzog, Andrea Musacchio. 2016. "Insights from biochemical reconstitution into the architecture of human kinetochores". *Nature* 8;537(7619):249-253. (* equal contribution)

Content

List of Figures

List of tables

List of abbreviations

AEBSF	4-2-(AminoEthyl) BenzeneSulfonyl Floride
APC	Anaphase promoting complex
AUC	Analytical ultra centrifugation
BSA	Bovine serum albumin
BUB	Budding uninhibited by benzomyl
B3BD	Bub3-binding domain
CATD	Centromere targeting domain
CCAN	Constitutive centromere associated network
CDK	Cyclin-dependent kinase
CENP	Centromeric protein
CLN-HD	CENP-L and CENP-N homology domain
CPC	Chromosomal passenger complex
CREST	Calcinosis, Raynaud's syndrome, Esophageal dysmotility, Sclerodactyly, Telangiectasia
Cryo-EM	Cryo-electron microscopy
DAPI	4,6-diamidin-2-phenylinodoldihydrochlorid
DMEM	Dulbecco's modified eagle's medium
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
FBS	Fetal bovine serum
GFP	Green fluorescent protein
GST	Glutathione S-transferase
HJURP	Holiday junction recognition protein
IPTG	Isopropyl-D-thiogalactopyranoside
ITC	Isothermal titration calorimetry
IF	Immunofluorescence
MCC	Mitotic checkpoint complex
MCS	Multiple cloning site
MSA	Multiple sequence alignment
PBS	Phosphate buffered saline
PEG	Polyethylene glycol
PMSF	Phenylmethylsulfonyl fluoride

1 Introduction

1.1 The Cell Cycle

The cell cycle is a ubiquitous and tightly regulated process involved in the growth and proliferation of the cells (Alberts 2017). Coordination of many regulatory proteins directs a cell through different phases of its division cycle. Morphologically, the cell cycle can be sub-divided into two phases: interphase and mitotic (M)-phase. Interphase itself consists of three phases: G1 (Gap 1), S (synthesis phase), and G2 (Gap 2) phases. The mitotic phase typically consists of two main events: nuclear division (mitosis) and cell division (cytokinesis) (Figure 1-1). The G1-phase corresponds to the interval (gap) between mitosis and DNA replication. During G1-phase, the cell is metabolically active and continuously grows. The G1-phase is followed by the S-phase, during which DNA replication takes place. S-phase is followed by the G2-phase in which cell growth continues and prepares to progress to the next stage of the cell cycle. In the M-phase, the cell segregates its chromosomes equally into the resulting two daughter cells. The G0-phase of the cell cycle (also known as quiescent stage) was originally used to describe the cells that are outside of the replicative cell cycle. Cells in the G0-phase remain metabolically active but can no longer proliferate unless triggered by extracellular signals (Alberts et al. 2002).

In order to ensure the fidelity of cell reproduction and growth the cell cycle progression is monitored at specific points, called checkpoints. Checkpoints can be defined as transition points at which the progression of the cells to the next phase can be halted until conditions are favorable. The central machines that drive the cell cycle progression are the "Cyclin-dependent kinases (CDKs)", which belongs to the serine-threonine kinase family. Cyclin-dependent kinases are activated by binding to proteins called Cyclins (Evans et al. 1983; Morgan 1995). Different types of Cyclins are made during different phases of the cell cycle which results in the formation of specific CDK/Cyclin complexes that trigger distinct cell cycle events (Coudreuse & Nurse 2010; Loog & Morgan 2005; Pines 1991).

The first checkpoint in the G1-phase is known as the "start" which is mainly influenced by the cell size, growth factors and cell nutrients (Pardee 1974; Hartwell et al. 1970). Once cells have passed this first checkpoint, they are committed to enter into S-phase for genome duplication. The G2/M checkpoint, also called the DNA replication checkpoint is the second checkpoint, which ensures that all the chromosomes have been accurately replicated without any DNA damage. The third checkpoint, also called the spindle assembly checkpoint (SAC) operates during mitosis. The SAC ensures the sister chromatid separation, only after proper attachment of chromosomes to the mitotic spindle (Santaguida & Musacchio 2009; Musacchio & Salmon 2007).
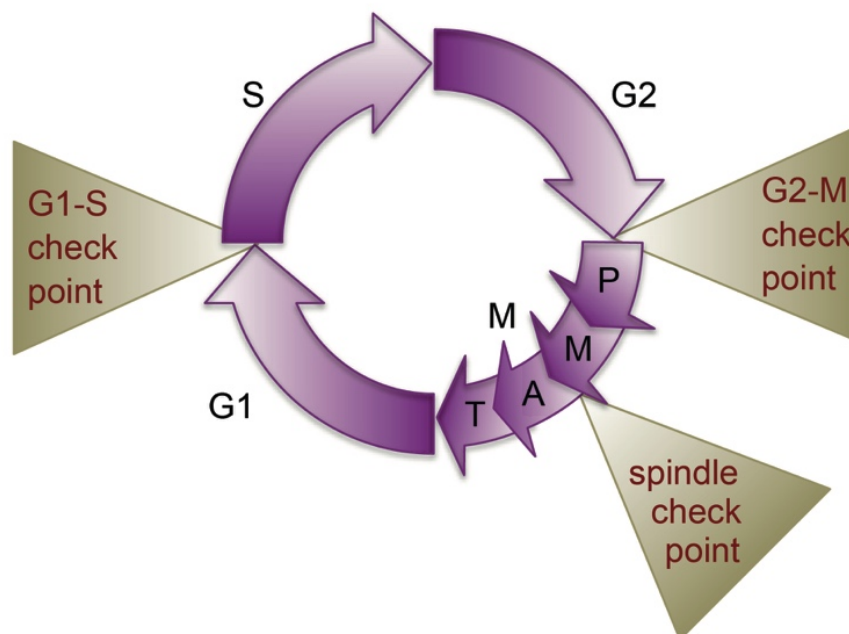
Figure 1-1 The cell cycle

The cell cycle is comprised of interphase (G1, S and G2 phases) and mitotic phase (M-phase). Mitotic phase is comprised of prophase (P), metaphase (M), anaphase (A) and telophase (T). The cell cycle is monitored by three checkpoints: G1/S, G2/M and Spindle assembly checkpoint (SAC), as depicted. Figure adapted from (Harashima et al. 2013).

1.2 Mitosis

The term mitosis was coined by Walther Flemming in 1880s (Flemming 1882). Mitosis can be described as a process during which previously duplicated chromosomes are equally distributed into two newly formed daughter cells. This process must be precisely regulated, as the maintenance of genomic integrity is crucial for the viability of the cell.

Mitosis can be subdivided into five distinct phases: prophase, pro-metaphase, metaphase, anaphase and telophase as depicted in Figure 1-2 (Alberts 2017). In prophase, chromosomes are condensed by a protein complex called condensin, the nuclear envelope breaks down and microtubules, filamentous intracellular structures, composed of alpha- and beta-tubulin subunits, start to nucleate from structures called centrosomes. During pro-metaphase, microtubules start to organize themselves into fibrous structures known as the mitotic spindle. Spindle fibers are composed of multiple microtubules and are able to capture chromosomes via large protein assemblies, known as kinetochores, on the centromere (Walczak & Heald 2008). In metaphase, the chromosomes align along the metaphase plate of the spindle apparatus. Key to proper chromosome segregation is the bi-orientation of the chromosomes, whereby each sister chromatid attaches to microtubules originating from opposite spindle poles. In anaphase, the cohesin complex, that holds the sister chromatids together is broken down, allowing the sister chromatids to be separated by the force produced by the depolymerizing microtubules moving them to the opposite spindle poles (Hays et al. 1982). In order to ensure high fidelity of cell division, the spindle assembly checkpoint (SAC) delays the anaphase onset until all the chromosomes achieve bi-orientation, this will be discussed further in section 1.5 (Musacchio & Salmon 2007). In telophase, a new nuclear membrane is formed around each set of chromosomes forming two identical daughter nuclei and the chromosomes begin to decondense, completing the process of mitosis. Cytokinesis marks the final step of the cell division, a contractile ring forms at the site of cell division, separating the two daughter cells (Alberts 2017).
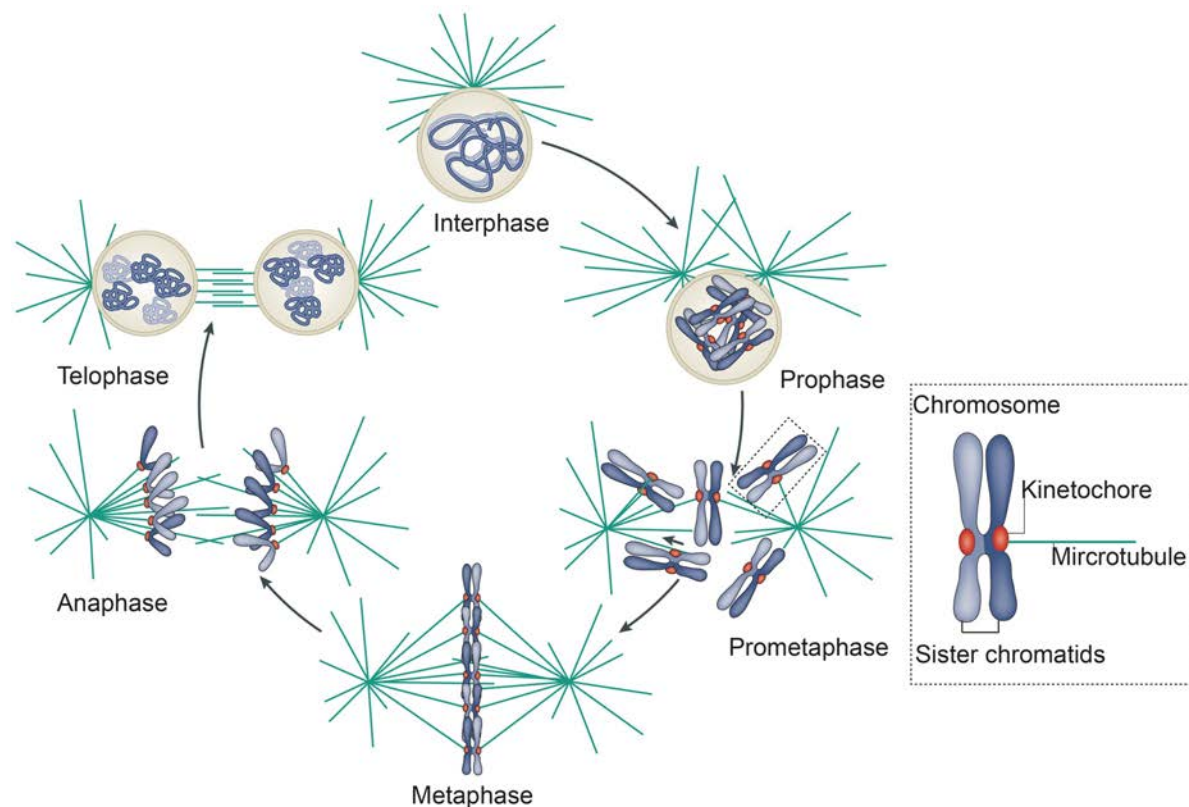
Figure 1-2 Schematics of mitotic cell division

Chromosomes are condensed during prophase. After nuclear envelope breakdown, kinetochores are attached by spindle microtubules. In metaphase, the chromosomes are bi-oriented and are aligned along the metaphase plate. In anaphase, the sister chromatids move towards the opposite spindle poles. During telophase, the chromatin decondenses and the nuclear envelope reforms in order to produce two daughter nuclei. Figure adapted from (Cheeseman & Desai 2008).

1.3   Centromere structure and specification

The centromere is a unique region on the chromosome where a large proteinaceous complex called the kinetochore is assembled. Centromeres can be broadly classified into three categories. Eukaryotic organisms like budding yeast, *Saccharomyces cerevisiae*, have point centromeres, which are defined by a 125-basepair (bp)-specific DNA sequence, that is sufficient to assemble the kinetochores (Hegemann & Fleig 1993; Pluta et al. 1995). The kinetochores of *S. cerevisiae* are regarded as the simplest kinetochores, because they bind to a single microtubule. In contrast to budding yeast, human kinetochores consist of a regional centromere, which occupy much longer DNA regions and bind to multiple microtubules (Kursel & Malik 2016; Fukagawa & Earnshaw 2014; Malik & Henikoff 2009). Human centromeres consist of repetitive DNA sequences called alpha-satellite repeats, which are based on a 171 bp monomer organized in a tandem head-to-tail fashion. The biological significance of these

repetitive sequences remains unclear. The formation of neo-centromeres on DNA devoid of apha-satellite repeats, suggests that the presence of centromeric DNA does not necessarily lead to the formation of a functional centromere. This in turn hints that the centromere specification is regulated by the sequence-independent epigenetic mechanism (Aldrup-Macdonald & B. A. Sullivan 2014). Finally, in organisms like *Caenorhabditis elegans*, the centromere extends over the entire length of the chromosome (holocentric) (Maddox et al. 2004).

1.3.1 Centromeric chromatin organization

Centromeres are epigenetically defined by the incorporation of the evolutionarily conserved histone H3-like variant called CENP-A (Earnshaw & Rothfield 1985). CENP-A was initially identified in the sera of patients suffering from Calcinosis, Raynaud's syndrome, Esophageal dysmotility, Sclerodactyly and Telangiectasia (CREST) syndrome (Earnshaw & Rothfield 1985). Immunoblots from these patients reveal three recurring bands, named CENP-A, CENP-B and CENP-C. CENP-A homologs have been identified in different species, including Cse4 in budding yeast, Cnp1 in fission yeast, and CID/CenH3 in *drosophila*, with all having been shown to be essential for chromosome segregation (Buchwitz et al. 1999; Henikoff et al. 2000; Takahashi et al. 2000).

CENP-A is a 16 kDa protein that consists of a histone fold domain with 62% sequence identity to that of a canonical histone H3 and an N-terminal tail that is highly divergent. The histone fold domain consists of three alpha helices connected via two loops (Figure 1-3). Further studies revealed that the L1 loop and the second alpha helix are crucial for the recruitment of CENP-A to the centromeres, therefore this region was named as the centromere targeting domain (CATD) (Black, Brock, et al. 2007). Studies on chimeric histone H3 that contains an engineered version of CATD (H3^{CATD}) demonstrates that CATD is sufficient to recruit CENP-A to the centromeres (Figure 1-3) (Black, Jansen, et al. 2007).

Several studies have demonstrated the importance of CENP-A. Inactivation or depletion of CENP-A results in the mis-segregation of chromosomes during the M-phase (Takahashi et al. 2000) (Warburton et al. 1997; Vafa & K. F. Sullivan 1997). CENP-A has also been shown to be required for the recruitment of several inner kinetochore proteins thus indicating its important role in kinetochore assembly (Stellfox et al. 2013). The fundamental unit of chromatin is the nucleosome, which consists of a

tetramer of H3 and H4 and two dimers of H2A and H2B. Previous studies on CENP-A nucleosomes have reported alternative forms of nucleosomes such as tetrasome (two copies of CENP-A and H4 but lacking H2A and H2B) and hemisome (a single copy of each histone (CENP-A, H2A, H2B and H4). However, recent studies demonstrate that the predominant form of a CENP-A nucleosome is octameric (consisting of a tetramer of CENP-A and H4 and two dimers of H2A and H2B) (Hasson et al. 2013; Tachiwana et al. 2011; Shelby et al. 1997).

Unlike canonical nucleosomes whose deposition is replication-coupled, the deposition of CENP-A is uncoupled from DNA replication. In human cells, the deposition of CENP-A takes place during the G1-phase of the cell cycle (Black, Jansen, et al. 2007). Like other histones, the deposition of CENP-A onto centromeric chromatin requires histone chaperones. Histone chaperones are broadly defined as a group of proteins that bind to histones and regulate the process of nucleosome assembly. Different histone chaperones exist for the assembly of different histone variants or isoforms. HJURP (Holliday junction recognition protein), a chaperone that binds to the CENP-A/H4 tetramer, has been shown to be required for the loading of new CENP-A onto centromeric chromatin (Shuaib et al. 2010). Structural analysis of HJURP in complex with a CENP-A/H4 tetramer reveals that HJURP recognizes the CATD domain of CENP-A via its N-terminus (Zasadzińska et al. 2013; Hu et al. 2011). Similar to HJURP, Scm3 (suppressor of chromosome mis-segregation) is also involved in the incorporation of Cse4^{CENP-A} nucleosomes in yeast (Sanchez-Pulido et al. 2009). Depletion of HJURP in human cells or Scm3 in yeast causes similar defects in CENP-A assembly. CENP-A deposition by HJURP requires an additional three-subunit Mis 18 complex (Hayashi et al. 2004) that consists of Mis18α, Mis18β and M18BP1 (Fujita et al. 2007; Maddox et al. 2007). Recent work from our laboratory has demonstrated that two copies of M18BP1 bind to a hexamer of Mis18α:Mis18β (4:2) (Pan et al. 2017). Previous studies have suggested that the deposition of CENP-A onto centromeres is regulated by inhibitory CDK phosphorylation of CENP-A, HJURP and M18BP1 (Yu et al. 2015; Müller et al. 2014; McKinley & Cheeseman 2014; Silva et al. 2012). Further studies have demonstrated that the phosphorylation of M18BP1 by CDK prevents the formation of Mis 18 complex, thus restricting CENP-A deposition (McKinley & Cheeseman 2014; Pan et al. 2017). CDK activity is high during the M-phase and declines with the destruction of Cyclin B at anaphase. The reduction of CDK activity

allows the formation of Mis 18 complex, which primes the centromere and prepares for the loading of new CENP-A during the G1-phase of the cell cycle.
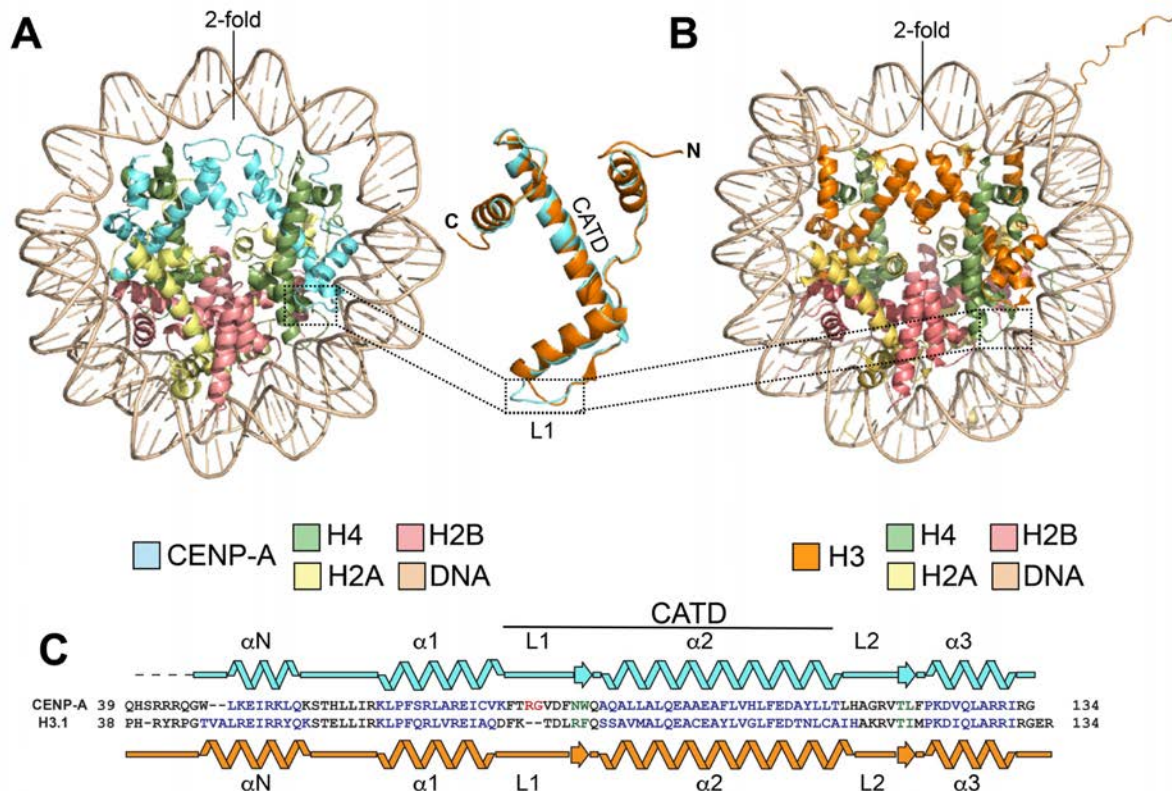
Figure 1-3 Comparison of CENP-A and H3 nucleosomes

(A-B) Front view of CENP-A (PDB ID 3AN2) (Tachiwana et al. 2011) and H3 nucleosome structures (PDB ID 1AOI) (Luger et al. 1997). (C) Protein sequences of CENP-A and H3 are aligned with the secondary structure elements. Figure adapted from (Musacchio & Desai 2017).

1.4   Kinetochore organization and function

Kinetochores were initially observed by Metzner in 1894, where the fibers of the spindle associated with a distinct region of the chromosome called the "kinetic region". Faithful segregation of chromosomes by kinetochores requires a coordinated action of ~100 individual proteins organized into several sub-complexes (Wittmann et al. 2001; Walczak & Heald 2008). The overall architecture of the kinetochores are conserved from yeast to mammals (Kitagawa & Hieter 2001; Meraldi et al. 2006). Kinetochores were first observed under electron microscopy as the trilaminar objects with two electron opaque layers separated by a middle translucent layer (Figure 1-4). The two electron-dense layers are called inner and outer kinetochore (McEwen et al. 2007;

McEwen et al. 1998). Based on their spatial localization and their differential functions, these trilaminar objects can be categorized into four modules (Santaguida & Musacchio 2009):

1)   A fibrous corona that contains proteins of the spindle assembly checkpoint, is only visible at unattached kinetochores. SAC is a feedback control mechanism that monitors the kinetochore-microtubule attachments and delays anaphase onset until all sister chromatids are properly attached to opposite spindle poles. The SAC is discussed in more detail in the section 1.5.

2)   The outer kinetochore or Knl1-Mis12-Ndc80 (KMN) network, that bridges the inner kinetochore with spindle microtubules. It also serves as the receptor for the spindle assembly checkpoint (SAC). The KMN network is discussed in more detail in section 1.7.

3)   The inner kinetochore or constitutive centromere associated network (CCAN), which provides a platform for the assembly of the kinetochore by connecting the centromeric chromatin to the outer kinetochore (Foltz et al. 2006; Okada et al. 2006). The CCAN is discussed in more detail in section 1.8.

4)   The inner centromere comprises the chromosomal passenger complex (CPC). The CPC is responsible for discriminating the improper from proper microtubule attachments and stabilizing the later in a tension-dependent manner (Carmena et al. 2012; Wang et al. 2010). The CPC is discussed in more detail in section 1.6.
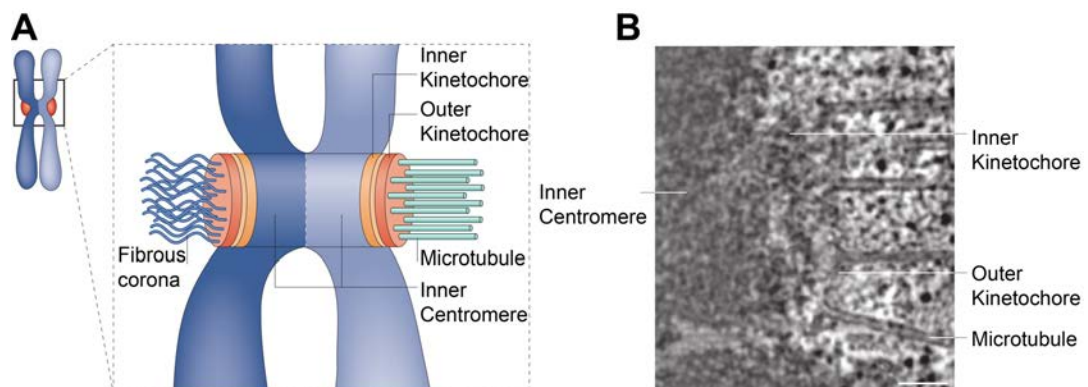
Figure 1-4 Architecture of the vertebrate kinetochore

(A) Schematic view of a mitotic chromosome. (B) Electron micrograph of a human kinetochore. The inner kinetochore, outer kinetochore, the inner centromere and the fibrous corona which are detectable on the unattached kinetochore, are shown. The electron micrograph represents a single slice from a tomographic volume of a high-pressure frozen mitotic cell. Scale bar: 100 nm. Figure adapted from (Cheeseman & Desai 2008).

1.5 Spindle assembly checkpoint

The Spindle Assembly Checkpoint (SAC) is a quality control mechanism that ensures the fidelity of chromosome segregation (Nezi & Musacchio 2009; Musacchio & Salmon 2007). Despite its name, the SAC monitors the status of kinetochore-microtubule attachments rather than spindle assembly. The key function of the SAC is to prevent the onset of anaphase until all the chromosomes are stably attached to the spindle (Lara-Gonzalez et al. 2012; Musacchio & Salmon 2007). The SAC is activated in the presence of unattached kinetochores and deactivated upon proper attachment of kinetochores to microtubules, resulting in the onset of anaphase.

Most of the components of the SAC, such as the budding uninhibited by benzimidazole genes Bub1 and Bub3, the mitotic arrest deficient (Mad) genes Mad1, Mad2 and Mad3 (BubR1 in humans), monopolar spindle protein 1 (MPS1) and Aurora B kinase were initially identified by genetic screens in *S. Cerevisiae* (R. Li & Murray 1991; Hoyt et al. 1991; Weiss & Winey 1996). Further studies on SAC components demonstrated that almost all SAC components are localized to unattached kinetochores (Musacchio & Salmon 2007; Lara-Gonzalez et al. 2012).

The downstream target of the SAC effectors is the anaphase promoting complex or Cyclosome (APC/C). The APC/C is a 1.2 MDa, multi-subunit E3-ubiquitin ligase that targets the degradation of cell-cycle regulatory proteins via the proteasome (Pines 2011). Activation of APC/C requires one of the two co-activators, Cdc20 or Cdh1. Mitotic APC/C is activated by Cdc20 (APC/C^{Cdc20}), while Cdh1 is responsible for the activation of APC/C after mitotic exit. The Cdc20 activated APC/C triggers mitotic exit by polyubiquitination of two crucial substrates Cyclin-B and Securin (Glotzer et al. 1991; Yamamoto et al. 1996). APC/C^{Cdc20} -mediated destruction of Cyclin-B results in the inactivation of CDK1, which leads to the dephosphorylation of CDK1 substrates thereby allowing mitotic exit. Degradation of Securin results in the activation of Seperase, which in turn, cleaves the Cohesin complex allowing the separation of sister chromatids and the onset of anaphase (Funabiki et al. 1996; Holloway et al. 1993; Sudakin et al. 1995). APC/C^{Cdc20} is inhibited by the formation of the SAC effector

complex known as the mitotic checkpoint complex (MCC). MCC is comprised of Cdc20/Mad2 and BubR1/Bub3 sub-complexes, in which BubR1 and Mad2 bind simultaneously to distinct sites on the same Cdc20 molecule (Sudakin et al. 2001; Hardwick et al. 2000; Fraschini et al. 2001). Most studies suggest that both BubR1 and Mad2 can inhibit the APC/C (Fraschini et al. 2001; Hardwick et al. 2000; Sudakin et al. 2001) (Figure 1-5). BubR1 promotes docking of MCC onto APC/C and blocks substrate recruitment to the APC/C through its KEN (Lys-Glu-Asn) and D (destruction) box (Alfieri et al. 2016; Lara-Gonzalez et al. 2011; Yamaguchi et al. 2016). Mad2 inhibits APC/C by competing for the same binding site on Cdc20 (Izawa & Pines 2012).

The formation of the MCC requires the hierarchical recruitment of the SAC proteins. The outer kinetochore or KMN network serves as the crucial platform for the SAC components. The Knl1 component within the KMN network, recruits Bub1/Bub3 and BuR1/Bub3 via its KI (lysine-isoleucine)-motifs (Krenn et al. 2012; Krenn et al. 2014; Kiyomitsu et al. 2007; Kiyomitsu et al. 2011). More recently it has been shown that the kinetochore recruitment of Bub1/Bub3 requires the phosphorylation of MELT (methionine-glutamic acid-leucine-threonine) repeats by the SAC kinase MPS1 (Shepperd et al. 2012; Yamagishi et al. 2012; Primorac et al. 2013). Bub1 is a Ser/Thr kinase that forms a stoichiometric complex with Bub3 which is required for its kinetochore localization. Previous studies on Bub1 have shown that the Bub1 kinase activity is dispensable for the activation of the SAC (Klebig et al. 2009; Perera et al. 2007). Bub1 localizes to kinetochores already in prophase and is involved in the recruitment of downstream SAC components, such as BubR1, Bub3, Mad1, Mad2 and Cdc20 (Boyarchuk et al. 2007; Klebig et al. 2009). Bub1 promotes the incorporation of BubR1/Bub3, Cdc20 and Mad2, into the MCC. Like Bub1, BubR1 also bind to Bub3 with a similar binding mechanism (Larsen et al. 2007). BubR1 is a crucial component of MCC, together with Bub3, Cdc20 and Mad2, thus contributing directly to the SAC function (Hardwick et al. 2000; Sudakin et al. 2001). Although both Bub1 and BubR1 display a high degree of sequence and structural similarity, they perform different functions in the SAC.

The SAC proteins Mad1/Mad2 and Bub1/Bub3, both phosphorylated by MPS1, constitute the minimal platform required for the assembly of MCC (Faesen et al. 2017). Mad2 can adopt two distinct conformations: an open (unbound) conformation O-Mad2 and a closed (bound) conformation C-Mad2 (De Antoni et al. 2005; Sironi et al. 2002; Mapelli et al. 2007). The primary step of the MCC catalytic assembly is the dimerization

of O-Mad2 to C-MAD2 bound to Mad1 that is essential for catalysis (Faesen et al. 2017). O-Mad2 is then converted to the C-Mad2 and simultaneously binds to Cdc20. The C-Mad2 within the Mad1:C-Mad2 complex acts as a template that promotes the conversion of an O-Mad2 to a C-Mad2 bound to Cdc20, this is commonly referred to as the Mad2-template model (De Antoni et al. 2005; Sironi et al. 2002). The MPS1 phosphorylated, Mad1:C-Mad2 serves as the catalyst for the accumulation of Cdc20:C-Mad2 complexes (Faesen et al. 2017). Subsequently, Mad2:Cdc20 binds BubR1, which is not required for the catalysis, but it is essential for the MCC stability and inhibition of the APC/C (Prinz et al. 2016; Luo et al. 2000; Chao et al. 2012). The MCC inhibits the APC/C in two ways; first, it sequesters Cdc20 away; and second, it reduces its ability to recruit its substrates such as Cyclin-B1 and Securin (Herzog et al. 2009; Chang et al. 2015). Within the MCC, BubR1 somehow interferes with the binding of these substrates to APC, which results in the inhibition of APC. Thus, once the MCC is assembled, both Mad2 and BubR1 inhibit the formation of APC/C^{Cdc20}, which, in turn delays anaphase onset.

Once correct bi-orientation is achieved, the SAC must be inactivated. SAC silencing occurs via several pathways: the dephosphorylation of the SAC components by the phosphatases counter acting the mitotic kinases (Lara-Gonzalez et al. 2012), the removal of the SAC components from the kinetochores by the corona, and MCC disassembly by p31 comet and ATPase Trip13 (Corbett 2017). Upon stable kinetochore-microtubule attachments, the minus-end directed motor protein Dynein removes SAC components by a process called stripping (Howell et al. 2001). The removal of Mad1:Mad2 complex from the attached kinetochores appears to be critical for efficient SAC inactivation (Maldonado & Kapoor 2011). The RZZ complex recruits Spindly to the kinetochores and has a role in the process of SAC silencing. The RZZ complex is comprised of three proteins, ROD, Zwilch and Zw10 (named after Rough Deal, Zwilch and Zeste white 10), that in turn recruit the adaptor protein Spindly. Together with Spindly, the RZZ complex is also required for the recruitment of cytoplasmic Dynein-Dynactin complexes to the kinetochore (Urnavicius et al. 2015; Mosalaganti et al. 2017; Scaërou n.d.). Consequently, Dynein removes the SAC proteins Mad1-C-Mad2, RZZ and Spindly from the attached kinetochores (Gassmann et al. 2008; Chan et al. 2009). Mutations within the conserved regions of Spindly abrogate kinetochore recruitment of Dynein-Dynactin, which in turn blocks corona shedding and SAC silencing (Howell et al. 2001; Varma et al. 2008). The Dynein-

dependent pathway does not appear to be conserved in yeast, implying that Dynein-independent pathways of SAC silencing might exist in different organisms (Funabiki & Wynne 2013). If the phosphorylation of the SAC components by mitotic kinases is crucial for the activation of SAC, phosphatases play an equally important role in SAC inactivation. The recruitment of protein phosphatase 1 (PP1) to the N-terminal region of Knl1 results in the dephosphorylation of MELT repeats within the Knl1 resulting in the removal of SAC components (Bub1/BubR1/Bub3) (D. Liu et al. 2010; Meadows et al. 2011; Rosenberg et al. 2011). Thus, PP1 performs two important functions: First, it stabilizes the kinetochore-microtubule attachments by counteracting kinases such as Aurora B. Second, it dephosphorylates the MELT repeats on Knl1, which is essential for the removal of the SAC components Bub1/BubR1/Bub3 protein complexes from the kinetochores (Rosenberg et al. 2011; Espert et al. 2014). PP2A^{B56} is also known to contribute to SAC silencing by promoting the recruitment of PP1 to the kinetochores, thereby counteracting Aurora B and MPS1 activity (Espert et al. 2014).
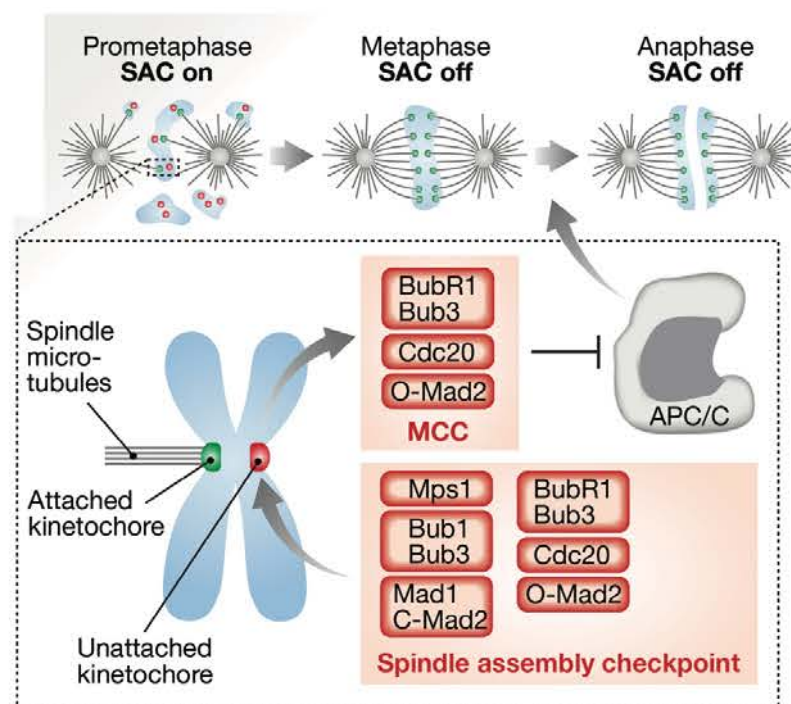
Figure 1-5 Spindle assembly checkpoint (SAC)

Schematic view of the spindle assembly checkpoint signaling. SAC proteins are recruited to the unattached kinetcohores (red), resulting in the formation of MCC. MCC inhibits APC/C complex which in turn prevents the progression into anaphase. Upon proper Kinetochore-microtubule attachments (green), the SAC turns off, leading to the onset of anaphase. Figure adapted from (Overlack et al. 2014).

1.6 Regulation of kinetochore-microtubule attachments

Accurate chromosome segregation involves the discrimination of correct (amphitelic, bi-oriented) and incorrect (syntelic, merotelic) or incomplete (monotelic) kinetochore-microtubule (kMT) attachments (Nicklas & Koch 1969; X. Li & Nicklas 1995). Kinetochore-microtubule attachments are regulated by the activity of the KMN network, coupled to the polymerization and depolymerization dynamics of the kMT plus-ends by microtubule-associated proteins (MAPs). The Spindle and kinetochore-associated (Ska) complex in metazoans and its homolog Dam1 complex in yeast have been identified as microtubule binding complexes (Hanisch et al. 2006; Daum et al. 2009; Miranda et al. 2005; Westermann et al. 2005). Both Dam1 and Ska complexes are dependent on Ndc80 complex for their kinetochore localization and have been shown to be required for the stabilization of kinetochore-microtubule attachments (Welburn et al. 2009; Gaitanos et al. 2009; Schmidt et al. 2012). This suggests that both Dam1 complex and SKA complex are proposed to work through the regulation of Ndc80 complex or the KMN network which is required for effective kinetochore-microtubule attachments (Gaitanos et al. 2009; Schmidt et al. 2012; Tien et al. 2010; Welburn et al. 2009).

Kinetochore-microtubule attachments are regulated by a balance between kinetochore-localized kinases and phosphatases that are involved in the "error correction" pathway. The basic principle of error correction is that the incorrect microtubule attachments are detached by the phosphorylation of key binding interfaces such as the KMN network, but as soon as the correct microtubule attachments are formed, these key binding interfaces are dephosphorylated, while the attachments are stabilized. Aurora B, a Ser/Thr kinase and protein phosphatase 2A (PP2A) are the main components in this pathway. Aurora B is a subunit of the chromosomal passenger complex (CPC), together with Survivin, INCENP and Borealin (Carmena et al. 2012). Recruitment of Aurora B to the centromeres is dependent on Haspin kinase, which phosphorylates histone H3 on Thr[3], along with the phosphorylation of histone H2A by Bub1. Aurora B localizes to centromeres and destabilizes the erroneous kinetochore-microtubule attachments (Lampson et al. 2004; Pinsky et al. 2006). Specifically, Aurora B phosphorylates the N-terminal tail of Hec1, thereby decreasing the microtubule binding affinity to the Ndc80 complexes (Ciferri et al. 2008; J. G. DeLuca & Musacchio 2012; Alushin et al. 2010). Phosphorylation of Dam1 and Ska complexes by Aurora B have been shown to negatively regulate their interaction with the Ndc80 complexes in

order to prevent stabilization of incorrect attachments (Chan et al. 2012; Schmidt et al. 2012). How Aurora-B activity is enhanced in the presence of incorrect attachments remains unclear. Current models suggest that tension on the bi-oriented kinetochores leads to spatial separation of Aurora B from its substrates, which results in differential phosphorylation levels of those substrates. In contrast, unattached or erroneously attached kinetochores exhibit low tension, whereby Aurora B is located closer to its kinetochore substrates, leading to a high level of phosphorylation (Wan et al. 2009; Joglekar et al. 2009; Maresca & Salmon 2009). Importantly, the Aurora B activity is counteracted by phosphatases $PP2A^{B56}$ and PP1. $PP2A^{B56}$, is recruited to unattached kinetochores and dephosphorylates Aurora B substrates upon an increase in the tension, while PP1 is most likely targeted to attached kinetochores and ensures lower level of phosphorylation on bi-oriented kinetochores (Foley & Kapoor 2013; D. Liu et al. 2010). However, the substrate specificity of PP1 and PP2A phosphatases remain unclear. Additional proteins such as Polo-like kinase 1 (Plk1) and SAC proteins (including MPS1, Bub1/BUB3 and BubR1/Bub3) are also known to regulate kinetochore-microtubule attachments (Lampson & Kapoor 2005; Maciejowski et al. 2010; Hewitt et al. 2010; Suijkerbuijk et al. 2012) suggesting that the error correction and spindle assembly checkpoint machinery work closely together (Saurin 2018; Santaguida et al. 2011).

1.7   The outer kinetochore

The outer kinetochore consists of a 10-subunit protein assembly referred as the KMN network, which consists of the Knl1 complex, the Mis12 complex and the Ndc80 complex (Varma & Salmon 2012; Gascoigne & Cheeseman 2013; De Wulf et al. 2003; Desai et al. 2003; Westermann et al. 2003; Cheeseman et al. 2004; X. Liu et al. 2005) (Figure 1-6). Two important functions of the KMN network are:

i)   Binding to either assembling or disassembling microtubules (Maiato et al. 2004; Davis & Wordeman 2007).

ii)   Recruitment of the SAC machinery (Kiyomitsu et al. 2011; Kiyomitsu et al. 2007; Martin-Lluesma et al. 2002; McAinsh et al. 2006; S. A. Miller et al. 2008).

1.7.1   The Ndc80 complex

The Ndc80 complex is highly elongated, with a long axis of ~55-60 nm, and is composed of four subunits, namely Ndc80 (also known as Hec1), Nuf2, Spc24 and Spc25, which consist of large segments of coiled-coil regions flanked by globular domains (Kops et al. 2005; Cheeseman et al. 2004). The Ndc80 complex is the main microtubule receptor of the kinetochores. Depletion of Ndc80 results in the loss of kinetochore-microtubule attachments and chromosome mis-segregation (J. G. DeLuca et al. 2002). The N-terminal regions of both Ndc80 and Nuf2 subunits are responsible for the microtubule binding, while the C-terminal regions of Spc24 and Spc25 are responsible for the kinetochore localization of Ndc80 complex (Ciferri et al. 2008; J. G. DeLuca & Musacchio 2012; Gascoigne & Cheeseman 2011; Joglekar et al. 2010; Tooley & Stukenberg 2011). The first high- resolution structural insights of the Ndc80 complex were obtained by generating engineered versions of the Ndc80 complex (Ndc80^{Bonsai} and Ndc80^{Dwarf}), in which most of the coiled-coil regions were removed (Ciferri et al. 2008). The crystal structure of Ndc80^{Bonsai} reveals that the Calponin Homology domain (CH domain) present at the N-terminus of Ndc80 and Nuf2 mediates the interactions with the plus-end of the microtubules (Ciferri et al. 2008). Mutations within the conserved CH domains of Ndc80 and Nuf2 results in the loss of stable kinetochore-microtubule attachments which arrests the cells in metaphase (Alushin et al. 2010; Tooley et al. 2011). Furthermore, these microtubule attachments are dynamically regulated through Aurora B and Aurora A kinases. Both Aurora B and Aurora A kinases phosphorylate the N-terminal tail of Ndc80 thereby neutralizing its positive charges and decreasing its affinity for microtubules (Cheeseman et al. 2006; Chan et al. 2012; Welburn et al. 2010) (K. F. DeLuca et al. 2018). Other kinases, such as MPS1 and Nek2A, are also implicated in the phosphorylation of Ndc80 complex (Chen et al. 2002; Wei et al. 2011). However, the precise regulation of these kinases remains unknown. High-resolution imaging suggests that the Ndc80 complex binds microtubules in at least two ways. First, the positively charged N-terminal tail of the Ndc80 subunit interact with the negatively charged C-terminal tails of tubulin monomers (E-hooks) (Ciferri et al. 2008; Tooley et al. 2011). Second, the CH domain of Hec1, recognizes both α- and β-tubulin at the intra and inter-tubulin interfaces. This is suggested to promote the oligomerization of the Ndc80 complexes on the

microtubules (Alushin et al. 2010; Tooley & Stukenberg 2011; Powers et al. 2009; Cheeseman et al. 2006).

1.7.2 The Knl1 complex

The Knl1 complex is comprised of two subunits, Knl1 (also called CASC5 or Blinkin in humans, Spc105 in yeast and Spc105R in fly) and Zwint or (Zwint-1). Knl1 is the largest outer kinetochore subunit; it consists of 2,316 residues, which are predicted to be mostly unstructured (Cheeseman et al. 2006). Depletion of knl1 results in the kinetochore-null phenotype, leading to the failure of chromosome segregation (Nekrasov et al. 2003; Kerres et al. 2007). The N-terminal region of Knl1 has been implicated in the microtubule binding as well as recruitment of SAC proteins. The extreme N-terminus of Knl1 recruits PP1 that counteracts the activity of Aurora B kinase, which is involved in the destabilization of kinetochore-microtubule attachments (Welburn et al. 2010; D. Liu et al. 2010; Meadows et al. 2011; Rosenberg et al. 2011). The crystal structure of the C-terminal domain of Knl1 reveals a tandem of RWD domains that mediate interactions with Mis12 complex and Zwint, which is required for its kinetochore targeting (Petrovic et al. 2014; Cheeseman et al. 2014.). Zwint comprises 277 residues and is implicated in the recruitment of ZW10 to kinetochores (Starr et al. 2000). ZW10 is part of the Rod-Zwilch-Zw10 (RZZ) complex, which has been shown to be involved in SAC activation as well as SAC silencing (Karess 2005; Scaërou et al. 1999; Scaërou n.d.). Thus, Knl1 recruits SAC proteins, via its N-terminal region and also via Zwint (Krenn et al. 2012; Primorac et al. 2013; Krenn et al. 2014; Kiyomitsu et al. 2007).

1.7.3 The Mis12 complex

The Mis12 complex is comprised of four subunits, Dsn1, Mis12, Nnf1 and Nsl1 (Petrovic et al. 2010; Maskell et al. 2010; Kline et al. 2006). Mis12 complex is elongated, with ~22 nm rod-like particles when observed under negative stain EM analysis. Crystal structure of the Mis12 complex reveals that all four subunits have similar topologies, and are arranged into two distinct sub-complexes, Mis12: PMF1 and DSN1: NSL1. However, in *Drosophila melanogaster*, the homolog of the Dsn1 subunit is missing, but it contains two paralogues of the KMN subunit Nnf1 (Nnf1a and Nnf1b) (Y. Liu et al. 2016). Specific mutations within the Mis12 complex results in the

defects in the kinetochore assembly and chromosome segregation (Kline et al. 2006). The Mis12 complex directly binds to the inner kinetochore subunits, CENP-C and CENP-T, which is required for its kinetochore localization (Screpanti et al. 2011; Gascoigne et al. 2011; Gascoigne & Cheeseman 2013). Binding of Mis12 complex to CENP-C and CENP-T is competitive, and therefore both CENP-C and CENP-T recruit Mis12 complex individually (Huis in 't Veld et al. 2016). Aurora B mediated phosphorylation of the Dsn1 subunit has been shown to increase the affinity of Mis12 complex with CENP-C (Petrovic et al. 2016). Besides, inner kinetochore subunits, Mis12 complex also binds to the Spc24 and Spc25 subunits of the Ndc80 complex, and to the C-terminus of Knl1, promoting the KMN assembly (Petrovic et al. 2010; Westermann et al. 2003; Cheeseman et al. 2004). Taken together, Mis12 complex acts as a central "hub" within the KMN network, since it binds to inner kinetochore component (CENP-C) as well as the other two KMN components, the Knl1complex and Ndc80 complex (Petrovic et al. 2010; Petrovic et al. 2014; Ciferri et al. 2008; Screpanti et al. 2011; Hornung et al. 2011; Malvezzi et al. 2013).
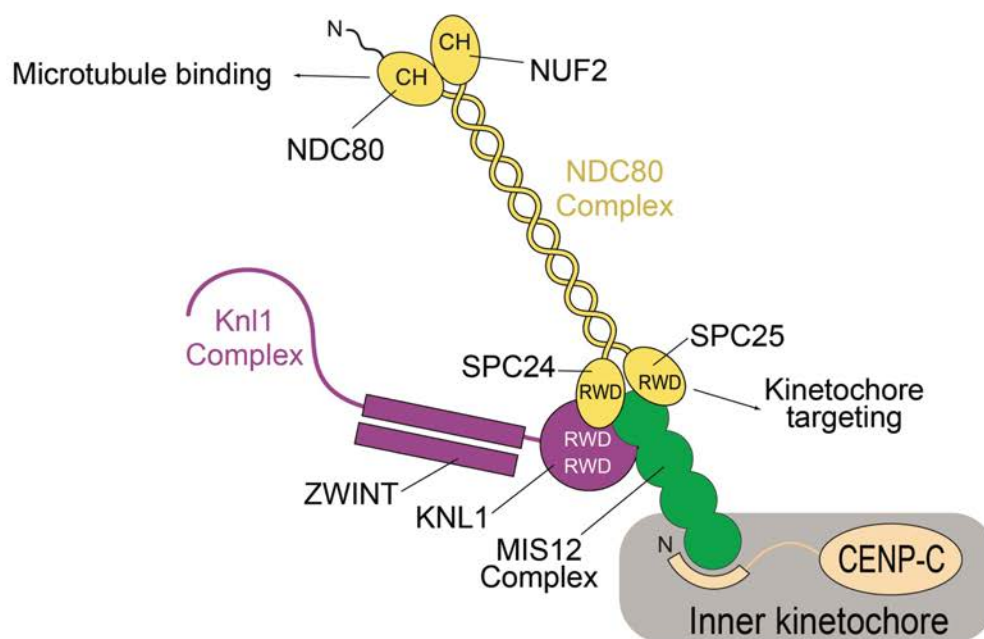
Figure 1-6 Organization of the KMN network

Schematics of the KMN network and its interaction with inner kinetochore (which are discussed in the text). Figure adapted from (Petrovic et al. 2016)

1.8   CCAN organization

The CCAN (which stands for constitutive centromere-associated network) forms the foundation for kinetochore assembly on the centromeric chromatin. As the name suggests, most of the CCAN members are constitutively present at the centromeres during the cell cycle (Perpelescu & Fukagawa 2011; Hori & Fukagawa 2012; Hori, Okada, et al. 2008; Nishino et al. 2012; Saitoh et al. 1992; Pesenti et al. 2016). The CCAN makes up the core of the kinetochore and is responsible for the establishment of the outer kinetochore, which in turn constitutes the microtubule binding interface. Extensive biochemical and genetic studies have identified different CCAN subunits, which are organized into different sub-complexes. CCAN proteins are commonly referred to as CENP (CENtromere Protein) (Okada et al. 2006; Izuta et al. 2006; Foltz et al. 2006). Most of the CCAN sub-complexes can be reconstituted in the absence of other subunits; these include; CENP-C, the CENP-L and CENP-N sub-complex (CENP-LN), the CENP-H, CENP-I, CENP-K, and CENP-M sub-complex (CENP-HIKM), the CENP-T, CENP-W, CENP-S, and CENP-X sub-complex (CENP-TWSX), as well as the CENP-O, CENP-P, CENP-Q, CENP-U, and CENP-R sub-complex (CENP-OPQUR) (Figure 1-7). Each of these sub-complexes will be discussed in the subsequent sections.
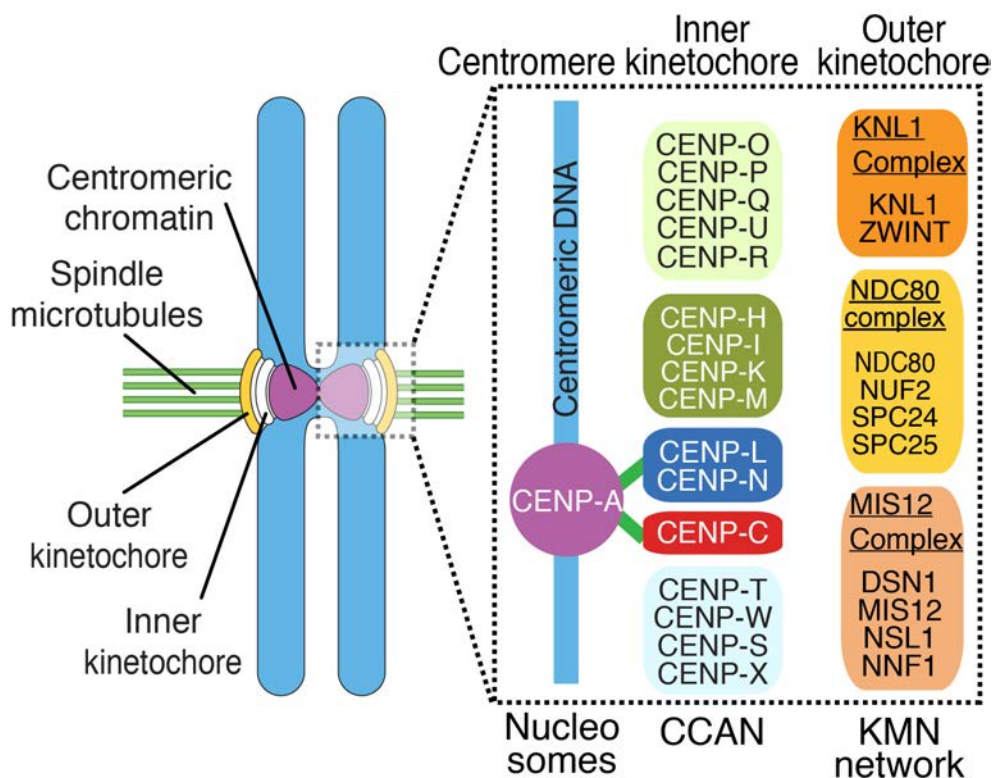
Figure 1-7 Organization of the inner kinetochore

Organization of the CCAN sub-complexes. Inner kinetochore or CCAN is a network of interactions involving 16 subunits that are organized into sub-complexes, as depicted. Within the CCAN network both CENP-LN and CENP-C are known to interact with CENP-A nucleosomes (Shown by green line). Figure modified from (Weir et al. 2016).

1.8.1 CENP-C

CENP-C was originally identified as an antigen from patients suffering from CREST syndrome (Earnshaw & Rothfield 1985). It was the first protein found to localize to the inner kinetochore region (Saitoh et al. 1992). CENP-C homologs are found in several organisms, including *S. cerevisiae* (Mif2) and *S. pombe* (Cnp3) (Saitoh et al. 1992; M. T. Brown et al. 1993; Tomkiel et al. 1994; Meluh & Koshland 1995; Fukagawa & W. R. Brown 1997) (Figure 1-8). Previous studies have shown that the depletion of CENP-C results in the chromosome misalignments and kinetochore assembly defects (Fukagawa & W. R. Brown 1997; Fukagawa et al. 1999; Kwon et al. 2007). Most of the studies carried out on CENP-C suggest that the functionally important domains of CENP-C are highly conserved, despite their low sequence similarity (Przewloka et al. 2011; Klare et al. 2015; Kato et al. 2013; Pesenti et al. 2016). Human CENP-C is a 943-amino acid protein that is predicted to be completely disordered and positively charged.

The N-terminal region of CENP-C consits of a conserved domain that is required to bind the Mis12 complex (Screpanti et al. 2011; Petrovic et al. 2016). Thus, CENP-C connects the inner kinetochore with outer kinetochore by binding to the Mis12 complex (Musacchio & Desai 2017). In some organisms, including *D. melanogaster* and *C. elegans*, CENP-C is the only identified CCAN subunit, pointing to its important role in bridging function (Y. Liu et al. 2016; Meraldi et al. 2006; Drinnenberg et al. 2014).

The middle region of CENP-C also contains a conserved region that is rich in PEST (proline, glutamate, serine, threonine) which is implicated in the binding to CENP-HIKM complex (Nagpal et al. 2015; Klare et al. 2015). Detailed biochemical studies on the interaction between CENP-C and CENP-HIKM complex have resulted in the identification of the residues (Leu^{265}, Phe^{266}, Leu^{267} and Tryp^{317}) involved in this specific interaction (Klare et al. 2015). It has also been demonstrated that CENP-C recruits CENP-HIKM complex by binding directly to CENP-HK within the CENP-HIKM complex (Klare et al. 2015).

CENP-C specifically recognizes centromeric nucleosome CENP-A via its central domain and its C-terminal CENP-C motif (Pesenti et al. 2016; Kato et al. 2013; Carroll et al. 2010). A recent structural analysis of rat CENP-C bound to CENP-A nucleosome reveals that the central domain of CENP-C binds to a hydrophobic region in the C-terminal tail of CENP-A, as well as to the acidic patch of histone H2A and H2B. It has been proposed that the conserved CENP-C motif uses a similar mechanism in order to bind the CENP-A nucleosome (Kato et al. 2013). Previous studies on CENP-C have proposed that the binding of CENP-C to CENP-A is important for the kinetochore targeting of CENP-C (Sugimoto et al. 1994; Yang et al. 1996). Furthermore, *in vitro* FRET (fluorescence resonance energy transfer) assay demonstrated that CENP-C rigidifies both internal and surface nucleosome structure, which affects the overall shape and stability of the CENP-A nucleosome (Falk et al. 2015). The C-terminal region of CENP-C consists of a Cupin domain that is involved in the dimerization (Sugimoto et al. 1994).

CENP-C depletion in HeLa cells leads to a near complete loss of CENP-HIKM, CENP-LN and CENP-TWSX, suggesting that CENP-C plays an important role in recruiting the inner kinetochore proteins (Klare et al. 2015; Nagpal et al. 2015; Weir et al. 2016; Milks et al. 2009). In chicken DT40 cells, CENP-H localization is not completely abolished upon CENP-C depletion, suggesting that the recruitment of CCAN sub-complexes might work in different ways in different species. In addition to providing stability to the CENP-A nucleosome, CENP-C is also involved in the replenishment of CENP-A at the centromeres (L. Y. Guo et al. 2017; Cao et al. 2018). CENP-C executes this function by recruiting Mis18BP1, a component of the Mis18 complex as already discussed before (Erhardt et al. 2008).

1.8.2 The CENP-LN complex

Orthologs of CENP-N and CENP-L have been identified in both *S.cerevisiae* (named Chl4 (CENP-N) and Iml3 (CENP-L)) and *S. pombe* (Meraldi et al. 2006) (Figure 1-8). Interestingly, in some insect cell lineages harboring holocentric chromosomes, CENP-LN is present, while both CENP-A and CENP-C are missing (Drinnenberg et al. 2014). CENP-L and CENP-N were identified as CENP-A associated proteins in human cells (Obuse et al. 2004; Foltz et al. 2006), while in chicken cells, CENP-L and CENP-N were identified as CENP-H and I interacting proteins (Okada et al. 2006). Knockout of

CENP-N resulted in severely disrupted chromosome alignment and mitotic arrest, suggesting an important role in mitosis (McKinley et al. 2015).

CENP-N was the first protein known to bind CENP-A nucleosomes specifically. The N-terminal region of CENP-N is sufficient for this interaction (Carroll et al. 2009). Binding of CENP-N to CENP-A nucleosomes is DNA sequence independent as CENP-N bound efficiently even when a different, non centromeric DNA (601-Widom) is wrapped around the histone octamer (Carroll et al. 2009; Carroll et al. 2010; Fang et al. 2015). The CATD region within the CENP-A has been found to be sufficient to bind to CENP-N (Carroll et al. 2009). Moreover, recent HXMS (hydrogen/deuterium exchange-mass spectrometry) experiments on CENP-N bound by CENP-A nucleosomes has revealed that the HX protection is conferred to the CATD region of CENP-A, corroborating previous findings (L. Y. Guo et al. 2017). In addition, ectopic targeting of the CATD region within the CENP-A is sufficient to recruit CENP-N to the LacO/LacI arrays (Fang et al. 2015). However, how CENP-N has a preferential selectivity for CATD of CENP-A nucleosomes had remained elusive. This will be discussed in detail in the following sections.

It has been found that, in cells depleted of CENP-A and then released into S-phase, new CENP-N fails to load on to centromeres, while pre-deposited CENP-N remained stably attached to kinetochores. This suggests that the loading of CENP-N depends on CENP-A (Hoffmann et al. 2016). Besides CENP-A, the loading of CENP-N seems to rely on higher order chromatin organization at the centromeres, whereby the chromatin undergoes a structural transition from closed or compacted chromatin in the G1-phase to an open chromatin in the S-phase. Moreover, CENP-N cannot bind to compacted chromatin, but it does bind to open chromatin (Fang et al. 2015). Collectively, these observations suggest that CENP-N is loaded during the S-phase where by the chromatin is in open confirmation and dissociates during the G2-phase (Fang et al. 2015; Hellwig et al. 2011). It was proposed that CENP-N is the only CCAN protein that undergoes a rapid turnover during the G1-phase and continues until the S-phase (Hellwig et al. 2011). Kinetochore targeting experiments lacking the C-terminal region of CENP-N resulted in the severe impairment of its localization, which suggests that the localization of CENP-N might depend on its interactions with other CCAN proteins (Carroll et al. 2009).

Sequence alignment of Chl4 and CENP-N, Iml3 and CENP-L shows some strong patches of similarity. Iml3, is structurally related to a bacterial recombination

associated protein, RdgC, and binds to the C-terminal domain of Chl4^{CENP-N} (Hinshaw & Harrison 2013). Crystal structure of the C-terminal region of Chl4^{CENP-N} in complex with Iml3^{CENP-L} reveals that the residues involved in the heterodimerization are conserved, suggesting a similar structural organization in the human CENP-N:CENP-L complex (Hinshaw & Harrison 2013). Besides hetero-dimerization, Iml3 can also homo-dimerize engaging the same surface, therefore heterodimer and homodimer formations are mutually exclusive (Q. Guo et al. 2013; Hinshaw & Harrison 2013). In addition to binding to Iml3, Chl4 is proposed to be involved in the binding to CENP-C ortholog Mif2 (Hinshaw & Harrison 2013). Similarly, CENP-LN complex has also been implicated in binding to other CCAN components, such as CENP-C and CENP-HIKM complex, in addition to binding to CENP-A nucleosomes. However, the importance of these specific interactions remains obscure (Weir et al. 2016; McKinley et al. 2015; Nagpal et al. 2015). Depletion of CENP-N results in the failure to load nascent CENP-A, as well as loss of other CCAN sub-complexes, which results in defects in kinetochore assembly (Carroll et al. 2009; McKinley et al. 2015; Foltz et al. 2006). Although it is well established that the CENP-LN complex binds CENP-A nucleosomes and is involved in the interactions with other CCAN members, many questions remain to be addressed. For instance, how does CENP-LN complex selectively recognize CENP-A nucleosome? Does CENP-N require additional CCAN components for its kinetochore localization? These questions are discussed in detail in the Results section.

1.8.3 The CENP-HIKM complex

CENP-H has been identified as a coiled-coil protein, that localizes to the kinetochores in mouse cells (Sugata et al. 1999). Further studies have identified human and chicken CENP-H homologs (Sugata et al. 2000; Fukagawa et al. 2001). Subsequently, CENP-I was identified as a CCAN subunit based on its co-localization with CENP-A and CENP-H. The interaction between CENP-H and CENP-I was first demonstrated using a yeast two–hybrid assay (Nishihashi et al. 2002). Immunoprecipitation experiments of CENP-H and CENP-I in both human and chicken cells have led to the identification of two more CCAN subunits; CENP-K and CENP-M (Okada et al. 2006). Similar, to other CCAN proteins, CENP-H, I and K orthologs have been found in both *S. cerevisiae* and *S. pombe* (Nishihashi et al. 2002; Schleiffer et al. 2012) (Figure 1-8). CENP-M was initially identified as a highly expressed protein in proliferating cells and tumors, and it

was therefore termed as proliferation-associated nuclear element-1 (PANE-1) (Renou et al. 2003; Bierie et al. 2004). Previous studies on the CENP-H, I, K and M suggests that these CCAN subunits are required for chromosome alignment and segregation (Okada et al. 2006; Nishihashi et al. 2002; Basilico et al. 2014). A recent structural and biochemical analysis of CENP-M reveals that it is evolutionarily and structurally related to small GTPases, but that it is incapable of GTP-binding and is hence referred to as pseudo GTPase (Basilico et al. 2014). Both *in vitro* and *in vivo* studies demonstrate that the CENP-H, I, K and M form a stable complex, stabilized by CENP-M (Basilico et al. 2014; McKinley et al. 2015). Depletion of each individual subunit within the CENP-HIKM complex leads to the depletion of other components, suggesting that their kinetochore localization is co-dependent (Okada et al. 2006; Basilico et al. 2014). In addition to interacting with CENP-C and CENP-LN complex, CENP-HIKM complex also interacts with another CCAN sub-complex, CENP-TWSX (Basilico et al. 2014; Weir et al. 2016). However, the molecular details of this interaction remain to be understood. Interestingly, a point mutant of CENP-M that affects the CENP-M/CENP-I interaction leads to the loss of CENP-I as well as CENP-TW, suggesting that the kinetochore localization of CENP-TW is dependent on the CENP-HIKM complex (Basilico et al. 2014). In addition to CENP-TW, the CENP-HIKM complex is also involved in the kinetochore recruitment of CENP-OPQUR complex (Okada et al. 2006; Izuta et al. 2006; Pesenti et al. 2018).

1.8.4 The CENP-OPQUR complex

Immunoprecipitation experiments with CENP-H/I and CENP-A chromatin have led to the identification of the additional CCAN components; CENP-O, P, Q and U (Okada et al. 2006; Foltz et al. 2006). CENP-U was initially identified as a constitutive centromeric component and named CENP-50 (Minoshima et al. 2005). Co-expression of these subunits results in the formation of a stable complex CENP-OPQU (Hori, Okada, et al. 2008). Further studies have led to the identification of CENP-R as an integral component of this complex (Meraldi et al. 2006; McClelland et al. 2007). The depletion of CENP-R has no effect on the kinetochore localization of CENP-OPQU, which suggests that CENP-R might act as a downstream component (Okada et al. 2006; Hori, Okada, et al. 2008). CENP-OPQUR occupies an outermost position of the inner kinetochore as its kinetochore recruitment requires the binding to CENP-LN and CENP-HIKM sub-complexes. Conversely, CENP-OPQUR is dispensable for the

recruitment of CENP-LN or CENP-HIKM (Pesenti et al. 2018; McKinley et al. 2015; Foltz et al. 2006). Biochemical studies on the CENP-OPQUR complex have demonstrated that the CENP-OP sub-complex binds to CENP-HIKMLN, while CENP-QU binds microtubules (Pesenti et al. 2018; Amaro et al. 2010). The EM analysis of CENP-HIKMLNOPQUR complex provides the first comprehensive analysis of the CCAN network. One of the important conclusions of the EM analysis was the compact and globular nature of the CCAN complex, in contrast to the fibrous organization of the KMN network (Pesenti et al. 2018). Disruption of any of these components causes mitotic delay and defects in the chromosome alignment along the metaphase plate and deficiency in recovering from spindle damages (Hori, Okada, et al. 2008). Cells depleted of CENP-U partially align their chromosomes, but fail to enter into anaphase, which suggests that CENP-U plays an important role in mitosis. Depletion of CENP-U in mouse embryonic stem cells results in the missegregation of chromosomes and cell death (Kagawa et al. 2014). The CENP-OPQU complex is conserved in budding yeast, where is referred to as the COMA (Ctf19-Okp1-Mcm21-Ame1) complex (De Wulf et al. 2003). Structural characterization of the Ctf19:Mcm21 complex reveals that the CENP-OP are structural paralogs each comprising RWD domains (RING finger, WD repeat, DEAD helicases) (Westermann & Schleiffer 2013). Interestingly, these RWD domains are also found in the C-terminus of Knl1, Spc24/25 and the SAC protein Mad1, which suggests that the RWD domains are recurrent in kinetochore proteins. Collectively, CENP-OPQUR contributes to chromosome alignment through a direct interaction with microtubules.

1.8.5 The CENP-TWSX complex

CENP-T was originally identified as a CENP-A interacting partner, while CENP-S was found to be an interacting partner of CENP-M and CENP-U (Izuta et al. 2006; Foltz et al. 2006). Immunoprecipitation experiments with CENP-T and CENP-S have led to the identification of two additional subunits, CENP-W and CENP-X (Hori, Amano, et al. 2008; Amano et al. 2009). CENP-TW, together with CENP-SX, forms an hetero-tetramer, with a nucleosome-like structure (Nishino et al. 2012). CENP-T, -W, -S and –X consist of a histone fold domain (HFD), which distinguishes them from other CCAN members (Hori, Amano, et al. 2008; Nishino et al. 2012). *In vitro* studies have indicated that the CENP-TW heterodimer binds to centromeric DNA via histone fold domain (Takeuchi et al. 2014). The N-terminal region of CENP-T directly binds to the Ndc80

complex, upon phosphorylation by the CDK1:Cyclin B complex (Huis in 't Veld et al. 2016; Gascoigne & Cheeseman 2013; Gascoigne et al. 2011). Since CENP-T binds to both centromeric DNA and the Ndc80 complex, which binds microtubules, CENP-T functions as a bridge between chromatin and microtubules (Nishino et al. 2012; Gascoigne et al. 2011; Huis in 't Veld et al. 2016).

Besides binding to DNA, the CENP-TWSX complex also binds to the CENP-HIKM complex, which is required for its centromeric localization (Basilico et al. 2014). Cells depleted of CENP-S or X display mild mitotic defects as observed in knockout cells of CENP-O complex proteins (Hori, Amano, et al. 2008; Nishino et al. 2012). Studies on the yeast CENP-T ortholog, Cnn1 have revealed a conserved function (Bock et al. 2012; Schleiffer et al. 2012). In chicken DT40 cells, ectopic localization of CENP-T at a non-centromeric locus recruit the Ndc80 complex while no other CCAN proteins were detected, suggesting that CENP-T recruits NDC80 independently of CENP-C (Hori et al. 2013; Volkov et al. 2018). Taken together, these results suggest the existence of two independent pathways to recruit the Ndc80 complex to kinetochores (Hori et al. 2013; Huis in 't Veld et al. 2016; Musacchio & Desai 2017; Gascoigne et al. 2011).

1.8.6 Biochemical reconstitution of the CCAN sub complexes

Until now, a comprehensive understanding of the organization of the 16 subunits of the CCAN has been missing. Recently, all sub-complexes of the vertebrate CCAN and the KMN network have been reconstituted using recombinant proteins (Weir et al. 2016; Pesenti et al. 2018). Similar studies on the kinetochore proteins purified from *S. cerevisiae* have also been performed (Akiyoshi et al. 2010; M. P. Miller et al. 2016). Combined with crosslink- mass spectrometry analysis, the complicated meshwork of interactions of the CCAN and the KMN network, has now been revealed (Weir et al. 2016). Recent work from our laboratory has led to the near complete reconstitution of the CCAN network. This includes CENP-C, and the CENP-LN, CENP-HIKM, and CENP-OPQUR sub-complexes (Weir et al. 2016; Pesenti et al. 2018). The binding of each individual CCAN sub-complex has been shown to increase the overall stability and selectivity of the reconstituted CCAN complex to CENP-A nucleosome, a manifestation of cooperativity (Weir et al. 2016). The reconstituted CCAN is bound to the CENP-A nucleosome through CENP-C and CENP-N. Analytical ultra-centrifugation (AUC) experiments on the reconstituted CCAN reveals that there are

two copies of CCAN sub-complexes bound by one CENP-A nucleosome (Weir et al. 2016). As CENP-C has the KMN binding site, the reconstituted CCAN can directly bind to the KMN network (Weir et al. 2016; Screpanti et al. 2011). This KMN network has also been shown to bind microtubules *in vitro* (Weir et al. 2016; Musacchio & Desai 2017).

Although, most of the CCAN sub-complexes have been reconstituted *in vitro*, the structural basis for the recognition of the centromeric nucleosomes by CENP-N has remained elusive. This study focuses on how CENP-N recognizes CENP-A and engages in the interaction with the other CCAN members within the CCAN network.
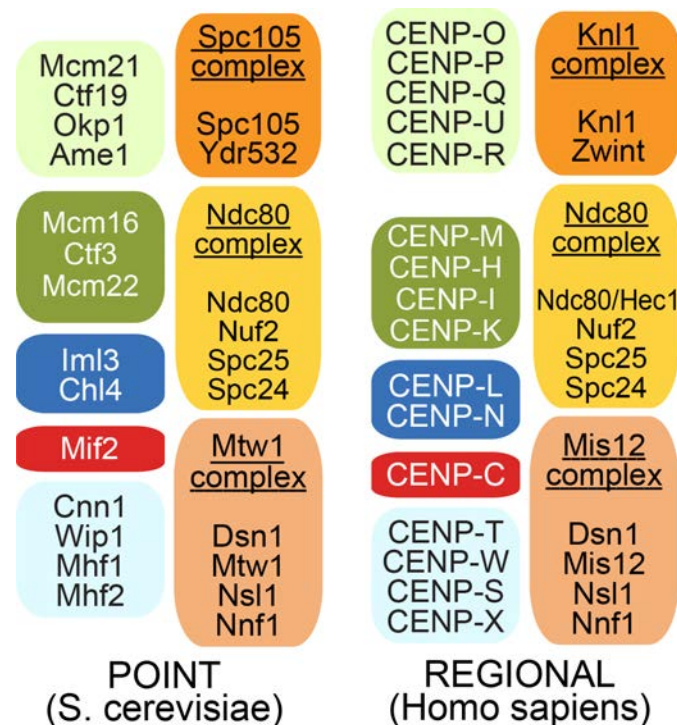
Figure 1-8 Comparison of CCAN sub-complexes in yeast and humans

The similar subunit composition of yeast and humans suggests conservation of a structural module that is present in only one copy at point centromeres and multiple copies at regional centromeres. Figure modified from (Weir et al. 2016)**.**

1.9   Objectives of the study

The CCAN acts as a bridge between the centromeric chromatin and the microtubule binding KMN network (Weir et al. 2016; Musacchio & Desai 2017). It is known that the primary step in the kinetochore assembly is the specific recognition of the centromeric CENP-A nucleosomes by CENP-C and the CENP-LN complex (Carroll et al. 2009; Carroll et al. 2010). Henceforth, understanding the molecular mechanism by which the CENP-LN complex and CENP-C recognize centromeric chromatin is of great importance to unravel the events of kinetochore assembly. The molecular basis of CENP-C binding to CENP-A nucleosomes has already been described in previous studies (Kato et al. 2013). Although CENP-LN interacts specifically with the CENP-A nucleosomes and other CCAN subunits such as CENP-C and the CENP-HIKM complex, the structural and molecular basis for these interactions remains poorly understood (Weir et al. 2016; McKinley et al. 2015). Moreover, siRNA depletion of CENP-N or CENP-L resulted in the loss of other CCAN components, including CENP-H, CENP-I, CENP-K and CENP-C, leading to kinetochore assembly defects. Importantly, the depletion of other CCAN subunits also resulted in a similar phenotype, suggesting a co-dependency of the CCAN members. Thus, a specific function for CENP-N or CENP-L cannot be inferred from siRNA-based studies. Therefore, *in vitro* reconstitution and structural studies have become crucial to delineate specific interactions within the CCAN network (Weir et al. 2016). That is why the main aim of my PhD project was to perform both biochemical and structural studies coupled with *in vivo* analysis, to understand the role of CENP-LN complex in kinetochore assembly. In order to achieve this aim, I have addressed the following questions:

I.   How does CENP-N specifically recognize CENP-A nucleosomes over canonical H3 nucleosomes?

II. Does CENP-LN complex interact with other CCAN components? If so, can we delineate these specific interaction interfaces?

III. What is the mechanism by which the CENP-LN complex is recruited to the kinetochores?

2 Materials and Methods

2.1 Materials

2.1.1 Chemicals and Consumables

All reagents used in this study are listed below:

Table 2-1 Reagents used in this study

Reagents/Enzymes	Supplier
Acetic Acid	Sigma-Aldrich, USA
Acrylamide (30%, Mix 37.5:1)	AppliChem GmbH, Germany
Adenosine-5'-triphosphate (ATP)	Sigma-Aldrich, USA
Agarose	Invitrogen, USA
Ammoniumperoxosulphate (APS)	Serva Electrophoresis, Germany
β−Mercaptoethanol	Serva Electrophoresis, Germany
Bovine Serum Albumin (BSA)	Carl Roth Chemie, Germany
Bradford Protein Assay	Bio-Rad Laboratories, USA
Bromophenol blue	Sigma-Aldrich, USA
Coomassie G250/R250	Serva Electrophoresis, Germany
Dithioerythritol (DTE)	Serva Electrophoresis, Germany
DNA polymerase Q5 (Master mix)	New England Biolabs
Taq DNA polymerase	Invitrogen, USA
DNA ladder (1Kb)	Fermentas, Germany
Ethanol	Thermoscientific, USA
Ethylenediaminetetraacetic acid (EDTA)	Gerbu Biotechnik,GmbH, Germany
FuGENE transfection Reagent	Promega Corp, USA

Glycerol	Gerbu Biotechnik GmbH, Germany
Hydrochloric acid (HCl)	AppliChem GmbH, Germany
4-(2-Hydroxyethyl)-piperazine-1-ethanesulfonic acid (HEPES)	Sigma-Aldrich, USA
Imidazole	Merck, Germany
Isopropyl-β-D-thiogalactopyranoside (IPTG)	Carl Roth Chemie GmbH, Germany
Magnesium Chloride	J.T.Baker Chemicals, USA
Methanol	Sigma-Aldrich, USA
Midori Green	Nippon Genetics, Germany
2-(N-morpholine)-ethane sulfonic acid (MES)	Sigma-Aldrich, USA
Phenylmethyl sulfonic acid (PMSF)	Serva Electrophoresis GmbH, Germany
Polyethylene glycol (PEG) 3350	Sigma-Aldrich, USA
Potassium Chloride (KCl)	Sigma-Aldrich, USA
Proteases (TEV, preScission)	Dortmund Protein Facility (DPF), MPI, Germany
Protease inhibitor	Serva Electrophoresis GmbH, Germany
Restriction Endonucleases	New England Bio labs, USA
Crystallization Suites (JCSG Core-I-IV, PACT, PEGs I and II)	Qiagen, Germany
Grids (1.2/1.3, 2/1, 2/2, carbon support)	Quantifoil, Germany
Sodium Acetate (NaAc)	Sigma-Aldrich, USA

Sodium Chloride (NaCl)	Sigma-Aldrich, USA
Sodium dodecyl sulphate (SDS)	Carl Roth Chemie GmbH, Germany
Sodium hydroxide (NaOH)	Waldeck GmbH & Co., Germany
Tris-(hydroxyethyl)-amino ethane (Tris)	Sigma-Aldrich, USA
Tris-(2-Carboxyethyl)-Phosphine (TCEP)	Sigma-Aldrich, USA
Triton-X-100	Serva Electrophoresis GmbH, Germany
5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal)	Thermo Fisher Scientific, USA
Uranyl formate	Sigma-Aldrich, USA

2.1.2 Culture Media

The following culture medias were used for the expression of proteins:

i) LB Media: 10 g Bacto-tryptone, 5 g Yeast Extract, 10 g NaCl dissolved in 1 l with H_2O

ii) TB Media: 12 g Bacto-tryptone, 24 g Yeast extract, 4 ml glycerol, 100 ml potassium phosphate monobasic filled upto 1 l with H_2O

iii) 2XTY Media: 16g Bacto-tryptone, 10 g Yeast extract, 5 g NaCl dissolved in 1 l with H_2O.

The above described culture medias were obtained from in-house facility at MPI, Dortmund. The culture media used for the expansion or expression of insect cells is sf-900 II SFM obtained from Thermo Fisher scientific.

2.1.3 Antibodies

The following primary antibodies (Table 2-2) and secondary antibodies (Table-2-3) were used in this study

Table 2-2 Primary antibodies

Antigen	Origin	Dilution	Supplier/Reference
CENP-C	Rabbit	1:1000	In-house made
CREST	Human	1:100	Antibodies Inc.
CENP-A	Mouse	1:500	Gene tex
CENP-HK	Rabbit	1:1000	In-house made
GFP- Boost		1:500	Chromotek
RFP-Boost		1:500	Chromotek

Table 2-3 Secondary antibodies

Antigen	Origin	Dilution	Fluorophore	Supplier
IgG α-Human	Donkey	1:10000	Alexa 405	Jackson immuno research labs
IgG α-Rabbit	Donkey	1:10000	Alexa 488	Jackson immuno research labs
IgG α-Mouse	Donkey	1:10000	Alexa 647	Invitrogen

2.1.4 Kits used in this study

All the kits used in this study are listed below

Table 2-4 Kits

Kit	Purpose	Supplier
Nucleo Spin Plasmid (NoLid)	Isolation of Plasmid DNA from small bacterial cultures.	Machery Nagel
QIAquick Gel extraction	Purification of DNA from gels	Qiagen
QIAquick PCR purification Kit	Purification of DNA from PCR	Qiagen

2.1.5 Antibiotics

The following antibiotics were used for bacteria (Table 2-5) and mammalian cells (Table 2-6)

Table 2-5 Antibiotics for bacteria

Antibiotic	Concentration
Ampicillin	100 µg/ml
Kanamycin	50 µg/ml
Gentamycin	10 µg/ml
Tetracyclin	7 µg/ml

Table 2-6 Antibiotics for mammalian cells

Antibiotic	Concentration
Penicillin	100 U/ml
Streptomycin	0.1 mg/ml
Blasticidin	5 µg/ml
Hygromycin	250 µg/ml

2.1.6 Strains used for transformation

The following strains of chemically competent cells were used in this study

Table 2-7 Competent cells used for transformation

Bacteria	Strain	Supplier
E. coli	OmniMax	Max-Planck Institute, DPF, Dortmund, Germany
E. coli	C41	Max-Planck Institute, DPF, Dortmund, Germany
E. coli	BL21	Max-Planck Institute, DPF, Dortmund, Germany
E. coli	Max Efficiency DH10BAC chemically competent cells (EMBACY)	Thermo Fisher Scientific (GIBCO), Waltham, USA

2.1.7 Instruments used in this study

The following instruments were used during the course of this study

Table 2-8 List of instruments used in this study

Application	Instrument	Company
Agarose gel electrophoresis	Agarose gel electrophoresis system	Carl Roth Chemie GmbH, Germany
Cell counting	Scepter	Merck Milllipore, Germany
	Neubauer counting chamber	Marienfeld-Superior, Germany
Cell lysis	Sonifier cell Disruptor	Branson Ultrasonics, USA
Centrifuges	Sorvall Centrifuge	Thermo Fisher scientific, USA
	Microcentrifuge	Eppendorf AG, Germany
	Avanti centrifuge	Beckman Coulter, USA
Concentrators	Amicon Ultra 3,10,30,50 kDa	Merck KGaA, Germany
DNA	Trans illuminator	Thermo Fisher scientific,USA
Gel Doc^{TM} XR	Gel documentation system	Bio-Rad
Chromatography	AKTA Purifier, micro	GE Healthcare, UK
Columns	GSTTrap	GE Healthcare, UK
	HisTrap	GE Healthcare, UK
	Heparin	GE Healthcare, UK
	Resource S	GE Healthcare, UK
	Superdex (S200 10/300, S75 10/100, S200 16/600)	GE Healthcare, UK
	Superose (S6 10/300, S6 16/600)	GE Healthcare, UK
Glass beads	Glass beads	Carl Roth GmbH, Germany
Incubator	Multritron	Infors AG, Switzerland
Microscopes	CX41 with Olympus camera	olympus corporation, Japan

	Leica with power supply K; 1500 LCD	Leica Microsystems, Germany
PCR	Professional Trio Thermocycler	Analytik, Germany
Pipettes	Eppendorf pipettes	Eppendorf AG, Germany
pH-meter	Basic Meter	Sartorius, Germany
Power Supply Unit	Power Pac	Bio-Rad laboratories, USA
ITC	ITC200 micro calorimeter	GE Healthcare, UK
Crystallography		
Imaging system	Rock Imager	Formulatrix Inc., USA
Crystallization Robot	Mosquito Nano dispenser	TTP Lab Tech Ltd., UK
Anode	MICROSTAR	Bruker AXS Corp, USA
Detector	Mar Image Plate detecter	Marresearch, Germany
Screening	96-well Microplate V-Shape	Grenier Bio-One GmbH, Germany
	Cryschem plate 24 well, sitting drop	Hampton Research, USA

2.1.8 Software's used in this study

The following software's have been used during the course of this study

Table 2-9 List of software's used in this study

Software	Version	Supplier
Illustrator CC 2017	21.0.1	Adobe
ApE- A Plasmid Editor	2.0.45	M.Wayne Davis
Image J	1.46r	National Institutes of Health
Excel	15.37	Microsoft
Word	15.37	Microsoft
Imaris	7.3.4	Bitplane
Eman2	2.0.7	Baylor college of medicine, USA

Sphire	1.2	MPI, Dortmund, Germany
Pymol	1.4	Scrodinger LLC, UK
SedFit	14.4d	National Institutes of Health, USA
Unicorn control software	5.31	GE Healthcare, UK
CCP4MG	2.9.0	Diamond light source, UK
Origin	5.0	Microcal, UK

2.1.9 List of Protein constructs used in this study

The following constructs have been used in this study

Table 2-10: List of constructs generated/used in this study

Encoded Protein	Construct	Source
CENP-LN	pFL-CENP-L-CENP-N	Musacchio lab
CENP-LN^{230-C}	pFL-CENP-L:CENP-N^{230-C}	This study
CENP-L	pFG-CENP-L	This study
CENP-LN^{R11A}	pFL-CENP-L-CENP-N	This study
CENP-N$^{1-212}$	pST50TR-DHFRHIS	Musacchio lab
CENP-N$^{1-212}$-GST	pST50TR-DHFRHIS	This study
CENP-N$^{1-235}$	pST50TR-DHFRHIS	Musacchio lab
CENP-N$^{1-212-E7A}$	pST50TR-DHFRHIS	This study
CENP-N$^{1-212-R11A}$	pST50TR-DHFRHIS	This study
CENP-N$^{1-212-K15A}$	pST50TR-DHFRHIS	This study
CENP-N$^{1-212-K45A}$	pST50TR-DHFRHIS	This study
CENP-N$^{1-212-Y147A}$	pST50TR-DHFRHIS	This study
CENP-N$^{1-212-R196A}$	pST50TR-DHFRHIS	This study
CENP-N$^{1-212-K15+Y147A}$	pST50TR-DHFRHIS	This study
CENP-N$^{1-212-K45+Y147A}$	pST50TR-DHFRHIS	This study
CENP-N$^{1-212}$-MBP	pST50TR-DHFRHIS	This study
CENP-C$^{2-545}$	pGEX-6P-2rbs	Musacchio lab
CENP-C$^{189-545}$	pGEX-6P-2rbs	Musacchio lab

CENP-C$^{225-364}$	pGEX-6P-2rbs	This study
CENP-C$^{2-545-E302, F303, I304, I305, D305A}$	pGEX-6P-2rbs	This study
CENP-C$^{225-364-E302, F303, I304, I305, D305A}$	pGEX-6P-2rbs	This study
CENP-C$^{1-71}$	pGEX-6P-2rbs	Musacchio lab
CENP-HIKM	pFL	Musacchio lab
CENP-HI^{57-C}KM	pFL	Musacchio lab
CENP-HI$^{57-702}$KM	pFL	This Study
CENP-HK	pFL	Musacchio lab
CENP-IM	pFL	Musacchio lab
CENP-N	pCDNA5-mCherry-IRES	This study
CENP-N^{E7A}	pCDNA5-mCherry-IRES	This study
CENP-N^{R11A}	pCDNA5-mCherry-IRES	This study
CENP-N^{K15A}	pCDNA5-mCherry-IRES	This study
CENP-N^{K45A}	pCDNA5-mCherry-IRES	This study
CENP-N^{Y147A}	pCDNA5-mCherry-IRES	This study
CENP-N^{R196A}	pCDNA5-mCherry-IRES	This study
CENP-N$^{K15+Y147A}$	pCDNA5-mCherry-IRES	This study
CENP-N$^{K45+Y147A}$	pCDNA5-mCherry-IRES	This study
CENP-C	pCDNA5-eGFP-IRES	Musacchio lab
CENP-C$^{E302, F303, I304, I305, D305A}$	pCDNA5-eGFP-IRES	This study

2.1.10 List of online software's used in this study

The following online tools were used in this study

Table 2-11 List of online tools used in this study

Online Tools	Application
ConSurf	Conservation mapping
PsiPred	Secondary structure prediction
PHD	Secondary structure prediction

ESPript	Multiple sequence alignment
T-coffee	Multiple sequence alignment
ClustalW	Multiple sequence alignment
Uniprot	Protein sequences
Sphire	EM data analysis
EMAN2	EM data analysis

2.2   Methods

2.2.1   Restriction based cloning

All plasmids used in this study were generated by restriction enzyme based cloning. Restriction enzyme based cloning involves the amplification of the gene of interest with the 5' and 3' restriction sites. To amplify the fragments consisting of gene of interest Q5 2X master mix was used along with the template DNA and primers as described in the following table 2-12. PCR conditions used to amplify the respective DNA fragments were described in table 2-13. Both insert (comprising of gene of interest) and vector were cut with restriction enzymes in order to generate complimentary ends. Restriction digestions were performed with restriction enzymes from New England Bio labs (NEB) for either ~3 - 5 h or overnight at 37 $^{\circ}$C. After the restriction digestion, both the insert and vector were run on an agarose gels supplemented with midori green. Agarose gels were prepared by dissolving either 0.8% or 1% agarose in 1X TAE (Tris-Acetate-EDTA) buffer along with 1:2000 midori green DNA stain. After examining the DNA insert and vector, the bands were excised and purified using Gel extraction kit. For ligation of the insert and vector, a rapid DNA ligation kit (Thermo scientific) was used. Typically, 10-100 ng of linearized vector was combined with a 3-6-fold molar excess of insert in a 20 µl reaction and incubated overnight. This ligation mix was subsequently transformed into competent OMNIMAX cells by heat-shock method and was plated on a LB-agar plates supplemented with appropriate antibiotics. These plates were incubated overnight at 37 $^{\circ}$C and positive colonies were picked and verified by DNA sequencing.

Table 2-12 Standard PCR reaction

Reaction Mix	Volume for 25 µl PCR reaction
DNA template	~30-60 ng
10 µM Forward primer	1.25 µl
10 µM Reverse primer	1.25 µl
Q5 2X Master mix	12.5 µl
ddH$_2$O	X µl

Table 2-13 Standard program used for PCR amplification

Step	Temperature	Time	Number of cycles
Denaturation	98 °C	45 s	
	98 °C	10 s	30X
Annealing	52 °C – 65 °C	30 s	
Elongation	72 °C	3 – 5 min	
	72 °C	3 – 5 min	
Pause	4 °C	~	

2.2.2 Plasmids for bacterial expressions

2.2.2.1 CENP-N

To generate CENP-N plasmids, a cDNA sequence encoding the human CENP-N isoform 2 was sub-cloned into the vector pST50Tr-DHFRHIS with the restriction enzymes NdeI and NGOMIV with a C-terminal cleavable 6x His tag. Different fragments of CENP-N namely CENP-N$^{1-212}$ and CENP-N$^{1-235}$ were generated in a similar way. In order to generate a C-terminal GST tagged or MBP tagged CENP-N, GST or MBP tag was amplified and cloned via gibson assembly into pST50Tr-DHFRHIS vector containing CENP-N. All CENP-N mutants were generated by site directed mutagenesis as described in section 2.3. Plasmid comprising of CENP-N$^{1-212}$ with a C-terminal His tag was generated by Dr. Kerstin Klare. All the plasmids of CENP-N either wildtype or mutants were sequence verified prior to usage.

2.2.2.2 CENP-C

To generate different fragments of CENP-C, a codon optimized (from Gene Art life technologies) human CENP-C was sub-cloned into a pGEX-6p-2rbs vector with an N-terminal cleavable GST-tag. Different fragments of CENP-C such as CENP-C$^{2-545}$ or CENP-C$^{225-364}$ were cloned into pGEX-6p-2rbs vector using BamHI and SalI sites. Plasmid comprising of CENP-C$^{2-545}$ with an N-terminal cleavable GST tag was generated by Dr. Sadasivam Jeganathan. CENP-C mutants were generated by site directed mutagenesis as described in section 2.3. All CENP-C plasmids either wild type or mutants were sequence verified prior to usage.

2.3   Site directed mutagenesis

Site directed mutagenesis is an *in vitro* method that uses custom designed primers in order to confer a desired mutation at a specific site within the gene of interest. Usually, the primers are designed in a way that the mutant codon is flanked by the overlapping oligonucleotides. Site-specific mutations were introduced by using a quick-change protocol from Stratagene. Briefly, the mutations were generated by PCR amplification of the wildtype DNA sequence using a mutated primer with overlapping oligonucleotides that helps to re-circularize the plasmid. After PCR reaction, 1 µl of DpnI enzyme was added in order to digest the template plasmid but not PCR product. This PCR product was then transformed into chemically competent OmniMax cells. Colonies were picked and DNA sequence was verified prior to usage.

2.4   Protein Expression and purification

All the protein purifications performed in this study were processed in 3 step purifications.

i) **Affinity purification**: Affinity purification is the first step in protein purification. All the proteins used in this study are either GST-tagged or His-tagged. Crude lysates harboring protein of interest with GST tag was allowed to bind to glutathione resins while proteins with His tag were allowed to bind to cOmplete His tag beads.

ii) **Ion-exchange chromatography**: Ion-exchange chromatography is process by which charged biological molecules such as proteins, peptides or nucleotides can be separated. Based on the charge of the protein, it was allowed to bind either a cation exchange or anion exchange columns or to a Hi Trap Heparin HP column.

iii) **Size-exclusion chromatography**: The last purification step is the size exclusion chromatography which separates proteins from other contaminants based on size. Different columns are used based on different sizes of proteins. Proteins were measured at 280nm using an Akta purifier system with a UNICORN control software. The eluted fractions were collected and run on a 12-14% SDS-PAGE gels and then flash frozen in liquid nitrogen until further usage.

2.4.1 (His) CENP-N constructs

Escherichia Coli (DE3) cells harboring CENP-N constructs of either CENP-N$^{1-212}$ or CENP-N$^{1-235}$ with a C-terminal 6X His tag was grown in TB media supplemented with 100 µg Ampicillin at 37 $^{\circ}$C. Temperature was reduced to 20 $^{\circ}$C when the OD$_{600}$ reached 0.8 - 1.0 followed by the addition of 0.2 mM IPTG and cultures were allowed to grow overnight. Following day, the cells were harvested at 4600 g for 15-20 mins. Following centrifugation, supernatant was discarded while the bacterial pellets were resuspended in the lysis buffer comprised of 50 mM HEPES (pH 7.5), 500 mM NaCl, 10 mM MgCl$_2$, 10% glycerol, 5 mM Imidazole, 2 mM β-mercaptoethanol. Resuspended cells were lysed by sonication followed by centrifugation at 100000 g at 4 $^{\circ}$C for 1 h. Upon centrifugation, the supernatant containing the protein of interest, was then allowed to bind to cOmplete^{TM} His tag beads (Sigma Aldrich) and incubated for 2 h at 4 $^{\circ}$C. After incubation, the resin was allowed to wash with 70 volumes of lysis buffer. After extensive washing, elution was carried out in lysis buffer supplemented with 300 mM Imidazole. As a second step of purification, a 6 ml Resource S cation exchange column was equilibrated with 15% Resource S Buffer B containing 20 mM HEPES pH 7.5, 1 M NaCl, 10 mM MgCl$_2$, 10% glycerol, 1 mM TCEP and 85% Resource S Buffer A containing 20 mM HEPES pH 7.5, 10 mM MgCl$_2$, 10% glycerol, 1 mM TCEP. The eluted CENP-N$^{1-212}$ or CENP-N$^{1-235}$ was diluted with buffer A so as to reach a final concentration of 150 mM NaCl, and then loaded onto a resource S cation exchange column that was pre-equilibrated. A linear gradient of buffer B from 150 to 1000 mM NaCl in 15 bed column volumes was applied in order to elute the CENP-N$^{1-212}$-His or CENP-N$^{1-235}$-His. Fractions containing either CENP-N$^{1-212}$-His or CENP-N$^{1-235}$-His were run on 14% SDS PAGE and then pooled. The pooled fractions were then allowed to concentrate in 10 kDa molecular weight cut-off concentrators and loaded on a Superdex 200 16/600 column pre-equilibrated in 20 mM HEPES pH 7.5, 2.5% glycerol, 300 mM NaCl and 1 mM TCEP. The eluted peak fractions were run on

a 14% SDS gel and visualized with Coomassie brilliant blue staining (2.5% Coomassie brilliant blue R250, 50% ethanol and 10% Acetic acid) and then concentrated using 10 kDa concentrator and flash frozen in liquid nitrogen until further use.

2.4.2 CENP-N$^{1\text{-}212}$ -GST

Purification of CENP-N$^{1\text{-}212}$ -GST was similar to that of the purification of CENP-N constructs with His tag but with minor changes. The culture conditions employed were similar as described above in section 2.4.3. Cells expressing CENP-N$^{1\text{-}212}$ -GST were harvested at 4600 g for 15-20 mins. Following centrifugation, the bacterial pellets were resuspended in the lysis buffer comprised of 50 mM HEPES (pH 7.5), 500 mM NaCl, 10 mM MgCl$_2$, 10% glycerol, 2 mM β-mercaptoethanol. After resuspension, the cell pellets were lysed by sonication followed by centrifugation at 100000 g at 4 °C for 1 h. Upon centrifugation, the supernatant was then allowed to bind to Glutathione-Agarose beads (Sigma Aldrich) and incubated for 2 h at 4 °C. After incubation, the resin was allowed to wash with 70 volumes of lysis buffer. After extensive washing, elution was carried out in lysis buffer supplemented with 30 mM reduced glutathione. Addition of reduced glutathione to lysis buffer decreases overall pH of the elution buffer therefore, the elution buffer was calibrated to pH 7.5. The eluted CENP-N$^{1\text{-}212}$ -GST fractions were run on a 14% SDS PAGE and then pooled. The pooled fractions were then allowed to concentrate in 10 kDa molecular weight cut-off vivaspin concentrators and loaded on a Superdex 200 16/600 column preequilibrated in 20 mM HEPES pH 7.5, 2.5 % glycerol, 300 mM NaCl and 1 mM TCEP. The eluted peak fractions were concentrated using 10 kDa concentrator and then flash frozen in liquid nitrogen until further use. All the CENP-N$^{1\text{-}212}$ -GST constructs wildtype or mutants were purified same as above.

2.4.3 (GST) CENP-C

Purification protocol of GST-CENP-C$^{2\text{-}545}$ was already made available in our laboratory by Dr. Kerstin Klare (Klare et al. 2015). All the constructs of CENP-C, either CENP-C$^{2\text{-}545}$ or CENP-C$^{225\text{-}364}$ wildtype or mutants were purified in a similar way. *Escherichia Coli* (DE3) cells harboring CENP-C constructs of either CENP-C$^{2\text{-}545}$ or CENP-C$^{225\text{-}364}$ with a N-terminal cleavable GST tag was grown in TB media supplemented with 100 µg Ampicillin at 37 °C. The temperature was reduced to 18 °C when the OD$_{600}$ reached 0.8-1.0 followed by the addition of 0.2 mM IPTG and cultures were allowed to grow

overnight. Cells were harvested at 4600 g for 20 mins. Following centrifugation, the bacterial pellets were resuspended in the lysis buffer comprised of 50 mM HEPES (pH 7.5), 500 mM NaCl, 10% glycerol, 2 mM β--mercaptoethanol. Cell pellets were lysed by sonication followed by centrifugation at 100000 g at 4 $^{\circ}$C for 1 h. Upon centrifugation, the supernatant was allowed to bind to Glutathione agarose beads (Sigma Aldrich) and incubated for 2 h at 4 $^{\circ}$C. After incubation, the resin was allowed to wash with 70 volumes of lysis buffer. Based on the application, GST-tagged CENP-C constructs were either eluted in lysis buffer supplemented with 30 mM reduced glutathione or subjected to overnight cleavage with 3C protease (kindly provided by Dortmund protein facility (DPF), MPI). A 5 ml HiTrap Heparin HP column was equilibrated with 15% buffer B containing 20 mM HEPES pH 7.5, 2 M NaCl, 10% glycerol, 1 mM TCEP and 85% buffer A containing 20 mM HEPES pH 7.5, 10 % glycerol, 1 mM TCEP. The eluted protein was diluted with buffer A so as to reach a final concentration of 200 mM NaCl, and then loaded onto a heparin column that was pre-equilibrated. A linear gradient of buffer B from 200 to 1200 mM NaCl in 20 bed column volumes was applied in order to elute the GST tagged CENP- C constructs or untagged CENP-C constructs. Fractions containing CENP-C constructs were run on 14% SDS PAGE and then pooled. The pooled fractions were then allowed to concentrate in 10 kDa molecular weight cut-off vivaspin concentrators and loaded on a Superdex 200 16/600 column preequilibrated in 20 mM HEPES pH 7.5, 2.5% glycerol, 300 mM NaCl and 1mM TCEP. The eluted peak fractions were concentrated using 30 kDa concentrator and then flash frozen in liquid nitrogen until further use.

2.4.4 CENP-A nucleosomes

purification of the *in vitro* reconstituted CENP-A nucleosomes was performed according to previously described protocol (Guse et al. 2012). Briefly, cDNA encoding for histones CENP-A and H4 were cloned into a polycistronic expression vector. This was followed by co-expression and subsequent purification of CENP-A/H4 tetramer. Similarly, both histone H2A and H2B were also cloned into different pET3 vectors expressed in *E. coli* and purified. The purified H2A and H2B histones were mixed in equal amounts to reconstitute H2A/H2B dimers. A cDNA fragment comprising of eight repeats of the 601-145 bp sequence was cloned into Puc57 vector and transformed into *E. Coli*. This was then amplified, purified and cleaved using restriction enzymes. Finally, CENP-A/H4 tetramers, H2A/H2B dimers were mixed in order to form histone

octamer. Centromeric DNA was finally added to purified histone octamer followed by a salt gradient dialysis in order to reconstitute CENP-A nucleosomes. CENP-A nucleosome assembly was performed by Doro Vogt.

2.5   Protein production with MultiBac expression system

Some of the proteins used in this study didn't yield stable products when expressed in *E. Coli*. This prompted us to use alternative expression system i.e. MultiBac expression system in insect cells. Baculovirus expression systems have emerged as a powerful eukaryotic vector systems for recombinant protein productions. Baculoviruses have the advantage of post-translationally modifying recombinant proteins in a manner similar to that of mammalian cells. These post translational modifications on the recombinant proteins might assist them in proper folding which yields in the soluble products (Jones & Morikawa 1996). Basically, the gene of interest is cloned into a transfer vector that harbors a baculovirus promoter that is flanked by baculaovirus DNA from a nonessential locus. After transfection of this recombinant transfer vector into insect cells, the gene of interest is inserted into the genome of the parent virus by homologous recombination.

2.5.1   Plasmids for insect cell expression

2.5.1.1   (GST) CENP-L

 A cDNA sequence encoding human CENP-L was cloned into the first MCS of a Multi-Bac pFG vector containing an N-terminal GST fusion with a 3C protease cleavage site. CENP-L was cloned using the restriction endonucleases BamHI and SalI that harbors polh promoter. The GST-CENP-L construct was sequence verified prior to usage.

2.5.1.2   (GST) CENP-LN

A cDNA sequence encoding human CENP-N was cloned into the second MCS of a MultiBac pFG vector that already contains a CENP-L gene cloned in its first MCS. CENP-N was cloned using the restriction endonucleases XmaI and XhoI that is placed under p10 promoter. A truncated version of CENP-N^{230-C} was also cloned in a similar way using the restriction endonucleases XmaI and XhoI. All the constructs were sequence verified prior to usage. Construct design was performed by Dr. John Weir.

2.5.1.3 CENP-HI$^{57-756}$KM

A codon-optimized cDNA sequence encoding CENP-I$^{57-756}$ was cloned into the first MCS of a pFH vector containing an N-terminal cleavable 6X His-Tag. CENP-I$^{57-756}$ was cloned using the restriction endonucleases BamHI and SalI. Similarly, a cDNA sequence encoding human CENP-M was cloned into the second MCS of the pFL vector under the control of p10 promoter using restriction endonucleases XmaI and NheI. Similarly, cDNA sequence encoding human CENP-H and CENP-K was cloned into the first and second MCS using the same restriction endonucleases as for CENP-I and CENP-M into the second pFL vector. pFL vector harboring CENP-I and CENP-M genes was linearized using the restriction endonuclease BstZ171 while pFL vector containing CENP-H and CENP-K genes was amplified using sequence and ligation independent cloning (SLIC) primers in order to ligate into the linearized vector. The SLIC reaction was carried out to generate a final pFL-based vector harboring a four-expression cassette CENP-H, CENP-K, CENP-I and CENP-M. Construct design was performed by Dr. John Weir.

2.6 Virus production, amplification and expression in insect cells

Baculoviruses were generated using a standard protocol (Trowitzsch et al. 2010). Briefly, plasmids encoding genes of interest were transformed into DH10EMBacY cells similar to that of a standard plasmid transformation except that the recovery time was extended to 6 h at 37 $^{\circ}$C. Transformed cells were plated onto a LB agar plates supplemented with 10 µg/ml Gentamycin, 7 µg/ml Tetracyclin, 50 µg/ml Kanamycin, 40 µg/ml IPTG and 100 µg/ml X-Gal. Selection of positive clones occurs via blue-white screening, positive clones are white. Single white clones were inoculated into the LB media supplemented with 10 µg/ml Gentamycin, 7 µg/ml Tetracyclin, 50 µg/ml Kanamycin and was allowed for growth overnight. Plasmid DNA was extracted using the miniprep kit, except that the DNA was allowed to precipitate with 50% Isopropanol at -20 $^{\circ}$C overnight. On the following day, the DNA was washed twice with 70% ethanol and resuspended in Tris-EDTA buffer. The DNA was then transfected with Fu GENE into Sf9-cells, an immortalized cell line isolated from *spodoptera frugiperda* at a density of $1*10^{6}$/ml. Cells were incubated for 48 h at 27 $^{\circ}$C followed by transferring of 1ml of the supernatant to a 10 cm dish with 9 ml of Sf9-cells plated at a density of $1*10^{6}$/ml. Virus amplification was allowed for 72 h at 27 $^{\circ}$C, the supernatant was labeled as virus

V_0 and stored at 4 $^{\circ}$C until further use. In order to test the expression of the protein of interest, a small-scale affinity purification was performed. Upon efficient protein expression, the virus V_0 was then allowed to amplify further to virus V1 and V2 in Sf9 cells at $1*10^6$/ml further to virus V1 and V2. After amplification of the virus V2, the virus titer was optimized for the large-scale protein expression.

2.6.1 Protein purification from insect cells

Recombinant protein purification from insect cells, were similar to that of the bacterial protein purifications except with minor modifications.

2.6.1.1 CENP-LN complex

Tnao38 cells were infected with the LN virus at a ratio of 1:40 (v/v) and incubated at 27 $^{\circ}$C for 72 h. Infected Tnao38 cells were harvested and proceeded immediately for protein purification, as the freeze thawed cell pellets has resulted in insoluble products. Tnao38 cells containing GST-tagged CENP-LN was resuspended in a lysis buffer comprising of 50 mM HEPES pH 7.5, 300 mM NaCl, 10% glycerol, 1 mM Mgcl2, 10 mM β-mercaptoethanol, 0.1 mM 4-(2- Aminoethyl) benzenesulfonyl fluoride) and 2.5 units/ml Benzonase. Cells were lysed by sonication and then cleared by centrifugation at 48,000 g for 1 h. The cleared supernatant was directly applied onto the GSH beads and incubated for ~2 h. Following incubation with the lysate, GSH beads were allowed to wash with ~10 column volumes of lysis buffer followed by ~5 column volumes of lysis buffer containing 1 M NaCl to remove the non-specifically bound DNA along with other contaminants. The GST-tagged CENP-LN was eluted by including 30 mM glutathione in the lysis buffer. All the fractions were analyzed by 14% tris tricine SDS-PAGE. The fractions containing GST-tagged CENP-LN complex were either subjected to GST cleavage by 3C protease or was directly concentrated by using a 30 kDa concentrator. The concentrated protein was then and then loaded onto Superdex 200 16/60 column. the corresponding eluted fractions were run on the 14% tris tricine gel and concentrated and stored in -80 $^{\circ}$C until further use. All other constructs of CENP-LN complex were purified in a similar way as described above.

2.7 GST pulldown assays

All GST pulldown experiments performed in this study were done using pre-blocked

GSH sepharose beads. ~1 ml of GSH sepharose beads were washed thrice with 1 ml of washing buffer comprising of 20 mM HEPES pH 7.5, 150 mM NaCl, 10 mM MgCl$_2$, 1 mM TCEP. The beads were then allowed to incubate overnight at 4 $^\circ$C with blocking buffer supplemented with 20 mM HEPES pH 7.5, 500 mM NaCl, 10 mM MgCl$_2$, 1 mM TCEP and 500 µg/ml BSA. The following day the beads were washed thrice with 1 ml of washing buffer and resuspended in ~500 µl of same buffer to yield a 50/50 slurry of beads. All GST pulldown experiments were performed under similar binding buffer conditions (10 mM HEPES pH 7.5, 300 mM NaCl, 2.5% glycerol, 1 mM TCEP) except that the concentration of NaCl was adjusted to 150 mM for GST-CENP-N/GST-LN-CENP-A nucleosome interactions. Briefly, GST-tagged proteins that serves as a bait were added to 10 µl of pre-blocked beads at a concentration of 1 µM. To this reaction, 3 µM of potential binding partners were added and the reaction was topped up to 40 µl with the binding buffer and incubated for 1 h at 4 $^\circ$C. The supernatant was decanted and the beads were washed twice with 200 µl of binding buffer supplemented with 0.05% Triton-X-100. After decanting the supernatant, 15 µl of leammli sample loading buffer was added to the beads and then boiled. Samples were then loaded on a 14% SDS-PAGE and the bands were visualized with Coomassie brilliant blue staining.

2.8   Analytical Size exclusion chromatography

All analytical size exclusion chromatography (SEC) experiments were performed using either Superdex 200 5/150 or Superose 6 5/150 on AKTAmicro system (GE Healthcare). Buffer conditions employed were 20 mM HEPES pH 7.5, 2.5% glycerol, 2 mM TCEP with either 300 mM or 150 mM NaCl. Recombinant proteins were mixed at a concentration of 6 µM in a total volume of 50 µl and then incubated on ice for 1 h. Samples were spun at 15,000 rpm for 15 mins prior to injection. All proteins were eluted under isocratic conditions using the corresponding SEC buffer (20 mM HEPES pH 7.5, 2.5% glycerol, 2 mM TCEP with either 300 mM or 150 mM NaCl). Elution of proteins was monitored at 280 nm. The corresponding peak fractions (100 µl) were loaded on a SDS-PAGE and the bands were visualized by Coomassie blue staining.

2.9   Isothermal titration Calorimetry

All ITC measurements were performed using ITC200 micro calorimeter (GE Healthcare). All recombinant protein samples were loaded onto a Superdex 200

10/300 column pre-equilibrated in 20 mM HEPES pH 7.5, 2.5% glycerol, 2 mM TCEP and 300 mM NaCl in order to resuspend all recombinant proteins in same buffer. All ITC measurements were performed at 25 °C. Typically, in each titration the protein in the cell is at a concentration 5 - 8 µM that was titrated against the ligand protein that is at a concentration of 50 - 80 µM. All ITC titrations were done with 19 x 2 µl injections at an interval of 180 s. The injections were continued beyond saturation point in order to allow for the determination of heats of ligand dilution. Data collected were fitted by least-square procedure to a single site binding model using ORIGIN software package (MicroCal, Malvern instruments, Worcestershire, UK)

2.10 Crytsallization of CENP-N$^{1-235}$

X-ray crystallography is one of the method that allows the structural determination of proteins or protein complexes at atomic resolution. The primary step of this method is to obtain protein crystals that are subjected to X-ray irradiation, and the diffracted rays are collected and recorded on a detector as a diffraction pattern. This diffraction pattern contains all the information that allows for the structural determination of the proteins. In this study, crystallization trials were performed using either CENP-N$^{1-212}$-His or CENP-N$^{1-235}$-His. Crystallization trials of CENP-N were setup at a concentration of ~6 mg/ml in a 96 well plate (Corning^{R} 3350 Protein crystallisation plates) using a Masquito Nanodispenser (TTP LabTech Ltd., Melbourn, UK) by sitting drop method. CENP-N crystals grew in various commercial screens obtained from Qiagen such as PEGSII conditions B11 (0.1 M HEPES pH 7.5, 0.2 M Sodium acetate and 20% (W/V) PEG 3000) and B12 (0.1 M Tris pH 8.5, 0.2 M lithium sulphate and 30% (W/V) PEG 3000) as well PEG H6 (0.2 M di-potassium phosphate, 20% (w/v) PEG 3350) and H8 (0.2 M di-ammonium phosphate, 20% (w/v) PEG 3350). All crystals grew within 24-48 h, reaching a maximum size in 5 - 7 days. Crystals of CENP-N$^{1-235}$His were further optimized in a 24-well plate by hanging drop method using a two-dimensional grid screen with different conditions of PEG3350 (from 6-16%) and 0.1 M HEPES pH (from 6.6 to 7.2). Since, the N-terminal region of CENP-N has not been crystallized earlier the structural determination of CENP-N by using molecular replacement method has not been possible, this has prompted us to grow selenomethionine crystals in order to obtain phase information which would help in the structure determination of CENP-N. Selenomethionine labeled CENP-N was obtained by growing CENP-N in minimal media lacking essential amino acids for ~6 h which was then followed by addition of

selenomethionine in order to incorporate selenomethionine inspite of methionine. Selenomethionie crystals grew under same conditions as that of native crystals. Fishing of the crystals were done in mother liquor solution containing 20% (V/V) glycerol.

2.10.1 Data collection and processing

Diffraction data of both native and SeMet crystals were 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 collected were integrated and scaled using XDS and Xscale (Kabsch 2010). Both native and selenomethionine crystals grew in space group $P4_1$ with two molecules of CENP-N per asymmetric unit. PHENIX (Adams et al. 2010) software was used to obtain phases from SeMet crystals, which identified 12 of the 14 possible SeMet sites. To improve the anomalous signal, two SeMet datasets from two different crystals were merged. Despite of the relative low resolution of SeMet CENP-N (3.3 Å), the quality of phases allowed for auto building of most alpha helices into the electron density. PHASER software from CCP4 (Collaborative Computational Project, Number 4 1994) was used for molecular replacement of this initial model into the native dataset (conservation resolution at I/sigma = 3 of 2.89 Å, data used to 2.74 Å). The sequence of CENP-N was allocated with the help of these anomalous peaks. Refinement was performed with REFMAC (Collaborative Computational Project, Number 4 1994) and PHENIX, which resulted in a model with very good geometry (98% residues in favored regions of the Ramachandran plot, 0% outliers) with R_{work}/R_{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 within the asymmetric unit are similar except for the loops 137 - 142 and 164 - 174. In monomer B, the loop 137 - 142 is pulled ~3 Å away from the remainder of the molecule while the loop 164-174 packs against a symmetry related molecule, which results in the stabilization of M167. Residues 166 - 168 have very weak density especially in chain A, which corroborates the weak and anomalous signal of SeM167, indicating multiple conformations of this loop. Structure of CENP-N has been deposited in the protein data bank (PDB) with code 6EQT.

2.11 Cryo electron microscopy

Cryo-EM studies on the CENP-N^{NT}:CENP-A^{NCP} complex were performed by Keda Zhou from Karolin Luger's laboratory at University of Colorado, USA. Typically, 4 µl of CENP-N^{NT}:CENP-A^{NCP} complex (~2.5 µM) was applied to quantifoil 2/2 grids that were glow discharged at 40 mA for 20s. The applied sample was incubated for 1 min and the plunge freezed into Ethane using a Vitrobot (FEI, MARK IV). Blot time was 4 s. All grids were stores in liquid nitrogen prior to imaging. CENP-N^{NT}:CENP-A^{NCP} complex was imaged using FEI Titan Krios (300kV), equipped with a Gatan K2 Summit direct detector with a nominal magnification of 29000x. The movies were captured in super resolution mode with electron dose rate at 10 electrons per pixel per second with a defocus range of -1.0 to -2.5 µm.

2.11.1 Single particle analysis of Cryo-EM

Motioncor2 was used for the alignment of images while GCTF was applied for constant transfer function (CTF) estimation (Zheng et al. 2017). Images were manually evaluated and particles were manually pocked for initial 2D classification in RELION 2.05 (Fernandez-Leiro & Scheres 2017). Initial 2D class averages were used for particle auto-picking in RELION 2.0. Reference free 2D classification of these particles resulted in the generation of 200 class averages. Noisy class averages were discarded. Particles from the retained class averages were used for unbiased 3D construction in cryoSPARC, which was used as a reference for 3D classification in RELION 2.0. Four 3D classes were generated. The particle images from the class at high resolution were used for 3D refinement in Relion 2.1b1. The resolution of the map was estimated by RELION (Kucukelbir et al. 2014). For model fitting, high-resolution structure of nucleosome with 601 DNA (PDB: 3LZ0) (Vasudevan et al. 2010), X-ray structure of CENP-A^{NCP} (PDB 3AN2) (Tachiwana et al. 2011) and CENP-N that was crystallized in this study were used. The map was locally refined in Coot and the final model was subjected to real space refinement in PHENIX. The coordinates were deposited in the protein data base (PDB: 6C0W and EMD-7326)

2.12 Cell culture

All cell lines used in this study were adherent cell lines grown in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% TET-free Fetal Bovine Serum, 2

mM L-Glutamine. For passaging of the cells, the media was removed and the cells were washed with 1x PBS (Phosphate buffer saline) followed by incubation with Trypsin at 37 $^{\circ}$C for 5 minutes. To stop trypsination, fresh medium containing serum was added where by 1/10 of the cells were seeded for every new passage. Cells were cultured in 10 cm dishes at 37 $^{\circ}$C supplemented with 5% CO_2 and were sub cultured at a confluence of 70-80%.

2.12.1 Plasmids for mammalian expression

DNA encoding for either human CENP-C sequence (codon optimized from Gene Art Life Technologies$^{TM)}$ was cloned into pcDNA5/FRT/TO-eGFP-ires vector, a derivative of pcDNA5/FRT/TO vector (Invitrogen, Carlsbad, CA) which was generated in-house as an N-terminal fusion to eGFP (Petrovic et al. 2010). Mutant CENP-C constructs were generated as described in section 2.3. Wildtype or mutant CENP-C were cloned using restriction endonucleases BamHI and XhoI. A cDNA sequence encoding human CENP-N was sub-cloned into pcDNA5/FRT/TO-mCherry-ires, a derivative of pcDNA5/FRT/TO vector which was generated in-house as an N-terminal fusion to mCherry. Wild type or mutant CENP-N constructs were cloned using the restriction endonucleases BamHI and XhoI. Mutant CENP-N constructs were generated as described in section 2.3. For transient transfections, both wild type and mutant CENP-N constructs was cloned into a pcDNA5-FRT-TO vector harboring a C-terminal mCherry tag.

2.13 Generation of stable cell line

Parental Flp-In-T-Rex Hela cell lines (a gift from Prof. Stephan S. Taylor's laboratory at the university of Manchester) were used to generate doxycycline inducible cell lines. Flp-In T-Rex Hela cells expressing mCherry-CENP-N were generated as described (Tighe et al. 2004). The cells were maintained in DMEM supplemented with 10% tetracycline-free FBS, 2 mM L-Glutamine, 250 μg/ml Hygromycin and 5 μg/ml Blasticidin. Cell lines were generated by co-transfection of the Flp-recombinase expression vector (Pog44) and a pcDNA5/FRT/TO expression plasmid consisting of mCherry CENP-N construct using X-tremeGENE^{TM} 9 DNA transfection reagent. Cells were selected in DMEM media after 48 h of transfection followed by exchange of the media every 2 days. After 2 weeks in selection, single cell clones were transferred to

separate dishes and then optimized for protein expression by adjusting Doxycycline concentrations (~5 - 25 ng/ml) and treatment time. Upon positive protein expression, the cells were frozen and stored.

2.13.1 Transient transfections of CENP-N constructs

Transient transfections were performed In U2OS cells, a gift from A. Bird (MPI-Dortmund, Germany). Cells were grown in DMEM media supplemented with 10% FBS (Clontech, Japan), penicillin and streptomycin (GIBCO, Carlsbad, CA) and 2 mM L-glutamine (PAN Biotech). Transient transfections were performed in asynchronously growing cells and expressed for 24 h before preparation for immunofluorescence analysis.

2.13.2 Small interfering RNA (siRNA) depletion of CENP-C

Depletion of endogenous CENP-C was performed by transient transfections of a small interfering RNA (5'-GGAUCAUCUCAGAAUAGAA-3') that targets the mRNA of endogenous CENP-C. Transfection was performed using HiPerFect transfection reagent according to manufacturer's instructions. To obtain efficient depletion of CENP-C, 60 nM siRNA was transfected 3 times within 72 h (Klare et al. 2015). Within the last 48 h cells were transiently transfected with either GFP-CENP-C^{WT} or GFP-CENP-C^{5A} mutant. To arrest the cells in mitosis, 16 h prior to fixation the cells were synchronized overnight with 0.33 µM of nocodazole (Sigma-Aldrich).

2.14 Immunofluorescence

Fixation of U2OS or FlpIn T-Rex hela cells were performed using 4% paraformaldehyde in phosphate buffered saline (PBS). Fixed cells were then permeabilised using 0.25% Triton X-100 in PBS followed by the addition of blocking buffer i.e. 3% BSA in PBS. U2OS cells were stained with the following antibodies anti-CENP-C (rabbit polyclonal) diluted 1:1000 and CREST or anti centromere antibody (Antibodies Inc. 15-2340001) diluted 1:100. FlpIn T-Rex Hela cells were stained for GFP (GFP-Boost, chromtek) diluted 1:500, mCherry (RFP-Boost, chromtek) diluted 1:500, CENP-C diluted 1:1000, CENP-HK diluted 1:1000, CREST or anti-centromere antibodies diluted 1:100 or CENP-A (Mouse, Gene Tex) diluted 1:500 in blocking buffer. Cells were allowed to incubate with the primary antibodies for 2-4 h followed

by the addition of secondary antibody. Secondary antibodies used in this study are donkey anti human Alexa Fluor 405, donkey anti-rabbit Alexa flour 488, donkey anti-human (Jackson immmunoResearch Laboratories, Inc., Westgrove, PA) and donkey anti-mouse (Invitrogen), DNA was stained with 0.5 µg/ml with DAPI (Serva), and coverslips were mounted with Mowiol mounting media (Calbiochem).

2.15   Image analysis, representation and quantification

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 a Z-sections (using the softWoRx software from Deltavision), all the images were then converted into a maximal intensity projections of TIFF files for illustrative purposes. FlpIn-T-Rex HeLa cells were imaged at room temperature with a spinning disk confocal microscopy of a 3i Marianas system equipped with an Axio Observer Z1 microscope, a CSU-X1 confocal scanner unit, Plan-Apochromat 63x or 100x/1.4NA objectives and Orca Flash 4.0 sCMOS camera. All the images were then converted into a maximal intensity projections of TIFF files for illustrative purposes. Quantification of the kinetochore signals were performed on unmodified Z-series images using Imaris 7.3.4 software 32-bit software (Bitplane, Zurich, Switzerland). After background subtraction, all signals were normalized to CREST while for experiments done on U2OS cells, the values obtained were further normalized to mean CENP-N-mCherry^{WT} construct. Quantifications are based on analyzing 7 - 10 cells and 177 to 393 kinetochores per condition were analyzed. For representative purposes, single cells were cropped and same channel were adjusted to the same levels in each cell. All channels collected at the microscope were exported separately as a 16-bit Tiff files along with a scale bar of 10 µm in order to compare the size of all cells. All images were opened in image J and adjusted to match the same contrast and brightness. These images were then converted to 8-bit Tiff files for representation. All the measurements were exported in Excel (Microsoft) and graphed using GraphPad Prism 6.0 (GraphPad software, San Diego California, USA).

3 Results

3.1 Purification of the recombinant kinetochore proteins

To perform biochemical and structural studies on the CENP-LN complex recombinant material was needed in sufficient amount. Therefore, the first step of my work was to purify different kinetochore proteins. To this end, both bacterial and insect cell expression systems were used to express proteins. Purification of full-length proteins were made whenever possible. However, under the conditions where the full-length proteins could not be generated, other strategies were used. These included co-expressions and shorter constructs based on secondary structural predictions. In addition, some of the constructs were also generated using different tags for various biochemical assays. Detailed protocols for the expression and purification of the CCAN sub-complexes from bacteria and insect cells can be found in section 2.4, 2.5 and 2.6. Fortunately, the expression and purification of certain components had previously been established in the laboratory.

3.1.1 Purification of CENP-LN complex:

Despite several attempts, we could not purify CENP-N^{FL} because of its stability issues. Previous studies on CENP-N have shown that CENP-N forms heterodimer with CENP-L (Weir et al. 2016; McKinley et al. 2015; Hinshaw & Harrison 2013). When CENP-N was coexpressed along with CENP-L, this complex was soluble in insect cells and yielded a stable product. Size exclusion chromatography (SEC) elution profile of GST-CENP-LN is depicted in Figure 3-1, left. The eluted peak fractions were run on a 14% SDS-PAGE and stained with Coommassie blue dye (Figure 3-1, right). CENP-LN purified by this method was assessed to be >95% pure.
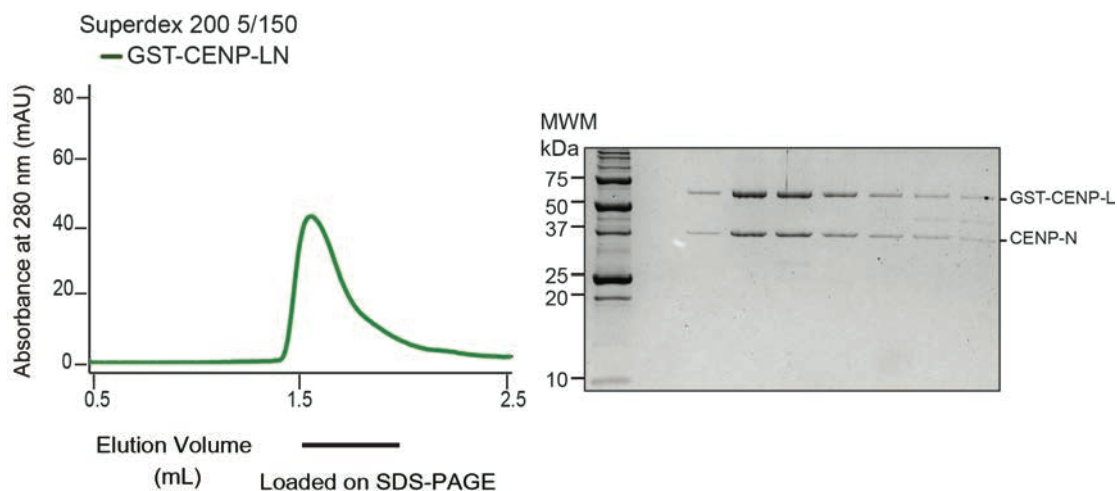
Figure 3-1 Purification of GST-CENP-LN complex

Size exclusion chromatography of GST-tagged CENP-LN complex expressed in insect cells. Right, SDS-PAGE analysis of the fractions indicated by a horizontal bar under the SEC elution profile.

3.1.2 Construct design for CENP-N

As mentioned previously, CENP-N^{FL} could not be expressed in either bacteria or insect cells. This prompted us to generate different fragments of CENP-N that could be purified *in vitro*. Since the crystal structure of human CENP-N was unavailable we took advantage of the PSIPRED online tool, which predicts the secondary structure of proteins, which was used to design different CENP-N constructs (Figure 3-2).
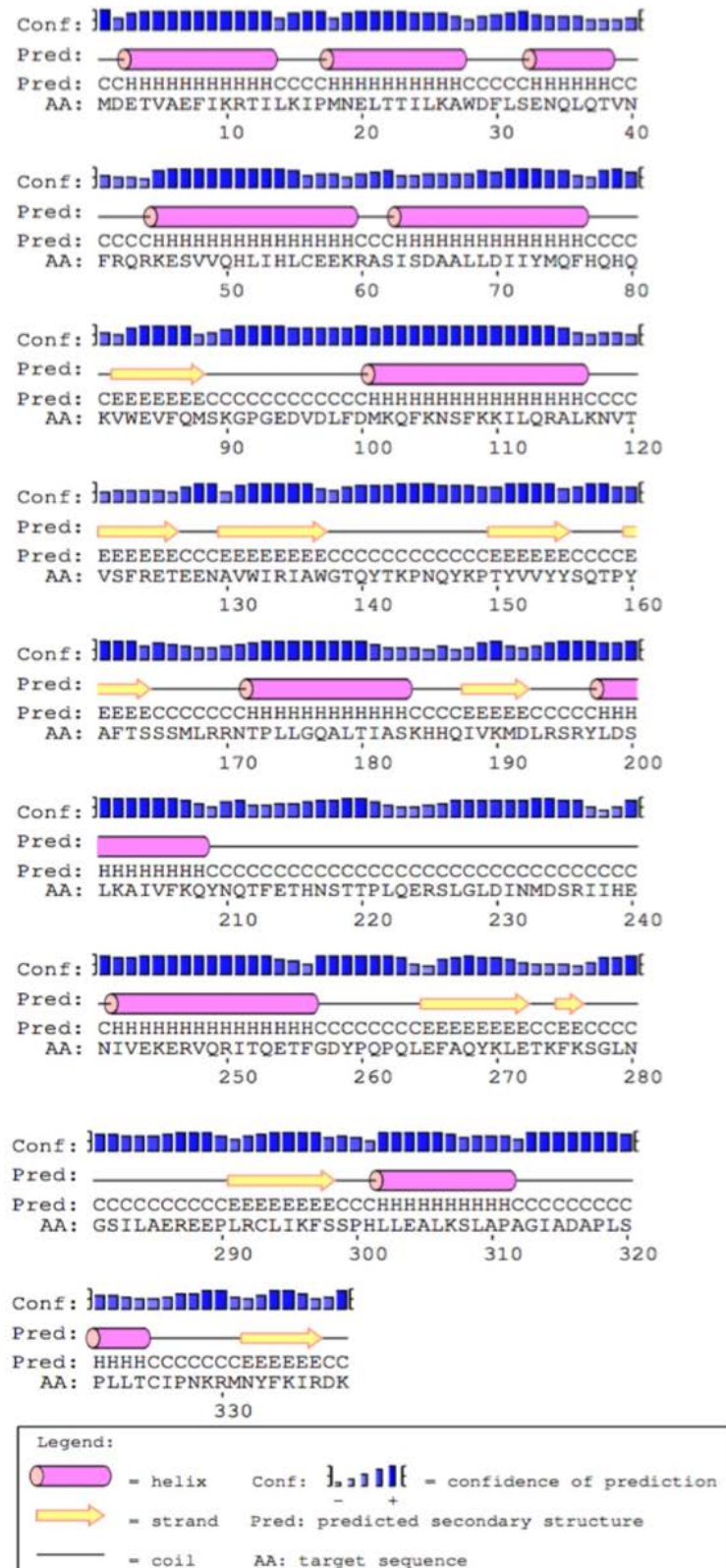
Figure 3-2 Secondary structure prediction of CENP-N by PSIPRED

PSIPRED incorporates two feed-forward neural networks that analyze the outputs obtained from PSI-BLAST (Position Specific Iterated – BLAST).

3.1.3 Purification of CENP-N$^{1-212}$

Based on PSIPRED secondary structure, a construct of CENP-N$^{1-212}$ had already been made available in the lab. An N-terminally tagged CENP-N$^{1-212}$ did not yield any soluble CENP-N when expressed in bacteria or insect cells. In contrast, when CENP-N$^{1-212}$ with a C-terminal-His tag was expressed in bacteria, I was able to purify CENP-N$^{1-212}$ with high yields (3 mg per liter of culture). Additionally, to perform a solid phase binding assay using CENP-N$^{1-212}$, we sub-cloned the GST coding sequence into the vector containing CENP-N-His in order to generate CENP-N$^{1-212}$-GST construct. Expression and purification of CENP-N$^{1-212}$-GST were performed in similar ways to that of CENP-N$^{1-212}$-His with some minor modifications. Size exclusion chromatography (SEC) elution profile of CENP-N$^{1-212}$-GST is depicted in Figure 3-3, left. The eluted peak fractions were run on a 14% SDS-PAGE and stained with Coommassie blue dye (Figure 3-3, right). CENP-N purified by this method was assessed to be >95% pure.
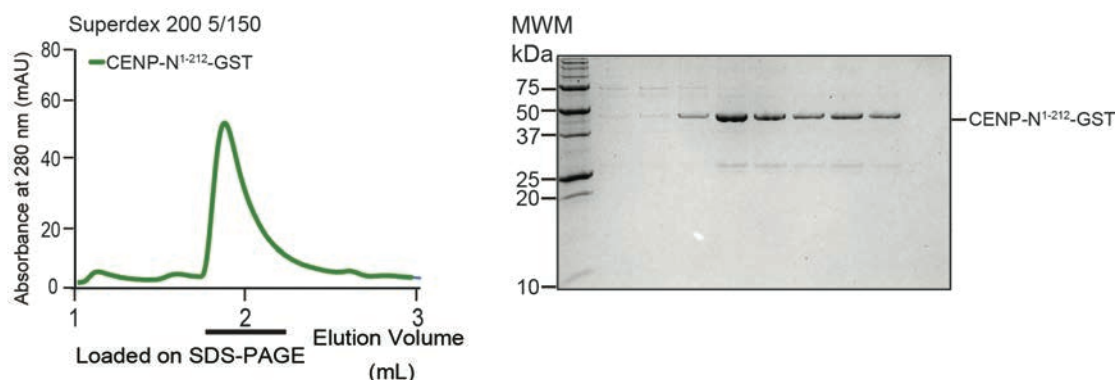

Figure 3-3 Purification of CENP-N$^{1-212}$-GST

SEC elution profile of CENP-N$^{1-212}$-GST expressed in bacteria. The corresponding peak fractions that are indicated by the horizontal bar under the profile were run on 14% SDS-PAGE in order to monitor the purity of the protein.

3.1.4 Purification of CENP-N$^{1-235}$

In addition to the above CENP-N$^{1-212}$ construct, I also generated a similar construct, CENP-N$^{1-235}$-His. The expression of CENP-N$^{1-235}$-His in bacteria enabled us to purify this CENP-N construct in higher yields. SEC elution profile of CENP-N$^{1-235}$-His

is depicted in Figure 3-4, left. The eluted peak fractions were run on a 14% SDS-PAGE and stained with Coommassie blue dye (Figure 3-4, right). CENP-N purified by this method was assessed to be >95% pure.
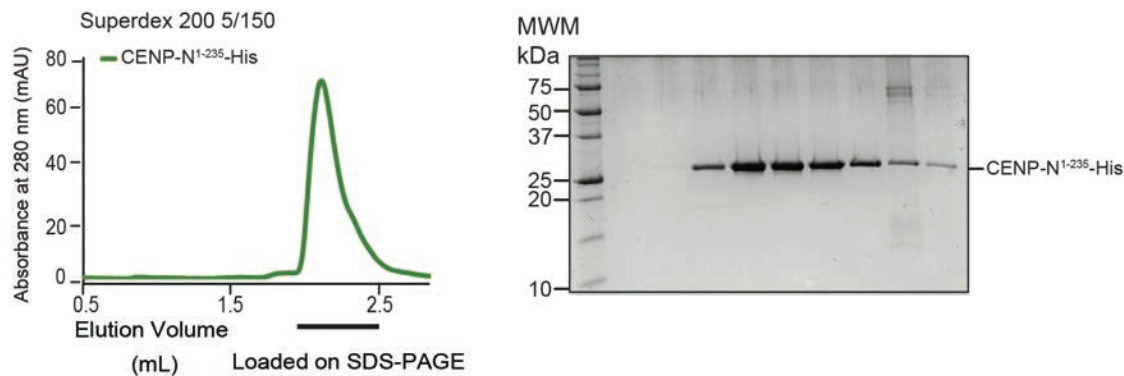

Figure 3-4 Purification of CENP-N$^{1-235}$-His

SEC elution profile of CENP-N$^{1-235}$-His expressed in bacteria. The corresponding peak fractions that are indicated by the horizontal bar under the profile were run on 14% SDS-PAGE in order to monitor the purity of the protein. In addition to CENP-N$^{1-212}$, CENP-N$^{1-235}$ can also be purified in larger quantities.

3.1.5 Purification of CENP-LN^{230-C}

Like the N-terminal region of CENP-N, the expression of the C-terminal region of CENP-N alone was also attempted. Despite several attempts, the expression of C-terminal region of CENP-N was unsuccessful. However, when the C-terminal region of CENP-N was coexpressed along with CENP-L, this complex was soluble in insect cells. Therefore, CENP-LN^{230-C} was purified using similar purification procedures as that of CENP-LN^{FL} complex. SEC elution profile of CENP-LN^{230-C} complex is depicted in Figure 3-5, left. The eluted peak fractions were run on a 14% SDS-PAGE and stained with Coommassie blue dye (Figure 3-5, right).
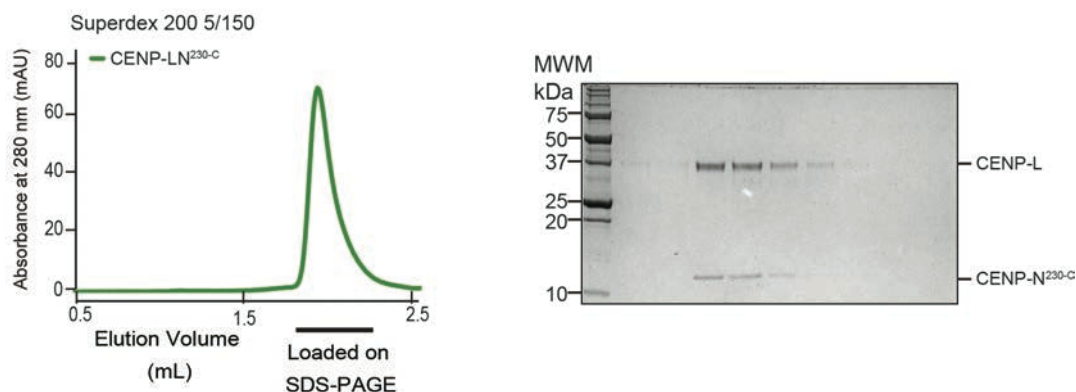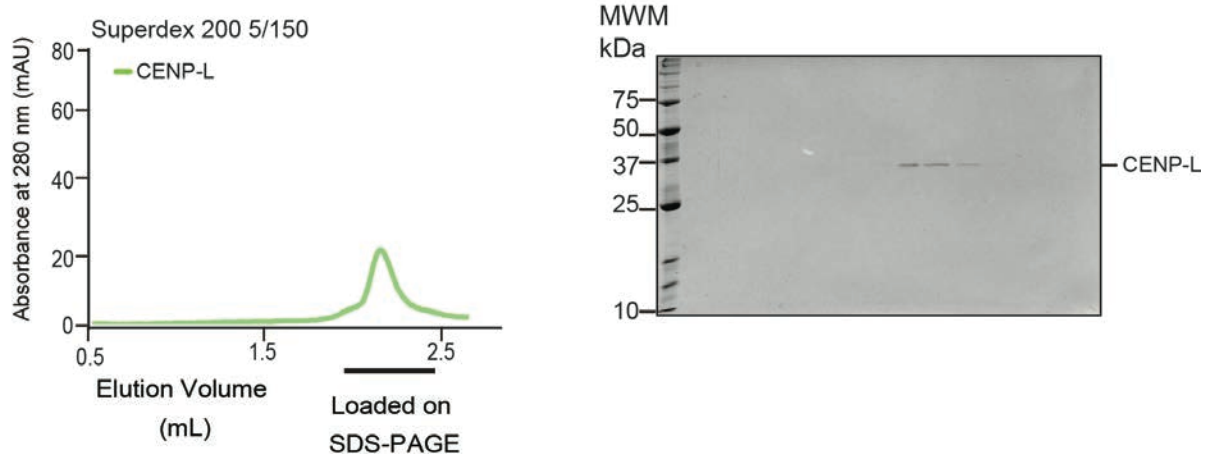
Figure 3-5 Purification of CENP-LN^{230-C}

SEC elution profile of CENP-LN^{230-C} expressed in insect cells. The corresponding peak fractions indicated by the horizontal bar under the profile were run on 14% SDS-PAGE in order to monitor the purity of the protein.

3.1.6 Purification of CENP-L

To investigate whether CENP-L binds to CCAN subunits other than CENP-N, the CENP-L coding sequence was sub-cloned into custom-made pFG vector as a GST fusion. The purification of CENP-L^{FL} was achieved by employing a similar expression and purification procedure to those used for the CENP-LN^{FL} complex. This shows that CENP-L^{FL} is stable when compared with CENP-N^{FL}, which is only stable when co-expressed with CENP-L. Studies on Iml3, a yeast ortholog of CENP-L, demonstrated that CENP-L forms both homodimers and heterodimers (with CENP-N) (Hinshaw & Harrison 2013). However, CENP-L^{FL} eluted at the predicted molecular weight (33 kDa) of a monomer from the gel filtration column, suggesting that it is a monomer, although we cannot exclude the possibility of a concentration-dependent dimerization (Figure 3-6).

Figure 3-6 Purification of CENP-L

SEC elution profile of CENP-L expressed in insect cells. The corresponding peak fractions that are indicated by the horizontal bar under the profile were run on 14% SDS-PAGE to monitor the purity of the protein.

3.1.7 Purification of additional CCAN sub-complexes

To investigate the binding of different CCAN proteins with CENP-LN complex, different CCAN sub-complexes were produced recombinantly. Some CCAN members have been reported to directly interact with the CENP-LN complex (Carroll et al. 2009; Weir et al. 2016; McKinley et al. 2015). Purification protocols for some of the CCAN subunits were already available in the lab (Weir et al. 2016; Klare et al. 2015; Basilico et al. 2014). Most of the CCAN sub-complexes and truncated constructs were produced in collaboration with other members of the group, including Dr. John Weir and Dr. Kerstin Klare, with the kind technical assistance from Doro Vogt, Sabina Wohlgemuth and Ingrid Hoffmann.

3.2 Recombinant CENP-LN complex binds to CENP-A^{NCP}

To test the selective binding of the recombinant CENP-LN complex to CENP-A^{NCP}, *in vitro* reconstituted nucleosome core particles (NCPs) with a 145 bp "601" DNA sequence containing H4, H2A, H2B and either CENP-A or canonical H3 histones were produced (Guse et al. 2012) and tested for their ability to interact with either GST-CENP-LN or GST-CENP-L alone. In GST pulldown assays, the recombinant CENP-LN complex (bait) bound selectively to CENP-A^{NCPs} (prey) and not canonical H3^{NCPs}, as described previously (Carroll et al. 2009). We did not observe any interaction of CENP-L with either CENP-A^{NCPs} or H3^{NCPs} (Figure 3-7 A-B). These results suggest that CENP-N is necessary to interact with CENP-A^{NCPs}.
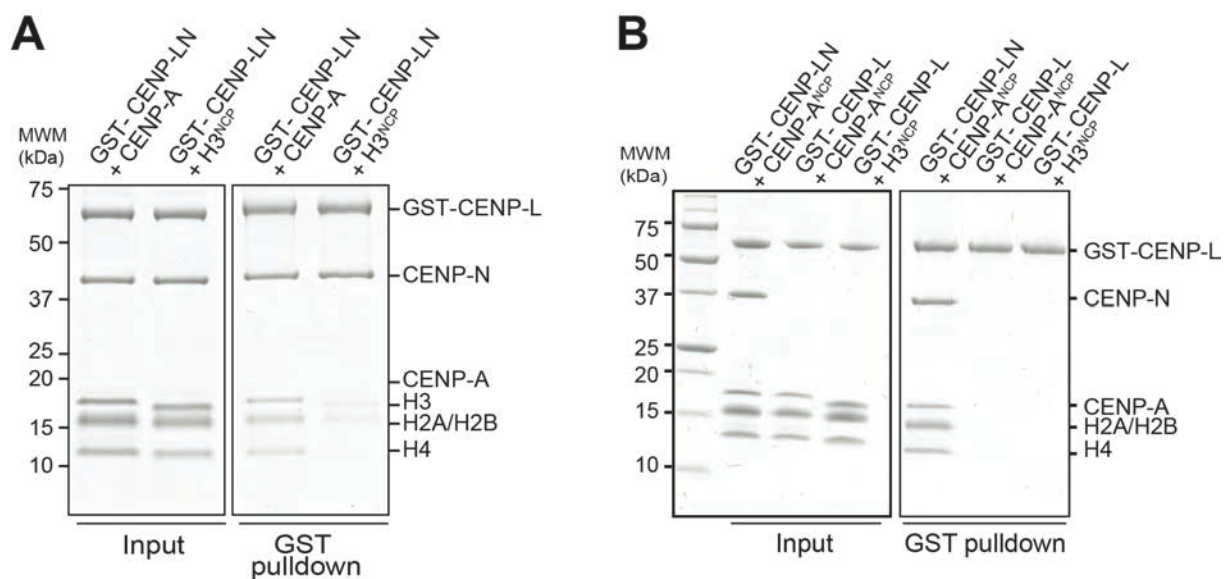
GST-pulldown assay of recombinant CENP-LN complex (A) or GST-CENP-L (B). The indicated GST baits were immobilized on glutathione-sepharose beads (1 μM) and incubated with 3 μM of either CENP-A^{NCPs} or canonical H3$^{NCPs.}$

3.3   N-terminal region of CENP-N is sufficient to bind to CENP-A^{NCP}

Because the CENP-N full-length protein is presumably unstable and cannot be expressed, several attempts were made to assess binding to CENP-A^{NCP} using truncated versions of CENP-N. Previous studies on CENP-N have indicated that CENP-N$^{1-289}$ binds to nucleosomes (Carroll et al. 2009). GST-pulldown assay reveals that CENP-N$^{1-212}$ binds specifically to CENP-A^{NCPs} over canonical H3^{NCPs} (Figure 3-8, A), while a construct that lacks the N-terminal residues (CENP-N$^{1-229}$) did not bind nucleosomes (Figure 3-8, B). Taken together, these results demonstrate that the N-terminal region of CENP-N (1-212) mediates the binding with CENP-A^{NCPs}.
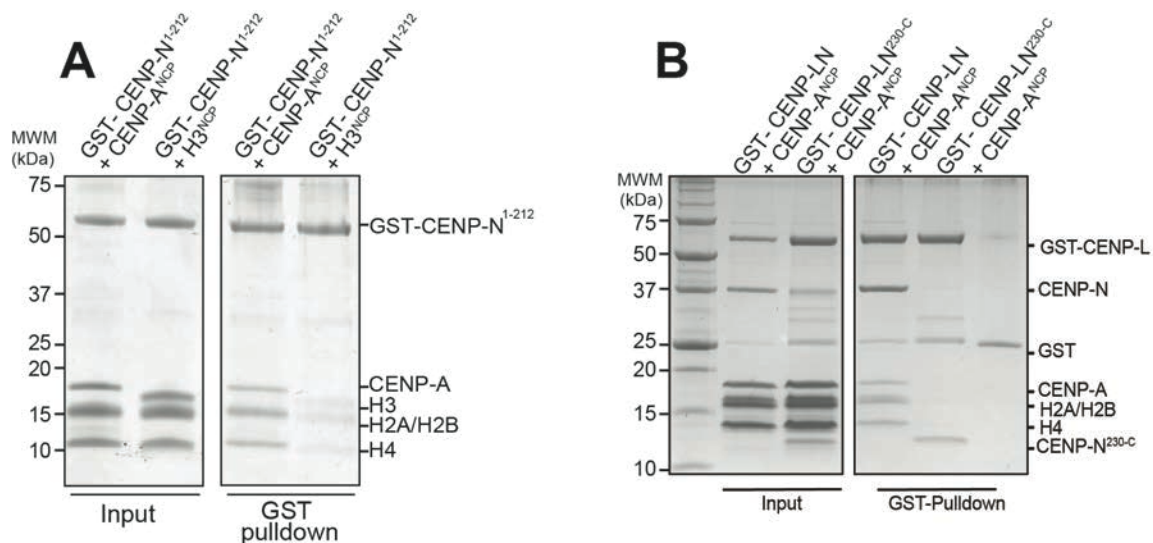

Figure 3-8 CENP-N$^{1-212}$ interacts with CENP-A^{NCP}

GST-pulldown assay of recombinant CENP-N$^{1-212}$ (A) or CENP-LN^{230-C} complex (B). The indicated GST baits were immobilized on glutathione-sepharose beads (1 μM) and incubated with 3 μM of either CENP-A^{NCP} or canonical H3$^{NCP.}$

3.4   Analytical ultracentrifugation

CENP-N directly binds to CENP-A^{NCPs} and CENP-A^{NCPs} are octameric in nature, consisting of two copies of each histones (H2A, H2B, CENP-A, H4). To confirm the expedited stoichiometry of the CENP-N in complex with CENP-A^{NCP}, the analytical

ultracentrifugation (AUC) method was used. The sedimentation velocity AUC method allows for the determination of molecular weight, shape and also the stoichiometry of protein complexes. AUC was performed on two different samples: CENP-A^{NCP} and CENP-LN:CENP-A^{NCP} complex. Analysis of the sedimentation distribution curves of both CENP-A^{NCP} and CENP-LN:CENP-A^{NCP} complex demonstrates that there is one predominant species for each of these samples. In agreement with the previous studies, the observed molecular weight of CENP-A^{NCP} is 206 kDa (Figure 3-9 A-B and E) suggesting that CENP-A^{NCP} is octameric in nature, while the derived molecular weight of CENP-LN:CENP-A^{NCP} complex is 360 kDa, which suggests that two copies of CENP-LN bind to single CENP-A^{NCP} (Figure 3-9 C-D and E).
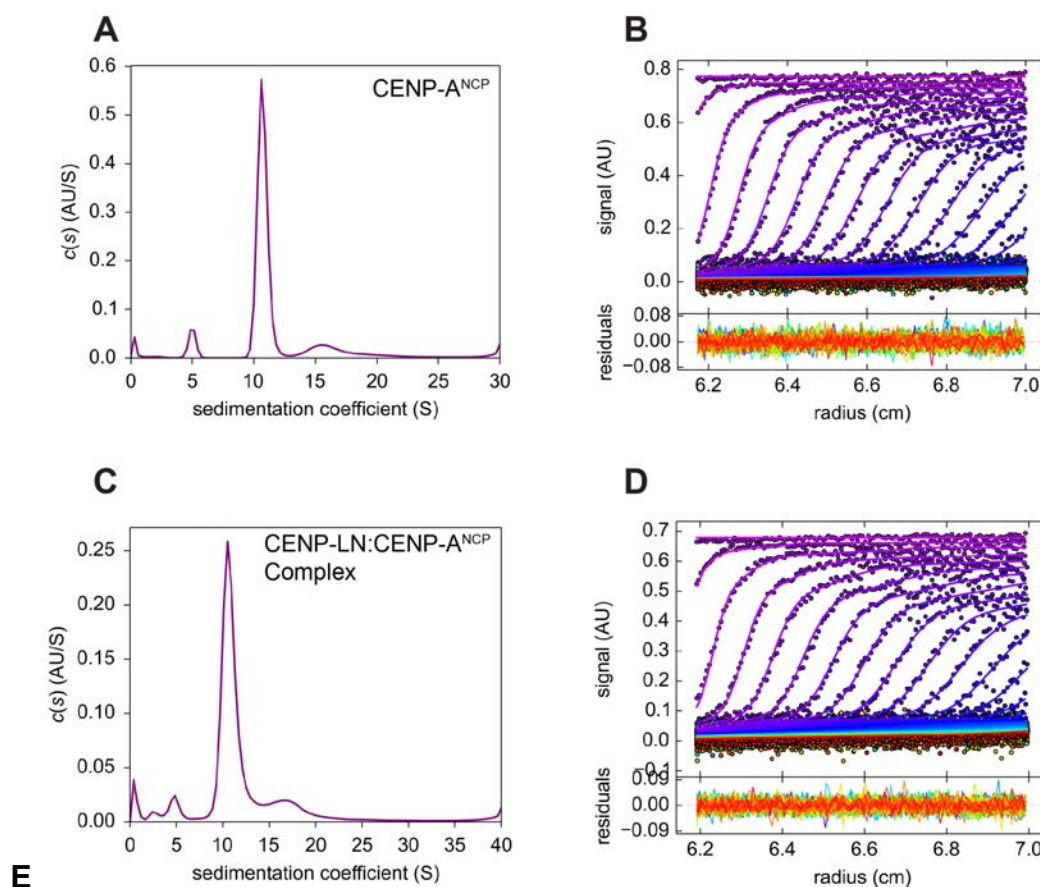
S.no.	Complex	Predicted mass (kDa)	Observed mass (kDa)	Frictional ratio	Sedimentation coefficient (S)	Predicted stoichiometry
1	CENP-A^{NCP}	200	206	1.4	10.7	Octamer
2	CENP-LN+ CENP-A^{NCP}	357	360	1.76	10.9	2 CENP-LN:1 Octamer

Figure 3-9 Analytical ultracentrifugation

Sedimentation velocity AUC results from CENP-A^{NCP} alone (A) and CENP-LN:CENP-A^{NCP} complex (C) with the best fit size distributions. Best-fitting results of the sedimentation velocity AUC data of CENP-A^{NCP} alone (B) and CENP-LN:CENP-A^{NCP} complex (D). Residuals represent the deviation of the continuous c(s) distribution model from the observed signals. (E) Summary table of the results obtained by AUC analysis.

3.5 Structural analysis of CENP-N

Previous studies on Chl4/Iml3, the yeast orthologs of CENP-N/L, resulted in the determination of the structure of the C-terminus of Chl4^{CENP-N} in complex with Iml3^{CENP-L} (Hinshaw & Harrison 2013). However, the structural characterization of human CENP-LN heterodimer remains unknown. Therefore, several attempts were made to determine the structure of the CENP-LN complex. The lack of crystal structure of the N-terminus of CENP-N or Chl4, prompted us to perform crystallization trials on the minimal CENP-N binding region of CENP-A^{NCP}.

3.5.1 Crystallization of CENP-N

Crystallization trials were performed using two constructs of CENP-N, CENP-N$^{1-212}$ and CENP-N$^{1-235}$. Using the sitting drop method, several conditions were screened using commercial kits obtained from Qiagen. Crystals of both CENP-N$^{1-212}$ and CENP-N$^{1-235}$ were obtained in different conditions such as PEGS II B11 (0.1 M HEPES pH 7.5, 20% PEG3000 and 0.2 M sodium acetate) and PEGS H6 (0.2 M di-potassium phosphate and 20% PEG 3350) at 4 °C. CENP-N crystals grew within 24 to 48 h, with a maximum growth reached in 5 to 7 days as depicted in Figure 3-10.
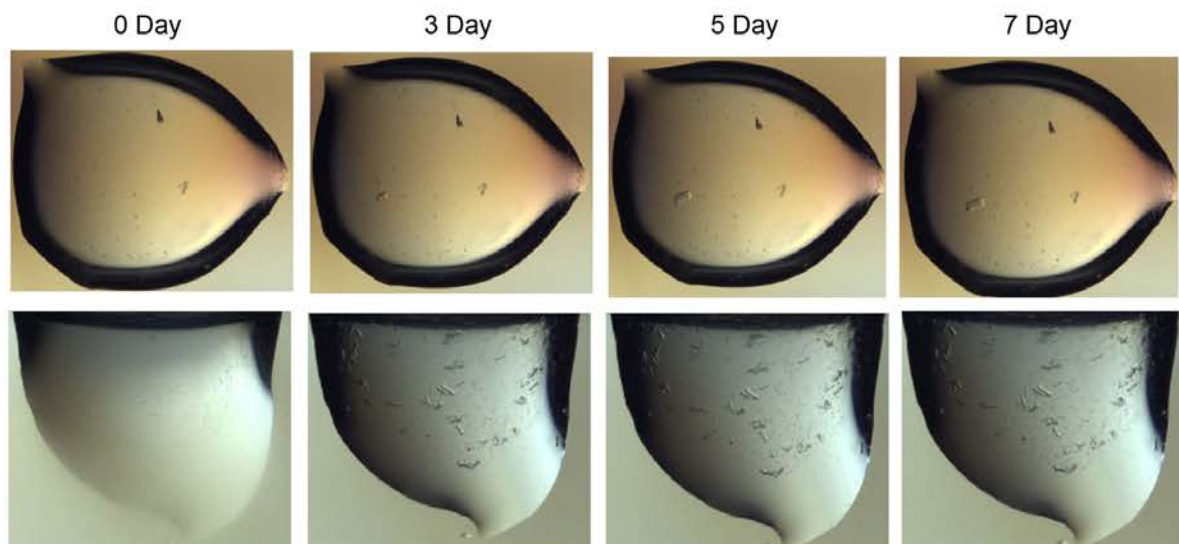
Figure 3-10 Crystallization conditions for CENP-N[1-235]

Representative images of the crystals of CENP-N[1-235] in PEGS condition B11 (upper panel) and PEGSII H6 (lower panel). Crystals grew at 4 ºC for up to 7 days.

Further optimization of the crystals was carried out with CENP-N[1-235]-His construct by the hanging drop method in a 24 well plate. In order to obtain phase information, SeMet CENP-N[1-235] (where most of the methionines are replaced by SeMet) was purified. SeMet has the characteristic scattering property compatible with single wavelength anomalous diffraction (SAD), which is used to derive experimental phases. In addition, to the native CENP-N[1-235] crystals, SeMet CENP-N[1-235] crystals were also obtained under similar screening conditions.

3.5.2 Structure determination of CENP-N[1-235]

SeMet crystals of CENP-N diffracted to 3.3 Å, while the native crystals diffracted to 2.8 Å. Both native and SeMet CENP-N crystalized with two molecules per asymmetric unit in space group $P4_1$. In order to improve the anomalous signal, SeMet datasets from two different crystals were merged. Initial model of CENP-N was derived from the phase information obtained from SeMet datasets. Molecular replacement of this initial model into the high-resolution native datasets allowed us to determine the structure of CENP-N. A detailed description of the structural determination of CENP-N is provided in section 2.10. The structure of CENP-N was solved by Dr. Ingrid Vetter, MPI, Dortmund. Data collection and refinement statistics are represented in Table-3-1

Table 3-1 Data collection and refinement statistics for the crystal structure of CENP-N

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	88.99 88.99 76.96	89.14 89.14	88.99 88.99
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/σI	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

3.5.3 Structural analysis of CENP-N$^{1\text{-}235}$

Although we used CENP-N$^{1\text{-}235}$, only the first 1-210 residues are clearly visible, suggesting that the last 25 residues are flexible. CENP-N$^{1\text{-}235}$ is composed of two closely associated domains that interact through an extended interface in order to form a single structural unit. Residues 1-77 are organized in five α-helices wrapped in a helical bundle, while residues 78-212 form a six-stranded anti-parallel β-sheet that is flanked by α-helices (Figure 3-11).
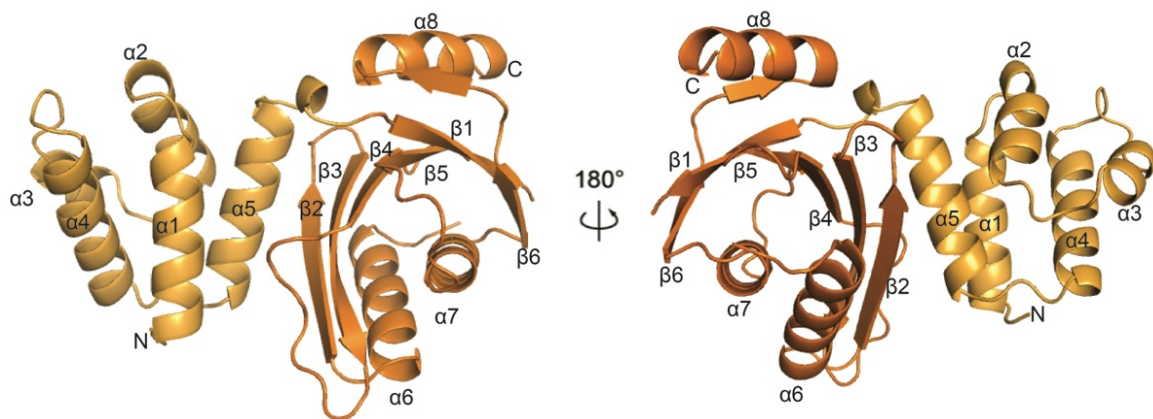

Figure 3-11 Crystal structure of CENP-N$^{1\text{-}235}$

CENP-N$^{1\text{-}235}$ exhibits a two-domain organization: an N-terminal domain (residues 1-77, bright orange) and a C-terminal domain (residues 78-212, orange) in a fixed orientation.

The crystal structure of CENP-N revealed that its first residue (Met1) is buried inside the hydrophobic core of the protein, likely to stabilize an interaction between the α1 and α4 helices (Figure 3-12). This explains why an N-terminal tagging of CENP-N was not viable in our biochemical experiments. However, the addition of a glycine linker (Gly-Gly)$_2$ at the Met1 position, was compatible with the N-terminal tagging of CENP-N. This
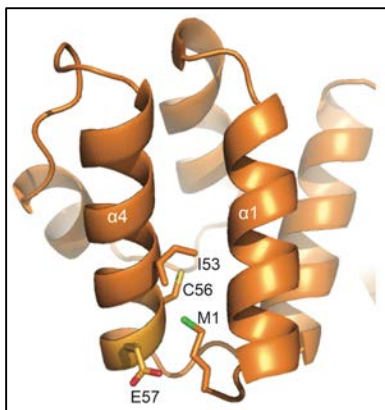
has allowed us to test if the C-terminal tagging of CENP-N$^{1\text{-}212}$ has any influence on the binding to nucleosomes. We performed GST-pulldowns using the N-terminal tag of CENP-N and found that it binds equivalently to CENP-A nucleosomes (Data not shown). These results demonstrate that both N-terminal and C-terminal tagging of CENP-N$^{1\text{-}212}$ has no effect on the binding to nucleosomes.

Figure 3-12 Zoom-in-view of the N-terminal domain of CENP-N$^{1\text{-}235}$

Met1 residue within the pyrin domain of CENP-N is buried inside the hydrophobic core of the protein.

A closer look at the interface of the two domains of CENP-N[1-235] reveals that the residues within the β3-β4 loop engage in contact with the residues within α1 and α5 helices, suggesting that the two domains are in a fixed orientation as depicted in Figure 3-13. This further suggests that the interface between the two domains of CENP-N[1-235] is unlikely to change even when CENP-N is bound to CENP-A[NCP].
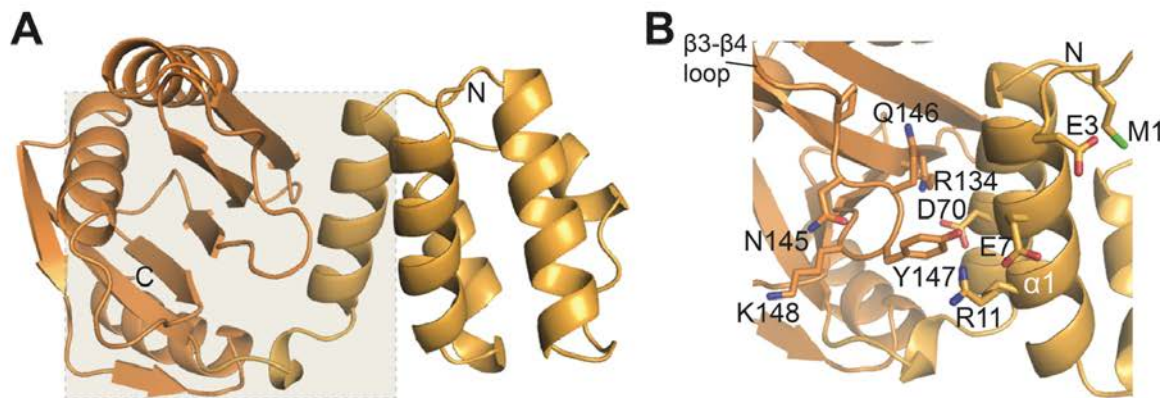

Figure 3-13 Zoom-in-view of the interface of CENP-N[1-235]

(A) Crystal structure of CENP-N[1-235]. (B) Zoom-in-view of the highlighted grey box in A. Residues within the pyrin domain interact with the CLN-HD domain, resulting in the fixed reciprocal orientation of these domains.

3.5.4 Structural similarities of CENP-N

The DALI server was used to identify the proteins with similar folds within the protein data bank (PDB). The DALI server identified structural homology of the first domain of CENP-N (residues 1-77) with the PYRIN domain (PDB 4O7Q) (Figure 3-14 A-B) and the second domain of CENP-N (residues 78-210) with the N-terminal domain of Iml3, the yeast ortholog of CENP-L (PDB 4KR1) (Figure 3-14 C-D). Pyrin domains (PYDs) are 'death fold' family domains that are implicated in protein-protein interactions. Most of the previous studies on PYDs suggest a role in inflammation and apoptosis. Pyrin domains was originally discovered in the protein PYRIN and was initially referred to as DAPIN or PAAD (Weber & Vincenz 2001). Most of the PYDs are exclusively present at the amino terminal ends of proteins, such as CENP-N. To date, PYDs have not been implicated in the interactions with DNA or chromatin; therefore, this is the first study to suggest a role for PYDs in the interaction with chromatin. Despite the low sequence similarity of CENP-N (residues 78-210) and Iml3[CENP-L], the DALI server identified structural similarity between the domains of CENP-N and CENP-L. Thus, we refer to domains of CENP-N and CENP-L as CLN-HD (for CENP-L and CENP-N homology domain). Collectively, these structural similarities between CENP-N and CENP-L

suggest that they are evolutionarily related. Despite the structural similarity of CENP-N and Iml3$^{\text{CENP-L}}$, CENP-L did not bind nucleosomes (Figure 3-7 B), indicating that CENP-N and CENP-L have different functions.
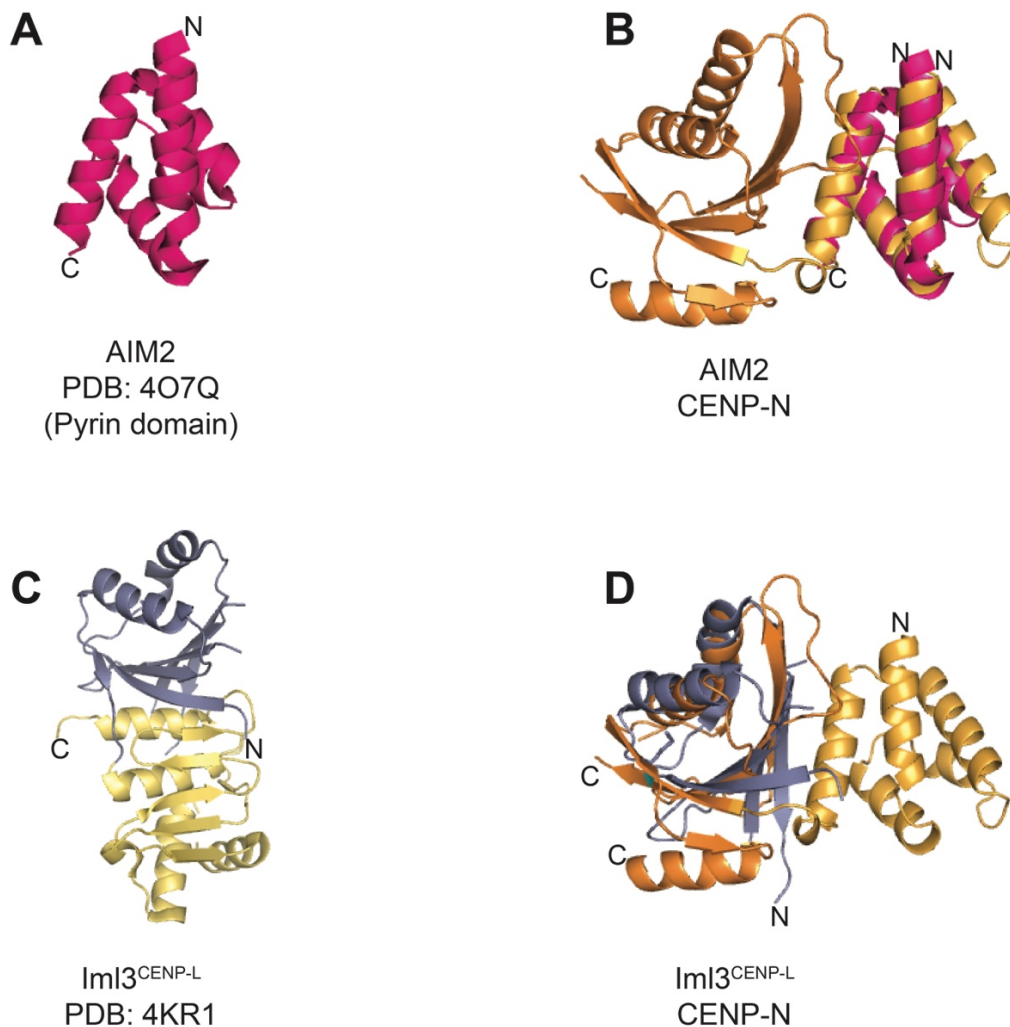

Figure 3-14 Identification of structurally similar proteins of CENP-N

(A) Cartoon representation of AIM2 (Pyrin domain containing protein) obtained from (PDB ID 4O7Q). (B) Overlay of AIM2 with CENP-N. (C) Cartoon representation of Iml3 (yeast homolog of CENP-L) obtained from (PDB ID 4KR1). (D) Overlay of Iml3 with CENP-N.

Despite our attempts to crystallize the C-terminal fragment of CENP-N in complex with CENP-L, we could not obtain crystals However, sequence alignment of CENP-N and Chl4 indicate sequence and structural conservation (Figure 3-15 A). Similarly, the sequence alignment of CENP-L and Iml3 shows a strong sequence similarity (Figure 3-15 B). Iml3 is comprised of an N-terminal domain and a C-terminal or insert domain

(Figure 3-15 C-D). Iml3 dimerizes with Chl4 through a subdomain within the insert domain (Hinshaw & Harrison 2013).
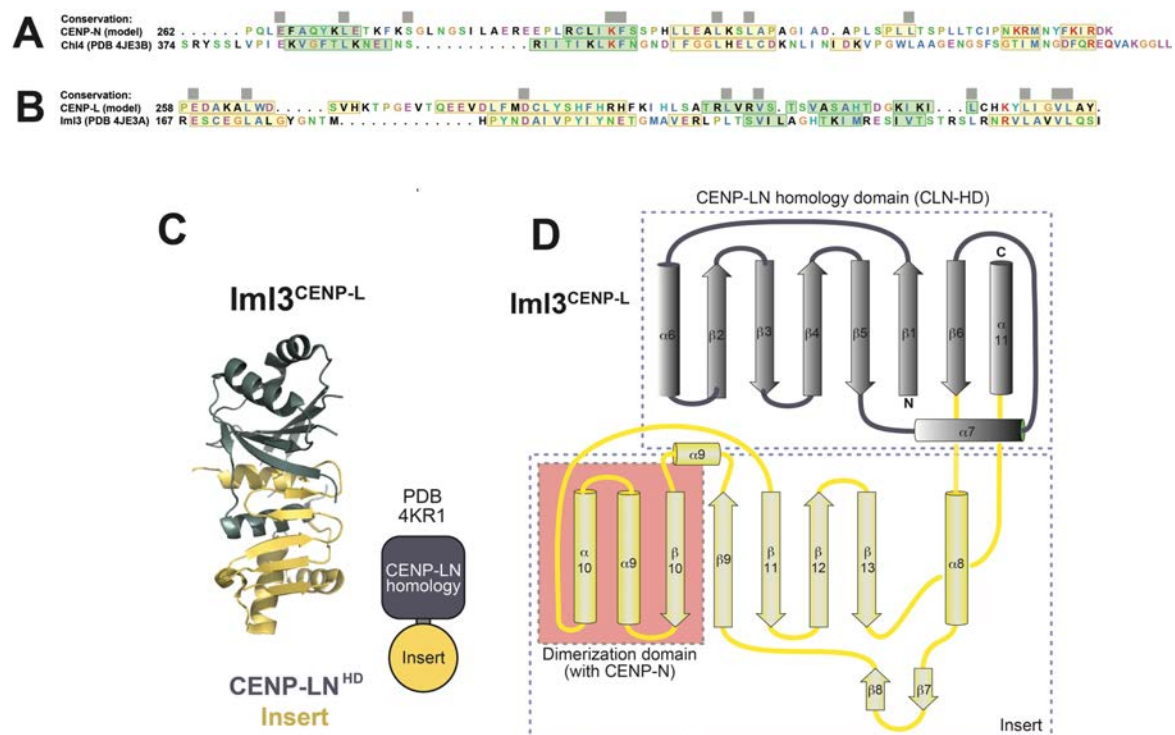

Figure 3-15 Topology of Iml3^{CENP-L}

(A)Structure based sequence alignment of the C-terminal region of CENP-N with the C-terminal region of Chl4. (B) Structure based sequence alignment of CENP-L with Iml3 (PDB ID 4KR1). (C) Overall domain organization of Iml3 with CLN-HD domain in dark grey and Insert domain in yellow. (D) Topology diagram of CENP-L based on the structure of Iml3.

Due to the strong sequence similarity of CENP-N with Chl4 and CENP-L with Iml3, the structure of CENP-N$^{1-235}$ reported in this study and the structure of Iml3:Chl4^C complex represent a comprehensive view of the structural organization of the CENP-L^{Iml3}:CENP-N^{Chl4} complex (Figure 3-16).
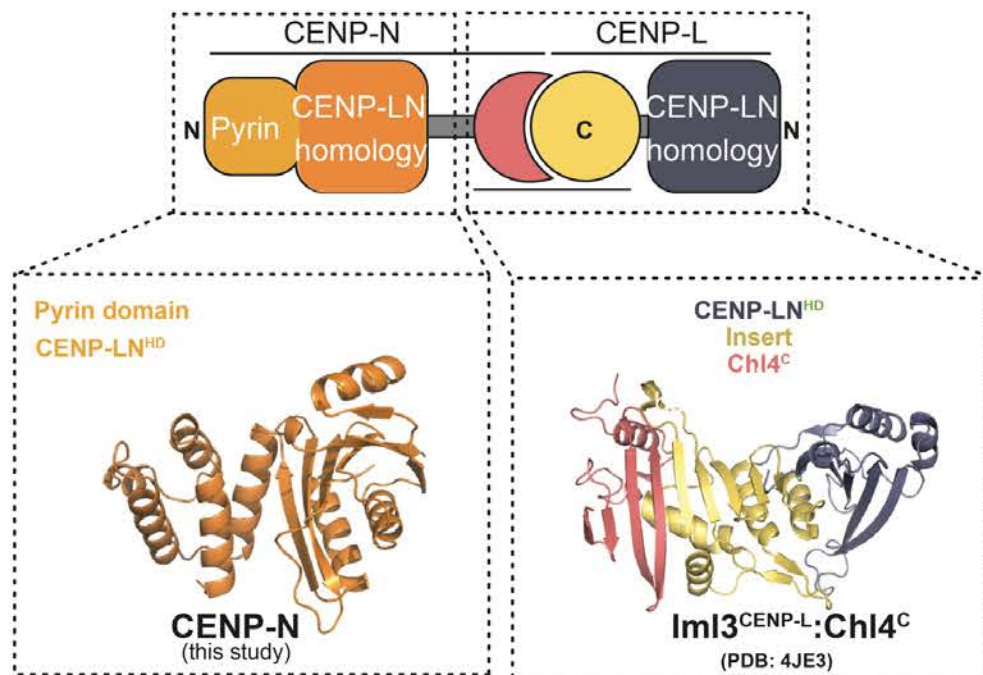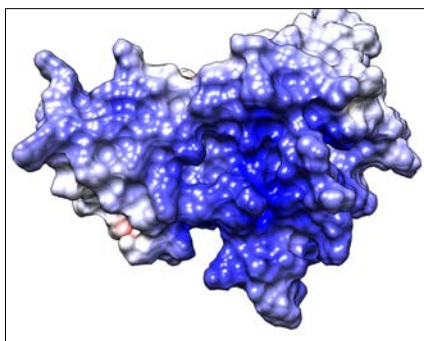
Figure 3-16 Organization of the CENP-LN complex

Schematics of the overall organization of CENP-LN complex, and complementarity of CENP-N[1-235] and Chl4[C]:Iml3 (PDB ID 4JE3) crystal structures.

Both the PYD and CLN-HD domains consist of a series of positively charged exposed residues (Arg^{42}, Arg^{44}, Lys^{148} and Arg^{196}). Moreover, mapping the electrostatic potential of the surface of CENP-N revealed a wide positively charged surface on one side of the protein. Many DNA-binding proteins consists of positively charged residues so that they can contact negatively charged DNA (Figure 3-17). Taken together, these observations suggest that CENP-N might interact with the DNA via its positively charged residues, which is discussed in detail in section 3.6.1.

Figure 3-17 Electrostatic potential of CENP-N

Electrostatic potential of CENP-N, displaying the charged surface (blue) with contour levels $\pm$ 4 k_BT/e (k_B, Boltzmann constant; T, absolute temperature; e, the magnitude of electron charge, calculated with the APBS Pymol plugin). CENP-N clearly displays a charged surface on one side, suggesting a possible involvement in DNA binding.

3.6 Cryo-EM structure of CENP-N:CENP-A nucleosome complex

Following the structural determination of CENP-N, several attempts were made to determine the structure of the CENP-N:CENP-A^{NCP} complex. During these ongoing investigations, we became aware that Keda Zhou from Prof. Karolin Lugers lab at the University of Colorado, United States, was also working on the structural determination of the same complex. Therefore, our teams collaborated in order to determine the structure of CENP-N bound to CENP-A^{NCP}.

3.6.1 Cryo-EM analysis

The structure of the CENP-N:CENP-A^{NCP} complex was determined using cryo-electron microscopy at a resolution of ~4.0 Å (Figure 3-18 and 3-19). One of the peculiar feature of this sample was the stacking of nucleosomes (lateral interactions between the nucleosomes) observed in raw images. The reasons for this stacking are not fully understood. A representative micrograph post-CTF (contrast transfer function) correction is depicted in Figure 3-18 A. After the particles were picked, they were classified into different 2D class averages as depicted in Figure 3-18 B. The cryo-EM structure of CENP-N:CENP-A^{NCP} complex was reconstructed, and the best three-dimensional class was selected and refined at 4.0 Å resolution (Figure 3-18 C-D). Data collection and the refinement statistics are presented in Table 3-2.
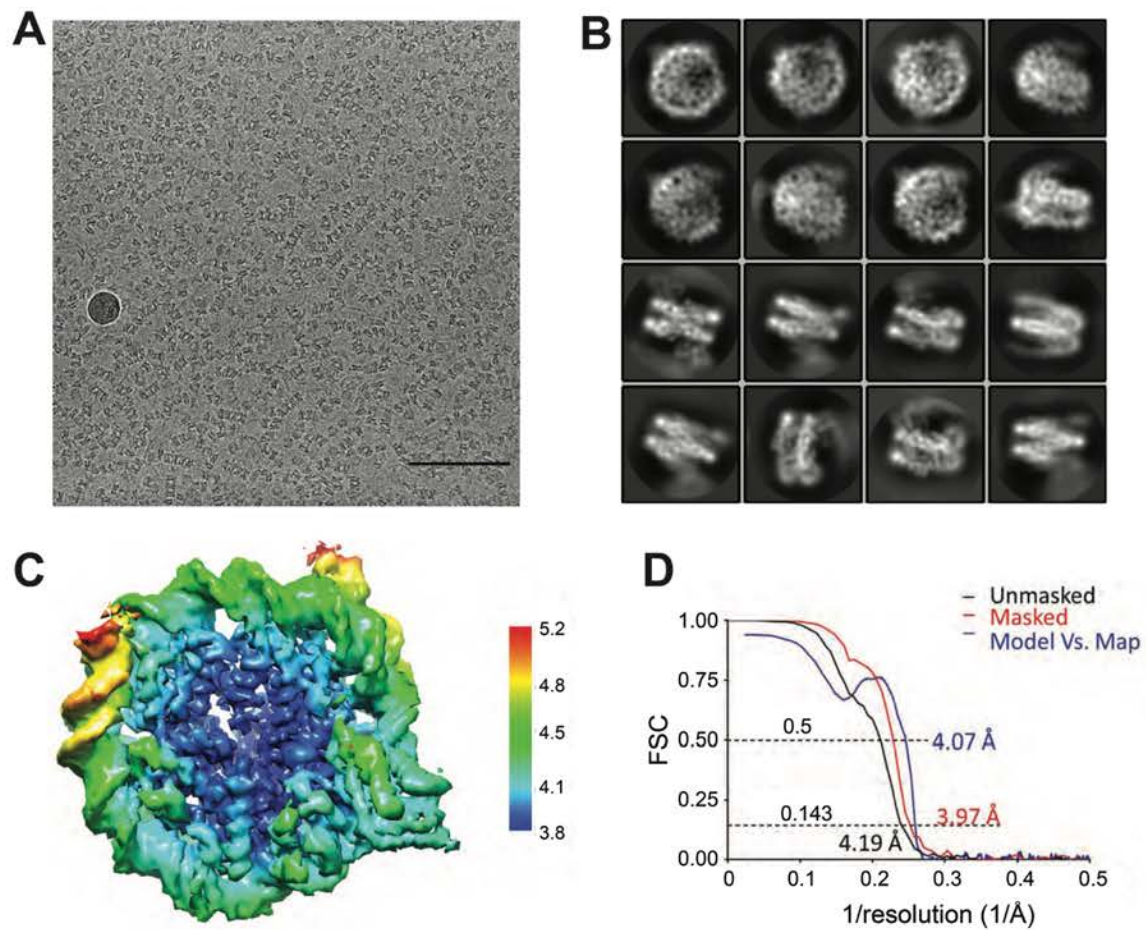
Figure 3-18 Cryo-EM analysis of CENP-N:CENP-A^{NCP} complex

(A) Representative micrograph after CTF correction at -2.5 μm defocus. Scale bar = 100 nm. (B) Representative 2D class averages from RELION 2D classification. (C) Fourier shell correction (FSC) curve for the EM maps. (D) Estimated local resolution for CENP-N$^{1-289}$:CENP-A^{NCP} by RELION. Unit for colour scale is Å.

Table 3-2: EM data collection, processing and refinement statistics

EM data collection, processing and refinement statistics	
Voltage (kv)	300
Magnification	290,000x
Defocus (μm, nominal)	-1.0 to -2.5
Pixel size (Å)	1.02
Electron dose rate (counts/pixel/s)	10
Total electron dose (e$^-$/ Å^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) (Å)	4.0/4.2
Map sharpened b-factor ($Å^2$)	-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

The structure of the CENP-N:CENP-A^{NCP} complex was deduced by fitting the high-resolution structures of CENP-A histone core (PDB ID 3AN2) (Tachiwana et al. 2011), combined with DNA derived from a nucleosomes reconstituted with the 145 bp of 601 DNA sequence (PDB ID 3LZ0) (Vasudevan et al. 2010) and the newly determined crystal structure of CENP-N$^{1-235}$, into the cryo-EM density maps. Though the CENP-N$^{1-289}$ construct had been used to determine the structure of the CENP-N:CENP-A^{NCP} complex, there is clear density only until residues ~210, indicating that the last ~79 C-terminal residues are flexible. Comparisons between the free and CENP-N in complex with CENP-A^{NCP} reveals that the CENP-N structure changes very little upon binding to CENP-A^{NCP}. The EM structure of CENP-N:CENP-A^{NCP} complex reveals that the CENP-N binds to CENP-A^{NCP} by engaging with both CENP-A histone and DNA (Figure 3-19). CENP-N interacts extensively with DNA from bp -21 to -35 relative to the twofold axis, or super helical location (SHL) -2 to -3 (Figure 3-20 A). Of the ~2,400 Å^2 of CENP-N and CENP-A^{NCP} surface that become buried in the complex, ~1,400 Å^2 was at the interface of CENP-N:DNA. CENP-N utilizes both its domains, PYD and CLN-HD, to interact with DNA. Four loops within the CLN-HD domain of CENP-N straddle the DNA double helix over ~8 bp, and the following 7 bp are bound by the PYRIN domain. The positively charged residues of CENP-N (Arg^{42}, Arg^{44}, Lys^{148} and Arg^{196}), which was discussed earlier, are located at this DNA binding interface, as depicted from the electrostatics of CENP-N:DNA binding interface (Figure 3-20 A).
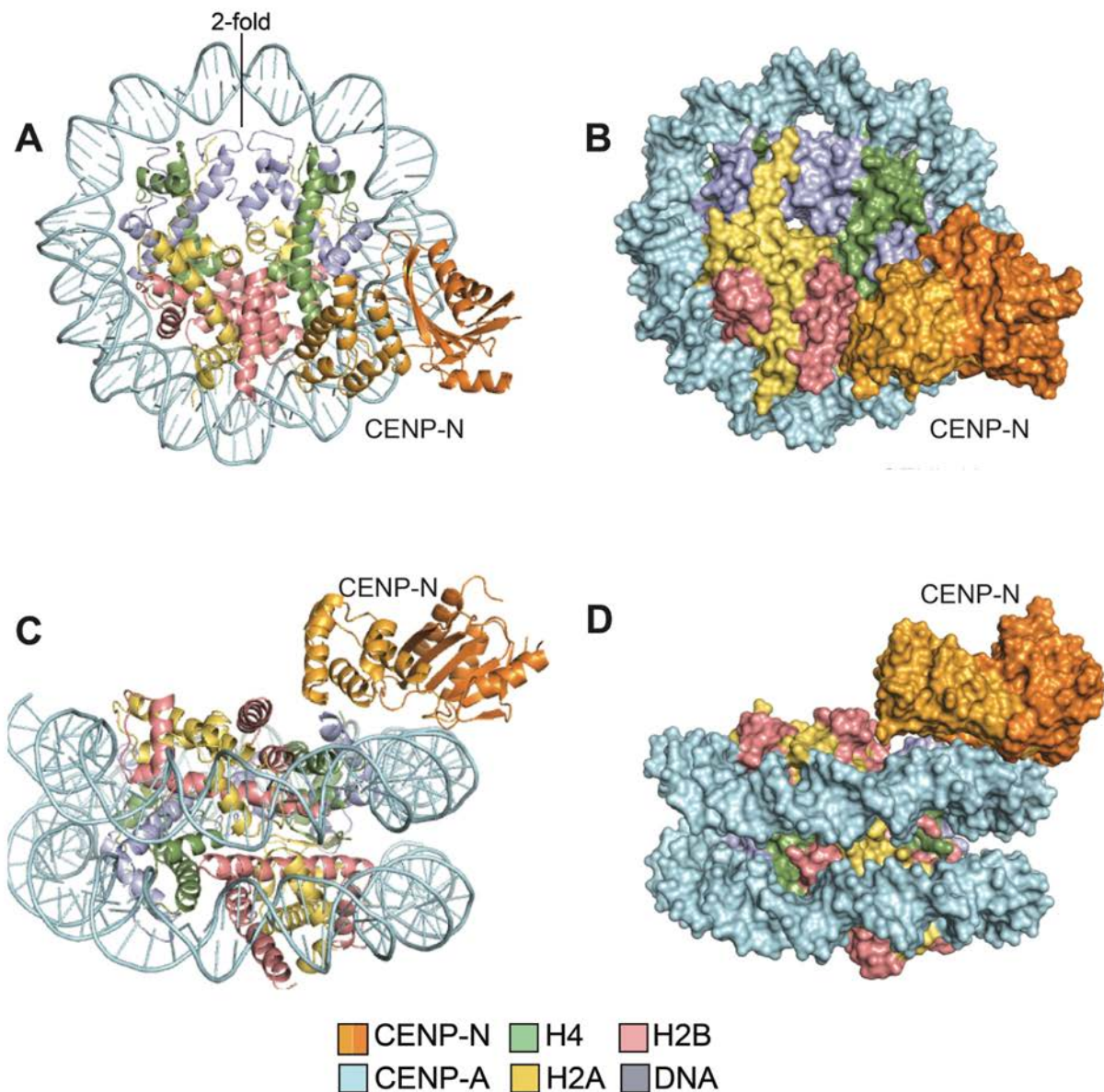
Figure 3-19 Cryo-EM structure of CENP-N:CENP-A^{NCP} complex

(A) Cartoon representation of CENP-N:CENP-A^{NCP} complex in front view (B). Surface representation of the CENP-N:CENP-A^{NCP} complex in front view. (C) Cartoon representation of CENP-N:CENP-A^{NCP} complex in top view (D) Surface representation of the CENP-N:CENP-A^{NCP} complex in top view.

One of the important features of the intermolecular contacts that bind DNA to the histones is the series of histone arginine residues that are inserted into the DNA minor groove. Similarly, some of the side chains of the residues from both the PYRIN domain (Arg^{44}) and CLN-HD (Lys^{148}, Met^{167}, Arg^{170}) are inserted into the minor grooves of DNA (Figure 3-20 A). The highly conserved Pro^{17} within the α2 helix of CENP-N is oriented in a way that allows interactions of the sidechains of residues Lys^{15}, Arg^{42}, Lys^{45} and Arg^{194} with DNA. Collectively, these observations explain why CENP-N failed to bind CENP-A/H4 tetramers lacking DNA (Carroll et al. 2009). Although CENP-N binds

exclusively to CENP-A^{NCP}, it retains a substantial affinity to H3^{NCP}, possibly due to its interactions with DNA. Collectively, these results suggest that DNA binding clearly enhances the binding affinity of CENP-N to nucleosomes; however, it is unlikely to contribute towards the specificity of CENP-A^{NCP} over canonical H3^{NCP}, as CENP-N bound selectively to CENP-A^{NCP} even when both CENP-A^{NCP} and H3^{NCP} were wrapped with the same DNA. Moreover, the binding of CENP-N seems to be independent of the DNA sequence, as CENP-N bound to the 601 DNA sequence *in vitro* (Figure 3-8 A).

3.6.2 CENP-N:CENP-A^{NCP} interface

Previous studies have demonstrated that CENP-N binds to a specialized region within CENP-A, known as the CATD region (centromere-targeting domain) (Carroll et al. 2009; Black, Jansen, et al. 2007). Corroborating the previous findings, our structure clearly demonstrates how and why the L1 loop of CENP-A is crucial for the binding selectivity of the CENP-A^{NCP} by CENP-N. CENP-N interacts with the L1 loop via a series of residues from the α1 helix in CENP-N^{PYD} and the β3$-\beta$4 loop in CENP-N^{CLN-HD}, suggesting that both domains of CENP-N$^{1-235}$ are involved in this specific interaction. At the interface of the L1 loop lie the conserved and solvent-exposed residues of CENP-N, including Glu^3, Glu^7 and Arg^{11} from CENP-N^{PYD} and Lys^{143}, Pro^{145}, Gln^{146} and Lys^{148} from CENP-N^{CLN-HD} (Figure 3-21). Tyr^{147}, located within the CLN-HD domain of CENP-N, contributes to the stabilization of both the domains of CENP-N$^{1-235}$. Interaction with the L1 loop of CENP-A, involves three residues within the α1 helix of CENP-N, Glu^3, Glu^7 and Arg^{11}, whose sidechains come in contact with the L1 loop of CENP-A (Figure 3-20 B). Sequence alignment of histones CENP-A and H3 revealed the insertion of two residues, Arg^{80} and Gly^{81}, within the L1 loop of CENP-A (Figure 3-20 B). Ectopic targeting experiments on CENP-N have demonstrated that both Arg^{80} and Gly^{81} of CENP-A are required for localization of CENP-N, as mutations within these two residues leads to loss of recruitment of CENP-N (Fang et al. 2015).
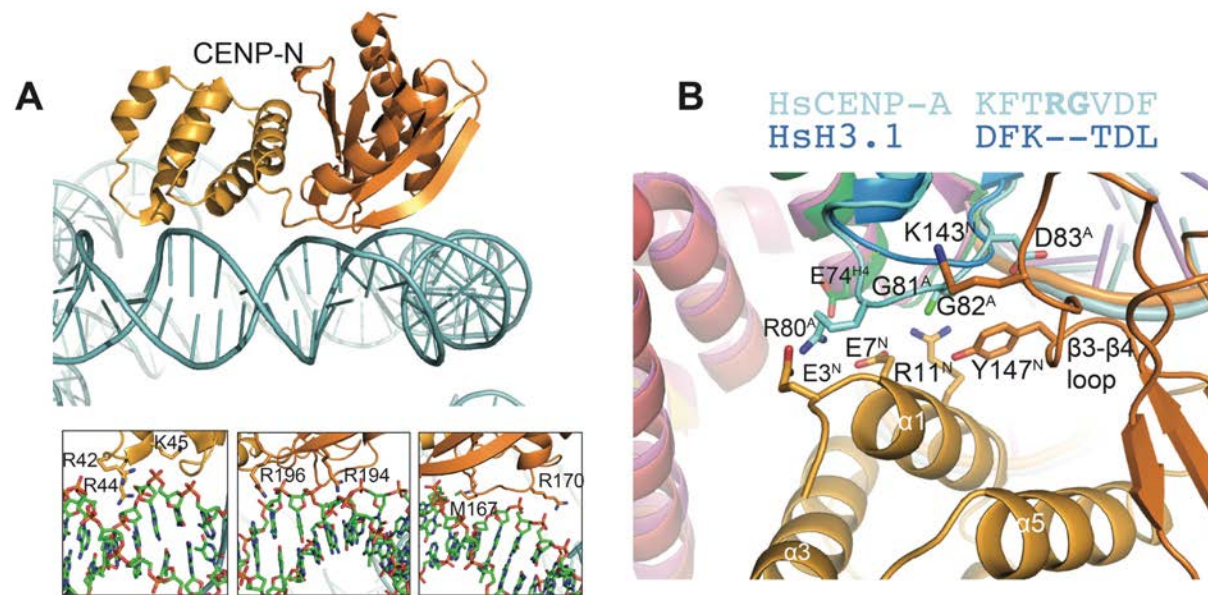
Figure 3-20 Structural features of CENP-N:CENP-A^{NCP} complex

(A) Interactions of CENP-N with DNA backbone along with close-up views of selected interactions of CENP-N with major and minor groves of DNA. (B) Interactions of CENP-N with the L1 loop of CENP-A and comparison with the superimposed H3.

Our structure clearly demonstrates why the insertion of Arg^{80} and Gly^{81}, within the L1 loop of CENP-A is crucial, as it allows Arg^{80} to form hydrogen bonds with Glu^{3} and Glu^{7} of CENP-N (Figure 3-20), while the absence of a side chain at Gly^{81} of CENP-A allows the L1 loop of CENP-A to insert deeply into a space formed between the two domains of CENP-N. The side chain of Tyr^{147} within the β3–β4 loop of CENP-N comes in close vicinity to Val^{82} of CENP-A (Figure 3-20 B). Previous studies on CENP-N have demonstrated that Arg^{11} of CENP-N is important for the interaction with CENP-A^{NCP}, as substitution of Arg^{11} to Ala resulted in the complete loss of the binding of CENP-N to CENP-A (Figure 3-21) (Carroll et al. 2009). The structure implies that Arg^{11} of CENP-N is inserted between the L1 loop of CENP-A and L2 loop of H4, where it might be involved in a double salt bridge with Glu^{74} of H4 and Glu^{7} of CENP-N (Figure 3-20 B). Taken together, these results suggest that CENP-N utilizes both its PYD and CLN-HD domains to interact with the CATD of CENP-A, and with the nucleosomal DNA. Moreover, while the CATD region confers specificity, the binding of DNA increases the binding affinity of CENP-N to CENP-A^{NCP}.
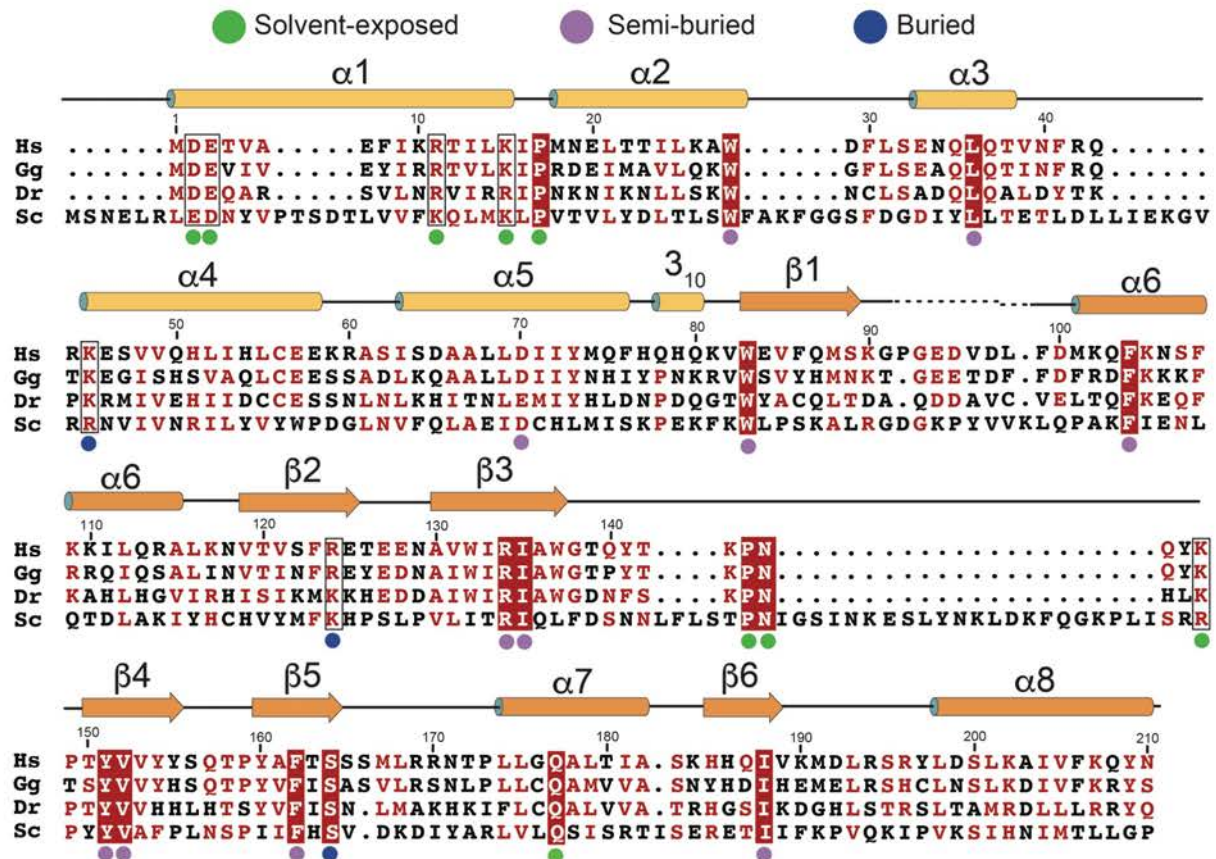
Figure 3-21 Multiple sequence alignment of CENP-N

Multiple sequence alignment of CENP-N from different species with secondary structure displaying solvent-exposed (green), semi-buried (purple) and buried residues (blue).

Additional features that were observed in our structure were that the binding of CENP-N to DNA seems to stabilize the DNA ends, as in the structure a clear density was observed for 139 of the 147 bp of DNA (Figure 3-22 A-B). In addition, the ordering of the flexible N-terminal helix of CENP-A was also observed (Figure 3-22 B). Both these features were largely disordered or invisible in previous structural studies on CENP-A^{NCP} (Figure 3-22) (Tachiwana et al. 2011). Taken together, these results clearly suggest that the binding of CENP-N to CENP-A^{NCP} may stabilize CENPA^{NCP}. In addition to this, we have also observed that the N-terminal region of H4 is clearly ordered, possibly due to its tight packing against the β3-β4 loop of CENP-N (Figure 3-22 C).
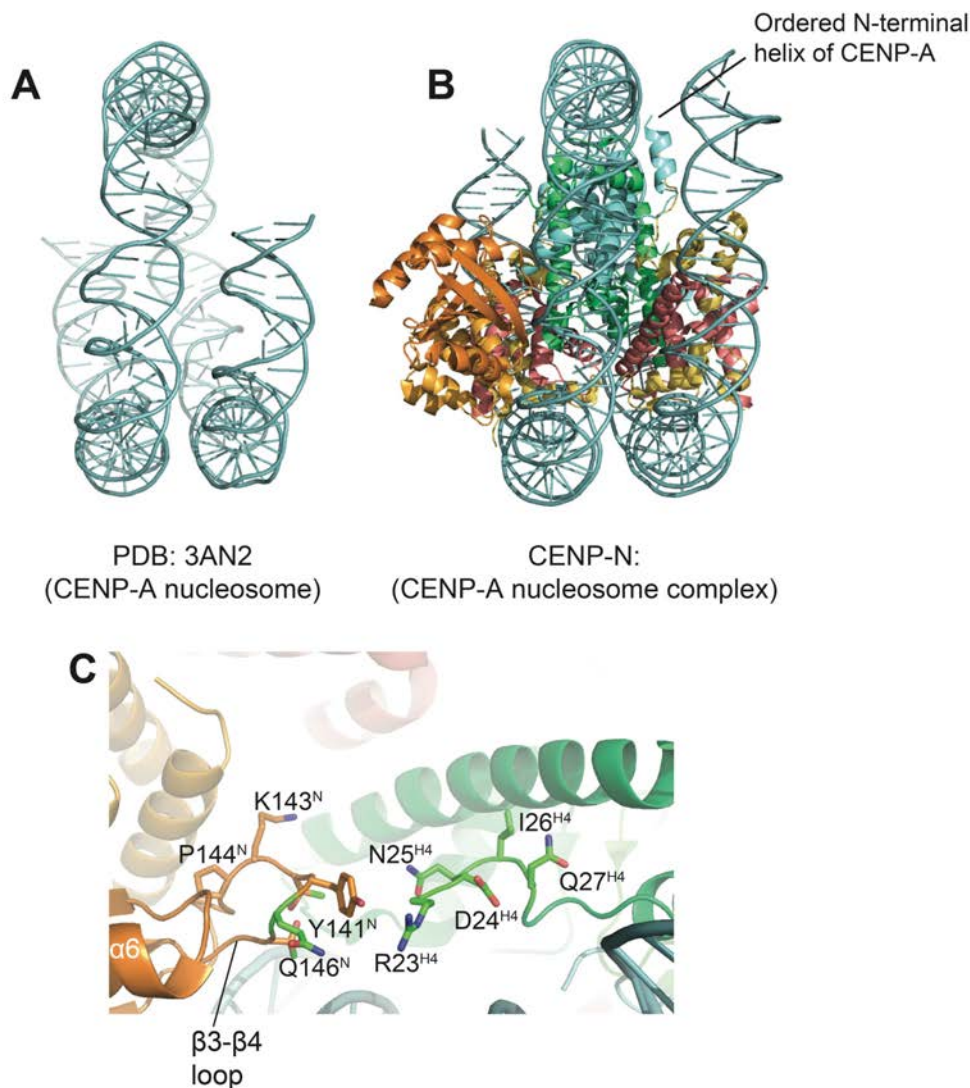
Figure 3-22 Additional features of the cryo-EM structure of CENP-N:CENP-A^{NCP} complex

Comparison of the DNA ends in the crystal structure of CENP-A^{NCP} (A) and the newly determined structure of CENP-N:CENP-A^{NCP} complex (B) showing the ordered N-terminal helix of CENP-A. (C) Zoom-in-view of the interface between the CENP-N and the ordered H4 N-terminal region starting at residue Arg[23].

3.7   Mutational validation of CENP-N:CENP-A^{NCP} complex

To delineate the residues involved in the interaction of CENP-N with CENP-A^{NCP}, I generated a series of CENP-N mutants based on the crystal structure of CENP-N and the cryo-EM structure of the CENP-N:CENP-A^{NCP} complex. In addition, ConSurf (a tool used to estimate the evolutionary conservation of the amino acids of the proteins based on orthologous sequences) analysis was performed to identify conserved residues exposed on the surface of CENP-N (Figure 3-23 A). To test the effects of the single alanine mutants of CENP-N in binding to CENP-A^{NCP}, GST-pulldown assays were performed, in which GST-CENP-N (WT or mutants) were used as bait and CENP-A^{NCP}

served as prey. Alanine mutations of CENP-N residues Lys[15] and Lys[45], which reside at the DNA-binding interface, and Tyr[147] and Glu[7]-Tyr[147], which are located in close proximity to the L1 loop of CENP-A, showed slightly decreased binding to CENP-A^{NCP} when compared with the wildtype CENP-N (Figure 3-23 B). To determine if reduced binding to CENP-A has any effect on the kinetochore localization of CENP-N, transient transfections of both CENP-N wildtype and alanine mutants were performed in asynchronous U2OS cells. *In vivo* transfections of CENP-N single mutants (Lys[15], Lys[45], Tyr[147] and Glu[7] - Tyr[147]) resulted in lower levels of CENP-N at the centromeres in comparison to CENP-N wildtype, corroborating the *in vitro* results (Figure 3-23 C). These results demonstrate that the above described single alanine mutants or a double mutant of Glu[7]-Tyr[147] of CENP-N have limited effect on the binding to CENP-A^{NCP}.
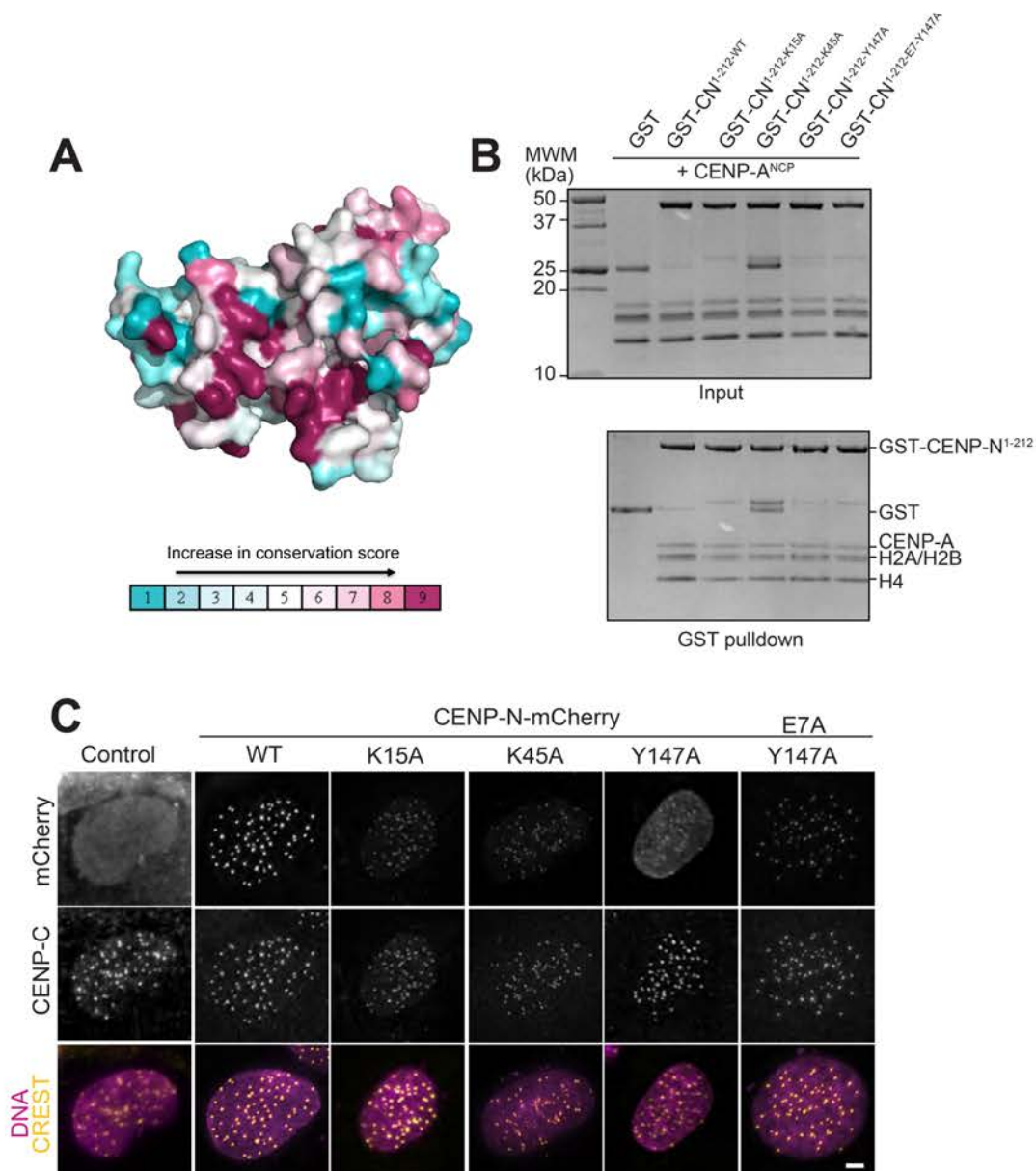
Figure 3-23 Validation of CENP-N: CENP-A^{NCP} complex

(A) ConSurf analysis of CENP-N crystal structure in surface representation. Conservation of CENP-N using the color-coded bar shown in figure. (B) *In vitro* GST-pulldown assays of CENP-N (WT and mutants) immobilized on solid phase with CENP-A^{NCP}. (C) Transient transfections of a wildtype CENP-N-mCherry and of its mutant variants.

To obtain a more penetrant phenotype, a series of double mutants of CENP-N (Lys^{15} - Tyr^{147A} and Lys^{45} - Tyr^{147A}) along with a single alanine mutant Arg^{11} were generated (Carroll et al. 2009). GST-pulldown assays using CENP-N$^{K15-Y147A}$, CENP-N$^{K45-Y147A}$ or CENP-N^{R11A} mutants resulted in a complete loss of binding of CENP-N to CENP-A^{NCP} (Figure 3-24 A). Therefore, combining the mutations (Lys^{15}- Tyr^{147A} and Lys^{45}- Tyr^{147A}) has an additive effect on the nucleosome binding *in vitro*. Kinetochore localization studies of the CENP-N double mutants severely impaired the ability to target to centromeres (Figure 3-24 B-C). Taken together, our results clearly demonstrate that the CENP-N binds to the L1 loop of CENP-A through the α1 helix in CENP-N^{PYD} and the β3$-\beta$4 loop in CENP-N^{CLN-HD} domain. Specific mutations of the residues that reside at these interfaces (CENP-N:CENP-A/CENP-N:DNA) lead to severe defects in the kinetochore localization of CENP-N. Collectively, these results suggest that the kinetochore localization of CENP-N depends on CENP-A^{NCP}, which is in agreement with the previous studies (Carroll et al. 2009).
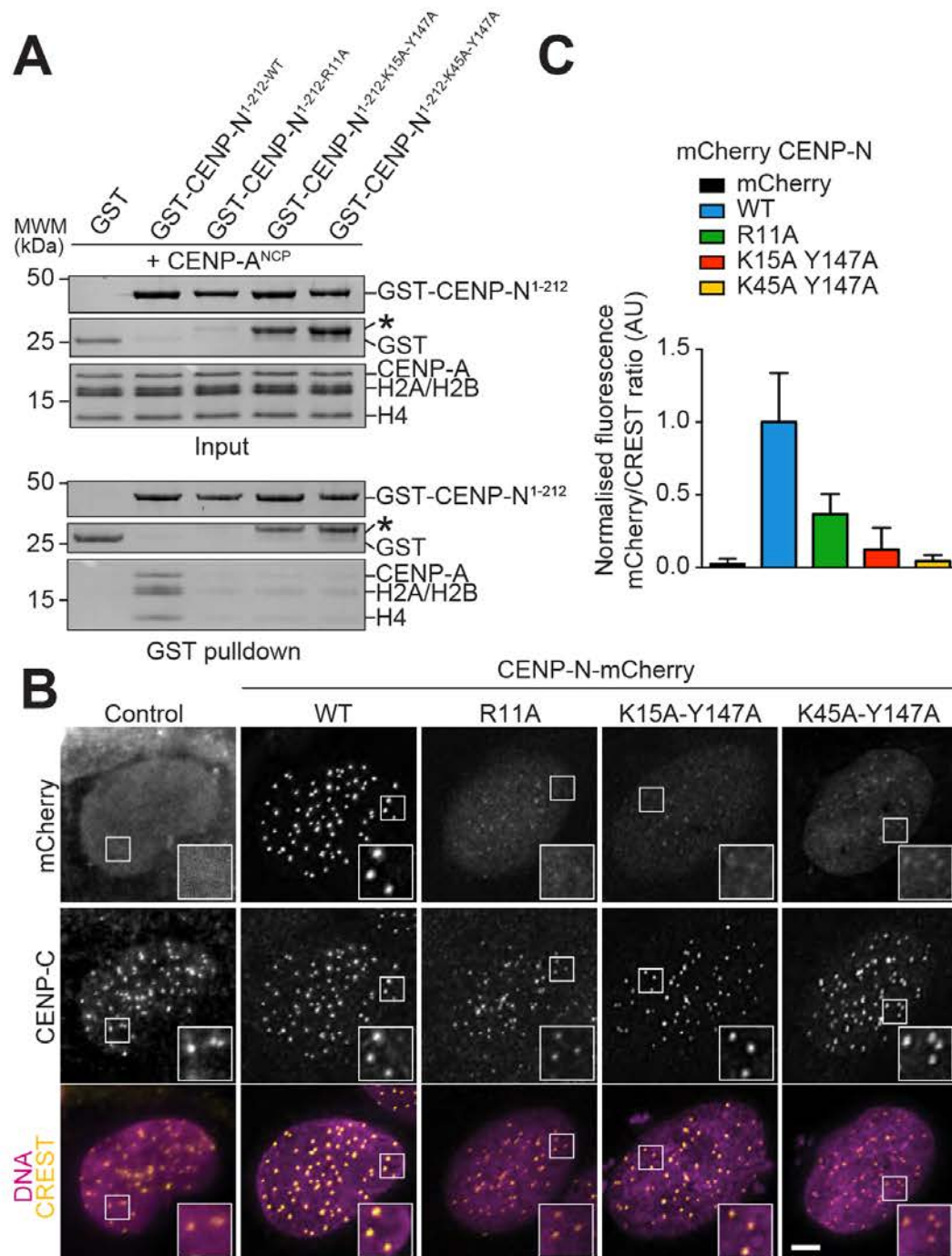
Figure 3-24 Kinetochore localization of CENP-N depends on CENP-A

(A) *In vitro* GST-pulldown assays of CENP-N (WT and mutants) immobilized on solid phase with CENP-A^{NCP}. (B) Transient transfections of a wild-type CENP-N-mCherry and of its mutant variants. (C) Quantifications of localization of CENP-N-mCherry constructs both wildtype and mutants. Error bar represents SD.

3.8 Recombinant CENP-LN complex directly binds CENP-C

As described in the introduction, CCAN is a network of interacting proteins. The CENP-LN complex is proposed to bind CENP-C, but the molecular details of this interaction remain unclear. Previous studies on CENP-C propose a central role for CENP-C in kinetochore assembly. Therefore, it is important to understand the interaction of CENP-C with CENP-LN complex, given the fact that they both bind CENP-A^{NCP} (Weir et al. 2016; Klare et al. 2015; Carroll et al. 2010). CENP-C (943 residues in humans) is an intrinsically disordered protein that harbors binding sites for different kinetochore proteins and for CENP-A^{NCP}. The N-terminal region of CENP-C contains a binding site for the Mis12 sub-complex of the KMN network, while the C-terminal region contains a dimerization domain known as the Cupin domain (Screpanti et al. 2011; Trazzi et al. 2009). The central motif and the CENP-C motif of CENP-C are involved in the binding to CENP-A^{NCP} (Kato et al. 2013). In addition to CENP-A^{NCP}, CENP-C is also involved in the binding to CENP-HIKM complex and CENP-LN complex (Klare et al. 2015; McKinley et al. 2015; Nagpal et al. 2015; Weir et al. 2016) (Figure 3-25). In order to map the interaction domain of CENP-C with the CENP-LN complex, different recombinant versions of both CENP-LN and CENP-C were tested for their ability to bind to CENP-C or CENP-LN complex respectively.
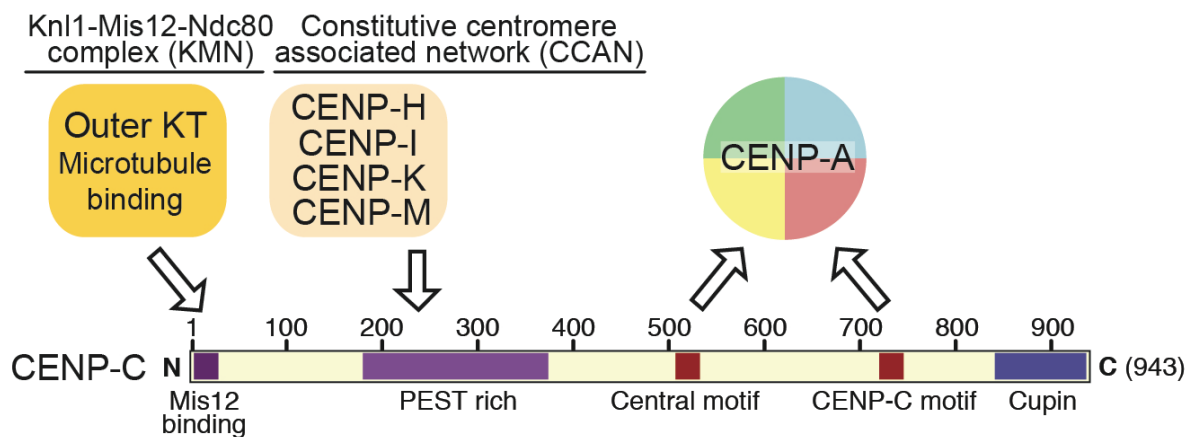

Figure 3-25 Schematic representation of the domain organization of human CENP-C

CENP-C interacts with the Mis12 complex within the KMN network at its N-terminus, CENP-HIKM complex within the PEST rich domain and CENP-A^{NCP} via central and CENP-C motif.

SEC analysis of CENP-LN complex with CENP-C$^{2-545}$ indicates a direct binding of CENP-C$^{2-545}$ to CENP-LN complex, made evident by the co-elution of the complex from a Superdex 200 column (Figure 3-26 A) (Weir et al. 2016; Nagpal et al. 2015; McKinley et al. 2015). GST-pulldown assays were performed using GST-CENP-LN complex as bait, which was incubated with either CENP-C$^{1-71}$ or CENP-C$^{2-545}$ that served as prey. As shown in Figure 3-26 B, CENP-LN complex directly bound to CENP-C$^{2-545}$ but not CENP-C$^{1-71}$, indicating that the binding site of CENP-LN on CENP-C is between residues 72-545 of CENP-C. We did not observe any binding of CENP-LN complex with a construct encompassing residues 545-943 of CENP-C (Data not shown).
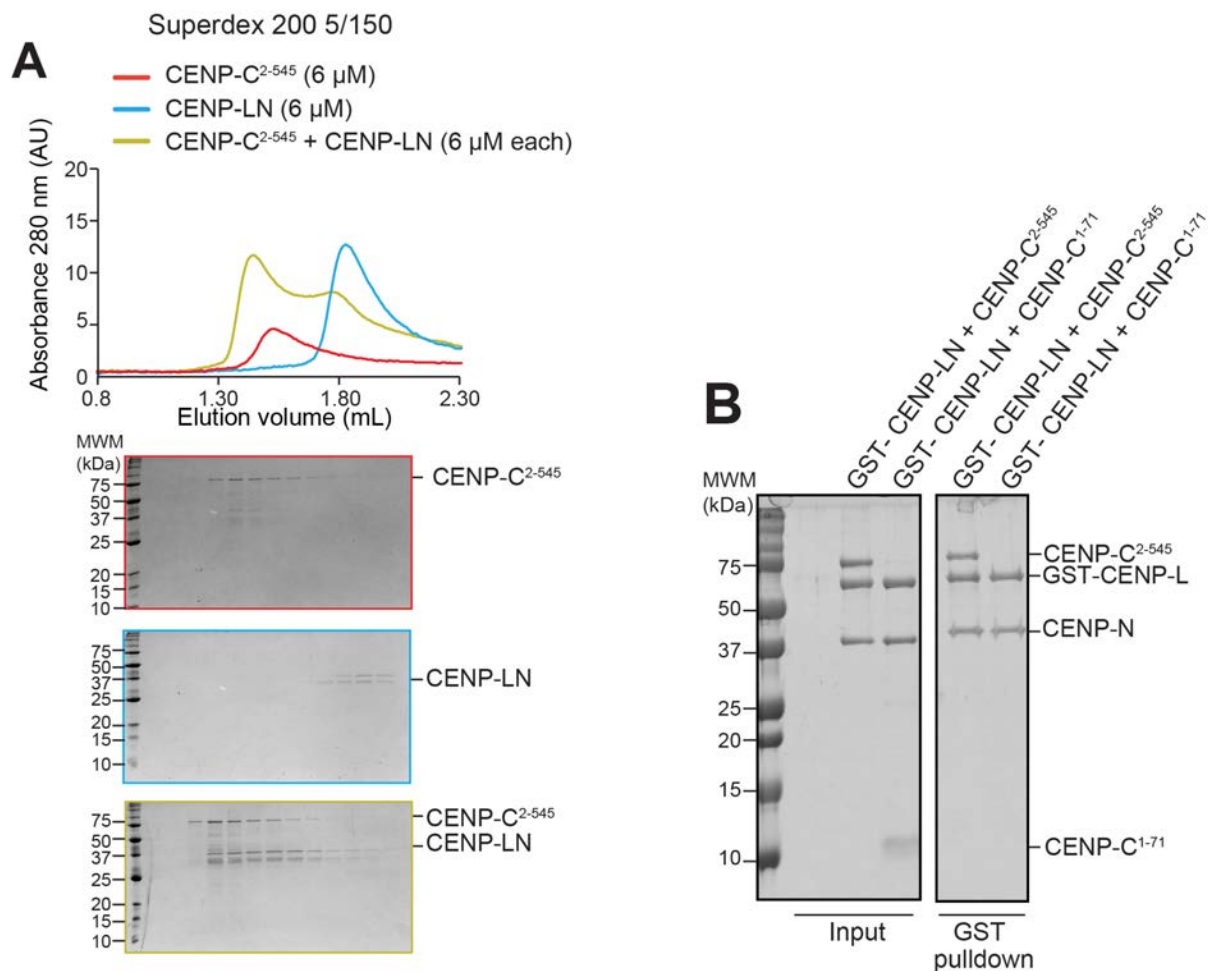

Figure 3-26 CENP-LN binds CENP-C directly

(A) Size exclusion chromatography (SEC) elution profile of CENP-LN complex with CENP-C$^{2-545}$ both at 6 µM. CENP-C$^{2-545}$ and CENP-LN forms a stoichiometric complex and co-elute in analytical SEC (B) GST-CENP-LN complex (bait) was immobilized on beads and incubated with either CENP-C$^{1-71}$ or CENP-C$^{2-545}$ (prey).

3.9 C-terminal region of CENP-N in complex with CENP-L binds CENP-C

To precisely define the domain of CENP-LN involved in the interaction with CENP-C, the binding of CENP-C$^{2\text{-}545}$ was tested with different constructs of recombinant CENP-LN. Initially, the nucleosome binding fragment of CENP-N was tested, and SEC assays revealed that CENP-N$^{1\text{-}235}$ does not bind CENP-C$^{2\text{-}545}$ (Figure 3-27). This suggests that the binding sites of CENP-C and CENP-A nucleosomes are distinct and non-overlapping on the CENP-LN complex. In contrast, Mif2 (yeast homolog of CENP-C) was implicated in the interaction with the N-terminus of Chl4 within the Chl4/Iml3 (CENP-N/CENP-L) complex suggesting that the interaction interfaces might be different from yeast to humans (Hinshaw & Harrison 2013). Despite several attempts, I was not able to purify the C-terminal fragment of CENP-N, due to its instability. Therefore, I could not perform any direct binding experiments with the C-terminus of CENP-N and CENP-C.
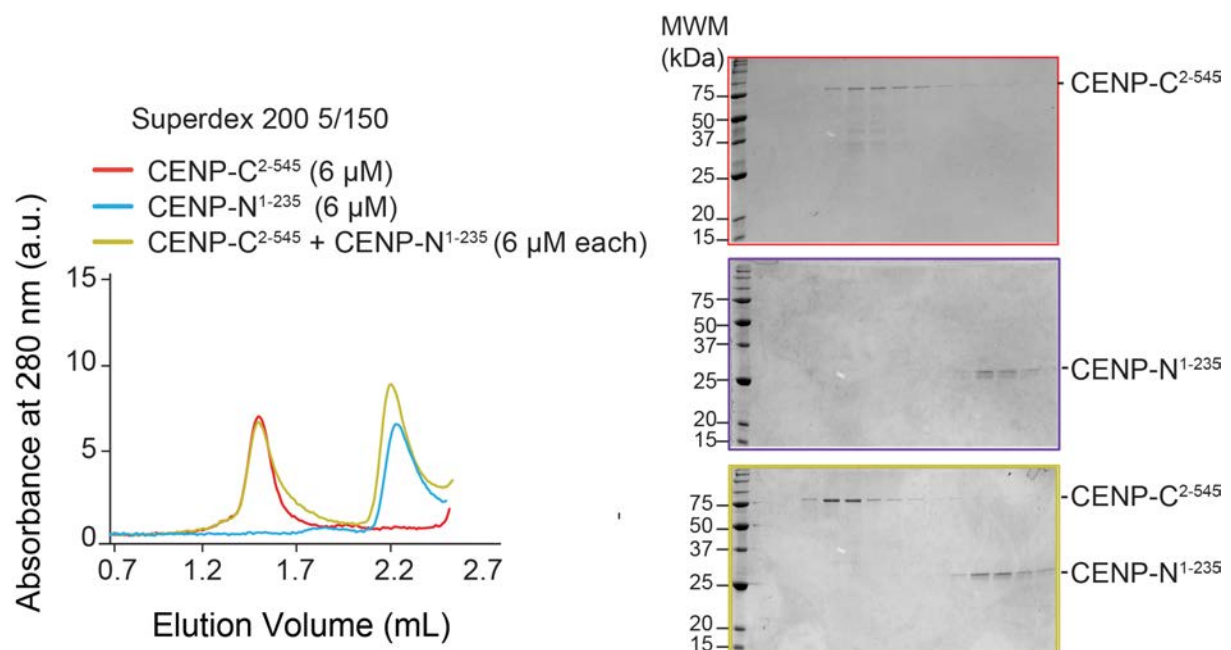

Figure 3-27 CENP-N$^{1\text{-}235}$ doesn't bind CENP-C

SEC elution profile of CENP-N$^{1\text{-}235}$ with CENP-C$^{2\text{-}545}$ both at 6 µM. CENP-N$^{1\text{-}235}$ doesn't bind CENP-C$^{2\text{-}545}$ as is evident from two distinct peaks.

However, co-expression of the C-terminus of CENP-N (CENP-N$^{230\text{-}C}$) in complex with CENP-L resulted in a soluble and stable product. When CENP-LN$^{230\text{-}C}$ was incubated with CENP-C$^{2\text{-}545}$, they co-eluted in a stoichiometric complex in SEC, indicating a direct interaction (Figure 3-28 A). In a GST-pulldown assay, GST-CENP-LN$^{230\text{-}C}$ complex

bound CENP-C$^{2-545}$ in a manner equivalent to that of GST-CENP-LN^{FL} (which served as a positive control), indicating that CENP-LN^{230-C} is sufficient for a tight interaction with CENP-C (Figure 3-28 B).
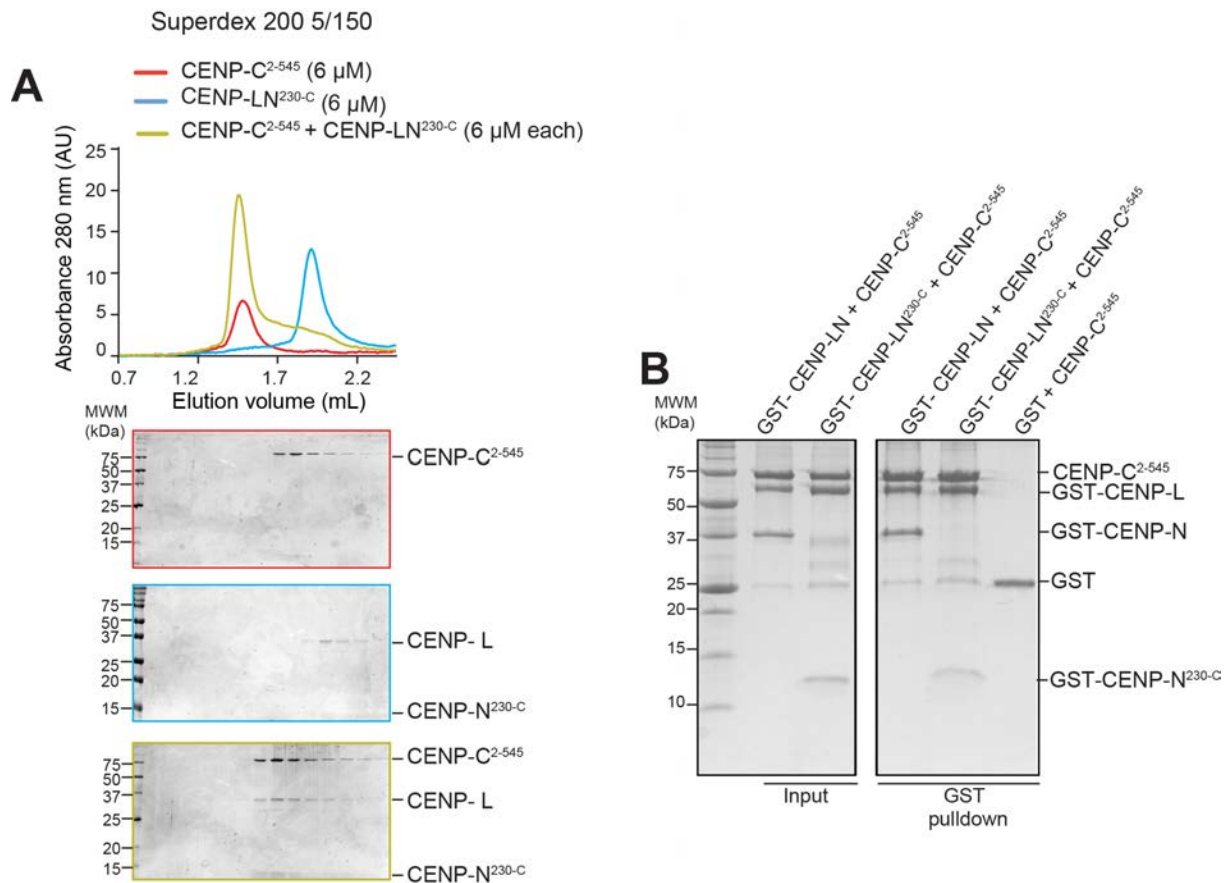

Figure 3-28 CENP-LN^{230-C} directly binds CENP-C$^{2-545}$

(A) SEC elution profile of CENP-LN^{230-C} with CENP-C$^{2-545}$, both at 6 µM. CENP-LN^{230-C} co-elutes in a stoichiometric complex with CENP-C$^{2-545}$, indicating the formation of a complex. (B) GST-CENP-LN^{FL} or CENP-LN^{230-C} baits were immobilized on beads and incubated with CENP-C$^{2-545}$.

Since CENP-LN^{230-C} bound CENP-C$^{2-545}$, we wanted to determine whether CENP-L^{FL} alone binds to CENP-C$^{2-545}$. SEC analysis of CENP-L^{FL} and CENP-C$^{2-545}$ indicated no binding as depicted in Figure 3-29. Although, CENP-L^{FL} cannot bind CENP-C$^{2-545}$, we cannot exclude the likely involvement of the dimerization domain of CENP-L with CENP-N in mediating the interaction with CENP-C. Although CENP-L^{FL} neither binds CENP-A^{NCP} nor CENP-C$^{2-545}$, it is still required to provide stability to CENP-N, as the C-terminal region of CENP-N is unstable without CENP-L. Collectively, these results demonstrate that the N-terminal region of CENP-N is involved in the interaction with CENP-A^{NCP}, while the C-terminal region of CENP-N in complex with CENP-L is

involved in the interaction with CENP-C, suggesting that both ends of CENP-N are indispensable for proper kinetochore assembly.
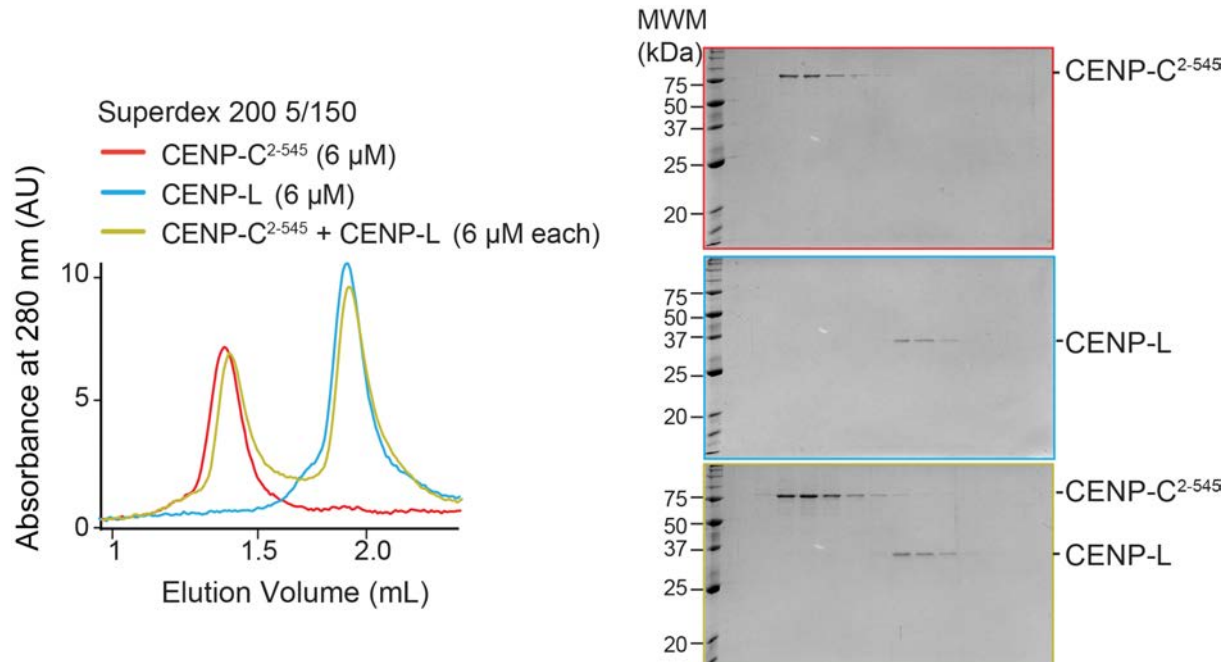

Figure 3-29 CENP-L does not bind CENP-C

SEC elution profile of CENP-L with CENP-C$^{2-545}$, both at 6 µM. CENP-L doesn't bind CENP-C$^{2-545}$, as is evident from the two distinct peaks

3.10 CENP-LN complex directly binds the PEST domain of CENP-C

Further dissection of the binding region of CENP-C on CENP-LN complex is not amenable because truncation of the C-terminal residues of CENP-N yields insoluble product, even when co-expressed with CENP-L. Therefore, I shifted my focus onto the CENP-LN binding region of CENP-C$^{2-545}$. Previous studies on CENP-C in our lab identified the PEST (Pro, Glu, Ser and Thr) rich domain of CENP-C (Leu^{265}, Phe^{266}, Leu^{267} and Trp^{317}) as a binding region for CENP-HIKM complex (Klare et al. 2015). To test the involvement of the PEST domain of CENP-C in the binding to CENP-LN complex, a recombinant construct of CENP-C encompassing the residues 225-364 was generated and tested for its binding to CENP-LN complex. Recombinant CENP-C$^{225-364}$ is sufficient to interact with the CENP-LN complex, as is evident from the shift towards higher molecular weight in the SEC assay (Figure 3-30 A). These results indicate that the PEST region of CENP-C acts as a hub for the binding of the CENP-

HIKM and CENP-LN complexes (Figure 3-30 B). The binding of both CENP-HIKM and CENP-LN complex to CENP-C is compatible, as demonstrated in the biochemical reconstitution of the CCAN sub-complexes (Weir et al. 2016).
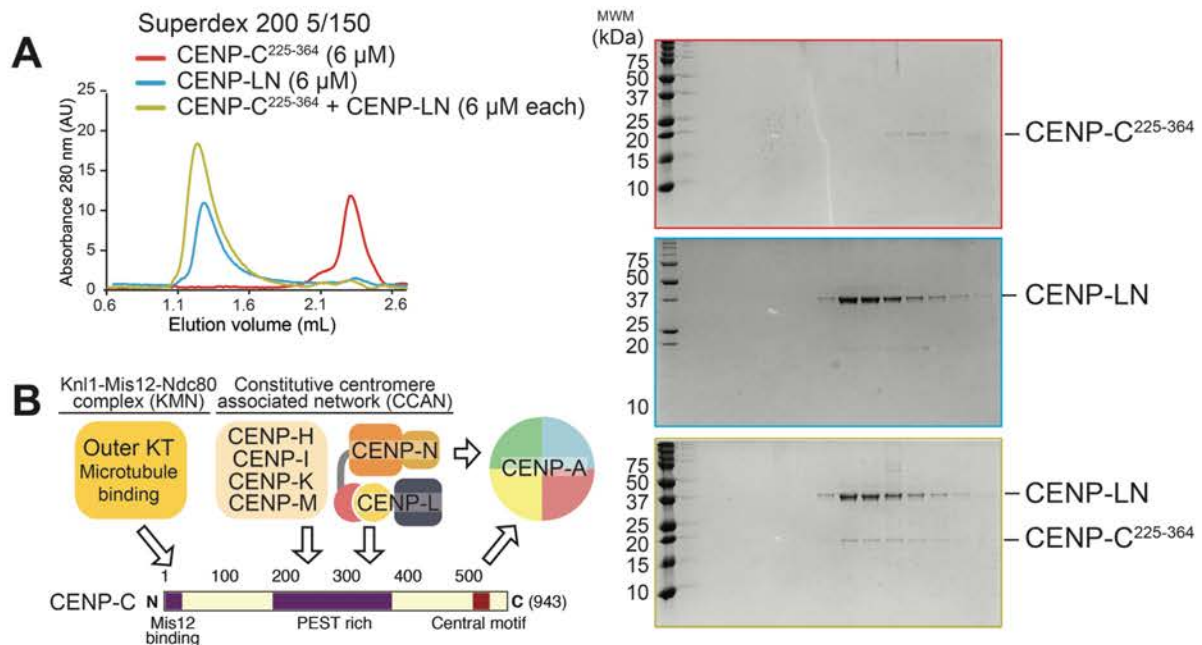

Figure 3-30 CENP-LN binds PEST domain of CENP-C

(A) SEC elution profile of CENP-LN with CENP-C$^{225-364}$, both at 6 µM. CENP-LN co-elutes in a complex with CENP-C$^{225-364}$. (B) Schematics of the domain organization of CENP-C depicting the binding regions of both CENP-HIKM and CENP-LN complex within the PEST domain of CENP-C.

To determine the residues of CENP-C involved in the interaction with CENP-LN complex, I obtained sequence alignments of CENP-C from various species to identify conserved residues within the PEST domain. Single alanine mutants of most of the conserved residues within the PEST domain of CENP-C did not affect the interaction with CENP-LN complex (Figure 3-31 A, residues highlighted in grey). After several attempts of permutations and combinations, we identified a combinatorial mutant of CENP-C encompassing residues Glu^{302A}, Phe^{303A}, Ile^{304A}, Ile^{305A} and Asp^{306A} (referred to as CENP-C^{5A}), that failed to interact with CENP-LN complex (Figure 3-31). SEC experiments indicated that the CENP-C^{5A} mutant abolished the interaction with the CENP-LN complex (Figure 3-31 C), suggesting that these specific residues within the PEST domain of CENP-C are involved in the binding to the CENP-LN complex. Corroborating the SEC analysis, the GST-CENP-C$^{2-545-5A}$ mutant completely abolished the interaction with CENP-LN complex when compared with GST-CENP-C$^{2-545-WT}$ in a GST-pull down assay (Figure 3-31 B).
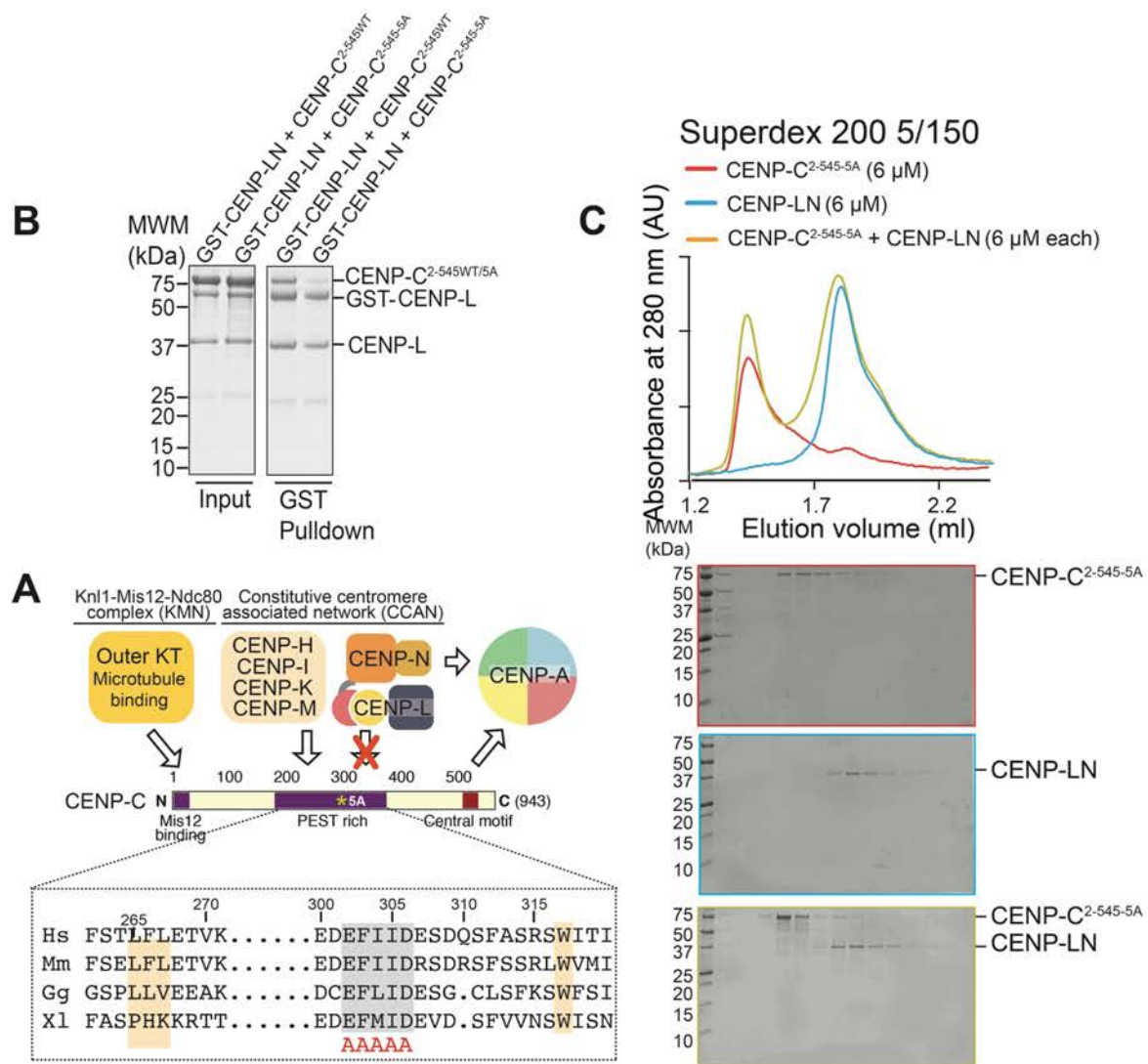
Figure 3-31 CENP-C$^{2-545-5A}$ mutant abolishes interactions with CENP-LN complex

(A) Schematics depicting the binding region of CENP-C involved in the interaction with CENP-LN complex. Multiple sequence alignment of CENP-C (lower panel). (B) GST-CENP-C$^{2-545-WT}$ or GST-CENP-C$^{2-545-5A}$ baits were immobilized on beads and incubated with CENP-LN complex. (C) SEC elution profile of CENP-LN with CENP-C$^{2-545-5A}$, both at 6 µM. CENP-C$^{2-545-5A}$ completely abolished the interaction with CENP-LN complex, as is evident from the two distinct peaks.

3.11 Specificity of CENP-C$^{2-545-5A}$ Mutant

As depicted in Figure 3-31 A, besides binding to CENP-LN complex, the N-terminal fragment of CENP-C also binds to Mis12 complex, CENP-A^{NCP} and CENP-HIKM complex (Petrovic et al. 2010; Klare et al. 2015; Kato et al. 2013). Given that the CENP-LN binding site is flanked by the binding regions of CENP-HIKM and CENP-A^{NCP}, I examined the specificity of CENP-C^{5A} mutant. To this end, the CENP-C^{5A} mutant was incubated with CENP-HIKM and CENP-A^{NCP}. As presented in Figure 3-32, SEC analysis demonstrated that CENP-C^{5A} co-eluted with the CENP-A^{NCP} (Figure 3-32 A) and also with the CENP-HIKM complex (Figure 3-32 B), indicating that the binding of CENP-A^{NCP} or CENP-HIKM is not affected by the CENP-C^{5A} mutant. These results suggest that the CENP-C^{5A} mutant specifically abolished interactions with the CENP-LN complex within the CCAN network.
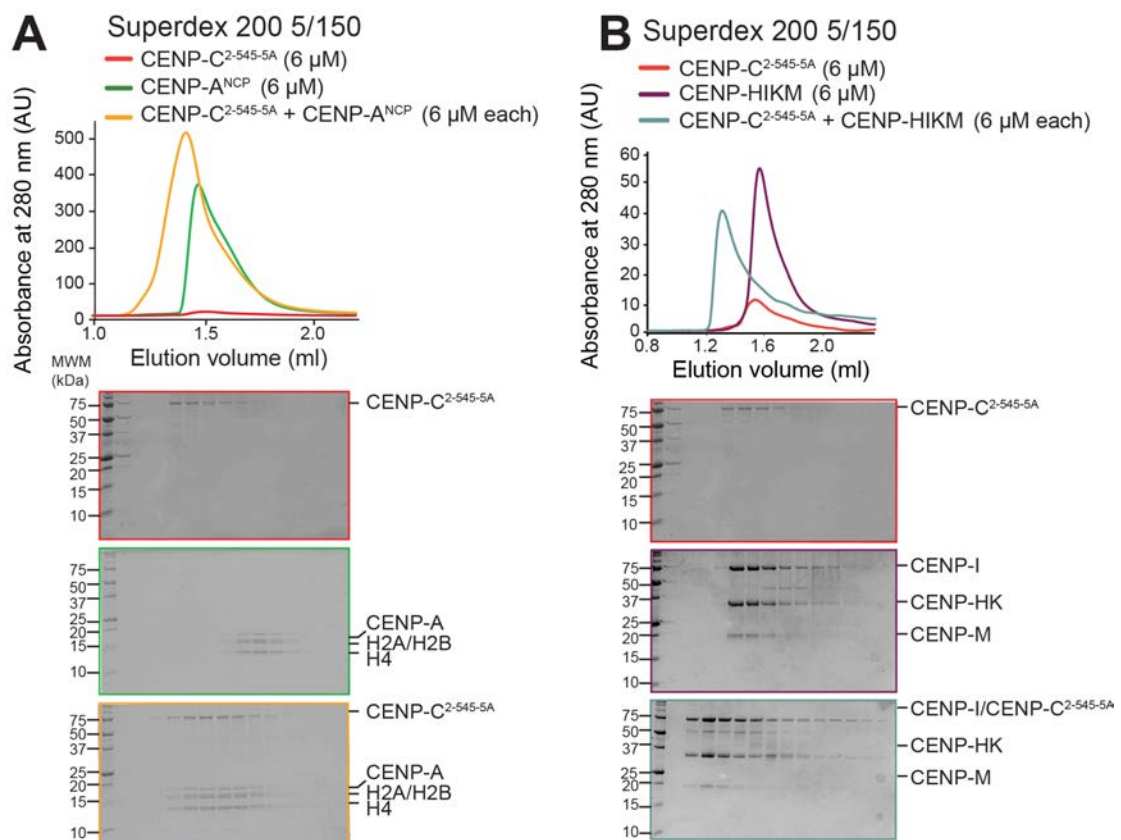

Figure 3-32 Specificity of CENP-C$^{2-545-5A}$

(A) SEC elution profile of CENP-A^{NCP} with CENP-C$^{2-545-5A}$, both at 6 µM. CENP-C$^{2-545-5A}$ co-elutes with CENP-A^{NCP}, indicating the formation of CENP-C$^{2-545-5A}$:CENP-A^{NCP} complex. (B) SEC elution profile of CENP-A^{NCP} with CENP-HIKM, both at 6 µM. CENP-C$^{2-545-5A}$ co-elutes with CENP-HIKM complex indicating the formation of CENP-C$^{2-545-5A}$:CENP-HIKM complex.

3.12 CENP-LN:CENP-C interaction affinity:

Isothermal titration calorimetry (ITC) assays were performed to measure the binding affinity of CENP-C$^{225-364}$ for CENP-LN^{230-C} complex. ITC is a quantitative technique that allows direct measurement of the binding affinities between biomolecules. It works by measuring the heat that is either released or absorbed during the binding events. Recombinant CENP-LN^{230-C} was loaded into a sample cell at a concentration of 6 µm. CENP-C$^{225-364}$, at 60 µM concentration, was then aspirated into a syringe. CENP-C$^{225-364}$ was injected stepwise at regular intervals of 2 secs and the binding events were measured. As presented in Figure 3-33 A, the dissociation constant of CENP-C$^{225-364}$ for CENP-LN complex is 1 µM while CENP-C$^{225-364-5A}$ showed no binding, which corroborates the previous results (Figure 3-33 B).
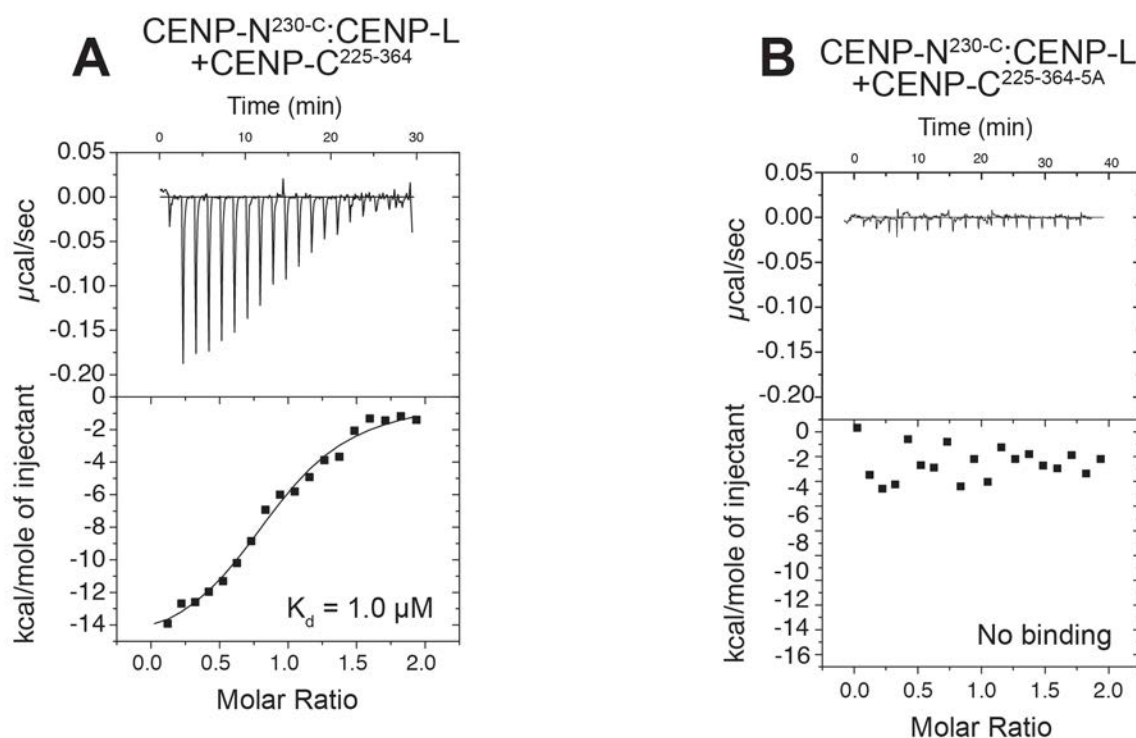

Figure 3-33 ITC assays of CENP-LN^{230-C} with CENP-C

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

3.13 Depletion of CENP-C leads to depletion of CENP-N

To investigate whether the kinetochore localization of CENP-LN complex also depends on CENP-C in addition to CENP-A, an mCherry tagged CENP-N cell line was generated. Expression of mCherry-CENP-N was monitored by immunofluorescence and western blotting. Depletion of CENP-C was done using CENP-C siRNA for 72 h, as described previously (Klare et al. 2015). Corroborating the previous studies on CENP-C, CENP-C depletion leads to a complete loss of CENP-N. This implies that CENP-N might depend on CENP-C in addition to CENP-A^{NCP} for kinetochore recruitment (Klare et al. 2015; Weir et al. 2016; Nagpal et al. 2015) (Figure 3-34). This hypothesis will be investigated in more detail in the further progress of this work.
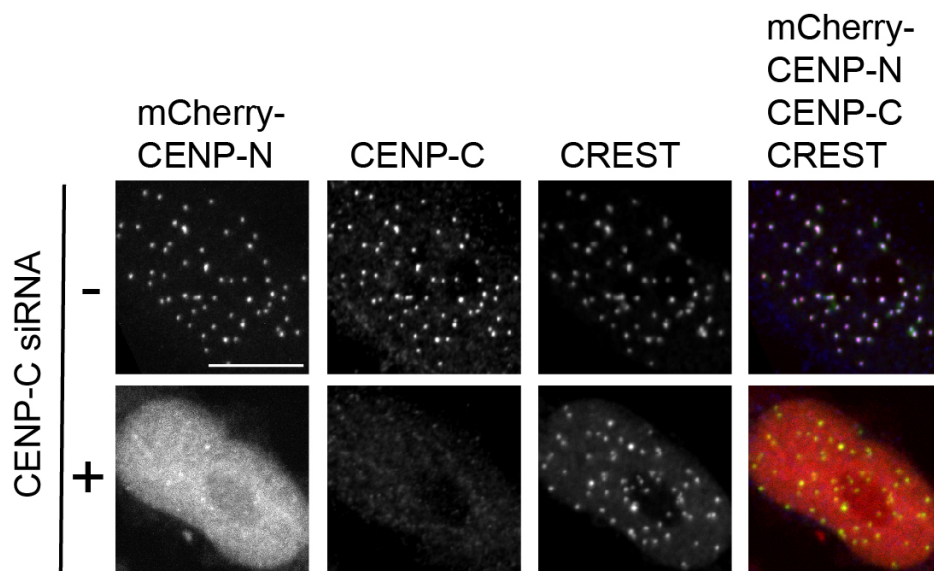

Figure 3-34 Expression of mCherry-CENP-N cell line

Cells expressing mCherry-CENP-N were treated with CENP-C siRNA. Cells were fixed and stained for CENP-C and CREST. Cells depleted of CENP-C show a complete loss of kinetochore localization of CENP-C and CENP-N.

3.14 Kinetochore localization of CENP-N depends on CENP-C:

As described previously, the loss of CENP-C from the kinetochores also leads to a loss of kinetochore localization of CENP-N. This prompted us to examine whether the kinetochore localization of CENP-N also depends on CENP-C in addition to CENP-A^{NCP}. To this end, *in vivo* experiments were performed using an inducible mCherry-CENP-N HeLa cell line in both asynchronous (interphase) and synchronous cells (mitotic). In asynchronous cells expressing mCherry-CENP-N full length, endogenous CENP-C was depleted for 48 h, followed by the transfection of either GFP-CENP-C^{WT}

or GFP-CENP-C^{5A} mutant. Both GFP-CENP-C^{WT} and GFP-CENP-C^{5A} rescued the loss of CENP-C completely, indicating that the CENP-C localization is not affected by the mutations that impair the binding to CENP-LN complex (Figure 3-35 A-B). Like CENP-C, CENP-A kinetochore levels also seem to be unaffected by the presence of CENP-C^{5A}, suggesting that both CENP-A and CENP-C may act upstream in the hierarchy of the CCAN assembly (Figure 3-35 A-D). Interestingly, GFP-CENP-C^{5A} mutant rescued >60% of CENP-N when compared with the CENP-C^{WT} indicating that the CENP-C^{5A} mutant has a moderate effect on the kinetochore localization of CENP-LN complex, while its binding to CENP-LN is completely abolished *in vitro* (Figure 3-35 A-C). Collectively, these observations suggest that in interphase, the mutations of CENP-N that impair the binding to CENP-A have strong effect on the kinetochore localization of CENP-N, when compared with the mutations that impair the binding of CENP-LN with CENP-C.
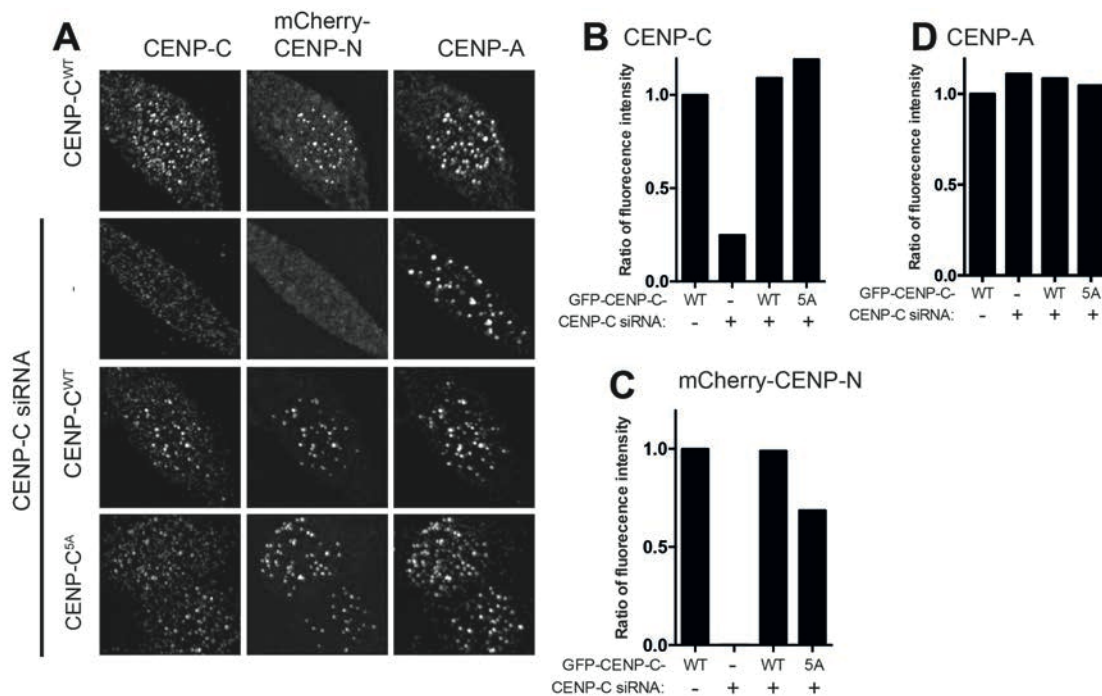

Figure 3-35 Kinetochore localization of CENP-N in interphase

(A**)** FlpIn TRex HeLa cells expressing mCherry-CENP-N were treated with CENP-C RNAi. CENP-C depleted cells were transfected with either GFP-CENP-C^{WT} or GFP-CENP-C^{5A}. Quantification of kinetochore fluorescence intensity levels of CENP-C (B) mCherry-CENP-N (C) and CENP-A (D) normalized to CENP-C.

To investigate the role of the GFP-CENP-C^{5A} in mitosis, CENP-C RNAi experiments were performed as described above, with the change that the cells were treated with nocodazole (in order to arrest the cells in mitosis) 16 h prior to fixation. Expression of GFP-CENP-C^{5A} largely rescued the kinetochore levels of both CENP-C and CENP-A

when compared with GFP-CENP-C^{WT} (Figure 3-36 A-B-D). Conversely, GFP-CENP-C^{5A} was unable to produce significant rescue of the kinetochore levels of CENP-N (Figure 3-36 A-C). Taken together these results suggest that that mitotic centromeric localization of CENP-N depends on CENP-C.
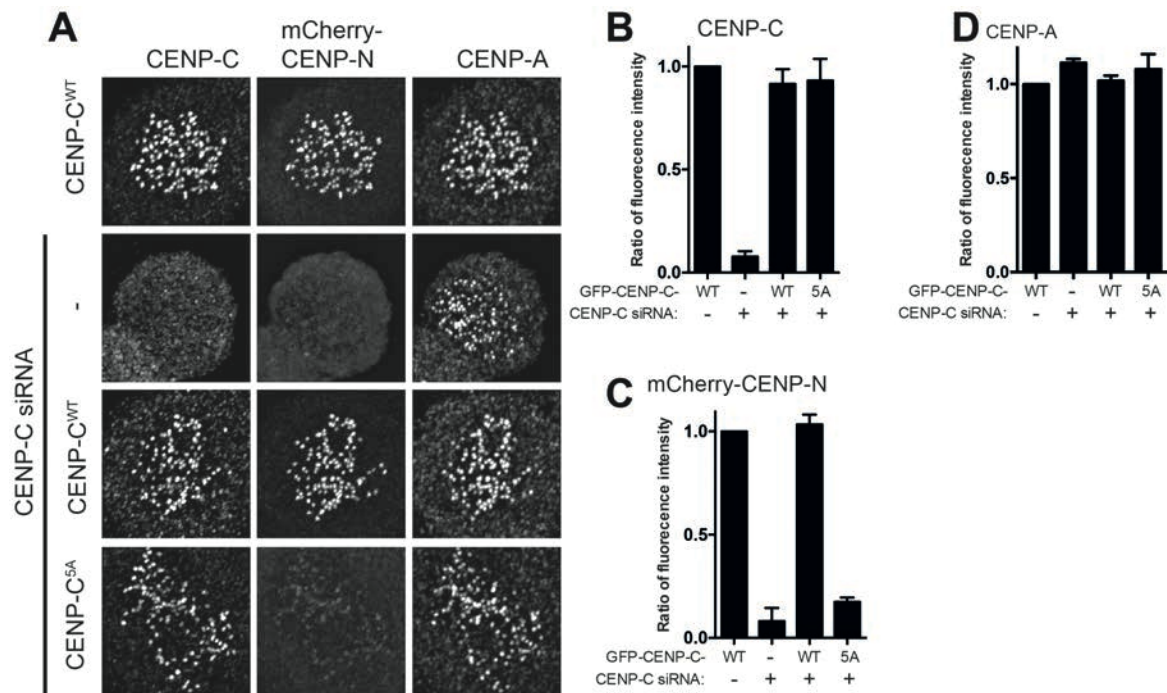

Figure 3-36 Kinetochore localization of CENP-N depends on CENP-C in mitosis

(A) FlpIn TRex HeLa Cells expressing mCherry-CENP-N were treated with CENP-C RNAi. CENP-C depleted cells were transfected with either GFP-CENP-C^{WT} or GFP-CENP-C^{5A}. Quantification of kinetochore fluorescence intensity levels of CENP-C (B) mCherry-CENP-N (C) and CENP-A (D) following the rescue of CENP-C with either GFP-CENP-C^{WT} or GFP-CENP-C^{5A} and normalized to CENP-C.

3.15 Co-operative assembly of CCAN onto CENP-A nucleosomes

A recent study from our laboratory demonstrated the biochemical reconstitution of a 7-subunit CCAN sub-complex comprising CENP-C, CENP-HIKM and CENP-LN (CHIKMLN complex) bound to CENP-A^{NCP} (Weir et al. 2016). Furthermore, it has been proposed that the CCAN binds the CENP-A nucleosomes in a cooperative manner. Cooperativity is a phenomenon in which the binding of one molecule to a ligand influences (increases or decreases) the binding affinity of a second molecule to the ligand. For instance, the binding affinity (k_d) of CENP-LN to CENP-A^{NCP} was ~450 nM, while the addition of CENP-CHIKM to the CENP-LN complex resulted in a 20-fold increase in binding affinity to CENP-A^{NCP} (K_d = 20nM) (Weir et al. 2016). Having

identified that CENP-LN binds to both CENP-C and CENP-A^{NCP}, we next wanted to determine the effects of the specific mutation on the overall stability of this complex. To this end, we performed GST-pull-down assays with GST-CENP-LN (bait) incubated with CENP-A^{NCP} and CENP-C^{WT} or CENP-C$^{2-545-5A}$ (prey). The CENP-LN bound to CENP-A nucleosomes much stronger in the presence of CENP-C$^{2-545-WT}$ than that of CENP-C$^{2-545-5A}$ mutant (Figure 3-37 A). This preliminary experiment suggested that both CENP-C and CENP-LN binds to the same nucleosome and the overall stability of the CCAN sub-complexes relies on each individual interaction. Moreover, when the crystal structure of CENP-C in complex with H3-CATD was superimposed with the cryo-EM structure of CENP-N:CENP-A^{NCP}, CENP-N and CENP-C fit without any steric clashes, implying that the binding of CENP-C and CENP-LN to the same nucleosome is compatible (Figure 3-37 B). Collectively, these results demonstrate that that CENP-C and CENP-LN acts cooperatively to stably anchor the CCAN onto CENP-A nucleosomes.
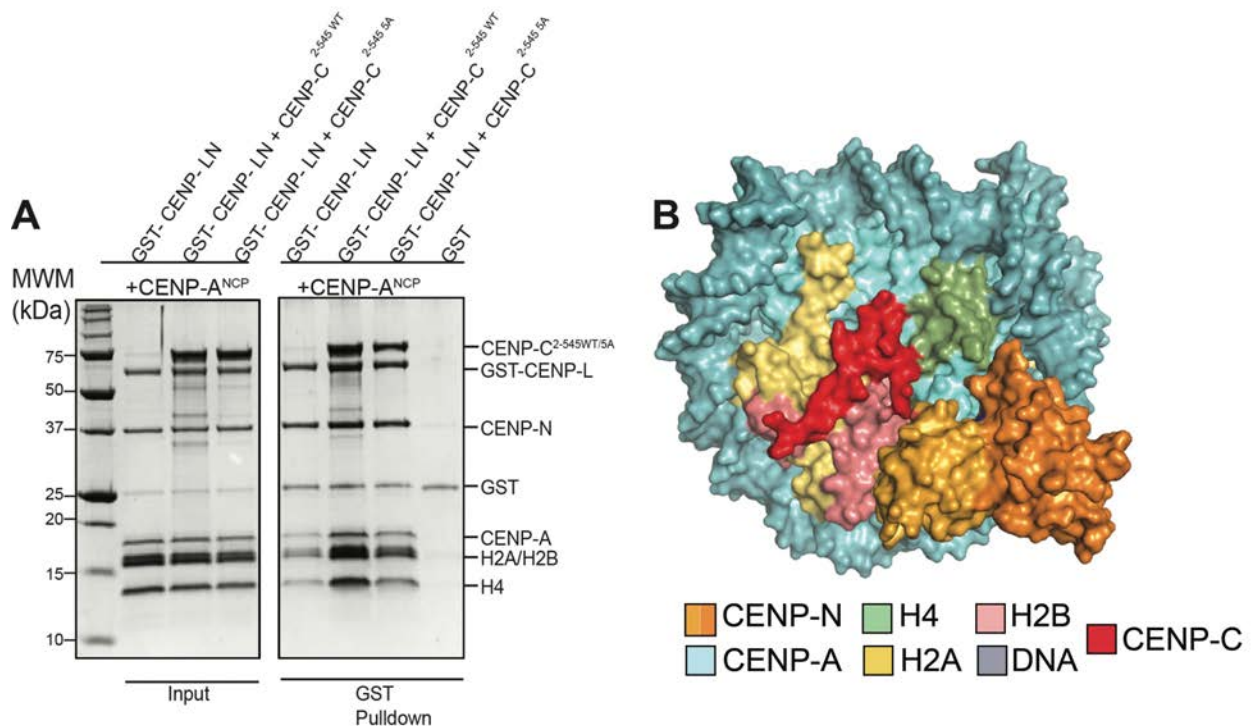

Figure 3-37 Co-operative binding of CCAN onto CENP-A nucleosome

A) GST-pulldown assay of CCAN subcomplexes. GST-CENP-LN complex was immobilized on GSH beads and incubated with either CENP-C$^{2-545-WT}$ or CENP-C$^{2-545-5A}$ and CENP-A^{NCP} (B) Surface representation of a composite model built by combining the coordinates of the CENP-C motif (red) bound by nucleosomes (PDB ID: 4X23) with those of the CENP-N:CENP-A^{NCP} complex.

3.16 Downstream recruitment of CCAN sub complexes

In addition to CENP-A^{NCP} and CENP-C, we have also identified a direct interaction between the CENP-LN and CENP-HIKM complex corroborating the previous studies (Figure 3-38), but the details of this interaction remain unclear. Moreover, It has been shown that the kinetochore localization of CENP-HIKM depends on CENP-C but the role of CENP-LN complex in its recruitment has not been investigated so far (Weir et al. 2016; McKinley et al. 2015; Klare et al. 2015). This prompted us to examine the mitotic kinetochore recruitment of the CENP-HIKM complex when the localization of CENP-N is compromised.
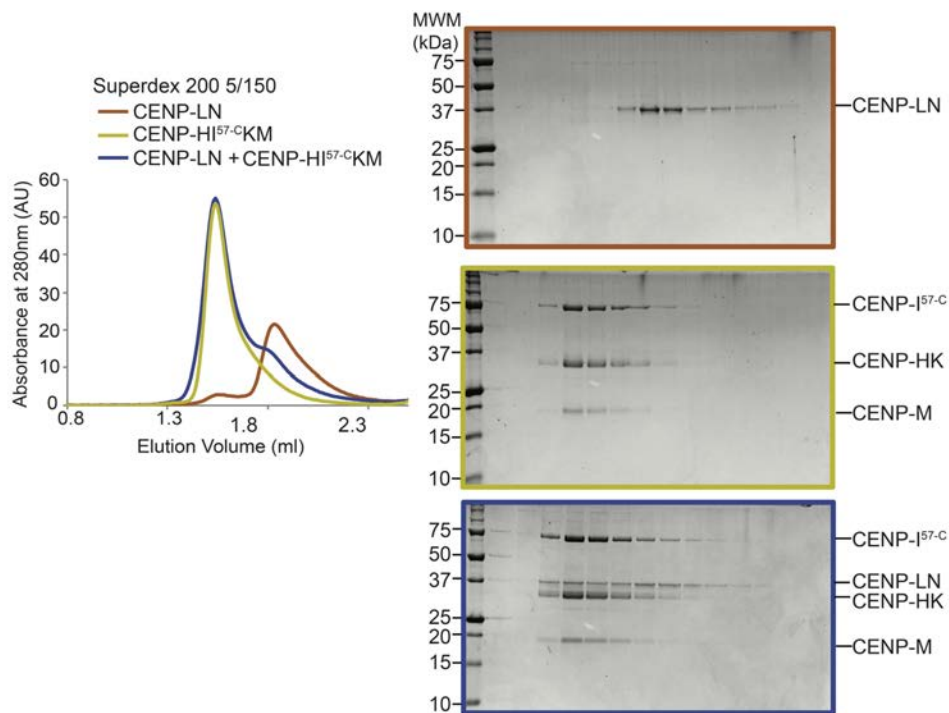

Figure 3-38 CENP-LN complex directly interacts with the CENP-HIKM complex

SEC elution profile of the CENP-LN complex and of its combination with the CENP-HIKM complex, both at 6 µM. CENP-LN and CENP-HIKM form a stoichiometric complex and co-elute in analytical SEC.

We performed rescue experiments using GFP-CENP-C^{WT} and GFP-CENP-C^{5A} in cells expressing mCherry-CENP-N. As described above, endogenous CENP-C was depleted by RNAi followed by the transfection of either GFP-CENP-C^{WT} or GFP-CENP-C^{5A}. Cells were arrested in mitosis and CENP-C, CENP-HK and mCherry-CENP-N levels were quantified. As expected, GFP-CENP-C^{5A} completely abrogated

the kinetochore localization of CENP-N, which corroborates the previous results of this study (Figure 3-39). Surprisingly, the kinetochore levels of CENP-HK were severely affected in the absence of CENP-N, suggesting a role for CENP-N in recruiting CENP-HIKM complex (Figure 3-38 A and C). These results demonstrate the importance of CENP-LN:CENP-C interaction in the kinetochore recruitment of the CENP-HIKM complex.
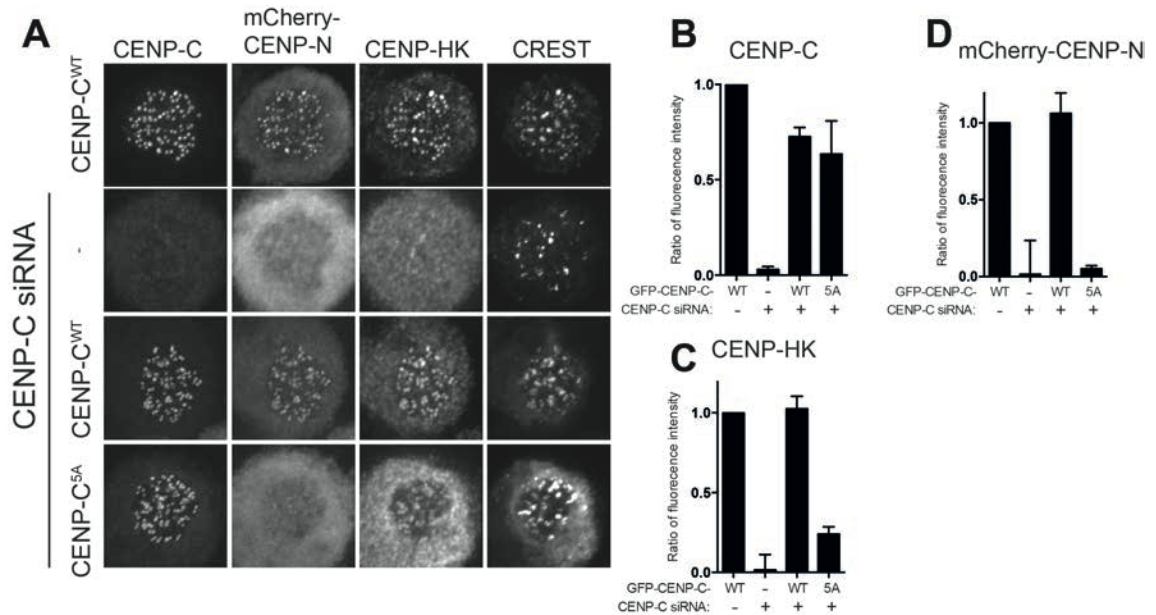

Figure 3-39 Dependency of CENP-HIKM on CENP-N in mitosis

(A) FlpIn TRex HeLa cells expressing mCherry-CENP-N were treated with CENP-C RNAi. CENP-C depleted cells were transfected with either GFP-CENP-C^{WT} or GFP-CENP-C^{5A}. Quantification of kinetochore fluorescence intensity levels of CENP-C (B) CENP-HK (C) and mCherry-CENP-N (D) following the rescue of CENP-C with either GFP-CENP-C^{WT} or GFP-CENP-C^{5A} and normalized to CREST.

4 Discussion

Despite several previous studies, it has been difficult to assign a specific function for each CCAN subunit, as depletion of one CCAN subunit leads to the loss of the other CCAN subunits due to codependence in their kinetochore localization. CENP-N associates with CENP-L to form a stable heterodimer of CENP-LN complex. Therefore, how CENP-N together with CENP-L, recognizes centromeric chromatin and engages in the physical interactions with other kinetochore subunits was poorly understood. In this study, I have presented biochemical and structural characterization of CENP-LN complex. Specifically, the structural characterization of CENP-N alone and in complex with CENP-A^{NCP} revealed the molecular basis of this interaction for the first time. My PhD work further identified the specific residues required for the interaction between CENP-LN complex, CENP-A^{NCP} and CENP-C. My work represents a successful case of "from structure to function" workflow, where the determination of the CENP-N structure together with CENP-A^{NCP}, paved the way for understanding the structural basis for the selectivity of CENP-A^{NCP} by CENP-N.

The determination of the structure of CENP-N revealed a two-domain organization, with an N-terminal Pyrin domain (residues 1-77) and a C-terminal CENP-L homology domain (residues 78-212) in a fixed orientation. Cryo-EM structure of CENP-N:CENP-A^{NCP} complex clearly demonstrated that CENP-N utilizes both the Pyrin and the CLN-HD domains to interact with both the CENP-A histone and the nuclesomal DNA. To date, this is the first study to suggest a role for Pyrin domain in the interaction with chromatin.

CENP-N directly binds at the CATD domain of CENP-A (Carroll et al. 2009). Specifically, the RG (Arg^{80} and Gly^{81}) loop within the CATD of CENP-A which is a unique marker that discriminates CENP-A from canonical H3 nucleosomes. This study reveals for the first time why the insertion of Arg^{80} and Gly^{81} (RG loop) is crucial for the preferential selectivity of CENP-A by CENP-N. In addition to recognition by CENP-N, the RG loop also appears to facilitate the folding of CENP-A arrays into a compact ladder-like chromatin structure (Fang et al. 2015). Previous studies have reported that the compact CENP-A chromatin impairs the binding of CENP-N to CENP-A by concealing the RG loop. In addition, previous studies have suggested that the

centromeric chromatin undergoes a structural transition from a compact state in the G1-phase to an open state in the S-phase of the cell cycle (Fang et al. 2015). Interestingly, majority of CENP-N is loaded onto centromeres during the S-phase of the cell cycle, when the chromatin is in the open confirmation (Hellwig et al. 2011; Fang et al. 2015; Hoffmann et al. 2016). Collectively, these results suggest that the CENP-N is loaded onto the centromeres by the direct recognition of the RG loop during the S-phase of the cell cycle. Moreover, the residues of CENP-N that are involved in the interaction with CENP-A are also conserved in yeast, suggesting the likely involvement of these residues in the binding of Chl4^{CENP-N} to Cse4^{CENP-A} nucleosomes. However, the binding of Chl4^{CENP-N} to Cse4^{CENP-A} nucleosomes remains poorly understood.

Comparison of the crystal structure of CENP-A^{NCP} with the cryo-EM structure of the CENP-N:CENP-A^{NCP} complex revealed that the DNA ends appear to be well ordered, suggesting a role for CENP-N in stabilizing CENP-A^{NCP}, whose DNA ends had been shown to be disordered in isolation (Tachiwana et al. 2011). Indeed, a recent study has demonstrated that the DNA ends of CENP-A^{NCP} are less accessible to micrococcal nuclease digestion when bound by CENP-N, in agreement with our results (Cao et al. 2018). One of the features observed in the CENP-N:CENP-A^{NCP} structure was a weak interaction of the $\beta 3$-$\beta 4$ loop of CENP-N with the N-terminal tail of histone H4. Interestingly, a recent study on the centromere-specific histone post-translational modifications identified H4K20 monomethylation (H4K20me1) to be specifically enriched at the centromeres (Hori et al. 2014). Although H4K20me1 is constitutively detected at the centromeres, this modification does not occur on CENP-A prior to its incorporation, suggesting that the CENP-A^{NCP} acquires this modification once stably incorporated into the centromeric chromatin (Hori et al. 2014). Given that the CCAN sub-complexes are also constitutively bound to centromeres and CENP-N interacts weakly with the N-terminal tail of histone H4, it remains to be established if H4K20me1 influences the binding of CENP-N to nucleosomes and facilitate kinetochore assembly.

4.1 CENP-N and CENP-C recognize distinct features of the CENP-A^{NCP}

The determination of the cryo-EM structure of CENP-N:CENP-A^{NCP} has allowed us to compare the differential binding mode of CENP-C and CENP-N with CENP-A nucleosomes. Comparison of the structures of CENP-C:H3$^{1-132}$-IEEGLG with CENP-N:CENP-A^{NCP} has revealed that CENP-N and CENP-C use different strategies to bind

to the CENP-A nucleosomes (Figure 4-1). The central motif and CENP-C motif of CENP-C recognize the acidic patch of H2A and H2B, along with the divergent C-terminal tail of CENP-A (Kato et al. 2013), while CENP-N recognizes only a small feature (RG loop) on the surface of the CENP-A nucleosome (Figure 4-1 A). Corroborating these findings, HXMS (hydrogen deuterium exchange mass spectrometry) experiments have shown that the only HX protection that was observed in the CENP-N^{NT}:CENP-A^{NCP} complex was confined to the L1 loop, in contrast to the widespread HX protection conferred to the CENP-A^{NCP} by CENP-C. The fact that CENP-N and CENP-C target different parts of CENP-A^{NCP} suggests that these proteins can bind to the same CENP-A nucleosome. Furthermore, superimposition of the cryo-EM structure of CENP-N:CENP-A^{NCP} on the crystal structure of CENP-C:H3$^{1-132}$-IEEGLG (PDB ID 4X23) indicates that both CENP-C and CENP-N can be accommodated without any steric clashes (Figure 4-1 C). Indeed, recent work from our own laboratory has demonstrated the assembly of a seven-subunit CCAN sub-complex containing CENP-C, CENP-LN and CENP-HIKM, on a centromeric CENP-A nucleosome, suggesting that the binding of the CCAN sub-complexes to a single CENP-A^{NCP} is compatible (Weir et al. 2016). The binding of CENP-C to CENP-A^{NCP} does not appear to affect the selectivity of CENP-N to CENP-A^{NCP}, but rather alters the nucleosome core and rigidifies the CENP-A nucleosome (Falk et al. 2015). Although CENP-C and CENP-N bind to CENP-A^{NCP} individually, the overall stability of CENP-LN, CENP-C and CENP-A^{NCP} seems to rely on each individual interaction, as CENP-LN together with CENP-C bound CENP-A^{NCP} much stronger than that of CENP-LN alone, suggesting a co-operative binding mechanism of the CCAN sub-complexes (Figure 3-37 A) (Cao et al. 2018; Weir et al. 2016). Therefore, these results suggest that the binding of CENP-LN and CENP-C to CENP-A nucleosomes might impart additive effects on the stability of the CCAN sub-complexes.
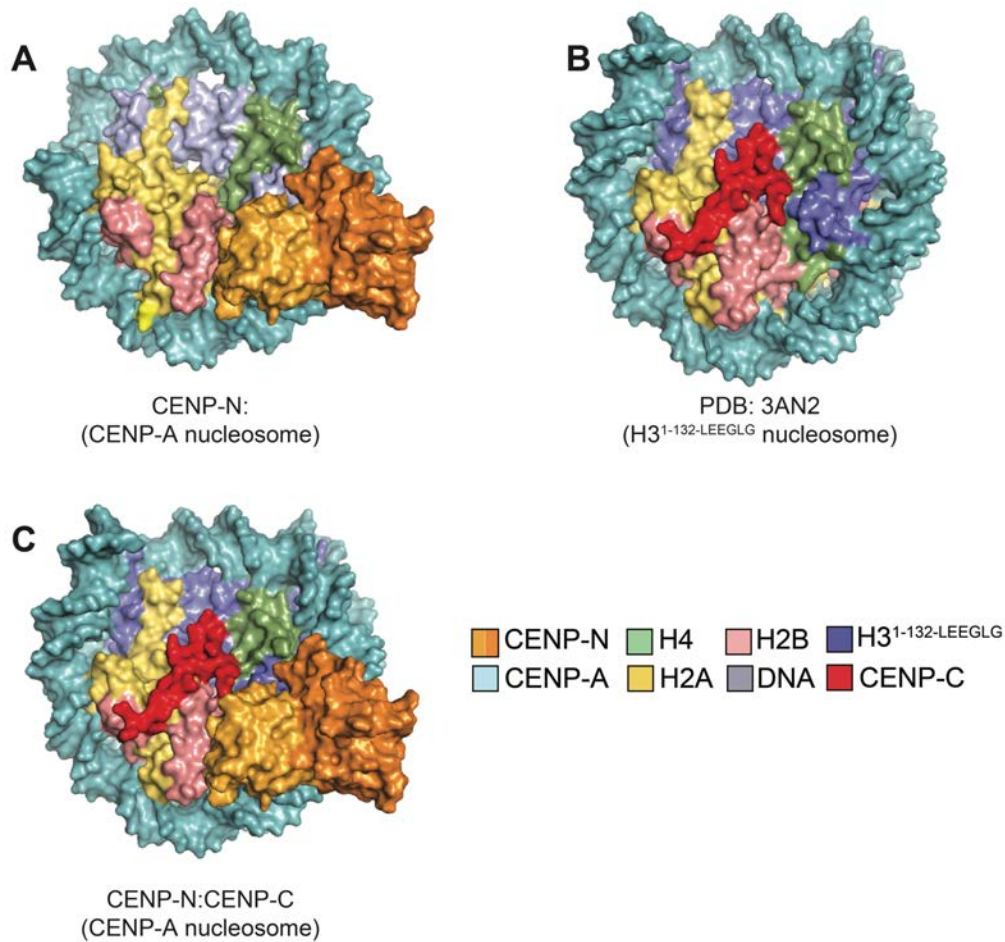
Figure 4-1 Comparison of the nucleosome binding modes of CENP-N and CENP-C

(A) Surface representation of the cryo-EM structure of CENP-N:CENP-A[NCP] complex determined in this study. (B) Surface representation of the crystal structure of CENP-C:H3[1-132]-IEEGLG complex (PDB ID 4X23). (C) Overlay of structures represented in A and B.

Stable centromere inheritance through many generations requires new CENP-A deposition specifically at the site of the pre-existing centromere. The deposition of new CENP-A relies heavily on the conserved CENP-A specific histone chaperone, HJURP (Holliday Junction recognition protein) (Zasadzińska et al. 2013; Hu et al. 2011; Dunleavy et al. 2009). Histone chaperones play key roles in facilitating the nucleosome assembly and disassembly processes. Interestingly, HJURP also recognizes the CATD region of CENP-A (Dunleavy et al. 2009). CENP-N recognizes the RG motif located within the CATD of CENP-A, while HJURP recognizes Ser[68] located just outside the CATD (Hu et al. 2011). Moreover, while the DNA binding clearly increases the affinity of CENP-N for CENP-A[NCP], HJURP cannot bind to CENP-A/H4 tetramers in the presence of DNA (Zasadzińska et al. 2013). In summary, both HJURP and CENP-N recognize distinct features of the CATD of CENP-A. HJURP binds to pre nucleosomal CENP-A while CENP-N binds to CENP-A once stably incorporated into

centromeric chromatin (Zasadzińska et al. 2013; Carroll et al. 2009). This hints that CENP-N might not be necessary for the deposition of new CENP-A but may be required for the maintenance of CENP-A.

Although the precise regulation of CENP-A deposition is not clearly understood, many studies have shown a role for chromatin remodeling enzymes in this process. Chromatin remodelers are multi-protein complexes that are involved in the regulation of chromatin accessibility and nucleosome positioning on DNA. In comparison with the chromatin remodelers SWI2/SNF2 or the BAH (bromo-adjacent homology domain) domain of Sir3, the binding of CENP-N to CENP-A^{NCP} shares some similarities (Farnung et al. 2017; X. Liu et al. 2017; Narlikar et al. 2013). Although both proteins (SWI2/SNF2 and BAH) engage in extensive contacts with DNA, the ATPase domain within the SWI2/SNF2 chromatin remodeler comes in close vicinity to the histone H3, without making any direct contacts. The BAH domain of SIR3 makes a direct contact with histone H4 N-terminal tail by recognizing the Lys^{16} of H4 and the acidic patch of H2A-H2B. Finally, CENP-N interacts directly with the CENP-A and weakly with the N-terminal tail of histone H4 (Figure 4-2). In short, CENP-N, like chromatin remodelers uses a unique combination of both histone and DNA binding. Previous studies on remodeling and spacing factor (RSF) (remodeling and spacing factor) show that this remodeler interacts with CENP-A and is involved in evicting histone H3 (Perpelescu et al. 2009). The precise mechanism by which the RSF interacts with CENP-A and evicts histone H3 remains unknown. Collectively, these observations suggest that CENP-N may protect CENP-A from eviction by masking the interface that could be recognized by remodelers, thereby restricting CENP-A in its place.
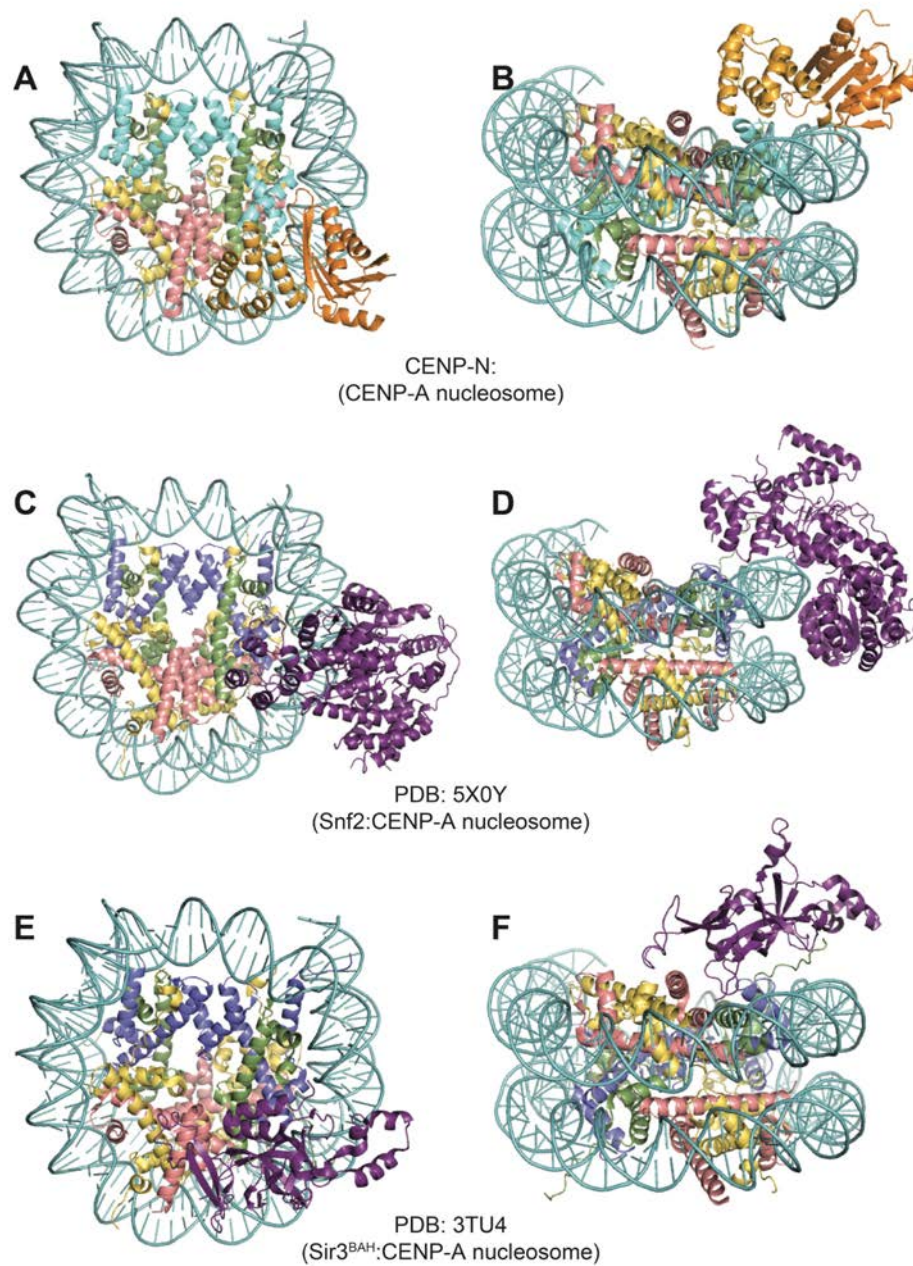
Figure 4-2 Comparison of the nucleosome binding modes of CENP-N with chromatin remodelers

(A-B) Front and top views of the structure of CENP-N:CENP-A^{NCP} complex. (C-D) Front and top views of the structure of SWI2/SNF2:H3^{NCP}complex showing similarities in the interaction with the nucleosome (X. Liu et al. 2017). (E-F) Front and top views of the structure of BAH:H3^{NCP}complex showing the interactions with acidic patch of H2A-H2B along with the H4 N-terminal tail and H3, while the interaction with DNA is very limited (Armache et al. 2011).

4.2   CENP-LN complex directly interacts with CENP-C

Besides binding to CENP-A^{NCP}, the CENP-LN complex has also been implicated in the direct binding to other CCAN sub-complexes such as CENP-C and the CENP-HIKM complex (Weir et al. 2016; McKinley et al. 2015). Several studies on CENP-C have demonstrated a linear organization of binding motifs within this largely disordered protein (Klare et al. 2015; Screpanti et al. 2011; Weir et al. 2016; Kato et al. 2013). Although both CENP-LN and CENP-C are known to bind to CENP-A^{NCP}, the binding of CENP-LN to CENP-C is independent of their binding to nucleosomes, as a construct of CENP-LN that lacks the N-terminal region still bound to CENP-C (Figure 3-28). Previous studies have reported that a truncation mutant lacking the C-terminus of CENP-N failed to localize to the kinetochores, suggesting that the interaction with other CCAN subunits is necessary for its kinetochore localization (Carroll et al. 2009). In this study, I have demonstrated that the C-terminal tail of CENP-N is sufficient to interact with both CENP-L and CENP-C (Figure 3-28) suggesting a role for both proteins in the kinetochore localization of CENP-N. In contrast to humans, the binding of Mif2^{CENP-C} seems to engage the N-terminus of Chl4^{CENP-N} (Hinshaw & Harrison 2013). However, the details of this interaction remain poorly understood. Despite several attempts we were unable to identify the residues on the CENP-LN complex that bind CENP-C. This prompted a shift in my focus to CENP-C. Previous studies on CENP-C have identified a direct binding of its PEST domain to CENP-HIKM complex and have implicated the residues Leu^{265}, Phe^{266}, Leu^{267} and Trp^{317} in this specific interaction (Klare et al. 2015). Through a series of recombinant approaches, I have identified that the PEST domain of CENP-C interacts with the CENP-LN complex. Moreover, this study has identified the specific residues within the PEST domain of CENP-C (Glu^{302}, Phe^{303}, Ile^{304}, Ile^{305}, and Asp^{306}) that are involved in the interaction with CENP-LN complex; whose mutations (CENP-C^{5A}) leads to complete abolishment of this interaction. PEST sequences are known to be associated with proteins that have short intracellular life and is therefore hypothesized that the PEST sequences might act as a signal peptide for proteolytic degradations. When a CENP-C lacking an N-terminal tail (residues 1-373) was transfected into BHK (Baby hamster kidney) cells, nuclei showed a distorted patterns implying that the PEST sequences might be required for proper centromere targeting of CENP-C (Lanini & McKeon 1995). Similar to humans, the PEST domain of CENP-C in chickens is also known to interact with CENP-HIKM and CENP-LN

complexes and was required for the kinetochore targeting of CENP-C in mitosis (Nagpal et al. 2015). Collectively, these results clearly demonstrate that CENP-HIKM and CENP-LN complexes occupy non-overlapping binding sites within the PEST domain of CENP-C. Thus, CENP-LN complex interacts with CENP-A^{NCP} and CENP-C independently of each of each other. A detailed summary of all these interactions is depicted in Figure 4-3.

Figure 4-3 Summary of interactions of CENP-LN complex with CCAN subunits

A detailed summary of interactions of different domains of CENP-LN complex with different CCAN sub-complexes. (green, indicates interaction, red, indicates no interaction).

4.3 Kinetochore localization of the CENP-LN complex depends on CENP-A^{NCP} and CENP-C

Previous studies as well as the current study have demonstrated that the kinetochore localization of the CENP-LN complex is dependent on CENP-A^{NCP} (Figure 3-24) (Carroll et al. 2009; McKinley et al. 2015). In order to test if the kinetochore localization of the CENP-LN complex also depends on other CCAN members, we carried out *in vivo* analysis using either GFP-CENP-C^{WT} or GFP-CENP-C^{5A} mutant and assessed for the ability of other CCAN members to localize to centromeres. The localization of the CENP-LN complex to kinetochores is severely affected in the presence of CENP-C^{5A} mutant in mitosis but moderately affected in interphase (Figure 3-36 and 3-35). Moreover, a recent study has shown that the N-terminal fragment of CENP-N is sufficient for its kinetochore localization in interphase but not in mitosis, suggesting that CENP-N depends on other CCAN members for its kinetochore localization in mitosis (McKinley et al. 2015; Hoffmann et al. 2016). Collectively, these observations

suggest that CENP-N switches its dependency from CENP-A^{NCP} to CENP-C as the cells progress from interphase to mitosis, a hypothesis that requires further investigations.

Kinetochore levels of both CENP-C and CENP-A are unaffected upon CENP-N depletion, suggesting that CENP-C and CENP-A act upstream of CENP-LN complex in the hierarchical assembly of the CCAN. Although previous studies have reported that the kinetochore localization of CENP-C depends on CENP-LN and CENP-HIKM complexes, we did not observe any dependency of CENP-C on CENP-LN complex (McKinley et al. 2015). However, a detailed understanding of how CENP-C is recruited to kinetochores remains an area of future research.

As described in this study, upon depletion of CENP-C, the CENP-C^{5A} mutant was not able to rescue the localization of CENP-N, but it does so with CENP-C. This allowed us to study the kinetochore localization of the CENP-HIKM complex in the absence of CENP-N. We observed that the loss of CENP-N, in the presence of the CENP-C^{5A} mutant, lead to a decrease in the levels of CENP-HK at the kinetochores, suggesting that CENP-N is involved in the recruitment of CENP-HIKM complex. Previous studies on CENP-HIKM complex reveals a direct interaction with CENP-LN and CENP-C, which is required for its kinetochore localization (Weir et al. 2016; McKinley et al. 2015; Klare et al. 2015). Despite the fact that CENP-LN directly binds to CENP-HIKM complex (Figure 3-38), the molecular details of this interaction remain poorly understood. A recent study from our own lab has demonstrated that the binding of the CENP-HIKM with the CENP-LN complex is a prerequisite for the binding of the CENP-OPQUR complex (Pesenti et al. 2018). Both CENP-HIKM and CENP-LN are responsible for the kinetochore recruitment of CENP-OPQUR complex. The EM analysis of CENP-HIKMLNOPQUR complex provided a first comprehensive structural analysis of the CCAN complexes (Figure 4-4). Specifically, the CENP-LN appears to be sandwiched between CENP-HIKM and CENP-OPQUR complexes, suggesting multiple contacts between CENP-HIKM and CENP-LN with CENP-OPQUR complex (Pesenti et al. 2018). One of the important conclusions drawn from the work of Pesenti *et.al.* (2018) was the compact and globular nature of CCAN sub-complexes, in contrast to the fibrous organization of the KMN network in the outer kinetochore. Collectively, these results suggest that the CENP-LN complex plays an important role in mediating

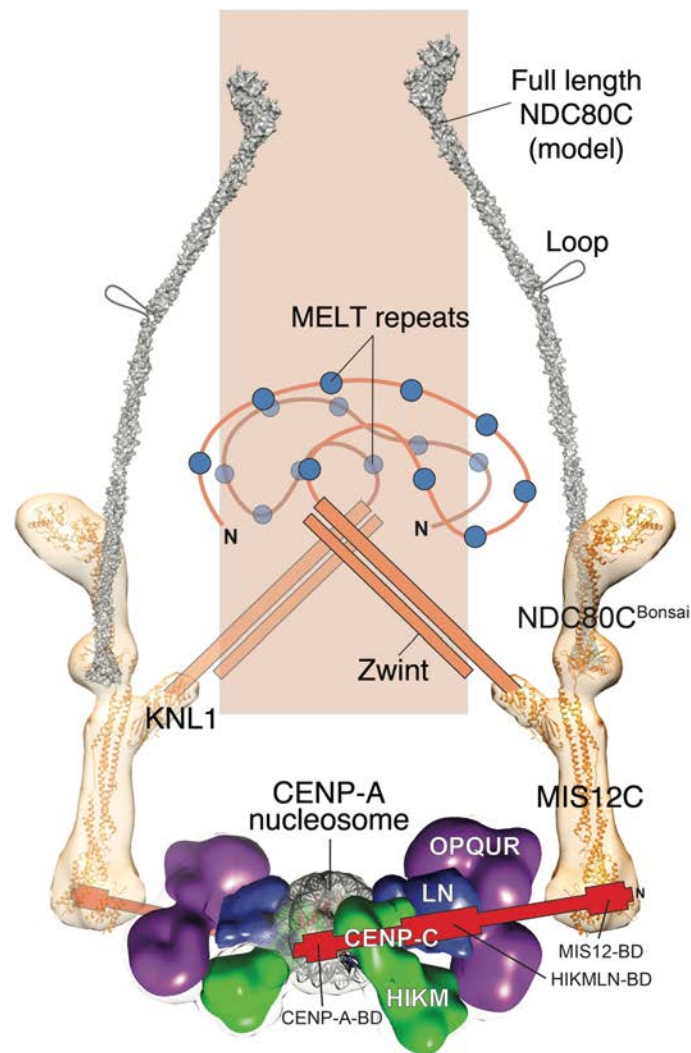the interactions with the other CCAN subunits required for proper kinetochore assembly.

Figure 4-4 Structural organization of the 26-subunit kinetochore (rKT 26).

A single CENP-A^{NCP} binds two copies of the CENP-CHIKMLNOPQUR complex (Pesenti et al. 2018). The CCAN complexes are compact and globular in nature, when compared with the fibrous organization of the KMN. Both CENP-LN complex and CENP-C binds CENP-A^{NCP} and other members within the CCAN network. CENP-C connects the inner kinetochore with the outer kinetochore by binding to Mis 12 complex. Figure adapted from (Pesenti et al. 2018).

The kinetochore localization of CENP-N and CENP-L seems to be mutually dependent, as cells depleted of either CENP-N or CENP-L exhibited similar phenotypes that resulted in the loss of either protein, along with the other CCAN subunits CENP-HIKM and CENP-OPQUR (Okada et al. 2006; McClelland et al. 2007; Weir et al. 2016).

However, unlike CENP-N^{FL}, CENP-L^{FL} can be expressed alone, suggesting that the stability of CENP-N depends on CENP-L. In some fungi, like *Candida glabarata*, CENP-L is not present and how CENP-N remains stable by itself is unknown. Moreover, dipterans and nematodes do not have CENP-N orthologs either (Schleiffer et al. 2012). In fact, most of the CCAN proteins are not universally conserved, for example, *Drosophila* has only CENP-C and the rest of the CCAN is missing. How these species manage accurate chromosome segregation with a minimal set of CCAN subunits remain unclear.

Collectively, my doctoral work has considerably extended our understanding on the organization and the role of the CENP-LN complex in kinetochore assembly. Furthermore, this study provides the structural basis for the recognition of the centromeric CENP-A nucleosome by CENP-N. In addition, this study clearly shows that the N-terminus of CENP-N binds to CENP-A^{NCP} and C-terminus of CENP-N, in complex with CENP-L, binds to CENP-C (Figure 4-5). Therefore, both the N-terminal and the C-terminal regions of CENP-N are required for proper kinetochore assembly. In conjunction with the progress in the reconstitution of CCAN subunits, structural analysis promises new crucial insights into the organization of the CCAN network. Moreover, centromere specific histone modifications such as H4K20me1, are emerging as important for kinetochore assembly. Given the fact that most of the CCAN sub-complexes can be reconstituted *in vitro*, it remains to be established how these centromere specific epigenetic modifications influence kinetochore assembly.
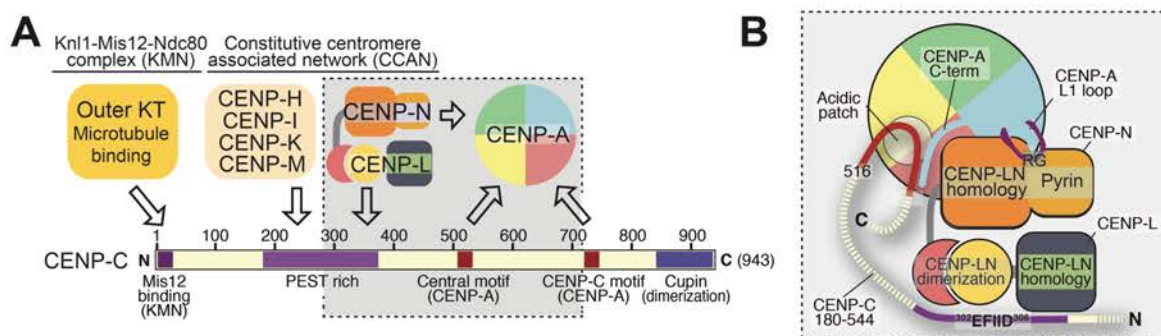

Figure 4-5 Schematics of the organization of CENP-LN and CENP-C with CENP-A^{NCP} complex

(A) Schematics of the organization of CENP-C showing binding of CENP-HIKM and newly identified CENP-LN complex to PEST rich domain of CENP-C along with other interactions. (B) An enlargement of the grey box in A, that summarizes the interactions of CENP-C, CENP-LN and CENP-A^{NCP}.

5 Summary

Accurate chromosome segregation requires the assembly of kinetochores, large multiprotein complexes that are built on the centromeres. A key step in this process involves the assembly of the constitutive centromere-associated network (CCAN) on CENP-A, a histone H3 variant that is enriched at the centromeres. The CCAN is composed of 16 protein components which can be subdivided into four functional groups: the CENP-LN complex, the CENP-HIKM complex, the CENP-OPQUR complex and the CENP-TWSX complex. Two kinetochore proteins, CENP-C and CENP-N are known to specifically recognize CENP-A^{NCP}. Although the CENP-LN complex is known to interact with CENP-A^{NCP}, the mechanistic basis for its interaction with CENP-A^{NCP} or other kinetochore proteins remains poorly understood.

This study provides insights into the organization of CENP-LN complex within the CCAN using biochemical, structural and *in vivo* approaches. Our results unravel the structural basis for the specific recognition of the CENP-A specific L1 loop by CENP-N. Additionally, we also have identified that the so-called PEST domain in the N-terminal half of CENP-C (a major component of the CCAN and a direct CENP-A binder), interacts directly with the CENP-LN complex. Furthermore, this study demonstrates that stable and sustained kinetochore localization of the CENP-LN complex requires its interactions with both CENP-A^{NCP} and CENP-C.

In summary, this work describes the mechanism by which the CENP-LN complex interacts with CENP-A^{NCP} and CENP-C. The obtained results significantly advance our understanding of the functional role of the CENP-LN complex within the CCAN network, which is required for the kinetochore assembly.

6 Zussamenfassung

Für die korrekte Trennung der Chromosomen während der Zellteilung müssen große Multiproteinkomplexe, sogenannte Kinetochore, an den Zentromeren der Chromosomen gebildet werden. Ein entscheidender Schritt dieses kritischen Prozesses ist die Bildung des konstitutiv Zentromer-assoziierten Netzwerks (constitutive centromere-associated network, CCAN) an dem Nukleosome CENP-A, einer Variante des Histons H3, welche vor allem an den Zentromeren lokalisiert ist. Das CCAN Netzwerk besteht aus insgesamt 16 Proteinen die in vier funktionelle Gruppen eingeteilt werden können: den CENP-LN Komplex, sowie die Proteinkomplexe CENP-HIKM, CENP-OPQUR und CENP-TWSX. Die Kinetochor proteine, CENP-C und CENP-N sind für die spezifische Erkennung des CENP-A^{NCP} Nukleosomes verantwortlich. Obwohl bekannt ist, dass der CENP-LN Komplex mit CENP-A^{NCP} direkt interagiert, ist über die mechanistische Grundlage dieser Interaktion ebenso wenig bekannt wie über die Interaktion zwischen CENP-A^{NCP} mit anderen Proteinen des Kinetochores.

Die hier vorliegende Arbeit nutzt biochemische, strukturbiologische und *in vivo* Verfahren um die Organisation des CENP-LN Komplexes innerhalb des CCAN Netzwerkes aufzuklären. Die erhaltenen Resultate erklären die strukturelle Basis für die spezifische Erkennung der L1-Region von CENP-A durch CENP-N. Zusätzlich kann gezeigt werden, dass die sogenannte PEST-Domäne in der N-terminalen Hälfte von CENP-C (ein wichtiges Protein des CCAN Netzwerkes und ein Interaktionspartner von CENP-A) direkt mit dem CENP-LN Komplex interagiert. Weiterhin wird ersichtlich, dass sowohl die Interaktion mit CENP-A^{NCP} als auch mit CENP-C die Voraussetzung für eine stabile und dauerhafte Lokalisierung des CENP-LN Komplexes innerhalb des Kinetochores ist.

Zusammengefasst beschreibt diese Arbeit den Interaktionsmechanismus des CENP-LN Komplexes mit CENP-A^{NCP} und CENP-C. Die präsentierten Ergebnisse fördern das Verständnis für die funktionelle Rolle des CENP-LN Komplexes innerhalb des CCAN Netzwerkes, welches für die ordentliche Assemblierung des gesamten Kinetochores notwendig ist.

7 Bibliography

Adams, P.D. et al., 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta crystallographica. Section D, Biological crystallography*, 66(Pt 2), pp.213–221.

Akiyoshi, B. et al., 2010. Tension directly stabilizes reconstituted kinetochore-microtubule attachments. *Nature*, 468(7323), pp.576–579.

Alberts, B., 2017. *Molecular Biology of the Cell*, Garland Science.

Alberts, B. et al., 2002. *Molecular Biology of the Cell 4th Edition*, Garland Science.

Aldrup-Macdonald, M.E. & Sullivan, B.A., 2014. The past, present, and future of human centromere genomics. *Genes*, 5(1), pp.33–50.

Alfieri, C. et al., 2016. Molecular basis of APC/C regulation by the spindle assembly checkpoint. *Nature*, 536(7617), pp.431–436.

Alushin, G.M. et al., 2010. The Ndc80 kinetochore complex forms oligomeric arrays along microtubules. *Nature*, 467(7317), pp.805–810.

Amano, M. et al., 2009. The CENP-S complex is essential for the stable assembly of outer kinetochore structure. *The Journal of cell biology*, 186(2), pp.173–182.

Amaro, A.C. et al., 2010. Molecular control of kinetochore-microtubule dynamics and chromosome oscillations. *Nature cell biology*, 12(4), pp.319–329.

Armache, K.-J. et al., 2011. Structural basis of silencing: Sir3 BAH domain in complex with a nucleosome at 3.0 Å resolution. *Science (New York, N.Y.)*, 334(6058), pp.977–982.

Basilico, F. et al., 2014. The pseudo GTPase CENP-M drives human kinetochore assembly. *eLife*, 3, p.e02978.

Bierie, B. et al., 2004. The proliferation associated nuclear element (PANE1) is conserved between mammals and fish and preferentially expressed in activated lymphoid cells. *Gene expression patterns : GEP*, 4(4), pp.389–395.

Black, B.E., Brock, M.A., et al., 2007. An epigenetic mark generated by the incorporation of CENP-A into centromeric nucleosomes. *Proceedings of the National Academy of Sciences of the United States of America*, 104(12), pp.5008–5013.

Black, B.E., Jansen, L.E.T., et al., 2007. Centromere identity maintained by nucleosomes assembled with histone H3 containing the CENP-A targeting domain. *Molecular cell*, 25(2), pp.309–322.

Bock, L.J. et al., 2012. Cnn1 inhibits the interactions between the KMN complexes of the yeast kinetochore. *Nature cell biology*, 14(6), pp.614–624.

Boyarchuk, Y. et al., 2007. Bub1 is essential for assembly of the functional inner centromere. *The Journal of cell biology*, 176(7), pp.919–928.

Brown, M.T., Goetsch, L. & Hartwell, L.H., 1993. MIF2 is required for mitotic spindle integrity during anaphase spindle elongation in Saccharomyces cerevisiae. *The Journal of cell biology*, 123(2), pp.387–403.

Buchwitz, B.J. et al., 1999. A histone-H3-like protein in C. elegans. *Nature*, 401(6753), pp.547–548.

Cao, S. et al., 2018. Constitutive centromere-associated network contacts confer differential stability on CENP-A nucleosomes in vitro and in the cell. *Molecular biology of the cell*, 29(6), pp.751–762.

Carmena, M. et al., 2012. The chromosomal passenger complex (CPC): from easy rider to the godfather of mitosis. *Nature reviews. Molecular cell biology*, 13(12), pp.789–803.

Carroll, C.W. et al., 2009. Centromere assembly requires the direct recognition of CENP-A nucleosomes by CENP-N. *Nature cell biology*, 11(7), pp.896–902.

Carroll, C.W., Milks, K.J. & Straight, A.F., 2010. Dual recognition of CENP-A nucleosomes is required for centromere assembly. *The Journal of cell biology*, 189(7), pp.1143–1155.

Chan, Y.W. et al., 2012. Aurora B controls kinetochore-microtubule attachments by inhibiting Ska complex-KMN network interaction. *The Journal of cell biology*, 196(5), pp.563–571.

Chan, Y.W. et al., 2009. Mitotic control of kinetochore-associated dynein and spindle orientation by human Spindly. *The Journal of cell biology*, 185(5), pp.859–874.

Chang, L. et al., 2015. Atomic structure of the APC/C and its mechanism of protein ubiquitination. *Nature*, 522(7557), pp.450–454.

Chao, W.C.H. et al., 2012. Structure of the mitotic checkpoint complex. *Nature*, 484(7393), pp.208–213.

Cheeseman, I.M. & Desai, A., 2008. Molecular architecture of the kinetochore-microtubule interface. *Nature reviews. Molecular cell biology*, 9(1), pp.33–46.

Cheeseman, I.M. et al., 2004. A conserved protein network controls assembly of the outer kinetochore and its ability to sustain tension. *Genes & development*, 18(18), pp.2255–2268.

Cheeseman, I.M. et al., KNL1 and the CENP-H/I/K Complex Coordinately Direct Kinetochore Assembly in Vertebrates.

Cheeseman, I.M. et al., 2006. The conserved KMN network constitutes the core microtubule-binding site of the kinetochore. *Cell*, 127(5), pp.983–997.

Chen, Y. et al., 2002. Phosphorylation of the mitotic regulator protein Hec1 by Nek2 kinase is essential for faithful chromosome segregation. *The Journal of biological chemistry*, 277(51), pp.49408–49416.

Ciferri, C. et al., 2008. Implications for kinetochore-microtubule attachment from the structure of an engineered Ndc80 complex. *Cell*, 133(3), pp.427–439.

Collaborative Computational Project, Number 4, 1994. The CCP4 suite: programs for protein crystallography. *Acta crystallographica. Section D, Biological crystallography*, 50(Pt 5), pp.760–763.

Corbett, K.D., 2017. Molecular Mechanisms of Spindle Assembly Checkpoint Activation and Silencing. *Progress in molecular and subcellular biology*, 56, pp.429–455.

Coudreuse, D. & Nurse, P., 2010. Driving the cell cycle with a minimal CDK control network. *Nature*, 468(7327), pp.1074–1079.

Daum, J.R. et al., 2009. Ska3 is required for spindle checkpoint silencing and the maintenance of chromosome cohesion in mitosis. *Current biology : CB*, 19(17), pp.1467–1472.

Davis, T.N. & Wordeman, L., 2007. Rings, bracelets, sleeves, and chevrons: new structures of kinetochore proteins. *Trends in cell biology*, 17(8), pp.377–382.

De Antoni, A. et al., 2005. The Mad1/Mad2 complex as a template for Mad2 activation in the spindle assembly checkpoint. *Current biology : CB*, 15(3), pp.214–225.

De Wulf, P., McAinsh, A.D. & Sorger, P.K., 2003. Hierarchical assembly of the budding yeast kinetochore from multiple subcomplexes. *Genes & development*, 17(23), pp.2902–2921.

DeLuca, J.G. & Musacchio, A., 2012. Structural organization of the kinetochore-microtubule interface. *Current opinion in cell biology*, 24(1), pp.48–56.

DeLuca, J.G. et al., 2002. hNuf2 inhibition blocks stable kinetochore-microtubule attachment and induces mitotic cell death in HeLa cells. *The Journal of cell biology*, 159(4), pp.549–555.

DeLuca, K.F. et al., 2018. Aurora A kinase phosphorylates Hec1 to regulate metaphase kinetochore-microtubule dynamics. *The Journal of cell biology*, 217(1), pp.163–177.

Desai, A. et al., 2003. KNL-1 directs assembly of the microtubule-binding interface of the kinetochore in C. elegans. *Genes & development*, 17(19), pp.2421–2435.

Drinnenberg, I.A. et al., 2014. Recurrent loss of CenH3 is associated with independent transitions to holocentricity in insects. *eLife*, 3.

Dunleavy, E.M. et al., 2009. HJURP is a cell-cycle-dependent maintenance and deposition factor of CENP-A at centromeres. *Cell*, 137(3), pp.485–497.

Earnshaw, W.C. & Rothfield, N., 1985. Identification of a family of human centromere proteins using autoimmune sera from patients with scleroderma. *Chromosoma*, 91(3-4), pp.313–321.

Erhardt, S. et al., 2008. Genome-wide analysis reveals a cell cycle-dependent mechanism controlling centromere propagation. *The Journal of cell biology*, 183(5), pp.805–818.

Espert, A. et al., 2014. PP2A-B56 opposes Mps1 phosphorylation of Knl1 and thereby promotes spindle assembly checkpoint silencing. *The Journal of cell biology*, 206(7), pp.833–842.

Evans, T. et al., 1983. Cyclin: a protein specified by maternal mRNA in sea urchin eggs that is destroyed at each cleavage division. *Cell*, 33(2), pp.389–396.

Faesen, A.C. et al., 2017. Basis of catalytic assembly of the mitotic checkpoint complex. *Nature*, 542(7642), pp.498–502.

Falk, S.J. et al., 2015. Chromosomes. CENP-C reshapes and stabilizes CENP-A nucleosomes at the centromere. *Science (New York, N.Y.)*, 348(6235), pp.699–703.

Fang, J. et al., 2015. Structural transitions of centromeric chromatin regulate the cell cycle-dependent recruitment of CENP-N. *Genes & development*, 29(10), pp.1058–1073.

Farnung, L. et al., 2017. Nucleosome-Chd1 structure and implications for chromatin remodelling. *Nature*, 550(7677), pp.539–542.

Fernandez-Leiro, R. & Scheres, S.H.W., 2017. A pipeline approach to single-particle processing in RELION. *Acta crystallographica. Section D, Structural biology*, 73(Pt 6), pp.496–502.

Flemming, W., 1882. *Zellsubstanz, Kern und Zelltheilung*,

Foley, E.A. & Kapoor, T.M., 2013. Microtubule attachment and spindle assembly checkpoint signalling at the kinetochore. *Nature reviews. Molecular cell biology*, 14(1), pp.25–37.

Foltz, D.R. et al., 2006. The human CENP-A centromeric nucleosome-associated complex. *Nature cell biology*, 8(5), pp.458–469.

Fraschini, R. et al., 2001. Bub3 interaction with Mad2, Mad3 and Cdc20 is mediated by WD40 repeats and does not require intact kinetochores. *The EMBO journal*, 20(23), pp.6648–6659.

Fujita, Y. et al., 2007. Priming of centromere for CENP-A recruitment by human hMis18alpha, hMis18beta, and M18BP1. *Developmental cell*, 12(1), pp.17–30.

Fukagawa, T. & Brown, W.R., 1997. Efficient conditional mutation of the vertebrate CENP-C gene. *Human molecular genetics*, 6(13), pp.2301–2308.

Fukagawa, T. & Earnshaw, W.C., 2014. The Centromere: Chromatin Foundation for the Kinetochore Machinery. *Developmental cell*, 30(5), pp.497–509.

Fukagawa, T. et al., 1999. CENP-C is necessary but not sufficient to induce formation of a functional centromere. *The EMBO journal*, 18(15), pp.4196–4209.

Fukagawa, T. et al., 2001. CENP-H, a constitutive centromere component, is required for centromere targeting of CENP-C in vertebrate cells. *The EMBO journal*, 20(16), pp.4603–4617.

Funabiki, H. & Wynne, D.J., 2013. Making an effective switch at the kinetochore by phosphorylation and dephosphorylation. *Chromosoma*, 122(3), pp.135–158.

Funabiki, H. et al., 1996. Cut2 proteolysis required for sister-chromatid seperation in fission yeast. *Nature*, 381(6581), pp.438–441.

Gaitanos, T.N. et al., 2009. Stable kinetochore-microtubule interactions depend on the Ska complex and its new component Ska3/C13Orf3. *The EMBO journal*, 28(10), pp.1442–1452.

Gascoigne, K.E. & Cheeseman, I.M., 2013. CDK-dependent phosphorylation and nuclear exclusion coordinately control kinetochore assembly state. *The Journal of cell*

biology, 201(1), pp.23–32.

Gascoigne, K.E. & Cheeseman, I.M., 2011. Kinetochore assembly: if you build it, they will come. *Current opinion in cell biology*, 23(1), pp.102–108.

Gascoigne, K.E. et al., 2011. Induced ectopic kinetochore assembly bypasses the requirement for CENP-A nucleosomes. *Cell*, 145(3), pp.410–422.

Gassmann, R. et al., 2008. A new mechanism controlling kinetochore-microtubule interactions revealed by comparison of two dynein-targeting components: SPDL-1 and the Rod/Zwilch/Zw10 complex. *Genes & development*, 22(17), pp.2385–2399.

Glotzer, M., Murray, A.W. & Kirschner, M.W., 1991. Cyclin is degraded by the ubiquitin pathway. *Nature*, 349(6305), pp.132–138.

Guo, L.Y. et al., 2017. Centromeres are maintained by fastening CENP-A to DNA and directing an arginine anchor-dependent nucleosome transition. *Nature communications*, 8, p.15775.

Guo, Q. et al., 2013. Structural insights into the role of the Chl4-Iml3 complex in kinetochore assembly. *Acta crystallographica. Section D, Biological crystallography*, 69(Pt 12), pp.2412–2419.

Guse, A., Fuller, C.J. & Straight, A.F., 2012. A cell-free system for functional centromere and kinetochore assembly. *Nature protocols*, 7(10), pp.1847–1869.

Hanisch, A., Silljé, H.H.W. & Nigg, E.A., 2006. Timely anaphase onset requires a novel spindle and kinetochore complex comprising Ska1 and Ska2. *The EMBO journal*, 25(23), pp.5504–5515.

Harashima, H., Dissmeyer, N. & Schnittger, A., 2013. Cell cycle control across the eukaryotic kingdom. *Trends in cell biology*, 23(7), pp.345–356.

Hardwick, K.G. et al., 2000. MAD3 encodes a novel component of the spindle checkpoint which interacts with Bub3p, Cdc20p, and Mad2p. *The Journal of cell biology*, 148(5), pp.871–882.

Hartwell, L.H., Culotti, J. & Reid, B., 1970. Genetic control of the cell-division cycle in yeast. I. Detection of mutants. *Proceedings of the National Academy of Sciences of*

the United States of America, 66(2), pp.352–359.

Hasson, D. et al., 2013. The octamer is the major form of CENP-A nucleosomes at human centromeres. *Nature structural & molecular biology*, 20(6), pp.687–695.

Hayashi, T. et al., 2004. Mis16 and Mis18 are required for CENP-A loading and histone deacetylation at centromeres. *Cell*, 118(6), pp.715–729.

Hays, T.S., Wise, D. & Salmon, E.D., 1982. Traction force on a kinetochore at metaphase acts as a linear function of kinetochore fiber length. *The Journal of cell biology*, 93(2), pp.374–389.

Hegemann, J.H. & Fleig, U.N., 1993. The centromere of budding yeast. *BioEssays : news and reviews in molecular, cellular and developmental biology*, 15(7), pp.451–460.

Hellwig, D. et al., 2011. Dynamics of CENP-N kinetochore binding during the cell cycle. *Journal of cell science*, 124(Pt 22), pp.3871–3883.

Henikoff, S. et al., 2000. Heterochromatic deposition of centromeric histone H3-like proteins. *Proceedings of the National Academy of Sciences of the United States of America*, 97(2), pp.716–721.

Herzog, F. et al., 2009. Structure of the anaphase-promoting complex/cyclosome interacting with a mitotic checkpoint complex. *Science (New York, N.Y.)*, 323(5920), pp.1477–1481.

Hewitt, L. et al., 2010. Sustained Mps1 activity is required in mitosis to recruit O-Mad2 to the Mad1-C-Mad2 core complex. *The Journal of cell biology*, 190(1), pp.25–34.

Hinshaw, S.M. & Harrison, S.C., 2013. An Iml3-Chl4 heterodimer links the core centromere to factors required for accurate chromosome segregation. *Cell reports*, 5(1), pp.29–36.

Hoffmann, S. et al., 2016. CENP-A Is Dispensable for Mitotic Centromere Function after Initial Centromere/Kinetochore Assembly. *Cell reports*, 17(9), pp.2394–2404.

Holloway, S.L. et al., 1993. Anaphase is initiated by proteolysis rather than by the inactivation of maturation-promoting factor. *Cell*, 73(7), pp.1393–1402.

Hori, T. & Fukagawa, T., 2012. Establishment of the vertebrate kinetochores. *Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology*, 20(5), pp.547–561.

Hori, T. et al., 2014. Histone H4 Lys 20 monomethylation of the CENP-A nucleosome is essential for kinetochore assembly. *Developmental cell*, 29(6), pp.740–749.

Hori, T. et al., 2013. The CCAN recruits CENP-A to the centromere and forms the structural core for kinetochore assembly. *The Journal of cell biology*, 200(1), pp.45–60.

Hori, T., Amano, M., et al., 2008. CCAN makes multiple contacts with centromeric DNA to provide distinct pathways to the outer kinetochore. *Cell*, 135(6), pp.1039–1052.

Hori, T., Okada, M., et al., 2008. CENP-O class proteins form a stable complex and are required for proper kinetochore function. *Molecular biology of the cell*, 19(3), pp.843–854.

Hornung, P. et al., 2011. Molecular architecture and connectivity of the budding yeast Mtw1 kinetochore complex. *Journal of molecular biology*, 405(2), pp.548–559.

Howell, B.J. et al., 2001. Cytoplasmic dynein/dynactin drives kinetochore protein transport to the spindle poles and has a role in mitotic spindle checkpoint inactivation. *The Journal of cell biology*, 155(7), pp.1159–1172.

Hoyt, M.A., Totis, L. & Roberts, B.T., 1991. S. cerevisiae genes required for cell cycle arrest in response to loss of microtubule function. *Cell*, 66(3), pp.507–517.

Hu, H. et al., 2011. Structure of a CENP-A-histone H4 heterodimer in complex with chaperone HJURP. *Genes & development*, 25(9), pp.901–906.

Huis in 't Veld, P.J. et al., 2016. Molecular basis of outer kinetochore assembly on CENP-T. *eLife*, 5.

Izawa, D. & Pines, J., 2012. Mad2 and the APC/C compete for the same site on Cdc20 to ensure proper chromosome segregation. *The Journal of cell biology*, 199(1), pp.27–37.

Izuta, H. et al., 2006. Comprehensive analysis of the ICEN (Interphase Centromere

Complex) components enriched in the CENP-A chromatin of human cells. *Genes to cells : devoted to molecular & cellular mechanisms*, 11(6), pp.673–684.

Joglekar, A.P., Bloom, K. & Salmon, E.D., 2009. In vivo protein architecture of the eukaryotic kinetochore with nanometer scale accuracy. *Current biology : CB*, 19(8), pp.694–699.

Joglekar, A.P., Bloom, K.S. & Salmon, E.D., 2010. Mechanisms of force generation by end-on kinetochore-microtubule attachments. *Current opinion in cell biology*, 22(1), pp.57–67.

Jones, I. & Morikawa, Y., 1996. Baculovirus vectors for expression in insect cells. *Current opinion in biotechnology*, 7(5), pp.512–516.

Kabsch, W., 2010. Integration, scaling, space-group assignment and post-refinement. *Acta crystallographica. Section D, Biological crystallography*, 66(Pt 2), pp.133–144.

Kagawa, N. et al., 2014. The CENP-O complex requirement varies among different cell types. *Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology*, 22(3), pp.293–303.

Karess, R., 2005. Rod-Zw10-Zwilch: a key player in the spindle checkpoint. *Trends in cell biology*, 15(7), pp.386–392.

Kato, H. et al., 2013. A conserved mechanism for centromeric nucleosome recognition by centromere protein CENP-C. *Science (New York, N.Y.)*, 340(6136), pp.1110–1113.

Kerres, A., Jakopec, V. & Fleig, U., 2007. The conserved Spc7 protein is required for spindle integrity and links kinetochore complexes in fission yeast. *Molecular biology of the cell*, 18(7), pp.2441–2454.

Kitagawa, K. & Hieter, P., 2001. Evolutionary conservation between budding yeast and human kinetochores. *Nature reviews. Molecular cell biology*, 2(9), pp.678–687.

Kiyomitsu, T., Murakami, H. & Yanagida, M., 2011. Protein interaction domain mapping of human kinetochore protein Blinkin reveals a consensus motif for binding of spindle assembly checkpoint proteins Bub1 and BubR1. *Molecular and cellular biology*, 31(5), pp.998–1011.

Kiyomitsu, T., Obuse, C. & Yanagida, M., 2007. Human Blinkin/AF15q14 is required for chromosome alignment and the mitotic checkpoint through direct interaction with Bub1 and BubR1. *Developmental cell*, 13(5), pp.663–676.

Klare, K. et al., 2015. CENP-C is a blueprint for constitutive centromere-associated network assembly within human kinetochores. *The Journal of cell biology*, 210(1), pp.11–22.

Klebig, C., Korinth, D. & Meraldi, P., 2009. Bub1 regulates chromosome segregation in a kinetochore-independent manner. *The Journal of cell biology*, 185(5), pp.841–858.

Kline, S.L. et al., 2006. The human Mis12 complex is required for kinetochore assembly and proper chromosome segregation. *The Journal of cell biology*, 173(1), pp.9–17.

Kops, G.J.P.L. et al., 2005. ZW10 links mitotic checkpoint signaling to the structural kinetochore. *The Journal of cell biology*, 169(1), pp.49–60.

Krenn, V. et al., 2014. KI motifs of human Knl1 enhance assembly of comprehensive spindle checkpoint complexes around MELT repeats. *Current biology : CB*, 24(1), pp.29–39.

Krenn, V. et al., 2012. Structural analysis reveals features of the spindle checkpoint kinase Bub1-kinetochore subunit Knl1 interaction. *The Journal of cell biology*, 196(4), pp.451–467.

Kucukelbir, A., Sigworth, F.J. & Tagare, H.D., 2014. Quantifying the local resolution of cryo-EM density maps. *Nature methods*, 11(1), pp.63–65.

Kursel, L.E. & Malik, H.S., 2016. Centromeres. *Current biology : CB*, 26(12), pp.R487–R490.

Kwon, M.-S. et al., 2007. CENP-C is involved in chromosome segregation, mitotic checkpoint function, and kinetochore assembly. *Molecular biology of the cell*, 18(6), pp.2155–2168.

Lampson, M.A. & Kapoor, T.M., 2005. The human mitotic checkpoint protein BubR1 regulates chromosome-spindle attachments. *Nature cell biology*, 7(1), pp.93–98.

Lampson, M.A. et al., 2004. Correcting improper chromosome-spindle attachments during cell division. *Nature cell biology*, 6(3), pp.232–237.

Lanini, L. & McKeon, F., 1995. Domains required for CENP-C assembly at the kinetochore. *Molecular biology of the cell*, 6(8), pp.1049–1059.

Lara-Gonzalez, P. et al., 2011. BubR1 blocks substrate recruitment to the APC/C in a KEN-box-dependent manner. *Journal of cell science*, 124(Pt 24), pp.4332–4345.

Lara-Gonzalez, P., Westhorpe, F.G. & Taylor, S.S., 2012. The spindle assembly checkpoint. *Current biology : CB*, 22(22), pp.R966–80.

Larsen, N.A. et al., 2007. Structural analysis of Bub3 interactions in the mitotic spindle checkpoint. *Proceedings of the National Academy of Sciences of the United States of America*, 104(4), pp.1201–1206.

Li, R. & Murray, A.W., 1991. Feedback control of mitosis in budding yeast. *Cell*, 66(3), pp.519–531.

Li, X. & Nicklas, R.B., 1995. Mitotic forces control a cell-cycle checkpoint. *Nature*, 373(6515), pp.630–632.

Liu, D. et al., 2010. Regulated targeting of protein phosphatase 1 to the outer kinetochore by KNL1 opposes Aurora B kinase. *The Journal of cell biology*, 188(6), pp.809–820.

Liu, X. et al., 2017. Mechanism of chromatin remodelling revealed by the Snf2-nucleosome structure. *Nature*, 544(7651), pp.440–445.

Liu, X. et al., 2005. Molecular analysis of kinetochore architecture in fission yeast. *The EMBO journal*, 24(16), pp.2919–2930.

Liu, Y. et al., 2016. Insights from the reconstitution of the divergent outer kinetochore of Drosophila melanogaster. *Open biology*, 6(2), p.150236.

Loog, M. & Morgan, D.O., 2005. Cyclin specificity in the phosphorylation of cyclin-dependent kinase substrates. *Nature*, 434(7029), pp.104–108.

Luger, K. et al., 1997. Crystal structure of the nucleosome core particle at 2.8 A

resolution. *Nature*, 389(6648), pp.251–260.

Luo, X. et al., 2000. Structure of the Mad2 spindle assembly checkpoint protein and its interaction with Cdc20. *Nature structural biology*, 7(3), pp.224–229.

Maciejowski, J. et al., 2010. Mps1 directs the assembly of Cdc20 inhibitory complexes during interphase and mitosis to control M phase timing and spindle checkpoint signaling. *The Journal of cell biology*, 190(1), pp.89–100.

Maddox, P.S. et al., 2004. "Holo"er than thou: chromosome segregation and kinetochore function in C. elegans. *Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology*, 12(6), pp.641–653.

Maddox, P.S. et al., 2007. Functional genomics identifies a Myb domain-containing protein family required for assembly of CENP-A chromatin. *The Journal of cell biology*, 176(6), pp.757–763.

Maiato, H. et al., 2004. The dynamic kinetochore-microtubule interface. *Journal of cell science*, 117(Pt 23), pp.5461–5477.

Maldonado, M. & Kapoor, T.M., 2011. Moving right along: how PP1 helps clear the checkpoint. *Developmental cell*, 20(6), pp.733–734.

Malik, H.S. & Henikoff, S., 2009. Major evolutionary transitions in centromere complexity. *Cell*, 138(6), pp.1067–1082.

Malvezzi, F. et al., 2013. A structural basis for kinetochore recruitment of the Ndc80 complex via two distinct centromere receptors. *The EMBO journal*, 32(3), pp.409–423.

Mapelli, M. et al., 2007. The Mad2 conformational dimer: structure and implications for the spindle assembly checkpoint. *Cell*, 131(4), pp.730–743.

Maresca, T.J. & Salmon, E.D., 2009. Intrakinetochore stretch is associated with changes in kinetochore phosphorylation and spindle assembly checkpoint activity. *The Journal of cell biology*, 184(3), pp.373–381.

Martin-Lluesma, S., Stucke, V.M. & Nigg, E.A., 2002. Role of Hec1 in spindle checkpoint signaling and kinetochore recruitment of Mad1/Mad2. *Science (New York,*

N.Y.), 297(5590), pp.2267–2270.

Maskell, D.P., Hu, X.-W. & Singleton, M.R., 2010. Molecular architecture and assembly of the yeast kinetochore MIND complex. *The Journal of cell biology*, 190(5), pp.823–834.

McAinsh, A.D. et al., 2006. The human kinetochore proteins Nnf1R and Mcm21R are required for accurate chromosome segregation. *The EMBO journal*, 25(17), pp.4033–4049.

McClelland, S.E. et al., 2007. The CENP-A NAC/CAD kinetochore complex controls chromosome congression and spindle bipolarity. *The EMBO journal*, 26(24), pp.5033–5047.

McEwen, B.F. et al., 1998. A new look at kinetochore structure in vertebrate somatic cells using high-pressure freezing and freeze substitution. *Chromosoma*, 107(6-7), pp.366–375.

McEwen, B.F., Dong, Y. & VandenBeldt, K.J., 2007. Using electron microscopy to understand functional mechanisms of chromosome alignment on the mitotic spindle. *Methods in cell biology*, 79, pp.259–293.

McKinley, K.L. & Cheeseman, I.M., 2014. Polo-like kinase 1 licenses CENP-A deposition at centromeres. *Cell*, 158(2), pp.397–411.

McKinley, K.L. et al., 2015. The CENP-L-N Complex Forms a Critical Node in an Integrated Meshwork of Interactions at the Centromere-Kinetochore Interface. *Molecular cell*, 60(6), pp.886–898.

Meadows, J.C. et al., 2011. Spindle checkpoint silencing requires association of PP1 to both Spc7 and kinesin-8 motors. *Developmental cell*, 20(6), pp.739–750.

Meluh, P.B. & Koshland, D., 1995. Evidence that the MIF2 gene of Saccharomyces cerevisiae encodes a centromere protein with homology to the mammalian centromere protein CENP-C. *Molecular biology of the cell*, 6(7), pp.793–807.

Meraldi, P. et al., 2006. Phylogenetic and structural analysis of centromeric DNA and kinetochore proteins. *Genome biology*, 7(3), p.R23.

Milks, K.J., Moree, B. & Straight, A.F., 2009. Dissection of CENP-C-directed centromere and kinetochore assembly. *Molecular biology of the cell*, 20(19), pp.4246–4255.

Miller, M.P., Asbury, C.L. & Biggins, S., 2016. A TOG Protein Confers Tension Sensitivity to Kinetochore-Microtubule Attachments. *Cell*, 165(6), pp.1428–1439.

Miller, S.A., Johnson, M.L. & Stukenberg, P.T., 2008. Kinetochore attachments require an interaction between unstructured tails on microtubules and Ndc80(Hec1). *Current biology : CB*, 18(22), pp.1785–1791.

Minoshima, Y. et al., 2005. The constitutive centromere component CENP-50 is required for recovery from spindle damage. *Molecular and cellular biology*, 25(23), pp.10315–10328.

Miranda, J.J.L. et al., 2005. The yeast DASH complex forms closed rings on microtubules. *Nature structural & molecular biology*, 12(2), pp.138–143.

Morgan, D.O., 1995. Principles of CDK regulation. *Nature*, 374(6518), pp.131–134.

Mosalaganti, S. et al., 2017. Structure of the RZZ complex and molecular basis of its interaction with Spindly. *The Journal of cell biology*, 216(4), pp.961–981.

Musacchio, A. & Desai, A., 2017. A Molecular View of Kinetochore Assembly and Function. *Biology*, 6(1).

Musacchio, A. & Salmon, E.D., 2007. The spindle-assembly checkpoint in space and time. *Nature reviews. Molecular cell biology*, 8(5), pp.379–393.

Müller, S. et al., 2014. Phosphorylation and DNA binding of HJURP determine its centromeric recruitment and function in CenH3(CENP-A) loading. *Cell reports*, 8(1), pp.190–203.

Nagpal, H. et al., 2015. Dynamic changes in CCAN organization through CENP-C during cell-cycle progression. *Molecular biology of the cell*, 26(21), pp.3768–3776.

Narlikar, G.J., Sundaramoorthy, R. & Owen-Hughes, T., 2013. Mechanisms and functions of ATP-dependent chromatin-remodeling enzymes. *Cell*, 154(3), pp.490–503.

Nekrasov, V.S. et al., 2003. Interactions between centromere complexes in Saccharomyces cerevisiae. *Molecular biology of the cell*, 14(12), pp.4931–4946.

Nezi, L. & Musacchio, A., 2009. Sister chromatid tension and the spindle assembly checkpoint. *Current opinion in cell biology*, 21(6), pp.785–795.

Nicklas, R.B. & Koch, C.A., 1969. Chromosome micromanipulation. 3. Spindle fiber tension and the reorientation of mal-oriented chromosomes. *The Journal of cell biology*, 43(1), pp.40–50.

Nishihashi, A. et al., 2002. CENP-I is essential for centromere function in vertebrate cells. *Developmental cell*, 2(4), pp.463–476.

Nishino, T. et al., 2012. CENP-T-W-S-X forms a unique centromeric chromatin structure with a histone-like fold. *Cell*, 148(3), pp.487–501.

Obuse, C. et al., 2004. Proteomics analysis of the centromere complex from HeLa interphase cells: UV-damaged DNA binding protein 1 (DDB-1) is a component of the CEN-complex, while BMI-1 is transiently co-localized with the centromeric region in interphase. *Genes to cells : devoted to molecular & cellular mechanisms*, 9(2), pp.105–120.

Okada, M. et al., 2006. The CENP-H-I complex is required for the efficient incorporation of newly synthesized CENP-A into centromeres. *Nature cell biology*, 8(5), pp.446–457.

Overlack, K., Krenn, V. & Musacchio, A., 2014. When Mad met Bub. *EMBO reports*, 15(4), pp.326–328.

Pan, D. et al., 2017. CDK-regulated dimerization of M18BP1 on a Mis18 hexamer is necessary for CENP-A loading. *eLife*, 6.

Pardee, A.B., 1974. A restriction point for control of normal animal cell proliferation. *Proceedings of the National Academy of Sciences of the United States of America*, 71(4), pp.1286–1290.

Perera, D. et al., 2007. Bub1 maintains centromeric cohesion by activation of the spindle checkpoint. *Developmental cell*, 13(4), pp.566–579.

Perpelescu, M. & Fukagawa, T., 2011. The ABCs of CENPs. *Chromosoma*, 120(5), pp.425–446.

Perpelescu, M. et al., 2009. Active establishment of centromeric CENP-A chromatin by RSF complex. *The Journal of cell biology*, 185(3), pp.397–407.

Pesenti, M.E. et al., 2018. Reconstitution of a 26-Subunit Human Kinetochore Reveals Cooperative Microtubule Binding by CENP-OPQUR and NDC80. *Molecular cell*, 71(6), pp.923–939.e10.

Pesenti, M.E., Weir, J.R. & Musacchio, A., 2016. Progress in the structural and functional characterization of kinetochores. *Current opinion in structural biology*, 37, pp.152–163.

Petrovic, A. et al., 2014. Modular assembly of RWD domains on the Mis12 complex underlies outer kinetochore organization. *Molecular cell*, 53(4), pp.591–605.

Petrovic, A. et al., 2016. Structure of the MIS12 Complex and Molecular Basis of Its Interaction with CENP-C at Human Kinetochores. *Cell*, 167(4), pp.1028–1040.e15.

Petrovic, A. et al., 2010. The MIS12 complex is a protein interaction hub for outer kinetochore assembly. *The Journal of cell biology*, 190(5), pp.835–852.

Pines, J., 2011. Cubism and the cell cycle: the many faces of the APC/C. *Nature reviews. Molecular cell biology*, 12(7), pp.427–438.

Pines, J., 1991. Cyclins: wheels within wheels. *Cell growth & differentiation : the molecular biology journal of the American Association for Cancer Research*, 2(6), pp.305–310.

Pinsky, B.A. et al., 2006. The Ipl1-Aurora protein kinase activates the spindle checkpoint by creating unattached kinetochores. *Nature cell biology*, 8(1), pp.78–83.

Pluta, A.F. et al., 1995. The centromere: hub of chromosomal activities. *Science (New York, N.Y.)*, 270(5242), pp.1591–1594.

Powers, A.F. et al., 2009. The Ndc80 kinetochore complex forms load-bearing attachments to dynamic microtubule tips via biased diffusion. *Cell*, 136(5), pp.865–875.

Primorac, I. et al., 2013. Bub3 reads phosphorylated MELT repeats to promote spindle assembly checkpoint signaling. *eLife*, 2, p.e01030.

Prinz, F. et al., 2016. Functional and Structural Characterization of Bub3·BubR1 Interactions Required for Spindle Assembly Checkpoint Signaling in Human Cells. *The Journal of biological chemistry*, 291(21), pp.11252–11267.

Przewloka, M.R. et al., 2011. CENP-C is a structural platform for kinetochore assembly. *Current biology : CB*, 21(5), pp.399–405.

Renou, J.-P. et al., 2003. Identification of genes differentially expressed in mouse mammary epithelium transformed by an activated beta-catenin. *Oncogene*, 22(29), pp.4594–4610.

Rosenberg, J.S., Cross, F.R. & Funabiki, H., 2011. KNL1/Spc105 recruits PP1 to silence the spindle assembly checkpoint. *Current biology : CB*, 21(11), pp.942–947.

Saitoh, H. et al., 1992. CENP-C, an autoantigen in scleroderma, is a component of the human inner kinetochore plate. *Cell*, 70(1), pp.115–125.

Sanchez-Pulido, L. et al., 2009. Common ancestry of the CENP-A chaperones Scm3 and HJURP. *Cell*, 137(7), pp.1173–1174.

Santaguida, S. & Musacchio, A., 2009. The life and miracles of kinetochores. *The EMBO journal*, 28(17), pp.2511–2531.

Santaguida, S. et al., 2011. Evidence that Aurora B is implicated in spindle checkpoint signalling independently of error correction. *The EMBO journal*, 30(8), pp.1508–1519.

Saurin, A.T., 2018. Kinase and Phosphatase Cross-Talk at the Kinetochore. *Frontiers in cell and developmental biology*, 6, p.62.

Scaërou, F., The ZW10 and Rough Deal checkpoint proteins function together in a large, evolutionarily conserved complex targeted to the kinetochore. *Journal of cell science*, 114(17), p.3103.

Scaërou, F. et al., 1999. The rough deal protein is a new kinetochore component required for accurate chromosome segregation in Drosophila. *Journal of cell science*, 112 (Pt 21), pp.3757–3768.

Schleiffer, A. et al., 2012. CENP-T proteins are conserved centromere receptors of the Ndc80 complex. *Nature cell biology*, 14(6), pp.604–613.

Schmidt, J.C. et al., 2012. The kinetochore-bound Ska1 complex tracks depolymerizing microtubules and binds to curved protofilaments. *Developmental cell*, 23(5), pp.968–980.

Screpanti, E. et al., 2011. Direct binding of Cenp-C to the Mis12 complex joins the inner and outer kinetochore. *Current biology : CB*, 21(5), pp.391–398.

Shelby, R.D., Vafa, O. & Sullivan, K.F., 1997. Assembly of CENP-A into centromeric chromatin requires a cooperative array of nucleosomal DNA contact sites. *The Journal of cell biology*, 136(3), pp.501–513.

Shepperd, L.A. et al., 2012. Phosphodependent recruitment of Bub1 and Bub3 to Spc7/KNL1 by Mph1 kinase maintains the spindle checkpoint. *Current biology : CB*, 22(10), pp.891–899.

Shuaib, M. et al., 2010. HJURP binds CENP-A via a highly conserved N-terminal domain and mediates its deposition at centromeres. *Proceedings of the National Academy of Sciences of the United States of America*, 107(4), pp.1349–1354.

Silva, M.C.C. et al., 2012. Cdk activity couples epigenetic centromere inheritance to cell cycle progression. *Developmental cell*, 22(1), pp.52–63.

Sironi, L. et al., 2002. Crystal structure of the tetrameric Mad1-Mad2 core complex: implications of a "safety belt" binding mechanism for the spindle checkpoint. *The EMBO journal*, 21(10), pp.2496–2506.

Starr, D.A. et al., 2000. HZwint-1, a novel human kinetochore component that interacts with HZW10. *Journal of cell science*, 113 (Pt 11), pp.1939–1950.

Stellfox, M.E., Bailey, A.O. & Foltz, D.R., 2013. Putting CENP-A in its place. *Cellular and molecular life sciences : CMLS*, 70(3), pp.387–406.

Sudakin, V. et al., 1995. The cyclosome, a large complex containing cyclin-selective ubiquitin ligase activity, targets cyclins for destruction at the end of mitosis. *Molecular biology of the cell*, 6(2), pp.185–197.

Sudakin, V., Chan, G.K. & Yen, T.J., 2001. Checkpoint inhibition of the APC/C in HeLa cells is mediated by a complex of BUBR1, BUB3, CDC20, and MAD2. *The Journal of cell biology*, 154(5), pp.925–936.

Sugata, N. et al., 2000. Human CENP-H multimers colocalize with CENP-A and CENP-C at active centromere--kinetochore complexes. *Human molecular genetics*, 9(19), pp.2919–2926.

Sugata, N., Munekata, E. & Todokoro, K., 1999. Characterization of a novel kinetochore protein, CENP-H. *The Journal of biological chemistry*, 274(39), pp.27343–27346.

Sugimoto, K. et al., 1994. Human centromere protein C (CENP-C) is a DNA-binding protein which possesses a novel DNA-binding motif. *Journal of biochemistry*, 116(4), pp.877–881.

Suijkerbuijk, S.J.E. et al., 2012. Integration of kinase and phosphatase activities by BUBR1 ensures formation of stable kinetochore-microtubule attachments. *Developmental cell*, 23(4), pp.745–755.

Tachiwana, H. et al., 2011. Crystal structure of the human centromeric nucleosome containing CENP-A. *Nature*, 476(7359), pp.232–235.

Takahashi, K., Chen, E.S. & Yanagida, M., 2000. Requirement of Mis6 centromere connector for localizing a CENP-A-like protein in fission yeast. *Science (New York, N.Y.)*, 288(5474), pp.2215–2219.

Takeuchi, K. et al., 2014. The centromeric nucleosome-like CENP-T-W-S-X complex induces positive supercoils into DNA. *Nucleic acids research*, 42(3), pp.1644–1655.

Tien, J.F. et al., 2010. Cooperation of the Dam1 and Ndc80 kinetochore complexes enhances microtubule coupling and is regulated by aurora B. *The Journal of cell biology*, 189(4), pp.713–723.

Tighe, A., Johnson, V.L. & Taylor, S.S., 2004. Truncating APC mutations have dominant effects on proliferation, spindle checkpoint control, survival and chromosome stability. *Journal of cell science*, 117(Pt 26), pp.6339–6353.

Tomkiel, J. et al., 1994. CENP-C is required for maintaining proper kinetochore size

and for a timely transition to anaphase. *The Journal of cell biology*, 125(3), pp.531–545.

Tooley, J. & Stukenberg, P.T., 2011. The Ndc80 complex: integrating the kinetochore's many movements. *Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology*, 19(3), pp.377–391.

Tooley, J.G., Miller, S.A. & Stukenberg, P.T., 2011. The Ndc80 complex uses a tripartite attachment point to couple microtubule depolymerization to chromosome movement. *Molecular biology of the cell*, 22(8), pp.1217–1226.

Trazzi, S. et al., 2009. The C-terminal domain of CENP-C displays multiple and critical functions for mammalian centromere formation. *PloS one*, 4(6), p.e5832.

Trowitzsch, S. et al., 2010. New baculovirus expression tools for recombinant protein complex production. *Journal of structural biology*, 172(1), pp.45–54.

Urnavicius, L. et al., 2015. The structure of the dynactin complex and its interaction with dynein. *Science (New York, N.Y.)*, 347(6229), pp.1441–1446.

Vafa, O. & Sullivan, K.F., 1997. Chromatin containing CENP-A and alpha-satellite DNA is a major component of the inner kinetochore plate. *Current biology : CB*, 7(11), pp.897–900.

Varma, D. & Salmon, E.D., 2012. The KMN protein network--chief conductors of the kinetochore orchestra. *Journal of cell science*, 125(Pt 24), pp.5927–5936.

Varma, D. et al., 2008. Direct role of dynein motor in stable kinetochore-microtubule attachment, orientation, and alignment. *The Journal of cell biology*, 182(6), pp.1045–1054.

Vasudevan, D., Chua, E.Y.D. & Davey, C.A., 2010. Crystal structures of nucleosome core particles containing the "601" strong positioning sequence. *Journal of molecular biology*, 403(1), pp.1–10.

Volkov, V.A. et al., 2018. Multivalency of NDC80 in the outer kinetochore is essential to track shortening microtubules and generate forces. *eLife*, 7.

Walczak, C.E. & Heald, R., 2008. Mechanisms of mitotic spindle assembly and

function. *International review of cytology*, 265, pp.111–158.

Wan, X. et al., 2009. Protein architecture of the human kinetochore microtubule attachment site. *Cell*, 137(4), pp.672–684.

Wang, F. et al., 2010. Histone H3 Thr-3 phosphorylation by Haspin positions Aurora B at centromeres in mitosis. *Science (New York, N.Y.)*, 330(6001), pp.231–235.

Warburton, P.E. et al., 1997. Immunolocalization of CENP-A suggests a distinct nucleosome structure at the inner kinetochore plate of active centromeres. *Current biology : CB*, 7(11), pp.901–904.

Weber, C.H. & Vincenz, C., 2001. The death domain superfamily: a tale of two interfaces? *Trends in biochemical sciences*, 26(8), pp.475–481.

Wei, R. et al., 2011. Phosphorylation of the Ndc80 complex protein, HEC1, by Nek2 kinase modulates chromosome alignment and signaling of the spindle assembly checkpoint. *Molecular biology of the cell*, 22(19), pp.3584–3594.

Weir, J.R. et al., 2016. Insights from biochemical reconstitution into the architecture of human kinetochores. *Nature*, 537(7619), pp.249–253.

Weiss, E. & Winey, M., 1996. The Saccharomyces cerevisiae spindle pole body duplication gene MPS1 is part of a mitotic checkpoint. *The Journal of cell biology*, 132(1-2), pp.111–123.

Welburn, J.P.I. et al., 2010. Aurora B phosphorylates spatially distinct targets to differentially regulate the kinetochore-microtubule interface. *Molecular cell*, 38(3), pp.383–392.

Welburn, J.P.I. et al., 2009. The human kinetochore Ska1 complex facilitates microtubule depolymerization-coupled motility. *Developmental cell*, 16(3), pp.374–385.

Westermann, S. & Schleiffer, A., 2013. Family matters: structural and functional conservation of centromere-associated proteins from yeast to humans. *Trends in cell biology*, 23(6), pp.260–269.

Westermann, S. et al., 2003. Architecture of the budding yeast kinetochore reveals a

conserved molecular core. *The Journal of cell biology*, 163(2), pp.215–222.

Westermann, S. et al., 2005. Formation of a dynamic kinetochore- microtubule interface through assembly of the Dam1 ring complex. *Molecular cell*, 17(2), pp.277–290.

Wittmann, T., Hyman, A. & Desai, A., 2001. The spindle: a dynamic assembly of microtubules and motors. *Nature cell biology*, 3(1), pp.E28–34.

Yamagishi, Y. et al., 2012. MPS1/Mph1 phosphorylates the kinetochore protein KNL1/Spc7 to recruit SAC components. *Nature cell biology*, 14(7), pp.746–752.

Yamaguchi, M. et al., 2016. Cryo-EM of Mitotic Checkpoint Complex-Bound APC/C Reveals Reciprocal and Conformational Regulation of Ubiquitin Ligation. *Molecular cell*, 63(4), pp.593–607.

Yamamoto, A., Guacci, V. & Koshland, D., 1996. Pds1p, an inhibitor of anaphase in budding yeast, plays a critical role in the APC and checkpoint pathway(s). *The Journal of cell biology*, 133(1), pp.99–110.

Yang, C.H. et al., 1996. Identification of overlapping DNA-binding and centromere-targeting domains in the human kinetochore protein CENP-C. *Molecular and cellular biology*, 16(7), pp.3576–3586.

Yu, Z. et al., 2015. Dynamic phosphorylation of CENP-A at Ser68 orchestrates its cell-cycle-dependent deposition at centromeres. *Developmental cell*, 32(1), pp.68–81.

Zasadzińska, E. et al., 2013. Dimerization of the CENP-A assembly factor HJURP is required for centromeric nucleosome deposition. *The EMBO journal*, 32(15), pp.2113–2124.

8 Acknowledgements

Acknowledgements

"The biography is not included in the online version for reasons of data protection"

Affidavit

Hiermit erkläre ich, Satya Krishna Pentakota, gem. § 7 Abs. (2) d) + f) der Promotionsordnung der Fakultät für Biologie zur Erlangung des Dr. rer. nat., dass ich die vorliegende Dissertation selbständig verfasst und mich keiner anderen als der angegebenen Hilfsmittel bedient, bei der Abfassung der Dissertation nur die angegebenen Hilfsmittel benutzt und alle wörtlich oder inhaltlich übernommenen Stellen als solche gekennzeichnet habe.

Datum: 26 |11|2018

Unterschrift:

Hiermit erkläre ich, Satya Krishna Pentakota, gem. § 7 Abs. (2) e) + g) der Promotionsordnung der Fakultät für Biologie zur Erlangung des Dr. rer. nat., dass ich keine anderen Promotionen bzw. Promotionsversuche in der Vergangenheit durchgeführt habe und dass diese Arbeit von keiner anderen Fakultät/Fachbereich abgelehnt worden ist.

Datum: 26|11|2018

Unterschrift:

Hiermit erkläre ich, Prof. Dr. Andrea Musacchio, gem. § 6 Abs. (2) g) der Promotionsordnung der Fakultät für Biologie zur Erlangung des Dr. rer. nat., dass ich das Arbeitsgebiet, dem das Thema "*Insight into the role of CENP-N in kinetochore structure and function*" zuzuordnen ist, in Forschung und Lehre vertrete und den Antrag von Satya Krishna Pentakota befürworte und die Betreuung auch im Falle eines Weggangs, wenn nicht wichtige Gründe dem entgegenstehen, weiterführen werde.

Datum:

Unterschrift: