The role of HJURP in centromere specification and inheritance throughout the cell cycle

Ewelina Zasadzińska
Włocławek, Poland

M.Sc. Biotechnology, Łódź University of Technology, 2012
M.Sc. Biochemistry, University of Virginia, 2014

A Dissertation Presented to the Graduate Faculty of the University of Virginia
in Candidacy for the Degree of Doctor of Philosophy

Department of Biochemistry and Molecular Genetics

University of Virginia
July 2018

Thanks and Dedications

Table of Contents:

List of Abbreviations:

APEX- engineered ascorbate peroxidase

Asf1- Anti-silencing function protein 1

BirA- Bifunctional ligase/repressor BirA

CAF-1- chromatin assembly factor-1

CATD- CENP-A targeting domain

CCAN- Constitutive centromere associated network

CDK- Cyclin-dependent kinase

CENP-A through -X Centromere Protein -A through -X

CID- Drosophila CENP-A homolog

Cnp1- S. pombe CENP-A homolog

Cse4- S. cerevisiae CENP-A homolog

FACT- Facilitates chromatin transcription

HCTD21/2-Hjurp carboxy terminal domain 1/2

HJURP- Holliday junction recognition protein

LacI- Lac repressor

LacO- Lac operator

LAP- Localization and Affinity Purification

MBP- Maltose binding protein

MCM2-7- Minichromosome maintenance protein complex

Mis12, 16 and 18 - Missegregation of chromosomes phenotype mutants

Mis18BP1- Mis18 Binding Protein 1

PCNA- Proliferating cell nuclear antigen

RF- Replication fork

SCM- Suppressor of chromosome missegregation

SILAC- Stable isotope labeling by/with amino acids in cell culture

SUMO- Small Ubiquitin-like Modifier

TIR1-Transport inhibitor response 1

List of Figures:

Figure 4.7. HJURP copurifies with the MCM2-7 helicase complex and simultaneously interact with MCM2-CENP-A/H4 proteins.

Supplemental Figure Legends

Figure 4.S1.

Figure 4.S2.

Figure 4.S3.

Figure 4.S4.

Figure 5.1. WDR18 interacts with HJURP and Mis18α *in vivo*

Figure 5.2. WDR18 depletion affects HJURP but not Mis18α protein levels

Figure 5.3. The effects of WDR18 depletion on cell cycle progression and HJURP transcription

Figure 5.4. WDR18 possibly regulates HJURP stability during interphase

Figure 5.5. WDR18 depletion possibly impairs CENP-A and Mis18α centromeric localization

Figure 5.6 WDR18 influence on the status of HJURP posttranslational modifications

Abstract:

The faithful chromosome segregation during mitosis and meiosis is critical for ensuring that each daughter cell inherits the correct number of chromosomes to maintain genomic stability. The events of chromosome missegregation can lead to aneuploidy which is a hallmark of many diseases such as birth defects and cancer. Centromeres are chromosomal domains that direct the process of chromosome segregation by recruiting kinetochore apparatus during mitosis. Centromere identity in most eukaryotes is specified epigenetically by the incorporation of a centromere specific nucleosomes in which canonical histone H3 variant is replaced by the Centromere Protein A (CENP-A). Therefore, the assembly and propagation of centromeric nucleosomes are critical for maintaining centromere identity. Assembly of centromere specific nucleosomes in humans requires the dedicated CENP-A chaperone HJURP, and the Mis18 complex to couple the deposition of new CENP-A to the site of the pre-existing centromere. New CENP-A deposition occurs specifically in early G1 and during DNA replication existing CENP-A containing nucleosomes are stably inherited and partitioned to daughter strands. In this dissertation, I will describe how HJURP plays a dual role in centromere specification and is implicated in both: new CENP-A deposition as well as inheritance of existing CENP-A nucleosomes. Chapter one will contains a general introduction to how centromere identity is dictated and inherited across different species. Chapter two is dedicated to the requirement of HJURP self-association for new CENP-A deposition. In chapter three, I will describe optimization of proximity based labelling assays that we employed for identification of a mechanism governing CENP-A inheritance. The chapter four is dedicated to the role of HJURP and MCM2 chaperones in CENP-A retention across S phase. In chapter five I will describe new preliminary data exploring a potential role of WDR18 protein in HJURP protein stability.

Chapter 1 – The establishment and propagation of centromeric chromatin.

This chapter is based on the previously published book chapter:

Zasadzińska, E., and Foltz, D.R. (2017). Orchestrating the Specific Assembly of Centromeric Nucleosomes. Progress in molecular and subcellular biology 56, 165-192.

Reproduced with permission from Springer Nature Publisher.

This chapter is based on the previously published review:

Srivastava, S., Zasadzińska, E., and Foltz, D.R. (2018). Posttranslational mechanisms controlling centromere function and assembly. Current opinion in cell biology 52, 126-135.

Reproduced with permission from ScienceDirect, Elsevier.

General Introduction

The faithful chromosome segregation is critical for governing genomic stability. This process ensures that each daughter cell inherits the correct number of chromosomes during every single cell division event. Defects in chromosome segregation lead to aneuploidy which is a hallmark of many diseases such as cancer and birth defects. Therefore, understanding mechanisms regulating chromosome segregation is of critical importance.

The kinetochore is an enormous protein structure that assembles on the centromeric chromatin and directs chromosome segregation during mitosis and meiosis. The kinetochore apparatus is a highly complex assembly in which the inner kinetochore proteins are recruited to the centromere, while the outer kinetochore proteins provide a platform for the attachment of microtubules that emanate from the opposite spindle poles (Figure 1.1B). Once correctly attached, chromosomes are then pulled to the opposite poles of the spindle and equally segregated between daughter cells.

As centromeres are the chromatin domains that drive kinetochore assembly, therefore, it is critical for the centromeric chromatin to be precisely specified and stably propagated throughout multiple generations to facilitate correct kinetochore recruitment and subsequent faithful chromosome segregation. Centromeres In most species are characterized by the presence of unique nucleosomes containing the histone H3 variant Centromere Protein A (CENP-A) (Figure 1.1 A). However, different organisms employ distinct strategies to specify the centromere location. Budding yeast contain point centromeres which location is determined by the presence of approximately 150 bp domain with three distinct sequences: CDEI, CDEII and CDEIII (Clarke and Carbon, 1980; Fitzgerald-Hayes et al., 1982). In budding yeast, these sequences are sufficient for the establishment of a functional centromere. However, the wide variation of centromere DNA repeat sequences across species, and indeed the lack of DNA repetitive elements in several species suggest that DNA

sequence elements may not be critical for centromere function in higher eukaryotes. Moreover, the existence of neocentromeres and pseudodicentromeric chromosomes (Scott and Sullivan, 2013) strongly suggest that centromeres do not depend on the underlying DNA sequence for their inheritance but are epigenetic loci that are stably inhered through epigenetic processes. Indeed, in higher eukaryotes the centromeric chromatin is defined by epigenetic chromatin features, primarily by the presence of a centromere specific CENP-A histone variant, rather than underlying DNA sequence. CENP-A specification of epigenetic centromeres means that the process of nucleosome assembly is a key event in inheritance of the locus.

All histone H3 variants employ distinct mechanisms, facilitated by histone chaperones, which selectively recognize them upon synthesis and escort to the site of nucleosome assembly. Similarly, CENP-A uses its own specific machinery that orchestrates the spatiotemporal assembly of centromeric chromatin during the cell cycle. In humans, new CENP-A incorporation is a multistep mechanism that involves identification of centromeric chromatin for new CENP-A incorporation, deposition of newly synthesized CENP-A/H4 and stabilization of CENP-A nucleosomes. Each of those steps requires the activity of multiple protein factors which work together to ensure that CENP-A nucleosomes are deposited specifically at the centromeric domain, at the correct timing and only once per cell cycle.

Histone chaperones and centromere assembly

Incorporation of histones into the chromatin requires assembly factors or chaperones that work together to facilitate nucleosome deposition (Burgess and Zhang, 2013; Ransom et al., 2010). Histone H3 variants use their specific independent chaperone complexes that govern a selective recognition and facilitate their deposition in replication-dependent (H3.1 variant) or replication-independent (H3.3 and CENP-A variants) nucleosome assembly pathways (Sarma and Reinberg, 2005; Szenker et al., 2011; Weber and Henikoff, 2014) (Figure 1.1 A). The major histone variant H3.1 is deposited into newly replicated naked DNA during DNA replication via the CAF-1 complex that include p150, p60 and p46/48 (Tagami et al., 2004; Tyler et al., 1999; Tyler et al., 2001) The H3.3 variant is regulated by two chaperone complexes distinct from the H3.1 replication-dependent chaperones responsible for H3.1 deposition. The Hira chaperone is devoted to the genome-wide deposition of histone H3.3 at active and repressed genes (Chow et al., 2005; Goldberg et al., 2010; Lewis et al., 2010; Mito et al., 2005; Szenker et al., 2011; Tagami et al., 2004; Tamura et al., 2009). DAXX also acts as a chaperone for H3.3 and mediates H3.3 deposition at telomeric and pericentric heterochromatin in conjunction with the H3K9 binding protein ATRX (Goldberg et al., 2010; Lewis et al., 2010).
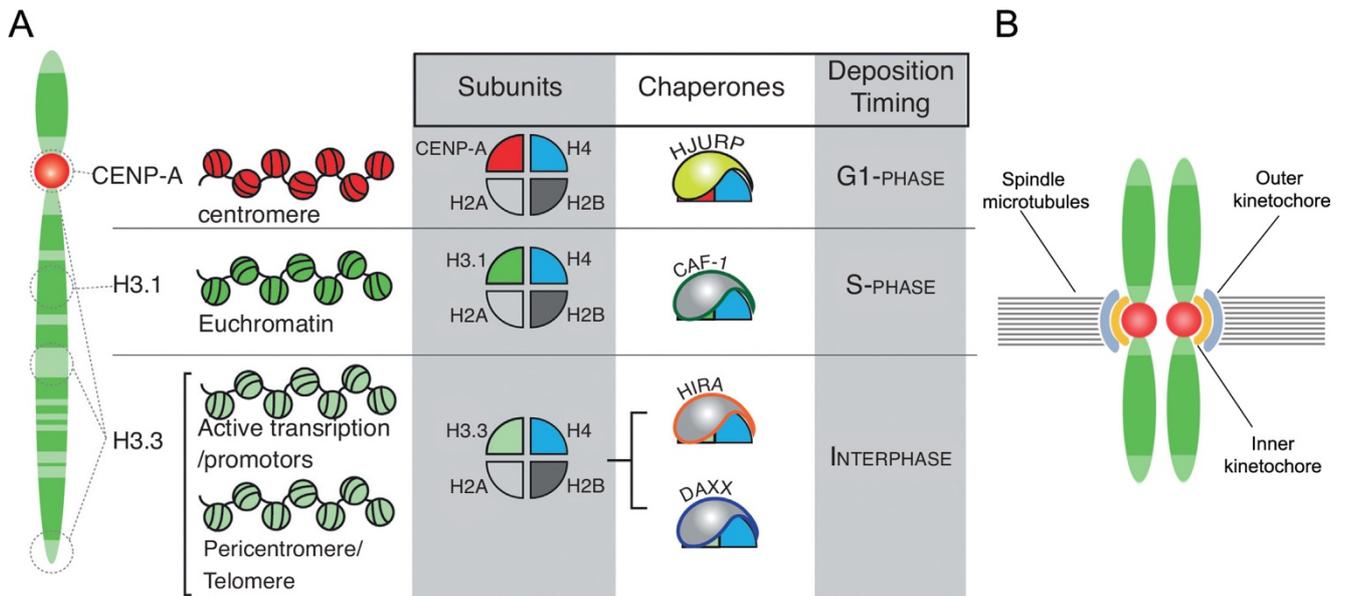
Figure 1.1. Centromeres are distinct chromatin domains that drive kinetochore assembly.

(A) The interaction of histone H3 variants with distinct chaperone proteins determines the timing and site of variant nucleosome deposition. (B) CENP-A containing centromeric chromatin specifies the site of the kinetochore assembly. Kinetochores contain inner and outer plates and are essential to facilitate microtubule attachments.

Figure modified from: Srivastava, S., Zasadzińska, E., and Foltz, D.R. (2018). Posttranslational mechanisms controlling centromere function and assembly. Current opinion in cell biology 52, 126-135.
Figure included with permission.

Similar to the other H3 variants, the centromere specific histone H3 variant CENP-A interacts with a dedicated chaperone prior to deposition into chromatin. Prenucleosomal human CENP-A associates with the Holliday junction recognition protein (HJURP) (Dunleavy et al., 2009; Foltz et al., 2009; Shuaib et al., 2010). HJURP is necessary for incorporation of CENP-A into the centromeric chromatin and is recruited to centromeres in early G1, when new CENP-A assembly is occurring (Fig 1.2) (Bernad et al., 2011; Dunleavy et al., 2009; Foltz et al., 2009; Jansen et al., 2007). Suppression of HJURP completely abolishes new CENP-A deposition, results in errors in kinetochore assembly and ultimately leads to a high rate of chromosome segregation defects (Dunleavy et al., 2009; Foltz et al., 2009).

The centromere targeting domain (CATD) of CENP-A is sufficient to determine the centromeric deposition of CENP-A. The CATD domain spans loop 1 and the α2 helix of CENP-A and when replaced with corresponding domain within canonical H3.1 was demonstrated to confer both HJURP binding and centromeric localization (Black et al., 2007; Foltz et al., 2009). His 104 and Leu112, residues within CATD C-terminal region together with either Asn85 or Gln89 within CATD N-terminus are sufficient to confer HJURP binding, but not sufficient to facilitate centromere incorporation (Bassett et al., 2012).

HJURP specifically recognizes the CATD domain of CENP-A through its N-terminal CENP-A binding domain (Fig. 1.3). The CENP-A binding domain of HJURP shares homology with the yeast Scm3 proteins that also act as CENP-A (Cse4, Cnp1) specific chaperone (Fig. 1.3, 1.4) (Camahort et al., 2007; Mizuguchi et al., 2007; Pidoux et al., 2009; Sanchez-Pulido et al., 2009; Stoler et al., 2007; Williams et al., 2009). Although the mechanism of centromere inheritance between budding yeast and humans is very different, both systems are dependent on a CENP-A specific histone chaperone. HJURP binds CENP-A through the conserved Scm3 domain. A number of residues within yeast Scm3 were proposed

to be essential for CENP-A$^{\text{Cnp1}}$ incorporation including Leucine 56 and Leucine 73. The fact that those key residues required for CENP-A$^{\text{Cnp1}}$ deposition are conserved as hydrophobic amino acids in other eukaryotes including humans, implies the mechanism by which CENP-A is selectively recognized and deposited at the centromeric chromatin by its chaperone is common in yeast and humans (Cho and Harrison, 2011; Pidoux et al., 2009).
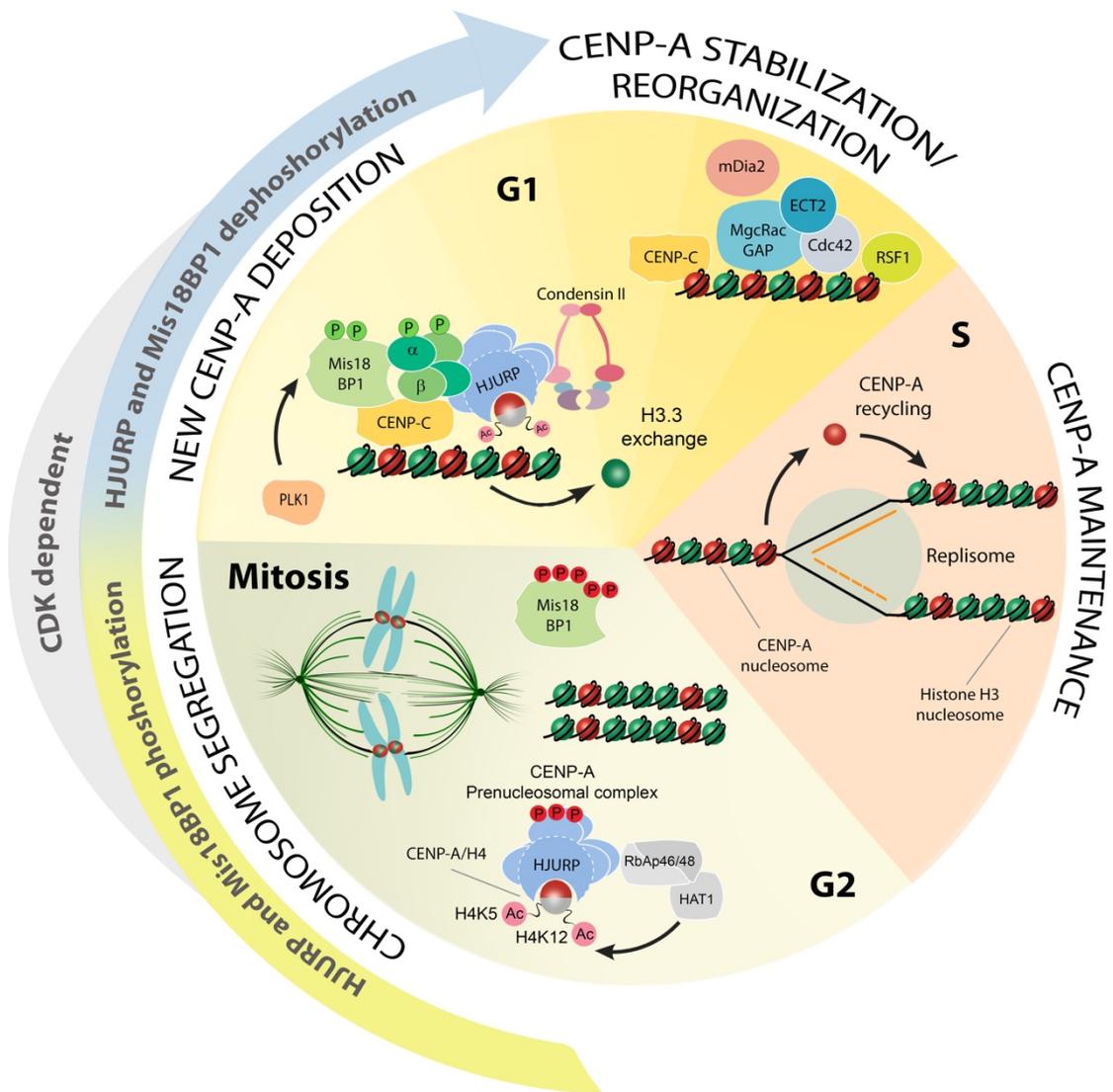
Figure 1.2. The model of cell cycle regulated CENP-A deposition in humans

Figure 1.2. The model of cell cycle regulated CENP-A deposition in humans

New CENP-A deposition occurs exclusively during early G1 and protein complexes involved are depicted in the model. HJURP binds the newly synthesized CENP-A/H4 CENP-A complex in prenucleosomal form (Dunleavy et al., 2009; Foltz et al., 2009; Shuaib et al., 2010). H4K5AC and H4K12Ac histone marks present in the CENP-A prenucleosomal complex are dependent upon RbAp46/48/Hat1 activity and required for CENP-A deposition (Shang et al., 2016). HJURP/CENP-A/H4 localization relies on the Mis18 complex (Barnhart et al., 2011; Fujita et al., 2007). The CENP-A deposition machinery is controlled by the CDK activity. Cell cycle regulated and CDK1/CDK2-dependent phosphorylation of Mis18BP1 and HJURP prevents from premature CENP-A loading during G2 and mitosis, and dephosphorylation of these proteins occurs prior new CENP-A deposition in G1 (Muller et al., 2014; Silva et al., 2012). G1 coupled and Plk1 mediated phosphorylation of the Mis18 complex promotes its centromeric localization and CENP-A deposition (McKinley and Cheeseman, 2014). Mis18BP1 is recruited to centromeres upon its direct interaction with CENP-C (Dambacher et al., 2012). Human Mis18α and Mis18β form a four subunit complex which is incorporated to the centromere through interaction of Mis18α with Mis18BP1 and Mis18β with CENP-C (Nardi et al., 2016; Stellfox et al., 2016). HJURP mediates deposition of CENP-A nucleosomes, and histone H3.3 placeholder is removed from the centromeric chromatin (Dunleavy et al., 2011). Following new CENP-A deposition centromeric nucleosomes are stabilized and protein factors involved in this process are depicted in the model. During DNA replication existing CENP-A nucleosomes are retained across the replication fork (Bodor et al., 2014; Jansen et al., 2007).

In contrast to HJURP, which is recruited to centromeres with a refined temporal window when new CENP-A nucleosomes assembly occurs, the fission yeast Scm3 protein remains associated with centromere through most of the cell cycle (Pidoux et al., 2009). This localization may provide a mechanism to insure the reassembly of CENP-A^{Cnp1} in the event of centromeric chromatin disruption or to block the ubiquitination and degradation of centromeric CENP-A^{Cnp1} alternatively Scm3 may provide additional function at the centromere beyond CENP-A^{Cnp1} deposition.

The crystal structures of both yeast Scm3/CENP-A^{Cse4}/H4 and human HJURP-Scm3/CENP-A/H4 complexes demonstrate that the association of CENP-A (Cse4) with its chaperone prevents CENP-A/H4 tetramer formation and precludes spontaneous DNA interactions by the histone complex in the prenucleosomal form (Cho and Harrison, 2011; Hu et al., 2011). The vertebrate HJURP is much larger than its yeast orthologue Scm3, and contains several domains that are absent from both the S. *pombe* and *Cerevisiae* orthologues (Fig. 1.4) (Sanchez-Pulido et al., 2009). Similarly to *S. cerevisiae* Scm3, humans HJURP was demonstrated to mediate an interaction with DNA through its "mid" domain (HMD), which is required for new CENP-A deposition (Fig. 1.3) (Muller et al., 2014; Xiao et al., 2011). It is not known whether in addition to its ability to bind DNA, HJURP also has a capacity to interact with RNA. Given the evidence that RNA plays a role in centromere specification and HJURP recruitment it is an outstanding question that awaits future studies (Bergmann et al., 2011; Quenet and Dalal, 2014a).
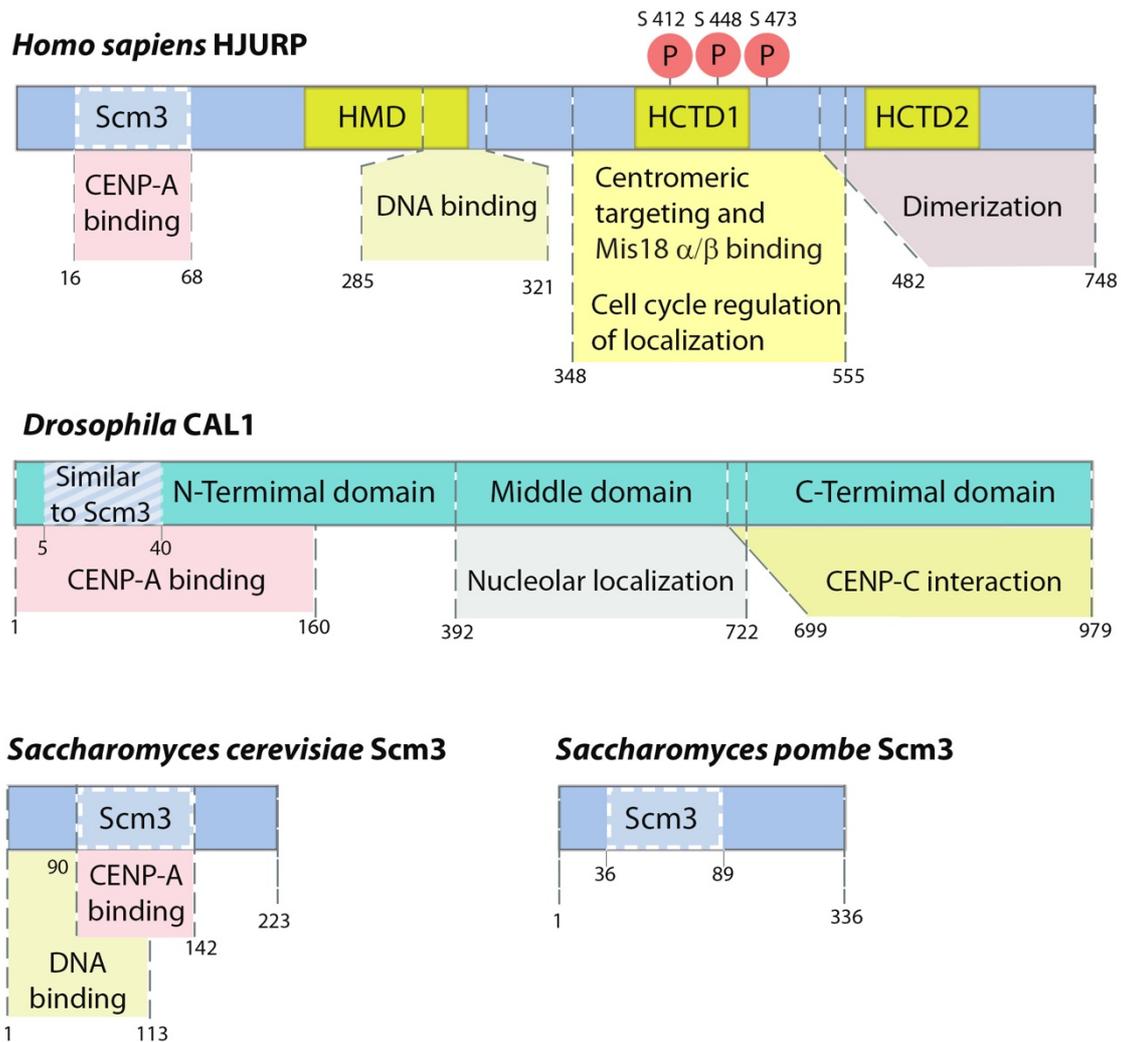
Figure 1.3. Molecular organization of CENP-A specific chaperone in different species
Domains identified within CENP-A chaperones among different species and their roles are depicted (Barnhart et al., 2011; Bassett et al., 2012; Cho and Harrison, 2011; Dechassa et al., 2011; Hu et al., 2011; Muller et al., 2014; Sanchez-Pulido et al., 2009; Schittenhelm et al., 2010; Shuaib et al., 2010; Wang et al., 2014; Zasadzinska et al., 2013). The Scm3 domain is conserved among eukaryotes except for the *Drosophila melanogaster* where the similarity was assessed based on both sequence and secondary structure similarity (Phansalkar et al., 2012; Sanchez-Pulido et al., 2009). HMD- HJURP mid domain; HCTD1-HJURP carboxy terminal domain 1; HCTD2-Hjurp carboxy terminal domain 2.
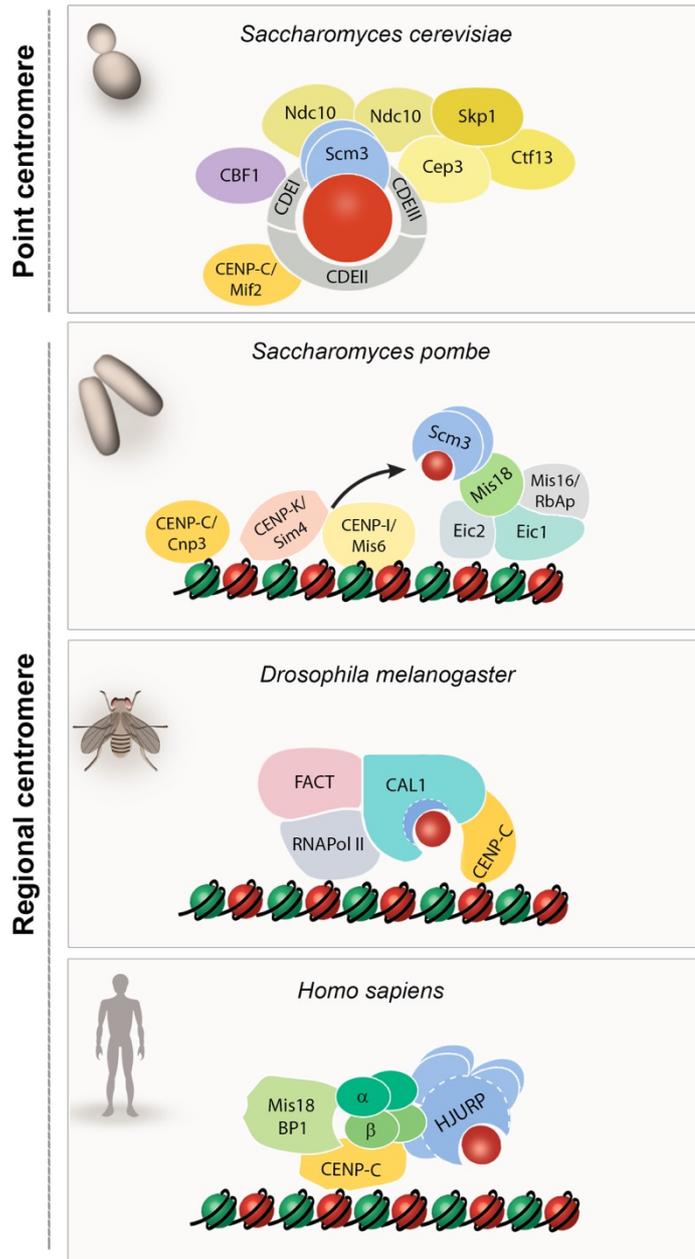
Figure 1.4. Protein complexes involved in CENP-A deposition pathway in eukaryotes

Figure 1.4. Protein complexes involved in CENP-A deposition pathway in eukaryotes

The comparison of CENP-A deposition machinery across species. All conserved proteins involved in CENP-A deposition pathway are colored similarly. Budding yeast point centromeres are specified by unique DNA elements: CDEI, CDEII and CDEIII (Clarke and Carbon, 1980; Fitzgerald-Hayes et al., 1982), which are required for recruitment of DNA binding proteins as depicted in the model. The regional centromeres in fission yeast and higher eukaryotes are specified by the presence of CENP-A containing nucleosomes. CENP-A incorporation into centromeric chromatin is mediated by its distinct histone chaperone - HJURP in vertebrates, Scm3 in yeast and CAL1 in *Drosophila melanogaster* (Barnhart et al., 2011; Bernad et al., 2011; Camahort et al., 2007; Dechassa et al., 2011; Dunleavy et al., 2009; Foltz et al., 2009; Mizuguchi et al., 2007; Pidoux et al., 2009; Shuaib et al., 2010; Stoler et al., 2007; Williams et al., 2009). HJURP and Scm3 share common ancestry, as depicted on the model, and the CAL1 share similarity to Scm3 based on the sequence and secondary structure similarity (Phansalkar et al., 2012; Sanchez-Pulido et al., 2009). CENP-C is conserved in all eukaryotes but it's essential role in centromere specification is restricted to higher eukaryotes where it is required for recruitment of the Mis18 complex (Dambacher et al., 2012; Moree et al., 2011). The role of Mis18 complex in CENP-A deposition pathway is conserved from fission yeast to humans, however, no Mis18 homologue was identified in *Drosophila* (Fujita et al., 2007; Maddox et al., 2007). The fission yeast have only one copy of Mis18 protein and the function of MIS18BP1 was replaced by the Eic1 protein (Hayashi et al., 2014; Subramanian et al., 2014). Human Mis18 complex is a multisubunit complex composing of Mis18α/β heterotetramer and Mis18BP1 (Fujita et al., 2007; Maddox et al., 2007; Nardi et al., 2016). The CENP-A deposition in Drosophila requires active transcription mediated by the FACT and RNA Polymerase II (Chen et al., 2015).

Centromeric recruitment of HJURP is independent of CENP-A binding and is mediated by the HJURP carboxyl terminal domain 1 (HCTD1) (Fig. 1.3) (Wang et al., 2014; Zasadzinska et al., 2013) (Chapter 2). The HJURP carboxyl terminal domain 2 (HCTD2) serves as a homo dimerization interface and facilitates HJURP self-association, consistent with formation of the budding yeast Scm3/CENP-A^{Cse4}/H4 hexamer and Scm3 self-association in fission yeast. In both species, the multimerization mediated by this domain is required for new CENP-A deposition (Mizuguchi et al., 2004; Pidoux et al., 2009; Wang et al., 2014; Zasadzinska et al., 2013) (Chapter 2). This evidence provides a mechanism by which prenucleosomal HJURP complex brings two CENP-A molecules to the site of CENP-A deposition consistent with the CENP-A nucleosomes forming an octamer. Alternatively, one HJURP present in the prenucleosomal complex brings newly synthetized CENP-A/H4 heterodimer, and the other HJURP molecule can recognize CENP-A present within centromeric chromatin, consistent with the hemisome hypothesis (Wang et al., 2014; Zasadzinska et al., 2013) (Chapter 2).

The proposed role of the histone chaperone has been to preclude the stochastic interactions between the histone protein and DNA prior to nucleosomes assembly. Consistent with this idea, the interaction of HJURP with the CENP-A/H4 heterotetramer blocks several key residues along the DNA interface of CENP-A (Cse4) (Cho and Harrison, 2011; Hu et al., 2011). In addition, histone chaperones are known to facilitate the assembly of histone subunits into nucleosomes. Both Scm3 and HJURP mediate CENP-A (Cse4) nucleosome assembly *in vitro* (Barnhart et al., 2011; Camahort et al., 2009; Dechassa et al., 2011; Shivaraju et al., 2011). Much consideration has been given to the whether the CENP-A nucleosome adopts non-canonical forms (Black and Cleveland, 2011; Quenet and Dalal, 2012), however, deposition experiments suggest that, while CENP-A may take on varied

conformations, the CENP-A chaperone facilitates the formation of an octameric nucleosomes with a left-handed wrap of the DNA.

Pre-nucleosomal Posttranslational modifications and CENP-A deposition

CENP-A is bound to its chaperone as a heterodimer with histone H4, thus modification of H4 therefore may contribute to CENP-A nucleosome assembly. Indeed, histone H4 is acetylated on K5ac and K12ac within the prenucleosomal complex, and these modifications are necessary for CENP-A deposition (Fig 1.2) (Shang et al., 2016).

Human RbAp46 (a.k.a. RBBP7) and RpAp48 (a.k.a. RBBP4) are highly homologous genes whose protein products are present in many chromatin remodeling complexes (Loyola and Almouzni, 2004). Mutants of the *S. pombe* homolog of the RbAp proteins, Mis16, cause chromosome segregation defects due to a failure to assemble CENP-A^{Cnp1} nucleosomes (Hayashi et al., 2004). RbAp46/48 co-purified with HJURP in the prenucleosomal CENP-A complex (Dunleavy et al., 2009; Shuaib et al., 2010). A crystal structure of the Mis16-Scm3-CENP-A^{Cnp1}/H4 complex shows that Mis16 contacts both the Scm3 chaperone and histone H4 (An et al., 2015). Depletion of RbAp proteins reduces HJURP recruitment and new CENP-A deposition (Dunleavy et al., 2009; Shang et al., 2016). K5 and K12 acetylation of the histone H4 bound to CENP-A within the prenucleosomal complex are dependent on RbAp48, and these modifications are required for CENP-A deposition *in vivo* (Fig.1.2) (Shang et al., 2016). In the *Xenopus* system H4K5 and H4K12 acetylation marks in prenucleosomal CENP-A complex are dependent upon HAT1 activity (Shang et al., 2016) and Hat1 activity is required for CENP-A deposition in *Drosophila* (Boltengagen et al., 2016). Therefore, a major role of RbAP48 may be the recruitment of the histone acetyltransferase required for modifying Histone H4. What components may read out the presence of H4 acetylation within the assembly pathway is not known.

RbAp46 and RbAp48 depletion results in reduced HJURP protein levels (Dunleavy 2009) and a second role for these proteins may be in regulating the stability of the CENP-A prenucleosomal complex (Mouysset et al., 2015). RbAP46 forms a complex with the CRL4 ubiquitin ligase, a member of the cullin-RING-ligase family, and DDB1 protein (where DDB1 mediates the association of CUL4 with its substrate specific receptor-RbAP46)(Lee and Zhou, 2007; Mouysset et al., 2015). RbAp46 is required for stabilizing CENP-A protein levels and the CRL4-RbAp46 complex activity promotes efficient new CENP-A deposition in humans (Mouysset et al., 2015). This is in contrast to studies in yeast and *Drosophila*, where the association of CENP-A with the SCF E3-ubiquitin ligase complex leads to CENP-A degradation (see below).

Two different posttranslational modifications of CENP-A are proposed be important for CENP-A deposition. These are phosphorylation of serine 68 and ubiquitylation of lysine 124 (Niikura et al., 2015; Yu et al., 2015). Both modifications are located outside of the CATD domain that is sufficient for HJURP binding, and situated on the helix α1 and helix α3 of CENP-A, respectively. However, both are proposed to influence HJURP binding to CENP-A. CENP-A lysine 124 (K124) in humans undergoes mono- and di-ubiquitylation mediated by the CUL4A-RBX1-COPS8 E3 ligase complex (Niikura et al., 2015). Downregulation of any of the CUL4A-RBX1-COPS8 subunits or mutation of Lys124 leads to loss of centromeric CENP-A in mitosis and interphase cells. Mutation of CENP-A lysine 124 weakens the interaction with CENP-A chaperone HJURP.

Phosphorylation at CENP-A-Ser68 is proposed to preclude its interaction with HJURP, negatively regulating new CENP-A deposition. CENP-A Ser68 phosphorylation depends on Cdk1/cyclin B activity during early mitosis and PP1a phosphatase dephosphorylates Ser68 in late mitosis, making CENP-A competent for HJURP binding and new incorporation in the following G1 (Yu 2015). While the phosphomimetic S68Q mutation appears to preclude

HJURP binding both *in vivo and in vitro* (Hu et al., 2011; Yu et al., 2015), Bassett et al reported that S68Q substitution within CENP-A has no effect on HJURP mediated targeting and subsequent incorporation into chromatin at non-centromeric sites. Moreover, that recombinant CENP-A containing the S68Q mutation forms a complex *in vitro* with HJURP with similar efficiency when compared to the wild-type form (Bassett et al., 2012). Despite the effects observed *in vivo* for the S68 and K124 mutations, both mutations are fully able to rescue CENP-A null cells, suggesting that these modifications are not essential for the process of centromere specification and inheritance (Fachinetti, 2016).

In budding yeast, Psh1 prevents ectopic localization of CENP-A^{Cse4} (Hewawasam et al., 2010; Ranjitkar et al., 2010). Psh1 is an E3 ubiquitin ligase that was identified as associated with yeast CENP-A^{Cse4} in immunoprecipitation experiments and characterized as a kinetochore and centromere associated protein. Psh1 regulates CENP-A^{Cse4} levels by ubiquitylating CENP-A^{Cse4} and targeting it for proteolysis; thus, preventing its accumulation outside the centromeric chromatin. Psh1 and Scm3 both recognize the CENP-A^{Cse4} -CATD domain; therefore, Scm3 appears to protect CENP-A^{Cse4} from the Psh1 mediated ubiquitination and subsequent degradation (Hewawasam et al., 2010; Ranjitkar et al., 2010). Drosophila CENP-A^{CID} directly interacts with the with the F-Box Protein Partner of Paired (Ppa), a variable component of a SCF E3-ubiquitin ligase complex in *Drosophila*. Ppa binds CENP-A^{CID} through the CATD domain and regulates its stability (Cardozo and Pagano, 2004; Moreno-Moreno et al., 2011; Nakayama and Nakayama, 2006; Schuh et al., 2007).

Coupling chaperone recruitment to existing centromeres

Human centromeres range from 0.3 to 5Mbp in size and account for less than 1% of the chromosome (Cleveland et al., 2003). The restriction of centromeres to a single locus ensures the stable inheritance of centromeres by avoiding situations where multiple centromeres on one chromosome could make attachments to opposing poles and result in chromosome breakage during mitosis.

The recruitment of the CENP-A specific histone chaperone to the existing centromere is an essential step in epigenetic inheritance. Mis18 is a key adapter protein that mediates the recruitment of the CENP-A chaperone to centromeres in several organisms (Fig 1.4, 1.5), but is absent from organisms with point centromeres. Mis18 was originally identified in a genetic screen in fission yeast to identify genes required for proper chromosome segregation (Hayashi et al., 2004). spMis18 mutants eliminate CENP-A[Cnp1] incorporation to centromeres and Mis18 directly interacts with Scm3 to determine its recruitment (Pidoux et al., 2009). In humans, Mis18 exists as a complex comprised of Mis18α, Mis18β and Mis18BP1 proteins (Fig. 1.4, 1.5). The Mis18 complex is essential for the recruitment of HJURP and CENP-A to the centromeric chromatin due to a direct interaction with the HJURP centromere targeting domain within the HCTD1 (Fig. 1.3) (Barnhart et al., 2011; Fujita et al., 2007; Nardi et al., 2016; Wang et al., 2014). Mis18 proteins do not require HJURP for recruitment, demonstrating that they are upstream components of the pathway (Barnhart et al., 2011; Bernad et al., 2011). Consistent with studies in yeast, depletion of the Mis18 complex subunits in human cells results in a high rate of chromosome segregation defects and loss of centromeric CENP-A (Fujita et al., 2007). The role of the Mis18 proteins in the CENP-A deposition pathway is evolutionarily conserved, as depletion of Mis18BP1 (KNL-2) homologues in *C. elegans* and *Xenopus* also leads to defects in CENP-A deposition in these

species, although as discussed below the pathway has undergone several permutations in different organisms (Maddox et al., 2007; Moree et al., 2011).

Since the Mis18 proteins are required for HJURP recruitment, the key question becomes how the Mis18 protein recognizes the existing centromere. CENP-A nucleosomes recruit the CCAN (constitutive centromere-associated network), a multiprotein complex comprised of 16 subunits, present at the centromere throughout the cell cycle, that serves as a structural core for kinetochore assembly during mitosis (Amano et al., 2009; Cheeseman and Desai, 2008; Earnshaw et al., 1986; Foltz et al., 2006; Izuta et al., 2006; McKinley and Cheeseman, 2016; Nishihashi et al., 2002; Okada et al., 2006; Saitoh et al., 1992; Sugata et al., 1999). The CENP-C component of the CCAN directly recognizes the CENP-A nucleosome (Carroll et al., 2010; Guse et al., 2011; Kato et al., 2013). New CENP-A nucleosomes within the alpha satellite DNA are assembled directly adjacent to the existing CENP-A (Ross et al., 2016). CENP-C plays a crucial role in recruiting the proteins required for CENP-A deposition (Fig. 1.2, 1.3), and thus links the existing centromere to the assembly of new CENP-A nucleosomes in early G1. CENP-C interacts directly with two proteins within the Mis18 complex, Mis18BP1 and Mis18β (Dambacher et al., 2012; Moree et al., 2011; Stellfox et al., 2016). CENP-C depletion causes defects in Mis18BP1 and HJURP recruitment and leads to loss of CENP-A chromatin assembly (Dambacher et al., 2012; Moree et al., 2011; Stellfox et al., 2016).

Species	Mis18	Mis18BP1/ counterpart	CENP-A chaperone	CENP-C	CENP-A deposition timing
S. cerevisiae	Not Identified	Not Identified	Scm3	Mif2	S phase
S. pombe	Mis18	Eic1	Scm3	Cnp3	S phase/ G2
Ustilago	Mis18	Not Identified	Not Identified	MIF2 related protein	Unknown
C. elegans	Not Identified	KNL-2	Not Identified	HCP-4	Mitosis
Drosophila	Not Identified	Not Identified	CAL1	CENP-C	Metaphase/ Anaphase
Zebrafish	α β	Mis18 BP1	HJURP	CENP-C	Unknown
Xenopus	α β	Mis18 BP1	HJURP	CENP-C	Early Interphase
Mouse	α β	Mis18 BP1	HJURP	CENP-C	Unknown
Human	α β	Mis18 BP1	HJURP	CENP-C	Late telophase/ early G1

Figure 1.5. Conservation of CENP-A deposition factors across species

Table detailing the conserved proteins involved in CENP-A deposition pathway as well as timing of CENP-A deposition in different model organisms (Bernad et al., 2011; Dunleavy et al., 2007; Jansen et al., 2007; Maddox et al., 2007; Mellone et al., 2011; Moree et al., 2011; Pearson et al., 2004; Schuh et al., 2007; Takayama et al., 2008)

The pivotal role that CENP-C plays in determining the site of centromeric chromatin assembly is exemplified by experiments in chicken DT40 cells, where the endogenous centromere is conditionally removed and the functional kinetochore assembled at an ectopic LacO locus These experiments show that tethering the LacI-fused HJURP or full length CENP-C are sufficient to recruit CENP-A in order to establish a functional epigenetic *de novo* centromere (Hori et al., 2013). Although tethering the CENP-C N terminus (1–643 aa) in this system is sufficient to recruit microtubule binding proteins and the CPC complex, it fails to incorporate CENP-A nucleosomes (Hori et al., 2013). This is consistent with the identification of the N-terminus of CENP-C as the region of interaction with Mis18BP1 and Mis18β (Dambacher et al., 2012; Moree et al., 2011; Stellfox et al., 2016). In contrast, CENP-C homologues in yeast (Mif2 and Cnp3) are not essential to facilitate CENP-A deposition (Fig. 1.4) (Meluh and Koshland, 1995, 1997; Westermann et al., 2003).

Additional factors in the CCAN also contribute to directing new CENP-A nucleosome deposition. Depletion of the CENP-HIKM complex in chicken cells compromise the incorporation of newly synthesized CENP-A (Okada et al., 2006). Consistent with this observation, fission yeast CENP-I^{Mis6} and CENP-K^{Sim4} are required for CENP-A nucleosome deposition (Fig 1.4) (Pidoux et al., 2009; Takahashi et al., 2000). Similar to CENP-C, tethering CENP-I to a non-centromeric site in chicken DT40 cells drives new CENP-A deposition and forms an epigenetic centromere (Hori et al., 2013). This suggests that the CCAN components play a dual role, and are required for both centromere specification in G1 and recruitment of kinetochore components during mitosis.

Budding yeast centromeres are determined by DNA sequence. And although they share a homologous CENP-A chaperone, Scm3, the mechanism by which Scm3 is recruited to centromeres is distinct from epigenetic centromeres (Fig. 1.4). The centromere determining elements (CDE) in budding yeast are essential for recruitment of a DNA binding protein Cbf1

specifically recognizing CDEI and a multisubunit protein complex: CBF3 (centromere binding factor 3), containing Ndc10, Cep3, Ctf13 and Skp1, associated with CDEIII DNA element (Cho and Harrison, 2011; Doheny et al., 1993; Goh and Kilmartin, 1993; Hyman et al., 1992; Lechner and Carbon, 1991; Mizuguchi et al., 2007; Russell et al., 1999; Strunnikov et al., 1995). The CBF3 subunit-Ndc10 is required for the recruitment of the Scm3 chaperone and subsequent deposition of the CENP-A^{Cse4} containing nucleosome (Camahort et al., 2007; Mizuguchi et al., 2007) .

The chromatin landscape influence on CENP-A deposition

CENP-A nucleosomes are interspersed with the canonical H3 nucleosomes within the centromeres of flies and humans (Blower et al., 2002). Centromeres were initially thought to be transcriptionally silent loci, a characteristic that is consistent with the posttranslational modifications found in the surrounding pericentric heterochromatin (Peters et al., 2001; Ribeiro et al., 2010; Rice et al., 2003). However, studies in human and *Drosophila* derived chromatin fibers demonstrated that H3K9me2 and H3K9me3 marks are absent from the CENP-A (CID) domain (Lam et al., 2006; Sullivan and Karpen, 2004). Furthermore, histone H3 nucleosomes found interspersed with CENP-A nucleosomes in humans are decorated with histone marks associated with active or poised chromatin, such as H3K4me1/2, H3K36me2/3 (Bergmann et al., 2011; Sullivan and Karpen, 2004). The histone H3K4 trimethylation, associated with actively transcribed regions, is absent from the centromeric core domain in humans and *Drosophila*, but is present at chicken centromeric DNA (Ribeiro et al., 2010; Sullivan and Karpen, 2004). Until recently the centrochromatin localized histones in higher eukaryotes were thought to be hypo-acetylated and lack acetylated marks found generally in euchromatin such as H3K9Ac, H4K5Ac, H4K8Ac, H4K12Ac or H4K16Ac. However, a recent study documented the presence of H4K5Ac and H4K12Ac

within CENP-A containing nucleosomes in chicken and humans. (Bailey et al., 2013; Shang et al., 2016; Sullivan and Karpen, 2004)

Transcripts from centromeric repeat sequences have been observed in multiple model organisms (Bergmann et al., 2011; Bouzinba-Segard et al., 2006; Carone et al., 2009; Carone et al., 2013; Chan et al., 2012; Choi et al., 2011; Eymery et al., 2009; Hall et al., 2012; Lam et al., 2006; May et al., 2005; Ohkuni and Kitagawa, 2011; Quenet and Dalal, 2014b; Stimpson and Sullivan, 2010; Topp et al., 2004; Wong et al., 2007). Active RNA Polymerase II is recruited to endogenous human centromeres during mitosis and early G1 (Chan et al., 2012; Quenet and Dalal, 2014b). Inhibition of RNA Polymerase II mediated transcription in HeLa cells leads to decreased ✔-satellite transcript levels in mitosis, loss of CENP-C recruitment to endogenous centromeres, and chromosome segregation defects (Chan et al., 2012). The mechanistic role of centromeric transcripts and the act of transcription in centromere function is not yet clear, although histone H3 eviction may be a key aspect.

Utilizing a synthetic human artificial chromosome (HAC), Bergmann et al. demonstrated that the presence of H3K4me2 and transcription events at the centromere play a critical role in CENP-A assembly and centromere function by altering the recruitment of CENP-A deposition machinery (Bergmann et al., 2011). Tethering a lysine-specific demethylase 1 (LSD1) to the HAC centromeric domain leads to removal of H3K4 methylation and results in loss of transcription of α-satellite DNA at this loci. This correlates with loss of HJURP localization, impaired CENP-A deposition, and ultimately leads to loss of kinetochore function (Bergmann et al., 2011).

Biochemical purification of RNA associated with prenucleosomal CENP-A/HJURP complex identified a 1.3 kb RNA product that co-localizes with a-satellite DNA and CENP-A, and hybridizes to centromeric α-satellite probes, suggesting it originated from α-satellite transcripts (Quenet and Dalal, 2014a). Targeting of α-satellite transcripts as well as other

centromere derived RNAs by siRNA *in vivo* results in reduced CENP-A and HJURP recruitment to the centromere, suggesting that the RNA component partially encoded within α-satellite DNA play a role in CENP-A deposition pathway (Quenet and Dalal, 2014a). Exactly how RNAs are associated with the CENP-A prenucleosomal complex is still unknown, as well as how this association would contribute mechanistically to CENP-A deposition.

A strong link between CENP-A deposition and transcription was demonstrated in *Drosophila*. Chen at al, using an inducible ectopic centromere approach, demonstrated that new CENP-A^{CID} deposition at the ectopic centromere requires transcription (Fig. 1.4)(Chen et al., 2015). The mass spec analysis of binding partners of the drosophila CENP-A^{CID} chaperone CAL1 *in vivo* identified two subunits of the FACT complex: Spt16 and SSRP1, both of which physically interact with CAL1. FACT was also previously found associated with centromere in human cells (Foltz et al., 2009; Obuse et al., 2004). FACT is involved in transcription elongation from chromatin templates *in vitro* and promoting deposition of histone H3.3 nucleosomes *in vivo* in *Drosophila* system (Orphanides et al., 1998). Spt16 and SSRP1 subunits colocalize with CENP-A^{CID} in *Drosophila* cells and downregulation of FACT leads to defects in CENP-A^{CID} recruitment at endogenous centromeres. CAL1 along with FACT facilitate RNA Polymerase II mediated transcription at the site of CENP-A^{CID} deposition which is required for CENP-A^{CID} incorporation to occur. In support to these findings other groups reported localization of the active form of RNA Polymerase II at endogenous centromeres in *Drosophila* system during mitosis, which is coincident with new CENP-A^{CID} deposition timing (Rosic et al., 2014).

In addition to the role of the Mis18 complex in the recognition of the CCAN and direct recruitment of HJURP, the Mis18 complex influences posttranslational modifications within the centromeric chromatin (Kim et al., 2012). Deletion of Mis18 in *S. pombe* leads to

increased levels of histone H3 and H4 acetylation at centromeres (Hayashi et al., 2004). In vertebrates, the Mis18 complex influences histone modifications and DNA methylation. Knockout of Mis18α in mice leads to reduced H3K9 and H3K4 methylation and increased acetylation within centromeric repeats (Kim et al., 2012). The *de novo* methyltransferase enzymes DNMT3a/b are also recruited to centromere by Mis18α/β (Kim et al., 2012). Downregulation of DNMT3b or Mis18α leads to increased transcription of centromeric repeats (Gopalakrishnan et al., 2009). However, the importance of DNMT3a/b in centromere function is unclear since cells lacking DNTM3a/b are viable (Reviewed in (Brown and Robertson, 2007)).

More recently Mis18BP1 was shown to recruit the KAT7 lysine methyltransferase complex to centromeres (Ohzeki et al., 2016). Disruption of the KAT7 complex leads to reduced CENP-A deposition. KAT7 in conjunction with RSF1 may regulate histone turnover to facilitate new CENP-A deposition in G1. In future work it will be important to determine exactly how the Mis18 complex may integrate multiple downstream chromatin modifying pathways to promote centromere deposition.

Licensing of centromere assembly

CENP-A incorporation into the centromeric chromatin is cell cycle regulated, although the timing of CENP-A deposition differs across species (Fig 1.2, 1.5) (Allshire and Karpen, 2008; Boyarchuk et al., 2011). Budding yeast CENP-A[Cse4] incorporation is coincident with DNA replication (Pearson et al., 2004; Wisniewski et al., 2014). Similarly, in fission yeast, CENP-A deposition occurs during early S phase, but also during G2 phase (Takayama et al., 2008). In vertebrates, new CENP-A incorporation is uncoupled from DNA replication and restricted to late telophase/early G1 phase (Bernad et al., 2011; Jansen et al., 2007; Silva et al., 2012).

The process of human CENP-A deposition occurs via a licensing mechanism that restricts deposition to the G1 phase and controls the assembly of CENP-A to ensure that only a limited amount of new CENP-A is assembled in each cell cycle. The timing of CENP-A deposition is restricted to the early G1 phase by inhibition of CENP-A deposition *through* CDK activity, which is high during S and G2-phase, drops rapidly following satisfaction of the mitotic checkpoint (Fig 1.2) (Silva et al., 2012). Although CENP-A transcript and protein levels accumulate from mid S phase into G2, CDK1/CDK2-dependent phosphorylation of Mis18BP1 prevents from premature CENP-A loading during this time (Silva et al., 2012). Mis18BP1 dephosphorylation occurs during early G1, coincident with new CENP-A deposition. The Plk1 kinase positively regulates CENP-A deposition. Plk1 phosphorylates the Mis18 complex during G1 to promote its recruitment to centromeres (Fig 1.2)(McKinley and Cheeseman, 2014). Inhibition of the Plk1 kinase activity abrogates new CENP-A deposition. The opposing functions of PLk1 and Cdk1 phosphorylation provide tight temporal control of CENP-A deposition by limiting Mis18 recruitment.

The assembly of CENP-A nucleosomes in G1 is limited by at least two mechanisms. The Mis18 complex forms a conserved tetramer (Nardi et al., 2016; Subramanian et al., 2016). In humans, this includes two copies of each Mis18 paralog (Mis18α/Mis18β). Mis18 binds the centromere stably in late telophase. Binding of HJURP to Mis18 disrupts the complex and eliminates the ability of Mis18 to continue to interact, essentially removing the signal for HJURP recruitment, and blocking further CENP-A deposition at that site. In addition, the Mis18β subunit undergoes ubiquitylation and degradation by the $SCF^{\beta TrCP}$ E3 ubiquitin ligase, thus degrading the signal for HJURP recruitment to centromeres (Kim et al., 2014).

Evolutionary diversity in CENP-A deposition pathways

Despite the high degree of conservation between the CENP-A binding domains within the HJURP and Scm3 chaperones, that spans billions of years of evolution, there is a great variety in the CENP-A deposition pathways across organisms (Fig 1.4, 1.5). This likely reflects the unique strategies for centromeric chromatin assembly that these organisms employ.

Drosophila species lack a clear HJURP homolog, but an siRNA screen for genes involved in CENP-A^{CID} centromere deposition in *Drosophila* S2 cells identified CAL1 (chromosome alignment defect 1) as a key factor (Erhardt et al., 2008). *Drosophila* CAL1 is a fly specific protein that functions as a CENP-A^{CID} chaperone (Chen et al., 2014; Erhardt et al., 2008; Goshima et al., 2007). In spite the small similarity to yeast Scm3 domain of Kluyveromyces lactis, determined based on the sequence and secondary structure similarity, CALI does not share common ancestry with yeast Scm3 and human HJURP (Phansalkar et al., 2012; Sanchez-Pulido et al., 2009). CAL1 directly binds to CENP-A^{CID} /H4 dimer and was shown to function as the CENP-A^{CID} specific assembly factor in fruit flies (Chen et al., 2014). Its depletion in *Drosophila* results in loss of centromeric CENP-A^{CID} localization and is associated with chromosome segregation defects (Chen et al., 2014; Erhardt et al., 2008; Goshima et al., 2007). Both HJURP and CAL1 are sufficient to promote *de novo* centromere establishment. Tethering HJURP to the chromosome arm or to a naïve alpha satellite array is sufficient to facilitate CENP-A deposition outside of the centromeric chromatin and results in formation of functional kinetochore at an ectopic site in human cells (Barnhart et al., 2011; Ohzeki et al., 2012). Similarly, targeting CAL1 to an ectopic site was demonstrated to mediate *de novo* CENP-A^{CID} deposition in *Drosophila,* which leads to formation of a de novo centromere outside of the endogenous centromeric loci (Chen et al., 2014). This *de novo* centromere is epigenetically maintained and serves as platform for recruitment of a functional

kinetochore (Chen et al., 2014). There are several organisms that contain CENP-A nucleosomes for which a functional chaperone has not been identified, including the well-studied nematode *C. elegans* (Fig. 1.5). *C. elegans* have holocentric chromosomes in which the centromere position may be variable and obfuscate the need for specific targeting of the CENP-A histone variant.

Conservation of the Mis18 complex is also highly variable across species. Species as divergent as *S. pombe* and humans possess Mis18, but in higher eukaryotes the Mis18 gene underwent duplication (Fig 1.5). The Mis18 paralogs, termed Mis18α and Mis18β, share about 30% sequence identity, but have diverged in their function in higher eukaryote centromeres (Fujita et al., 2007; Hayashi et al., 2004; Stellfox et al., 2016). The Mis18 complex has not been identified in *Drosophila* or *S. cerevisiae* (Fig 1.5). In both cases, these organisms have devised alternative strategies to couple the CENP-A chaperones to the existing centromere. CALI binds CENP-C in *Drosophila* and the Ndc10 complex, which directly recognizes DNA, recruits the Scm3 chaperone in budding yeast (Camahort et al., 2007; Doheny et al., 1993; Erhardt et al., 2008; Goh and Kilmartin, 1993; Jiang et al., 1993; Lechner and Carbon, 1991; Mellone et al., 2011; Sorger et al., 1995).

While *S. pombe* possess a Mis18 homolog, it lacks the vertebrate Mis18BP1 orthologue (Fig 1.5). The Mis18BP1 function in *S. pombe* is replaced by the Eic1 protein (a.k.a Mis19) (Fig 1.4). The Eic1 and Eic2 proteins co-purified with the spMis18 and exhibit a similar temporal pattern of centromeric localization throughout the cell cycle (Hayashi et al., 2014; Subramanian et al., 2014). Eic1 was demonstrated to be essential for the recruitment of the Mis18, Mis16 and Scm3 proteins to the centromere and for CENP-A[Cnp1] incorporation. However, Eic2 is dispensable for recruitment of CENP-A[Cnp1] to the centromere. This suggests Eic1 is functionally analogous to the Mis18BP1 subunit in recruitment of CENP-A

deposition, although Eic1 is evolutionary distinct and does not share any apparent sequence homology to Mis18BP1 (Hayashi et al., 2014; Subramanian et al., 2014).

Centromere stabilization and re-organization

The recruitment of CENP-A to centromeres via HJURP and Mis18 is not sufficient for the stability of CENP-A, but requires additional proteins that may potentially reorganize centromeric chromatin to increase stability. These factors include the Rho GTPase MgcRACGAP, the formin protein mDia and the Rsf-1 remodeling complex and appear to be recruited to centromere later than Mis18 and HJURP (Fig. 1.2) (Izuta et al., 2006; Lagana et al., 2010; Liu and Mao, 2016; Obuse et al., 2004; Perpelescu et al., 2009).

MgcRacGap co-purifies with centromeric chromatin and with Mis18BP1 from HeLa cells (Izuta et al., 2006; Lagana et al., 2010; Perpelescu et al., 2009). MgcRacGAP localizes to centromeres in late G1. Although the exact timing between MgcRacGAP recruitment and HJURP recruitment has not been established, it appears that MgcRacGAP is recruited later, after new CENP-A incorporation is accomplished. Depletion of MgcRacGAP or its binding partner, ECT2 (guanine nucleotide exchange factor) results in loss of newly incorporated CENP-A, while existing CENP-A is not affected. This suggests that new and old CENP-A populations during G1 are in some way unique. Furthermore, the Cdc42, a small GTPase identified as a target of MgcRacGAP–ECT2 complex, is also recruited to the centromeres during interphase. The Cdc42 activity requires GTPase cycling mediated by MgcRacGAP-ECT2, proposing that a GTPase switch is implicated in the maturation of the newly deposited CENP-A containing nucleosomes (Lagana et al., 2010). mDia2 is a downstream effector of Rho signaling (Gasman et al., 2003; Lammers et al., 2008). mDia2 depletion leads to defects in new CENP-A deposition. The constitutively active form of mDia2 restores CENP-A levels at the centromere resulting from MgcRacGAP downregulation, consistent with its role

downstream of MgcRacGAP in this process. Interestingly, mDia2 depletion leads to prolonged HJURP association with the centromere, suggesting that the processes of HJURP recruitment and MgcRacGAP stabilization are mechanistically linked (Liu and Mao, 2016).

The RSF (remodeling and spacing factor), comprised of the Rsf-1 and SNF2h subunits, has been characterized as an ATP-dependent nucleosome remodeling and spacing factor that together with the FACT complex is implicated in transcription initiation (LeRoy et al., 1998; Orphanides et al., 1998). The RSF complex co-purified with CENP-A nucleosomes prepared from interphase cell extracts (Izuta et al., 2006; Obuse et al., 2004; Perpelescu et al., 2009). RSF centromere localization peaks during the middle of G1 phase. RSF1 can reconstitute and space CENP-A nucleosomes on a naked DNA template, and is required for stability of CENP-A nucleosomes within the centromeric chromatin (Perpelescu et al., 2009). This argues that an energy-dependent remodeling events are involved in stabilization of newly deposited CENP-A nucleosomes.

Condensation of centromeric chromatin is a potentially important step in efficient CENP-A deposition. The Condensin complexes are involved in ATP-dependent chromosome condensation of during mitosis, and are also implicated in centromere establishment in yeast and humans (Hagstrom et al., 2002; Ono et al., 2004; Samoshkin et al., 2009; Wignall et al., 2003; Yong-Gonzalez et al., 2007). Of the two partially overlapping Condensin complexes that have been characterized (Condensin I and II) the Condensin II complex is selectivity involved in centromeric chromatin assembly (Barnhart-Dailey et al., 2016; Bernad et al., 2011; Hirano, 2005). Downregulation of common components to the Condensin complexes (SMC2 and SMC4) or the Condensin II specific subunits (CapH2 and CapD3) leads to reduced assembly of new CENP-A nucleosomes in humans and *Xenopus* extracts (Barnhart-Dailey et al., 2016; Bernad et al., 2011; Samoshkin et al., 2009). CAPH2 was found at human

centromeres in early G1, coincident with new CENP-A deposition, and its recruitment is HJURP-dependent (Fig.1.2) (Barnhart-Dailey et al., 2016).

In chicken cells FACT subunits: SSRP1 and SPT16 co-purified with CENP-A and localize to the centromeric chromatin. FACT interacts with ATP-dependent chromatin remodeling factor CHD1, and the centromeric recruitment of these proteins throughout the cell cycle is dependent upon the CENP-H-I-K-M complex. The downregulation of FACT or CHD1 factors leads to loss of new CENP-A deposition, demonstrating that chromatin remodeling activity of FACT and CHD1 complex plays a critical role in CENP-A deposition (Okada 2009). It remains elusive whether the FACT and CHD1 complex require active transcription in order to play their role in CENP-A incorporation in chicken system.

The process of DNA replication

DNA replication is one of the fundamental processes which occur within the cell and governs accurate and efficient duplication of the genome. In all eukaryotes, DNA replication is thought to be initiated at the replication origins, specific sites selectively recognized by proteins involved in replication initiation. In contrast to buddying yeast, where replication origin sequences share common DNA elements and are well characterized, in mammalian species replication initiation sites are not fully defined (Fangman and Brewer, 1991; Hsiao and Carbon, 1979; Martin et al., 2011; Petryk et al., 2016; Rao et al., 1994; Schaarschmidt et al., 2004; Stinchcomb et al., 1979; Sugimoto and Fujita, 2017; Valenzuela et al., 2011; Vashee et al., 2003; Wyrick et al., 2001).

DNA-binding origin recognition complex (ORC) is a multisubunit complex that recognizes and localizes to the replication origins and promote the recruitment of subsequent protein factors required for replication initiation. The ORC complex comprises of six subunits ORC1, ORC2, ORC3, ORC4, ORC5, ORC6 (Bell and Stillman, 1992; Diffley and Cocker, 1992). The origins of replication are thought occupied by the ORC complex components throughout the cell cycle and it is the DNA replication licensing; however, also known as the prereplicative complex (pre-RC) assembly, that occurs during G1 phase and triggers DNA replication start. (Aparicio et al., 1997; Bell and Stillman, 1992; Costa et al., 2013; Diffley and Cocker, 1992; Diffley et al., 1995; Foss et al., 1993; Fox et al., 1995; Fujita et al., 1998).

The early step of that mechanism relies on binding of the Cdt6 initiation factor to the ORC assembled onto DNA, resulting in formation of Cdt6-ORC-DNA complexes that are ready to recruit the MCM2-7 replicative helicase (Remus et al., 2009). The MCM2-7 helicase forms a hexameric ring, composed of MCM2, MCM3, MCM4, MCM5, MCM6, MCM7 subunits, that acts a ATP-dependent molecular motor required for the DNA unwinding

(Bochman and Schwacha, 2007; Chong et al., 1995; Costa et al., 2011; Patel and Picha, 2000; Remus et al., 2009). The subsequent step of replication initiation involves the sequential recruitment of two replicative helicases bound by Cdt1, which collaborate with ORC and Cdt6, that results in formation of MCM double hexamer surrounding the double stranded DNA at the replication origin (Bell and Labib, 2016; Bochman and Schwacha, 2007; Chen and Bell, 2011; Chen et al., 2007; Chong et al., 1995; Coleman et al., 1996; Costa et al., 2013; Costa et al., 2011; Madine et al., 1995; Maiorano et al., 2000; Nishitani et al., 2000; Remus et al., 2009; Takara and Bell, 2011; Ticau et al., 2015).The double hexamers are recruited and assembled in a head-to-head orientation, and this conformation facilitates bidirectional replication initiation (Costa et al., 2014; Evrin et al., 2009; Remus et al., 2009; Sun et al., 2014). Once the MCM2 loading is completed, the origins of replication are „licensed" and await to be activated. Importantly, while cells enter S phase, the ability of ORC, CDC6, and Cdt1proteins to facilitate MCM recruitment and pre-RC assembly are inhibited due to increased CDK activity, in order to prevent events of chromatin re-replication. The recruitment of the other replisome components, however, is promoted to facilitate origin firing (Diffley, 2004).

At the S phase entry, the double hexamer is converted into the CMG complex, comprising of MCM2-7 helicase bound to Cdt45, MCM10 and GINS proteins (Sld5, Psf1, Psf2, Psf3) (Costa et al., 2011; Ilves et al., 2010; Kang et al., 2012; Kanke et al., 2012; Moyer et al., 2006; Pacek et al., 2006; Takayama et al., 2003; van Deursen et al., 2012). The CMG complex activation requires the DDK and CDK kinases that act on the MCM2-7 helicase complex and its accessory proteins to allow for Cdc45 interaction with the Treslin protein (Araki, 2010; Labib, 2010; Tanaka et al., 2007; Zegerman and Diffley, 2007). This in turn leads to the recruitment of Sld2/RecQL4 and DPB11/TopBP1 proteins to the complex (Bruck et al., 2015; Marks et al., 2017; Masai et al., 2010; Remus et al., 2009; Zegerman and Diffley,

2007). The replication protein A and DNA polymerases are also recruited to form a replisome ready to facilitate nascent DNA synthesis (Abid Ali and Costa, 2016; Kanemaki and Labib, 2006; Marks et al., 2017; Takayama et al., 2003)

The active CMG helicase complexes unwind dsDNA and travels along each of the leading strands exposing ssDNA to RPA binding and making it available as a template for new DNA synthesis (Aparicio et al., 1997; Kanemaki and Labib, 2006; Kanemaki et al., 2003; Tercero et al., 2000; Yuzhakov et al., 1999). The leading strand is being synthesized in a continuous fashion while the lagging strand synthesis occurs discontinuously via Okazaki fragments that need to be ligated (Burgers and Kunkel, 2017; Howes and Tomkinson, 2012; Okazaki et al., 1968; Tye et al., 1977)

In eukaryotes, there are three polymerases that are shown to be engaged at the DNA replication fork including Pol α, Pol ε, Pol δ (Burgers, 2009; Waga et al., 1994). The leading strand synthesis was proposed to be mediated by Pol ε, while the lagging strand synthesis rely on the activity of the Pol δ synthesizing the discontinuous Okazaki fragments (Lujan et al., 2013; Miyabe et al., 2011; Nick McElhinny et al., 2010). The Pol α primase complex activity is required for the initiation of DNA synthesis on both strands, as it provides short RNA primers complementary to a ssDNA template required for replication start to occur (Johansson and Dixon, 2013; Muzi-Falconi et al., 2003).

The efficient and coordinated synthesis of both leading and lagging strands during DNA replication elongation requires on the activity of PCNA, a processivity factor acting as a sliding clamp, and the RFC complex, acting as a PCNA clamp loader (Ohashi and Tsurimoto, 2017; Prelich and Stillman, 1988; Tsurimoto and Stillman, 1989). The topoisomerase I or II proteins are also recruited and implicated in active DNA replication as their role is to remove supercoils generated by DNA unwinding (Yang et al., 1987). Due to discontinuous nature of lagging strand synthesis, the elongation stage is more challenging and

requires additional factors. DNA polymerase Pol δ carries the extension of the lagging strand until it collides with the proceeding Okazaki fragment. Then the Pol δ collaborates with the flap endonuclease 1 (FEN1) and this process leads to removal of the initiating RNA primer (also referred to as nick translation) (Balakrishnan and Bambara, 2013; Grasby et al., 2012; Lin et al., 2013; Stith et al., 2008; Stodola and Burgers, 2016). DNA ligase I is recruited to the replisome and the newly synthesized fragments of a lagging strand can be sealed (Howes and Tomkinson, 2012; Waga et al., 1994).

Replication of chromatin states

While DNA replication, a process of duplicating genomic DNA, is fairly well understood, the mechanisms that regulate inheritance of epigenetic chromatin states, and therefore govern maintenance of gene expression profiles and cell identity, remain poorly understood. During DNA replication nucleosomes are disassembled ahead of the replication machinery in order to allow for new DNA synthesis, and this process presents a challenge for inheritance of parental nucleosomes. Canonical H3-H4 heterotetramers are recycled during DNA replication and multiple studies indicate that during this process old histones are not mixed with newly synthesized dimers during nucleosome re-formation following DNA replication (Leffak, 1984; Leffak et al., 1977; Yamasu and Senshu, 1990).

The clear mechanism facilitating the retention of parental nucleosome is lacking, however; it was proposed that CAF1 complex, which functions as a histone H3 chaperone involved in *de novo* nucleosome assembly, also plays a role in maintenance of parental H3 containing nucleosomes during S phase. Deletion of the CAF1 complex subunits in yeast was found to be associated with defects in the silencing of genes present in the heterochromatin regions, transcriptionally repressed telomeric heterochromatin and genes present at the silent

HM loci (Enomoto and Berman, 1998; Enomoto et al., 1997; Kaufman et al., 1998; Kaufman et al., 1997; Monson et al., 1997).

Recent reports from multiple groups described the MCM2 subunit of the MCM2-7 helicase complex as an important player also implicated in histone recycling. The mutations within MCM2 domain that confers histone binding leads to defects chromatin silencing in yeast (Foltman et al., 2013). Recent studies proposed that inheritance of canonical histone H3 containing parental nucleosomes rely on the activity of MCM2 chaperone. MCM2 was shown to interact with H3/H4 tetramers as well as H3/H4 dimers in complex with Asf1, and these interactions were proposed to be implicated in the stable inheritance of parental histones across DNA replication (Huang et al., 2015; Richet et al., 2015).

The FACT chaperone, that in humans comprises of Spt16 and SSRP1 subunits, was also proposed to be associated with parental histones during DNA replication. FACT is known for its role in chromatin disruption ahead of the RNA polymerase and chromatin reassembly after DNA transcription is completed (Hammond et al., 2017). FACT chaperon was also shown to interact with multiple components of the replication machinery (including MCM2, MCM4, Polα and RPA1), to travel along with the replisome, and was proposed to be required for replisome progression (Alabert et al., 2014; Foltman et al., 2013; Gambus et al., 2006; Kurat et al., 2017; Tan et al., 2006; VanDemark et al., 2006; Wittmeyer et al., 1999; Zhou and Wang, 2004). FACT has a capacity to interact with H2A-H2B histone dimers as well as H3-H4 tetramers. FACT complex was also found to interact with MCM2 through parental histones that have been released from chromatin solubilized with benzonase treatment (Belotserkovskaya et al., 2003; Foltman et al., 2013; Orphanides et al., 1999; Tsunaka et al., 2016).

Patterns of CENP-A nucleosome inheritance

CENP-A containing nucleosomes are not being turnover, therefore facilitating faithful centromere propagation through multiple generations. The timing of centromeric chromatin replication is not precisely defined as DNA replication was shown to occur asynchronously between centromeres of different chromosomes. It is proposed however, that centromeres undergo DNA replication during mid to late S phase (O'Keefe et al., 1992; Ten Hagen et al., 1990). Mechanistically, how centromeric DNA is replicated is not well understood. Recent study showed that DNA2 is implicated in efficient replication of centromeric DNA. DNA2 protein is known for its helicase and endonuclease activities, and is involved in replication fork processing and double-strand break resection. The DNA2 helicase/nuclease activity was also shown to be required for resolving DNA secondary structures arising at centromeres during DNA replication due to AT rich nature of a-satellite repeat sequences (Budd et al., 1995; Li et al., 2018)

Jansen et al, using a covalent fluorescent pulse-labeling with SNAP tagging, which allows for specific labeling if an existing pool of particular protein *in vivo*, demonstrated that the new CENP-A deposition in humans, in contrast to bulk histones, is uncoupled from DNA synthesis and occurs only once per cell cycle after cells complete mitosis. Remarkably, assembled CENP-A containing nucleosomes are equally partitioned to sister chromatids while cells undergo DNA replication such that each sister centromere receives approximately 50% of existing CENP-A (Jansen 2007, Ross 2016). This holds true during each S phase and regardless of either CENP-A levels dispersed during previous replication events or the size of a-satellite arrays present within centromeric DNA that differ among chromosomes (Ross 2016). While the existing CENP-A nucleosomes are distributed into daughter strands during S phase, the nucleosome gaps within CENP-A occupied domains are filled by newly deposited H3.1 and H3.3 containing nucleosomes. The H3.3 histone variant was proposed to

be a placeholder for CENP-A because its centromeric levels were significantly reduced in early G1, consistent with replacement by newly synthesized CENP-A (Dunleavy 2011). The two-single molecule optical mapping strategy on extended chromatin fibers was used to analyze the dynamics of CENP-A occupied domains within centromeric DNA. These efforts revealed that the CENP-A domain is precisely positioned at the a-satellite array at the same location throughout the cell cycle. The global analysis of centrochromatin dynamics, where pattern of distinct pools of nascent CENP-A sequentially incorporated over multiple cell cycles was analyzed on chromatin fibers, demonstrated that new CENP-A nucleosomes and deposited adjacent to existing CENP-A nucleosomes. All this evidence suggests that the dynamics of centrochromatin during CENP-A redistribution in S phase or new deposition following mitotic exit are similar among all human chromosomes.

The presence of the CATD domain within CENP-A nucleosomes was proposed to be sufficient for governing its stability and retention at the centromere throughout multiple generations (Bodor 2014). H3^{CATD} chimeric protein was fused to the SNAP-tag and the turnover rate of this chimeric protein was analyzed. These experiments showed that the CATD domain within CENP-A nucleosomes is sufficient for governing its stability and retention at the centromere throughout multiple generations (Bodor et al., 2014). In spite the fact that HJURP specifically recognizes CENP-A through CATD domain, strikingly the siRNA mediated downregulation experiments suggest however, that HJURP is dispensable for the inheritance of exogenously expressed CENP-A. These data suggest that there is mechanism that specifically recognizes CENP-A nucleosomes through the CATD domain and facilitates stable transition of centromeric nucleosomes throughout DNA replication, independently of CENP-A chaperone HJURP.

Recent study proposed that canonical H3.1 parental nucleosomes require the activity of MCM2 together with Asf1 chaperones in order to be stably retained across S phase. Huang et al., 2015 demonstrates that CENP-A, similarly to its cousin histone H3.1, can directly interact with MCM2 chaperone. This group also showed that the CENP-A specific chaperone HJURP interacts with MCM2 *in vivo*. In contrast to findings proposed by Bodor 2014, Huang et al., 2015 speculated that CENP-A inheritance during S phase might be governed by an analogous mechanism to H3 parental nucleosomes, where HJURP-MCM2 co-chaperone complex might be required for CENP-A recycling. However, the exact mechanism facilitating CENP-A inheritance throughout DNA replication w remains poorly understood and lie in wait to be explored.

CENP-C binding to the CENP-A containing nucleosomes was demonstrated to introduce surface and internal changes in the nucleosome structure that was proposed to confer CENP-A nucleosome stability. Downregulation of CENP-C in human tissue culture cells leads to rapid loss of existing CENP-A nucleosomes (Falk et al., 2016). CENP-C was also demonstrated to be required for new CENP-A deposition therefor it appears to be a crucial player in both centromere establishment and propagation (Erhardt et al., 2008), (Moree et al., 2011), (Falk et al., 2016).

Chapter 2: Dimerization of the CENP-A assembly factor HJURP is required for centromeric nucleosome deposition

Abstract

The epigenetic mark of the centromere is thought to be a unique centromeric nucleosome that contains the histone H3 variant, CENP-A. The deposition of new centromeric nucleosomes requires the CENP-A specific chromatin assembly factor HJURP. Crystallographic and biochemical data demonstrate that the Scm3-like domain of HJURP binds a single CENP-A—histone H4 heterodimer. However, several lines of evidence suggest that HJURP forms an octameric CENP-A nucleosome. How an octameric CENP-A nucleosome forms from individual CENP-A/histone H4 heterodimers is unknown. Here we show that HJURP forms a homodimer through its second HJURP_C domain. HJURP exists as a dimer in the soluble preassembly complex and at chromatin when new CENP-A is deposited. Dimerization of HJURP is essential for the deposition of new CENP-A nucleosomes. The recruitment of HJURP to centromeres occurs independent of dimerization and CENP-A binding. These data provide a mechanism whereby the CENP-A prenucleosomal complex achieves assembly of the octameric CENP-A nucleosome through the dimerization of the CENP-A chaperone HJURP.

Introduction

The equal distribution of chromosomes into daughter cells during mitosis depends on the proper assembly of a centromere on each chromosome. Centromere assembly occurs independently of DNA sequence, with the exception of budding yeast point centromeres (Allshire and Karpen, 2008; Cleveland et al., 2003; Stellfox et al., 2012). All eukaryotes use a conserved, CenH3 containing, centromere-specific nucleosome to determine (or mark) the site of the centromere. Human centromere-specific nucleosomes contain centromere protein-A (CENP-A) in place of histone H3. CENP-A containing nucleosomes are found interspersed with canonical histone H3 nucleosomes within human centromeres (Blower et al., 2002). CENP-A nucleosomes direct the recruitment of a constitutive centromere associated network (CCAN) and the kinetochore proteins that together orchestrate the attachment of chromosomes to the mitotic spindle and regulate cycle progression through the mitotic checkpoint.

Existing CENP-A is quantitatively retained at centromeres following DNA replication and redistributed to sister centromeres (Jansen et al., 2007). Thus, continuous inheritance of centromere position requires that new CENP-A deposition occur every cell cycle in order to maintain a sufficient number of CENP-A nucleosomes to specify the centromeric locus. The assembly of new centromeric nucleosomes depends on the CENP-A-specific chromatin assembly factor, HJURP (Holliday junction recognition protein) (Bernad et al., 2011; Dunleavy et al., 2009; Foltz et al., 2009). CENP-A interacts with HJURP as a soluble pre-nucleosomal complex. The deposition of centromeric nucleosomes in yeast requires the HJURP homolog, Scm3 (Camahort et al., 2007; Dechassa et al., 2011; Mizuguchi et al., 2007; Sanchez-Pulido et al., 2009; Stoler et al., 2007; Williams et al., 2009). HJURP and Scm3 share close to 69% homology within a small 52 amino acid region in the amino terminus of HJURP, which is required for CENP-A binding (Sanchez-Pulido et al., 2009;

Shuaib et al., 2010). The Scm3 domain of HJURP is sufficient to facilitate the formation of CENP-A nucleosomes *in vitro* and *in vivo* (Barnhart et al., 2011). The recruitment of HJURP and the deposition of CENP-A occur during early G1 (Dunleavy et al., 2009; Foltz et al., 2009; Jansen et al., 2007; Schuh et al., 2007). HJURP recruitment to centromeres depends on the activity of the Mis18 complex (Barnhart et al., 2011; Moree et al., 2011), which influences the histone modification and DNA methylation status of centromeres (Fujita et al., 2007; Kim et al., 2012). However, the mechanism by which Mis18 directs HJURP to centromeres remains unclear.

The crystal structures of the CENP-A—histone H4 heterotetramer, containing two copies each of CENP-A and H4, as well as the CENP-A octameric nucleosome have been solved (Sekulic et al., 2010; Tachiwana et al., 2011). Additional evidence suggests that the CENP-A nucleosome may transition from an octameric nucleosome to hemisome, containing a single copy of CENP-A and H4, as a cell progresses through the cell cycle (Bui et al., 2012; Shivaraju et al., 2012). Similar to the H3-H3 interface in the canonical nucleosome, dimerization of CENP-A is required for stable CENP-A deposition. Mutants of human CENP-A or the *Drosophila* homolog, CID, in which the CENP-A-CENP-A dimerization interface is disrupted, are unable to form stable nucleosomes *in vivo* (Bassett et al., 2012; Zhang et al., 2012), suggesting that formation of a CENP-A octamer is required for stable nucleosome formation. Human HJURP and yeast Scm3 mediate the formation of octameric nucleosomes *in vitro* (Barnhart et al., 2011; Dechassa et al., 2011; Kingston et al., 2011; Shivaraju et al., 2011). Interestingly, several recent biochemical studies of HJURP/Scm3 in complex with CENP-A have demonstrated that CENP-A interacts with HJURP as a heterodimer containing a single copy of CENP-A and histone H4 (Bassett et al., 2012; Cho and Harrison, 2011; Feng et al., 2011; Zhou et al., 2011). These observations raise the

question of how an octameric CENP-A nucleosome may be assembled from a heterodimeric intermediate.

Vertebrate HJURP proteins are significantly larger than their yeast orthologs and contain additional conserved domains (Sanchez-Pulido et al., 2009). Human HJURP contains two HJURP_C-terminal domains (HCTD) within the carboxyl terminal half of the protein. HJURP_C domains are also found in the myocyte enhancer factor 2 (MEF2) transcription factors (Potthoff and Olson, 2007), but a functional role for this domain in MEF2 has not been determined. Only the Scm3 domain has been implicated in CENP-A deposition.

Here we demonstrate that HJURP, in the prenucleosomal complex, forms a dimer through its carboxyl terminus. We show that the targeting of HJURP to centromeres occurs independently of HJURP dimerization and requires a region of HJURP between the conserved domain (CD) and the second HJURP C-terminal domain (HCTD2) (Sanchez-Pulido et al., 2009). Importantly, we also find that dimerization of HJURP is essential for the assembly of CENP-A nucleosomes at centromeres. These data identify the region of HJURP sufficient for centromere targeting and provide a potential mechanism by which octameric CENP-A nucleosomes are assembled from a heterodimeric intermediate.

Results

Centromeric localization of HJURP through the carboxyl terminus.

A set of non-overlapping HJURP truncation proteins were examined to determine which domains of HJURP mediate its specific recruitment to centromeres. Centromeric localization was assessed in early G1 cells (midbody positive) at 24 hours post-transfection. Full-length HJURP localized to centromeres in approximately 82% (±5) of G1 cells (n>60 cells, 2 independent experiments) (Figure 2.1A,B). The Scm3 domain alone was not recruited to centromeres, as shown previously (Barnhart et al., 2011). The HJURP Conserved domain (CD) is a distinguishing feature of vertebrate HJURP orthologs (Sanchez-Pulido et al., 2009) but is absent from HJURP/Scm3 orthologs in fungi. The HJURP$^{1-348}$ deletion mutant, which contains the Scm3 and CD domains of HJURP, did not localize to centromeres; however, the complementary deletion mutant, containing the remaining carboxyl half (HJURP$^{352-end}$), was recruited to centromeres (Figure 2.1A,B). The HJURP$^{352-end}$ mutant was as efficient at centromere recruitment in G1 cells as the full-length protein (91%±1). When the recruitment was compared between all cells (asynchronously dividing) and just those in G1 (based on the presence of a mid-body), HJURP$^{352-end}$ was enriched at centromeres 3-fold during G1 similar to full-length HJURP. These data demonstrate that centromere targeting is controlled by the carboxyl half of HJURP and that this protein truncation is under the same cell-cycle control as endogenous HJURP. The carboxyl half of HJURP contains the HJURP_C terminal domains (HCTD) and based on these results also contains the specific centromere-targeting domain. We conclude that centromere targeting does not require the CD domain or CENP-A binding through the Scm3 domain.
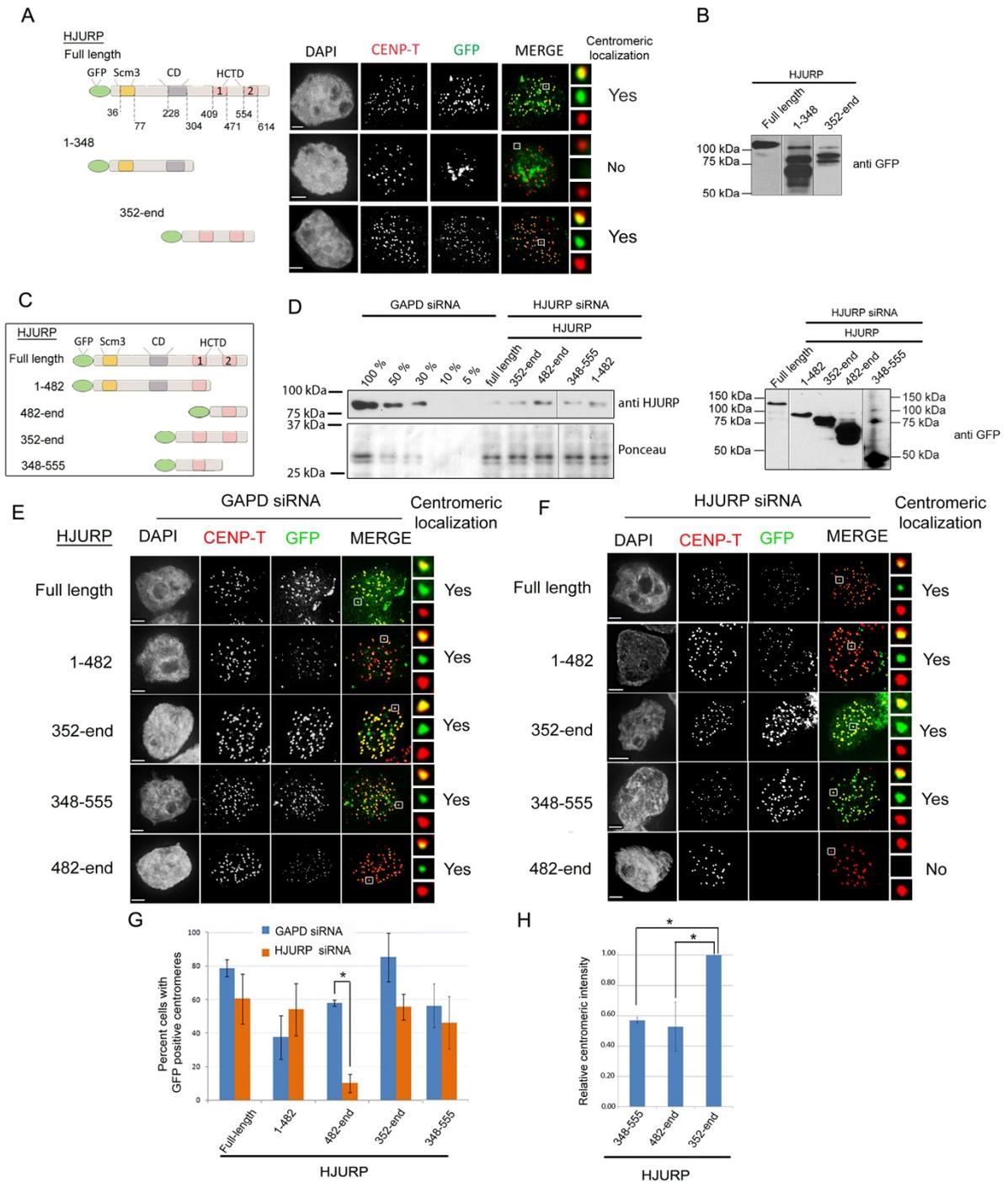
Figure 2.1. Identification of CIS-acting elements within HJURP required for centromere recruitment

Figure 2.1. Identification of CIS-acting elements within HJURP required for centromere recruitment (A) Schematic representations of transfected HJURP fragments (left panel) and corresponding representative images of cells transfected with GFP tagged HJURP fragments (right panel). DNA was visualized by DAPI staining; anti-CENP-T is shown in red, the GFP-tagged fragments are in green. Merge includes CENP-T and GFP signals. Scale bar is 2μm in all panels. Boxed regions are magnified to the right of merged images. (B) Anti-GFP western blot shows the expression of GFP-HJURP fusion proteins expressed in A. (C) Schematic representations of transfected GFP-HJURP fragments used in D-F. (D) Western blot showing the efficiency of HJURP depletion (left blot) and expression of transfected constructs (right blot) used in E and F. Efficiency of HJURP siRNA treatment was assessed by anti-HJURP antibody, the expression efficiency was assessed by anti-GFP antibody, ponceau staining serves as a loading control. (E) (F) Representative images of cells expressing GFP-HJURP fragments and treated with either GAPD (E) or HJURP (F) siRNA. DNA was visualized by DAPI; CENP-T is shown in red, the GFP-tagged fragments in green. (G) Quantification of the percentage of G1 cells in which GFP-HJURP was recruited to centromeres. Data are from at least 2 independent experiments, >60 cells per condition. Error bars represent the standard deviation. * indicates $p<0.01$. (H) Relative centromeric intensity of HJURP fragments. n>180 centromeres per condition. * indicates $p<0.05$.

Recruitment of HJURP to centromeres through direct recruitment and dimerization.

To further refine the centromere targeting domain, a pair of HJURP truncation mutants (HJURP$^{1-482}$ and HJURP$^{482-end}$) was expressed, which divided HJURP into amino and carboxyl halves (Figure 2.1C, D). By separating the two C-terminal repeats (HCTD's) into two fragments, we predicted only one of these two truncation mutants would localize to centromeres. However, we observed that both HJURP$^{1-482}$ and HJURP$^{482-end}$ could localize to centromeres (Figure 2.1E).

We hypothesized that the recruitment of both non-overlapping HJURP carboxyl-terminal fragments to centromeres may occur directly through the centromere-targeting domain and indirectly through dimerization with endogenous HJURP. The recruitment of an HJURP fragment via dimerization should be dependent on endogenous HJURP for centromere localization. Conversely, the recruitment of a fragment containing the direct centromere-targeting domain should localize to centromeres independently of endogenous HJURP. Therefore, we tested the recruitment of a series of HJURP truncation mutants in cells where endogenous HJURP was depleted by siRNA.

Endogenous HJURP was depleted for 24 hours using an siRNA directed against the 3'UTR of HJURP and followed by expression of HJURP truncation mutants (Figure 2.1D,E). Endogenous HJURP expression was decreased by siRNA treatment to approximately 20% of GAPD-treated levels (Figure 2.1D). Centromeric localization of the exogenously expressed HJURP fragments was analyzed in early G1 phase cells (Figure 2.1E,F). Exogenous full-length HJURP efficiently associated with the centromeric chromatin in control (GAPD) and HJURP siRNA treated cells. HJURP$^{352-end}$ was recruited to centromeres in control and HJURP siRNA treated cells (Figure 2.1E,F). The number of cells that recruited HJURP$^{352-end}$

under control and HJURP siRNA conditions recapitulated the same degree of recruitment as the exogenously expressed full-length HJURP under similar conditions (Figure 2.1G).

Centromeric recruitment of the complementary pair of HJURP deletion mutants, HJURP[1-482] and the reciprocal fragment, HJURP[482-end], was tested following siRNA treatment. HJURP[1-482] was recruited to centromeres with similar efficiency in control and HJURP siRNA treated cells. It was also recruited to centromeres in a similar fraction of cells as exogenous full-length HJURP in the HJURP depletion condition (Figure 2.1G). In contrast, HJURP depletion abolished centromeric recruitment of HJURP[482-end] despite its recruitment to centromeres in the control GAPD depletion (Figure 2.1E-G). The dependence of HJURP[482-end] on endogenous HJURP for its recruitment is consistent with dimerization of HJURP at centromeres. Together, these data suggest that dimerization and recognition of the centromere are mediated by distinct domains within HJURP.

We compared the centromeric intensities between HJURP[352-end] (HCTD1 & HCTD2) with HJURP[348-555] (HCTD1) and HJURP[482-end] (HCTD2), each of which contains only a single HCTD. HJURP[352-end] contains both repeats and should therefore be recruited by both direct targeting to the centromere and through dimerization. Consistent with this, HJURP[352-end] GFP signal was 1.5x more intense at centromeres than either HJURP[348-555] or HJURP[482-end,] which contain the individual HCTD1 and HCTD2 regions, respectively (Figure 2.1E,H). Therefore, the HCTD2 domain mediates a multimerization of HJURP, and amino acids 348 to 482 directly recruit HJURP to centromeres.

Dimerization of HJURP through the carboxyl terminus.

The ability of HJURP to self-associate was examined using a LacI/LacO based *in vivo* interaction assay to directly assess if HJURP multimerizes *in vivo* (Figure 2.2A). Full-length or HJURP[352-end] was fused to the lac repressor (LacI) and expressed as bait in cells that have a stably integrated LacO array (Barnhart et al., 2011; Janicki et al., 2004). The interaction between HJURP proteins was tested by expressing GFP-HJURP fragments as prey. Tethering HJURP to the LacO array resulted in GFP-HJURP recruitment (Figure 2.2B,C). Furthermore, LacI-HJURP was able to recruit carboxyl-terminal fragments of HJURP (Figure 2.2B,C). This interaction only required the HJURP carboxyl terminus because tethering the LacI-HJURP[352-end] fragment to the array was sufficient to recruit GFP-carboxyl fragments, containing amino acids 352-end and 482-end (Figure 2.2D,E). Full-length and HJURP[352-end] showed minimal recruitment of an HJURP fragment containing amino acids 1-482 to the array, showing that HCTD2 in the carboxyl terminus is the primary sight of HJURP self-association. The HCTD1 domain present in HJURP[348-555] was unable to be efficiently recruited by either the full length or HJURP[352-end] bait protein. We conclude that HJURP fragments containing the second HCTD2 domain of HJURP are sufficient to mediate self-interaction *in vivo*.
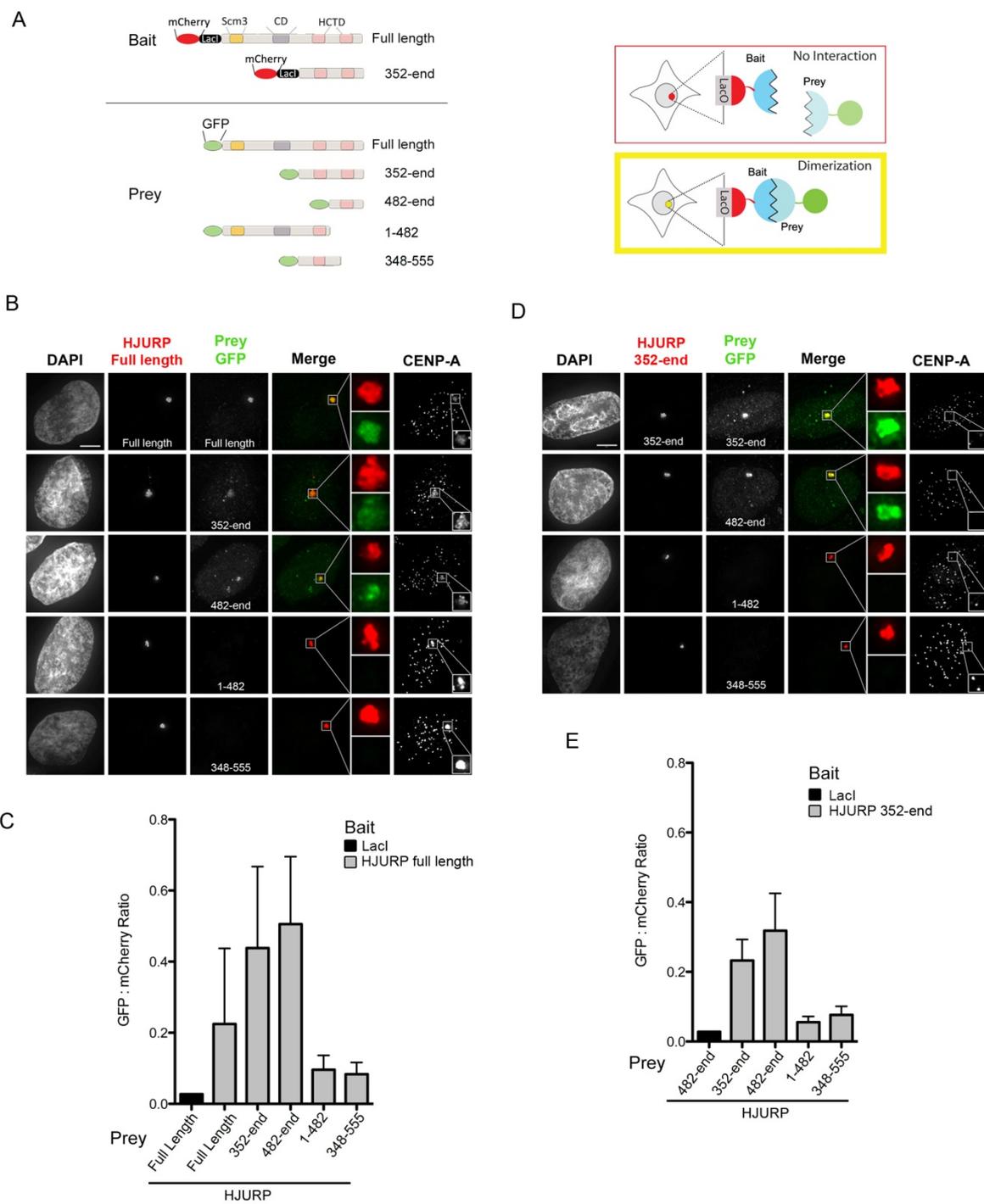
Data in this figure was generated by Dr. M.C. Barnhart-Dailey.

Figure 2.2. *In vivo* recruitment of HJURP through the carboxyl terminus

Figure 2.2. *In vivo* recruitment of HJURP through the carboxyl terminus. (A) Schematic of LacO-LacI interaction assay and the bait and prey constructs used in the study. (B) U2OS-LacO cells were co-transfected with mCherry-LacI-HJURP^{FullLength} and indicated GFP-tagged prey fragments. DNA is stained with DAPI. Centromere staining and endogenous CENP-A recruitment to the arrays is shown using anti-CENP-A antibody. Scale bar represents 5 μM. All images are scaled equally. Boxed regions are magnified to the right of merged images. (C) Quantitation of prey protein recruitment to the array when HJURP^{FullLength} (grey) or control mCherry-LacI alone (black) is targeted. Recruitment is expressed as the ratio of GFP to mCherry integrated intensity at the array. (D) Cells co-transfected with mCherry-LacI-HJURP$^{352-end}$ or mCherry-LacI alone as bait with the indicated GFP-tagged prey fragments. Cells were stained with DAPI to visualize DNA and anti-CENP-A as in (B). Scale bar represents 5 μM (E) Prey protein recruitment to the array in response to mCherry-LacI-HJURP$^{352-end}$ or control mCherry-LacI targeting is quantified as in (C). The GFP:mCherry ratios are plotted as the mean of n=3 experiments at ≥30 arrays per condition. Error bars represent standard deviation.

In order to elucidate whether HJURP self-association through its carboxyl terminus is a direct interaction that occurs without any additional factors, the putative dimerization domain (HJURP$^{482\text{-end}}$) was expressed and purified from bacteria (Figure 2.3A,2.S1A). His-tagged HJURP$^{482\text{-end}}$ protein migrates on a denaturing SDS-PAGE gel as a 36 kDa protein, which is consistent with its calculated molecular weight of 30kDa. Based on size exclusion chromatography, the Stokes radius of His-HJURP$^{482\text{-end}}$ was calculated as 5.46nm (Figure 2.3B-D, 2.S1B), which is twice as large as the expected 2.48 nm Stokes radius of a 30kDa globular protein. This larger than expected Stokes radius may indicate that either HJURP$^{482\text{-end}}$ is an elongated protein or that it exists as a multimer. For comparison, MBP-tagged HJURP$^{352\text{-}482}$, which should not dimerize, and is a larger protein than HJURP$^{482\text{-end}}$ (58kDa versus 30kDa), elutes from the size exclusion column with a smaller Stokes radius of 4.46 nm (Figure 2.3B,C). The sedimentation coefficient for HJURP$^{482\text{-end}}$ was determined by sucrose gradient (Figure 2.S1B-D). Based on these analyses, the molecular weight of HJURP$^{482\text{-end}}$ was calculated as 59 kDa, consistent with the formation of a dimer (Siegel and Monty, 1966) (Figure 2.S1D). An MBP pull-down assay using two differently tagged recombinant proteins, MBP-HJURP$^{482\text{-end}}$ and His-HJURP$^{482\text{-end}}$, demonstrates that the carboxyl terminus of HJURP self-associates (Figure 2.3E). The combined *in vitro* and *in vivo* data lead us to conclude that HJURP forms a dimer through a direct self-interaction mediated by the HCTD2 domain.
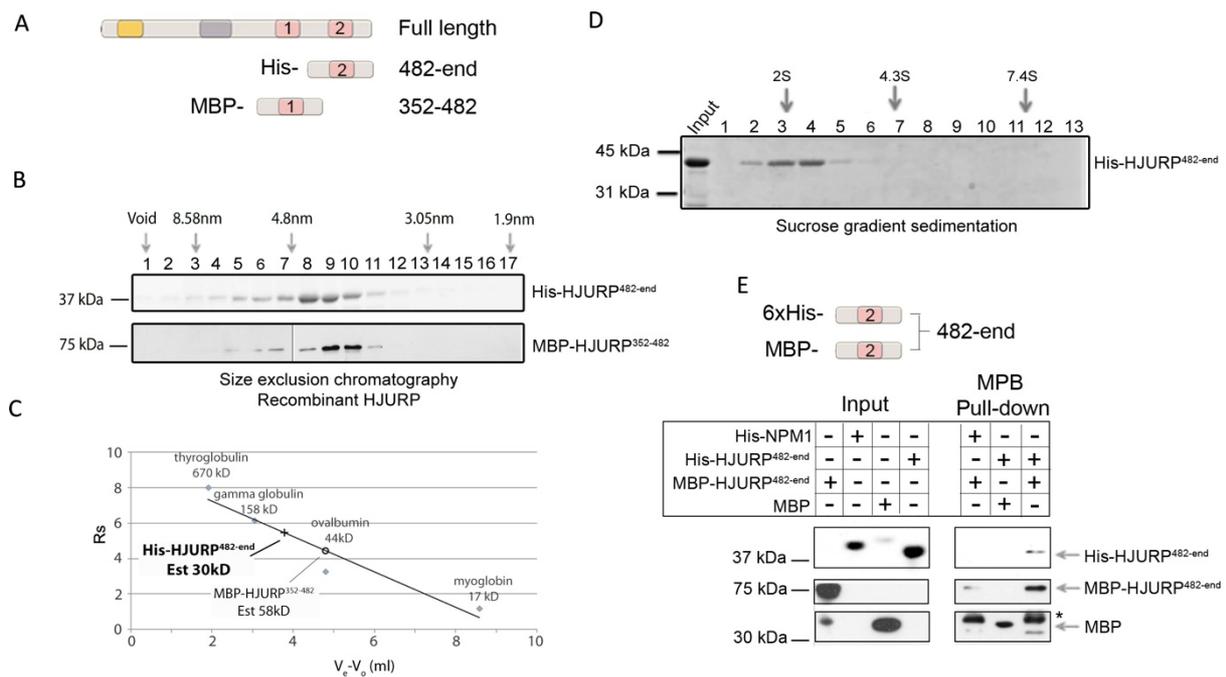
Figure 2.3. Dimerization of HJURP *in vitro*. (A) Schematic of recombinant carboxyl terminal HJURP fragments expressed in bacteria. (B) Recombinant MBP-HJURP[352-482] and His-HJURP[482-end] were analyzed by size exclusion chromatography. The arrows indicate the elution of protein standards. His-HJURP[482-end] has a predicted molecular weight of 30 KDa and was detected by Coomassie stain. MBP-tagged HJURP[352-482] has a predicted molecular weight of 58 KDa and was detected by immunoblot using anti-MBP antibody. (C) The elution of the standards and HJURP fragments are plotted relative to their stokes radius (Rs). (D) Coomassie stained SDS-PAGE gel of fractions collected after sucrose gradient ultracentrifugation of His-HJURP[482-end] fragment. (E) Schematic of recombinant carboxyl terminal HJURP fragments expressed in bacteria (top panel). MBP pull down experiment demonstrating direct interaction between differently tagged HJURP 482-end fragments (bottom panel). His and MBP-tagged proteins were visualized by antibody staining. Asterisk indicates MBP-HJURP breakdown product.

A
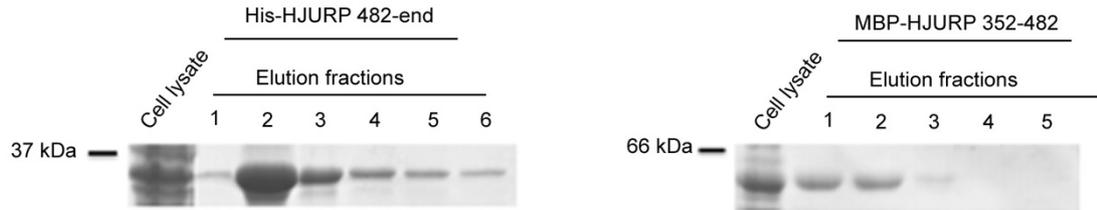
B

	MBP-HJURP$^{352-482}$	His-HJURP$^{482-748}$
MW based on amino acid composition	58 kDa	30 kDa
Expected Rs	3.23 nm	2.48 nm
Calculated Rs	4.46 nm	5.46 nm
Calculated Rs : Expected Rs	1.30	2.2

C

D
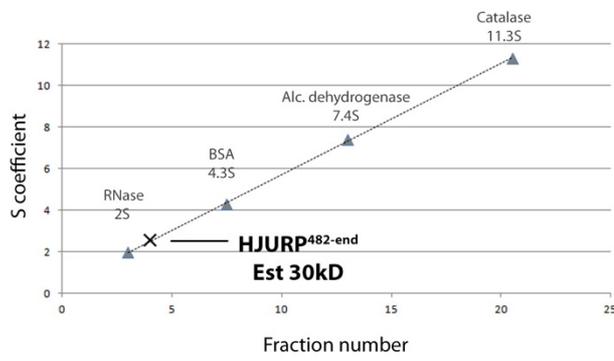
	Theretical MW [kDa]	Rs [nm]	S	Native MW [kDa]	Native MW: theoretical MW
His-HJURP 482-end	30.00	5.46	2.55	59	1.95

Figure 2.S1 (A) Coomassie stained SDS-PAGE gel of elution fractions eluted from purification of bacterially expressed HJURP fragments. (B) Table showing the calculated Rs of analyzed proteins from size exclusion chromatography and expected Rs for globular proteins with corresponding molecular mass (right panel). (C) Graph correlating peak fraction of the protein standards and His-HJURP$^{482-end}$, resulting from sucrose gradient separation, with their respective sedimentation coefficients. (D) Table showing the calculated molecular weight of His-HJURP$^{482-end}$ fragment.

Dimerization of HJURP forms a high molecular weight prenucleosomal complex.

To determine whether HJURP is present as a dimer in the prenucleosomal CENP-A complex we determined the native molecular weight of the HJURP complex from cells transiently expressing GFP-tagged full length and truncated HJURP. The prenucleosomal complex containing full-length GFP-tagged HJURP protein migrates on a sucrose gradient as a large complex with a sedimentation coefficient of 10.3S. This is slightly higher than the 10S reported previously for the endogenous CENP-A prenucleosomal complex, possibly due to the addition of the 30 kDa GFP-tag (Foltz et al., 2009). The calculated molecular weight of the prenucleosomal complex based on sedimentation and size exclusion chromatography was approximately 347kDa (Figure 2.4A,B), 2.5 times larger than the expected size of a heterotrimer-containing a single copy of CENP-A, histone H4, and GFP-HJURP (approximately 138 kDa).

Amino acids 482-748 were sufficient to interact with full-length HJURP in the LacO/LacI interaction assay and formed a dimer *in vitro*, making this region a prime candidate to mediate multimerization of the complex *in vivo*. To determine if the dimerization domain is responsible for forming the high molecular weight HJURP prenucleosomal complex, we examined the size of the soluble HJURP complex formed by HJURP[1-482], which lacks the dimerization domain. As we expected, HJURP[1-482] formed a significantly smaller complex with a molecular weight of approximately 130 kDa (Figure 2.4A,B, 2.S2A,B). This value is consistent with a complex that contains only a single HJURP, along with the CENP-A, histone H4 heterodimer, calculated to be 109kDa. We found the dimerization domain alone (HJURP[482-end]) was sufficient to form a dimer *in vivo*. We calculated the native molecular weight of HJURP[482-end] as 155 kDa (Figure 2.4A,B). The theoretical molecular mass for this fragment is expected to be 56kDa.

To determine if the HJURP prenucleosomal complex contains more than one HJURP molecule we co-transfected cells with constructs expressing GFP-tagged full-length HJURP and HA-tagged full-length HJURP and immunoprecipitated using anti-GFP antibodies. Under these conditions, GFP-HJURP associated with HA-HJURP (Figure 2.4C). The immunoprecipitations did not contain histone H2B and were therefore not chromatin-associated complexes but represent soluble HJURP multimers (Figure 2.S2C). We observed an interaction between full-length HJURP and the HJURP$^{482\text{-end}}$ fragment in co-immunoprecipitations from cells transiently transfected with GFP-HJURP$^{482\text{-end}}$ and HA-tagged HJURP protein, but not in control immunoprecipitations (Figure 2.4C, 2.S2D). Together these data demonstrate the HJURP carboxyl terminal tail (amino acids 482-748) is required for the formation of the multimeric HJURP prenucleosomal complex.
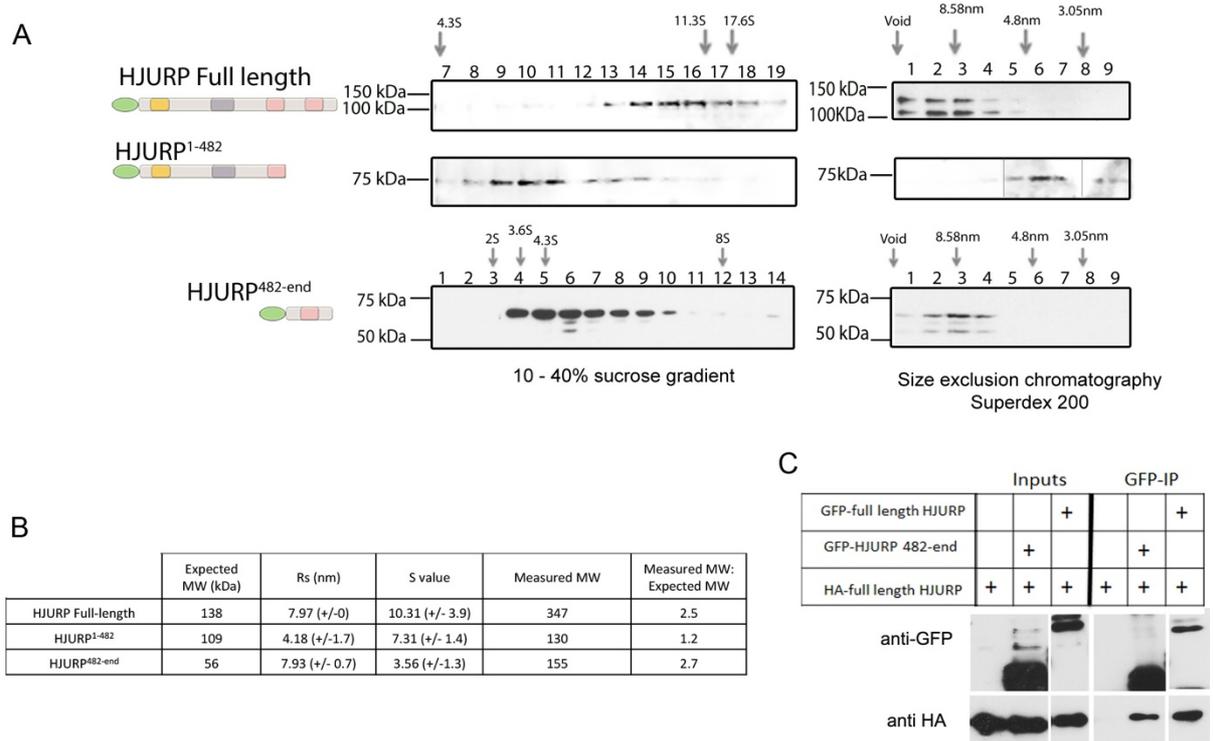
Figure 2.4. *In vivo* dimerization of HJURP in the pre-nucleosomal complex. (A) Anti-GFP immunoblots of fractions collected after sucrose gradient ultracentrifugation (left) or size exclusion chromatography (S.E.C) (right) of chromatin-free extracts from GFP-HJURP full-length, GFP-HJURP[1-482] and GFP-HJURP[482-end] expressing HEK293 cells. The arrows indicate the migration of protein standards. (B) Table showing the measured Rs, sedimentation coefficient (S), and expected molecular weights of the HJURP proteins analyzed by S.E.C and sucrose gradient sedimentation. Expected molecular weights include CENP-A and histone H4 (28kDa) for proteins that contain the CENP-A binding domain. (C) Cell extracts co-expressing full-length GFP-HJURP or GFP-HJURP[482-end] with HA-tagged full-length HJURP were subjected to anti-GFP immunoprecipitation followed by immunoblot using antibodies against HA and GFP.
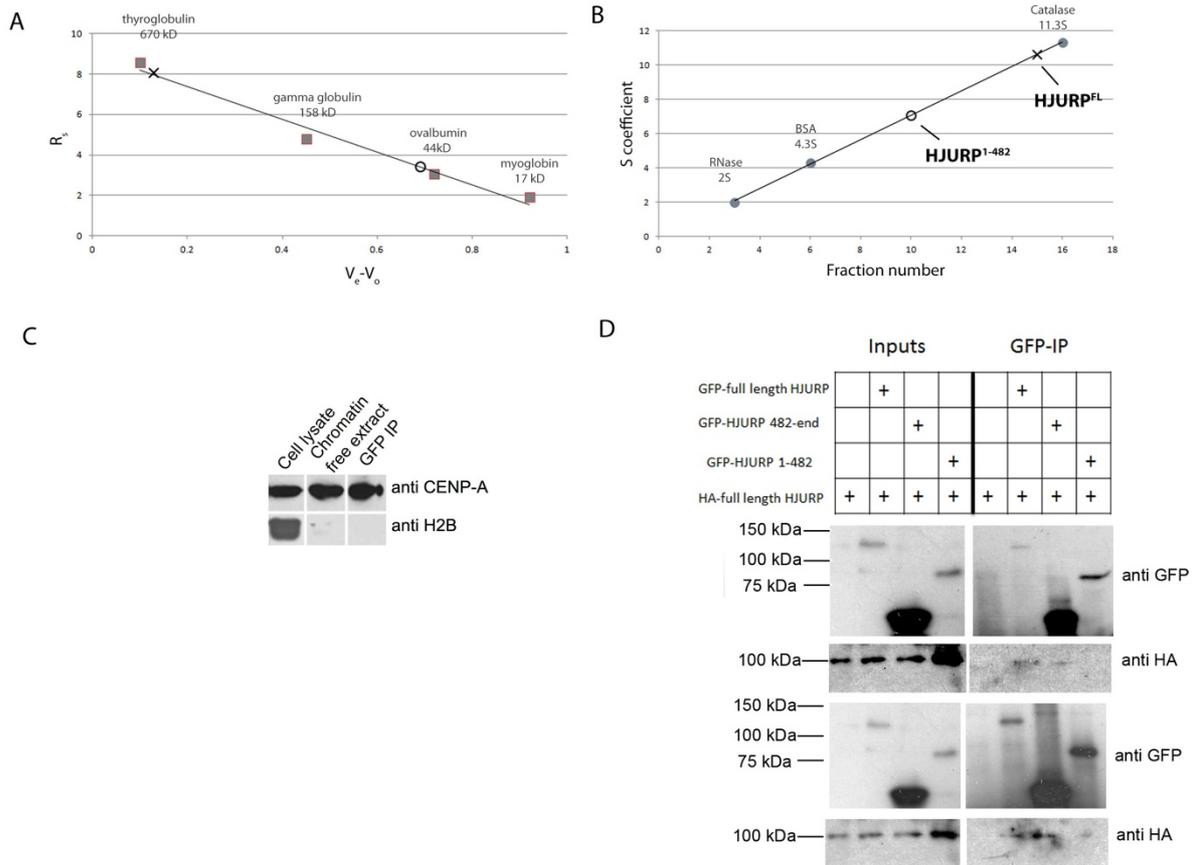
Figure 2.S2 (A) Graph correlating the Ve-Vo and Rs values of the protein standards and HJURP proteins analyzed by size exclusion chromatography. The molecular weights of the standard proteins are also displayed. (B) Graph correlating the peak fraction of the globular protein standards, resulting from sucrose gradient separation, with their sedimentation coefficients. The values of sedimentation coefficient corresponding to the standard proteins are displayed. B and C were used for calculating the native molecular mass of the ectopically expressed HJURP mutants. (C) Western blot from GFP immunoprecipitation showing that inputs and immunoprecipitation were prepared from chromatin free extracts, as the immunoprecipitations containing YFP-CENP-A did not contain histone H2B. (C) Cell extracts co-expressing full-length GFP-HJURP, GFP-HJURP$^{1-482}$ or HJURP$^{482-end}$ with HA-tagged full-length HJURP were extracted in RIPA buffer plus 20mM MgCl$_2$, treated with DNase and subjected to anti-GFP immunoprecipitation followed by immunoblot using antibodies against HA and GFP.

Dimerization of HJURP is required for CENP-A deposition.

Since dimerization of HJURP occurs on centromeric chromatin, we hypothesized that new CENP-A deposition requires HJURP dimerization. A CENP-A SNAP-tag assay was used to determine if new CENP-A was recruited to centromeres when endogenous HJURP was depleted by siRNA and replaced with exogenous HJURP that lacked the dimerization domain (HJURP$^{1\text{-}482}$). The SNAP-tag assay specifically follows the incorporation of new SNAP-tagged CENP-A nucleosomes by blocking detection of existing CENP-A with a non-fluorescent SNAP substrate and labeling new CENP-A with a SNAP substrate that is fluorescent (Foltz et al., 2009; Jansen et al., 2007). SNAP-tagged CENP-A cells were treated with HJURP 3'UTR siRNA for 24 hours to deplete endogenous HJURP followed by expression of either full-length or truncated HJURP replacement fragments (Figure 2.5A). Cells were given 24 hours to express the HJURP replacement fragments, and then new CENP-A assembly was assayed over the following 24 hours during which time CENP-A deposition was dependent on the replacement construct (Figure 2.5A). As expected, HJURP siRNA treatment of CENP-A SNAP cells significantly decreased the percentage of cells with new SNAP labeled CENP-A at centromeres and reduced the amount of new SNAP-labeled CENP-A at centromeres (Figure 2.5B-D, 2.S3A). The expression of full-length HJURP restored new CENP-A assembly in HJURP siRNA treated cells (Figure 2.5 C,D). As a negative control, we expressed the HJURP$^{202\text{-end}}$ fragment, which lacks the CENP-A binding domain and therefore should not rescue CENP-A deposition. New CENP-A recruitment to centromeres in cells transfected with the HJURP$^{202\text{-end}}$ was significantly impaired relative to the GAPD siRNA control (Figure 2.5C,D, 2.S3A).
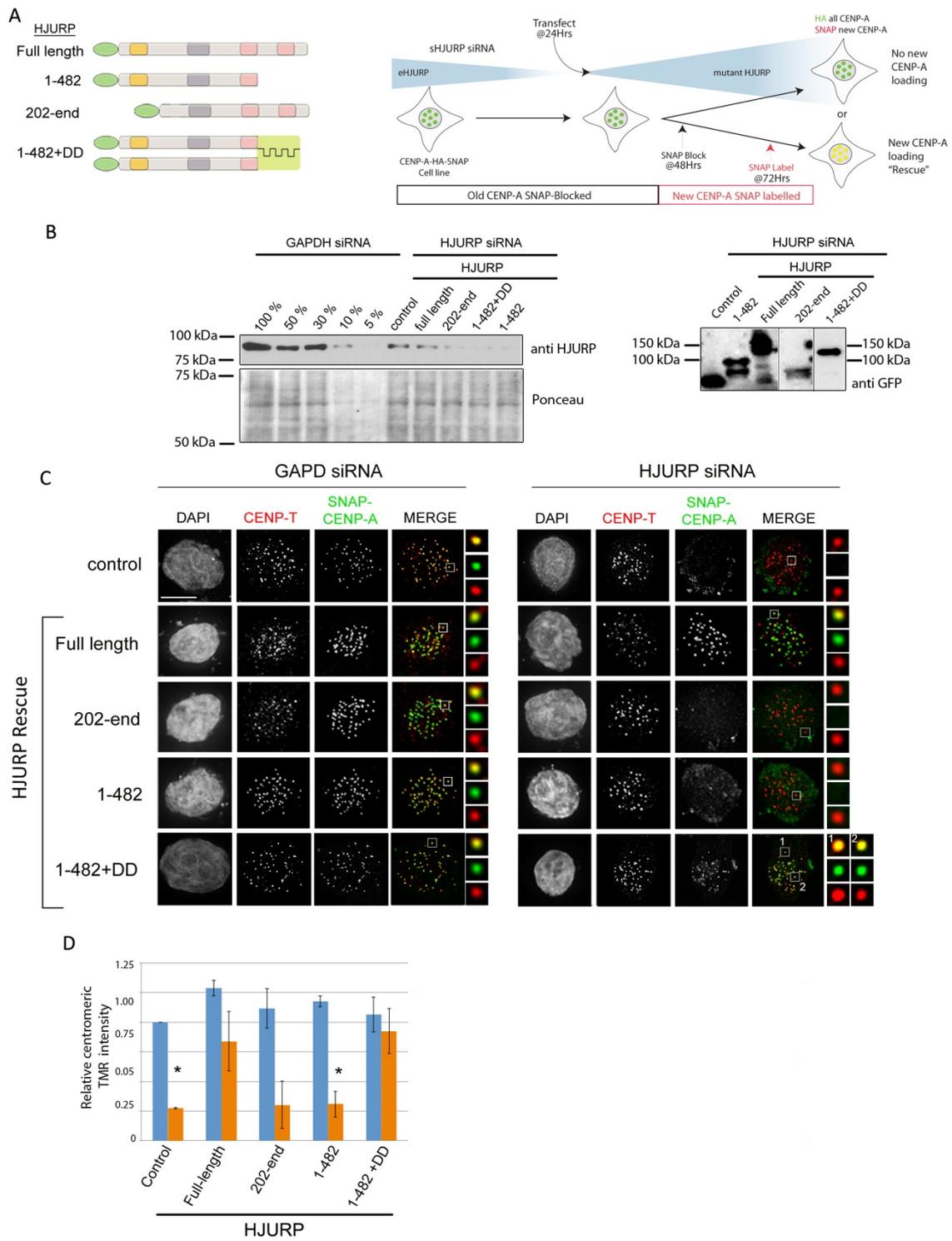
Figure 2.5. CENP-A assembly requires HJURP dimerization

Figure 2.5. CENP-A assembly requires HJURP dimerization. (A) Schematic showing the GFP-HJURP constructs used and the design of the SNAP-tag experiment testing new CENP-A recruitment. The 1-482+DD (dimerization domain) mutant contains the Lac repressor fused the carboxyl terminus of HJURP$^{1-482}$. (B) Western blot showing the efficiency of HJURP depletion and expression of transfected constructs used in (C) and (D). HJURP siRNA treatment efficiency was assessed using an anti-HJURP antibody (left). Ponceau staining is a loading control. Expression efficiency of HJURP truncation mutants was assessed by anti-GFP antibody (right). (C) Representative images of new CENP-A loading in the SNAP-tagging experiment. Cells were treated with either GAPD (left panel) or HJURP siRNA (right panel). New CENP-A (TMR-star labeled SNAP-CENP-A, green in merge) is recruited to centromeres in control and HJURP rescue conditions. Immunostaining for CENP-T (red) identifies centromeres. Scale bar is 5μm in all panels. Boxed regions are magnified to the right of merged images. Two boxed regions (1,2) are shown for HJURP siRNA treated HJURP$^{1-482-+DD}$. (D) Fluorescence intensity of centromeric TMR-star labeled SNAP-CENP-A was measured relative to GAPD siRNA treated control. n>180 centromeres per condition from 2 independent experiments. * indicates p<0.05.
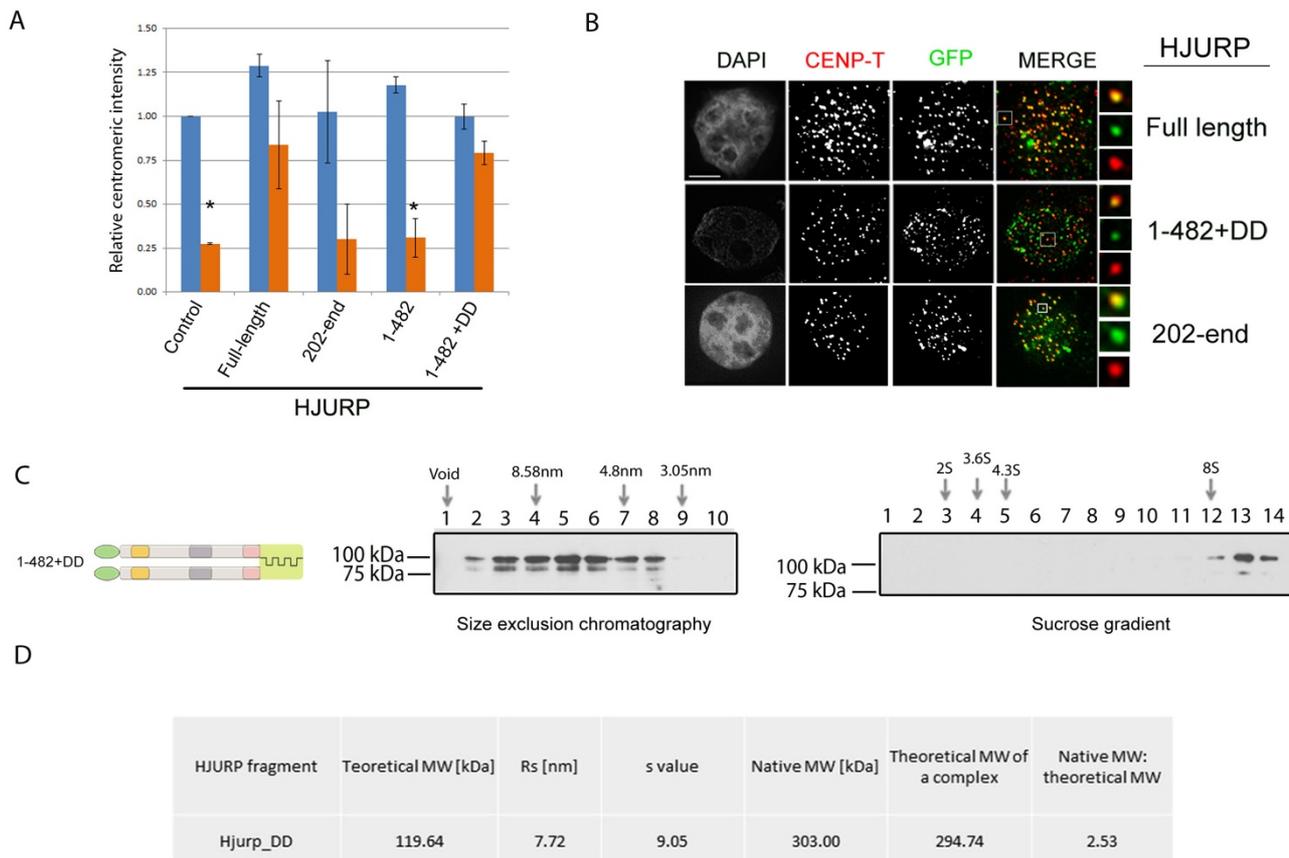
Figure 2.S3 (A) Quantitation of SNAP-tag experiments graphed as percent of SNAP-positive cells. Results represent at least 2 independent experiments, > 60 cells per experiment. Errors bars are standard deviation. * indicates p<0.01. (B) The recruitment of both HJURP[1-482+DD] and HJURP[202-end] mutants to centromeres. DNA is stained with DAPI and CENP-T antibody (red) marks centromeres. GFP-HJURP fragments are in green. (C) Chromatin-free extracts derived from HEK293 cells expressing GFP-HJURP[1-482+DD] were subjected to size exclusion chromatography and sucrose gradient followed by immunoblotting of the elution fraction. The arrows indicate the elution (left panel) and migration (right panel) of protein standards. Immunoblotting was performed with an anti-GFP antibody. (D) Table showing the calculated Rs, sedimentation coefficient, measured and expected molecular weight of GFP-HJURP[1-482+DD].

If dimerization is required for deposition, we expect expression of HJURP$^{1-482}$, lacking the dimerization domain, to reduce new CENP-A deposition, similarly to the CENP-A binding domain mutant (HJURP$^{202-end}$). New CENP-A deposition was not affected in GAPD siRNA treated cells co-transfected with HJURP$^{1-482}$; however, new CENP-A deposition was significantly decreased when HJURP was depleted (Figure 2.5C,D, 2.S3A). The reduction in new CENP-A assembly was similar to that observed in the control and the CENP-A binding mutant HJURP$^{202-end}$. These results demonstrate that the dimerization domain of HJURP is required for CENP-A deposition.

In order to determine if HJURP dimerization is the primary function of the HJURP carboxyl terminus, amino acids 483-748 were replaced with an exogenous dimerization domain (HJURP$^{1-482+DD}$). We used the Lac repressor, which has been previously engineered to form a dimer (Chen and Matthews, 1992). The HJURP$^{1-482+DD}$ fusion protein was recruited to centromeres as expected (Figure 2.S3B). We determined the native molecular weight of the HJURP$^{1-482+DD}$ protein expressed in cells based on hydrodynamic analysis to be 303 kDa, close to twice its predicted size (148kDa when complexed with CENP-A and histone H4) and consistent with formation of a dimer (Figure 2.S3C,D).

We tested if the addition of the dimerization domain was sufficient to rescue the CENP-A deposition defect of HJURP$^{1-482}$. HJURP$^{1-482+DD}$ rescued new CENP-A deposition in HJURP siRNA treated cells when compared to HJURP that lacked the dimerization domain (HJURP$^{1-482}$) (Figure 2.5C,D, 2.S3A). The percentage of cells with new CENP-A deposited at centromeres in HJURP siRNA treated cells was similar to GAPD siRNA controls when HJURP$^{1-482+DD}$ was expressed (Figure 2.5C,D, 2.S3A). Importantly, HJURP$^{1-482+DD}$ was able to rescue the degree of new CENP-A deposition per centromere to the same level as full-length HJURP (Figure 2.5G). These data demonstrate the primary function of the carboxyl

terminus of HJUPR is to form an HJURP dimer and that dimerization is required for the

stable assembly of new CENP-A nucleosomes.

Discussion

Crystal structures of the CENP-A/HJURP prenucleosomal complex demonstrate that HJURP precludes the formation of a pre-nucleosomal CENP-A/histone H4 heterotetramer by blocking the CENP-A self-dimerization domain (Hu et al., 2011). Yet, the HJURP-mediated deposition of stable centromeric nucleosomes requires an intact dimerization surface within CENP-A (Bassett et al., 2012; Zhang et al., 2012). Additionally, *in vitro* chromatin assembly assays show that human HJURP and yeast Scm3 mediate the assembly of octameric nucleosomes onto DNA templates (Barnhart et al., 2011; Dechassa et al., 2011; Shivaraju et al., 2011). Together these observations suggest that the *in vivo* deposition of CENP-A nucleosomes by HJURP results in a CENP-A nucleosome that contains two copies of CENP-A, which is consistent with an octameric nucleosome. Therefore, the formation of a CENP-A nucleosome requires two HJURP proteins to be recruited to each site of new CENP-A deposition. Here we show that HJURP forms a homodimer through the second HCTD repeat in its carboxyl terminus (Figure 2.6A). HJURP is dimerized in the prenucleosomal and chromatin-associated complexes. HJURP dimerization is required for new CENP-A deposition, providing a mechanism by which an octameric nucleosome is assembled at the centromere from two new CENP-A—histone H4 heterodimers (Figure 2.6B, nucleosome model).

Recent work in both human and budding yeast suggests that CENP-A nucleosomes occupy two distinct states during the cell cycle, an octameric and hemisome (tetrameric) form (Bui et al., 2012; Shivaraju et al., 2012). The hemisome form contains a single copy of each histone: CENP-A, H4, H2A and H2B. Based on these observations an alternative model exists whereby HJURP dimerization links the existing CENP-A hemisome to the incoming new CENP-A—histone H4 heterodimer (Figure 2.6B, hemisome model). In this model, the Scm3 domain of one HJURP dimer subunit interacts with the existing centromeric hemisome.

The Scm3 domain of the second HJURP dimer subunit binds a new CENP-A/H4 heterodimer. In this way, octameric CENP-A nucleosome formation can be coupled to the pre-existing CENP-A hemisome. Cell cycle analysis suggests that the CENP-A hemisome may be present at centromeres at the time when HJURP is recruited (Bui et al., 2012). However, CENP-A deposition does not absolutely require an existing hemisome as a substrate for new CENP-A deposition as CENP-A nucleosomes can be deposited at initially non-centromeric loci (Barnhart et al., 2011). These two models are not mutually exclusive, and it is possible that both modes of CENP-A deposition occur at centromeres.

Human HJURP contains two HCTD repeat domains within its carboxyl terminus. Duplication of the repeat domain is an evolutionarily recent event, which is restricted to mammals (excluding the egg laying monotremes) (Sanchez-Pulido et al., 2009). The two HCTD repeats are more similar between species than between the two repeats present within a species (Figure 2.S4A). HJURP_C-terminal domains are also found in the myocyte enhancer factor 2 (MEF2) transcription factor family (Potthoff and Olson, 2007). HCTD2 of HJURP is more similar to the HJURP_C-terminal domain of the MEF2 transcription factors (Figure 2.S4). Our data suggest that HCTD2 is sufficient to mediate dimerization and without this domain HJURP does not form a multimeric pre-nucleosomal complex or efficiently deposit new CENP-A.
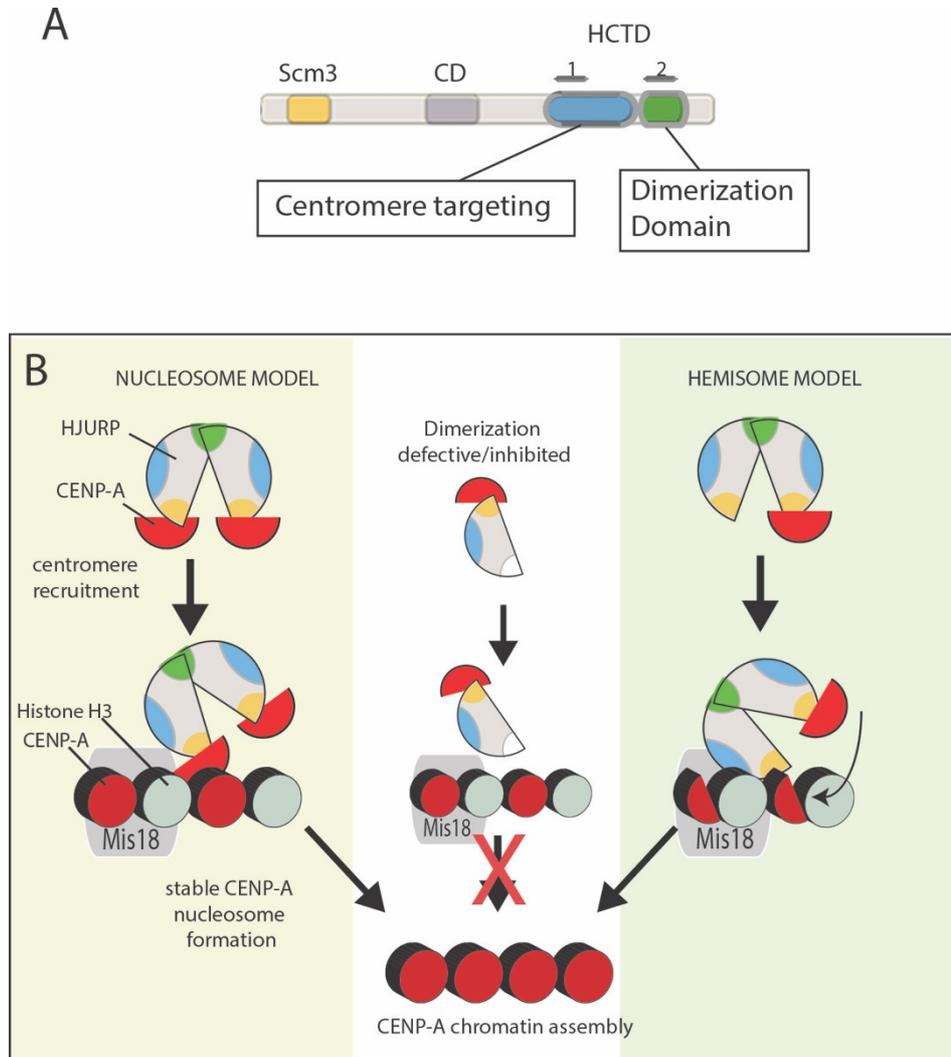
Figure 2.6. Model of HJURP dimerization in CENP-A deposition.

(A) Distinct regions within the carboxyl terminus of HJURP mediate centromere targeting and dimerization. Amino acids 482-end of HJURP include the HCTD2 and are sufficient to mediate HJURP dimerization. The amino acids that contribute to direct centromere targeting of HJURP are between residues 352 and 452 and include HCTD1. (B) Centromeric chromatin contains both CENP-A and histone H3 nucleosomes. During G1, new CENP-A nucleosomes are assembled and may displace existing H3 nucleosomes. HJURP recruitment depends on the Mis18 complex through an unknown process (gray ellipse). Dimerization of HJURP facilitates the assembly of CENP-A nucleosomes. HJURP dimerization may be required to bring two new CENP-A-H4 heterodimers to centromeres in order to form an octameric nucleosome *de novo* (left side). Alternatively, dimerization of HJURP may be required for formation of CENP-A nucleosomes from a pre-existing hemisome and new a CENP-A—histone H4 heterodimer (right side).
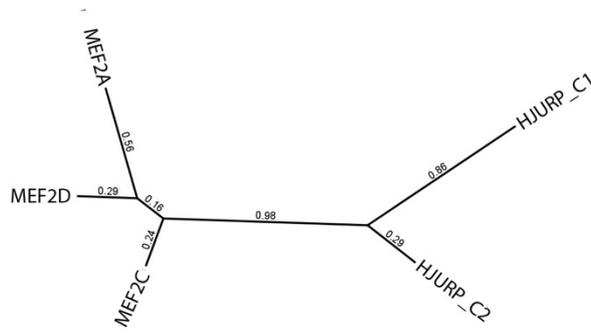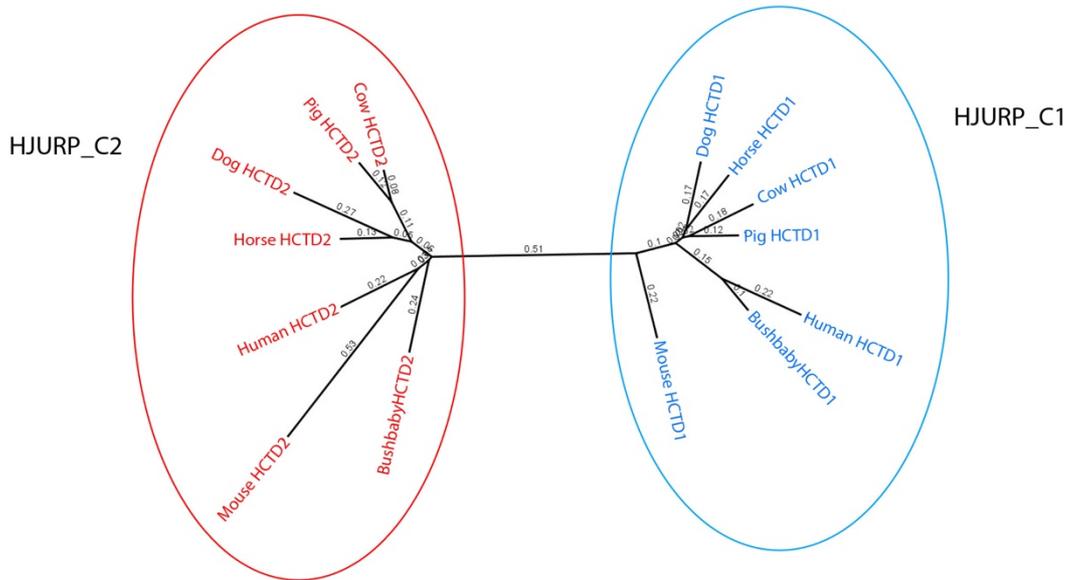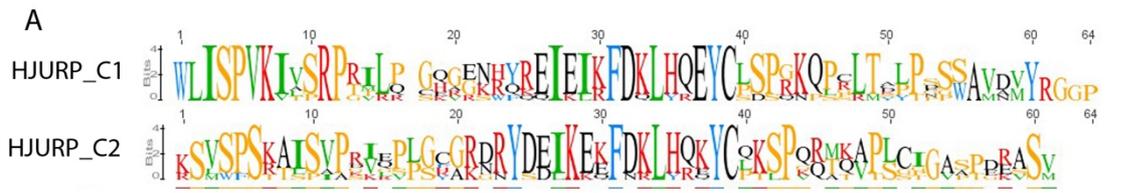
Figure 2.S4.

Figure 2.S4. Sequence alignment of the HJURP_C repeats. (A) Sequences of the HJURP_C-terminal domains (PFAM: 12347) were aligned with HJURP_C-terminal repeats found in the human MEF2 (myoctye enhancer factor 2) transcription factors using ClustalW. Color code corresponds to Clustal color convention. Trees were constructed using the Neighbor joining method. Numbers along each node represent substitutions per site. (B) Sequence conservation logos were generated for each individual repeat based on the sequences from human (NP_060880), mouse (NP_941054), cow (XP_874813), horse (XP_001915875), pig (XP_001925760), dog (XP_003433403) and bushbaby (XP_003803049) using the Geneious 5.0.4 software package. The tree display shows that the HJURP_C-terminal repeats are more similar within repeat number (HCTD1 vs. HCTD2) than within species.

Previous experiments demonstrated that the Scm3 domain was sufficient to deposit CENP-A nucleosomes *in vivo* at non-centromeric loci (Barnhart et al., 2011). These *in vivo* experiments were conducted using a LacI-tagged HJURP-Scm3 domain, and we show here, exogenous dimerization driven by LacI is sufficient for the deposition of CENP-A at endogenous centromeres in the absence of the endogenous HJURP dimerization domain. The number of HJURP binding sites within the endogenous centromeres is unknown. Binding domains for exogenously expressed HJURP may be very dense in the LacO array compared to endogenous centromeres and the high concentration of HJURP may circumvent the need for HJURP dimerization. Alternatively, dimerization of the LacI tag may functionally substitute for the HJURP dimerization domain to facilitate CENP-A deposition at the LacO array, as we demonstrated for CENP-A deposition at endogenous centromeres using $HJUPR^{1-482+DD}$. Previously reported *in vitro* experiments used purified proteins that did not dimerize, and we hypothesize that *in vitro* assembly may occur without the need for dimerization because these assays were conducted with a very high histone/HJURP protein to DNA ratios, which would favor the frequent interaction of HJURP—CENP-A heterotrimers and DNA.

Restricting CENP-A deposition to the centromere depends on the specific recruitment of HJURP. The Mis18 complex is required for HJURP recruitment to centromeres (Barnhart et al., 2011; Moree et al., 2011), although we know very little about how this process occurs. Our data suggest that the direct recognition of the CENP-A nucleosome or hemisome through the HJURP Scm3 domain is not the mechanism by which HJURP is recruited to centromeres, since the presence of the Scm3 domain is not sufficient to recruit HJURP to centromeres (Figure 2.1A,B). Instead we have demonstrated targeting of HJURP to centromeres depends on amino acids 348 to 482 of HJURP, a region that includes the HCTD1 domain (Figure 2.1E,F). The direct recruitment of HJURP to centromeres, presumably through Mis18, is therefore independent of the HJURP dimerization.

Histone chaperone dimerization may also be involved in the assembly of general chromatin. Histone chaperones involved in canonical histone H3 delivery and deposition also interact with a histone H3—H4 heterodimer at a one-to-one stoichiometric ratio similar to HJURP. These include ASF1, Vps75, NAP1, and NASP (Campos et al., 2010; English et al., 2006; Natsume et al., 2007; Su et al., 2011). Several of these chaperones bring two histone H3 heterodimers into a single complex through dimerization of the chaperone. For example, Vps75 forms an α-β earmuff structure which contains a dimerization domain within the amino terminus (Tang et al., 2008). Moreover, Vsp75 and ASF1 form a complex with the Rtt109 histone acetyltransferase and facilitate the acetylation of prenucleosomal histone H3 on Lysine 56 (Driscoll et al., 2007; Han et al., 2007; Schneider et al., 2006; Su et al., 2011).

The ability of centromeric nucleosome assembly factors to dimerize may be a conserved mechanism from yeast. HJURP and its yeast ortholog Scm3 share homology within the CENP-A binding domain (Sanchez-Pulido et al., 2009). Human HJURP is significantly larger than Scm3 and the dimerization domain of HJURP, which we identified here, is not found in Scm3 proteins. Despite the lack of conservation, the dimerization of yeast Scm3 proteins may also occur. Mizuguchi and colleagues (2007) observed that in high-salt conditions the Scm3-Cse4-histone H4 complex formed a hexamer with a 1:1:1 stoichiometry. A subsequent study by Cho and Harrison (2011) using physiologically salt concentrations observed a stable trimer. Since the heterotetramerization of Cse4 and histone H4 is precluded by Scm3 binding, the hexamer formation observed under high salt conditions may be formed by an interaction between Scm3 proteins. Dimerization of human HJURP requires the HCTD2 domain, which is absent from yeast Scm3. Therefore, yeast Scm3 proteins may also dimerize similar to HJURP, albeit through a distinct mechanism. Our study demonstrates dimerization of human HJURP is required for the stable deposition of CENP-A nucleosomes at centromeres and provides a mechanism by which octameric CENP-A

nucleosomes may be formed at the centromere from heterdimeric subunits, a mechanism that may also apply to canonical nucleosome formation.

Materials and Methods

DNA Constructs. All constructs were generated by PCR amplification using Vent polymerase, digested by restriction enzyme and ligated into the indicated plasmid. Primers used for PCR amplifications, restriction enzymes used for cloning and parent vectors are listed in Supplementary table 1. GFP-tagged plasmids were constructed using the pIC113 plasmid (Cheeseman and Desai, 2005). LacI-mCherry fusions were constructed as previously published (Barnhart et al., 2011). GFP-HJURP1-482-LacI construct was created by a two-step cloning approach. The HJURP 1-482 amino acid fragment lacking a stop codon was PCR amplified (See supplemental table 1 for primers) and cloned into pIC113 vector (Cheeseman and Desai, 2005) using Not1 and XhoI restriction sites. LacI was amplified from a vector provided by T. Misteli and was cloned into the Xho1 and KpnI sites of pIC113 containing HJURP1-482.

Transfection. Cells were cultured under standard conditions. DNA and siRNA transfections cells were seeded onto 6-well plates at a density of 2.25×10^5 (Hela and HEK293) or 4.5×10^5 cells/well 24 hours prior to transfection. DNA transfection was conducted using Effectene (Qiagen) with 0.4μg of plasmid DNA per well. siRNA rescue experiments were conducted by treating cells with siRNA 24 hours after plating using RNAiMAX (Invitrogen). For each condition either 10nM of HJURP 3'UTR siRNA (5'GAGAUAACCUCGAGUUCUUUU 3') (Dharmacon) or GAPD control siRNA (Invitrogen). Following 24 hours of siRNA treatment cell s were transfected using EFfectene (Qiagen) as indicated. Immunoblots were conducted using previously established protocols. Antibodies used: anti-HJURP 1:5000 (Bethyl Inc.), anti-GFP 1:1000 (Covance), HA 1.1-1:1000, Anti-tubulin (AA2) 1:100, anti-MBP-1:1000, H2B1:2000 (Millipore).

Indirect Immunofluorescence and SNAP labeling. Cells were plated to poly-lysine coated glass coverslips prior to transfection. Following transfection cells were pre-extracted with

0.1% Triton-X in PBS, fixed with 4% formaldehyde and quenched with 100mM Tris, pH 7.5. Cells were blocked in 0.3% Triton-X in PBS, 2% BSA, 2% FBS for 1.5 hr at room temperature and incubated with primary antibody for 1.5 hr. Anti-CENP-T (Barnhart et al., 2011) was used in 1:3000 dilution and detected using fluorescently conjugated secondary antibodies (Cy3 or Cy5, Jackson Immuno Inc.). DNA was stained with 0.2g/ml DAPI in PBS and coverslips were mounted in Prolong Gold (Invitrogen).

A stable cell line expressing SNAP-tagged CENP-A (Jansen et al., 2007) was treated with siRNA for 24 hours prior to transfection of the HJURP rescue constructs. After an additional 24 hours the pre-assembled CENP-A was blocked with 10 μM O^6-BG (BG-block; Covalys) for 30 min at 37°C followed by a PBS wash and three washes with DMEM over 30 minutes. Cells were incubated in DMEM for 24 hours and labeled with 2 μM TMR-*Star* (Covalys) in complete growth medium for 60 min at 37°C. Labeling was followed by one wash each wtih PBS, and DMEM and incubated for 30 minutes, and washed with PBS prior to fixation.

Images were collected using a 100x oil-immersion Olympus objective lens on a DeltaVision microscope (Applied Precision Inc.) using a Photometrics CoolSNAP HQ[2] camera and Softwrox acquisition software. Images were deconvolved and presented as maximum stacked images. Within siRNA experiments, GFP and TMR-star images for presentation and analysis were collected with identical exposure times and were scaled equally. Integrated intensities were measured from raw images using ImageJ. Intensities of GFP and TMR-star at centromeres were measured using a consistent set area for each experiment. Centromeres were identified based on the presence of the centromere marker (CENP-T). All quantitation of centromere recruitment was restricted to transfected cells by selecting only GFP-expressing cells. GFP intensities were averaged and background corrected using local background correction (Howell et al., 2000). TMR-star intensities were

background corrected using an average background calculated from three non-centromeric sites within the nucleus. G1-phase cells were identified by the presence of a mid-body, apparent by DIC optics.

Immunoprecipitation. Cells were lysed 24 hours post-transfection in RIPA buffer (150 mM NaCl, 1% NP40, 0.3% deoxycolate, 0.15% SDS, 50 mM Tris HCl pH 7.5, 1mM EDTA, 10% glycerol, Protease Inhibitors (Roche), 200 M NaV, 0.5 mM PMSF, 5 mM NaF, 50 mM -mercaptoethanol, 5 M microcystin) on ice for 15 minutes with occasional vortexing. Extracts were DNAseI (1:200, NEB Biolabs) treated where indicated. Lysates were centrifuged at 18000xg for 10 minutes at 4°C and pre-cleared with Protein A agarose (Biorad) for 2 hours at 4°C. Pre-cleared lysates were incubated with anti-GFP antibody (1:1000, Cell Signaling) at 4°C overnight. Antibody-bound complexes were recovered on Protein A Dynabeads (Invitrogen) at room temperature for 45 minutes, washed with RIPA buffer followed three times in PBS including 0.1-0.5% Tween-20. Complexes were eluted by boiling in SDS sample buffer.

Purification of recombinant proteins HIS-HJURP[482-748] and MBP-HJURP[352-482] were expressed in Rosetta BL21 (pLsyS). Cultures were grown in LB at 37°C to OD$_{600nm}$ of 0.6 and induced with 0.5 mM IPTG for 3 hours. Bacteria were lysed by French press and sonication in 25 mM Tris-Cl, pH 7.2, 200 mM NaCl, 20 mM MgCl$_2$, 10% Glycerol, 5 mM -mercaptoethanol, 10 mM -glycerophosphate, 0.2 mM PMSF, 1mM benzamidine, Protease Inhibitors (Roche). Lysis buffer for HIS-HJURP[482-748] purification was suplemented with 10mM imidazole. Lysates were centrifuged at 27,000 rpm for 15 at 4°C and the protiens were purified on Ni-NTA agarose (Qiagen) or Amylose resin (BioLabs). Proteins were eluted with lysis buffer plus 250 mM imidzole (Ni-NTA) or 10mM maltose (amylose).

Sucrose gradient and size exclusion chromatography. Chromatin-free extracts were prepared from transfected HEK293 as described previously (Foltz et al., 2009) except

chromatin isolation buffer (CIB) contained 300mM NaCl. Chromatin-free extracts were applied to a 14 ml 5%–40% sucrose gradient in CIB buffer except digitonin was replaced with 0.05% NP-40. Sucrose gradients were centrifuged at 4°C for 20 hr at 40,000 rpm in SW41 Ti swinging bucket rotor and the gradient was separated into 0.5 ml fractions using a BioComp Gradient Station. Sedimentation standards included Catalase (11.3S), Alcohol Dehydrogenase (7.45), BSA (4.3S), and RNaseA (2S). For size exclusion chromatography, chromatin free extracts were separated on an AKTA-Micro using a Superdex200 PC 3.2/30 column (GE Healthcare) in 3.75 mM Tris, pH 7.5, 300 mM KCl, 0.5 mM EDTA. Recombinant proteins were analyzed on a Superdex200 10/300 GL column (GE Healthcare) in 25mM Tris-Cl,pH 7.2, 200mM NaCl, 20mM MgCl2, 5mM ⅌-mercaptoethanol. Peak fractions of HIS-HJURP[482-748] were applied to a 14 ml 5-20% sucrose gradient and analyzed as described above. Apparent molecular weights were calculated using the Siegel and Monty equation (Siegel and Monty, 1966).

Author Contributions E.Z., M.C.B.D and P.H.L.K. performed experiments. D.R.F, E.Z., P.H.L.K. and M.C.B.D. designed the experiments and analyzed the data. D.R.F and E.Z. wrote the paper.

Chapter 3: Identification of Centromere associated proteins during DNA replication

Abstract

Epigenetic information that controls gene expression and defines chromosome domains with unique function is encoded in posttranslational modifications (PTMs) of histones and the incorporation of histone variants to make unique nucleosomes. Centromeres are unique chromatin domains on each chromosome which identity rely solely on the incorporation of the centromere specific CENP-A histone variant to ensure proper recruitment of the kinetochore and subsequent equal chromosome segregation. Once established, centromeric chromatin needs to be stably retained through multiple cell divisions, what makes centromeres to be an ideal model for studying the concept of epigenetic inheritance thought the cell cycle. Epigenetic memory is transmitted through the passage of replication fork in order for cells to maintain its epigenetic status, however, understanding the mechanisms that facilitate this phenomenon has been of a great interest. In this chapter, we will focus on transmission of the centromeric CENP-A histone variant at the centromeres. Our hypothesis is that CENP-A retention at the centromere during S phase is facilitated by a unique, yet undiscovered mechanism that distinguishes CenpA nucleosomes from bulk chromatin, and this process is absolutely critical to carrying out CENP-A centromere function. Parental nucleosomes might be maintained through similar mechanism. We aimed to identify novel proteins as well as complexes known to be actively involved in DNA replication that are required for reassembly of parental nucleosomes. We also intended to identify novel protein factors associated with the centromeric chromatin. We adapted quantitative mass spectrometry methodology coupled with recently developed BioID technique in which the promiscuous biotin ligase is fused to a target protein to mediate proximity-dependent biotinylation of neighboring proteins. In this chapter I will describe the development and optimization of this experimental strategy. I will compare the efficiency of two methods for proximity based in vivo labeling that we employed: BirA* and APEX. I will

discuss our experimental design, introduce the results of conducted screens. While the next chapter is dedicated solely to the mechanism governing CENP-A inheritance that we discovered using this strategy, in this chapter I will focus on other novel candidate proteins that we identified. I will introduce the results of following validation experiments that we conducted and discuss future directions.

Introduction

Epigenetic information that controls genome functions is encoded in posttranslational modifications of histones or the incorporation of nucleosomes containing histone variants. New DNA synthesis presents a challenge for the inheritance od epigenetic marks as DNA replication requires the disassembly of existing chromatin and reassembly of nucleosomes following passage of the replication fork. However, how exactly epigenetic information transits the replication fork is not well understood. It is suggested that CAF1 complex, which functions as a histone H3 chaperone, play a role in maintenance of parental H3 containing nucleosomes during S phase. Deletion of the CAF1 complex subunits in yeast leads to defects in the silencing of genes present in the heterochromatin regions, transcriptionally repressed telomeric heterochromatin and genes present at the silent HM loci (Enomoto and Berman, 1998; Enomoto et al., 1997; Kaufman et al., 1998; Kaufman et al., 1997; Monson et al., 1997). Recent studies described the MCM2 subunit of the MCM2-7 helicase complex an important player implicated in histone recycling. The mutations within MCM2 region that confers histone binding result in defective chromatin silencing in yeast (Foltman et al., 2013). Recent studies proposed that inheritance of canonical H3.1 parental nucleosomes rely on the activity of MCM2 chaperone. MCM2 can bind H3/H4 tetramers as well as H3/H4 dimers in complex with Asf1, and these interactions were proposed to contribute to the stable inheritance of parental histones across DNA replication (Huang et al., 2015; Richet et al., 2015). The FACT chaperone, that in humans comprises of Spt16 and SSRP1 subunits, is known for its role in chromatin disruption ahead of the RNA polymerase and chromatin reassembly after DNA transcription is completed (Hammond et al., 2017). FACT complex was also demonstrated to interact with multiple components of the replication machinery (including MCM2, MCM4, Polα and RPA1), to travel with the replication machinery, and was proposed to be required for replisome progression (Alabert et al., 2014; Foltman et al.,

2013; Gambus et al., 2006; Kurat et al., 2017; Tan et al., 2006; VanDemark et al., 2006; Wittmeyer et al., 1999; Zhou and Wang, 2004). FACT has a capacity to interact with H2A-H2B histone dimers as well as H3-H4 tetramers, and it was also found to interact with MCM2 through parental histones that have been released from chromatin solubilized with benzonase treatment (Belotserkovskaya et al., 2003; Foltman et al., 2013; Orphanides et al., 1999; Tsunaka et al., 2016).

CENP-A present within centromeric chromatin acts as the epigenetic mark required for specification of functional centromeres as centromeres in humans are formed independently of DNA sequence. In contrast to the canonical H3.1 histone variant, CENP-A deposition is independent of DNA replication, and occurs in early G1(Dunleavy et al., 2009; Foltz et al., 2009; Jansen et al., 2007). The mechanism of CENP-A deposition depends on CENP-A specific histone chaperone HJURP that recognizes CENP-A-CATD domain (CENP-A targeting domain) and binds CENP-A/H4 in prenucleosomal form to facilitate its deposition at the centromere (Bernad et al., 2011; Dunleavy et al., 2009; Foltz et al., 2009). CENP-A incorporation also requires the activity of the Mis18 complex comprising on Mis18α, Mis18β and Mis18BP1 subunits (Fujita et al., 2007; Kim et al., 2012). The Mis18 complex localizes to the centromeric chromatin independently of and prior CENP-A prenucleosomal complex. The presence of Mis18 proteins is required for HJURP recruitment that is facilitated by a direct interaction with the HJURP centromere targeting domain within the HCTD1(Barnhart et al., 2011; Nardi et al., 2016; Wang et al., 2014).

New DNA synthesis presents a challenge for the inheritance od epigenetic marks such as CENP-A containing nucleosomes. Previous experiments using SNAP-tag labelling assays elegantly demonstrated that during this processes existing CENP-A nucleosomes are completely retained at the centromere (Bodor et al., 2014; Jansen et al., 2007). Remarkably, assembled CENP-A containing nucleosomes are equally partitioned to sister chromatids

while cells undergo DNA replication and upon deposition, CENP-A nucleosomes are stably inherited throughout multiple generations (Bodor et al., 2014; Jansen et al., 2007; Ross et al., 2016). These reports suggest that existing CENP-A is specifically reassembled onto centromeric DNA following DNA synthesis; however, the mechanism that regulates CENP-A maintenance is currently unknown and there is no evidence in the literature regarding the contribution of canonical chaperones in CENP-A inheritance during S-phase.

The presence of the CATD domain within CENP-A nucleosomes was proposed to be sufficient for governing its stability and retention at the centromere throughout multiple generations (Bodor 2014). The siRNA based downregulation experiments suggest however, that the stability of already assembled CENP-A nucleosomes is independent of HJURP, which specifically binds the CATD domain (Bodor 2014). Recent study proposed that canonical H3.1 parental nucleosomes require the activity of MCM2 together with Asf1 chaperones in order to be stably retained across S phase. Since CENP-A can directly interact with MCM2 and the CENP-A chaperone- HJURP coimmunoprecipitates with MCM2 it was proposed that analogues retention mechanism applies to CENP-A as well, where HJURP-MCM2 co-chaperone complex might be required for CENP-A inheritance during DNA replication. The exact mechanism facilitating CENP-A inheritance throughout DNA replication remains vastly unknown and lie in wait to be explored.

We set out to identify mechanisms governing the inheritance of epigenetic information, such as CENP-A containing nucleosomes, during DNA replication. Based on the existing literature we hypothesized that there is an unknown mechanism that recognizes and selectively forms a transient interaction with CENP-A containing nucleosomes to separate them from bulk chromatin and to govern its maintenance at the centromere. The retention of canonical H3.1 containing nucleosomes may be similarly regulated in order to facilitate the inheritance of PTMs and ensure inheritance of epigenetic memory. We hypothesize that

proteins actively involved in regulation of DNA replication may also be involved in maintenance of parental nucleosomes.

Results

Development of a biotin-ligase mediated proximity labeling approach to identify proteins associated with parental nucleosomes

In order to identify the mechanism that governs the maintenance of parental nucleosomes while cells undergo DNA replication we employed a proximity based *in vivo* labeling assays coupled with quantitative Mass spectrometry methodology. We adopted the recently developed BioID and APEX techniques in which the promiscuous biotin ligase (BirA* or APEX) is fused to a target protein. These enzymes mediate the proximity-dependent biotinylation of lysine residues exposed on the surface of proteins in close association with the fusion protein (Fig. 3.1A) (Rhee et al., 2013; Roux et al., 2012).

Previous experiments identified CENP-A binding proteins using affinity purification strategies (Dunleavy et al., 2009; Foltz et al., 2009); however, these approaches, due to harsh purification conditions, require highly stable interactions and therefore are challenging for detecting transient interactions. We expect the interactions mediating the passage of CENP-A nucleosomes across the replication fork to be transient. In our strategy, the BirA* or APEX enzyme mediates a covalent biotin attachment to lysine residues of stable and transiently associated proteins. Biotinylated proteins are then purified under denaturing conditions using streptavidin-beads and analyzed by mass spectrometry (MS) (Fig. 3.1 A). We reasoned that the biotinylation approach will allow us to capture proteins that are essential for the inheritance of parental nucleosomes during DNA replication but also transiently interact with CENP-A or H3 histones and could not be detected in previous studies employing tandem affinity purifications strategies (Dunleavy et al., 2009; Foltz et al., 2009).
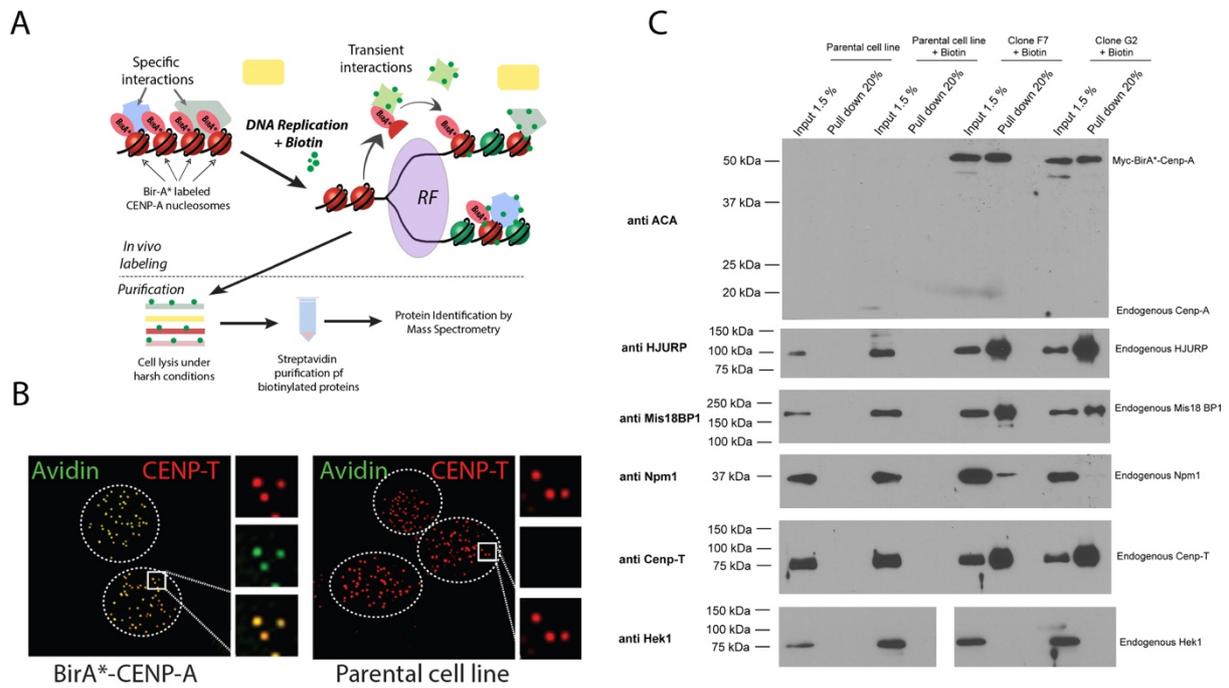
Figure 3.1 Development of a biotin-ligase mediated proximity labeling approach to identify proteins associated with parental nucleosomes

(A) Schematic representation of the experimental approach employing BirA* mediated biotinylation of proteins in close proximity. (B) Representative IF images of parental cells or cells expressing Myc-BirA*-CENP-A. CENP-T is shown in red, biotinylated proteins are shown in green. (C) Streptavidin purification of biotinylated proteins from indicated cells lines analyzed by western blot with use of indicated antibodies. Cells were incubated with medium supplemented with or without biotin for 24 hours.

To identify proteins associated with CENP-A during DNA replication the biotin ligase protein tag was fused to CENP-A. We generated stable HeLa cell lines expressing Myc-BirA*-CENP-A and performed a pilot experiment to test for the efficiency of CENP-A mediated biotinylation *in vivo*. When these cells are cultured in media supplied with biotin, proteins that are in close proximity or interact with BirA*-CENP-A will be biotinylated (Fig 3.1A). Myc-BirA*-CENP-A or parental cells were treated with biotin for 6 hours and subsequently analyzed by immunofluorescence microscopy (Fig 3.1 B). In the CENP-A expressing cell line we observe Myc-BirA*-CENP-A mediated biotinylation of proteins localized to centromeric chromatin (Fig 3.1 B, left panel), as assessed by labeling with fluorescently conjugated avidin. This demonstrates the highly localized and specific labeling achieved using Myc-BirA*-CENP-A.

In order to validate the BioID approach we isolated biotinylated proteins using magnetic streptavidin conjugated beads from cells treated with biotin for 24 hours. When we analyzed purified proteins by immunoblot we were able to identify several proteins known to be associated with CENP-A including HJURP, Mis18BP1 and CENP-T. As expected, we did not observe biotinylation of the outer kinetochore protein Ndc80, which is localized to centromeres but not in close proximity to the CENP-A nucleosome. Furthermore, we did not detect biotinylation of tested proteins in parental cell line cultured in biotin containing medium.

We also wanted to analyze the localization profile of Myc-BirA*-CENP-A fusion protein in our cells lines, however, the staining using Myc antibody resulted in high background signal in both parental and Myc-BirA*-CENP-A expressing cells. We therefore generated new constructs where CENP-A or histone H3.1 were fused to the BirA* biotin ligase and stable cell lines were generated expressing the fusion proteins (Fig. 1B). Biotin addition to the culture medium was used to induce CENP-A or H3.1 mediated labeling.

Biotinylated proteins were visualized by Cy3-conjugated streptavidin and analyzed by Western blot (Fig. 1C,D). The CENP-A–BirA*-HA cellular localization as well as its biotinylation profile co-localize with centromere marker CENP-T, indicating that we can specifically biotinylate proteins associated with centromeric chromatin. The H3.1-BirA*-HA localizes to the bulk chromatin and mediates biotinylation of protein factors associated with general chromatin (Fig.1B,C). We isolated proteins biotinylated by either CENP-A or H3.1 from randomly cycling cells using streptavidin purification. By immunoblot, we identified factors known to be closely associated with CENP-A and H3.1 histone including HJURP or Asf1a, respectively (Fig. 1D). We also detected histone H2B in the pull-down fractions, suggesting that we can induce biotinylation mediated by nucleosomal CENP-A-BirA*HA and H3.1-BirA*-HA (Fig. 1D).

Cells were cultured in media supplemented with biotin for 6 hours and the localization pattern of CENP-A and H3.1 fused to BirA*-HA as well as the biotinylation profile mediated by these histone variants were analyzed using anti HA antibody and Cy3-conjugated avidin, respectively (3.2 A,B). As expected, the CENP-A-BirA*-HA fusion protein localized the centromeres and showed efficient biotinylation of centromere associated proteins. The H3.1-BirA*-HA fusion localized to the bulk chromatin and showed the biotinylation of chromatin within the entire nuclei consistent with the role of histone H3.1 as a major histone H3 variant in human cells. The BirA*-HA fusion protein was robustly overexpressed and highly stable (Fig 3A-C) and this cell line showed very high levels of biotinylation within nuclei. Importantly fusing BirA*-HA to CENP-A was sufficient to restrict the localization specifically to the centromeric chromatin. As expected the biotinylation pattern mediated by either CENP-A or histone H3.1 demonstrated high specificity. When we performed Western blot analysis of purified biotinylated proteins we could detect CENP-A, HJURP, and CENP-I among CENP-A specific interactions, and Asf1α among H3.1 specific interactions.

Surprisingly, although the mutations introduced within the BirA* enzyme are thought to prevent it self-association, we find the Bi-A*-HA fusion protein did biotinylated itself (Fig 3.2 C).

We then experimentally tested the duration of the biotin treatment. The goal of this optimization was to select a duration of the treatment where the signal to noise ratio was the highest. Cells expressing CENP-A-BirA*-HA were treated with biotin for 0, 1, 2, 4, 6 or 9 hours, and were then analyzed by IF with the use of cy-3 conjugated avidin and CENP-T antibody. The signal to noise ratio was calculated by dividing the background corrected intensity of avidin corresponding to centromeres by the signal corresponding to the chromatin region outside of the centromeric domains. We observed that 1 hour of biotin treatment was sufficient to induce the CENP-A mediated biotinylation of centromere associated proteins. Biotin treatment for 6 hours results in the highest signal to noise ratio, and the biotin treatment for longer that 6 hours resulted in increased levels of biotinylation corresponding to non-centromeric chromatin. Based on this experiment we concluded that the 6 hours treatment with the biotin is optimal for conducting further experiments.
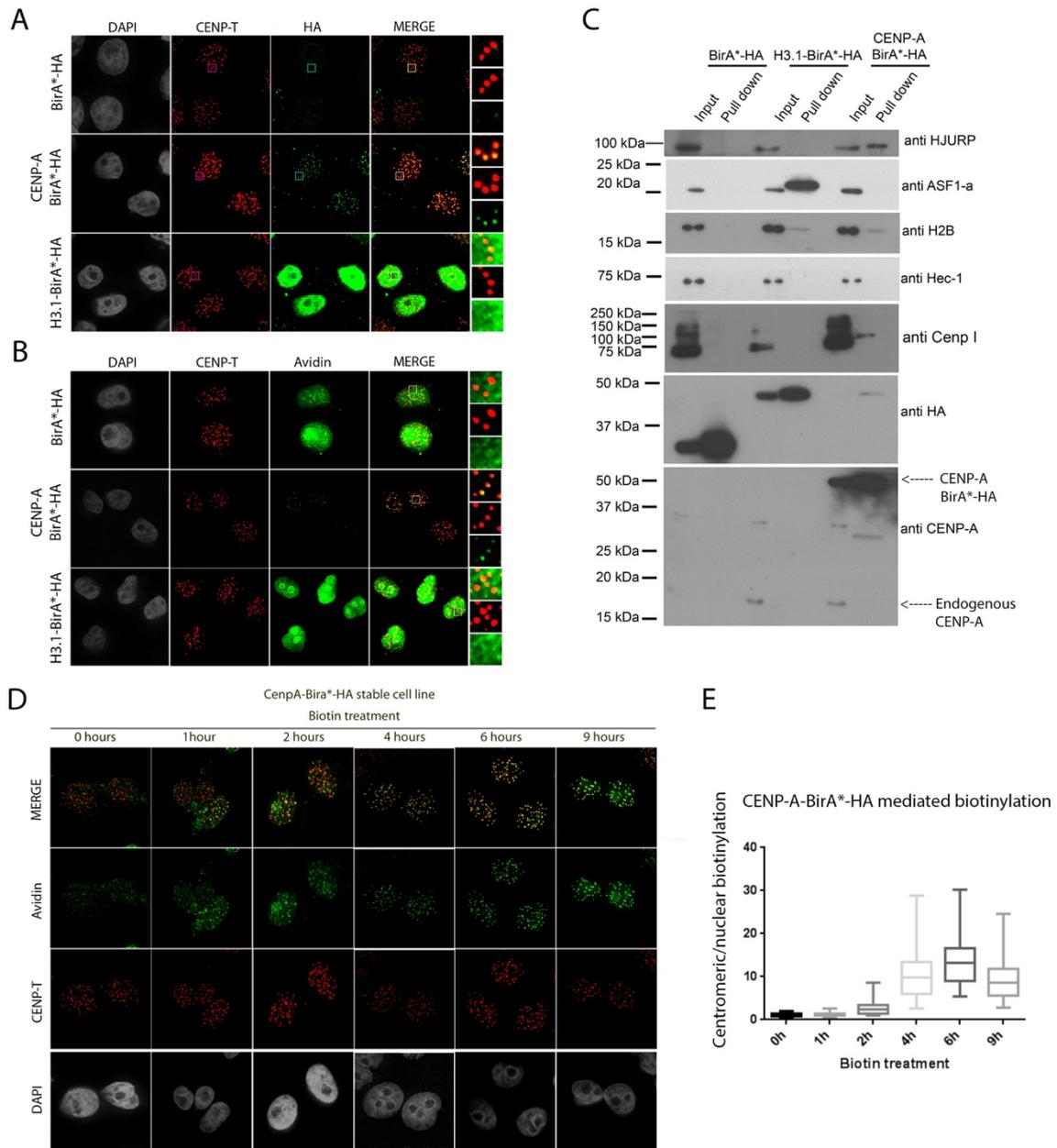
Figure 3.2. Optimization of BirA* mediated in vivo labelling assay

Figure 3.2. Optimization of BirA* mediated in vivo labelling assay

(A) (B) Representative images of cells stably expressing indicated proteins fused to the BirA* ligase and HA tag. Cells were incubated with medium supplemented with or without biotin for 5 hours. DNA is visualized by DAPI staining, CENP-T is shown in red, BirA*-HA fusion proteins (A) or biotinylated proteins (B) are shown in green. (C) Streptavidin purification of biotinylated proteins from indicated cells lines analyzed by western blot with use of indicated antibodies. Cells were incubated with medium supplemented with for 24 hours. (D) The representative IF images demonstrating the efficiency of biotinylation at indicated time points. Cells were incubated with medium supplemented with or without biotin for indicated time. DNA is visualized by DAPI staining, CENP-T is shown in red, biotinylated proteins are shown in green. (E) Quantification of D.

We also wanted to compare the efficiency of biotinylation mediated by BirA* and APEX enzymes. We therefore generated a set of cell lines expressing APEX fusion proteins including: CENP-A-APEX-HA, H3.1-APEX-HA and APEX-HA (Fig 3.3 A). In the CENP-A expressing cell line we observe efficient localization of this fusion protein to the to centromeric chromatin, as assessed by staining with anti HA antibody and centromere marker CENP-T. The H3.1-APEX-HA localized bulk chromatin, and the APEX-HA enzyme alone did not show prominent localization profile.

The APEX mediated biotinylation was previously reported to be highly efficient and occur as rapidly as within 1 minute (Rhee et al., 2013). APEX enzyme utilizes the biotin-phenol substrate and mediates generation of biotin-phenoxyl radicals in the presence of hydrogen peroxide that leads to biotinylation of proteins in close proximity (less than 20nm) (Rhee et al., 2013). We generated the biotin phenol substrate with the help of Dr Michael Hilinski, and Dr Conor Pierce form the Chemistry Department at UVA. The synthesized product was analyzed by NMR and confirmed to be biotin-phenol (Fig 3.3 B). The CENP-A-APEX-HA expressing cells were subjected to treatment with biotin phenol and hydrogen peroxide at 500 μm, 1mM, 1.5 mM or 2 mM concentration for 1 or 4 hours, and subsequently analyzed by IF using cy-3 conjugated Avidin and CENP-T antibody (Fig 3.3 C). The biotinylation profile mediated by CENP-A-APEX-HA fusion showed very weak signal even in cells treated with substrate and hydrogen peroxide for as long as 1 hour, with a significant number of centromeres lacking any detectable biotinylation signal. The prolonged treatment with the labelling reagent for up to 4 hours, although increased the efficiency of biotinylation, but also resulted in high toxicity and cell death. We also tested the efficiency of APEX mediated biotinylation by western blot.

We concluded that APEX mediated biotinylation was not as efficient as compared to BirA* mediated biotinylation. Based on these observations we decided to pursue our future studies using BirA* mediated proximity labeling strategy.
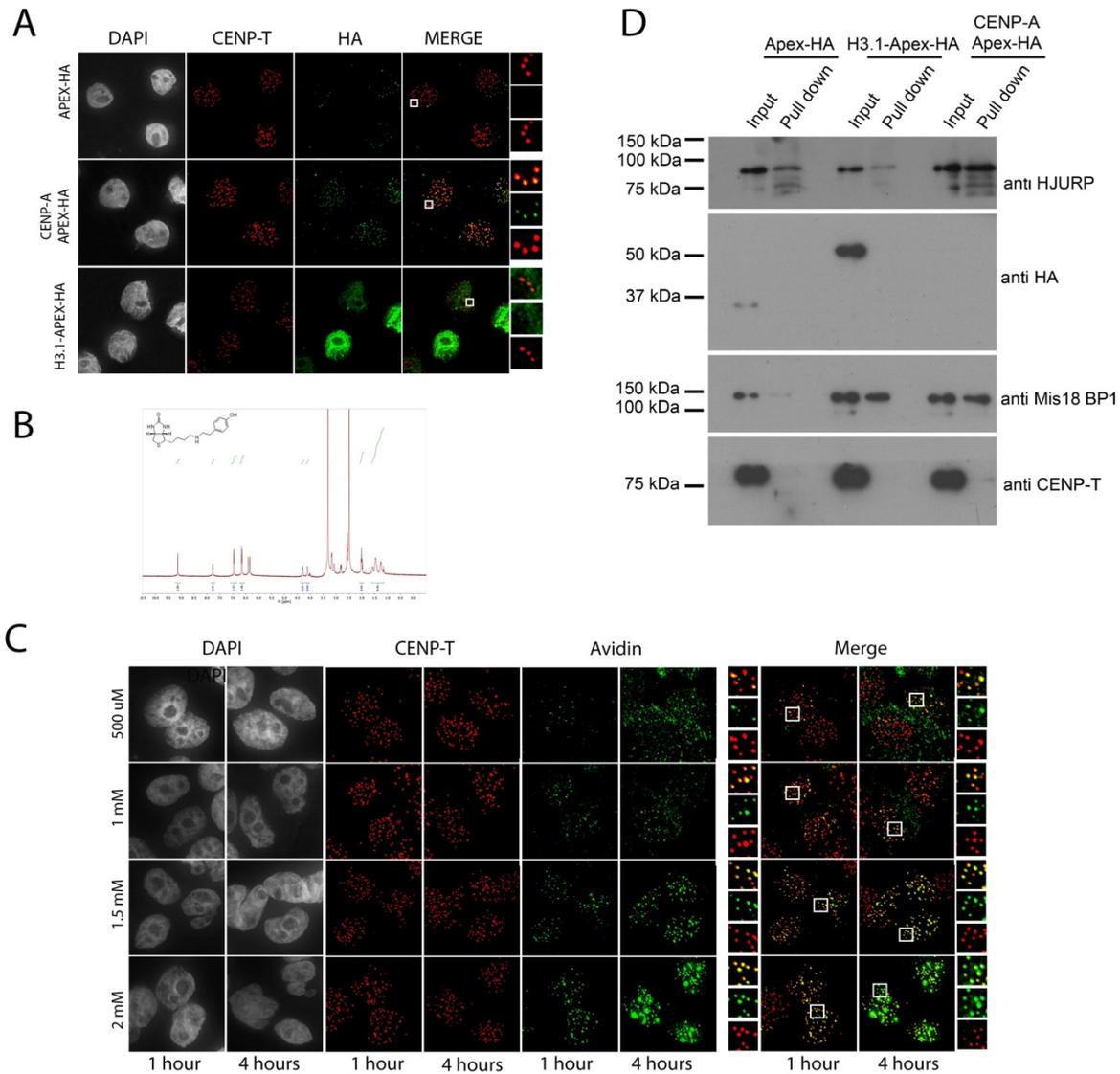
Figure 3.3 Optimization of APEX mediated in vivo labelling assay.

(A) Representative images of cells stably expressing indicated proteins fused to the APEX* ligase and HA tag. DNA is visualized by DAPI staining, CENP-T is shown in red, APEX-HA fusion proteins are shown in green. (B) Proton NMR spectrum of synthesized biotin-phenol substrate. (C) Representative images of cells incubated with medium supplemented with biotin-phenol and hydrogen peroxide at indicated times and concentrations. DNA is visualized by DAPI staining, CENP-T is shown in red, biotinylated proteins are shown in green. (D) Streptavidin purification of biotinylated proteins from indicated cells lines analyzed by western blot with use of indicated antibodies. Cells were incubated with medium supplemented with biotin phenol substrate for 60 min.

Identification of novel proteins associated with centromeres during S phase

In order to identify novel proteins involved in maintenance of CENP-A and H3.1 containing nucleosomes during S phase we designed an unbiased approach where we used the described above BirA* mediated proximity based *in vivo* labelling coupled with the quantitative Mass Spectrometry strategy. We expected that the differences between H3.1 and CENP-A may not be all-or-none. Therefore, we carried out our experiments using SILAC (stable isotope labeling by amino acids in cell culture) to allow us to make quantitative comparisons between CENP-A-BirA* and controls. We adopted cells expressing CENP-A-BirA*-HA to heavy medium where Lysine and Arginine amino acids were replaced by their heavy analogs. We then designed our screen where we included a comparison of CENP-A-BirA*-HA biotinylation profile in S phase with H3.1-BirA*-HA in S phase to identify CENP-A specific interactions. A comparison of CENP-A-BirA*-HA biotinylation profile in S phase with Parental cell line in S phase was included to identify endogenously biotinylated proteins. We also designed a comparison of CENP-A-BirA*-HA mediated biotinylation in S phase with a biotinylation of CENP-A-BirA*-HA in asynchronous cells to identify CENP-A specific interactions during S phase (Table 3.1)

Heavy medium	Phase	Vs.	Light medium	Phase	Ratio	MG132
CENP-A-BirA*	S	-->	Parental	S	1:1	No
CENP-A-BirA*	S	-->	H3.1-BirA*	S	1:1	No
CENP-A-BirA*	S	-->	CENP-A-BirA*	AS	1:1	No
CENP-A-BirA*	S	-->	Parental	S	1:1	Yes
CENP-A-BirA*	S	-->	H3.1-BirA*	S	1:1	Yes

Table 3.1. Conditions to be tested by SILAC.

HeLa cells stably expressing either BirA*-tagged CENP-A or H3.1 were synchronized by double thymidine block and release (Fig. 3.3A). Following the second thymidine block cells were released and biotin labeling was induced for 6 hours while cells underwent DNA replication (Fig 3.4 A). We also included a subset of conditions where cells were treated with the MG132 proteasome inhibitor because we were concerned that by the time the cells are harvested, the biotinylated proteins might be subjected to proteasome mediated degradation (Table 3.1). We compered the cell cycle profiles of cells used for the Mass Spec sample preparation. Both cell lines expressing CENP-A-BirA*-HA or H3.1-BirA responded to thymidine block and release as efficiently as the parental HeLa T-rex cells (Fig 3.4 B). We also made an observation that the MG132 addition, however, resulted in S-phase progression inhibition. Heavy (H) and light (L) cell lysates derived from the cell types and cell cycle points listed in Table 3.1 were mixed at 1:1 ratio. Biotinylated proteins were isolated from mixed cell lysates containing heavy and light components using streptavidin conjugated beads and subjected to trypsin digestion and identified by MudPIT-MS (Fig 3.4 A).

We used the Heavy/Light (H/L) ratio, calculated based on the abundance of unique peptides corresponding to each protein detected in the sample, as criteria to assess the specificity of the interactions. We used the Protein Discoverer software for the H/L score assignment and the R software for data analysis. The H/L score was subjected to natural log transformation and plotted (Fig 3.3C). As CENP-A-Bir*A-HA cell line was grown in heavy medium and released into S phase in the presence of biotin (Table 3.1) therefore all identified proteins with the score H/L >1 were classified as CENP-A specific and S-phase specific interactions. In the case where H3.1-BirA* is used, proteins with the score H/L < 1 were classified as H3.1 specific interactions. Proteins with the score H/L=1 or close to 1 were considered to be endogenously biotinylated or interactions common for both H3.1 and

CENP-A. CENP-A and H3.1 mediated patterns of biotinylation were compared with each other in order to trace the differences in mechanisms regulating stable inheritance of these two histone variants.
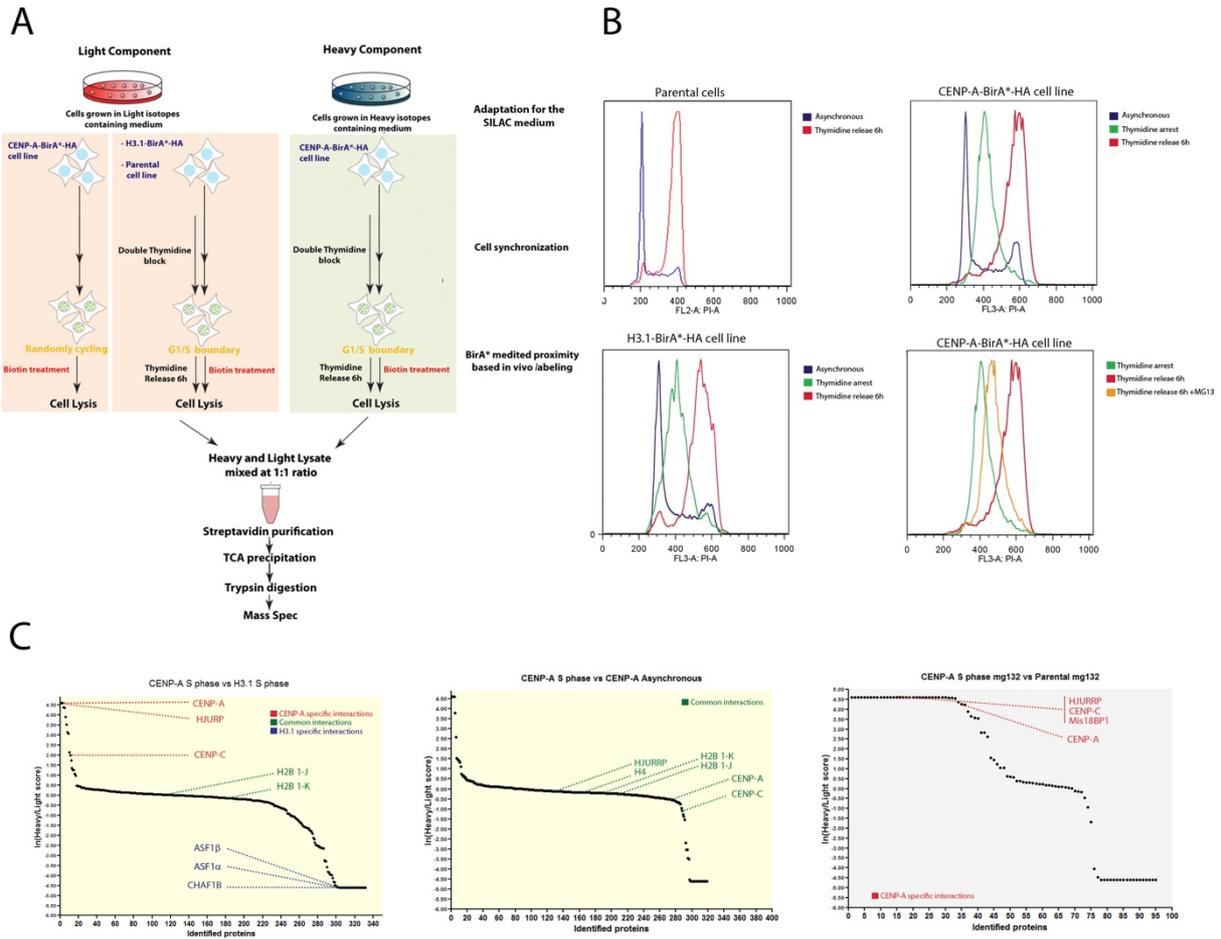
Figure 3.4 BirA* mediated proximity labeling approach to identify proteins associated with parental nucleosomes: experimental design.

Figure 3.4 BirA* mediated proximity labeling approach to identify proteins associated with parental nucleosomes: experimental design.

(A) Schematic representation of experimental designed to identify CENP-A associated proteins at indicated cell cycle stages. (B) DNA content analysis of cells used as an input for purification. (C) Graphs demonstrating proteins identified in indicated samples and their log2 transformed H/L score. The sample in A was derived from lysates prepared from cells expressing either CENP-A-BirA*-HA (heavy component) or H3.1-BirA*-HA (light component) released into S phase in the presence of biotin. The sample in B was derived from lysates prepared from cells expressing CENP-A-BirA*-HA either released into S phase (heavy component) or randomly cycling (light component) and treated with biotin. The sample in C was derived from lysates prepared from cells expressing CENP-A-BirA*-HA (heavy component) or parental cell line (light component) released into S phase in the presence of biotin and treated with MG132. The heavy and light components were mixed at 1:1 ratio. The Heavy/Light ratio (H/L) was calculated for each protein detected in the sample. Arrowheads are pointing proteins identified in the screen and known to be physically associated with CENP-A or histone H3.1, or both histone proteins.

We detected both prenucleosomal and chromatin associated complexes known to interact with CENP-A or H3.1 histone proteins in our experiments. Among CENP-A specific biotinylated proteins we identified: HJURP, CENP-A, CENP-C, CENP-T, Mis18BP1, NPM1, canonical histone proteins (H4, H2B, H1) (Appendix 1). In H3.1-BirA*-HA derived sample we detected several known H3.1 chaperones involved in H3.1 deposition during S-phase, including ASF1α, ASF1β, HAT, CAF1-p60, CAF1-p150, NASP and canonical histone proteins (H3, H4, H2B, H1) (Appendix 1). Therefore, we were able to identify the majority of proteins previously shown by our lab and others to associate with CENP-A, that were identified using more laborious affinity purification approaches.

We generated a list of proteins identified in each sample and rank ordered the identified proteins based on their assigned H/L score. Our quantitative mass spec analysis identified common components of the CENP-A and H3.1 nucleosomes, such as histone H2A, and H1 with a H/L ratio approximately 1. As expected endogenously biotinylated proteins that were present in both heavy and light components also showed H/L score close to 1 (including: Pyruvate carboxylase, Acetyl-CoA carboxylase, Propionyl-CoA carboxylase alpha chain and Acetyl-CoA carboxylase 2) (Appendix 1). These data suggest the lysates used as an input for purification were indeed mixed at 1:1 ratio. We detected proteins constitutively associated with centromeric chromatin such as CENP-C and CENP-T among CENP-A specific interactions (H/L>1). To our surprise, we have also identified components of the G1 coupled CENP-A deposition machinery, including HJURP and Mis18BP1 among CENP-A interactions during S phase (Fig 3.4 C). This led us to hypothesize that these factors are associated with CENP-A during DNA replication and perhaps play a dual role in CENP-A deposition in G1 and the retention in S-phase. We designed a set of follow up experiments to test this hypothesis and described the results in **Chapter 4** "**The inheritance of CENP-A nucleosomes during DNA replication".**

We have identified a subset of novel S phase specific CENP-A associated candidate proteins in our screen. We selected candidates based on their H/L score from across all performed experiments (the H/L score above 1.75). This criterion allowed us to select 68 candidate genes for further analysis (including HJURP and Mis18BP1) (Appendix 2). It is important to note that in contrast to HJURP, which was present among CENP-A specific interactions during S phase in every experiment that we performed, the selected novel candidate proteins were not consistently detected in all performed experiments (Appendix 2). The chapter 4 is dedicated to the role of HJURP in CENP-A inheritance across DNA replication, where we demonstrate that the mechanism of S phase coupled retention of the CENP-A nucleosomes requires CENP-A specific deposition machinery including HJURP together with the activity of MCM2. However, we also decided to follow up on selected novel candidate proteins because we reasoned that perhaps these proteins are transiently bound to CENP-A and these potential binding partners might play a role in centromeric inheritance or specification as well.

We performed the GO Enrichment Analysis on candidate interactions and found that these proteins are categorized to be involved in : CENP-A containing nucleosome assembly, kinetochore assembly, chromatin silencing at rDNA, histone H3-K27 trimethylation, nucleosome positioning, DNA replication-dependent nucleosome assembly, regulation of gene silencing by miRNA, SRP-dependent co-translational protein targeting to membrane, regulation of megakaryocyte differentiation, nuclear-transcribed mRNA catabolic process, nonsense-mediated decay, translational initiation, regulation of chromosome organization, cell division and cell cycle biological processes.

We then set out to validate the candidates and subject them to secondary screens including Fluorescence Microscopy and Chromatin Immunoprecipitation We searched the orfeome cDNA library for the presence of the genes of interest selected based on our screen.

We proceeded with 11 novel candidates found in the cDNA library including: KLHL35, SND1, PAK1IP1, TUFM, GNL2, SERPINH1, KAISO, ASF1β, ATPIF1, U2AF2. One of the candidate proteins U2AF2, has a binding partner-U2AF1, that was also present in our screen however, the H/L score for this protein was below 1.75. We decided to include this protein in our validation experiments due to its close association with U2AF2 protein. We also included HJURP, Mis18BP1 and its binding partner Mis18α proteins in our panel of interactions subjected to the validation experiments. We cloned those genes into a GFP containing vector, and generated stable cell lines expressing GFP-fused target proteins to assess the association with centromeric DNA in S phase. The cells expressing GFP-fusion proteins were synchronized with double thymidine block and release. Cells were released into S phase in the presence of MG132, and 3 hours post release fixed and subjected to ChIP using GFP antibody and IgG control. The eluted DNA was subsequently analyzed by qPCR using primer set specifically amplifying α-satellite DNA of chromosome 7. The data was normalized to IgG control within each cell line and plotted. Cells expressing YFP-CENP-A were used as a positive control. This experiment confirmed that although at different levels, 13 of all selected candidates showed significant association at the α-satellite centromeric DNA during S phase, among which 10 proteins are novel centromere associated proteins (Figure 3.5 A).

We also assessed the localization profile of the candidate proteins using fluorescence microscopy. Cells expressing GFP-fusion proteins were synchronized with double thymidine block and release into S phase for 3 hours in the presence of MG132. Cells were then fixed and analyzed by immunofluorescence microscopy using CENP-T as a centromere marker. We were able to detect the localization of 3 of proteins at the centromere in cells undergoing S phase, although the quantification of that data showed very low efficiency of localization,

which perhaps could be explained by dynamic nature of identified candidate proteins with centromeres in S phase (Figure 3.5 B, C).
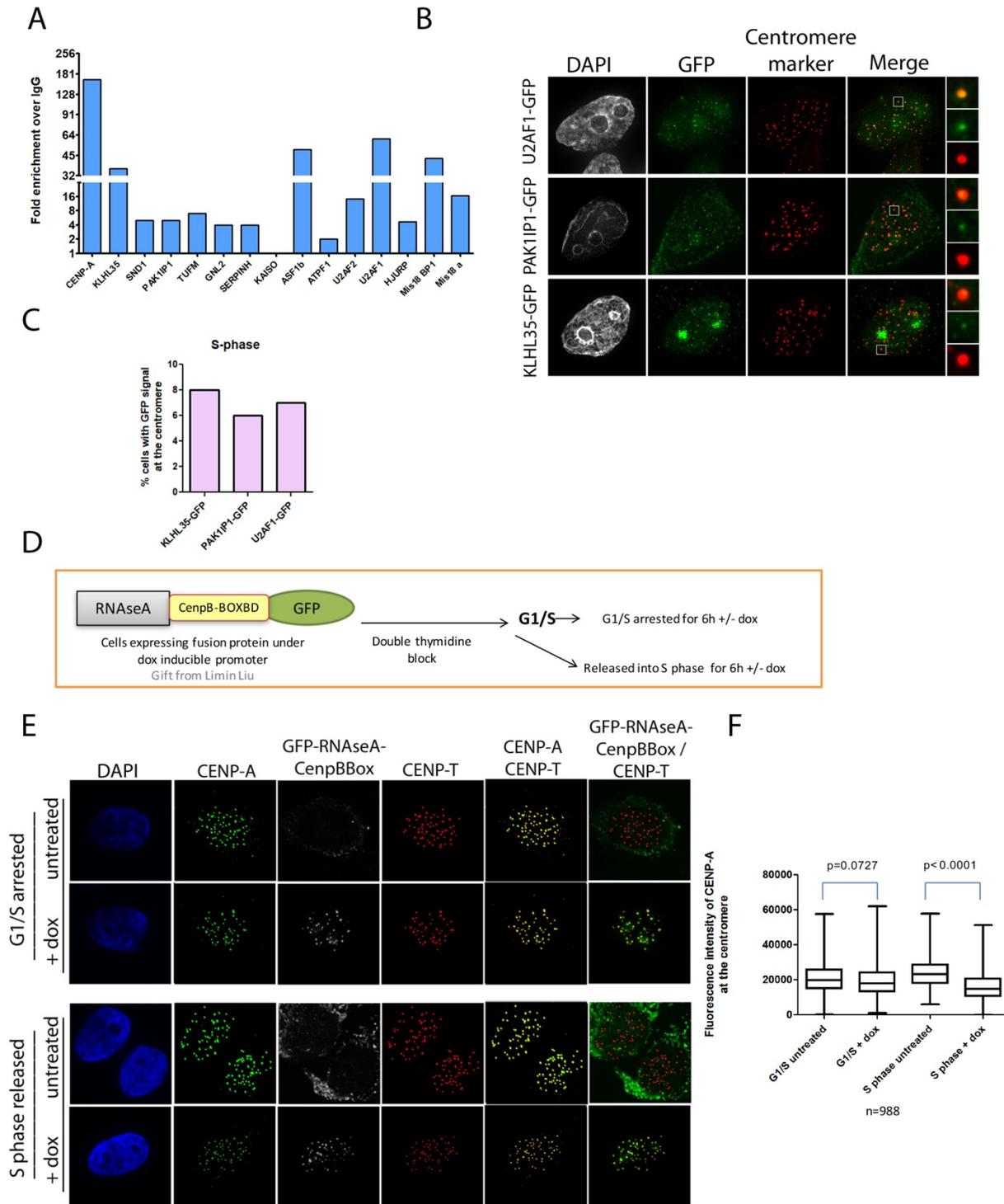
Figure 3.5 Validation of identified novel centromere associated proteins

Figure 3.5 Validation of identified novel centromere associated proteins

(A) Chromatin Immunoprecipitation from cells expressing indicated GFP-tagged proteins in S phase. ChIP was performed using anti GFP antibody and normal rabbit IgG. RT-PCR was performed using primers against α-satellite DNA specific for chromosome 7. (B) Localization profile of synchronized cell lines overexpressing indicated GFP fusion proteins and undergoing S phase. Cells were stained with DAPI to visualize DNA, and a centromere marker. (C) Quantification of B. (D) Schematic representation of the experiment shown in E. (E) Representative IF images of cells expressing RNAseA-CenpBBox-GFP fusion protein at indicated cell cycle stages. Cells were stained with DAPI to visualize DNA, CENP-T is shown in green and CENP-A is shown in red. (F) Quantification of E.

The role of RNA in CENP-A maintenance

One of the proteins that we identified as CENP-A interaction-U2AF1 is linked to spliceosome machinery. This suggests that perhaps U2AF1 protein is recruited to the centromere through the interaction with α-satellite transcripts and potentially contributes to the α-satellite RNA processing. This prompted us to ask whether the presence, if any, of α-satellite transcripts within centromeric chromatin during DNA replication is involved in the maintenance of centromeric nucleosomes. We therefore tethered RNAseA (a ribonuclease that catalyzes the degradation of RNA) specifically to the centromeric chromatin and analyzed the levels of endogenous CENP-A. We used a stable cell line expressing RNAseA-CenpBBox-GFP fusion protein under control of doxycycline inducible promoter. Cells were synchronized with double thymidine block and release and subsequently threated +/- dox while arrested at the G1/S boundary or released into S phase for 6 hours (Fig 3.5 D, E). Cells were fixed and we analyzed the levels of endogenous CENP-A in control cells (-dox) in cells where RNAseA-CenpBBox-GFP was present at the centromere (+dox) (Fig 3.5 E). We observed statistically significant decrease of CENP-A in cells with RNAseA-CenpBBox-GFP present at the centromere during S phase (Fig 3.5 F). This experiment demonstrated that the recruitment of RNAseA to the centromere negatively affected endogenous CENP-A levels in cells undergoing DNA replication, while did not have significant effects on CENP-A levels in thymidine arrested cells.

Conclusions

The BirA* mediated proximity based labelling approach coupled with quantitative mass spectrometry allowed us to detect CENP-A specific interactions as well as identify novel proteins potentially interacting with CENP-A nucleosomes during DNA replication. It is important to mention that, as we learned during conducting these experiments, the BirA* strategy, although very efficient in labelling proteins in close proximity, has some limitations. This includes the requirement for the neighboring protein to have lysine residues exposed on its surface in order to be labelled. This perhaps contributes to inconsistencies in the identified proteins detected in our experiments that might interact with CENP-A or other centromere bound proteins transiently and indirectly. We realized as well that endogenously biotinylated proteins were highly abundant in samples that we subjected to Mass spectrometry analysis, and concluded that the SILAC approach was very beneficial in our experimental design.

We identified 68 novel proteins potentially associated with centromeres during DNA replication. Our validation experiments performed on selected 11 candidates further support these results, as 90% of tested proteins did indeed show association with α-satellite DNA in cells undergoing DNA replication. It remains to be tested whether these proteins can interact with centromeric DNA at different cell cycle stages as well or whether they are recruited specifically in S phase.

We found that 3 proteins out of the selected candidates we followed up on, KLHL35, PAK1IP1 and U2AF1 showed centromeric localization profile as assessed by immunofluorescence microscopy upon stable overexpression. These observations require further validation; however, as the percent of cells with apparent centromeric localization was very low for all 3 cell lines. It is possible that the levels of expression of tested proteins is low and therefore difficult to visualize. These proteins may also interact

with the centromere in a dynamic nature during DNA replication which may be difficult to capture using standard immunofluorescence. It is also possible that KLHL35, PAK1IP1 and U2AF1 proteins can localize to the centromere more efficiently at different cell cycle stages and the localization profile of these proteins awaits further evaluation. One strategy that can improve the localization analysis of candidate proteins would be to fuse those proteins to the BirA* enzyme. This approach will allow to analyze the chemical trace of the localization in the form of biotin mark painted by the fusion protein, and test if the biotinylation signal co-localize with a centromere marker.

These observations however, if confirmed and further explored by additional experiments, might lead to characterization of pathways essential for governing centromere function. The U2AF1 protein that we found associated with α-satellite DNA belongs to the splicing factor SR family of genes which play significant roles in constitutive pre-mRNA splicing and alternative splicing. U2AF1 is a 35 kDa protein that forms a heterodimer with U2AF2 (65kDa) and associates with the 3′ splice site (3′ss) during mRNA splicing. The presence of centromeric transcripts was reported in multiple model organisms and this phenomenon was proposed to have instrumental roles in centromere assembly and function in yeast and humans (Bergmann et al., 2011; Bouzinba-Segard et al., 2006; Carone et al., 2009; Chan et al., 2012; Choi et al., 2011; Eymery et al., 2009; Hall et al., 2012; May et al., 2005; Ohkuni and Kitagawa, 2011; Quenet and Dalal, 2014a; Stimpson and Sullivan, 2010; Topp et al., 2004; Wong et al., 2007). Furthermore, the active form of the RNA Polymerase II was found at the endogenous centromeres in humans during mitosis and early G1, and centromeric transcripts were also proposed to be implicated in new CENP-A deposition pathway (Chan et al., 2012; Quenet and Dalal, 2014a). Biochemically purified RNA associated with prenucleosomal CENP-A/HJURP complex led to identification of 1.3 kb RNA product that co-localizes with a-satellite DNA and CENP-A

and hybridizes to centromeric α-satellite probes, suggesting it originated from a-satellite transcripts. The efforts to map the identified RNA were inconclusive as it could not be assigned to any specific locus within human genome suggesting that presumably it was a product of RNA processing. Downregulation of this RNA species *in vivo* results in defect in CENP-A and HJURP recruitment to the centromere demonstrating that the RNA component partially encoded within α-satellite DNA is an important new factor that play a regulatory role in CENP-A deposition pathway (Quenet and Dalal, 2014a). Yet it remains elusive whether RNA processing machinery plays a role in this process and what is the origin of RNA required for CENP-A incorporation. The presence of U2AF1 at the centromere that we detected in our experiments suggests that perhaps α-satellite transcripts are indeed processed and perhaps this process is facilitated by the by spliceosome machinery.

To further explore this hypothesis, it needs to be tested whether U2AF1 co-purifies with transcripts derived form a-satellite DNA. This can be achieved by mapping *in vivo* RNA-U2AF2 interactions using RNA-IP assay. The U2AF35 is essential for viability of both yeast and higher eukaryotes therefore it will be difficult to assess whether depletion of this gene will result in defects in centromere function. However, perhaps the auxin-based degron system for the rapid depletion that specifically targets AID fused fortein for proteasome mediated degradation could be employed to assess the functional role of U2AF1 localization at the centormere. This approach can be utilized provided that the AID fusion will not interfere with the U2AF1 function. Also it is essential to test whether other spliceosome components or other RNA processing factors are also recruited to the centromeric DNA, and if so what is the cell cylce timing when this occurs.

We tested the role of RNA in CENP-A inheritance by tethering RNAseA to the centromere during DNA replication and in thymidine arrested cells. This experiment demonstrates that

the presence of RNAseA at the centromere specifically during S phase results in significant reduction of CENP-A from centromeric chromatin. This data suggests that there is a RNA component present at the centromere that is required for centromere maintenance. This raises a question whether the RNA exists in DNA/RNA hybrids with the centromeric chromatin or forms a complex with centromere associated proteins. Targeting the RNAseH, non-sequence-specific endonuclease that catalyze the cleavage of RNA in RNA/DNA substrate, to the centromere and analyzing the levels of CENP-A and other centromere bound proteins should be employed in a similar experimental setup to address these questions. Importantly, we cannot ignore the possibility that the presence of RNAseA at the centromere results in loss of CENP-A due to replication stress, as short RNA primers are required for the replication to occur. It is apparent however, that the involvement of RNA in the context of centromere specification, maintenance and function is a poorly understood process and awaits more detailed studies.

We also identified two more novel proteins that localize to the centromeric chromatin: PAK1IP1 and KLHL35. PAK1IP1 protein was previously demonstrated to negatively regulate the PAK1 kinase, which has been shown to regulate various cellular activities, including cell proliferation, cell survival, mitosis and transcription. The overexpression of PAK1IP1 protein leads to inhibition of cell proliferation by inducing p53-dependent G1 cell-cycle arrest. While the role of PAK1IP1 is known, the function of KLHL35 protein is poorly understood. The sequence analysis of KLHL35 shows the presence of BTB domain, BACK domain, 6 x Kelch domains. The identification of the BACK domain in BTB and Kelch proteins suggest an important function for this domain with a possible role in substrate orientation in Cullin3-based E3 ligase complexes. It still remains to be tested what is the functional role of accumulation of these two proteins at the centromere during DNA replication. It remains to be tested whether these proteins can

interact with centromeric DNA at different cell cycle stages. This can be tested by depletion of these proteins and assessing the levels of CENP-A at the centromere. Importantly, it will also be essential to test whether the depletion of candidate proteins will not result in general inhibition of DNA replication or cell cycle progression and defects in fidelity of chromosome segregation. It is also important to test whether these proteins are required for proper centromere formation, and recruitment of CCAN components.

Materials and Methods

BioID and Mass Spectrometry. Affinity purification of biotinylated proteins was performed as previously described in (Roux et al., 2012). In brief, cells were incubated in DMEM 10% FBS media supplemented with 50 µM biotin for 6 hours (25x stock solution of biotin was prepared in DMEM at 1.25 mM concentration). Cells were washed three times with PBS and harvested. Cell pellets were lysed at 25°C in 1 ml lysis buffer (50 mM Tris, pH 7.4, 500 mM NaCl, 0.4% SDS, 5 mM EDTA, 1 mM DTT, and Complete protease inhibitor [Roche]). Cell lysates were sonicated, subsequently supplemented with Triton X-100 to 2% final concentration, and subjected to another round of sonication. Subsequently cell lysates were diluted with an equal volume of cold (4°C) 50 mM Tris (pH 7.4) and subjected to additional sonication. Cell lysates were then spun down at 10,000 relative centrifugal force for 5 mins at 4°C. Supernatants were collected and protein concentration was measured using the BCA assay. For the Mass Spectrometry analysis, heavy and light components were mixed at 1:1 ratio. Supernatants were incubated with Streptavidin Magnetic Beads (BioLabs) for overnight. Beads were then collected and washed twice with 1 ml buffer containing 2% SDS in dH$_2$O. The beads were then washed once with buffer containing 0.1% deoxycholate, 1% Triton X-100, 500 mM NaCl, 1 mM EDTA, 50 mM Hepes, pH 7.5, and once with buffer containing 250 mM LiCl, 0.5% NP-40, 0.5% deoxycholate, 1 mM EDTA, and 10 mM Tris, pH 8.1. Beads were then washed twice with buffer containing 50 mM Tris, pH 7.4, and 50 mM NaCl. Biotinylated proteins were then eluted from the beads with 100 µl of 2x Laemmli SDS-sample buffer saturated with biotin at 98°C. For the Mass Spectrometry analysis, the protein concentration was measured using the BCA assay.

Eluted samples were diluted with water up to 1ml and supplemented with 250 ul 100% TCA (final concentration of TCA 20%) and incubated for overnight at 4°C. Samples were spun down at 16000 rpm for 30 mins and protein pellet was washed with 1 ml of ice cold acetone 5

times. Protein pellets were dried in speed vac. Protein pellet was resuspended in buffer containing 100 mM ammonium bicarbonate (pH 8), 0.1% Rapigest and 10% ACN. Samples were supplemented with DTT at 5mM final concentration, and incubated at room temp for 1 hour. Iodacetamide was added at 12.5 mM final concentration, and samples were incubated in the dark for 1 hour. Proteins were digested with mass spec grade Trypsin (Trysin Gold from Promega). Trypsin was added at 1/20 ratio based on the amount of proteins measured in elutes. Samples were incubated for 15 hours at 37°C with shaking. The digestion was quenched with mass spec grade formic acid at 1% final concentration.

Sample mixtures were digested in 9ul volumes and injected directly onto an Easy Spray nano HPLC column ES801 (Thermo Scientific), packed with PepMap RSLC C18 media (2um, 100A, 50 um x 15 cm). An Easy nano LC (Thermo Scientific) delivered mobile phases A (0.1% formic acid in water) and B (0.1% formic acid in acetonitrile) as a gradient of 2 - 25% B over 60 min and 25 -50% B in 30 min at a flow rate of 250 nL / min. Mas spectra were collected using a Q Exactive Plus Orbitrap (Thermo Scientific) mass spectrometer at a resolution setting of 70,000 (FWHM @ 200 m/z) in full MS mode scanning from 300 - 2000 m/z and performing data-dependent MS/MS acquisition (top 10) with a resolution setting of 17,500 (FWHM @ 200 m/z). LC-MS data were analyzed using Proteome Discoverer software, version 1.4 (Thermo Scientific). MS and MS/MS spectra were searched using the Sequest HT algorithm. Trypsin-generated peptides were identified using a FASTA database of human protein sequences (Unirpot, October 2015) as well as a decoy database with scrambled sequences. False positives were filtered using a false discovery rate of 1%. All peptides were quantified in a label-free manner using the MS1 extracted ion chromatogram (XIC) peak area with a tolerance of 2 ppm. Ratios of [heavy: light] peptides were calculated and averaged for each identified protein in order to perform SILAC relative quantitation of proteins.

Chromatin immunoprecipitation. ChIP was performed as previously described in (Mayo et al., 2003). In brief, cells were synchronized using double thymidine block and released into S phase for 3 hours. Cells were then cross-linked on the plate by adding formaldehyde at a final concentration of 1% for 10 mins at 37°C.   Glycine was added at a final concentration of 0.125 M to stop the cross-linking reaction. Cells were then washed twice with PBS, harvested and stored in -80 °C. Cell pellets were thawed on ice and lysed in 1ml of Farnham Lysis Buffer (5mM PIPES (KOH) pH 8.0, 85 mM KCl, 0.5 % NP40, and protease inhibitors [Roche]). Cell lysates were incubated for 10 mins on ice with shaking. Nuclei were pelleted at 800g for 2 minutes and resuspended in 250 ul of Lysis Buffer (1% SDS, 10mM EDTA, 50 mM Tris-HCl pH 8.0, protease inhibitors [Roche]). Lysates were incubated on ice for 30 minutes with shaking, and subsequently sonicated. Lysates were spun down at 13000 rpm for 10 minutes; supernatants were collected and measured for the protein concentration. An equal amount of protein per sample was then diluted 10 times with the Dilution Buffer (1.1% TritonX100, 1.2mM EDTA, 16.7mM Tris-HCl pH8.0, 167 mM NaCl, protease inhibitors [Roche]). Lysates were pre cleared with Protein A agarose beads and IgG for 30 minutes at 4°C. Agarose beads were then spun down at 1300 rpm for 2 minutes and the supernatants were supplemented with GFP antibody or rabbit IgG and incubated for 17 hours at 4°C. Protein A agarose beads were added and samples were incubated for 1 hour at 4°C. Agarose beads were then pelleted by spinning down at 1300 rpm for 1 minute at 4°C, and subsequently washed twice with Low sat Wash Buffer (0.1% SDS, 1% Triton X100, 2mM EDTA, 20mMTris HCl pH8.0, protease inhibitors [Roche]). Beads were then washed once with High Salt Wash Buffer (0.1% SDS, 1% Triton X100, 2mMEDTA, 20mM Tris-HCl pH 8.0,500 mM NaCl, protease inhibitors [Roche]), twice with LiCl Wash Buffer (0.25M LiCl, 1% NP40, 1% deoxycholate, 1mM EDTA, 10 mM Tris-HCl pH 8.0), and twice with TE

buffer. Each wash was performed for 5 minutes at RT. Beads were incubated with 75 ul of Elution Buffer (0.1M NaHCO3, 1% SDS) at RT for 15 minutes. Samples were then spun down at 2000 rpm for 2 minutes and eluates were collected. The elution step was repeated; the elution fractions were combined and supplemented with NaCl to a final concentration of 0.3 M following by 17 hours incubation at 65 °C. DNA was purified with PCR purification kit (Qiagen) and stored at -20C. Following ChIP, DNA was quantified by qPCR using standard procedures on a StepOne™ Real-Time PCR System. Primers for qPCR were used as previously described in Ohzeki et al., 2012: Forward: GGCATATGTGCAAGTGGATATAC; Reverse: TATCCACTTGCAGAC.

Cell Culture and Transfection. HeLa T-rex cells and lines derived from this parental cell lines were cultured in a 37 °C incubator in 5% CO_2 in DMEM supplemented with 10% FBS (OPTIMA) and 1% Pen/Strep. Cells for the DNA transfection were seeded onto six-well plate prior to transfection and transfected when they reached 60% of confluency. DNA transfection was conducted with the Lipofectamine 2000 reagent using standard protocol (Thermo Fisher Scientific) with 2ug of plasmid DNA per well.

Cell synchronization. Cells were synchronized with double thymidine block and release. Thymidine was added to culture medium at 20mM for 18 hours. Following the first thymidine treatment cells were washed twice with PBS and released into S phase for 8 hours, and subsequently treated with second thymidine block. Cells were released form the second thymidine arrest as indicated in the text.

DNA content analysis. Cells were synchronized and subsequently harvested using PBS supplemented with 3 mM EDTA. Cells were then washed with PBS and spun down at 1000

rpm for 5 min. Cell pellets were resuspended in 200ul PBS, fixed with 5mls of 70% Ethanol and stored at 4°C. Fixed cells were centrifuged at 1600 rpm for 5 min and washed with PBS + 1% FBS. Cell pellets were then resuspended in fresh PI/RNaseA solution (10ug/ml propidium iodide, 250ug/ml RNase A in PBS + 1% FBS) and incubated at 37°C for 30 minutes. Samples were analyzed for their DNA content using flow cytometry.

Imaging and Quantification. Images were acquired using a × 100 oil-immersion Olympus objective lens on a DeltaVision microscope. Collected images are demonstrated as maximum stacked images. Images were subjected to deconvolution prior stacking. Integrated intensities were derived from raw images subjected to ImageJ (using the CRAQ plugin) and the centromere marker was used as a reference. Quantification data was analyzed in GraphPad Prism software, the statistical significance was assessed using unpaired t-test. The graphs were generated using GraphPad Prism software and displayed percentiles are as indicated in figure legends.

Stable cell lines. All cell lines expressing BirA-HA* fusion proteins were generated using Flp-In™ T-REx™ System.

Stable isotope labeling. SILAC labeling with light and heavy analogs of Lysine and Arginine was performed in DMEM Media for SILAC (Thermo scientific) supplemented with either Arginine- HCl and Lysine- 2HCl or 13C6-Arginine HCl and 13C6-Lysine HCl (13C Molecular), respectively. The medium was supplemented with 10% Dialyzed Fetal Bovine Serum (JR Scientific). Cells were adopted for heavy and light medium for 20 cell divisions.

Indirect immunofluorescence. Cells were pre-extracted with 0.1% Triton-X in PBS for 3 minutes, fixed with 4% formaldehyde for 10 minutes and subsequently quenched with 100mM Tris, pH 7.5 for 5 minutes. Fixed cells were incubated in blocking solution (0.1% Triton-X in PBS, 0.2% BSA, 2% FBS) for 1.5 h at RT, and incubated with indicated primary antibodies for 1.5 h. Anti-CENP-T, anti-CENP-A and anti-HA antibodies were used at 1:5000, 1:1000, and 1:1000 dilution, respectively and detected using fluorophore conjugated secondary antibodies (Cy3, Cy5 or FITC, Jackson Immuno Inc.). Cy3-conjigates streptavidin (Jackson Immuno Inc.) was used at 1:1000 dilution. DNA was visualized with 0.2mg/ml DAPI in PBS and coverslips were mounted in Prolong Gold (Invitrogen).

Synthesis of biotin phenol. Step 1: Synthesis of biotin succinimidyl ester. 500 mg biotin (MW 244.31 g/mol) was mixed with 1.2 equivalents of disuccinimidyl carbonate (256.17 g/mol, SigmaAldrich) and 2.0 equivalents of triethylamine (101.19 g/mol; d = 0.7255 g/mL) in 2 mL DMF. The reagents were stirred at room temperature for 1 hour. Disuccinimidyl carbonate was warmed up to room temperature before opening to avoid condensation. The DMF mixture was diluted into 80 mL PBS buffer, pH 7.4 to minimize hydrolysis of product. Immediately upon addition of PBS, a white precipitate was formed (this was the majority of the product). The solution was transferred to a separation funnel and 1 volume of ethyl acetate was added (the white precipitate was associated with the organic fraction). The aqueous layer was removed, and the organic layer was vacuum filtered. The solid was dried on vacuum, and collected. This process was repeated (adding additional volumes of ethyl acetate, up to 6 times total) until all of the precipitate has been collected. Then the ethyl acetate layers were combined and dried over sodium sulfate. The product was analyzed product by TLC, with the use of 1:4 methanol: ethyl acetate mixture as solvent and stained with DMAC (product Rf is 0.5). The ethyl acetate was then removed under vacuum and a

white solid was obtained, and combine with the white solid obtained by vacuum filtration in the earlier steps. <u>Step 2: Conjugation of biotin succinimidyl ester (biotinNHS) to tyramine</u>. The biotinNHS (MW 341.38 g/mol) was mixed with 1.2 equivalents of tyramine and 2.0 equivalents of triethylamine in DMF. The reagents were stirred at room temperature for 1 hour. The reaction was diluted with 40 volumes of water, then acidified with HCl until pH reached approximately 6. The extract with 40 mL ethyl acetate, was subsequently preformed and repeated three times. The combined organic layers were dried with sodium sulfate and analyzed by TLC, using 1:4 methanol: ethyl acetate mixture as solvent (product Rf is 0.3), and subsequently visualized with UV and stained with DMAC. The solvent was then removed under vacuum, and a yellow/white solid was obtained. The solid was then dissolved in warm (~40 C) methanol or 1:4 ethyl acetate: methanol. The silica column was run with 1 column volume of 100% ethyl acetate, then switched to 1:4 methanol: ethyl acetate. The fractions were collected and analyzed by TLC. The product fractions were then consolidated and characterized by proton and carbon NMR, and subsequently aliquoted as 500 mM (1000x) stocks in DMSO, and sored at 80 C.

Chapter 4: Inheritance of CENP-A nucleosomes during DNA replication

This chapter is based on the manuscript under review:

Abstract

Centromeres are chromosomal loci that define the site of kinetochore formation and ensure faithful chromosome segregation during mitosis. Centromeric identity is epigenetically specified by the incorporation of CENP-A nucleosomes. DNA replication presents a challenge for the inheritance of epigenetic information, including the retention of centromeric identity because CENP-A nucleosomes must be removed for replication fork progression. Despite this challenge, CENP-A nucleosomes are stably retained across S-phase. We used BioID to label and identify proteins transiently associated with CENP-A during DNA replication. Contrary to prior models, we found that pre-existing CENP-A binds to its chaperone HJURP during S-phase, suggesting a novel and surprising role for HJURP in the retention of CENP-A nucleosomes. We observed that HJURP is transiently associated with S-phase centromeres. Using auxin-based degradation to specifically address the role of HJURP during DNA replication, we demonstrate that HJURP binding to CENP-A is required for centromeric nucleosome inheritance. HJURP co-purifies with the MCM2-7 helicase complex and together with the MCM2 subunit binds CENP-A simultaneously. Therefore, pre-existing CENP-A nucleosomes requires a novel S-phase function of the HJURP chaperone, and interaction with MCM2, to ensure the faithful inheritance of centromere identity through DNA replication.

Introduction

Centromeres are unique chromatin domains present on each chromosome that facilitate recruitment of the constitutive centromere-associated network (CCAN) and kinetochore that work together to ensure equal chromosome segregation during mitosis (Amano et al., 2009; Cheeseman and Desai, 2008; Earnshaw et al., 1986; Foltz et al., 2006; Izuta et al., 2006; McKinley and Cheeseman, 2016; Nishihashi et al., 2002; Okada et al., 2006; Saitoh et al., 1992; Sugata et al., 1999). In most eukaryotes, deposition of centromere specific nucleosomes containing the histone H3 variant CENP-A serves as an epigenetic mark critical for centromere specification and inheritance, independent of the underlying DNA sequence (Allshire and Karpen, 2008; Cleveland et al., 2003).

In contrast to the canonical H3.1 histone variant, new CENP-A incorporation is uncoupled from DNA replication and occurs in early G1. The vertebrate Holliday Junction Recognition Protein (HJURP) and its homologs, Scm3 in yeast and CAL1 in *Drosophila,* specifically recognize prenucleosomal CENP-A and facilitate its deposition at the centromere (Barnhart et al., 2011; Bernad et al., 2011; Camahort et al., 2007; Chen et al., 2014; Dechassa et al., 2011; Dunleavy et al., 2009; Erhardt et al., 2008; Foltz et al., 2009; Goshima et al., 2007; Mizuguchi et al., 2007; Pidoux et al., 2009; Shuaib et al., 2010; Stoler et al., 2007; Williams et al., 2009). HJURP is known for its role in early G1 deposition of CENP-A nucleosomes (Dunleavy et al., 2009; Foltz et al., 2009). Centromeric recruitment of HJURP or budding yeast Scm3 in complex with CENP-A/H4 requires the Mis18 complex (Barnhart et al., 2011; Camahort et al., 2007; Fujita et al., 2007; Hayashi et al., 2004; Mizuguchi et al., 2007; Moree et al., 2011; Pidoux et al., 2009; Stoler et al., 2007; Williams et al., 2009). In humans, the Mis18 complex is comprised of Mis18α, Mis18β and Mis18 BP1 subunits (Fujita et al., 2007; Maddox et al., 2007). Mis18 complex directs HJURP to centromeres through physical interaction between HJURP centromere targeting domain and the Mis18α-β

C-terminal coiled-coil domains (Nardi et al., 2016; Wang et al., 2014). The recruitment of CENP-A deposition pathway proteins to centromeres is cell cycle regulated and both HJURP and Mis18 complex are controlled by Cdk activity or Cdk and Plk1 mediated phosphorylation, respectively (McKinley and Cheeseman, 2016; Muller et al., 2014; Pan et al., 2017; Silva et al., 2012; Spiller et al., 2017; Stankovic et al., 2017).

During DNA replication new DNA synthesis requires the disassembly of existing chromatin ahead of the replication fork, and reassembly of evicted histones once new DNA synthesis is completed (Annunziato et al., 1981; McKnight and Miller, 1977; Sogo et al., 1986). Despite this challenge, some proportion of parental H3.1 nucleosomes are inherited across DNA replication to preserve chromatin states, and chromatin landscape regulating gene expression, and therefore govern cellular identity and genome integrity. Although a mechanism of how parental nucleosomes are retained is not fully understood, experiments in yeast and human cells suggested that PCNA in complex with CAF-1 as well as the MCM2 chaperone in complex with ASF1 or FACT are important players in recycling parental nucleosomes at the replication fork (Alabert and Groth, 2012; Burgess and Zhang, 2013; Foltman et al., 2013; Gerard et al., 2006; Huang et al., 2015; Ransom et al., 2010; Shibahara and Stillman, 1999). MCM2-7 helicase complex facilitates DNA unwinding at replication origins during replication initiation, and travels with the replication forks to unwind DNA ahead of the progressing replisome (Bochman and Schwacha, 2009; Boos et al., 2012). MCM2 co-purifies with nucleosomal H3-H4, and directly binds all histone H3 variants, including CENP-A, through its N-terminal histone biding motif (HBD) (Foltman et al., 2013; Groth et al., 2007; Huang et al., 2015; Jasencakova et al., 2010; Richet et al., 2015).

Centromeric nucleosomes are stably inherited across multiple generations suggesting that CENP-A nucleosomes are faithfully retained through DNA replication (Bodor et al., 2013; Falk et al., 2015; Guo et al., 2017; Jansen et al., 2007; Mellone et al., 2011). CENP-A

nucleosomes are dispersed into daughter strands, the gaps resulting from CENP-A dilution are occupied by histone H3.1 and H3.3 nucleosomes (Bodor et al., 2013; Dunleavy et al., 2011; Falk et al., 2015; Jansen et al., 2007; Ross et al., 2016). This suggests that existing CENP-A is specifically reassembled onto centromeric DNA following DNA synthesis; however, the mechanism that regulates CENP-A maintenance is vastly unknown. We hypothesize that CENP-A retention at centromeres during S phase is facilitated by a unique mechanism that specifically distinguishes CENP-A containing nucleosomes from bulk histones and this process is essential for ensuring centromere identity.

Here we employed BioID in an unbiased experimental approach to identify proteins associated with CENP-A during DNA replication. We identify a novel role for HJURP chaperone outside of the G1-phase and show that HJURP plays a role in maintaining centromere identity during DNA replication. Using this approach, we demonstrated that HJURP and Mis18BP1 are strongly associated with existing CENP-A nucleosomes during DNA replication. To delineate a specific role for the protein in S-phase we used Auxin-inducible degron strategy for HJURP depletion and show that loss of HJURP leads to loss of CENP-A retention during DNA replication. We demonstrate that CENP-A also requires interaction with the MCM2 protein for stable inheritance. We show a CENP-A$^{RK->AA}$ mutation that is deficient in MCM2 binding *in vitro* fails to be efficiently retained at the centromere throughout DNA replication. Furthermore, HJURP and MCM2 can bind CENP-A simultaneously. These data demonstrate that the mechanism of S phase coupled retention of the CENP-A nucleosomes requires CENP-A specific deposition machinery including HJURP together with the activity of MCM2.

Results

Identification of proteins associated with CENP-A during DNA replication

CENP-A nucleosomes are dispersed between daughter DNA strands during replication; however, the mechanism that facilitates CENP-A inheritance is largely unknown. We hypothesize that interactions involved in CENP-A retention during S-phase occur transiently; therefore, in order to identify the proteins involved in this process we used the BioID proximity based *in vivo* labelling assay coupled with Mass Spectrometry (MS)(Roux et al., 2012) described in chapter 3. We used the BioID approach to biotinylate and purify proteins specifically and transiently associated with CENP-A during DNA replication, and identified the proteins by mass spectrometry. Cell lines expressing BirA*-fusion proteins were synchronized by double thymidine block and release and the biotinylation profiles mediated by CENP-A-BirA* versus histone H3.1-BirA* during DNA replication were quantitatively compared using SILAC (Fig. 4.1A). Heavy/Light ratios above 1 indicate an enrichment of the protein in the CENP-A-BirA* condition. This allowed us to identify proteins specifically associated with CENP-A during DNA replication relative to general chromatin. Using this strategy, we observed strong enrichment for the known S-phase histone chaperone protein Asf1 ✔, Asf1 ⚙, Chaf1b and NASP with histone H3.1, thus validating our approach (Fig. 4.1 A). In addition, we observed no enrichment for Macro H2A.1, Topo1 or histone H2B, which are common to both CENP-A and histone H3.1 containing chromatin. Surprisingly, our analysis shows that the CENP-A specific chaperone HJURP is associated with CENP-A during DNA replication. Comparisons of CENP-A-BirA* labeled proteins in S-phase arrested cells and randomly cycling cells shows no enrichment, suggesting that CENP-A associates with HJURP in S-phase as well as in G1 (Fig. 4.1 B).
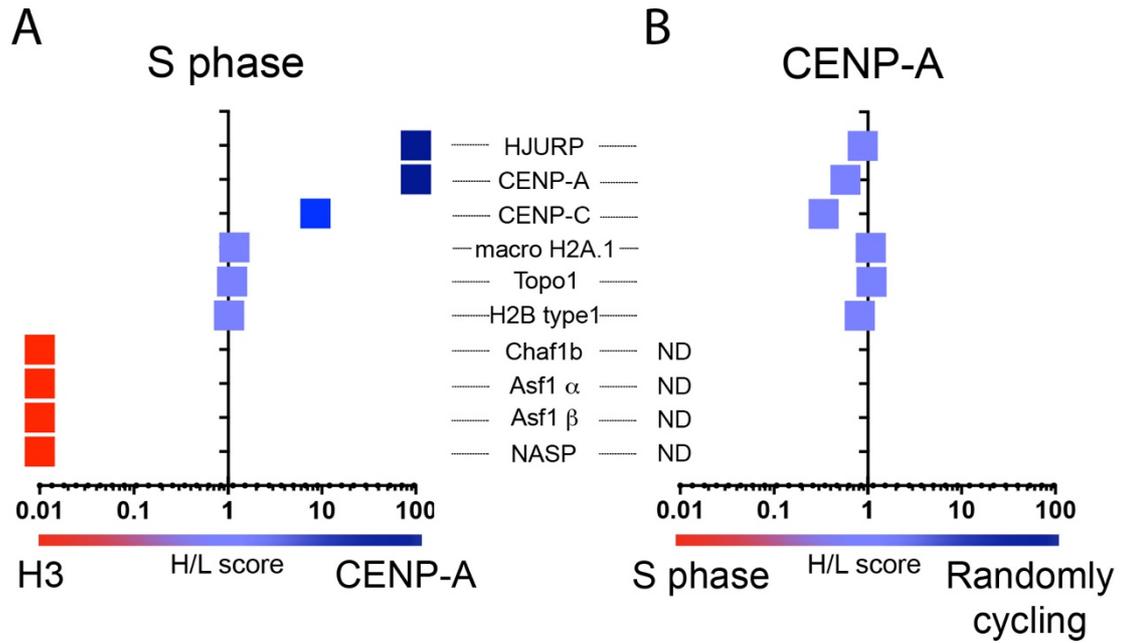
This figure includes data generated with the help of Dr. A.O. Bailey.

Figure 4.1. Labeling of proteins transiently associated with CENP-A and H3.1 nucleosomes.

(A)(B) Graphs demonstrating selected proteins identified in indicated samples and their corresponding H/L scores. The specificity of interaction is demonstrated by the heat map. SILAC comparisons are between CENP-A-BirA*-HA biotinylation profile in cells undergoing S phase (heavy component) versus biotinylation profile of H3.1-BirA-HA* in S phase (light component) (A) or biotinylation profile of CENP-A in asynchronous cells (light component) (B).

HJURP and Mis18BP1 are associated with chromatin assembled CENP-A during S-phase

The association of HJURP with CENP-A during DNA replication suggests that HJURP may be involved in facilitating the retention of pre-existing CENP-A as centromere nucleosomes are disrupted during replication. In order to demonstrate that pre-existing CENP-A nucleosomes already present within chromatin associate with HJURP during DNA replication we generated stable cell lines expressing CENP-A-BirA*-HA or H3.1-BirA*-HA under a Dox inducible promoter. Cells were treated with Dox for 96 hours and then Dox was washed out to shut down CENP-A expression and ensure the sole source of CENP-BirA* was from already incorporated CENP-A nucleosomes. Biotinylation was induced in cells undergoing S-Phase, arrested in early S-phase or asynchronous cell populations by providing biotin for 6 hours (Fig. 4.2 A). We observed that both HJURP and Mis18BP1 were biotinylated by nucleosomal CENP-A in asynchronous cells (Fig. 4.2 B right panel). Consistent with our MS purification above, we detected an interaction of nucleosomal CENP-A with HJURP and Mis18BP1 during early S phase, and the degree of biotinylation increased in cells that underwent DNA replication (Fig. 4.2 B left panel).

To further support the association of HJURP with centromeric chromatin during S-phase we performed ChIP for chromatin associated HJURP. Using CRISPR-Cas9 mediated gene editing we tagged endogenous HJURP with the AID-YFP tag in DLD1 cell line stably expressing the E3 ubiquitin ligase Tir1 (Fig. 4.2 C top panel) (see below). ChIP for endogenously-tagged HJURP-AID-YFP was performed using anti-GFP antibody from asynchronous population, cells blocked in early S-phase or those progressing through S-phase. qPCR analysis was conducted using primers amplifying α-satellite DNA from chromosome 7 (Tsuda et al., 1997). The amount of endogenous HJURP at the centromere increased as cells progressed through S-phase, compared to S-phase blocked cells or

asynchronously dividing population (Fig. 4.2 C). A similar enrichment of HJURP and

Mis18BP1 was observed at chromatin during S-phase in cells stably expressing GFP-HJURP

and GFP-Mis18BP1 created by lentiviral transduction (Fig. 4.S1.C).
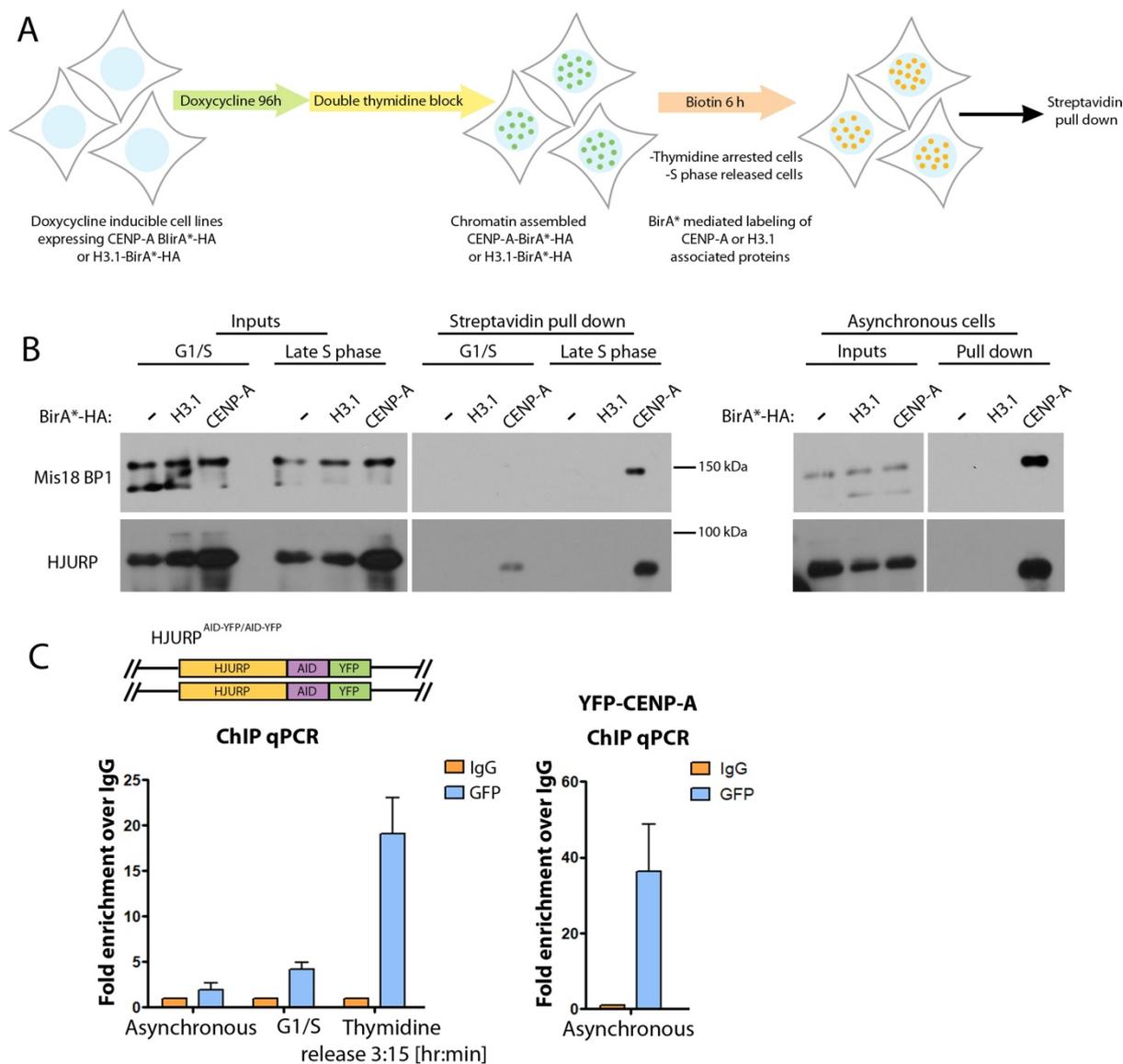
Figure 4.2. CENP-A deposition proteins are associated with centromeres during DNA replication. (A) Schematic representation of the experimental approach in B. (B) Cells expressing BirA* fused proteins under doxycycline inducible promoter were treated as shown in A. Biotinylated proteins were isolated by streptavidin purification following by Western blot analysis with use of indicated antibodies. (C) Schematic representation DLD1-Tir1 cell line where HJURP was endogenously tagged on both alleles with AID-YFP (top). Chromatin Immunoprecipitation from HJURP-AID-YFP cells (bottom left) or YFP-CENP-A cells (bottom right) at indicated time points. ChIP was performed using anti GFP antibody and normal rabbit IgG. RT-PCR was performed using primers against α-satellite DNA specific for chromosome 7. The graph represents an average of two independent experiments.
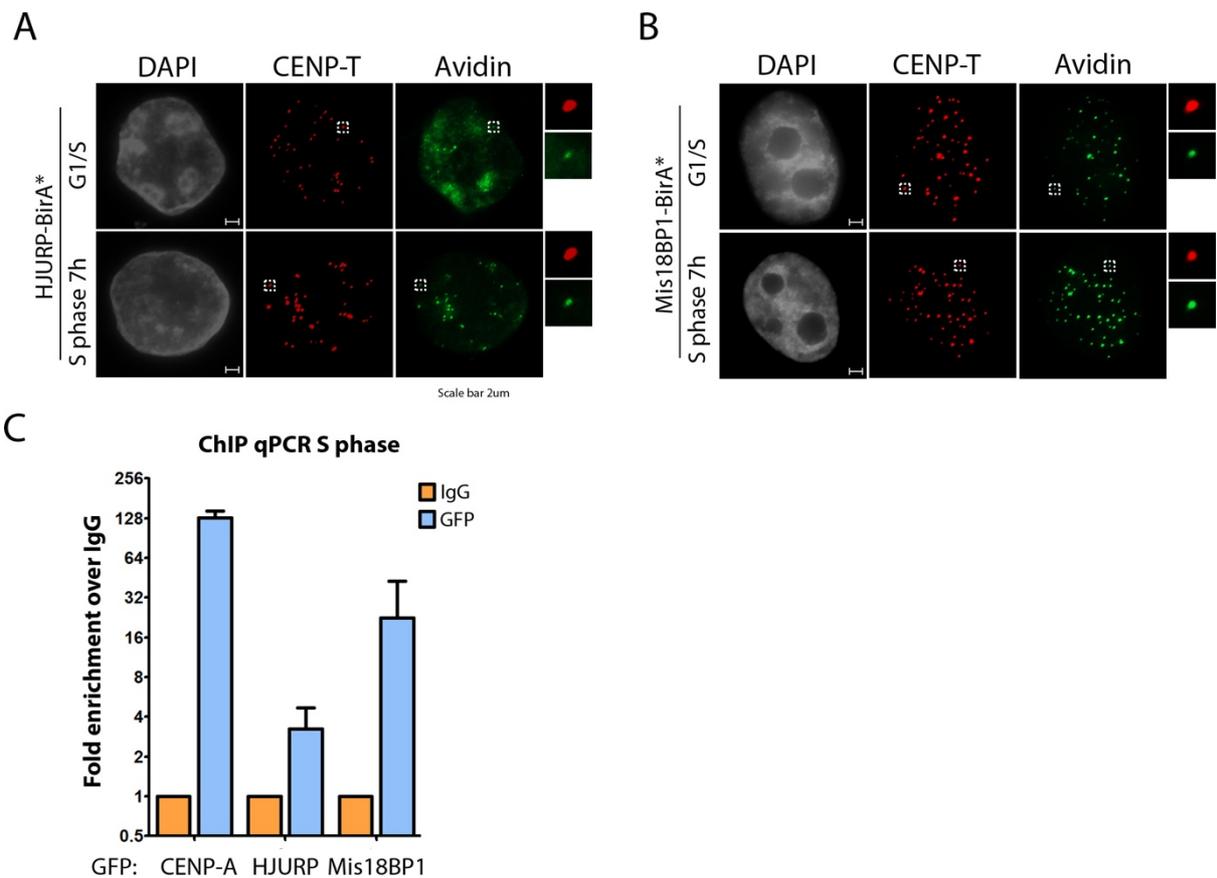
Figure 4.S1. (A)(B) Biotinylation pattern mediated by HJURP (A) or Mis18BP1 (B) during thymidine arrest and S phase release. Cells were treated with biotin for 7 hours at indicated cell cycle stages. (C) Chromatin Immunoprecipitation from cells expressing indicated GFP-tagged proteins in S phase. ChIP was performed using anti GFP antibody and normal rabbit IgG. RT-PCR was performed using primers against α-satellite DNA specific for chromosome 7. The graph represents an average of two independently performed experiments.

The accumulation of HJURP at centromeres has been observed during early G1 using immunofluorescence (Dunleavy et al., 2009; Foltz et al., 2009); however, HJURP has not been previously observed at centromeres during DNA replication (Bui et al., 2012). We reasoned that HJURP may interact with the centromere in a dynamic nature during DNA replication and may be difficult to detect using standard immunofluorescence. Therefore, we generated HJURP-BirA*-HA and Mis18BP1-BirA*-HA constructs which can be used to create a biotin trace of transient occupancy of the protein at the centromere during S-phase. Biotinylation mediated by HJURP-BirA* or Mis18BP1-BirA* was induced in cells arrested in early S phase or cells undergoing DNA replication (Fig. 4.S1 A, B). Biotinylation of centromere associated proteins was apparent in HJURP-BirA* and Mis18BP1-BirA* expressing cells in blocked cells as detected by fluorophore conjugated streptavidin. The signal increased in cells undergoing DNA replication, indicating HJURP and Mis18BP1 are recruited to the centromere transiently while cells are undergoing S phase. Consistent with a role for the G1 CENP-A deposition pathway in S-phase retention of CENP-A nucleosomes, all three Mis18 complex subunits (Mis18α, Mis18β and Mis18BP1) were previously found to be associated with nascent DNA in human cells (Alabert et al., 2014).

We hypothesized that protein-turnover may contribute to the transient association of HJURP with the centromere during S-phase, as other proteins are actively degraded during DNA replication to control their function (Roseaulin et al., 2013a; Roseaulin et al., 2013b). Therefore, we assessed HJURP-GFP, Mis18α-GFP and Mis18BP1-GFP recruitment in stably expressing cells during S-phase in the presence of the proteasome inhibitor MG132. Cells were released into S-phase and 3 hours post release treated with MG132 for 2 hours. GFP-tagged HJURP, Mis18α and Mis18BP1 all accumulated specifically at centromeres in response to the addition of MG132 (Fig. 4.3 A, B). Although GFP-tagged proteins levels increased at centromeres, this was not due to a direct increase in stability of either HJURP,

Mis18α or Mis18BP1 in response to MG132 treatment (Fig. 4.3 C). Live cell imaging showed that the accumulation of GFP-HJURP occurred within 20 minutes of MG132 treatment (Fig. 4.3 D, 4.S2), was unique to cells undergoing S-phase, and did not occur at high frequency in randomly cycling population (Fig. 4.3 E).
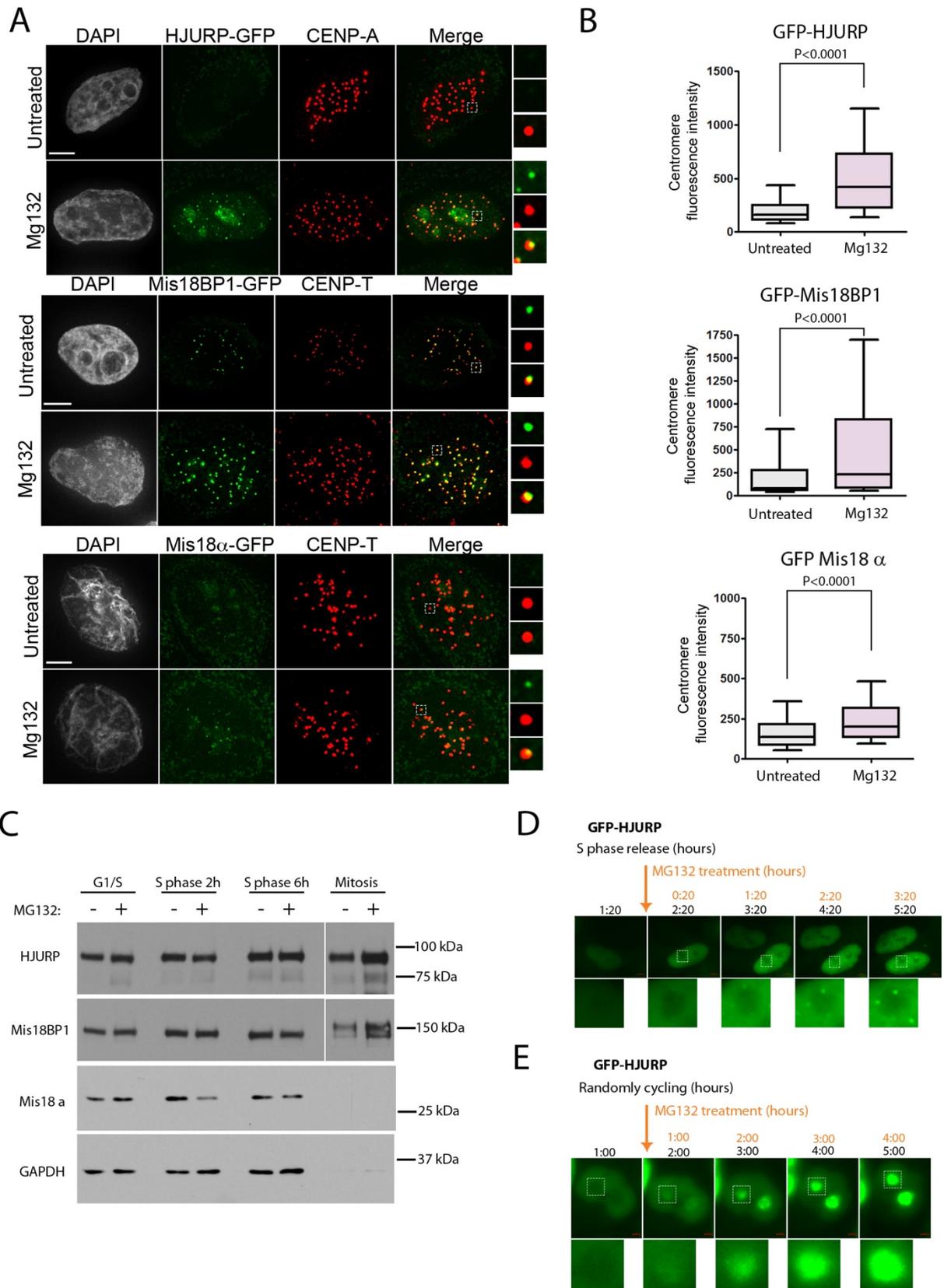
Figure 4.3. CENP-A deposition proteins accumulate at centromeres during DNA replication in response to MG132 treatment.

Figure 4.3. CENP-A deposition proteins accumulate at centromeres during DNA replication in response to MG132 treatment. (A) Representative images of cells expressing GFP fused HJURP, Mis18 BP1 or Mis18α during S phase with or without MG132 treatment. (B) Quantification of the fluorescence intensity of GFP fused proteins shown in A. Data was plotted using box-and-whisker plot: 5-95 percentile. The statistical significance was calculated using unpaired t-test and the p values are indicated. (C) Western blot analysis of HJURP, Mis18BP1 and Mis18α protein levels in response to MG132 treatment at indicated cell cycle stages. (D)(E) Montage of live cell movies of cells expressing GFP-HJURP in S phase released (D) or in asynchronous (E) cell populations. Cells were treated with MG132 during imaging starting from the time points indicated by the arrow.
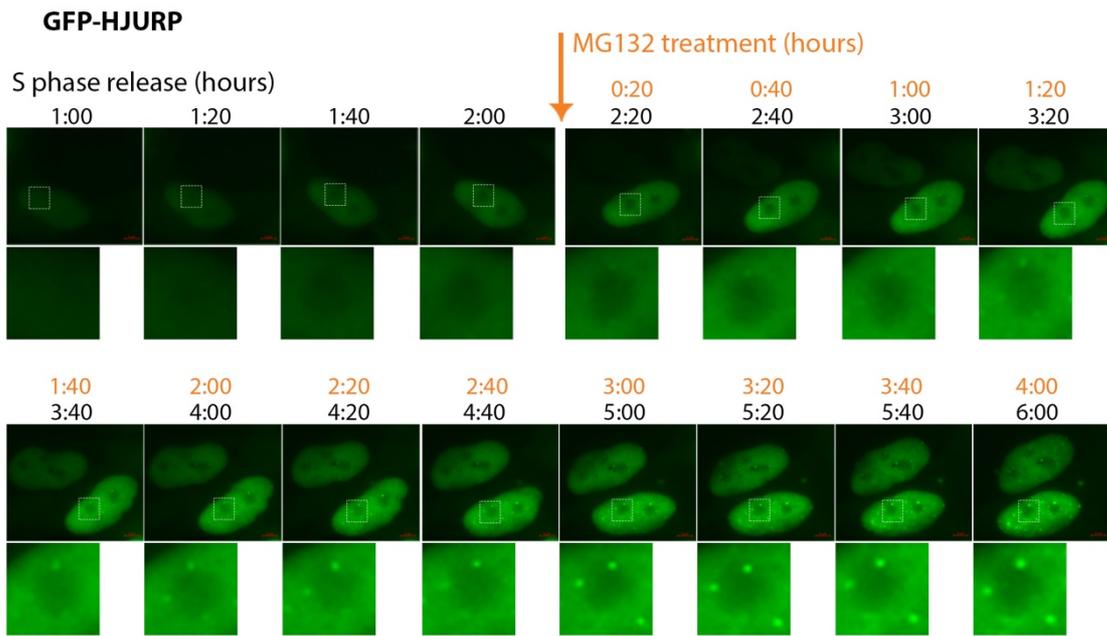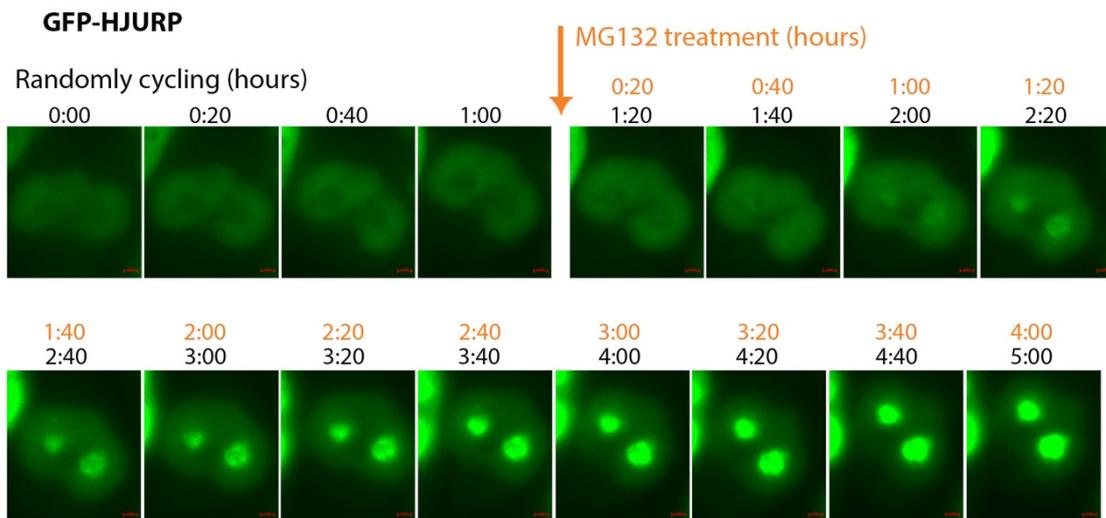
Figure 4.S2. (A)(B) Montage of live cell movies of cells expressing GFP-HJURP in S phase released (A) or in asynchronous (B) cell populations. Cells were treated with MG132 during the imaging and the treatment started from the time points indicated by the arrow.

.

HJURP is required for CENP-A inheritance during DNA replication

The association of HJURP with CENP-A during S-phase suggests that, in addition to its role in G1-coupled new CENP-A deposition, the HJURP chaperone may also contribute to CENP-A inheritance during DNA replication. In order to specifically test if HJURP is required for CENP-A retention during S phase, apart from its role in new CENP-A deposition, we employed an auxin-based degron (AID) system for the rapid depletion of proteins upon treatment with IAA (Nishimura et al., 2009) in HJURP-AID-YFP cell line that we generated (Fig. 4.2 C). IAA treatment for 24 hours results in reduction of HJURP protein levels below the limit of detection (Fig. 4.4 B). Consistent with the essential role of HJURP, cell survival was drastically reduced in response to the IAA treatment in HJURP-AID-YFP cells, whereas the parental DLD1-Tir1 cells were unaffected (Fig. 4.4 C).

Prior work has demonstrated that existing CENP-A nucleosomes show no turnover across the cell cycle (Bodor et al., 2013; Jansen et al., 2007). Consistent with these reports, we observed a similar amount of CENP-A at centromeres in cells blocked in early S-phase compared with centromeres that had progressed 7-hours through DNA replication (Fig. 4.S3 C, E). Therefore, CENP-A nucleosome are quantitatively retained during DNA replication. To directly test the role of HJURP in CENP-A inheritance, we examined the stability of endogenous CENP-A at the centromere through S–phase when endogenous HJURP-AID-YFP was degraded by addition of IAA. Parental DLD1-Tir1 or HJURP-AID-YFP cell lines were synchronized and treated with or without IAA for 90 mins prior to S-phase release in order to degrade existing HJURP in these cells (Fig. 4.4 E). DNA replication occurs asynchronously between centromeres of different chromosomes during mid to late S phase (O'Keefe et al., 1992; Ten Hagen et al., 1990). Therefore, the retention of endogenous CENP-A was analyzed in G2-phase cells. Fully replicated late G2 centromeres were easily identified as separated sisters. These cells had not yet entered mitosis based on the decondensed state of

the chromatin (Fig. 4.4 D, F). Under control conditions CENP-A levels decreased by approximately 50% at G2 centromere compared with centromeres in early S-phase (Fig. 4.4 D, F, G), which is consistent with redistribution of centromeric nucleosomes into daughter strands as cells progress through DNA replication followed by separation of sister centromeres. Parental DLD1-Tir1 cell line showed no change of CENP-A levels in G2 cells upon IAA treatment. However, the IAA treatment of HJURP-AID-YFP cells resulted in decreased centromeric CENP-A in G2 cells (Fig. 4.4 F, G). Similarly, analysis of mitotic cell populations revealed that CENP-A inheritance was reduced only in cells where HJURP was degraded (Fig. 4.4 H, I). Consistent with these results, the siRNA mediated depletion of HJURP prior to S phase entry also resulted in significant reduction of endogenous CENP-A at the centromeres (Fig. 4.S3 A-F). These data suggest that CENP-A nucleosomes require HJURP activity to faithfully transit the replication fork.
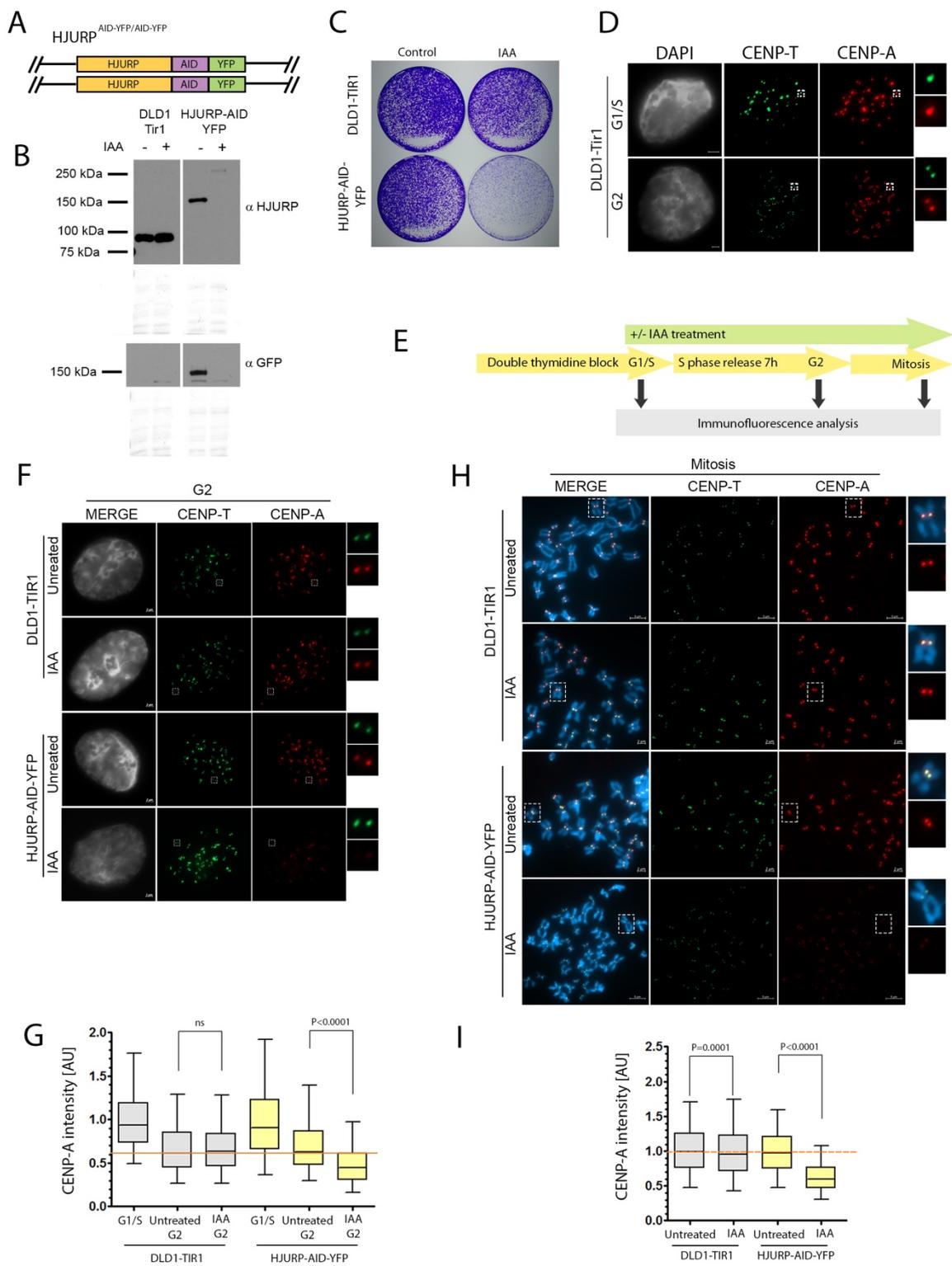
Figure 4.4. HJURP is required for CENPA retention across S phase

Figure 4.4. HJURP is required for CENPA retention across S phase (A) Schematic representation DLD1-Tir1 cell line where HJURP was endogenously tagged with AID-YFP at both alleles. (B) Western blot analysis of the efficiency of IAA-dependent HJURP degradation demonstrated by staining with indicated antibodies. Ponceau was used as a loading control. (C) Clonogenic assay using the parental DLD1-Tir1 and HJURP-AID-YFP cells plus or minus IAA treatment for 10 days. (D) Representative images of DLD1-Tir1 cells in G1/S arrest and G2. Insets are showing single centromeres and sister centromeres in G1/S and G2 cells, respectively. DNA was visualized by DAPI, CENP-T is shown in green and CENP-A is shown in red. (E) Schematic representation of the experiment in F and H. (F)(H) Representative images of cells in G2 and mitosis, respectively and treated as shown in D. DNA was visualized by DAPI, CENP-T is shown in green and CENP-A is shown in red. (G)(I) Quantification of F and H, respectively. The data normalized to G1/S condition (G) and untreated condition (I) within cell lines. Normalized data from four (G) or three (I) independent experiments was plotted using box-and-whisker plot: 5-95 percentile, n at least 8415 (G) and 4661 (I). The statistical significance was calculated using unpaired t-test and the p values are indicated.

A

Double thymidine block and release → siRNA treatment 38h / G1/S arrest in thymidine → S phase release 6h → Mitotic arrest 16h

Immunofluorescence analysis

B

Control siRNA — 50% / 25% / 10% / HJURP siRNA / Control siRNA 100%

150 kDa
100 kDa
75 kDa

anti HJURP

50kDa

anti Tubulin

C

Control siRNA — G1/S — 6.5 h release

HJURP siRNA — G1/S — 6.5 h release

D

DAPI / CENP-T / CENP-A / Merge

Control siRNA

HJURP siRNA

E S phase

Fluorescence intensity of endogenous CENP-A

P=0.0378   P<0.0001

G1/S   S phase   G1/S   S phase
Control siRNA   HJURP siRNA

F Mitosis

Fluorescence intensity of endogenous CENP-A

P<0.0001

Control siRNA   HJURP siRNA

Figure 4.S3.

Figure 4.S3. (A) Schematic representation of the experiment in C and D. Western blot analysis of the efficiency of siRNA mediated HJURP depletion demonstrated by staining with anti HJURP antibody. Tubulin staining was used as a loading control. (C)(D) Representative images of cells at indicated time points and treated as shown in panel A. DNA was visualized by DAPI, CENP-T is shown in green and CENP-A is shown in red. (E)(F) Quantification of C and D, respectively. Fluorescent intensity of Centromeric CENP-A was plotted using box-and-whisker plot: 5-95 percentiles. The statistical significance was calculated using unpaired t-test and the p values are indicated.

In order to specifically examine the retention of pre-existing CENP-A nucleosomes, we generated a stable cell line in which endogenous CENP-A was tagged with the SNAP tag (Jansen et al., 2007) and endogenous HJURP was tagged with the AID-YFP (Fig. 4.5 A). This dual tagged-protein approach allowed us to label existing CENP-A nucleosomes and determine their fate across the cell cycle when HJURP is degraded. HJURP-AID-YFP localizes to centromeres in telophase cells and is efficiently depleted upon IAA treatment (Fig. 4.5 B, C). In synchronized cells, pre-existing CENP-A was labeled with TMR* in the thymidine arrested population, cells were then released into S phase with or without IAA treatment and assayed at the following G1/S boundary 20 hours later. In agreement with previous studies, we observed approximately 50% reduction of pre-existing CENP-A levels following 1 cell cycle in control cells when compared to the initial G1/S TMR* signal (Jansen et al., 2007). This reduction is attributed to CENP-A dilution into sister chromatids during DNA replication (Fig. 4.5 D,E). HJURP-AID degradation resulted in significant loss of pre-existing CENP-A from the chromatin (Fig. 4.5 D, E). Previous work suggested that HJURP may be involved in the DNA damage response (Kato et al., 2007). In order to preclude a potential role for DNA damage induced by thymidine arrest, pre-existing CENP-A was labeled with TMR* in asynchronous cell population (CENP-A initial signal), and CENP-A-TMR* levels were analyzed 24 (or 27) hours later in cells subjected to IAA or control treatments (Fig. 4.5 F, G). Similarly, to synchronized cells, the degradation of HJURP-AID led to a loss of pre-existing CENP-A nucleosomes. We therefore conclude that, independently of its roles in new CENP-A deposition, HJURP is required for the inheritance of pre-existing CENP-A nucleosomes across DNA replication.
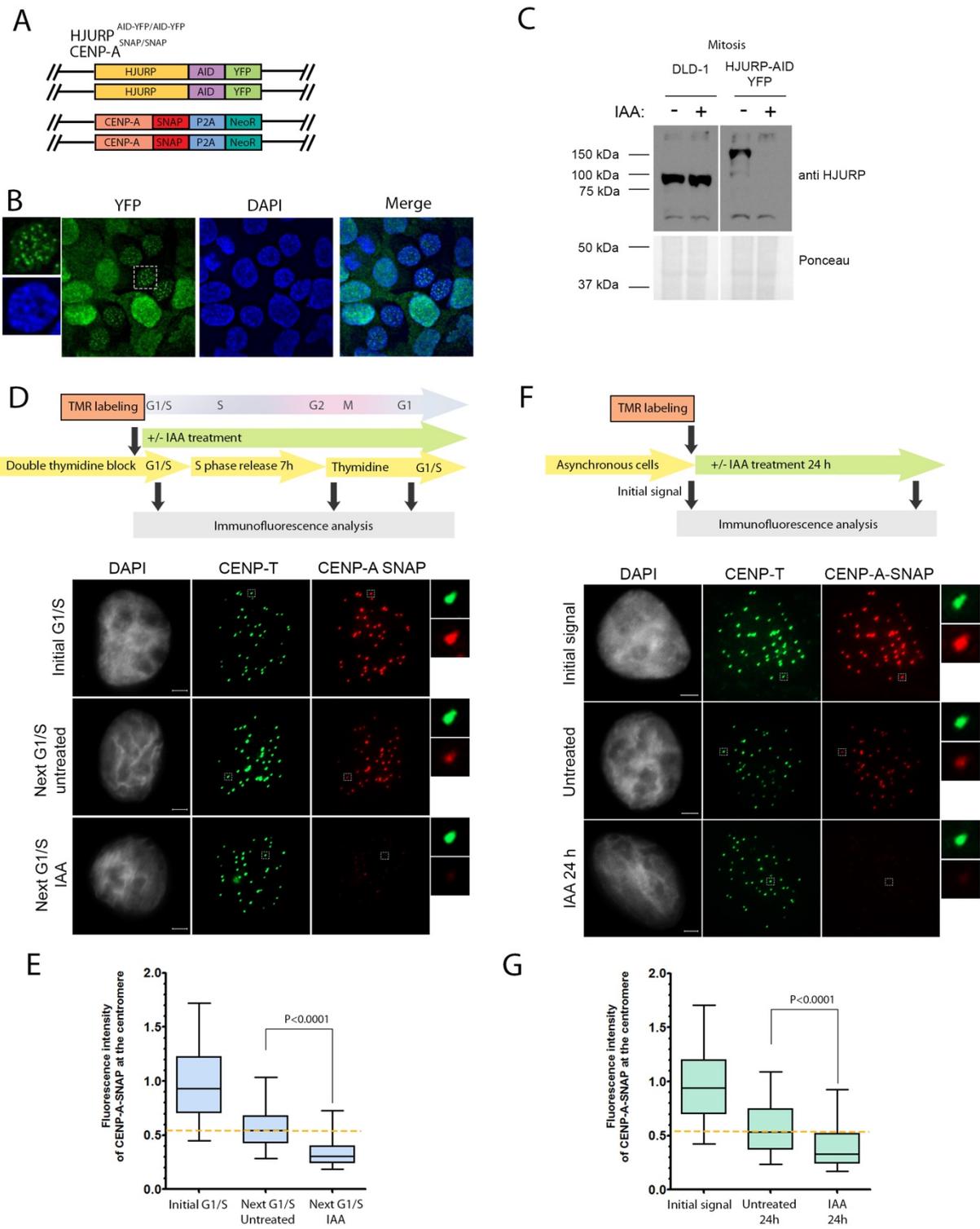
Data in this figure was generated with the help of Dr. L.Y. Guo from Dr. B.E. Black lab.

Figure 4.5. HJURP is required for CENPA inheritance of existing CENP-A nucleosomes

Figure 4.5. HJURP is required for CENPA inheritance of existing CENP-A nucleosomes. (A) Schematic representation DLD1-Tir1 cell line where HJURP was endogenously tagged with AID-YFP and the CENP-A was endogenously tagged with SNAP tag. (B) The IF images of the localization profile of HJURP-AID-YFP in cell line shown in A. (C) Western blot analysis of the efficiency of IAA-dependent HJURP degradation demonstrated by staining with HJURP antibody. Ponceau staining was used as a loading control. (D) (F) Schematic representation of the experiment (top). Representative images of cells at indicated time points and treated as shown in the top panel. DNA was visualized by DAPI, CENP-T is shown in green and CENP-A is shown in red (bottom). (E)(G) Quantification of D and F, respectively. The data was normalized to initial G1/S condition (E) or initial signal (G) within each individual experiment. Normalized data from two (E) or three (G) independent experiments was plotted using box-and-whisker plot: 5-95 percentile, n at least 2295 (E) and 4674 (G). The statistical significance was calculated using unpaired t-test and the p values are indicated.

CENP-A interaction with MCM-2 is required for retention during DNA replication

The MCM2 protein of the replicative MCM2-7 helicase complex directly binds to histones and is proposed to contribute to the retention of parental H3 nucleosomes during DNA replication (Huang et al., 2015; Richet et al., 2015). Mcm2 has been shown to bind both histone H3.1 and CENP-A, and amino acids R63 and K64 within H3.1 contribute to the interaction with MCM2 (Huang et al., 2015). The R63 and K64 residues are conserved among all human histone H3 variants (Fig. 4.S4 A). Therefore, we introduced alanines at positions R63 and K64 of CENP-A (CENP-A$^{RK->AA}$) and asked whether these mutations affect the interaction with MCM2 (Fig. 4.6 A, 4.S4 A). We generated recombinant CENP-A$^{RK->AA}$ plus histone H4 complex, wild type CENP-A and histone H4 complex and the MCM2-HBD fragment (a.a. 43-160) previously reported to be sufficient for interacting with the histone H3 and H4 complex. We tested the efficiency of interaction of MCM2-HBD with histone variants by *in vitro* pull down and SPR (Fig. 4.6 B, C). Consistent with previous studies we detected a physical interaction of MCM2-HBD with both histone H3/H4 and WT CENP-A/H4 complexes; however, histone H3 showed higher affinity for MCM2 binding when compared to CENP-A. The CENP-A$^{RK->AA}$ mutant failed to bind MCM2-HBD as efficiently as the WT form (Fig. 4.6 B, C).
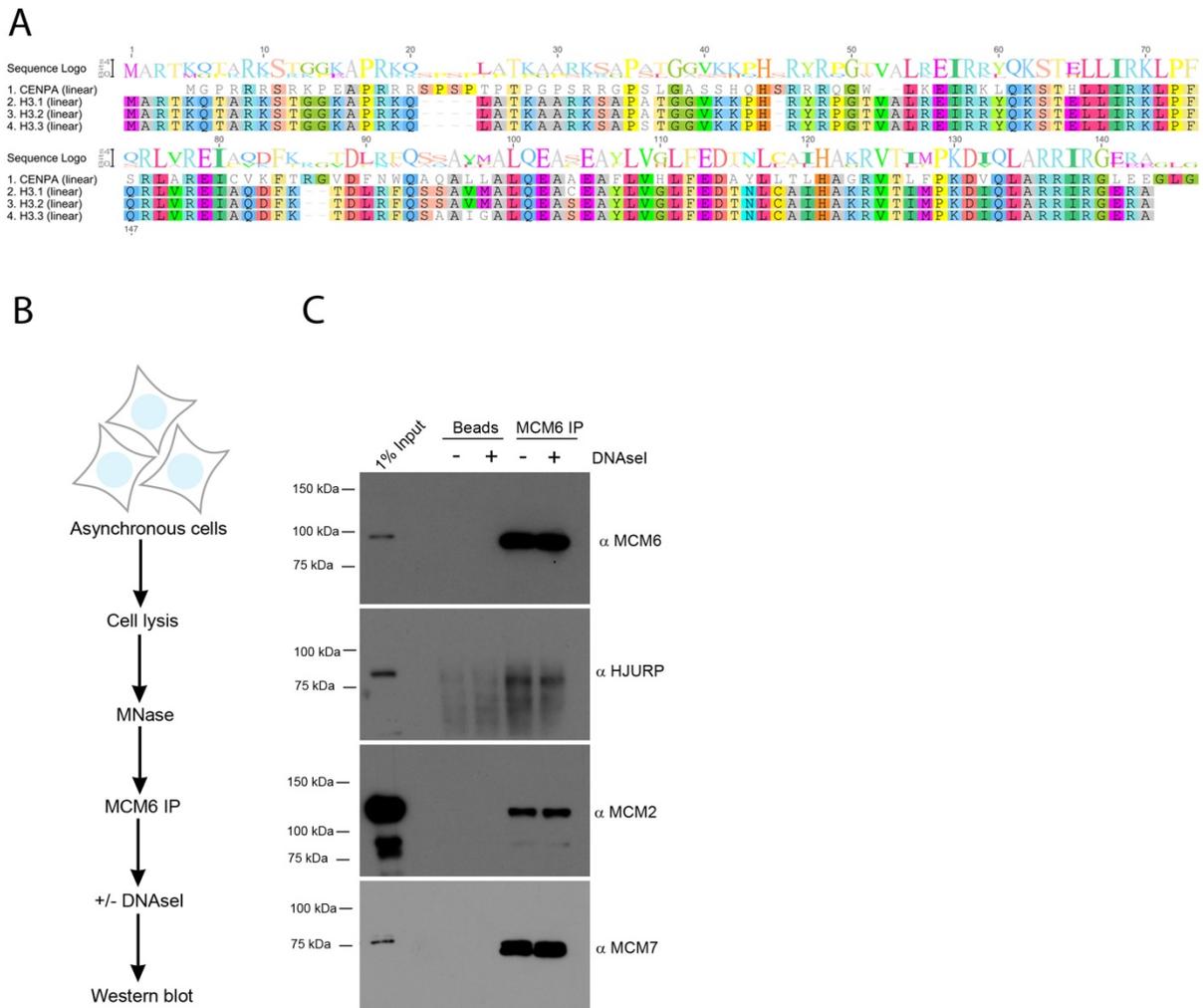
Figure 4.S4. (A) Sequence alignment of all human histone H3 variants demonstrating the conserved nature of RK motif implicated in mediating the interaction with MCM2 chaperone. **(B)** Schematic representation of the immunoprecipitation experiment shown in C. **(C)** Western blot analysis of MCM6 IP experiment performed from HEK derived cell lysates treated as indicated in panel B, demonstrating that the HJURP and MCM6 interaction is independent of DNA. Samples were analyzed with indicated antibodies.

In order to test whether the interaction of MCM2 with CENP-A is important for retention of centromeric nucleosomes during DNA replication, we tested the efficiency of inheritance across DNA replication of nucleosomes containing the CENP-A$^{RK->AA}$ mutant. Cell lines stably expressing either CENP-A^{WT}-GFP or CENP-A$^{RK->AA}$-GFP were synchronized using double thymidine block and release. Both CENP-A wild type and CENP-A$^{RK->AA}$ mutant were efficiently incorporated into centromeric chromatin, indicating that the CENP-A mutation did not alter *de novo* CENP-A deposition in G1 (Fig. 4.6 D). Fluorescence intensity of GFP-tagged CENP-A at the centromere was compared between thymidine arrested cells and cells that had been released into S phase for 7 hours. While the level of CENP-A^{WT} at the centromere was nearly identical before and after S-phase, levels of the CENP-A$^{RK->AA}$ mutant were significant reduced following S-phase, indicating that the CENP-A$^{RK->AA}$ mutant failed to be efficiently retained through DNA replication. Although reduced, the CENP-A$^{RK->AA}$ mutant was not completely lost from the centromeres when cells underwent DNA replication, which is likely due to the fact that MCM2 also makes contacts with the histone H4 and is consistent with our *in vitro* experiments as well as others (Huang et al., 2015; Richet et al., 2015).
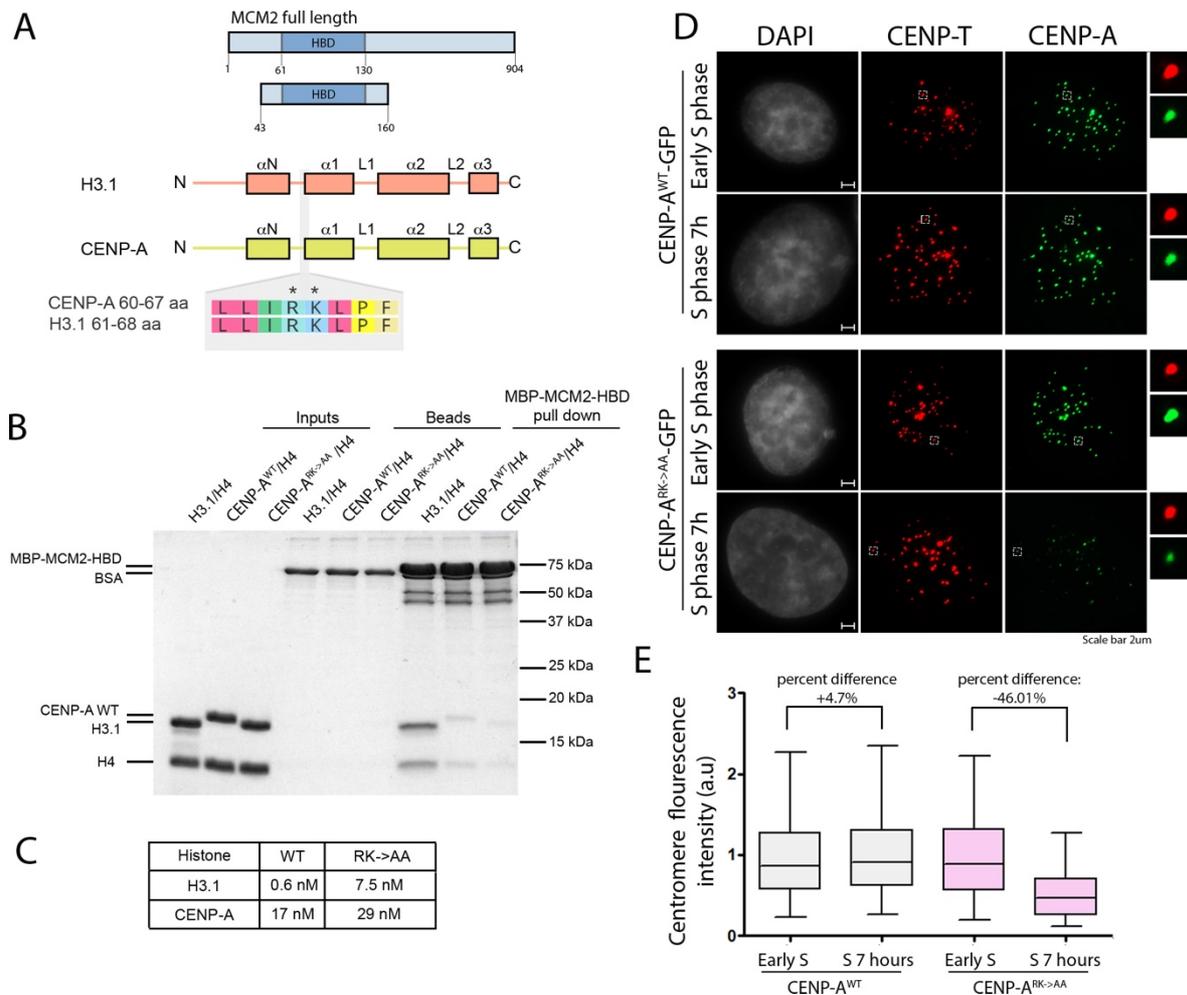
Data in panels B and C was generated by Dr. J. Huang.

Figure 4.6. MCM2 binds CENP-A and is involved in its maintenance during DNA replication (A) Schematic representation of constructs used in B and C. The CENP-A and H3.1 domain structure is shown. The alignment of 8 amino acid stretch corresponding to both histones demonstrates the conservation of Arginine 63 and Lysine 64 within the variants. (B) MBP-MCM2-HBD *in vitro* pull down demonstrating the interaction with indicated histone variants in the wild type and mutant form. (C) Table indicating Kd values measured to assess the strengths of interactions between MBP-MCM2-HBD and indicated histone variants in the wild type and mutant form. (D) Representative images of HeLa cells expressing either CENP-A^{WT}-GFP or CENP-A$^{RK->AA}$-GFP mutant at indicated cell cycle stages. DNA was visualized by DAPI, CENP-T is shown in red. (E) Quantification of D. The data vas plotted using box-and-whisker plot: 5-95 percentile, n at least 1943. The percent change of the levels of centromeric CENP-A^{WT}-GFP and CENP-A$^{RK->AA}$-GFP forms between experimental time points is indicated.

ASF1 together with MCM2 was proposed to facilitate recycling of canonical histones during DNA replication (Huang et al., 2015; Richet et al., 2015) . We therefore hypothesized that perhaps HJURP, analogous to ASF1, interacts with MCM2 to facilitate recycling of CENP-A containing nucleosomes specifically at centromeres. To determine whether Mcm2 and HJURP interact *in vivo*, endogenous HJURP-AID-GFP was immunoprecipitated using anti GFP antibody. Endogenous Mcm2 was co-immunoppreciated with HJURP (Fig. 4.7 A), consistent with previous experiments (Huang et al., 2015). Likewise, HEK cells transiently transfected with GFP-HJURP showed a similar interaction with endogenous Mcm2 (Fig. 4.7 B). To determine if the HJURP associates with the chromatin bound intact MCM complex, the endogenous MCM complex was immunoprecipitated from MNase digested cell lysates using an anti-MCM6 antibody. In addition to the other MCM components, the MCM6 co-immunoprecipitated HJURP and we confirmed that this interaction is DNA independent by treating immunopreciptates with DNAse prior to elution (Fig. 4.7 C, Fig.4.S4 B, C).

HJURP and MCM2 were both shown previously to directly bind the CENP-A Histone H4 heterodimer, (Cho and Harrison, 2011; Dunleavy et al., 2009; Foltz et al., 2009; Hu et al., 2011; Huang et al., 2015) and we confirm those observations here (Fig. 4.6 and 4.7). In order to determine whether MCM2 and HJURP are able to simultaneously bind CENP-A and histone H4 we tested if we could assemble the HJURP-MCM2-CENP-A/H4 complex *in vitro*. Recombinant MCM2-HBD and MBP-HJURP[1-208] fragments sufficient for CENP-A binding were purified from bacteria (Fig. 4.7 D). MBP pull down assays were performed in the presence and absence of recombinant CENP-A/Histone H4. MCM2-HBD copurified with MBP-HJURP[1-208] fragment only in the presence of CENP-A and histone H4 heterodimer, indicating that HJURP and MCM2 can bind CENP-A simultaneously (Fig. 4.7 E).

In order to understand the interaction between HJURP-CENP-A-H4-MCM2 complex we superimposed the crystal structures of CENP-A/H4/HJURP[1-80] (3R45) (Hu et al., 2011)

and MCM2$^{61-130}$/H3.3/H4 (5BNX) (Huang et al., 2015) using H3 histone variants as a reference (Fig. 4.7 F). This model structure demonstrates that the critical binding interfaces conferring MCM2 interactions (MCM2 residues L72, D80, D88, Y81, Y90, R110, E114, M117, R120, D121) with histone H3 and H4 heterodimer (CENP-A R63-K64; H4 R35-R36) are compatible with HJURP binding. However, the MCM2 L1 loop and HJURP β-sheet domain and the L1 loop—which blocks the DNA binding interface of CENP-A/H4—appear to occupy similar regions in the model. This perhaps implies that HJURP adopts different conformation when bound to CENP-A/H4 in the presence of MCM2, or vice versa. Furthermore, the Q89, H104 and L112 residues of CENP-A were previously demonstrated to be sufficient to confer HJURP recognition, and in our structural model this interface is accessible for HJURP binding in the presence of MCM2 (Fig 4.7 F bottom panel) (Bassett et al., 2012). The N85, H104, L112 residues of CENP-A were also shown to be sufficient for HJURP recognition, however, the N85 residue seems to be inaccessible in our model and may suggest that this mode of binding is not utilized when CENP-A is bound to MCM2 (Bassett et al., 2012). Collectively our data demonstrate that MCM2 is required for inheritance of CENP-A nucleosomes, and MCM2 together with HJURP chaperone bind CENP-A to facilitate its retention during DNA replication.
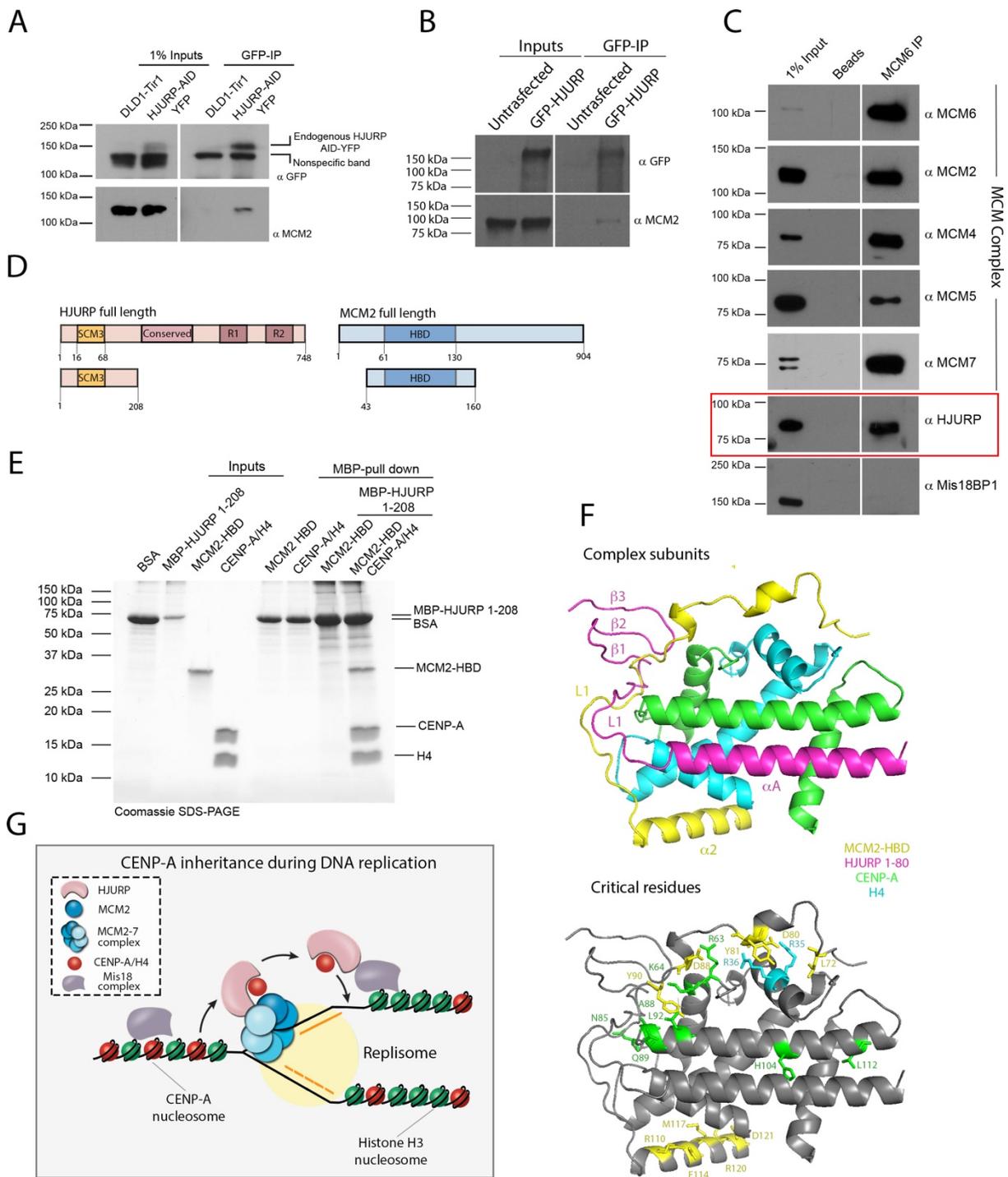
Figure 4.7. HJURP copurifies with the MCM2-7 helicase complex and simultaneously interact with MCM2-CENP-A/H4 proteins

Figure 4.7. HJURP co-purifies with the MCM2-7 helicase complex and simultaneously interact with MCM2-CENP-A/H4 proteins. (A) Western blot analysis of GFP IP experiment demonstrating the interaction of endogenous HJURP with endogenous MCM2. HJURP-AID-YFP cells were used as an input for the IP experiment and samples were analyzed with indicated antibodies. (B) Western blot analysis of GFP IP experiment performed from HEK cells overexpressing HJURP-GFP. Samples were analyzed with indicated antibodies. (C) Western blot analysis of MCM6 IP experiment performed from HEK derived cell lysates treated with Micrococcal nuclease. Samples were analyzed with indicated antibodies. (D) Schematic representation of constructs used in E. (E) MBP-HJURP[1-208] *in vitro* pull down demonstrating the interaction with MCM2-HBD only in the presence of CENP-A/H4 heterodimer. (F) The model of superimposed 3R45 and 5BNX crystal structures where H3.3 and CENP-A were used as a reference. MCM2-HBD is shown in yellow, HJURP[1-80], CENP-A and H4 are shown in pink, green, and aqua, respectively. Residues critical for facilitating the interaction of MCM2 with histones and mediating CENP-A recognition by HJURP are depicted in bottom panel. (G) The model of inheritance of CENP-A nucleosomes across DNA replication. MCM2-7 helicase complex is involved in unwinding chromatin ahead of the replication fork. HJURP is associated with MCM2-7 complex, and both MCM2 and HJURP can bind CENP-A nucleosomes simultaneously. The ability of CENP-A to be recognized by HJURP through CATD domain and MCM2 through the R63-K64 motif are both essential for facilitating CENP-A retention across S phase and maintaining centromere identity.

Discussion

During DNA replication nucleosomes are disassembled ahead of the replication machinery in order to allow for new DNA synthesis, and this process presents a challenge for CENP-A nucleosomes. Existing CENP-A is stably retained at the centromere for multiple cell divisions, suggesting that CENP-A nucleosomes are uniquely recognized, while chromatin is disrupted, and stably propagated to confer centromere identity (Bodor et al., 2013; Falk et al., 2015; Jansen et al., 2007). In this study, we identified the mechanism that facilitates inheritance of existing CENP-A nucleosomes during S phase. We found a novel function for HJURP outside of its known G1-phase role and demonstrated the association of HJURP with centromeric chromatin and parental CENP-A containing nucleosomes during DNA replication. Our work revealed that HJURP collaborates with the MCM2-7 helicase complex to facilitate retention of evicted CENP-A nucleosomes ahead of the replication machinery (Fig. 4.7 G).

Our experiments demonstrate the transient association of CENP-A nucleosomes with HJURP during DNA replication (Fig. 4.2 A-C, S1 A,C). Prior work showed that the CATD region, the domain of CENP-A which binds HJURP, is essential to confer CENP-A nucleosome stability, and is consistent the roles we propose for HJURP binding in retaining CENP-A during DNA replication(Bodor et al., 2013). Furthermore, using a degron strategy for rapid depletion of endogenous HJURP, we provide evidence that the association of HJURP with the centromeres in S phase is required to facilitate the retention of existing CENP-A nucleosomes (Fig. 4.4 E-I, 4.5 D-G, 4.S3).

DNA replication is a highly dynamic process and, although previous IF based studies relying on the accumulation of multiple copies of the protein at one locus, did not detect the presence of HJURP at centromeres in S phase cells or chromatin fibers, our more sensitive approaches using BioID labeling and ChIP have allowed us to capture a transient association

of HJURP with the replicating centromere (Bui et al., 2012; Dunleavy et al., 2009; Foltz et al., 2009). Our ChIP experiments detected more abundant association of endogenous HJURP with centromeric DNA of chromosome 7 in cells released from thymidine block for 3h when compared to thymidine arrested cells. This profile is consistent with reports showing that human centromeres are being replicated from mid to late S phase (O'Keefe et al., 1992; Ten Hagen et al., 1990).

CENP-A mRNA transcripts rise specifically in G2, and peak in mitosis. However, prior experiments using CENP-A driven by a constitutive promoter and thus providing new CENP-A throughout the cell cycle observed no loading during S phase(Shelby et al., 1997). Since we observe an interaction with HJURP and CENP-A in S-phase, this suggests that HJURP preferentially interacts with existing CENP-A rather than new CENP-A during S-phase. Consistent with this idea, the expression of HJURP Ser412, Ser448, and Ser473 mutants that prematurely localize to centromeres in S and G2 phase assembled new CENP-A only in G2 but not in S phase (Muller et al., 2014; Stankovic et al., 2017). Ser412, Ser448, and Ser473 are key residues of HJURP that undergo a decrease in phosphorylation at the M/G1 transition. How HJURP may delineate new and old CENP-A is not yet known.

Parental H3-h4 heterotetramers are recycled during DNA replication and old histones are not mixed with newly synthesized dimers during nucleosome re-formation following replication (Leffak, 1984; Leffak et al., 1977; Yamasu and Senshu, 1990). Regardless of this pattern of inheritance, existing H3 nucleosomes require the ASF1 chaperone, which is shown to disrupt the H3-H4 heterotetramer, and bind the H3-H4 heterodimer (English et al., 2005; Groth et al., 2007). CAF1 interacts with PCNA and in yeast it was shown to play a role in inheritance of epigenetic chromatin states (Enomoto and Berman, 1998; Moggs et al., 2000; Monson et al., 1997; Shibahara and Stillman, 1999). CAF-1 binds two H3-H4 dimers and promotes formation of a H3-H4 heterotetramer (Winkler et al., 2012) and perhaps it

collaborates with ASF1 to facilitate retention of H3/H4 tetramers. Likewise, HJURP binds to the CENP-A/H4 heterodimer(Cho and Harrison, 2011; Zhou et al., 2011). We demonstrate that HJURP is required for CENP-A retention during S-phase. CENP-A may also undergo a CENP-A/H4 heterodimer intermediate during DNA replication similar to canonical nucleosomes. Previously, we showed that HJURP dimerization is required for assembly of new CENP-A nucleosomes (Zasadzinska et al., 2013). The ability of HJURP to dimerize might also be required to facilitate the inheritance of pre-existing CENP-A/H4 heterotetramers.

We and others have demonstrated that CENP-A directly binds the MCM2 chaperone (Huang et al., 2015)(Fig. 4.6 C). We also show that MCM2 and HJURP can simultaneously interact with CENP-A/H4 *in vitro* (Fig. 4.7). Furthermore, disrupting the MCM2 binding interface within CENP-A by two amino acid substitutions (R63A and K64A) impairs the stable inheritance of the CENP-A nucleosome. We propose a model whereby CENP-A nucleosomes evicted ahead of the replication fork are recycled through the collaboration of HJURP and MCM2 chaperones (Fig. 4.7 G). Although we did not identify the MCM2-7 helicase components in our BioID screen, MCM4 was previously co-purified with CENP-A in human cells (Huttlin et al., 2017) suggesting that CENP-A is associated with MCM2 within the context of the intact MCM2-7 complex. *In vivo*, HJURP co-immunoprecipitates with the MCM complex, and additional direct contacts between HJURP and other MCM2-7 helicase subunits may underlie this interaction. Alternatively, the interaction may be mediated by the common binding of MCM2 and HJURP to the CENP-A/h4 heterodimer.

We demonstrated that MCM2 has a significantly lower affinity for CENP-A relative to histone H3. Huang et al. demonstrate that $H3^{RK->AA}$ mutants were poorly retained in general chromatin during S-phase, and we observed that MCM2 affinity for CENP-was lower than the poorly retained $H3^{RK->AA}$ mutant. Therefore, the lower binding affinity of MCM2 for

CENP-A relative to histone H3 may contribute to a lower CENP-A retention in the arms of chromosomes during replication, and HJURP binding in collaboration with MCM2 enhances stability of CENP-A specifically at centromeres to ensure stable centromere inheritance. MCM2 also collaborates with the FACT complex to recycle parental canonical histones that have been evicted from chromatin during DNA replication and transcription in yeast (Foltman et al., 2013) and it will be interesting in determine how additional canonical chaperones contribute to CENP-A inheritance in S-phase.

Acknowledgements

Author contributions

Materials and Methods

BioID and Mass Spectrometry. Affinity purification of biotinylated proteins was performed as previously described in (Roux et al., 2012). In brief, cells were incubated in DMEM 10% FBS media supplemented with 50 μM biotin for 6 hours (25x stock solution of biotin was prepared in DMEM at 1.25 mM concentration). Cells were washed three times with PBS and harvested. Cell pellets were lysed at 25°C in 1 ml lysis buffer (50 mM Tris, pH 7.4, 500 mM NaCl, 0.4% SDS, 5 mM EDTA, 1 mM DTT, and Complete protease inhibitor [Roche]). Cell lysates were sonicated, subsequently supplemented with Triton X-100 to 2% final concentration, and subjected to another round of sonication. Subsequently cell lysates were diluted with an equal volume of cold (4°C) 50 mM Tris (pH 7.4) and subjected to additional sonication. Cell lysates were then spun down at 10,000 relative centrifugal force for 5 mins at 4°C. Supernatants were collected and protein concentration was measured using the BCA assay. For the Mass Spectrometry analysis, heavy and light components were mixed at 1:1 ratio. Supernatants were incubated with Streptavidin Magnetic Beads (BioLabs) for overnight. Beads were then collected and washed twice with 1 ml buffer containing 2% SDS in dH$_2$O. The beads were then washed once with buffer containing 0.1% deoxycholate, 1% Triton X-100, 500 mM NaCl, 1 mM EDTA, 50 mM Hepes, pH 7.5, and once with buffer containing 250 mM LiCl, 0.5% NP-40, 0.5% deoxycholate, 1 mM EDTA, and 10 mM Tris, pH 8.1. Beads were then washed twice with buffer containing 50 mM Tris, pH 7.4, and 50 mM NaCl. Biotinylated proteins were then eluted from the beads with 100 μl of 2x Laemmli SDS-sample buffer saturated with biotin at 98°C. For the Mass Spectrometry analysis, the protein concentration was measured using the BCA assay.

Eluted samples were diluted with water up to 1ml and supplemented with 250 ul 100% TCA (final concentration of TCA 20%) and incubated for overnight at 4°C. Samples were spun down at 16000 rpm for 30 mins and protein pellet was washed with 1 ml of ice cold acetone 5

times. Protein pellets were dried in speed vac. Protein pellet was resuspended in buffer containing 100 mM ammonium bicarbonate (pH 8), 0.1% Rapigest and 10% ACN. Samples were supplemented with DTT at 5mM final concentration, and incubated at room temp for 1 hour. Iodacetamide was added at 12.5 mM final concentration, and samples were incubated in the dark for 1 hour. Proteins were digested with mass spec grade Trypsin (Trysin Gold from Promega). Trypsin was added at 1/20 ratio based on the amount of proteins measured in elutes. Samples were incubated for 15 hours at 37°C with shaking. The digestion was quenched with mass spec grade formic acid at 1% final concentration.

Sample mixtures were digested in 9ul volumes and injected directly onto an Easy Spray nano HPLC column ES801 (Thermo Scientific), packed with PepMap RSLC C18 media (2um, 100A, 50 um x 15 cm). An Easy nano LC (Thermo Scientific) delivered mobile phases A (0.1% formic acid in water) and B (0.1% formic acid in acetonitrile) as a gradient of 2 - 25% B over 60 min and 25 -50% B in 30 min at a flow rate of 250 nL / min. Mas spectra were collected using a Q Exactive Plus Orbitrap (Thermo Scientific) mass spectrometer at a resolution setting of 70,000 (FWHM @ 200 m/z) in full MS mode scanning from 300 - 2000 m/z and performing data-dependent MS/MS acquisition (top 10) with a resolution setting of 17,500 (FWHM @ 200 m/z). LC-MS data were analyzed using Proteome Discoverer software, version 1.4 (Thermo Scientific). MS and MS/MS spectra were searched using the Sequest HT algorithm. Trypsin-generated peptides were identified using a FASTA database of human protein sequences (Unirpot, October 2015) as well as a decoy database with scrambled sequences. False positives were filtered using a false discovery rate of 1%. All peptides were quantified in a label-free manner using the MS1 extracted ion chromatogram (XIC) peak area with a tolerance of 2 ppm. Ratios of [heavy: light] peptides were calculated and averaged for each identified protein in order to perform SILAC relative quantitation of proteins.

Chromatin immunoprecipitation. ChIP was performed as previously described in (Mayo et al., 2003). In brief, cells were synchronized using double thymidine block and released into S phase for 3 hours. Cells were then cross-linked on the plate by adding formaldehyde at a final concentration of 1% for 10 mins at 37°C. Glycine was added at a final concentration of 0.125 M to stop the cross-linking reaction. Cells were then washed twice with PBS, harvested and stored in -80 °C. Cell pellets were thawed on ice and lysed in 1ml of Farnham Lysis Buffer (5mM PIPES (KOH) pH 8.0, 85 mM KCl, 0.5 % NP40, and protease inhibitors [Roche]). Cell lysates were incubated for 10 mins on ice with shaking. Nuclei were pelleted at 800g for 2 minutes and resuspended in 250 ul of Lysis Buffer (1% SDS, 10mM EDTA, 50 mM Tris-HCl pH 8.0, protease inhibitors [Roche]). Lysates were incubated on ice for 30 minutes with shaking, and subsequently sonicated. Lysates were spun down at 13000 rpm for 10 minutes; supernatants were collected and measured for the protein concentration. An equal amount of protein per sample was then diluted 10 times with the Dilution Buffer (1.1% TritonX100, 1.2mM EDTA, 16.7mM Tris-HCl pH8.0, 167 mM NaCl, protease inhibitors [Roche]). Lysates were pre cleared with Protein A agarose beads and IgG for 30 minutes at 4°C. Agarose beads were then spun down at 1300 rpm for 2 minutes and the supernatants were supplemented with GFP antibody or rabbit IgG and incubated for 17 hours at 4°C. Protein A agarose beads were added and samples were incubated for 1 hour at 4°C. Agarose beads were then pelleted by spinning down at 1300 rpm for 1 minute at 4°C, and subsequently washed twice with Low sat Wash Buffer (0.1% SDS, 1% Triton X100, 2mM EDTA, 20mMTris HCl pH8.0, protease inhibitors [Roche]). Beads were then washed once with High Salt Wash Buffer (0.1% SDS, 1% Triton X100, 2mMEDTA, 20mM Tris-HCl pH 8.0,500 mM NaCl, protease inhibitors [Roche]), twice with LiCl Wash Buffer (0.25M LiCl, 1% NP40, 1% deoxycholate, 1mM EDTA, 10 mM Tris-HCl pH 8.0), and twice with TE

buffer. Each wash was performed for 5 minutes at RT. Beads were incubated with 75 ul of Elution Buffer (0.1M NaHCO3, 1% SDS) at RT for 15 minutes. Samples were then spun down at 2000 rpm for 2 minutes and eluates were collected. The elution step was repeated; the elution fractions were combined and supplemented with NaCl to a final concentration of 0.3 M following by 17 hours incubation at 65 °C. DNA was purified with PCR purification kit (Qiagen) and stored at -20C. Following ChIP, DNA was quantified by qPCR using standard procedures on a StepOne™ Real-Time PCR System. Primers for qPCR were used as previously described in Ohzeki et al., 2012: Forward: GGCATATGTGCAAGTGGATATAC; Reverse: TATCCACTTGCAGAC.

Cell Culture and Transfection. HEK293, HeLa, DLD1-Tir1 and lines derived from these parental cell lines were cultured in a 37 °C incubator in 5% CO_2 in DMEM supplemented with 10% FBS (OPTIMA) and 1% Pen/Strep. Cells for the DNA or siRNA transfection were seeded onto six-well plate prior to transfection and transfected when they reached 60% of confluency (HeLa, HEK293, DLD1-Tir1 cells). DNA transfection was conducted with the Lipofectamine 2000 reagent using standard protocol (Thermo Fisher Scientific) with 2ug of plasmid DNA per well. siRNA transfection was performed using the RNAiMAX transfection reagent (Invitrogen) using standard protocol. Cells were treated with either 2.5nM of HJURP siRNA (Dharmacon) or GAPD control siRNA (Invitrogen) for 38-40 hours.

Cell synchronization. Cells were synchronized with double thymidine block and release. Thymidine was added to culture medium at 20mM for 18 hours. Following the first thymidine treatment cells were washed twice with PBS and released into S phase for 8 hours, and subsequently treated with second thymidine block. Cells were released form the second thymidine arrest as indicated in the text.

DNA content analysis. Cells were synchronized and subsequently harvested using PBS supplemented with 3 mM EDTA. Cells were then washed with PBS and spun down at 1000 rpm for 5 min. Cell pellets were resuspended in 200ul PBS, fixed with 5mls of 70% Ethanol and stored at 4°C. Fixed cells were centrifuged at 1600 rpm for 5 min and washed with PBS + 1% FBS. Cell pellets were then resuspended in fresh PI/RNaseA solution (10ug/ml propidium iodide, 250ug/ml RNase A in PBS + 1% FBS) and incubated at 37°C for 30 minutes. Samples were analyzed for their DNA content using flow cytometry.

SNAP labelling. DLD1-Tir1 cells expressing endogenously tagged CENP-A-SNAP were plated on poly-lysine-coated glass coverslips. Asynchronous population or thymidine arrested cells was incubated in DMEM 10% FBS and labelled with 2 uM TMR-Star (Covalys) in complete growth medium for 20 min at 37°C. Cells were subsequently washed twice with each PBS, and DMEM and incubated for 30 min. Following incubation cells were washed twice with each PBS and DMEM. Asynchronous population was then incubated from 24 to 27 hours with or without IAA. Thymidine arrested population was treated +/- IAA for 60-90 minutes and subsequently released into S phase in the presence or absence of IAA. Cells were then pre-extracted with 0.1% Triton-X in PBS (3 minutes), fixed with 4% formaldehyde (10 minutes) and quenched with 100mM Tris, pH 7.5 (5 minutes), stained and analyzed by immunofluorescence microscopy.

Indirect immunofluorescence. Cells were pre-extracted with 0.1% Triton-X in PBS for 3 minutes, fixed with 4% formaldehyde for 10 minutes and subsequently quenched with 100mM Tris, pH 7.5 for 5 minutes. Fixed cells were incubated in blocking solution (0.1% Triton-X in PBS, 0.2% BSA, 2% FBS) for 1.5 h at RT, and incubated with indicated primary

antibodies for 1.5 h. Anti-CENP-T, anti-CENP-A and anti-HA antibodies were used at 1:5000, 1:1000, and 1:1000 dilution, respectively and detected using fluorophore conjugated secondary antibodies (Cy3, Cy5 or FITC, Jackson Immuno Inc.). Cy3-conjigates streptavidin (Jackson Immuno Inc.) was used at 1:1000 dilution. DNA was visualized with 0.2mg/ml DAPI in PBS and coverslips were mounted in Prolong Gold (Invitrogen).

Immunoprecipitation. Cells were harvested stored at -80C. In case of overexpression experiments, cells were harvested 24 h post transfection and stored at -80C. Cell pellets were thawed on ice and resuspended in RIPA lysis buffer (150mM NaCl, 1% NP-40, 0.3% deoxycholate, 0.15% SDS, 50mM Tris HCl pH 7.5, 1mM EDTA, 10% glycerol, protease inhibitors, 0.1mM PMSF, 5mM NaF, 10mM β-Glycerophosphate, 0.2mM NaV). For MCM6 IP lysis buffer was supplemented with 1mM ATP. Cell lysates were incubated for 15 minutes with rigorous vortexing periodically. In case pf MCM6 IP experiments cell lysates were diluted with an equal volume of dilution buffer (50 mM Tris HCl pH 7.5, 1mM ATP, protease inhibitors, 0.1mM PMSF, 5mM NaF, 10mM β-Glycerophosphate, 0.2mM NaV, 5mM CaCl$_2$) and subjected to MNase digestion for 4 minutes. Mnase treated lysates were quenched with 6mM EGTA. Lysates were centrifuged at 10000 rpm for 5 min at 4°C, and pre-cleared with Protein A agarose (Biorad) for 1 h at 4°C. Precleared extracts were then supplemented with anti-GFP antibody (1:1000, Cell Signaling) and incubated for 17 hours at 4°C. Extracts were subsequently incubated with Protein A Dynabeads (Invitrogen) on ice for 45 min, and washed once with RIPA buffer, followed by three washes in PBST (PBS + 0.1% Tween). For MCM6 IP the beads were washed 3 times with RIPA buffer supplemented with 1mM ATP and 3 times with was buffer 2 (PBS supplemented with: 0.1% Tween, 1mM ATP, protease inhibitors, 0.1mM PMSF, 5mM NaF, 10mM β-Glycerophosphate, 0.2mM NaV). Purified proteins were eluted by boiling in SDS sample buffer for 10 minutes**.**

Stable isotope labeling.   SILAC labeling with light and heavy analogs of Lysine and Arginine was performed in DMEM Media for SILAC (Thermo scientific) supplemented with either Arginine- HCl and Lysine- 2HCl or 13C6-Arginine HCl and 13C6-Lysine HCl (13C Molecular), respectively. The medium was supplemented with 10% Dialyzed Fetal Bovine Serum (JR Scientific). Cells were adopted for heavy and light medium for 20 cell divisions.

Surface Plasmon Resonance. SPR was performed on a Reichert4SPR instrument (Reichert Technologies). Biotinylated MCM2[43-160] was immobilized on a NeutrAvidin SPR sensor chip. Sufficient immobilization was obtained by flowing 100 μL 1 mg/ml protein through channel in SPR buffer: 20 mM HEPES pH 7.5, 500 mM NaCl, 1 mM DTT, 10 mM $MgCl_2$, 5 % glycerol, 0.1 % Triton X-100 and 0.1 mg/ml BSA. Serial dilutions of histone proteins (H3/H4 and CENP-A/H4) in SPR buffer were injected over the chip in cycles, and the bindings were monitored. After each injection, the chip was washed with regeneration buffer (1M Tris HCl pH4) to dissociate all histone proteins from hMCM2. Data was processed by Reichert's Surface Plasmon Resonance software. Kinetic constants were fitted to the binding curves, by a global fitting of all curves using 1:1 binding model.

Imaging and Quantification. Images were acquired using a × 100 oil-immersion Olympus objective lens on a DeltaVision microscope or the × 100 oil-immersion objective lens on a Zeiss microscope. Collected images are demonstrated as maximum stacked images. Images in Figures 4.3A, 4.5, 4.6, 4.S1 and 4.S3 were subjected to deconvolution prior stacking. Integrated intensities were derived from raw images subjected to ImageJ (using the CRAQ plugin) and the centromere marker was used as a reference. Quantification data was analyzed in GraphPad Prism software, the statistical significance was assessed using unpaired t-test.

The graphs were generated using GraphPad Prism software and displayed percentiles are as indicated in figure legends.

Stable cell lines. All cell lines expressing BiRA-HA* fusion proteins were generated using Flp-In™ T-REx™ System. The CENP-A-GFP CENP-A-R63AK64A-GFP stable cell line was generated using lentiviral transduction in HeLa cell line. HJURP$^{AID-YFP}$ and HJURP$^{AID-YFP}$/CENP-A^{SNAP} cell lines were generated using transient transfection of DLD1-Tir1 cells. In brief, using DLD-1 Flp-In T-Rex cells stably expressing Tir1 (Holland et al., 2012) as a starting point, CENP-A was tagged with a C-terminal SNAP tag using CRISPR-Cas9 as published (Guo et al., 2017), and a clone with CENP-A tagged on both alleles was used for experiments after verification by sequencing and CENP-A immunoblot. HJURP was tagged with a C-terminal AID-YFP tag using CRISPR-Cas-9, using the oligonucleotides (5'-CACCGAAACTAAAAGTGTGTAGCT-3' and 5'-AAACAGCTACACACTTTTAGTTTC-3') to target its 3' UTR. To generate the repair template, the AID-YFP sequence was amplified from the published pcDNA5-FRT-TO-H2B-AID-YFP construct (Holland et al., 2012), and 5′ and 3′ HJURP homology arms of~800 bp each were amplified from DLD-1 genomic DNA. All three pieces were inserted into a pUC19 backbone using HiFi DNA Assembly (NEB) and co-transfected with the HJURP gRNAs using published conditions (Guo et al., 2017). After transfection, YFP positive cells were isolated by FACS into 96-well plates, and clones were screened by microscopy for YFP signal. Cells were harvested and FACS sorted into single clones. Clones were verified using Western blot and genotyping methodology.

Mitotic chromosome spreads. Cells were arrested for 17 hours in 0.1 μg/ml nocodazole in DMEM Mitotic cells were harvested by mitotic shake off and spun down. Cell pellets were subsequently washed with 1ml of PBS and spun down. Cells were resuspended in hypotonic

solution (20 mM HEPES, pH 7.0, 1 mM $MgCl_2$, 0.2 mM $CaCl_2$, 20 mM KCl, LPC, and 0.5 µg/ml Colcemid) and incubated on ice for 10 mins. Cells were then spun down onto glass slides using cytospin, washed with PBS and subsequently fixed with 4% formaldehyde for 10 mins. Cells were then quenched with 100mM Tris pH 7.5 for 5 minutes and stored at 4°C.

Recombinant protein purification. Biotinylated human MCM2Δ1 (residue 43-160) was prepared by expressing Avitag-His6 tagged hMCM2Δ1 in BL21(DE3)pLysS cells at 18℃ for 18 hours after induced by 0.2 mM IPTG and the addition of 50 µM of D-biotin. The protein was purified by cobalt affinity chromatography and Superdex 200 size exclusion chromatography. MBP-His6-tagged hMCM2Δ1 was expressed in BL21(DE3)pLysS cells at 18℃. The fusion protein was purified by cobalt affinity chromatography.

Histone H3[RK->AA] mutant was generated by site-directed mutagenesis. Recombinant human histone H3 and H4 were expressed, reconstituted and purified according to (Dyer et al., 2004). Human CENP-A (wild type or mutant) and histone H4 were expressed by bicistronic expression vector in Rosetta cells at 37 °C for 3 hours (Tan et al., 2005). The cells were sonicated and cleared by centrifugation in 45mM sodium phosphate pH 6.8, 900mM NaCl, 1mM PMSF and 5mM BME. The CENP-A/H4 tetramer complex was then applied to hydroxyapatite resin (Bio-Rad), and eluted in buffer containing 45mM sodium phosphate pH 6.8, 3M NaCl, 1mM PMSF and 5mM BME. The eluted pool was changed to 20mM HEPES pH 7.5, 600mM NaCl and 1mM DTT by dialysis, before further purified by Source S chromatography.

In vitro recombinant protein pull-downs Thirty microliters of 50 µM MBP-His-MCM2[43-160] was incubated with 20 µl of amylose resin (New England Biolabs) in P300 buffer (NaP pH 7.0, 300mM NaCl). After washing the resin twice with 150 µl P300 buffer and equilibration

with PD buffer (20 mM HEPES pH 7.5, 250mM NaCl, 5 mM BME, 10 mM MgCl2, 5 % glycerol, 0.5 % NP40 and 0.1 mg/ml BSA). 50 μl of 15 μM recombinant histone proteins were added and allowed to bind to the immobilized MBP-His-MCM2$^{43-160}$ protein. Unbound histone proteins were removed by washing with PD buffer. The immobilized MBP-His-MCM2$^{43-160}$ bound protein complexes were then eluted from the resin by adding 25 μl 2x gel sampling buffer and heating at 95°C for 5 min. Samples were fractionated on an 15% acrylamide SDS-PAGE gel and visualized by Coomassie Blue staining.

Recombinant MBP-HJURP$^{1-208}$, MCM2-HBD proteins were purified by size exclusion chromatography using Superose 6 column and stored in SEC buffer (300mM Nacl, 20mM HEPES pH 7.5, 1mM DTT, 10% glycerol). MBP-HJURP1-208 and biotinylated MCM2$^{43-160}$ were then diluted with equal volumes of 2x pull down buffer (200 mM NaCl, 20mM HEPES ph 7.5, 1mM DTT, 20mM MgCl$_2$, 1% NP40, 0.2 mg/ml BSA). Amylose beads (New England Biolabs) were washed 3 times with 2x pull down buffer and subsequently incubated with MBP-HJURP$^{1-208}$ fragment for 1 hour at 4°C. The biotinylated MCM2$^{43-160}$ was then added at equimolar stoichiometry with respect to MBP-HJURP$^{1-208}$ fragment. The samples were subsequently supplemented with either recombinant CENP-A/H histones at equimolar stoichiometry or dilution buffer (250 mM NaCl, 20mM HEPES pH 7.5, 1mM DTT, 10 mM MgCl$_2$, 0.5% NP40, 0.1 mg/ml BSA). Protein complexes were incubated for 4 hours at 4°C, and subsequently washed 3 times with dilution buffer. The beads were then resuspended in 2xSB, boiled and analyzed by Coomassie stained SDS-PAGE.

Chapter 5: Mechanism regulating HJURP stability

Abstract

The CENP-A deposition pathway is cell cycle regulated, and depends on HJURP and the Mis18 complex. The levels of these proteins accumulate during late G2, persist throughout CENP-A deposition in early G1, and are subsequently diminished. We hypothesize that the stability of HJURP and the Mis18 complex is protected from degradation during G2 and mitosis and there is a regulatory mechanism in play required for degradation of these proteins following CENP-A deposition, in late G1. In human cells mechanisms controlling the stability of Mis18β and Mis18BP1 subunits have been demonstrated. Cul1 mediates the ubiquitylation of the Mis18β during interphase but not mitosis, and this modification targets Mis18β for proteasomal degradation. Mis18BP1 was shown to be co-modified by SUMO and ubiquitin during mitosis, which leads to its degradation via RNF4-dependent pathway. However, it remains unclear what mechanism governs HJURP protein levels in humans. Two independent mass spectrometry screens revealed that the WDR18 protein is associated with Mis18α and HJURP during mitosis. In this chapter I will focus on the potential role of WDR18 protein in HJURP stability. I will discuss my preliminary results which allowed me to form a hypothesis that WDR18 protein might be involved regulating HJURP turnover in interphase cells. These results however, were not easily reproducible. I will discuss here the collected data, pitfalls and future directions.

Introduction

The stability of proteins involved in CENP-A deposition pathway

The activity of proteins known to be involved in CENP-A deposition pathway, including HJURP and the Mis18 complex, is restricted to late G2-early G1 and late mitosis-early G1, respectively (Foltz et al., 2009; Kim et al., 2014). In addition, to the CDK mediated regulation of CENP-A deposition timing described in chapter 1 , it is thought that the cell cycle specific regulation of CENP-A deposition might also rely on controlling the protein stability. This notion is supported by the discovery that the protein levels of both HJURP and Mis18β are elevated prior to and throughout the timing of CENP-A deposition, while they are significantly diminished after CENP-A deposition is thought to be completed. Recently the mechanism controlling the stability of Mis18β has been identified. Mis18β was shown to be associated with the βTrCP-containing Skp1/Cul1/F-box protein complex specifically during interphase but not mitosis (Kim et al., 2014). Cul1 is a E3 ubiquitin ligase, that upon association with the Mis18β subunit, mediates its ubiquitylation and targets it for proteasomal degradation. This mechanism therefore contributes to the interphase inactivation of Mis18 complex (Kim et al., 2014). The MIS18BP1 subunit of the Mis18 complex was demonstrated recently to be co-modified by SUMO and ubiquitin and these modifications were enriched specifically during mitosis, leading to Mis18BP1 degradation via RNF4-dependent pathway (Cuijpers et al., 2017).

In human cells HJURP protein levels peak during late G2 and mitosis and are diminished at the end of G1 (Foltz et al., 2009). This suggests that the fluctuation of cellular HJURP levels might be a result of a tightly controlled expression levels as well as regulation of protein stability throughout the cell cycle. Importantly HJURP was found among proteins co-modified by SUMO and ubiquitin upon MG132 treatment (Cuijpers et al., 2017). This suggests that a crosstalk between SUMO and ubiquitin modifications leading to proteasomal

mediated degradation is implicated in regulation of HJURP stability in humans. However, the mechanism regulating HJURP protein in has not been identified.

Small protein PTMs, SUMO, and ubiquitin

Conjugation of ubiquitin or ubiquitin-like moieties to lysine residues of a protein substrate can change its fate within the cell by altering its cellular localization, or the affinity to its binding partners or, only in case of ubiquitylation, targeting it for proteasomal degradation. In addition to its role in protein turnover, ubiquitylation is also associated with various cellular processes including DNA repair pathways, regulation of the cell cycle, signaling pathways and endocytosis (Haglund and Dikic, 2005). Ubiquitin attachment to a substrate depends on the activity of three enzymes: E1, E2 and E3. E1 is required for ATP-dependent ubiquitin activation. E2 is essential for subsequent ubiquitin conjugation. The E3 ubiquitin ligase is responsible for the transfer of the activated ubiquitin moiety onto substrate protein. Targeting proteins for proteasomal recognition and subsequent degradation requires attachment of long ubiquitin chains, whereas conjugation of short ubiquitin moieties or monoubiquitylation can alter the fate of the modified protein substrate but does not lead to its degradation (Kerscher et al., 2006). The removal of ubiquitin occurs though a catalytic reaction mediated by processing proteases called deubiquitylating enzymes (DUBs) (Nijman et al., 2005; Reyes-Turcu et al., 2009).

Sumoylation is an ubiquitin-like, covalent posttranslational modification of lysine residues (Geiss-Friedlander and Melchior, 2007). SUMO modification has been shown to mediate various functions, including a role in nuclear-cytosolic transport, transcriptional regulation, altering the cellular localization of a target protein, affecting the target's affinity with other binding partners or altering protein stability (Johnson, 2004).

Sumoylation is mediated by a cascade of enzymes that resembles the ubiquitylation pathway. The attachment of the SUMO modification (in humans there are three forms of SUMO protein: SUMO1, SUMO2, SUMO3) to the substrate requires the activity of 3 enzymes: an E1 activating enzyme, an E2 conjugating enzyme and an E3 ligase. Sumoylation is a reversible modification that can be removed by SUMO-specific proteases belonging to the SENP family of enzymes (Hay, 2007).

The structure and function of human WDR18 protein

The members of the WD-repeats-containing protein family have been found only in eukaryotes and were shown to be involved in multiple cellular processes including cell division, apoptosis, cell fate determination, transcription, transmembrane signaling pathways and RNA processing (Smith, 2008). The WD repeat is an approximately 40 amino acid structural motif that usually ends with Trp-Asp sequence. The WD-repeat containing proteins differ not only in their function but also within their domain structure and they contain from 4 to 8 or more copies of the WD motif. Importantly the function of WD containing proteins was found to be predominantly associated with the regulatory mechanisms and mediating protein-protein interactions rather that performing enzymatic reactions (Smith, 2008).

One of the members of the WD repeat containing family is the human WDR18 protein. WDR18 is a 47 kDa protein (432 aa) that has been shown to localize to the nucleoli, nucleoplasm, and cytoplasm. Low levels of this protein were detected in the chromatin fraction of cells. WDR18 sequence analysis reveals the presence of 6 repeats of the WD motif that span the whole protein sequence, as well as a c-terminal coiled coil domain.

The function of WDR18 was linked to the desumoylation of transcription factors and proteins involved in ribosome biogenesis (Castle et al., 2012; Fanis et al., 2012; Finkbeiner et al., 2011). WDR18 together with other two components: PELP1 and Text10 has been shown to be associated with SENP3 (-a sumo specific protease), which is essential for maturation of the large ribosome subunit in nucleoli. SENP3-mediated desumoylation of PELP1 was demonstrated to be crucial for ribosome biogenesis, and the PELP1-Tex10-WDR18 complex was proposed to play a regulatory role in the ribosome formation. WDR18 in complex with SENP3, PELP1, TEX10, LASIL and Chtop was shown to associate with the transcription factor Zbp-89. The role of this interaction is to remove a sumoylation mark from Zbp-89 in a

SENP3 mediated manner, which changes the activity of Zbp-89 and promotes transcription (Fanis et al., 2012).

WDR18 was also proposed to function in the regulation of a DNA damage checkpoint signaling (Yan and Willis, 2013)(89). WDR18 directly interacts with TopBP1 (Topoisomerase II Binding Protein), which is required to co-activate ATR kinase in response to DNA double strand breaks. Upon its activation, ATR can phosphorylate its downstream targets such as Chk1 (Chen et al., 2009; Chen and Poon, 2008). Phosphorylated Chk1 then mediates the regulation of the cell cycle progression (Boddy et al., 1998). WDR18 was demonstrated to directly interact with both TopBP1 and Chk1, facilitating ATR mediated Chk1 phosphorylation (Yan and Willis, 2013).

The MitoCheck database demonstrated that human HJURP is associated in a complex with WDR18 during mitosis, as assessed by mass spectrometry analysis. The BioGrid database demonstrates the association of human WDR18 with the Cul3 ubiquitin ligase and USPU2 (ubiquitin specific protease involved in deubiquitylation). Furthermore, WDR18 in Drosophila, which shares 83% sequence similarity with its human homolog, was found to be associated with RNF2-E3 ubiquitin ligase, as indicated by the BioGrid database. Based on these interactions we hypothesize that the WDR18 protein contributes to the mechanism regulating the stability of HJURP in the preassembly complex.

Results

WDR18 interacts with HJURP and Mis18α *in vivo*

The mass spectrometry analysis of proteins interacting with the human Mis18 complex during mitosis that was done in our lab revealed an association of Mis18α subunit with the WDR18 protein. We further examined the interaction of WDR18 and the Mis18 complex by immunoprecipitation experiments. HA tagged Mis18α and HA-Mis18β or the HA-Mis18α alone were co-transfected with the WDR18-GFP into HEK293 cells. 24 hours postransfection, cells were harvested and upon lysis in RIPA buffer subjected to immunoprecipitation with use of an anti-GFP antibody (Fig 5.1 A). Cells expressing HA-Mis18α and HA-Mis18β alone or co-expressing GFP-WDR18 with either HA-H3.1 or HA-CENP-A were used as a negative control in this experiment. Protein complexes associated with the WDR18 were analyzed by Western blot using anti-GFP and anti-HA antibodies. WDR18 co-immunoprecipitated with Mis18α and the Mis18α/β complex but not with H3.1 or CENP-A (Fig 5.1 A). Importantly the HA-Mis18α/β complex, when expressed alone, was not detected in the IP fraction, demonstrating that WDR18 and the MIS18α/β interact *in vivo* (Fig 5.1 A).

The MitoCheck database indicates that WDR18 is associated with HJURP during mitosis (as assessed by Mass Spec). In order to further investigate whether HJURP indeed interacts with WDR18 *in vivo* we performed an immunoprecipitation experiment, where HA-HJURP was co-transfected with WDR18-GFP. Cells expressing HA-HJURP alone were used as a negative control in this experiment. HA-HJURP was present in the IP fraction only when co-expressed with WDR18-GFP, demonstrating an association between these two proteins (Fig 5.1 B). It remains to be tested whether HJURP and the Mis18 complex interact with WDR18 independently or form a large multi-subunit complex.
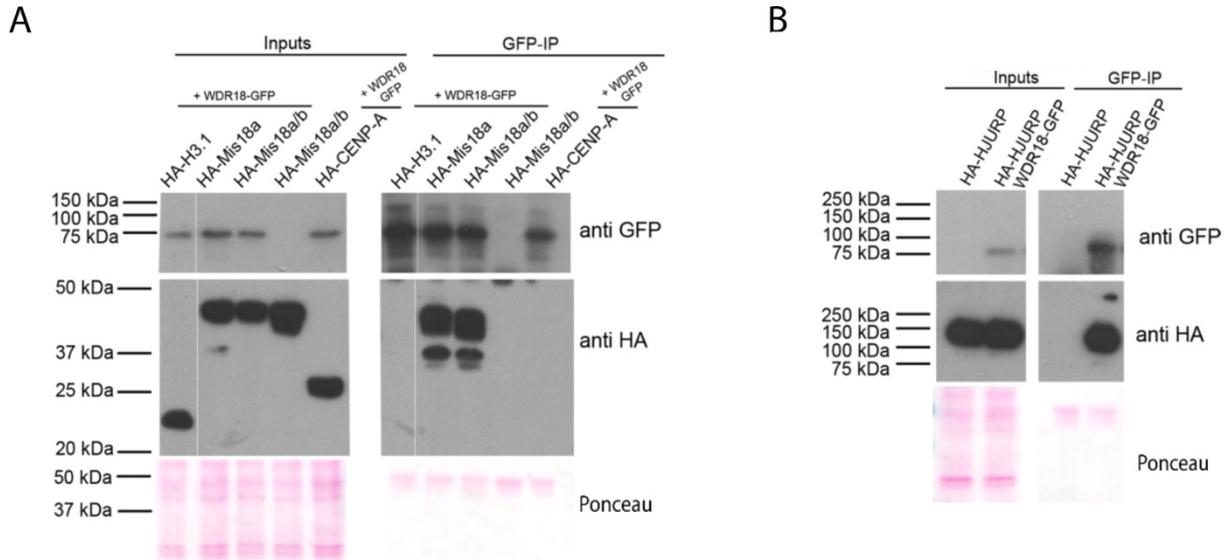
Figure 5.1. WDR18 interacts with HJURP and Mis18α *in vivo*

(A) (B) Cells co-expressing indicated HA-tagged fusion proteins together with WDR18 GFP, were used as an input for anti GFP immunoprecipitation followed by immunoblot using antibodies against HA and GFP. Ponceau staining was used as a loading control.

WDR18 depletion affects HJURP but not Mis18α protein levels

Based on our observations we aimed to ask whether WDR18 is implicated in the CENP-A deposition pathway. In order to address this question, we set out to deplete endogenous WDR18 protein expression using shRNA in human cells. We generated a stable cell line expressing WDR18-LAP which allowed us to test the efficacy of WDR18 depletion using anti-GFP antibody (Fig 5.2 A). HeLa T-Rex cells stably expressing WDR18-LAP were transfected with either WDR18 shRNA at two different concentrations, or a control vector, for 72 hours. Cells were then harvested and analyzed by Western blot with use of anti-GFP, anti HJURP, anti Mis18β antibodies (Fig 5.2 A). The analysis of the Western blot revealed that upon shRNA treatment the WDR18-LAP protein level was reduced by close to 75% when compared to cells treated with the empty vector, indicating that the shRNA treatment works efficiently. Strikingly, the levels of HJURP protein, but not Mis18β, were reduced by close to 50% in this experiment (Fig 5.2 A). In order to test whether WDR18 depletion affects Mis18β levels, we used cells stably expressing Mis18α-LAP and treated them with either control or WDR18 shRNA (Fig 5.2 B). This approach allowed us to monitor Mis18α protein levels. We concluded that upon WDR18 depletion the levels of Mis18α-LAP remain unaffected (Fig 5.2 B). It is possible however, that the GFP fusion increases the stability and renders Mis18α insensitive to WDR18 depletion. Therefore, these results need to be verified by analyzing the endogenous Mis18α protein levels upon WDR18 downregulation using Mis18α antibody.
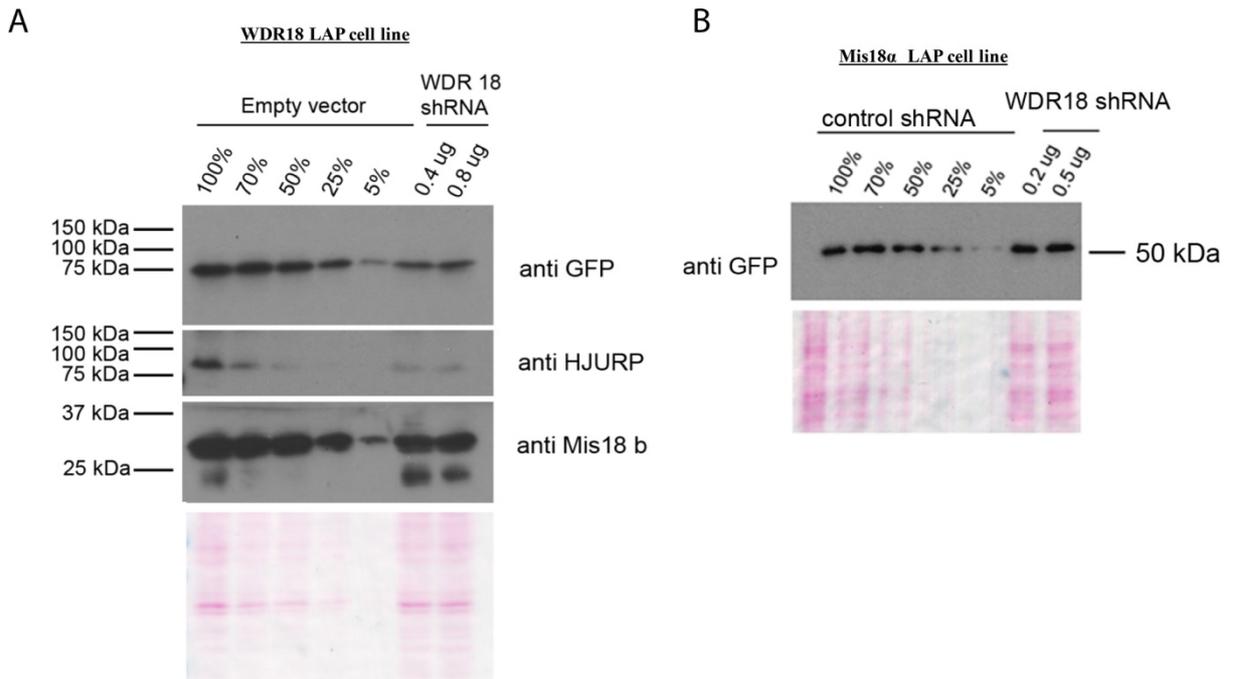
Figure 5.2. WDR18 depletion affects HJURP but not Mis18α protein levels

(A)(B) Western blot showing the effects of WDR18 depletion. Cells stably expressing WDR18-LAP (A) or Mis18α-LAP (B) were transfected with control or plasmid encoding WDR18 shRNA at two concentrations. Whole cell extracts were analyzed by Western blot using indicated antibodies. The efficiency of WDR18 shRNA was assessed by anti-GFP antibody in panel A. Ponceau staining was used as a loading control.

WDR18 was previously shown to mediate the sumoylation of transcription factors implicated in ribosome biogenesis pathway. We hypothesized that depletion WDR18 could interfere with the regulation of HJURP transcription. In order to test this hypothesis, we asked whether WDR18 depletion affects HJURP mRNA levels. We isolated RNA from cells treated with either control or WDR18 shRNA or siRNA and assessed the mRNA levels corresponding to WDR18 and HJURP. We designed two sets of primers for HJURP mRNA targeting either the boundary of exon 1 and 2 or the boundary of exons 4 and 5 (Fig 5.3 A, B). The WDR18 transcript levels were significantly downregulated in response to the shRNA and siRNA treatments when compared to the control, suggesting that the knockdown works efficiently in our experiments (Fig 5.3 A, B). The analysis of HJURP transcription profile upon siRNA mediated WDR18 depletion did not result in significant decrease in mRNA levels when compared to the control siRNA treatment (Fig 5.3 A, B). Strikingly, the WDR18 shRNA treatment resulted in increased HJURP mRNA levels when compared to the control shRNA, and this was true for both primer sets we used (Fig 5.3 A) This perhaps can be explained by the existence of a potential feedback loop, that activates HJURP expression when HJURP protein levels are low due to WDR18 downregulation. Importantly, this result was only observed in case of shRNA but not siRNA treatment and should be verified by another strategy for WDR18 depletion, such as auxin-based degron system for the rapid depletion of proteins or CRISPR Cas9 mediated gene silencing.

HJURP expression has been previously demonstrated to be cell cycle regulated. Therefore, the observed decrease of HJURP protein levels due to WDR18 depletion might be a result of the cell cycle arrest. We set out to analyze the effects of WDR18 downregulation on cell cycle progression. I order to test that, we performed DNA content analysis in cells treated with either control or WDR18 shRNA and synchronized

with double thymidine block and release. 72 hours after shRNA treatment cells were harvested and stained with propidium iodide, and subsequently analyzed by flow cytometry. We looked at randomly cycling, thymidine arrested, and S phase released cell. We did not observe any significant changes in the cell cycle profiles when comparing cells treated with control or WDR18 shRNAs. These data suggest that WDR18 depletion do not affect cell cycle progression. Overall this data suggests that the decrease of HJURP protein levels observed in our previous experiments is not due to altered expression or cell cycle arrest, and can be attributed to altered HJURP protein stability.
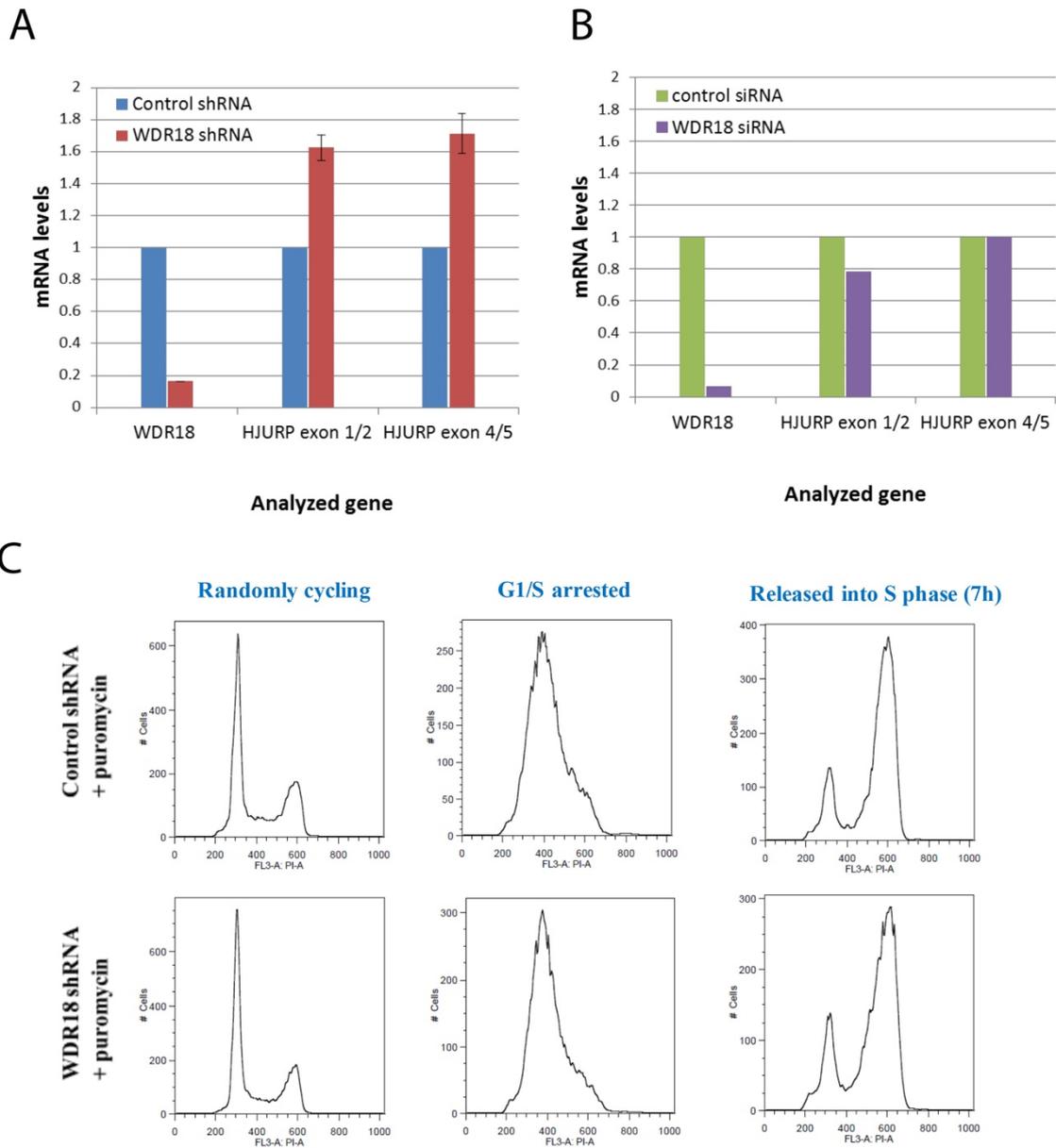
Figure 5.3. The effects of WDR18 depletion on cell cycle progression and HJURP transcription

(A)(B) HeLa cells were transfected with control or WDR18 shRNA (A) or siRNA (B) for 72 hours. Cells were harvested and the RNA was isolated following by real time PCR analysis of the indicated genes. The mRNA levels were normalized to IL8 gene. **(C)** HeLa cells were transfected with control or WDR18 shRNA for 72 hours and synchronized with double thymidine block and release. Cells were harvested at indicated cell cycle stages and stained with PI, following by FACS analysis.

WDR18 possibly regulates HJURP stability during interphase

Our observations from previous experiments suggest that WDR18 may play a role in a positive regulation of HJURP cellular levels. In order to test that we employed a cycloheximide pulse chase assay to monitor the turnover rates of CENP-A and HJURP proteins (Fig 5.4 A). HeLa cells were treated with WDR18 siRNA or WDR18 ShRNA for 72 hours following by treatment with cycloheximide for 0, 1/2, 1, 3 or 6 hours. Cells were then harvested and subsequently analyzed by Western blot with use of anti-HJURP, anti-NPM1 or anti-tubulin antibodies (Fig 5.4 A, B). Upon shRNA mediated WDR18 depletion the HJURP protein levels were decreased, and undetectable after 6 hours of cycloheximide treatment, whereas the levels of NPM1 were not significantly affected (Fig 5.4 A). In contrast, upon siRNA mediated WDR18 depletion the HJURP protein levels were unchanged in condition where no cycloheximide treatment was included (Fig 5.4 B). We did observe however, a higher turnover rate of HJURP protein in response to the cycloheximide treatment when WDR18 was depleted with the siRNA (Fig 5.4 B). We repeated these experiments including 3 hour cycloheximide treatment alone or in combination with MG132 proteasome inhibitor. These experiments confirmed that HJURP protein levels were less sensitive to WDR18 siRNA treatment when compared to WDR18 shRNA treatment, even though both approaches allowed as to deplete WDR18 mRNA levels to a similar degree (Fig 5.4 C, D) (Fig 5.3 A, B). These observations raised our concerns regarding the role of WDR18 in HJURP stability and whether these effects are specific to WR18 depletion.

Since the mitocheck database demonstrates the association of HJURP with WDR18 during mitosis, we asked whether HJURP levels are sensitive to WDR18 depletion specifically in mitosis. We turned to the cycloheximide pulse chase assay to monitor HJURP protein levels specifically in mitosis. HeLa cells were treated with

WDR18 shRNA for 72 hours following by treatment with or without nocodazole for 17 hours. In addition, we also treated the cells with cycloheximide alone or in combination with MG132 proteasome inhibitor. The western blot analysis of HJURP levels in these cells revealed that HJURP was subjected to proteasomal mediated degradation in asynchronous cells but not mitosis (Fig 5.4 E).
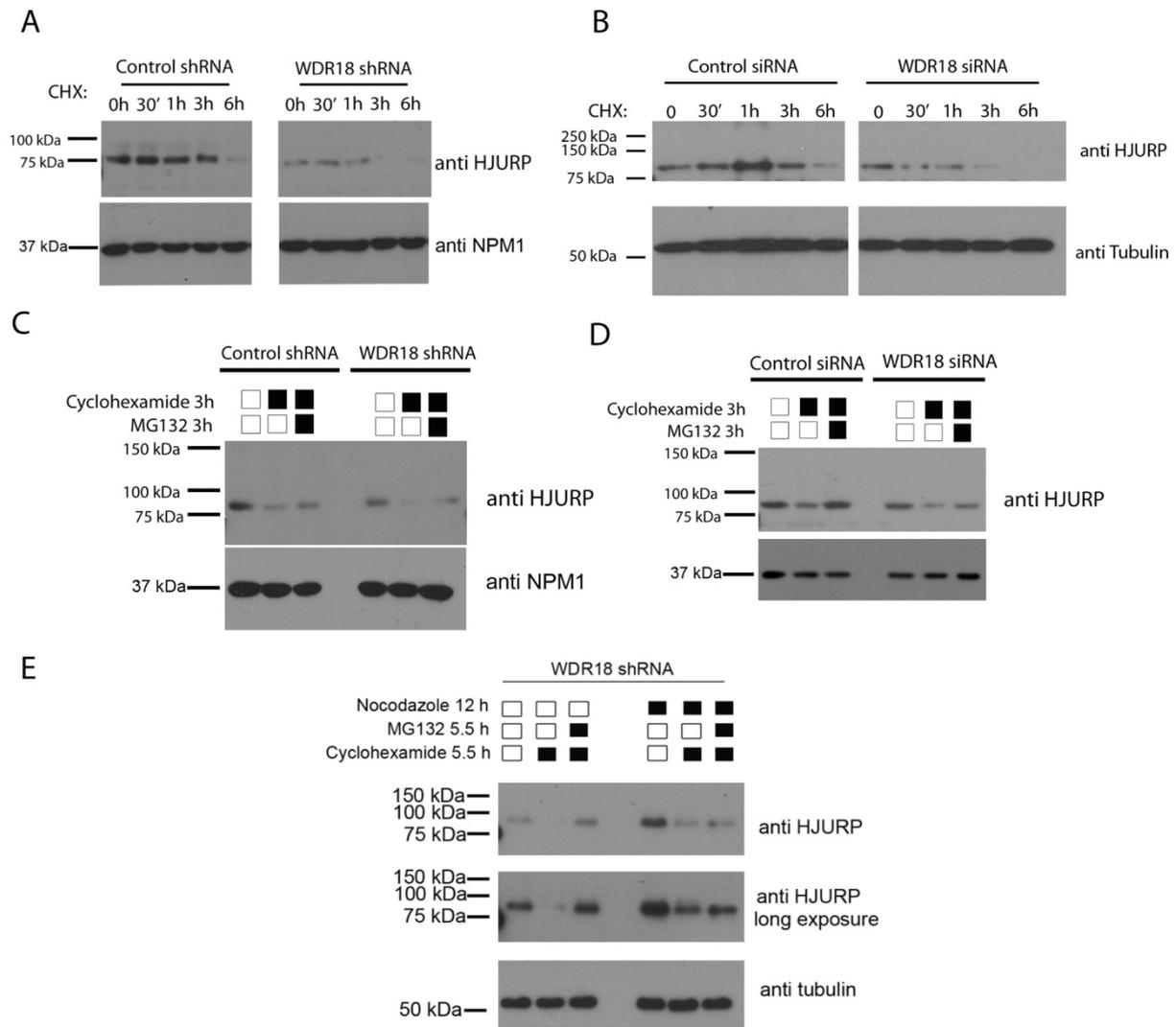
Figure 5.4. WDR18 possibly regulates HJURP stability during interphase

(A)(B) HeLa cells were transfected with control or WDR18 shRNA (A) and siRNA (B) for 72 hours and subsequently treated with cycloheximide for 0, 30, 60, 180 and 360 minutes. Whole cell extracts were analyzed by immunoblot using indicated antibodies. Anti-tubulin or anti NPM1 antibody staining were used as a loading control. **(C)(D)** HeLa cells were transfected with control or WDR18 shRNA (C) and siRNA (D) for 72 hours and subsequently treated with cycloheximide for 180 minutes with or without MG132. Whole cell extracts were analyzed by immunoblot using indicated antibodies. Anti NPM1 antibody staining was used as a loading control. **(E)** HeLa cells were transfected with WDR18 shRNA for 72 hours and treated with cycloheximide, MG132 and nocodazole as indicated. Whole cell extracts were analyzed by immunoblot using indicated antibodies.

WDR18 depletion possibly impairs CENP-A and Mis18α centromeric localization

We then asked whether WDR18 plays a role in the CENP-A deposition pathway. In order to address this question, we used shRNA mediated depletion of WDR18 in HeLa cells, and 72 hours after treatment, we analyzed the levels of endogenous CENP-A at the centromeres by immunofluorescence microscopy (Fig 5.5 A). A scrambled shRNA was used as a negative control. WDR18 depletion resulted in statistically significant decrease in endogenous CENP-A, but not CENP-T, levels at the centromere when compared to the control in one of our experiments (Fig 5.5 A, B-left panel). The replicate of this experiment did not demonstrate similar trend (Fig 5.5 C -left panel).

Given that WDR18 depletion results in HJURP degradation we aimed to ask whether defects in CENP-A localization that we saw in one of our experiments can be rescued by HJURP overexpression. We employed shRNA mediated depletion of WDR18 in cells overexpressing either GFP-HJURP and analyzed CENP-A levels at the centromere in response to the shRNA treatment. The CENP-A levels at the centromere observed in parental cells treated with WDR18 shRNA could not be rescued by HJURP overexpression (Fig 5.5 B-right panel). The replicate of this experiment did not demonstrate similar trend (Fig 5.5 C-right panel).

We also wanted to confirm these results using siRNA mediated depletion of WDR18. HeLa cells or cells overexpressing GFP-HJURP were treated with either control or WDR18 siRNA for 72 hours and subsequently fixed and stained with anti CENP-A in order to analyze CENP-A levels at the centromere. The siRNA mediated WDR18 depletion did not affect CENP-A levels at the centromere in these cells when compared to the control siRNA treatment. The HJURP overexpression did result slightly in increased CENP-A levels at the centromere (Fig 5.5 D).

We also tested whether depletion of WDR18 affected the centromeric localization of the Mis18 complex. To assess this, we used a cell line stably expressing Mis18α fused to a LAP tag and analyzed the centromeric localization of Mis18α-LAP in cells treated with shRNA against WDR18 (Fig 3D). WDR18 depletion did not significantly affect the number of cells with Mis18α localized at the centromere, but resulted in statistically significant decrease of the intensity of Mis18α-LAP at the centromere (Fig 5.5 E, F). However, we need to verify these results with the use of different approach to downregulate WDR18.
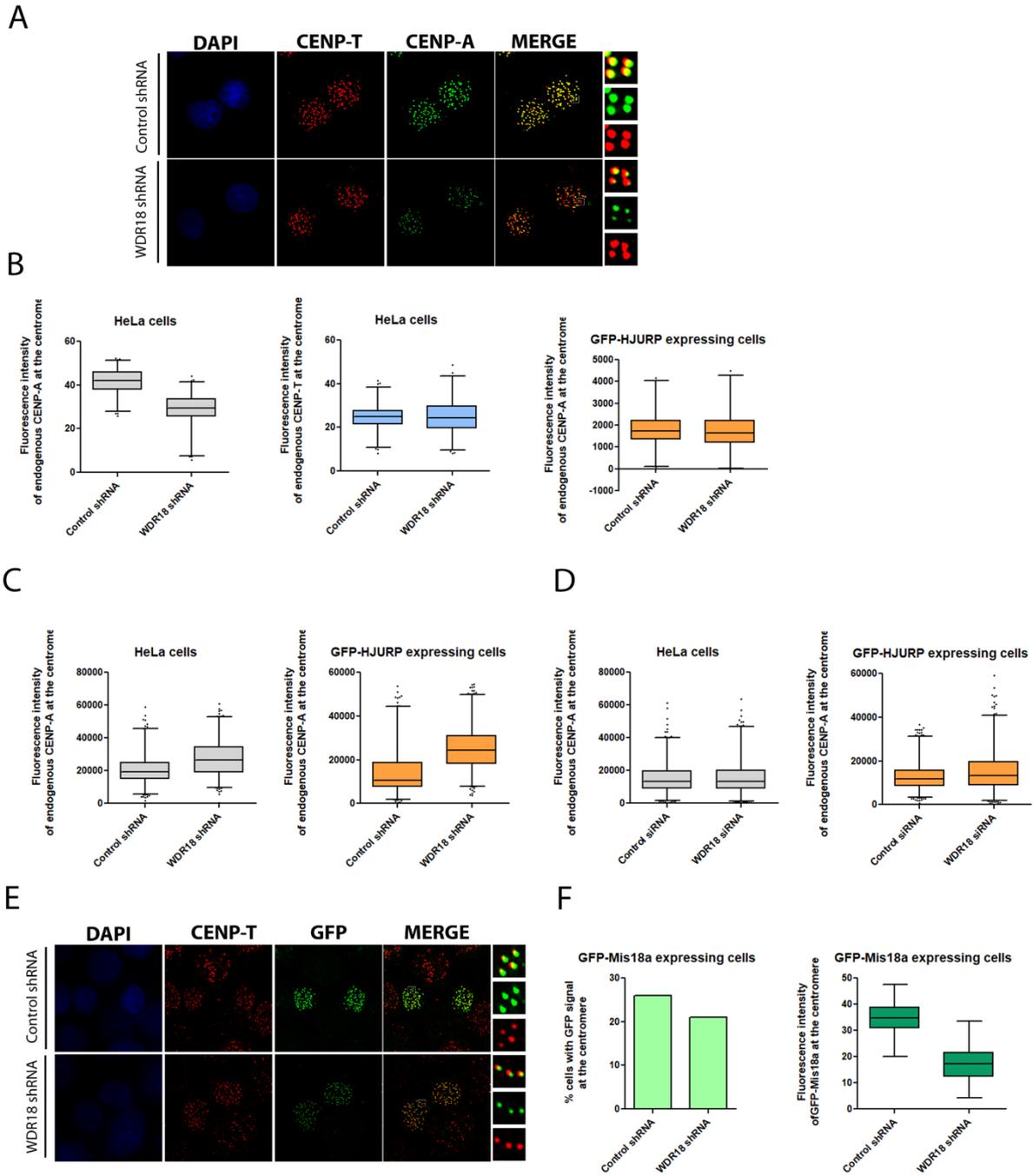
Figure 5.5. WDR18 depletion possibly impairs CENP-A and Mis18α centromeric localization

Figure 5.5. WDR18 depletion possibly impairs CENP-A and Mis18α centromeric localization

(A) Representative images of HeLa cells transfected with either control or WDR18 shRNA for 72 hours. DNA is visualized by DAPI staining, CENP-T is shown in red and endogenous CENP-A is shown in green. (B) Quantification of the integrated intensity of endogenous CENP-A or CENP-T at the centromere in HeLa or GFP-HJURP expressing cells treated with control or WDR18 shRNA. (C)(D) Quantification of the integrated intensity of endogenous CENP-A at the centromere in HeLa or GFP-HJURP expressing cells treated with control or WDR18 shRNA (C) or siRNA (D). (E) Representative images of HeLa cells stably expressing Mis18α-LAP transfected with either control or WDR18 shRNA for 72 hours. DNA is visualized by DAPI staining, CENP-T is shown in red and GFP-Mis18α is shown in green. (F) Quantification of percent cells with GFP positive centromeres treated as indicated in (E) (G) Quantification of the integrated intensity of Mis18α-LAP at the centromere in cells treated as indicated in panel (E).

WDR18 influence on the status of HJURP posttranslational modifications.

Our experiments suggest that the WDR18 protein potentially contributes to the mechanism controlling the stability of HJURP in the preassembly complex. WDR18 was previously demonstrated to interact with both desumoylating proteins and ubiquitin ligases. Therefore, we hypothesized that WDR18 possibly contributes to the posttranslational modification status of HJURP, and in turn influences the CENP-A deposition pathway. In order to test this hypothesis, we treated cells with the control or WDR18 shRNA for 48 hours and subsequently transfected with either control or Flag-ubiquitin encoding plasmid DNA (Fig 5.6 A,B). To prevent degradation of poliubiquinylated proteins we also included a condition where transfected cells were treated with the MG132 proteasome inhibitor for 6 hours. Cells were then harvested and subjected to Flag immunocprecipitation followed by Western blot analysis. When we used anti HJURP antibody to analyze the Western blot, a single band of a size consistent to monoubiquitynylated HJURP was detected in samples derived from both control and WDR18 shRNA treated cells. The band likely represents monoubiquitynylated HJURP as it was detected solely in IP fractions derived from cells transfected with Flag-ubiquitin, and was absent in samples obtained from untransfected cells Fig 5.6 B). Although we detected higher levels of the monoubiquitynylated HJURP in samples treated with MG132 and where WDR18 was depleted, we did not detect any poliubiqutinylated forms of HJURP (Fig 5.6 B). These results do not support the hypothesis that the presence of WDR18 prevents proteasome mediated degradation of HJURP. Interestingly however, the presence of WDR18 contributes to levels of monoubiquitynylated form of HJURP in MG132-dependent manner. This result suggests that perhaps WDR18 governs the stability of an unknown protein factor involved in HJURP monoubiquitynylation.

We also wanted to test whether WDR18 contributes to the HJURP sumoylation status. In order to address this question, we treated cells with the control or WDR18 shRNA for 48

hours and subsequently transfected with either control or His-Flag-SUMO1 and His-Flag-SUMO2 encoding plasmids (Fig 5.6 C, D). To prevent potential degradation of sumoylated proteins we also included a condition where transfected cells were treated with the MG132 proteasome inhibitor for 6 hours. Cells were then harvested and subjected to His pull down, followed by Western blot analysis. We used anti HJURP antibody to analyze the pull-down fractions by Western blot. A single band of a size consistent with endogenous HJURP was detected in all tested conditions, suggesting unspecific binding to the beads (Fig 5.6 D). We did not detect any bands consistent with modified forms of HJURP in this experiment. We therefore cannot conclude whether WDR18 contributes to HJURP sumoylation status based on this experiment.
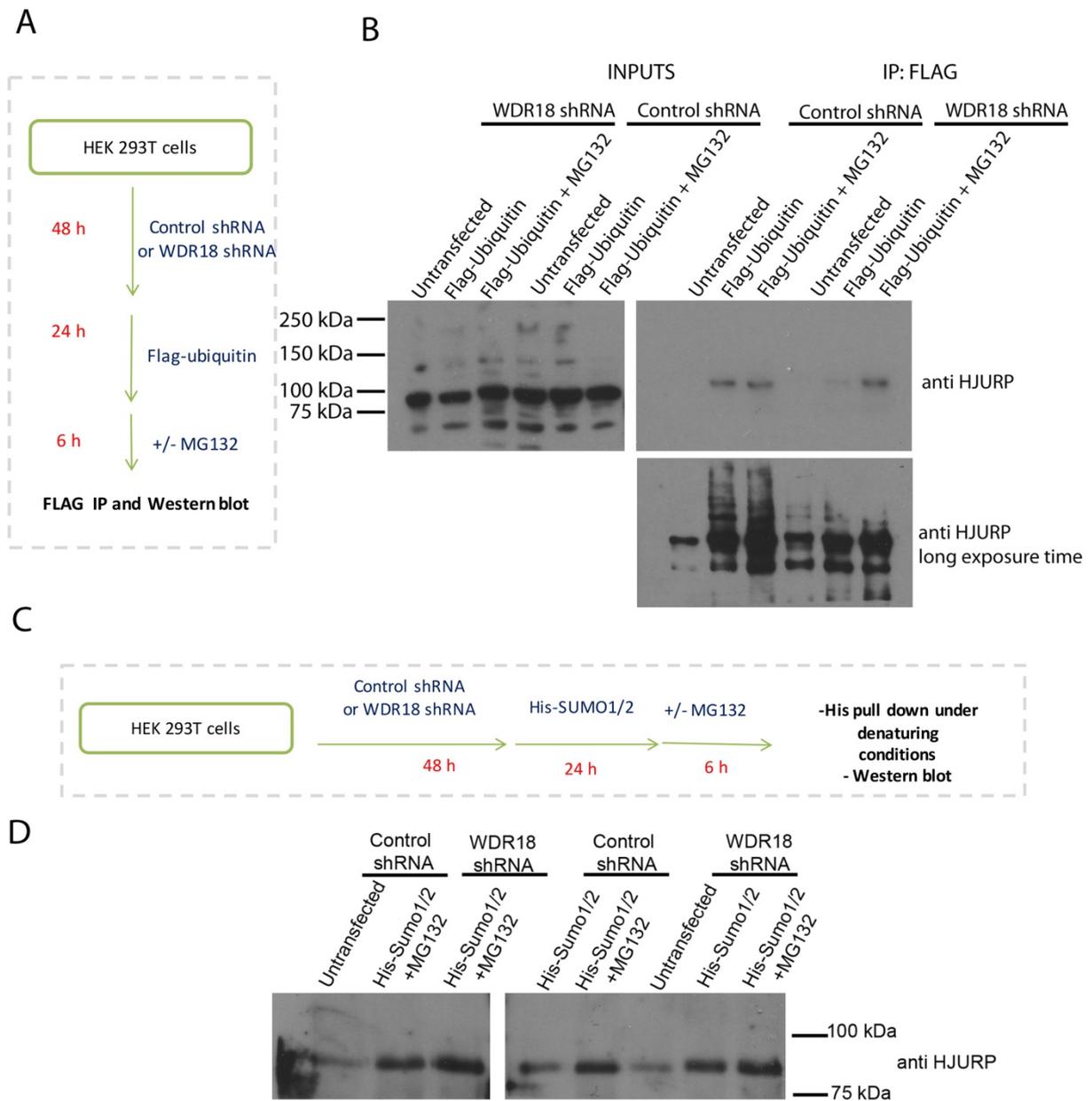
Figure 5.6 WDR18 influence on the status of HJURP posttranslational modifications
(A)(C) Schematic representation of experimental treatment in panel B and D, respectively.
(B)(D) Cells treated as indicated in panel A and C, respectively were used as an input for anti-Flag (B) or His (C) pull down and analyzed by Western blot using indicated antibodies.

Discussion

The function of proteins involved in CENP-A deposition pathway, such as HJURP and Mis18 complex, is cell cycle regulated. It is hypothesized that the mechanism regulating these protein factors involves the control of protein stability. The protein levels of both HJURP and components of the Mis18 complex increase prior to and during CENP-A deposition, and are significantly reduced throughout the rest of the cell cycle. A mechanism controlling Mis18β turnover has been proposed. The regulation of Mis18β stability involves Cul1 mediated ubiquitylation and subsequent proteasomal degradation of Mis18β during interphase but not mitosis. Importantly the MIS18BP1 subunit of the was shown recently to be co-modified by SUMO and ubiquitin and these modifications were enriched specifically during mitosis and implicated in Mis18BP1 degradation via RNF4-dependent pathway, following mitotic exit. In contrast to the Mis18 complex components, it remains unclear what mechanism controls HJURP protein levels in humans.

Two independent mass spectrometry screens revealed that WDR18 protein is associated with the Mis18 complex during interphase, and HJURP during mitosis. These interactions were confirmed by our immunoprecipitation experiments. We demonstrated that shRNA mediated depletion of WDR18 in randomly cycling cells reduces cellular levels of HJURP and centromeric levels of Mis18α which ultimately affects the CENP-A levels at the centromere. It is important to note that the effects of WDR18 depletion on centromeric CENP-A need to be validated, as we had difficulties with replicating this result. We also observed that cells treated with shRNA against WDR18 show rapid reduction of HJURP when protein synthesis is blocked with cycloheximide treatment. The BioGrid database demonstrates the association of WDR18 with the Cul3 ubiquitin ligase, USPU2 (ubiquitin specific protease) and there are several studies indicating an interaction between WDR18

with SENP3 (Castle et al., 2012; Fanis et al., 2012; Finkbeiner et al., 2011). These data allowed us to form the hypothesis that WDR18 contributes to the status of posttranslational modifications of HJURP and perhaps also the Mis18α protein. WDR18 might contribute to ubiquitylation or sumoylation of HJURP, and these events could protect HJURP from degradation.

In contrast to the shRNA mediated downregulation of WDR18, the siRNA treatment, in spite efficient WDR18 depletion, did not result in the same phenotype. When compared to the control siRNA, WDR18 siRNA treatment did not result in loss of centromeric CENP-A and overall did not affect HJURP protein stability significantly. These results suggest that the effects on centromeric CENP-A levels and HJURP stability that we observed in our shRNA experiments might not be specific to WDR18 depletion. These observations need to be verified by using another strategy for WDR18 depletion, such as CRISPR Cas9 mediated gene knockout or auxin-based degron system for the rapid depletion of proteins.

Our shRNA experiments suggest that WDR18 depletion does not increase the HJURP polyubiquitination, which is not consistent with the hypothesis that WDR18 controls HJURP stability. Our data indicate however, that the presence of WDR18 contributes to HJURP monoubiquitination status in the MG132-dependent manner. This result suggests that perhaps WDR18 governs the stability of an unknown protein factor involved in HJURP monoubiquitination. This event in turn might be involved the regulation of HJURP activity in the CENP-A deposition or inheritance. These results also need to be verified by replicating the experiment and by using another strategy for WDR18 depletion.

Our experiments clearly demonstrate that HJURP is subjected to monoubiquitination, however, it still remains to be testes what site within HJURP is modified, and more importantly what is the role of this modification. If HJURP monoubiquitination contributes to

its stability it would be critical to determine at what stage during the cell cycle this modification is placed and how it affects HJURP function. It also remains to be tested what is the mechanism that facilitates HJURP monoubiquitination and which protein modifiers and their accessory factors play critical role in that process.

In addition to interacting with HJURP, WDR18 was found to be associated with the Mis18α/β complex in our mass spectrometry screen and immunoprecipitation experiments. It remains to be tested whether HJURP and the Mis18 complex interact with WDR18 independently or form a large complex. It still remains unclear whether WDR18 affects the posttranslational modification status of the Mis18 complex components. The WDR18 depletion with the shRNA treatment did not alter Mis18 complex protein levels but resulted in decrease of Mis18α at the centromeres. Although these results need to be verified by another strategy for WDR18 depletion, this data suggests that WDR18 might play a regulatory role in the function of the Mis18 complex.

Materials and Methods

Cell Culture and Transfection. HeLa T-rex cells and lines derived from this parental cell lines were cultured in a 37 °C incubator in 5% CO_2 in DMEM supplemented with 10% FBS (OPTIMA) and 1% Pen/Strep. Cells for the DNA transfection were seeded onto six-well plate prior to transfection and transfected when they reached 60% of confluency. DNA transfection was conducted with the Lipofectamine 2000 reagent using standard protocol (Thermo Fisher Scientific) with 2ug of plasmid DNA per well. siRNA transfection was performed using the RNAiMAX transfection reagent (Invitrogen) using standard protocol. Cells were treated with either 2.5nM of WDR18 siRNA (Dharmacon) or GAPD control siRNA (Invitrogen) for 38-40 hours.

Cell synchronization. Cells were synchronized with double thymidine block and release. Thymidine was added to culture medium at 20mM for 18 hours. Following the first thymidine treatment cells were washed twice with PBS and released into S phase for 8 hours, and subsequently treated with second thymidine block. Cells were released form the second thymidine arrest as indicated in the text.

Cycloheximide Chase Assay. Cells were cultured and treated as indicated in the figure legends. The cycloheximide was used at 100 μg/ml concentration and MG132 was used at 10μM concentration for indicated time.

DNA content analysis. Cells were synchronized and subsequently harvested using PBS supplemented with 3 mM EDTA. Cells were then washed with PBS and spun down at 1000 rpm for 5 min. Cell pellets were resuspended in 200ul PBS, fixed with 5mls of 70% Ethanol

and stored at 4°C. Fixed cells were centrifuged at 1600 rpm for 5 min and washed with PBS + 1% FBS. Cell pellets were then resuspended in fresh PI/RNaseA solution (10ug/ml propidium iodide, 250ug/ml RNase A in PBS + 1% FBS) and incubated at 37°C for 30 minutes. Samples were analyzed for their DNA content using flow cytometry.

Imaging and Quantification. Images were acquired using a × 100 oil-immersion Olympus objective lens on a DeltaVision microscope. Collected images are demonstrated as maximum stacked images. Images were subjected to deconvolution prior stacking. Integrated intensities were derived from raw images subjected to ImageJ (using the CRAQ plugin) and the centromere marker was used as a reference. Quantification data was analyzed in GraphPad Prism software, the statistical significance was assessed using unpaired t-test. The graphs were generated using GraphPad Prism software and displayed percentiles are as indicated in figure legends.

Stable cell lines. All cell lines expressing LAP fusion proteins were generated using Flp-In™ T-REx™ System.

Indirect immunofluorescence. Cells were pre-extracted with 0.1% Triton-X in PBS for 3 minutes, fixed with 4% formaldehyde for 10 minutes and subsequently quenched with 100mM Tris, pH 7.5 for 5 minutes. Fixed cells were incubated in blocking solution (0.1% Triton-X in PBS, 0.2% BSA, 2% FBS) for 1.5 h at RT, and incubated with indicated primary antibodies for 1.5 h. Anti-CENP-T, anti-CENP-A and anti-HA antibodies were used at 1:5000, 1:1000, and 1:1000 dilution, respectively and detected using fluorophore conjugated secondary antibodies (Cy3, Cy5 or FITC, Jackson Immuno Inc.). DNA was visualized with 0.2mg/ml DAPI in PBS and coverslips were mounted in Prolong Gold (Invitrogen).

Immunoprecipitation. Cells were harvested stored at -80C. In case of overexpression experiments, cells were harvested 24 h post transfection and stored at -80C. Cell pellets were thawed on ice and resuspended in RIPA lysis buffer (150mM NaCl, 1% NP-40, 0.3% deoxycholate, 0.15% SDS, 50mM Tris HCl pH 7.5, 1mM EDTA, 10% glycerol, protease inhibitors, 0.1mM PMSF, 5mM NaF, 10mM β-Glycerophosphate, 0.2mM NaV). Lysates were centrifuged at 10000 rpm for 5 min at 4°C, and pre-cleared with Protein A agarose (Biorad) for 1 h at 4°C. Precleared extracts were then supplemented with indicated antibody (1:1000, Cell Signaling) and incubated for 17 hours at 4°C. Extracts were subsequently incubated with Protein A Dynabeads (Invitrogen) on ice for 45 min, and washed once with RIPA buffer, followed by three washes in PBST (PBS + 0.1% Tween). Purified proteins were eluted by boiling in SDS sample buffer for 10 minutes.

References

Abid Ali, F., and Costa, A. (2016). The MCM Helicase Motor of the Eukaryotic Replisome. Journal of molecular biology *428*, 1822-1832.

Alabert, C., Bukowski-Wills, J.C., Lee, S.B., Kustatscher, G., Nakamura, K., de Lima Alves, F., Menard, P., Mejlvang, J., Rappsilber, J., and Groth, A. (2014). Nascent chromatin capture proteomics determines chromatin dynamics during DNA replication and identifies unknown fork components. Nature cell biology *16*, 281-293.

Alabert, C., and Groth, A. (2012). Chromatin replication and epigenome maintenance. Nature reviews Molecular cell biology *13*, 153-167.

Allshire, R.C., and Karpen, G.H. (2008). Epigenetic regulation of centromeric chromatin: old dogs, new tricks? Nat Rev Genet *9*, 923-937.

Amano, M., Suzuki, A., Hori, T., Backer, C., Okawa, K., Cheeseman, I.M., and Fukagawa, T. (2009). The CENP-S complex is essential for the stable assembly of outer kinetochore structure. The Journal of cell biology *186*, 173-182.

An, S., Kim, H., and Cho, U.S. (2015). Mis16 Independently Recognizes Histone H4 and the CENP-ACnp1-Specific Chaperone Scm3sp. J Mol Biol *427*, 3230-3240.

Annunziato, A.T., Schindler, R.K., Thomas, C.A., Jr., and Seale, R.L. (1981). Dual nature of newly replicated chromatin. Evidence for nucleosomal and non-nucleosomal DNA at the site of native replication forks. The Journal of biological chemistry *256*, 11880-11886.

Aparicio, O.M., Weinstein, D.M., and Bell, S.P. (1997). Components and dynamics of DNA replication complexes in S. cerevisiae: redistribution of MCM proteins and Cdc45p during S phase. Cell *91*, 59-69.

Araki, H. (2010). Cyclin-dependent kinase-dependent initiation of chromosomal DNA replication. Current opinion in cell biology *22*, 766-771.

Bailey, A.O., Panchenko, T., Sathyan, K.M., Petkowski, J.J., Pai, P.J., Bai, D.L., Russell, D.H., Macara, I.G., Shabanowitz, J., Hunt, D.F.*, et al.* (2013). Posttranslational modification of CENP-A influences the conformation of centromeric chromatin. Proc Natl Acad Sci U S A *110*, 11827-11832.

Balakrishnan, L., and Bambara, R.A. (2013). Flap endonuclease 1. Annual review of biochemistry *82*, 119-138.

Barnhart, M.C., Kuich, P.H., Stellfox, M.E., Ward, J.A., Bassett, E.A., Black, B.E., and Foltz, D.R. (2011). HJURP is a CENP-A chromatin assembly factor sufficient to form a functional de novo kinetochore. The Journal of cell biology *194*, 229-243.

Barnhart-Dailey, M.C., Trivedi, P., Stukenberg, P.T., and Foltz, D.R. (2016). HJURP interaction with the condensin II complex during G1 promotes CENP-A deposition. Molecular biology of the cell.

Bassett, E.A., DeNizio, J., Barnhart-Dailey, M.C., Panchenko, T., Sekulic, N., Rogers, D.J., Foltz, D.R., and Black, B.E. (2012). HJURP uses distinct CENP-A surfaces to recognize and to stabilize CENP-A/histone H4 for centromere assembly. Dev Cell *22*, 749-762.

Bell, S.P., and Labib, K. (2016). Chromosome Duplication in Saccharomyces cerevisiae. Genetics *203*, 1027-1067.

Bell, S.P., and Stillman, B. (1992). ATP-dependent recognition of eukaryotic origins of DNA replication by a multiprotein complex. Nature *357*, 128-134.

Belotserkovskaya, R., Oh, S., Bondarenko, V.A., Orphanides, G., Studitsky, V.M., and Reinberg, D. (2003). FACT facilitates transcription-dependent nucleosome alteration. Science (New York, NY) *301*, 1090-1093.

Bergmann, J.H., Rodriguez, M.G., Martins, N.M., Kimura, H., Kelly, D.A., Masumoto, H., Larionov, V., Jansen, L.E., and Earnshaw, W.C. (2011). Epigenetic engineering shows H3K4me2 is required for HJURP targeting and CENP-A assembly on a synthetic human kinetochore. The EMBO journal *30*, 328-340.

Bernad, R., Sanchez, P., Rivera, T., Rodriguez-Corsino, M., Boyarchuk, E., Vassias, I., Ray-Gallet, D., Arnaoutov, A., Dasso, M., Almouzni, G.*, et al.* (2011). Xenopus HJURP and condensin II are required for CENP-A assembly. The Journal of cell biology *192*, 569-582.

Black, B.E., Brock, M.A., Bedard, S., Woods, V.L., Jr., and Cleveland, D.W. (2007). An epigenetic mark generated by the incorporation of CENP-A into centromeric nucleosomes. Proc Natl Acad Sci U S A *104*, 5008-5013.

Black, B.E., and Cleveland, D.W. (2011). Epigenetic centromere propagation and the nature of CENP-a nucleosomes. Cell *144*, 471-479.

Blower, M.D., Sullivan, B.A., and Karpen, G.H. (2002). Conserved organization of centromeric chromatin in flies and humans. Dev Cell *2*, 319-330.

Bochman, M.L., and Schwacha, A. (2007). Differences in the single-stranded DNA binding activities of MCM2-7 and MCM467: MCM2 and MCM5 define a slow ATP-dependent step. The Journal of biological chemistry *282*, 33795-33804.

Bochman, M.L., and Schwacha, A. (2009). The Mcm complex: unwinding the mechanism of a replicative helicase. Microbiology and molecular biology reviews : MMBR *73*, 652-683.

Boddy, M.N., Furnari, B., Mondesert, O., and Russell, P. (1998). Replication checkpoint enforced by kinases Cds1 and Chk1. Science (New York, NY) *280*, 909-912.

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

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

Boltengagen, M., Huang, A., Boltengagen, A., Trixl, L., Lindner, H., Kremser, L., Offterdinger, M., and Lusser, A. (2016). A novel role for the histone acetyltransferase Hat1 in the CENP-A/CID assembly pathway in Drosophila melanogaster. Nucleic Acids Res *44*, 2145-2159.

Boos, D., Frigola, J., and Diffley, J.F. (2012). Activation of the replicative DNA helicase: breaking up is hard to do. Current opinion in cell biology *24*, 423-430.

Bouzinba-Segard, H., Guais, A., and Francastel, C. (2006). Accumulation of small murine minor satellite transcripts leads to impaired centromeric architecture and function. Proc Natl Acad Sci U S A *103*, 8709-8714.

Boyarchuk, E., Montes de Oca, R., and Almouzni, G. (2011). Cell cycle dynamics of histone variants at the centromere, a model for chromosomal landmarks. Current opinion in cell biology *23*, 266-276.

Brown, K.D., and Robertson, K.D. (2007). DNMT1 knockout delivers a strong blow to genome stability and cell viability. Nat Genet *39*, 289-290.

Bruck, I., Perez-Arnaiz, P., Colbert, M.K., and Kaplan, D.L. (2015). Insights into the Initiation of Eukaryotic DNA Replication. Nucleus (Austin, Tex) *6*, 449-454.

Budd, M.E., Choe, W.C., and Campbell, J.L. (1995). DNA2 encodes a DNA helicase essential for replication of eukaryotic chromosomes. The Journal of biological chemistry *270*, 26766-26769.

Bui, M., Dimitriadis, E.K., Hoischen, C., An, E., Quenet, D., Giebe, S., Nita-Lazar, A., Diekmann, S., and Dalal, Y. (2012). Cell-cycle-dependent structural transitions in the human CENP-A nucleosome in vivo. Cell *150*, 317-326.

Burgers, P.M. (2009). Polymerase dynamics at the eukaryotic DNA replication fork. The Journal of biological chemistry *284*, 4041-4045.

Burgers, P.M.J., and Kunkel, T.A. (2017). Eukaryotic DNA Replication Fork. Annual review of biochemistry *86*, 417-438.

Burgess, R.J., and Zhang, Z. (2013). Histone chaperones in nucleosome assembly and human disease. Nat Struct Mol Biol *20*, 14-22.

Camahort, R., Li, B., Florens, L., Swanson, S.K., Washburn, M.P., and Gerton, J.L. (2007). Scm3 is essential to recruit the histone h3 variant cse4 to centromeres and to maintain a functional kinetochore. Mol Cell *26*, 853-865.

Camahort, R., Shivaraju, M., Mattingly, M., Li, B., Nakanishi, S., Zhu, D., Shilatifard, A., Workman, J.L., and Gerton, J.L. (2009). Cse4 is part of an octameric nucleosome in budding yeast. Mol Cell *35*, 794-805.

Campos, E.I., Fillingham, J., Li, G., Zheng, H., Voigt, P., Kuo, W.H., Seepany, H., Gao, Z., Day, L.A., Greenblatt, J.F.*, et al.* (2010). The program for processing newly synthesized histones H3.1 and H4. Nat Struct Mol Biol *17*, 1343-1351.

Cardozo, T., and Pagano, M. (2004). The SCF ubiquitin ligase: insights into a molecular machine. Nature reviews Molecular cell biology *5*, 739-751.

Carone, D.M., Longo, M.S., Ferreri, G.C., Hall, L., Harris, M., Shook, N., Bulazel, K.V., Carone, B.R., Obergfell, C., O'Neill, M.J.*, et al.* (2009). A new class of retroviral and satellite encoded small RNAs emanates from mammalian centromeres. Chromosoma *118*, 113-125.

Carone, D.M., Zhang, C., Hall, L.E., Obergfell, C., Carone, B.R., O'Neill, M.J., and O'Neill, R.J. (2013). Hypermorphic expression of centromeric retroelement-encoded small RNAs impairs CENP-A loading. Chromosome Res *21*, 49-62.

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

Castle, C.D., Cassimere, E.K., and Denicourt, C. (2012). LAS1L interacts with the mammalian Rix1 complex to regulate ribosome biogenesis. Molecular biology of the cell *23*, 716-728.

Chan, F.L., Marshall, O.J., Saffery, R., Won Kim, B., Earle, E., Choo, K.H., and Wong, L.H. (2012). Active transcription and essential role of RNA polymerase II at the centromere during mitosis. Proc Natl Acad Sci U S A *109*, 1979-1984.

Cheeseman, I.M., and Desai, A. (2005). A combined approach for the localization and tandem affinity purification of protein complexes from metazoans. Sci STKE *2005*, pl1.

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

Chen, C.C., Bowers, S., Lipinszki, Z., Palladino, J., Trusiak, S., Bettini, E., Rosin, L., Przewloka, M.R., Glover, D.M., O'Neill, R.J.*, et al.* (2015). Establishment of Centromeric Chromatin by the CENP-A Assembly Factor CAL1 Requires FACT-Mediated Transcription. Dev Cell *34*, 73-84.

Chen, C.C., Dechassa, M.L., Bettini, E., Ledoux, M.B., Belisario, C., Heun, P., Luger, K., and Mellone, B.G. (2014). CAL1 is the Drosophila CENP-A assembly factor. The Journal of cell biology *204*, 313-329.

Chen, J., and Matthews, K.S. (1992). Deletion of lactose repressor carboxyl-terminal domain affects tetramer formation. J Biol Chem *267*, 13843-13850.

Chen, S., and Bell, S.P. (2011). CDK prevents Mcm2-7 helicase loading by inhibiting Cdt1 interaction with Orc6. Genes & development *25*, 363-372.

Chen, S., de Vries, M.A., and Bell, S.P. (2007). Orc6 is required for dynamic recruitment of Cdt1 during repeated Mcm2-7 loading. Genes & development *21*, 2897-2907.

Chen, Y., Caldwell, J.M., Pereira, E., Baker, R.W., and Sanchez, Y. (2009). ATRMec1 phosphorylation-independent activation of Chk1 in vivo. The Journal of biological chemistry *284*, 182-190.

Chen, Y., and Poon, R.Y. (2008). The multiple checkpoint functions of CHK1 and CHK2 in maintenance of genome stability. Frontiers in bioscience : a journal and virtual library *13*, 5016-5029.

Cho, U.S., and Harrison, S.C. (2011). Recognition of the centromere-specific histone Cse4 by the chaperone Scm3. Proc Natl Acad Sci U S A *108*, 9367-9371.

Choi, E.S., Stralfors, A., Castillo, A.G., Durand-Dubief, M., Ekwall, K., and Allshire, R.C. (2011). Identification of noncoding transcripts from within CENP-A chromatin at fission yeast centromeres. The Journal of biological chemistry *286*, 23600-23607.

Chong, J.P., Mahbubani, H.M., Khoo, C.Y., and Blow, J.J. (1995). Purification of an MCM-containing complex as a component of the DNA replication licensing system. Nature *375*, 418-421.

Chow, C.M., Georgiou, A., Szutorisz, H., Maia e Silva, A., Pombo, A., Barahona, I., Dargelos, E., Canzonetta, C., and Dillon, N. (2005). Variant histone H3.3 marks promoters of transcriptionally active genes during mammalian cell division. EMBO Rep *6*, 354-360.

Clarke, L., and Carbon, J. (1980). Isolation of a yeast centromere and construction of functional small circular chromosomes. Nature *287*, 504-509.

Cleveland, D.W., Mao, Y., and Sullivan, K.F. (2003). Centromeres and kinetochores: from epigenetics to mitotic checkpoint signaling. Cell *112*, 407-421.

Coleman, T.R., Carpenter, P.B., and Dunphy, W.G. (1996). The Xenopus Cdc6 protein is essential for the initiation of a single round of DNA replication in cell-free extracts. Cell *87*, 53-63.

Costa, A., Hood, I.V., and Berger, J.M. (2013). Mechanisms for initiating cellular DNA replication. Annual review of biochemistry *82*, 25-54.

Costa, A., Ilves, I., Tamberg, N., Petojevic, T., Nogales, E., Botchan, M.R., and Berger, J.M. (2011). The structural basis for MCM2-7 helicase activation by GINS and Cdc45. Nature structural & molecular biology *18*, 471-477.

Costa, A., Renault, L., Swuec, P., Petojevic, T., Pesavento, J.J., Ilves, I., MacLellan-Gibson, K., Fleck, R.A., Botchan, M.R., and Berger, J.M. (2014). DNA binding polarity, dimerization, and ATPase ring remodeling in the CMG helicase of the eukaryotic replisome. eLife *3*, e03273.

Cuijpers, S.A.G., Willemstein, E., and Vertegaal, A.C.O. (2017). Converging Small Ubiquitin-like Modifier (SUMO) and Ubiquitin Signaling: Improved Methodology Identifies Co-modified Target Proteins. Molecular & cellular proteomics : MCP *16*, 2281-2295.

Dambacher, S., Deng, W., Hahn, M., Sadic, D., Frohlich, J., Nuber, A., Hoischen, C., Diekmann, S., Leonhardt, H., and Schotta, G. (2012). CENP-C facilitates the recruitment of M18BP1 to centromeric chromatin. Nucleus *3*, 101-110.

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

Diffley, J.F. (2004). Regulation of early events in chromosome replication. Current biology : CB *14*, R778-786.

Diffley, J.F., and Cocker, J.H. (1992). Protein-DNA interactions at a yeast replication origin. Nature *357*, 169-172.

Diffley, J.F., Cocker, J.H., Dowell, S.J., Harwood, J., and Rowley, A. (1995). Stepwise assembly of initiation complexes at budding yeast replication origins during the cell cycle. Journal of cell science Supplement *19*, 67-72.

Doheny, K.F., Sorger, P.K., Hyman, A.A., Tugendreich, S., Spencer, F., and Hieter, P. (1993). Identification of essential components of the S. cerevisiae kinetochore. Cell *73*, 761-774.

Driscoll, R., Hudson, A., and Jackson, S.P. (2007). Yeast Rtt109 promotes genome stability by acetylating histone H3 on lysine 56. Science *315*, 649-652.

Dunleavy, E.M., Almouzni, G., and Karpen, G.H. (2011). H3.3 is deposited at centromeres in S phase as a placeholder for newly assembled CENP-A in G(1) phase. Nucleus *2*, 146-157.

Dunleavy, E.M., Pidoux, A.L., Monet, M., Bonilla, C., Richardson, W., Hamilton, G.L., Ekwall, K., McLaughlin, P.J., and Allshire, R.C. (2007). A NASP (N1/N2)-related protein, Sim3, binds CENP-A and is required for its deposition at fission yeast centromeres. Mol Cell *28*, 1029-1044.

Dunleavy, E.M., Roche, D., Tagami, H., Lacoste, N., Ray-Gallet, D., Nakamura, Y., Daigo, Y., Nakatani, Y., and Almouzni-Pettinotti, G. (2009). HJURP is a cell-cycle-dependent maintenance and deposition factor of CENP-A at centromeres. Cell *137*, 485-497.

Dyer, P.N., Edayathumangalam, R.S., White, C.L., Bao, Y., Chakravarthy, S., Muthurajan, U.M., and Luger, K. (2004). Reconstitution of nucleosome core particles from recombinant histones and DNA. Methods in enzymology *375*, 23-44.

Earnshaw, W., Bordwell, B., Marino, C., and Rothfield, N. (1986). Three human chromosomal autoantigens are recognized by sera from patients with anti-centromere antibodies. J Clin Invest *77*, 426-430.

English, C.M., Adkins, M.W., Carson, J.J., Churchill, M.E., and Tyler, J.K. (2006). Structural basis for the histone chaperone activity of Asf1. Cell *127*, 495-508.

English, C.M., Maluf, N.K., Tripet, B., Churchill, M.E., and Tyler, J.K. (2005). ASF1 binds to a heterodimer of histones H3 and H4: a two-step mechanism for the assembly of the H3-H4 heterotetramer on DNA. Biochemistry *44*, 13673-13682.

Enomoto, S., and Berman, J. (1998). Chromatin assembly factor I contributes to the maintenance, but not the re-establishment, of silencing at the yeast silent mating loci. Genes & development *12*, 219-232.

Enomoto, S., McCune-Zierath, P.D., Gerami-Nejad, M., Sanders, M.A., and Berman, J. (1997). RLF2, a subunit of yeast chromatin assembly factor-I, is required for telomeric chromatin function in vivo. Genes & development *11*, 358-370.

Erhardt, S., Mellone, B.G., Betts, C.M., Zhang, W., Karpen, G.H., and Straight, A.F. (2008). Genome-wide analysis reveals a cell cycle-dependent mechanism controlling centromere propagation. The Journal of cell biology *183*, 805-818.

Evrin, C., Clarke, P., Zech, J., Lurz, R., Sun, J., Uhle, S., Li, H., Stillman, B., and Speck, C. (2009). A double-hexameric MCM2-7 complex is loaded onto origin DNA during licensing of eukaryotic DNA replication. Proceedings of the National Academy of Sciences of the United States of America *106*, 20240-20245.

Eymery, A., Callanan, M., and Vourc'h, C. (2009). The secret message of heterochromatin: new insights into the mechanisms and function of centromeric and pericentric repeat sequence transcription. Int J Dev Biol *53*, 259-268.

Fachinetti, D.L., G.A., Abdullah, A., Selzer, E.B., Cleveland, D.W., Black, B.E. (2016). CENP-A modifications on Ser68 and Lys124 are dispensable for establishment, maintenance, and long-term function of human centromeres

. Dev Cell *In Press*.

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

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

Fangman, W.L., and Brewer, B.J. (1991). Activation of replication origins within yeast chromosomes. Annual review of cell biology *7*, 375-402.

Fanis, P., Gillemans, N., Aghajanirefah, A., Pourfarzad, F., Demmers, J., Esteghamat, F., Vadlamudi, R.K., Grosveld, F., Philipsen, S., and van Dijk, T.B. (2012). Five friends of methylated chromatin target of protein-arginine-methyltransferase[prmt]-1 (chtop), a complex linking arginine methylation to desumoylation. Molecular & cellular proteomics : MCP *11*, 1263-1273.

Feng, H., Zhou, Z., Zhou, B.R., and Bai, Y. (2011). Structure of the budding yeast Saccharomyces cerevisiae centromeric histones Cse4-H4 complexed with the chaperone Scm3. Proc Natl Acad Sci U S A *108*, E596.

Finkbeiner, E., Haindl, M., and Muller, S. (2011). The SUMO system controls nucleolar partitioning of a novel mammalian ribosome biogenesis complex. The EMBO journal *30*, 1067-1078.

Fitzgerald-Hayes, M., Clarke, L., and Carbon, J. (1982). Nucleotide sequence comparisons and functional analysis of yeast centromere DNAs. Cell *29*, 235-244.

Foltman, M., Evrin, C., De Piccoli, G., Jones, R.C., Edmondson, R.D., Katou, Y., Nakato, R., Shirahige, K., and Labib, K. (2013). Eukaryotic replisome components cooperate to process histones during chromosome replication. Cell Rep *3*, 892-904.

Foltz, D.R., Jansen, L.E., Bailey, A.O., Yates, J.R., 3rd, Bassett, E.A., Wood, S., Black, B.E., and Cleveland, D.W. (2009). Centromere-specific assembly of CENP-a nucleosomes is mediated by HJURP. Cell *137*, 472-484.

Foltz, D.R., Jansen, L.E., Black, B.E., Bailey, A.O., Yates, J.R., 3rd, and Cleveland, D.W. (2006). The human CENP-A centromeric nucleosome-associated complex. Nat Cell Biol *8*, 458-469.

Foss, M., McNally, F.J., Laurenson, P., and Rine, J. (1993). Origin recognition complex (ORC) in transcriptional silencing and DNA replication in S. cerevisiae. Science (New York, NY) *262*, 1838-1844.

Fox, C.A., Loo, S., Dillin, A., and Rine, J. (1995). The origin recognition complex has essential functions in transcriptional silencing and chromosomal replication. Genes & development *9*, 911-924.

Fujita, M., Hori, Y., Shirahige, K., Tsurimoto, T., Yoshikawa, H., and Obuse, C. (1998). Cell cycle dependent topological changes of chromosomal replication origins in Saccharomyces cerevisiae. Genes to cells : devoted to molecular & cellular mechanisms *3*, 737-749.

Fujita, Y., Hayashi, T., Kiyomitsu, T., Toyoda, Y., Kokubu, A., Obuse, C., and Yanagida, M. (2007). Priming of centromere for CENP-A recruitment by human hMis18alpha, hMis18beta, and M18BP1. Dev Cell *12*, 17-30.

Gambus, A., Jones, R.C., Sanchez-Diaz, A., Kanemaki, M., van Deursen, F., Edmondson, R.D., and Labib, K. (2006). GINS maintains association of Cdc45 with MCM in replisome progression complexes at eukaryotic DNA replication forks. Nature cell biology *8*, 358-366.

Gasman, S., Kalaidzidis, Y., and Zerial, M. (2003). RhoD regulates endosome dynamics through Diaphanous-related Formin and Src tyrosine kinase. Nat Cell Biol *5*, 195-204.

Geiss-Friedlander, R., and Melchior, F. (2007). Concepts in sumoylation: a decade on. Nature reviews Molecular cell biology *8*, 947-956.

Gerard, A., Koundrioukoff, S., Ramillon, V., Sergere, J.C., Mailand, N., Quivy, J.P., and Almouzni, G. (2006). The replication kinase Cdc7-Dbf4 promotes the interaction of the p150 subunit of chromatin assembly factor 1 with proliferating cell nuclear antigen. EMBO reports *7*, 817-823.

Goh, P.Y., and Kilmartin, J.V. (1993). NDC10: a gene involved in chromosome segregation in Saccharomyces cerevisiae. The Journal of cell biology *121*, 503-512.

Goldberg, A.D., Banaszynski, L.A., Noh, K.M., Lewis, P.W., Elsaesser, S.J., Stadler, S., Dewell, S., Law, M., Guo, X., Li, X.*, et al.* (2010). Distinct factors control histone variant H3.3 localization at specific genomic regions. Cell *140*, 678-691.

Gopalakrishnan, S., Sullivan, B.A., Trazzi, S., Della Valle, G., and Robertson, K.D. (2009). DNMT3B interacts with constitutive centromere protein CENP-C to modulate DNA methylation and the histone code at centromeric regions. Hum Mol Genet *18*, 3178-3193.

Goshima, G., Wollman, R., Goodwin, S.S., Zhang, N., Scholey, J.M., Vale, R.D., and Stuurman, N. (2007). Genes required for mitotic spindle assembly in Drosophila S2 cells. Science (New York, NY) *316*, 417-421.

Grasby, J.A., Finger, L.D., Tsutakawa, S.E., Atack, J.M., and Tainer, J.A. (2012). Unpairing and gating: sequence-independent substrate recognition by FEN superfamily nucleases. Trends in biochemical sciences *37*, 74-84.

Groth, A., Corpet, A., Cook, A.J., Roche, D., Bartek, J., Lukas, J., and Almouzni, G. (2007). Regulation of replication fork progression through histone supply and demand. Science *318*, 1928-1931.

Guo, L.Y., Allu, P.K., Zandarashvili, L., McKinley, K.L., Sekulic, N., Dawicki-McKenna, J.M., Fachinetti, D., Logsdon, G.A., Jamiolkowski, R.M., Cleveland, D.W.*, et al.* (2017). Centromeres are maintained by fastening CENP-A to DNA and directing an arginine anchor-dependent nucleosome transition. Nature communications *8*, 15775.

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

Haglund, K., and Dikic, I. (2005). Ubiquitylation and cell signaling. The EMBO journal *24*, 3353-3359.

Hagstrom, K.A., Holmes, V.F., Cozzarelli, N.R., and Meyer, B.J. (2002). C. elegans condensin promotes mitotic chromosome architecture, centromere organization, and sister chromatid segregation during mitosis and meiosis. Genes Dev *16*, 729-742.

Hall, L.E., Mitchell, S.E., and O'Neill, R.J. (2012). Pericentric and centromeric transcription: a perfect balance required. Chromosome Res *20*, 535-546.

Hammond, C.M., Stromme, C.B., Huang, H., Patel, D.J., and Groth, A. (2017). Histone chaperone networks shaping chromatin function. Nature reviews Molecular cell biology *18*, 141-158.

Han, J., Zhou, H., Horazdovsky, B., Zhang, K., Xu, R.M., and Zhang, Z. (2007). Rtt109 acetylates histone H3 lysine 56 and functions in DNA replication. Science *315*, 653-655.

Hay, R.T. (2007). SUMO-specific proteases: a twist in the tail. Trends in cell biology *17*, 370-376.

Hayashi, T., Ebe, M., Nagao, K., Kokubu, A., Sajiki, K., and Yanagida, M. (2014). Schizosaccharomyces pombe centromere protein Mis19 links Mis16 and Mis18 to recruit CENP-A through interacting with NMD factors and the SWI/SNF complex. Genes Cells *19*, 541-554.

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

Hewawasam, G., Shivaraju, M., Mattingly, M., Venkatesh, S., Martin-Brown, S., Florens, L., Workman, J.L., and Gerton, J.L. (2010). Psh1 is an E3 ubiquitin ligase that targets the centromeric histone variant Cse4. Mol Cell *40*, 444-454.

Hirano, T. (2005). Condensins: organizing and segregating the genome. Curr Biol *15*, R265-275.

Holland, A.J., Fachinetti, D., Han, J.S., and Cleveland, D.W. (2012). Inducible, reversible system for the rapid and complete degradation of proteins in mammalian cells. Proceedings of the National Academy of Sciences of the United States of America *109*, E3350-3357.

Hori, T., Shang, W.H., Takeuchi, K., and Fukagawa, T. (2013). The CCAN recruits CENP-A to the centromere and forms the structural core for kinetochore assembly. The Journal of cell biology *200*, 45-60.

Howell, B.J., Hoffman, D.B., Fang, G., Murray, A.W., and Salmon, E.D. (2000). Visualization of Mad2 dynamics at kinetochores, along spindle fibers, and at spindle poles in living cells. J Cell Biol *150*, 1233-1250.

Howes, T.R., and Tomkinson, A.E. (2012). DNA ligase I, the replicative DNA ligase. Sub-cellular biochemistry *62*, 327-341.

Hsiao, C.L., and Carbon, J. (1979). High-frequency transformation of yeast by plasmids containing the cloned yeast ARG4 gene. Proceedings of the National Academy of Sciences of the United States of America *76*, 3829-3833.

Hu, H., Liu, Y., Wang, M., Fang, J., Huang, H., Yang, N., Li, Y., Wang, J., Yao, X., Shi, Y., *et al.* (2011). Structure of a CENP-A-histone H4 heterodimer in complex with chaperone HJURP. Genes Dev *25*, 901-906.

Huang, H., Stromme, C.B., Saredi, G., Hodl, M., Strandsby, A., Gonzalez-Aguilera, C., Chen, S., Groth, A., and Patel, D.J. (2015). A unique binding mode enables MCM2 to chaperone histones H3-H4 at replication forks. Nat Struct Mol Biol *22*, 618-626.

Huttlin, E.L., Bruckner, R.J., Paulo, J.A., Cannon, J.R., Ting, L., Baltier, K., Colby, G., Gebreab, F., Gygi, M.P., Parzen, H., *et al.* (2017). Architecture of the human interactome defines protein communities and disease networks. Nature *545*, 505-509.

Hyman, A.A., Middleton, K., Centola, M., Mitchison, T.J., and Carbon, J. (1992). Microtubule-motor activity of a yeast centromere-binding protein complex. Nature *359*, 533-536.

Ilves, I., Petojevic, T., Pesavento, J.J., and Botchan, M.R. (2010). Activation of the MCM2-7 helicase by association with Cdc45 and GINS proteins. Molecular cell *37*, 247-258.

Izuta, H., Ikeno, M., Suzuki, N., Tomonaga, T., Nozaki, N., Obuse, C., Kisu, Y., Goshima, N., Nomura, F., Nomura, N., *et al.* (2006). Comprehensive analysis of the ICEN (Interphase Centromere Complex) components enriched in the CENP-A chromatin of human cells. Genes Cells *11*, 673-684.

Janicki, S.M., Tsukamoto, T., Salghetti, S.E., Tansey, W.P., Sachidanandam, R., Prasanth, K.V., Ried, T., Shav-Tal, Y., Bertrand, E., Singer, R.H., *et al.* (2004). From silencing to gene expression: real-time analysis in single cells. Cell *116*, 683-698.

Jansen, L.E., Black, B.E., Foltz, D.R., and Cleveland, D.W. (2007). Propagation of centromeric chromatin requires exit from mitosis. The Journal of cell biology *176*, 795-805.

Jasencakova, Z., Scharf, A.N., Ask, K., Corpet, A., Imhof, A., Almouzni, G., and Groth, A. (2010). Replication stress interferes with histone recycling and predeposition marking of new histones. Molecular cell *37*, 736-743.

Jiang, W., Lechner, J., and Carbon, J. (1993). Isolation and characterization of a gene (CBF2) specifying a protein component of the budding yeast kinetochore. The Journal of cell biology *121*, 513-519.

Johansson, E., and Dixon, N. (2013). Replicative DNA polymerases. Cold Spring Harbor perspectives in biology *5*.

Johnson, E.S. (2004). Protein modification by SUMO. Annual review of biochemistry *73*, 355-382.

Kanemaki, M., and Labib, K. (2006). Distinct roles for Sld3 and GINS during establishment and progression of eukaryotic DNA replication forks. The EMBO journal *25*, 1753-1763.

Kanemaki, M., Sanchez-Diaz, A., Gambus, A., and Labib, K. (2003). Functional proteomic identification of DNA replication proteins by induced proteolysis in vivo. Nature *423*, 720-724.

Kang, Y.H., Galal, W.C., Farina, A., Tappin, I., and Hurwitz, J. (2012). Properties of the human Cdc45/Mcm2-7/GINS helicase complex and its action with DNA polymerase epsilon in rolling circle DNA synthesis. Proceedings of the National Academy of Sciences of the United States of America *109*, 6042-6047.

Kanke, M., Kodama, Y., Takahashi, T.S., Nakagawa, T., and Masukata, H. (2012). Mcm10 plays an essential role in origin DNA unwinding after loading of the CMG components. The EMBO journal *31*, 2182-2194.

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

Kato, T., Sato, N., Hayama, S., Yamabuki, T., Ito, T., Miyamoto, M., Kondo, S., Nakamura, Y., and Daigo, Y. (2007). Activation of Holliday junction recognizing protein involved in the chromosomal stability and immortality of cancer cells. Cancer research *67*, 8544-8553.

Kaufman, P.D., Cohen, J.L., and Osley, M.A. (1998). Hir proteins are required for position-dependent gene silencing in Saccharomyces cerevisiae in the absence of chromatin assembly factor I. Molecular and cellular biology *18*, 4793-4806.

Kaufman, P.D., Kobayashi, R., and Stillman, B. (1997). Ultraviolet radiation sensitivity and reduction of telomeric silencing in Saccharomyces cerevisiae cells lacking chromatin assembly factor-I. Genes & development *11*, 345-357.

Kerscher, O., Felberbaum, R., and Hochstrasser, M. (2006). Modification of proteins by ubiquitin and ubiquitin-like proteins. Annual review of cell and developmental biology *22*, 159-180.

Kim, I.S., Lee, M., Park, J.H., Jeon, R., Baek, S.H., and Kim, K.I. (2014). betaTrCP-mediated ubiquitylation regulates protein stability of Mis18beta in a cell cycle-dependent manner. Biochemical and biophysical research communications *443*, 62-67.

Kim, I.S., Lee, M., Park, K.C., Jeon, Y., Park, J.H., Hwang, E.J., Jeon, T.I., Ko, S., Lee, H., Baek, S.H.*, et al.* (2012). Roles of Mis18alpha in Epigenetic Regulation of Centromeric Chromatin and CENP-A Loading. Mol Cell.

Kingston, I.J., Yung, J.S., and Singleton, M.R. (2011). Biophysical characterization of the centromere-specific nucleosome from budding yeast. J Biol Chem *286*, 4021-4026.

Kurat, C.F., Yeeles, J.T.P., Patel, H., Early, A., and Diffley, J.F.X. (2017). Chromatin Controls DNA Replication Origin Selection, Lagging-Strand Synthesis, and Replication Fork Rates. Molecular cell *65*, 117-130.

Labib, K. (2010). How do Cdc7 and cyclin-dependent kinases trigger the initiation of chromosome replication in eukaryotic cells? Genes & development *24*, 1208-1219.

Lagana, A., Dorn, J.F., De Rop, V., Ladouceur, A.M., Maddox, A.S., and Maddox, P.S. (2010). A small GTPase molecular switch regulates epigenetic centromere maintenance by stabilizing newly incorporated CENP-A. Nat Cell Biol *12*, 1186-1193.

Lam, A.L., Boivin, C.D., Bonney, C.F., Rudd, M.K., and Sullivan, B.A. (2006). Human centromeric chromatin is a dynamic chromosomal domain that can spread over noncentromeric DNA. Proc Natl Acad Sci U S A *103*, 4186-4191.

Lammers, M., Meyer, S., Kuhlmann, D., and Wittinghofer, A. (2008). Specificity of interactions between mDia isoforms and Rho proteins. The Journal of biological chemistry *283*, 35236-35246.

Lechner, J., and Carbon, J. (1991). A 240 kd multisubunit protein complex, CBF3, is a major component of the budding yeast centromere. Cell *64*, 717-725.

Lee, J., and Zhou, P. (2007). DCAFs, the missing link of the CUL4-DDB1 ubiquitin ligase. Mol Cell *26*, 775-780.

Leffak, I.M. (1984). Conservative segregation of nucleosome core histones. Nature *307*, 82-85.

Leffak, I.M., Grainger, R., and Weintraub, H. (1977). Conservative assembly and segregation of nucleosomal histones. Cell *12*, 837-845.

LeRoy, G., Orphanides, G., Lane, W.S., and Reinberg, D. (1998). Requirement of RSF and FACT for transcription of chromatin templates in vitro. Science (New York, NY) *282*, 1900-1904.

Lewis, P.W., Elsaesser, S.J., Noh, K.M., Stadler, S.C., and Allis, C.D. (2010). Daxx is an H3.3-specific histone chaperone and cooperates with ATRX in replication-independent chromatin assembly at telomeres. Proc Natl Acad Sci U S A *107*, 14075-14080.

Li, Z., Liu, B., Jin, W., Wu, X., Zhou, M., Liu, V.Z., Goel, A., Shen, Z., Zheng, L., and Shen, B. (2018). hDNA2 nuclease/helicase promotes centromeric DNA replication and genome stability. The EMBO journal.

Lin, S.H., Wang, X., Zhang, S., Zhang, Z., Lee, E.Y., and Lee, M.Y. (2013). Dynamics of enzymatic interactions during short flap human Okazaki fragment processing by two forms of human DNA polymerase delta. DNA repair *12*, 922-935.

Liu, C., and Mao, Y. (2016). Formin-mediated epigenetic maintenance of centromere identity. Small GTPases, 1-6.

Loyola, A., and Almouzni, G. (2004). Histone chaperones, a supporting role in the limelight. Biochim Biophys Acta *1677*, 3-11.

Lujan, S.A., Williams, J.S., Clausen, A.R., Clark, A.B., and Kunkel, T.A. (2013). Ribonucleotides are signals for mismatch repair of leading-strand replication errors. Molecular cell *50*, 437-443.

Maddox, P.S., Hyndman, F., Monen, J., Oegema, K., and Desai, A. (2007). Functional genomics identifies a Myb domain-containing protein family required for assembly of CENP-A chromatin. The Journal of cell biology *176*, 757-763.

Madine, M.A., Khoo, C.Y., Mills, A.D., and Laskey, R.A. (1995). MCM3 complex required for cell cycle regulation of DNA replication in vertebrate cells. Nature *375*, 421-424.

Maiorano, D., Lemaitre, J.M., and Mechali, M. (2000). Stepwise regulated chromatin assembly of MCM2-7 proteins. The Journal of biological chemistry *275*, 8426-8431.

Marks, A.B., Fu, H., and Aladjem, M.I. (2017). Regulation of Replication Origins. Advances in experimental medicine and biology *1042*, 43-59.

Martin, M.M., Ryan, M., Kim, R., Zakas, A.L., Fu, H., Lin, C.M., Reinhold, W.C., Davis, S.R., Bilke, S., Liu, H.*, et al.* (2011). Genome-wide depletion of replication initiation events in highly transcribed regions. Genome research *21*, 1822-1832.

Masai, H., Matsumoto, S., You, Z., Yoshizawa-Sugata, N., and Oda, M. (2010). Eukaryotic chromosome DNA replication: where, when, and how? Annual review of biochemistry *79*, 89-130.

May, B.P., Lippman, Z.B., Fang, Y., Spector, D.L., and Martienssen, R.A. (2005). Differential regulation of strand-specific transcripts from Arabidopsis centromeric satellite repeats. PLoS Genet *1*, e79.

Mayo, M.W., Denlinger, C.E., Broad, R.M., Yeung, F., Reilly, E.T., Shi, Y., and Jones, D.R. (2003). Ineffectiveness of histone deacetylase inhibitors to induce apoptosis involves the transcriptional activation of NF-kappa B through the Akt pathway. The Journal of biological chemistry *278*, 18980-18989.

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

McKinley, K.L., and Cheeseman, I.M. (2016). The molecular basis for centromere identity and function. Nature reviews Molecular cell biology *17*, 16-29.

McKnight, S.L., and Miller, O.L., Jr. (1977). Electron microscopic analysis of chromatin replication in the cellular blastoderm Drosophila melanogaster embryo. Cell *12*, 795-804.

Mellone, B.G., Grive, K.J., Shteyn, V., Bowers, S.R., Oderberg, I., and Karpen, G.H. (2011). Assembly of Drosophila centromeric chromatin proteins during mitosis. PLoS Genet *7*, e1002068.

Meluh, P.B., and 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*, 793-807.

Meluh, P.B., and Koshland, D. (1997). Budding yeast centromere composition and assembly as revealed by in vivo cross-linking. Genes Dev *11*, 3401-3412.

Mito, Y., Henikoff, J.G., and Henikoff, S. (2005). Genome-scale profiling of histone H3.3 replacement patterns. Nat Genet *37*, 1090-1097.

Miyabe, I., Kunkel, T.A., and Carr, A.M. (2011). The major roles of DNA polymerases epsilon and delta at the eukaryotic replication fork are evolutionarily conserved. PLoS genetics *7*, e1002407.

Mizuguchi, G., Shen, X., Landry, J., Wu, W.H., Sen, S., and Wu, C. (2004). ATP-driven exchange of histone H2AZ variant catalyzed by SWR1 chromatin remodeling complex. Science (New York, NY) *303*, 343-348.

Mizuguchi, G., Xiao, H., Wisniewski, J., Smith, M.M., and Wu, C. (2007). Nonhistone Scm3 and histones CenH3-H4 assemble the core of centromere-specific nucleosomes. Cell *129*, 1153-1164.

Moggs, J.G., Grandi, P., Quivy, J.P., Jonsson, Z.O., Hubscher, U., Becker, P.B., and Almouzni, G. (2000). A CAF-1-PCNA-mediated chromatin assembly pathway triggered by sensing DNA damage. Molecular and cellular biology *20*, 1206-1218.

Monson, E.K., de Bruin, D., and Zakian, V.A. (1997). The yeast Cac1 protein is required for the stable inheritance of transcriptionally repressed chromatin at telomeres. Proceedings of the National Academy of Sciences of the United States of America *94*, 13081-13086.

Moree, B., Meyer, C.B., Fuller, C.J., and Straight, A.F. (2011). CENP-C recruits M18BP1 to centromeres to promote CENP-A chromatin assembly. The Journal of cell biology *194*, 855-871.

Moreno-Moreno, O., Medina-Giro, S., Torras-Llort, M., and Azorin, F. (2011). The F Box Protein Partner of Paired Regulates Stability of Drosophila Centromeric Histone H3, CenH3(CID). Curr Biol *21*, 1488-1493.

Mouysset, J., Gilberto, S., Meier, M.G., Lampert, F., Belwal, M., Meraldi, P., and Peter, M. (2015). CRL4(RBBP7) is required for efficient CENP-A deposition at centromeres. J Cell Sci *128*, 1732-1745.

Moyer, S.E., Lewis, P.W., and Botchan, M.R. (2006). Isolation of the Cdc45/Mcm2-7/GINS (CMG) complex, a candidate for the eukaryotic DNA replication fork helicase. Proceedings of the National Academy of Sciences of the United States of America *103*, 10236-10241.

Muller, S., Montes de Oca, R., Lacoste, N., Dingli, F., Loew, D., and Almouzni, G. (2014). Phosphorylation and DNA binding of HJURP determine its centromeric recruitment and function in CenH3(CENP-A) loading. Cell Rep *8*, 190-203.

Muzi-Falconi, M., Giannattasio, M., Foiani, M., and Plevani, P. (2003). The DNA polymerase alpha-primase complex: multiple functions and interactions. TheScientificWorldJournal *3*, 21-33.

Nakayama, K.I., and Nakayama, K. (2006). Ubiquitin ligases: cell-cycle control and cancer. Nat Rev Cancer *6*, 369-381.

Nardi, I.K., Zasadzinska, E., Stellfox, M.E., Knippler, C.M., and Foltz, D.R. (2016). Licensing of Centromeric Chromatin Assembly through the Mis18alpha-Mis18beta Heterotetramer. Mol Cell *61*, 774-787.

Natsume, R., Eitoku, M., Akai, Y., Sano, N., Horikoshi, M., and Senda, T. (2007). Structure and function of the histone chaperone CIA/ASF1 complexed with histones H3 and H4. Nature *446*, 338-341.

Nick McElhinny, S.A., Kumar, D., Clark, A.B., Watt, D.L., Watts, B.E., Lundstrom, E.B., Johansson, E., Chabes, A., and Kunkel, T.A. (2010). Genome instability due to ribonucleotide incorporation into DNA. Nature chemical biology *6*, 774-781.

Niikura, Y., Kitagawa, R., Ogi, H., Abdulle, R., Pagala, V., and Kitagawa, K. (2015). CENP-A K124 Ubiquitylation Is Required for CENP-A Deposition at the Centromere. Dev Cell *32*, 589-603.

Nijman, S.M., Luna-Vargas, M.P., Velds, A., Brummelkamp, T.R., Dirac, A.M., Sixma, T.K., and Bernards, R. (2005). A genomic and functional inventory of deubiquitinating enzymes. Cell *123*, 773-786.

Nishihashi, A., Haraguchi, T., Hiraoka, Y., Ikemura, T., Regnier, V., Dodson, H., Earnshaw, W.C., and Fukagawa, T. (2002). CENP-I is essential for centromere function in vertebrate cells. Dev Cell *2*, 463-476.

Nishimura, K., Fukagawa, T., Takisawa, H., Kakimoto, T., and Kanemaki, M. (2009). An auxin-based degron system for the rapid depletion of proteins in nonplant cells. Nature methods *6*, 917-922.

Nishitani, H., Lygerou, Z., Nishimoto, T., and Nurse, P. (2000). The Cdt1 protein is required to license DNA for replication in fission yeast. Nature *404*, 625-628.

O'Keefe, R.T., Henderson, S.C., and Spector, D.L. (1992). Dynamic organization of DNA replication in mammalian cell nuclei: spatially and temporally defined replication of chromosome-specific alpha-satellite DNA sequences. J Cell Biol *116*, 1095-1110.

Obuse, C., Yang, H., Nozaki, N., Goto, S., Okazaki, T., and Yoda, K. (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 Cells *9*, 105-120.

Ohashi, E., and Tsurimoto, T. (2017). Functions of Multiple Clamp and Clamp-Loader Complexes in Eukaryotic DNA Replication. Advances in experimental medicine and biology *1042*, 135-162.

Ohkuni, K., and Kitagawa, K. (2011). Endogenous transcription at the centromere facilitates centromere activity in budding yeast. Curr Biol *21*, 1695-1703.

Ohzeki, J., Shono, N., Otake, K., Martins, N.M., Kugou, K., Kimura, H., Nagase, T., Larionov, V., Earnshaw, W.C., and Masumoto, H. (2016). KAT7/HBO1/MYST2 Regulates CENP-A Chromatin Assembly by Antagonizing Suv39h1-Mediated Centromere Inactivation. Dev Cell *37*, 413-427.

Ohzeki, J.I., Bergmann, J.H., Kouprina, N., Noskov, V.N., Nakano, M., Kimura, H., Earnshaw, W.C., Larionov, V., and Masumoto, H. (2012). Breaking the HAC Barrier: Histone H3K9 acetyl/methyl balance regulates CENP-A assembly. The EMBO journal.

Okada, M., Cheeseman, I.M., Hori, T., Okawa, K., McLeod, I.X., Yates, J.R., 3rd, Desai, A., and Fukagawa, T. (2006). The CENP-H-I complex is required for the efficient incorporation of newly synthesized CENP-A into centromeres. Nat Cell Biol *8*, 446-457.

Okazaki, R., Okazaki, T., Sakabe, K., Sugimoto, K., and Sugino, A. (1968). Mechanism of DNA chain growth. I. Possible discontinuity and unusual secondary structure of newly synthesized chains. Proceedings of the National Academy of Sciences of the United States of America *59*, 598-605.

Ono, T., Fang, Y., Spector, D.L., and Hirano, T. (2004). Spatial and temporal regulation of Condensins I and II in mitotic chromosome assembly in human cells. Molecular biology of the cell *15*, 3296-3308.

Orphanides, G., LeRoy, G., Chang, C.H., Luse, D.S., and Reinberg, D. (1998). FACT, a factor that facilitates transcript elongation through nucleosomes. Cell *92*, 105-116.

Orphanides, G., Wu, W.H., Lane, W.S., Hampsey, M., and Reinberg, D. (1999). The chromatin-specific transcription elongation factor FACT comprises human SPT16 and SSRP1 proteins. Nature *400*, 284-288.

Pacek, M., Tutter, A.V., Kubota, Y., Takisawa, H., and Walter, J.C. (2006). Localization of MCM2-7, Cdc45, and GINS to the site of DNA unwinding during eukaryotic DNA replication. Molecular cell *21*, 581-587.

Pan, D., Klare, K., Petrovic, A., Take, A., Walstein, K., Singh, P., Rondelet, A., Bird, A.W., and Musacchio, A. (2017). CDK-regulated dimerization of M18BP1 on a Mis18 hexamer is necessary for CENP-A loading. Elife *6*.

Patel, S.S., and Picha, K.M. (2000). Structure and function of hexameric helicases. Annual review of biochemistry *69*, 651-697.

Pearson, C.G., Yeh, E., Gardner, M., Odde, D., Salmon, E.D., and Bloom, K. (2004). Stable kinetochore-microtubule attachment constrains centromere positioning in metaphase. Curr Biol *14*, 1962-1967.

Perpelescu, M., Nozaki, N., Obuse, C., Yang, H., and Yoda, K. (2009). Active establishment of centromeric CENP-A chromatin by RSF complex. The Journal of cell biology *185*, 397-407.

Peters, A.H., O'Carroll, D., Scherthan, H., Mechtler, K., Sauer, S., Schofer, C., Weipoltshammer, K., Pagani, M., Lachner, M., Kohlmaier, A.*, et al.* (2001). Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability. Cell *107*, 323-337.

Petryk, N., Kahli, M., d'Aubenton-Carafa, Y., Jaszczyszyn, Y., Shen, Y., Silvain, M., Thermes, C., Chen, C.L., and Hyrien, O. (2016). Replication landscape of the human genome. Nature communications *7*, 10208.

Phansalkar, R., Lapierre, P., and Mellone, B.G. (2012). Evolutionary insights into the role of the essential centromere protein CAL1 in Drosophila. Chromosome Res *20*, 493-504.

Pidoux, A.L., Choi, E.S., Abbott, J.K., Liu, X., Kagansky, A., Castillo, A.G., Hamilton, G.L., Richardson, W., Rappsilber, J., He, X.*, et al.* (2009). Fission yeast Scm3: A CENP-A receptor required for integrity of subkinetochore chromatin. Mol Cell *33*, 299-311.

Potthoff, M.J., and Olson, E.N. (2007). MEF2: a central regulator of diverse developmental programs. Development *134*, 4131-4140.

Prelich, G., and Stillman, B. (1988). Coordinated leading and lagging strand synthesis during SV40 DNA replication in vitro requires PCNA. Cell *53*, 117-126.

Quenet, D., and Dalal, Y. (2012). The CENP-A nucleosome: a dynamic structure and role at the centromere. Chromosome Res.

Quenet, D., and Dalal, Y. (2014a). A long non-coding RNA is required for targeting centromeric protein A to the human centromere. *3*, e03254.

Quenet, D., and Dalal, Y. (2014b). A long non-coding RNA is required for targeting centromeric protein A to the human centromere. Elife *3*, e03254.

Ranjitkar, P., Press, M.O., Yi, X., Baker, R., MacCoss, M.J., and Biggins, S. (2010). An E3 ubiquitin ligase prevents ectopic localization of the centromeric histone H3 variant via the centromere targeting domain. Mol Cell *40*, 455-464.

Ransom, M., Dennehey, B.K., and Tyler, J.K. (2010). Chaperoning histones during DNA replication and repair. Cell *140*, 183-195.

Rao, H., Marahrens, Y., and Stillman, B. (1994). Functional conservation of multiple elements in yeast chromosomal replicators. Molecular and cellular biology *14*, 7643-7651.

Remus, D., Beuron, F., Tolun, G., Griffith, J.D., Morris, E.P., and Diffley, J.F. (2009). Concerted loading of Mcm2-7 double hexamers around DNA during DNA replication origin licensing. Cell *139*, 719-730.

Reyes-Turcu, F.E., Ventii, K.H., and Wilkinson, K.D. (2009). Regulation and cellular roles of ubiquitin-specific deubiquitinating enzymes. Annual review of biochemistry *78*, 363-397.

Rhee, H.W., Zou, P., Udeshi, N.D., Martell, J.D., Mootha, V.K., Carr, S.A., and Ting, A.Y. (2013). Proteomic mapping of mitochondria in living cells via spatially restricted enzymatic tagging. Science (New York, NY) *339*, 1328-1331.

Ribeiro, S.A., Vagnarelli, P., Dong, Y., Hori, T., McEwen, B.F., Fukagawa, T., Flors, C., and Earnshaw, W.C. (2010). A super-resolution map of the vertebrate kinetochore. Proc Natl Acad Sci U S A *107*, 10484-10489.

Rice, J.C., Briggs, S.D., Ueberheide, B., Barber, C.M., Shabanowitz, J., Hunt, D.F., Shinkai, Y., and Allis, C.D. (2003). Histone methyltransferases direct different degrees of methylation to define distinct chromatin domains. Mol Cell *12*, 1591-1598.

Richet, N., Liu, D., Legrand, P., Velours, C., Corpet, A., Gaubert, A., Bakail, M., Moal-Raisin, G., Guerois, R., Compper, C*., et al.* (2015). Structural insight into how the human helicase subunit MCM2 may act as a histone chaperone together with ASF1 at the replication fork. Nucleic Acids Res *43*, 1905-1917.

Roseaulin, L.C., Noguchi, C., Martinez, E., Ziegler, M.A., Toda, T., and Noguchi, E. (2013a). Coordinated degradation of replisome components ensures genome stability upon replication stress in the absence of the replication fork protection complex. PLoS genetics *9*, e1003213.

Roseaulin, L.C., Noguchi, C., and Noguchi, E. (2013b). Proteasome-dependent degradation of replisome components regulates faithful DNA replication. Cell cycle (Georgetown, Tex) *12*, 2564-2569.

Rosic, S., Kohler, F., and Erhardt, S. (2014). Repetitive centromeric satellite RNA is essential for kinetochore formation and cell division. The Journal of cell biology *207*, 335-349.

Ross, J.E., Woodlief, K.S., and Sullivan, B.A. (2016). Inheritance of the CENP-A chromatin domain is spatially and temporally constrained at human centromeres. Epigenetics Chromatin *9*, 20.

Roux, K.J., Kim, D.I., Raida, M., and Burke, B. (2012). A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. The Journal of cell biology *196*, 801-810.

Russell, I.D., Grancell, A.S., and Sorger, P.K. (1999). The unstable F-box protein p58-Ctf13 forms the structural core of the CBF3 kinetochore complex. The Journal of cell biology *145*, 933-950.

Saitoh, H., Tomkiel, J., Cooke, C.A., Ratrie, H., 3rd, Maurer, M., Rothfield, N.F., and Earnshaw, W.C. (1992). CENP-C, an autoantigen in scleroderma, is a component of the human inner kinetochore plate. Cell *70*, 115-125.

Samoshkin, A., Arnaoutov, A., Jansen, L.E., Ouspenski, I., Dye, L., Karpova, T., McNally, J., Dasso, M., Cleveland, D.W., and Strunnikov, A. (2009). Human condensin function is essential for centromeric chromatin assembly and proper sister kinetochore orientation. PLoS One *4*, e6831.

Sanchez-Pulido, L., Pidoux, A.L., Ponting, C.P., and Allshire, R.C. (2009). Common ancestry of the CENP-A chaperones Scm3 and HJURP. Cell *137*, 1173-1174.

Sarma, K., and Reinberg, D. (2005). Histone variants meet their match. Nature reviews Molecular cell biology *6*, 139-149.

Schaarschmidt, D., Baltin, J., Stehle, I.M., Lipps, H.J., and Knippers, R. (2004). An episomal mammalian replicon: sequence-independent binding of the origin recognition complex. The EMBO journal *23*, 191-201.

Schittenhelm, R.B., Althoff, F., Heidmann, S., and Lehner, C.F. (2010). Detrimental incorporation of excess Cenp-A/Cid and Cenp-C into Drosophila centromeres is prevented by limiting amounts of the bridging factor Cal1. J Cell Sci *123*, 3768-3779.

Schneider, J., Bajwa, P., Johnson, F.C., Bhaumik, S.R., and Shilatifard, A. (2006). Rtt109 is required for proper H3K56 acetylation: a chromatin mark associated with the elongating RNA polymerase II. J Biol Chem *281*, 37270-37274.

Schuh, M., Lehner, C.F., and Heidmann, S. (2007). Incorporation of Drosophila CID/CENP-A and CENP-C into centromeres during early embryonic anaphase. Curr Biol *17*, 237-243.

Scott, K.C., and Sullivan, B.A. (2013). Neocentromeres: a place for everything and everything in its place. Trends Genet.

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

Shang, W.H., Hori, T., Westhorpe, F.G., Godek, K.M., Toyoda, A., Misu, S., Monma, N., Ikeo, K., Carroll, C.W., Takami, Y.*, et al.* (2016). Acetylation of histone H4 lysine 5 and 12 is required for CENP-A deposition into centromeres. Nat Commun *7*, 13465.

Shelby, R.D., Vafa, O., and Sullivan, K.F. (1997). Assembly of CENP-A into centromeric chromatin requires a cooperative array of nucleosomal DNA contact sites. J Cell Biol *136*, 501-513.

Shibahara, K., and Stillman, B. (1999). Replication-dependent marking of DNA by PCNA facilitates CAF-1-coupled inheritance of chromatin. Cell *96*, 575-585.

Shivaraju, M., Camahort, R., Mattingly, M., and Gerton, J.L. (2011). Scm3 is a centromeric nucleosome assembly factor. The Journal of biological chemistry *286*, 12016-12023.

Shivaraju, M., Unruh, J.R., Slaughter, B.D., Mattingly, M., Berman, J., and Gerton, J.L. (2012). Cell-cycle-coupled structural oscillation of centromeric nucleosomes in yeast. Cell *150*, 304-316.

Shuaib, M., Ouararhni, K., Dimitrov, S., and Hamiche, A. (2010). HJURP binds CENP-A via a highly conserved N-terminal domain and mediates its deposition at centromeres. Proc Natl Acad Sci U S A *107*, 1349-1354.

Siegel, L.M., and Monty, K.J. (1966). Determination of molecular weights and frictional ratios of proteins in impure systems by use of gel filtration and density gradient centrifugation. Application to crude preparations of sulfite and hydroxylamine reductases. Biochim Biophys Acta *112*, 346-362.

Silva, M.C., Bodor, D.L., Stellfox, M.E., Martins, N.M., Hochegger, H., Foltz, D.R., and Jansen, L.E. (2012). Cdk activity couples epigenetic centromere inheritance to cell cycle progression. Dev Cell *22*, 52-63.

Smith, T.F. (2008). Diversity of WD-repeat proteins. Sub-cellular biochemistry *48*, 20-30.

Sogo, J.M., Stahl, H., Koller, T., and Knippers, R. (1986). Structure of replicating simian virus 40 minichromosomes. The replication fork, core histone segregation and terminal structures. Journal of molecular biology *189*, 189-204.

Sorger, P.K., Doheny, K.F., Hieter, P., Kopski, K.M., Huffaker, T.C., and Hyman, A.A. (1995). Two genes required for the binding of an essential Saccharomyces cerevisiae kinetochore complex to DNA. Proc Natl Acad Sci U S A *92*, 12026-12030.

Spiller, F., Medina-Pritchard, B., Abad, M.A., Wear, M.A., Molina, O., Earnshaw, W.C., and Jeyaprakash, A.A. (2017). Molecular basis for Cdk1-regulated timing of Mis18 complex assembly and CENP-A deposition. EMBO reports *18*, 894-905.

Stankovic, A., Guo, L.Y., Mata, J.F., Bodor, D.L., Cao, X.J., Bailey, A.O., Shabanowitz, J., Hunt, D.F., Garcia, B.A., Black, B.E.*, et al.* (2017). A Dual Inhibitory Mechanism Sufficient to Maintain Cell-Cycle-Restricted CENP-A Assembly. Mol Cell *65*, 231-246.

Stellfox, M.E., Bailey, A.O., and Foltz, D.R. (2012). Putting CENP-A in its place. Cell Mol Life Sci.

Stellfox, M.E., Nardi, I.K., Knippler, C.M., and Foltz, D.R. (2016). Differential Binding Partners of the Mis18alpha/beta YIPPEE Domains Regulate Mis18 Complex Recruitment to Centromeres. Cell Rep *15*, 2127-2135.

Stimpson, K.M., and Sullivan, B.A. (2010). Epigenomics of centromere assembly and function. Current opinion in cell biology *22*, 772-780.

Stinchcomb, D.T., Struhl, K., and Davis, R.W. (1979). Isolation and characterisation of a yeast chromosomal replicator. Nature *282*, 39-43.

Stith, C.M., Sterling, J., Resnick, M.A., Gordenin, D.A., and Burgers, P.M. (2008). Flexibility of eukaryotic Okazaki fragment maturation through regulated strand displacement synthesis. The Journal of biological chemistry *283*, 34129-34140.

Stodola, J.L., and Burgers, P.M. (2016). Resolving individual steps of Okazaki-fragment maturation at a millisecond timescale. Nature structural & molecular biology *23*, 402-408.

Stoler, S., Rogers, K., Weitze, S., Morey, L., Fitzgerald-Hayes, M., and Baker, R.E. (2007). Scm3, an essential Saccharomyces cerevisiae centromere protein required for G2/M progression and Cse4 localization. Proc Natl Acad Sci U S A *104*, 10571-10576.

Strunnikov, A.V., Kingsbury, J., and Koshland, D. (1995). CEP3 encodes a centromere protein of Saccharomyces cerevisiae. The Journal of cell biology *128*, 749-760.

Su, D., Hu, Q., Zhou, H., Thompson, J.R., Xu, R.M., Zhang, Z., and Mer, G. (2011). Structure and histone binding properties of the Vps75-Rtt109 chaperone-lysine acetyltransferase complex. J Biol Chem *286*, 15625-15629.

Subramanian, L., Medina-Pritchard, B., Barton, R., Spiller, F., Kulasegaran-Shylini, R., Radaviciute, G., Allshire, R.C., and Arockia Jeyaprakash, A. (2016). Centromere localization and function of Mis18 requires Yippee-like domain-mediated oligomerization. EMBO Rep.

Subramanian, L., Toda, N.R., Rappsilber, J., and Allshire, R.C. (2014). Eic1 links Mis18 with the CCAN/Mis6/Ctf19 complex to promote CENP-A assembly. Open Biol *4*, 140043.

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

Sugimoto, N., and Fujita, M. (2017). Molecular Mechanism for Chromatin Regulation During MCM Loading in Mammalian Cells. Advances in experimental medicine and biology *1042*, 61-78.

Sullivan, B.A., and Karpen, G.H. (2004). Centromeric chromatin exhibits a histone modification pattern that is distinct from both euchromatin and heterochromatin. Nat Struct Mol Biol *11*, 1076-1083.

Sun, J., Fernandez-Cid, A., Riera, A., Tognetti, S., Yuan, Z., Stillman, B., Speck, C., and Li, H. (2014). Structural and mechanistic insights into Mcm2-7 double-hexamer assembly and function. Genes & development *28*, 2291-2303.

Szenker, E., Ray-Gallet, D., and Almouzni, G. (2011). The double face of the histone variant H3.3. Cell Res *21*, 421-434.

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

Tagami, H., Ray-Gallet, D., Almouzni, G., and Nakatani, Y. (2004). Histone H3.1 and H3.3 complexes mediate nucleosome assembly pathways dependent or independent of DNA synthesis. Cell *116*, 51-61.

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

Takara, T.J., and Bell, S.P. (2011). Multiple Cdt1 molecules act at each origin to load replication-competent Mcm2-7 helicases. The EMBO journal *30*, 4885-4896.

Takayama, Y., Kamimura, Y., Okawa, M., Muramatsu, S., Sugino, A., and Araki, H. (2003). GINS, a novel multiprotein complex required for chromosomal DNA replication in budding yeast. Genes & development *17*, 1153-1165.

Takayama, Y., Sato, H., Saitoh, S., Ogiyama, Y., Masuda, F., and Takahashi, K. (2008). Biphasic Incorporation of Centromeric Histone CENP-A in Fission Yeast. Molecular biology of the cell *19*, 682-690.

Tamura, T., Smith, M., Kanno, T., Dasenbrock, H., Nishiyama, A., and Ozato, K. (2009). Inducible deposition of the histone variant H3.3 in interferon-stimulated genes. The Journal of biological chemistry *284*, 12217-12225.

Tan, B.C., Chien, C.T., Hirose, S., and Lee, S.C. (2006). Functional cooperation between FACT and MCM helicase facilitates initiation of chromatin DNA replication. The EMBO journal *25*, 3975-3985.

Tan, S., Kern, R.C., and Selleck, W. (2005). The pST44 polycistronic expression system for producing protein complexes in Escherichia coli. Protein expression and purification *40*, 385-395.

Tanaka, S., Umemori, T., Hirai, K., Muramatsu, S., Kamimura, Y., and Araki, H. (2007). CDK-dependent phosphorylation of Sld2 and Sld3 initiates DNA replication in budding yeast. Nature *445*, 328-332.

Tang, Y., Meeth, K., Jiang, E., Luo, C., and Marmorstein, R. (2008). Structure of Vps75 and implications for histone chaperone function. Proc Natl Acad Sci U S A *105*, 12206-12211.

Ten Hagen, K.G., Gilbert, D.M., Willard, H.F., and Cohen, S.N. (1990). Replication timing of DNA sequences associated with human centromeres and telomeres. Molecular and cellular biology *10*, 6348-6355.

Tercero, J.A., Labib, K., and Diffley, J.F. (2000). DNA synthesis at individual replication forks requires the essential initiation factor Cdc45p. The EMBO journal *19*, 2082-2093.

Ticau, S., Friedman, L.J., Ivica, N.A., Gelles, J., and Bell, S.P. (2015). Single-molecule studies of origin licensing reveal mechanisms ensuring bidirectional helicase loading. Cell *161*, 513-525.

Topp, C.N., Zhong, C.X., and Dawe, R.K. (2004). Centromere-encoded RNAs are integral components of the maize kinetochore. Proc Natl Acad Sci U S A *101*, 15986-15991.

Tsuda, H., Takarabe, T., Inazawa, J., and Hirohashi, S. (1997). Detection of Numerical Alterations of Chromosomes 3, 7, 17 and X in Low-grade Intracystic Papillary Tumors of the Breast by Multi-color Fluorescence In Situ Hybridization. Breast cancer (Tokyo, Japan) *4*, 247-252.

Tsunaka, Y., Fujiwara, Y., Oyama, T., Hirose, S., and Morikawa, K. (2016). Integrated molecular mechanism directing nucleosome reorganization by human FACT. Genes & development *30*, 673-686.

Tsurimoto, T., and Stillman, B. (1989). Purification of a cellular replication factor, RF-C, that is required for coordinated synthesis of leading and lagging strands during simian virus 40 DNA replication in vitro. Molecular and cellular biology *9*, 609-619.

Tye, B.K., Nyman, P.O., Lehman, I.R., Hochhauser, S., and Weiss, B. (1977). Transient accumulation of Okazaki fragments as a result of uracil incorporation into nascent DNA. Proceedings of the National Academy of Sciences of the United States of America *74*, 154-157.

Tyler, J.K., Adams, C.R., Chen, S.R., Kobayashi, R., Kamakaka, R.T., and Kadonaga, J.T. (1999). The RCAF complex mediates chromatin assembly during DNA replication and repair. Nature *402*, 555-560.

Tyler, J.K., Collins, K.A., Prasad-Sinha, J., Amiott, E., Bulger, M., Harte, P.J., Kobayashi, R., and Kadonaga, J.T. (2001). Interaction between the Drosophila CAF-1 and ASF1 chromatin assembly factors. Mol Cell Biol *21*, 6574-6584.

Valenzuela, M.S., Chen, Y., Davis, S., Yang, F., Walker, R.L., Bilke, S., Lueders, J., Martin, M.M., Aladjem, M.I., Massion, P.P.*, et al.* (2011). Preferential localization of human origins of DNA replication at the 5'-ends of expressed genes and at evolutionarily conserved DNA sequences. PloS one *6*, e17308.

van Deursen, F., Sengupta, S., De Piccoli, G., Sanchez-Diaz, A., and Labib, K. (2012). Mcm10 associates with the loaded DNA helicase at replication origins and defines a novel step in its activation. The EMBO journal *31*, 2195-2206.

VanDemark, A.P., Blanksma, M., Ferris, E., Heroux, A., Hill, C.P., and Formosa, T. (2006). The structure of the yFACT Pob3-M domain, its interaction with the DNA replication factor RPA, and a potential role in nucleosome deposition. Molecular cell *22*, 363-374.

Vashee, S., Cvetic, C., Lu, W., Simancek, P., Kelly, T.J., and Walter, J.C. (2003). Sequence-independent DNA binding and replication initiation by the human origin recognition complex. Genes & development *17*, 1894-1908.

Waga, S., Bauer, G., and Stillman, B. (1994). Reconstitution of complete SV40 DNA replication with purified replication factors. The Journal of biological chemistry *269*, 10923-10934.

Wang, J., Liu, X., Dou, Z., Chen, L., Jiang, H., Fu, C., Fu, G., Liu, D., Zhang, J., Zhu, T.*, et al.* (2014). Mitotic Regulator Mis18beta Interacts with and Specifies the Centromeric Assembly of Molecular Chaperone Holliday Junction Recognition Protein (HJURP). The Journal of biological chemistry *289*, 8326-8336.

Weber, C.M., and Henikoff, S. (2014). Histone variants: dynamic punctuation in transcription. Genes Dev *28*, 672-682.

Westermann, S., Cheeseman, I.M., Anderson, S., Yates, J.R., 3rd, Drubin, D.G., and Barnes, G. (2003). Architecture of the budding yeast kinetochore reveals a conserved molecular core. The Journal of cell biology *163*, 215-222.

Wignall, S.M., Deehan, R., Maresca, T.J., and Heald, R. (2003). The condensin complex is required for proper spindle assembly and chromosome segregation in Xenopus egg extracts. The Journal of cell biology *161*, 1041-1051.

Williams, J.S., Hayashi, T., Yanagida, M., and Russell, P. (2009). Fission yeast Scm3 mediates stable assembly of Cnp1/CENP-A into centromeric chromatin. Mol Cell *33*, 287-298.

Winkler, D.D., Zhou, H., Dar, M.A., Zhang, Z., and Luger, K. (2012). Yeast CAF-1 assembles histone (H3-H4)2 tetramers prior to DNA deposition. Nucleic acids research *40*, 10139-10149.

Wisniewski, J., Hajj, B., Chen, J., Mizuguchi, G., Xiao, H., Wei, D., Dahan, M., and Wu, C. (2014). Imaging the fate of histone Cse4 reveals de novo replacement in S phase and subsequent stable residence at centromeres. Elife *3*, e02203.

Wittmeyer, J., Joss, L., and Formosa, T. (1999). Spt16 and Pob3 of Saccharomyces cerevisiae form an essential, abundant heterodimer that is nuclear, chromatin-associated, and copurifies with DNA polymerase alpha. Biochemistry *38*, 8961-8971.

Wong, L.H., Brettingham-Moore, K.H., Chan, L., Quach, J.M., Anderson, M.A., Northrop, E.L., Hannan, R., Saffery, R., Shaw, M.L., Williams, E.*, et al.* (2007). Centromere RNA is a key component for the assembly of nucleoproteins at the nucleolus and centromere. Genome Res *17*, 1146-1160.

Wyrick, J.J., Aparicio, J.G., Chen, T., Barnett, J.D., Jennings, E.G., Young, R.A., Bell, S.P., and Aparicio, O.M. (2001). Genome-wide distribution of ORC and MCM proteins in S. cerevisiae: high-resolution mapping of replication origins. Science (New York, NY) *294*, 2357-2360.

Xiao, H., Mizuguchi, G., Wisniewski, J., Huang, Y., Wei, D., and Wu, C. (2011). Nonhistone Scm3 binds to AT-rich DNA to organize atypical centromeric nucleosome of budding yeast. Mol Cell *43*, 369-380.

Yamasu, K., and Senshu, T. (1990). Conservative segregation of tetrameric units of H3 and H4 histones during nucleosome replication. Journal of biochemistry *107*, 15-20.

Yan, S., and Willis, J. (2013). WD40-repeat protein WDR18 collaborates with TopBP1 to facilitate DNA damage checkpoint signaling. Biochemical and biophysical research communications *431*, 466-471.

Yang, L., Wold, M.S., Li, J.J., Kelly, T.J., and Liu, L.F. (1987). Roles of DNA topoisomerases in simian virus 40 DNA replication in vitro. Proceedings of the National Academy of Sciences of the United States of America *84*, 950-954.

Yong-Gonzalez, V., Wang, B.D., Butylin, P., Ouspenski, I., and Strunnikov, A. (2007). Condensin function at centromere chromatin facilitates proper kinetochore tension and ensures correct mitotic segregation of sister chromatids. Genes Cells *12*, 1075-1090.

Yu, Z., Zhou, X., Wang, W., Deng, W., Fang, J., Hu, H., Wang, Z., Li, S., Cui, L., Shen, J.*, et al.* (2015). Dynamic phosphorylation of CENP-A at Ser68 orchestrates its cell-cycle-dependent deposition at centromeres. Dev Cell *32*, 68-81.

Yuzhakov, A., Kelman, Z., Hurwitz, J., and O'Donnell, M. (1999). Multiple competition reactions for RPA order the assembly of the DNA polymerase delta holoenzyme. The EMBO journal *18*, 6189-6199.

Zasadzinska, E., Barnhart-Dailey, M.C., Kuich, P.H., and Foltz, D.R. (2013). Dimerization of the CENP-A assembly factor HJURP is required for centromeric nucleosome deposition. The EMBO journal *32*, 2113-2124.

Zegerman, P., and Diffley, J.F. (2007). Phosphorylation of Sld2 and Sld3 by cyclin-dependent kinases promotes DNA replication in budding yeast. Nature *445*, 281-285.

Zhang, W., Colmenares, S.U., and Karpen, G.H. (2012). Assembly of Drosophila centromeric nucleosomes requires CID dimerization. Mol Cell *45*, 263-269.

Zhou, Y., and Wang, T.S. (2004). A coordinated temporal interplay of nucleosome reorganization factor, sister chromatin cohesion factor, and DNA polymerase alpha facilitates DNA replication. Molecular and cellular biology *24*, 9568-9579.

Zhou, Z., Feng, H., Zhou, B.R., Ghirlando, R., Hu, K., Zwolak, A., Miller Jenkins, L.M., Xiao, H., Tjandra, N., Wu, C*., et al.* (2011). Structural basis for recognition of centromere histone variant CenH3 by the chaperone Scm3. Nature *472*, 234-237.