The role of the Mis18α-β complex and its interactions with HJURP and CENP-A in human centromeric chromatin establishment

Isaac Kaufman Nardi, VA

B.S. Biology, Virginia Tech, 2010 M.S. Biochemistry and Molecular Genetics, University of Virginia, 2012

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 February, 2016

Table of Contents

Abstract

The centromere, in higher eukaryotes, is an epigenetically specified locus that is the site kinetochore formation on each chromosome during mitosis. The epigenetic mark that defines a centromere as of this writing is the histone H3 variant Centromere Protein-A (CENP-A). The establishment and maintenance of CENP-A at centromeres is absolutely critical for proper chromosome segregation and ploidy in cells. Several key factors have been identified as crucial machinery involved in CENP-A deposition. Understanding how a specific subset of these factors maintains CENP-A specifically at centromeric chromatin is the theme of this work. The second chapter of this work demonstrates the CENP-A chaperone HJURP is directly targeted to centromeres by the Mis18 complex. Furthermore, binding of Mis18 by HJURP mislocalizes the complex from centromeres giving cells control over where CENP-A deposition occurs and how much CENP-A gets consigned at centromeres. The third chapter characterizes a direct interaction between Mis18 and CENP-A that has never been characterized before. This work, though in its infancy, alludes to an important role of Mis18 as being an integral part of the CENP-A deposition machinery into centromeric chromatin. The fourth chapter describes the CENP-T/W/S/X complex and its ability to form “nucleosome-like” structures that may be involved in further differentiating the centromeres from outside chromatin. I also go over a potential role of HJURP in the DNA damage response which was touched on several years ago but not fully followed up ever since HJURPs role in CENP-A deposition has been discovered.

CHAPTER I: GENERAL INTRODUCTION

General Perspective and Significance

Accurate chromosome segregation is absolutely critical for proper cell ploidy, viability, and survival. To ensure this process happens properly every cell cycle the centromere, an epigenetically specified locus in higher eukaryotes, is absolutely required to be positioned and properly maintained on every chromosome to allow for the proteinacous kinetochore apparatus to form during mitosis. Chromosome mis-segregation and the resulting aneuploidy is a hallmark of a majority of cancers and is shown to be heavily involved in a variety of birth defects. Therefore, understanding how proper and improper chromosome segregation occur and how the centromere is involved in this process is of the utmost importance in not only understanding the process but also for the identification of novel targets for future chemotherapeutics. The work presented here is significant to the centromere field and cancer research because it furthers our understanding of the essential steps of CENP-A assembly into chromatin through a never before seen, and hugely critical interaction, between HJURP and Mis18.

The Epigenetic Propagation of the Centromere

Centromeric Chromatin

The centromere is a region on every chromosome that is differentiated by a single histone H3 variant, CENP-A. CENP-A is a key mark that differentiates the centromere from neighboring heterochromatin. Like H3, CENP-A is deposited into nucleosomes that wrap and organize the DNA (Cse4 in *S. Cerevisiae*, Cnp1 in *S. Pombe*, CID in *Drosophila*). Propagation of centromeres is absolutely essential for chromosome viability and inheritance. This region is the foundation for kinetochore formation, subsequent spindle microtubule attachments, and chromosome segregation during mitosis (Figure 1-1). Despite the essential role for centromeres in stable genetic inheritance the mechanisms for how this chromosomal region is propagated every cell cycle is not fully understood. Furthermore, centromere specification varies widely between some organisms adding even more depth to an already complicated field. We will discuss the different centromere environments that are present across species with an emphasis on hallmarks of epigenetically specified centromeres found in higher eukaryotes.

The simplest centromeres currently known are from budding yeast *Saccharomyces cerevisiae.* Centromere position in these organisms is not determined epigenetically but rather genetically by a 125 base pair DNA sequence that harbors a single CENP-A/Cse4 nucleosome. The underlying DNA sequences contains centromere DNA elements (CDEs) that recruit a variety of sequence specific DNA binding complexes that harbor the CENP-A/Cse4 deposition machinery and are necessary to recruit the proper kinetochore proteins.

Centromeres in *Schizosaccharomyces pombe* and higher eukaryotes such as mice and humans are determined epigenetically, as in the underlying DNA sequence, unlike *S. cerevisiae,* is neither necessary nor sufficient to form these regional centromeres that are found only on a single part of the chromosome. This is in contrast to such organisms as the nematode *C. elegans* that harbor holocentric chromosomes that assemble centromeres and kinetochores along the entire length of the chromosome during mitosis. Although the centromere in higher eukaryotes can be termed monocentric in that there is only one of centromeric region per chromosome, the region is quite dispersed. Indeed, centromeres in higher eukaryotes actually contain both H3 and CENP-A nucleosomes and are quite expansive. Despite the centromeres in higher eukaryotes being specified epigenetically they are generally assembled on repetitive DNA sequences, though there are examples in chickens where centromeres are formed on non-repetitive regions [1]. The underlying DNA in mammalian centromeres is composed of approximately 0.5-5 Mb of repetitive, tandemly repeating alpha-satellite DNA[2]. An exception to the large tandem repeats is in *Drosophila*, where the functionally mapped centromere consists of simple [5 bp] satellite sequences interspersed with various transposons[3]. As was mentioned above, CENP-A nucleosomes are not uniformly distributed along centromeres. Stretched chromatin fibers reveal CENP-A interspersed with H3 containing nucleosomes[4]. Some research suggests the interspersing of these different nucleosome flavors at the centromere is important to form a three-dimensional binding surface during mitosis. The model predicts CENP-A nucleosome regions are pushed out and together as a single sheet where they are exposed as a platform to form the mitotic kinetochore, while the H3 nucleosome regions are folded beneath near the inner centromere region[5].

A central question for the three-dimensional centromere organization mentioned above is how this process occurs. One explanation could be the unique epigenetic landscape found at centromeres. Strikingly, centromeres contain a mix of epigenetic marks commonly involved in both active and silenced chromatin. Despite CENP-A and H3 nucleosomes looking quite similar in crystal structures they are divergent in sequence homology especially when you compare the CENP-A N-terminal tail to that of H3. For CENP-A, this tail is modified by Aurora B at Ser-7 (Ser-10 for H3). Further work has revealed that there is a trimethylation event on CENP-A at Gly1 and two phosphorylation events at Ser16 and Ser18[6]. H3 containing nucleosomes at the centromere have an array of post translational modifications such as hypoacetylation and dimethylation at Lys-4, 9, and 36[7,8]. The biological significance of these modifications in reference to centromere propagation is unknown, but many of these post-translational modifications, if prevented, cause lagging chromosomes during anaphase in human cells[6]. The unique epigenetic landscape of the centromere is clearly important for defining where centromeres are established and propagated but the full mechanism is still not known.

Although alpha-satellite DNA is a common hallmark of centromeric chromatin in higher eukaryotes it is neither necessary nor sufficient for centromere formation. Mitotically stable neocentromeres have been observed that form outside of alpha-satellite DNA that contain no canonical centromeric sequences at all[9]. *Drosophila* and fission yeast studies have been performed where the endogenous centromere is removed and the cells are forced to form neocentromeres on these chromosomes in order to survive. The neocentromeres do indeed form but are completely devoid of centromeric DNA sequences[10,11]. Others have also used a LacO/LacI targeting system in humans and in

Drosophila[12–15] and successfully formed competent chromatin that is stable for kinetochore formation at a completely non-centromeric location. These data strongly suggest that the eukaryotic centromere is epigenetically specified and that CENP-A is one of the key epigenetic marks that define human centromeres.

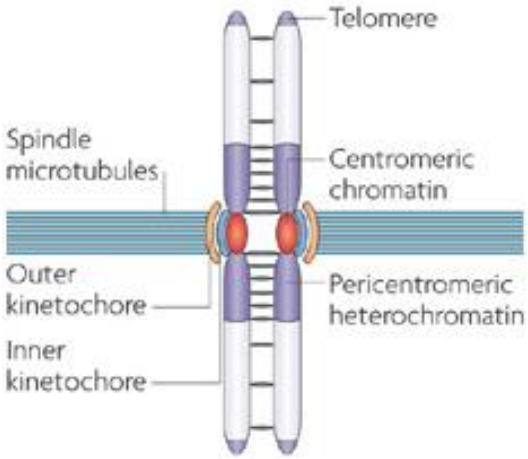


Figure 1-1: Schematic representation of the vertebrate centromere

The centromere is a unique chromosomal locus that is epigenetically defined by the presence of nucleosomes containing the histone H3 variant CENP-A. It is the assembly site for the kinetochore to form atop during mitosis. The kinetochore interacts with microtubules and mediates sister chromatid separation to each daughter cell during mitosis.
Figure included with permission, (Allshire & Karpen et al., 2008).

The CCAN: A Platform for Kinetochore Formation

The Constitutive Centromere Associated Network (CCAN) is a growing set of centromere and associated proteins that specifically localize to the centromere throughout the cell cycle and are required for kinetochore formation and centromere propagation. Knockdown of any one of these proteins causes centromere and kinetochore dysfunction and subsequent chromosome mis-segregation. Many of these CCAN components co-purify with CENP-A and form sub-complexes with eachother. The sub-complexes are recruited in a hierarchical manner and are often dependent on one another for proper localization. The sub groups of CENPs closes to CENP-A are referred to as the Nucleosome Associated Complex (NAC) all the way to the CENP-A Distal (CAD) groups that interact with the outer kinetochore subunits including spindle microtubules. The CCAN components are highly conserved from yeast to humans. Some organisms, including *Drosophila* and *C. Elegans,* completely lack all of the CCAN proteins except for CENP-C[16].

CENP-C is a very important CENP that is highly conserved across species. It along with CENP-N are the only known CCAN proteins that directly bind and interact with CENP-A nucleosomes[17,18]. CENP-C$^{aa422-537}$ was shown to be able to bind to CENP-A nucleosomes *in vitro* and co-purified with CENP-A nucleosomes dipped into *Xenopus Laevis* egg extracts. In particular, it was found CENP-C specifically recognized the last 6 amino acids of CENP-A, as when CENP-A nucleosomes were reconstituted harboring the homologous H3 residues, binding of CENP-C$^{aa422-537}$ was abolished. This result has not been recapitulated in human cells as of yet[17]. CENP-N does not physically interact with CENP-C but directly interacts with the CENP-A Targeting Domain (CATD) of CENP-A

nucleosomes and acts in a dual recognition pathway to specifically identify centromeric chromatin[17,18]. CENP-C, along with CENPs A and B, was previously defined as the prekinetochore complex on which the kinetochore assembles on human cells[19]. CENP-C has been placed as one of the furthest upstream components of CCAN recruitment, knockdown of this protein causes massive chromosome alignment and kinetochore defects[20]. More specifically, CENP-C has been shown to be involved in recruitment of centromere and kinetochore proteins such as CENP-K, CENP-E, Mad2, the Mis12 complex, and KNL1[20–22]. Mis12 is a member of the KMN network, which is involved at the kinetochore in microtubule binding. The N-terminus of CENP-C in humans and flies was shown to be specifically required for the recruitment of the Mis12 complex[14]. CENP-N and CENP-C are clearly key building blocks for CENP-A recruitment and building of the CCAN. The exact mechanisms of how they are carrying out their functions are still not known.

CENP-T was originally identified as a component of the NAC. Upon its knockdown, as is usual with key CCAN components, massive chromosome misalignment and kinetochore defects were observed[23]. CENP-T is a large protein with a histone fold domain and a long N-terminal tail that stretches out to the outer kinetochore. CENP-W is a small protein composed entirely of a histone fold domain and is tightly associated in complex with CENP-T. Unsurprisingly, the histone fold domains of both CENP-T and CENP-W are necessary for DNA binding and centromere recruitment. Knockdown of the CENP-T/W complex in chicken DT40 cells places the complex upstream of a majority of the CCAN components and puts it parallel with CENP-C recruitment[24]. The N-terminal region of CNEP-T has been shown to bind directly to the Ndc80 complex but again the

C-terminal histone fold domain is absolutely required for centromere recruitment. This implies a dual role for CENP-T in centromeric recruitment and propagation along with a function in kinetochore formation. Ectopic recruitment of CENP-T lacking its C-terminal histone fold domain induces a kinetochore structure that can binds microtubules at this ectopic region[14]. This research indicates that CENP-T has a motile structure involved in both inner and outer kinetochore activity.

CENP-S was identified with CENP-M and CENP-U/50 from affinity purifications[23,25] and found to be a binding partner of Stra13, otherwise known as CENP-X. Both of these proteins are almost entirely composed of histone fold domains and their crystal structures look strikingly like that of canonical histones[26]. Depletion of either CENP-S or CENP-X abolishes localization of its partner but not that of CENP-T or CENP-W, placing them downstream of these T/W complex. Furthermore, mitotic abnormalities typically follow when these components are knocked down broadly suggesting they are somehow necessary for proper CCCAN formation. Using high sensitivity mass spectrometry, CENP-S was found to be associated with CENP-T in immunoprecipitates[25]. CENP-S and CENP-X are also known as MHF1/MHF2 respectively and are associated with FANCM. FANCM is a member of the Fanconi Anemia (FA) complex[27]. Although CENP-S and CENP-X are identical to MHF1 and MHF2 no centomeric localization of FANCM or any of the FA complex members has been observed. As of now, the relationship between FANCM and CENP-S/X in relation to centromere function is currently not known.

Recent research has crystallized the structures of both the CENP-T histone fold domain in complex with CENP-W along with the CENP-S/X complex. It was found CENP-T/W forms a dimer (similar to H2A/H2B) while the CENP-S/X complex form a

tetramer. When these complexes are mixed together, a CENP-T/W/S/X heterotetramer forms consisting of one copy of each CENP. Strikingly, the individual complexes along with the T/W/S/X complex protect roughly 100bp of DNA *in vitro.* Mutants designed to compromise heterotetramerization or the DNA-protein contacts around the heterotetramer strongly reduce the DNA binding and supercoiling activities *in vitro* and compromise kinetochore assembly *in vivo*. These data suggest that the CENP-T-W-S-X complex forms a unique nucleosome-like structure to generate contacts with DNA. This unique complex could be yet another epigenetic mark that helps differentiate centromeres from the surrounding chromatin[26].

The CENP-H/I/K complex is located more distal than the above complexes and CENPs mentioned. The H/I/K complex is located in the CAD and is downstream of CENP-T recruitment[17,18,23]. Despite being upstream of the "nodal" CENPs involved in CCAN formation, the H/I/K complex was shown to be necessary for proper CENP-A deposition at centromeres which was shown to occur through CHD1, a member of the FACT complex[28]. The H/I/K complex has also been shown to be involved in microtubule attachment and Mad2 localization to unattached kinetochores in the absence of Aurora B[29,30].

The CCAN is composed of a vast array of inter-dependent CENPs that have an even wider breadth of functions that are still not fully understood in regards to centromere formation and kinetochore function during mitosis. Although knockdown of any one of these CCAN components result in centromere/kinetochore dysfunction and chromosomal abnormalities, the exact role of these individual CCAN components is still not fully understood. The known CCAN subgroups and their relative location between

the centromere and kinetochore are depicted in Figure 1-2. Overall, the CCAN functions along with CENP-A nucleosomes to propagate and define centromeric chromatin every cell cycle.

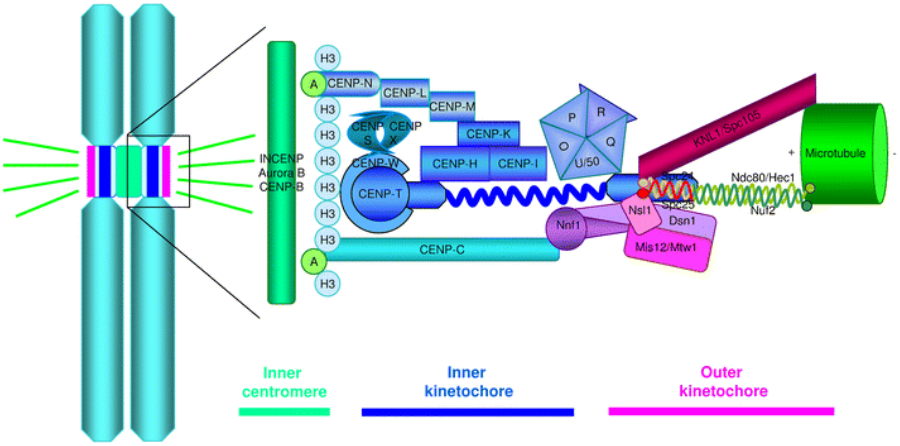


Figure 1-2. Molecular organization of the vertebrate constitutive centromere network and the kinetochore.

Schematic showing a vertebrate sister chromatid that is further detailed to show components of the outer centromere, inner kinetochore, and outer kinetochore with the different components making up the three regions. CENP-A nucleosomes are absolutely required at the base of the CCAN for direct interaction with both CENP-C and CENP-N while CENP-T "reaches" out and bridges the inner centromere to the outer kinetochore. CENP-S/X are involved with CENP-T/W to form a unique nucleosome-like complex that may aid in giving the centromere unique characteristics through novel DNA wrapping through non-canonical histone like proteins.
Figure included with permission, (Perpelescu and Fukagawa, 2011).

CENP-A Structural Characteristics

CENP-A is a Histone H3 variant that is specifically found only at centromeres and is thought to act as *the* epigenetic mark specifying centromeres. Like H3, it forms a nucleosome at centromeres and it is thought that particular characteristics of CENP-A may be involved in defining the centromere as a unique chromosomal locus for kinetochore formation. Having crystallized the CENP-A nucleosome it is striking how similar it looks compared to H3 nucleosomes[31,32] yet there are key differences in the primary sequence of CENP-A that are important for differentiating it from H3. Some studies have suggested that CENP-A nucleosomes are uniquely compacted when compared to H3.1 nucleosomes[33–36]. The N-terminal tail of CENP-A is also highly divergent from that of H3 and has been shown to contain a unique landscape of post-trsnslational modifications in cells. The unique post-translational modifications made to the CENP-A N-terminal tail as opposed to H3 aid in its ability to recruit the specific CCAN components as well as propagate the centromere and specify its location every cell cycle.

A defining region of CENP-A that differs from its H3.1 counterpart can be found in the CENP-A targeting domain (CATD). This region is located within the loop1 and alpha-2 helix of the histone fold of CENP-A. Several of the residues in this region differ from H3.1 and if these residues are replaced with their H3.1 counterparts, CENP-A completely loses its centromeric localization. Replacing H3.1 with the CATD is sufficient to target it to centromeres and rescue CENP-A depletion in cells[33,37,38]. The histone fold of Cse4 (CENP-A homolog) in *S. cerevisiae* is also sufficient for H3 centromere targeting[39], which shows how important and conserved this domain is in

CENP-A across species. The CATD region of CENP-A is also important for direct recruitment and binding of CENP-N. Furthermore, this region specifically binds to the Scm3 domain of HJURP[17,40]. It is clear from this data the CATD region in human cells is highly important in being recruited to and identifying the site for centromere assembly.

Others have suggested CENP-A differs from H3 in its chromatin bound state in that it forms not a nucleosome but rather a heterotetramer with single copies of all four histones[41,42]. This research suggesting a heterotetramer was proposed to take place in humans and flies. In fission yeast data supports a hexameric CENP-A nucleosome that lacks histones H2A and H2B but includes CENP-A/Cse4 chaperone Scm3[41,42]. It should be noted that the data for flies and humans containing single copies of each histone was generated using atomic force microscopy, which has been shown to detect a reduced height of octameric nucleosomes. This implies that the suggested height difference in the experiments may be in fact due to the experimental method of atomic force microscopy that may not be biologically relevant[36]. Evidence for an octameric CENP-A nucleosome has been demonstrated showing the dimerization interface between CENP-A/Cse4 is required for stable deposition at centromeres in budding yeast, flies, and humans[43–45].

In summation, assembled CENP-A nucleosomes have been agreed by the community at large to exist as octamers. Significant differences exist between CENP-A and H3.1 nucleosome on sequential and structural level that may contribute to the specificity of CENP-A nucleosomes being assembled exclusively at centromeres and their unique ability to recruit centromeric proteins.

The CENP-A Deposition Pathway

CENP-A deposition differs from canonical histone deposition in that it is uncoupled from DNA replication, instead taking place during early G1-phase[38,46]. This means that during DNA replication in S-phase CENP-A is divided and diluted between sister centromeres, which proceed to G2 and the concurrent mitosis with roughly half of their initial complement of CENP-A[33,47]. Although CENP-A gets diluted by 50% during S-phase, whether or not CENP-A is equally partitioned between sister centromeres and if the total amount of CENP-A is stably inherited is still not known.

In *Drosophila* CENP-A deposition slightly differs as CENP-A/CID loading takes place during anaphase or metaphase in S2 cells independently of spindle microtubule binding[48,49]. Budding yeasts allow for CENP-A/Cse4 incorporation in S-phase only[50]. Adding even more depth, fission yeast can deposit CENP-A/Cnp1 in both S and G2-phase[51,52]. CENP-A/CENH3 deposition in plants occurs during late G2-phase as well[53].

The CENP-T/W and S/X complexes turn over every cell cycle and has been shown to be at least indirectly involved in CENP-A deposition[54,55]. The H/I/K complex has also been shown to affect CENP-A deposition. Lastly, both CENP-C and CENP-N are necessary for CENP-A recruitment and directly interact with CENP-A nucleosomes, moreover they are even more stable during S-phase which is when CENP-A is diluted[56,57]. Given that CENP-A is diluted in S-phase, the stabilization of the CCAN components during this time is likely critical for maintaining centromeric chromatin and subsequent CENP-A replenishment during the next G1-phase.

Temporal Regulation of CENP-A Deposition

CENP-A deposition is tightly regulated in humans and there are several mechanisms cells employ to carry out this process in a specific window (G1-phase) during the cell cycle. The CCAN, mentioned above, has several components that are known to localize to the centromere in a cell cycle specific manner and are regulated as such just like CENP-A. A key regulator of CENP-A deposition and CCAN recruitment at specific windows in the cell cycle are cyclin-dependent kinases (Cdk's). These proteins, and the subsequent cyclins they bind, regulate a cells progression through the cell cycle. Through G1 and S-phase, cyclins accumulate resulting in increased Cdk activity. Upon the G2/M transition, Cdk levels are at their highest causing a massive phosphorylation of all of their substrates. Once mitosis occurs, and all sister chromatids are properly oriented on the metaphase plate, the spindle assembly checkpoint is passed and the cyclins are rapidly degraded by the anaphase-promoting complex. After this the cell enters G1 with low Cdk activity[58].

Given that CENP-A loading takes place at the end of anaphase in human cells, one could surmise that Cdk's mat play a role in the CENP-A deposition pathway, specifically the timing of CENP-A deposition. Inhibition of Cdk activity in both human and chicken cells cause premature loading of the CENP-A deposition machinery and CENP-A during G2-phase[59]. Giving even more credibility to this, it was found that HJURP is phosphorylated by Cdk1 which weakens its interaction with Mis18β, and the interaction of HJURP with the Mis18 complex is key for its recruitment to centromeres[60]. Other factors outside of Cdk's have also been shown to be important for CENP-A deposition in G1-phase. Polo-

like kinase-1 (Plk1) acts independently of Cdk activity but has been shown to be absolutely required for Mis18 complex localization and subsequent CENP-A deposition[61]. Taken together, these results suggest cells employ several mechanisms to regulate CENP-A deposition during the cell cycle. This shows that timing is incredibly important for the CENP-A deposition pathway, which is illustrated in Figure 1-3.

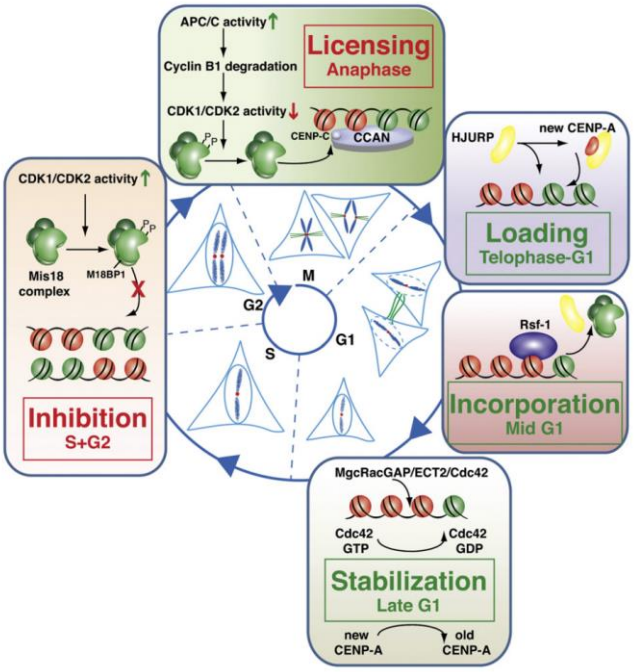


Figure 1-3. Temporal regulation of CENP-A during the cell cycle in humans.

CENP-A deposition is tightly regulated to G1-phase during the cell cycle. This temporal regulation is controlled by the overall levels of Cdk's and their substrates called cyclins. Cdk levels consistently rise all the way to mitosis where they rapidly fall after the spindle assembly checkpoint has been satisfied. This occurs along with a rapid degradation of the cyclin proteins, leading to low Cdk activity going into G1-phase. Deregulation of this process can cause misloading of CENP-A in other parts of the cell cycle. During early G1-phase the Mis18 complex localizes to centromeres and "primes" the chromatin causing HJURP recruitment that is in tow with new CENP-A/H4 to be deposited at the centromeres. In late G1 remodeling factors such as Rsf-1 and CHD1 localize to the centromeres are involved in fully incorporating the new CENP-A nucleosomes into chromatin. MgcRacGAP and Ect2 also localize to centromeres in late G1 and stabilize the new CENP-A nucleosomes by a mechanism involving GTP cycling.
Figure included with permission, (Nechemia-Arbely et al., 2012).

Centromere Priming Components: CENP-C

CENP-C is a constitutive centromere protein that has the ability to bind DNA[62,63]. As mentioned above, CENP-C has been found to be located closest to CENP-A nucleosomes upon K-SHREC analysis[64]. *In vitro* analysis has revealed CENP-C directly binds to CENP-A nucleosomes along with CENP-N, albeit at two different interfaces of CENP-A[17,18]. Further research has shown CENP-C is absolutely required for CENP-A deposition[17,65,66]. CENP-C has also been shown to interact with other centromere licensing proteins such as the Mis18 complex. In particular Mis18BP1 interaction with CENP-C has been shown in humans, mice, and Xenopus[65,67]. Indeed CENP-C is so conserved and important for CENP-A deposition it is even still present and required for CENP-A/CID loading in *Drosophila,* an organism that lacks the Mis18 complex altogether. In flies, the HJURP homologue CAL1 has been shown to directly interact with CENP-C at centromeres and this physical interaction is absolutely necessary for CENP-A/CID deposition[13,68].

Although CENP-C has been shown to interact with Mis18BP1 and is supposedly the most upstream component in CCAN recruitment lying closest to CENP-A nucleosomes, its depletion does not fully abolish Mis18BP1 localization to centromeres in mice. In *Xenopus* there are two isoforms of Mis18BP1 that localize to the centromeres, one of these isoforms is not affected by CENP-C knockdown but is not conserved across species[65,67]. While these researchers put forth a model that CENP-C recruits Mis18BP1, which then recruits both Mis18α and Mis18β, work from our lab shows this is not the whole story. Our lab has shown Mis18BP1 does indeed bind CENP-C but this binding also requires CENP-C directly interacting with and binding to Mis18β. Furthermore, we

have shown Mis18α binds directly and specifically to Mis18BP1 and this interaction is indispensable for Mis18BP1 localization to centromeres. Additionally, CENP-C targeting to a non-centromeric locus is not sufficient to initiate CENP-A chromatin formation at that site[14]. These data strongly suggest an interplay between CENP-C and the Mis18 complex for proper recruitment of one another to centromeres. Moreover, although CENP-C does play a central role in recruiting the CENP-A deposition machinery, it is neither solely necessary nor sufficient for recruitment of the Mis18 complex and CENP-A deposition at centromeres.

Centromere Priming Components: HJURP

HJURP (Holliday junction-recognizing protein) was first recognized as a protein that was highly upregulated upon DNA damage in a variety of human cancer cell lines; in particular it was upregulated when double-stranded breaks occurred. The authors named the protein HJURP based on the observation that it had an affinity for binding holliday junction DNA *in vitro.* HJURP was found to colocalize with key DNA damage pathway proteins including NBS1 and hMSH5. The researchers also found HJURP levels were at least indirectly regulated by ATM, as upon knockdown of this protein in U2OS cells HJURP protein levels were non-detectable. These results altogether, according to the authors, implied that HJURP was somehow involved in DNA repair, in particular homologous recombination[69]. As of this writing no further work on HJURP, and it's potential role in the DNA repair pathway, has been published.

Later on HJURP was co-purified with prenucleosomal CENP-A suggesting it might be chaperoning CENP-A to centromeres, as other canonical histones are bound with

prenucleosomal chaperones before they are deposited into chromatin. HJURP was found to localize specifically to centromeres during a short period in early G1-phase just after cells went through anaphase, which is the same time CENP-A is deposited[33,70,71]. Upon depletion of HJURP, CENP-A was lost at centromeres and it was confirmed that newly synthesized CENP-A (using a SNAP labeling method) was completely dependent on HJURP for deposition[70,71]. *In vitro* pull-down analysis using recombinant GST-HJURP along with CENP-A/H4 tetramer and H2A/H2B revealed that HJURP specifically and directly bound to CENP-A/H4[71]. Taken together, these results show very thoroughly that HJURP is the CENP-A specific chaperone.

As of this writing there is no crystal structure for full-length HJURP but the protein has been broken down into several domains, each with a unique function. Starting at the N-terminus is a highly conserved domain of HJURP known as the Scm3 domain. This domain spans the first 80 amino acids of HJURP and is required for CENP-A/H4 binding in both yeast and humans[12,43,72,73]. In *S. cerevisiae* and *S. Pombe*, the protein Scm3 functions analogously (and is highly sequentially conserved) in place of HJURP in that it is required for Cse4*S.c.*/Cnp1*S.p.* maintenance and assembly at centromeres. A second domain downstream of the Scm3 domain in humans and other vertebrates is known as the Conserved domain. Meghan Barnhart-Dailey, a post-doctoral researcher in our lab, has identified this region of HJURP as an RNA binding domain that is required for CENP-A deposition. Lastly, there are two conserved C-terminal regions of HJURP. Recent work in our lab has identified and named these domains as the R1 and R2 repetitive regions. The R1 domain (aa 348-555) was found to be required for HJURP

centromere localization while the R2 containing domain (aa 555-758) is required for HJURP homo-dimerization[74].

Although the crystal structure of full length HJURP has not been solved, as of this writing, the crystal structure of hScm3/Scm3 binding to the CENP-A/Cse4-H4 in humans, budding yeast, and *K. lactis* has been solved[40,75,76]. Strikingly, there is incredible conservation when comparing the structure of the Scm3 domain across species. Scm3 is found to form an N-terminal α-helix followed by a 3-stranded β-sheet. The α-helix runs along next to and interacts with the α-2 helix within the histone fold domain of CENP-A/Cse4, which is located within the CATD. This binding of HJURP/Scm3 precludes the α-2 helix of CENP-A/Cse4 from self-dimerizing and forming a tetramer, which is seen in the CENP-A/H4 heterotetramer structure that is not bound to the Scm3 domain of HJURP[34]. The β-sheet of HJURP/Scm3 interacts along the positively charged groove of CENP-A/Cse4 that is necessary for DNA binding. Thus, HJURP/Scm3 also functions to preclude CENP-A/Cse4 form binding DNA prematurely[40,75,76]. As stated above, previous work in our lab has found HJURP has a C-terminal self dimerization domain, this domain is thought to allow the two heterotrimeric HJURP-CENP/H4 complexes to localize to centromeres in conjunction with one another, thereby allowing for octameric nucleosome assembly[74].

As mentioned above, in *S. cerevisiae* and *S. Pombe*, the protein Scm3 functions analogously in place of HJURP in that it is required for Cse4*S.c.*/Cnp1*S.p.* maintenance and assembly at centromeres[72,77–80]. Although the primary sequences between HJURP and Scm3 are diverged they both share a highly conserved 62 amino acid CENP-A

binding region (Scm3 domain in hHJURP) in the N-terminus despite Scm3 in yeast having none of the C-terminal components found in human HJURP[78].

In humans and in fission yeast, hHJURP/*S.p.* Scm3 requires Mis18 to be recruited to their regional centromeres. In budding yeast there are no Mis18 homologs found as of yet, Scm3 instead localizes their point centromeres by binding to Ndc10, a protein that recognizes the CDEII sequence at the budding yeast centromere[77,81]. In this system Scm3 also binds to AT-rich DNA (which is enriched at budding yeast point centromeres) which could make this another avenue for Scm3 recruitment to centromeres[82]. Our lab (in specific the research in this paper) has found Mis18 and Scm3 interact in fission yeast and this interaction is conserved to humans where HJURP directly binds to the Mis18α-β complex[60]. We will show in this thesis that the physical interaction between Mis18 and HJURP/Scm3 is absolutely required for proper HJURP centromere recruitment and CENP-A deposition.

In vitro chromatin assembly assays using just the Scm3 domain of hHJURP have found it is sufficient to assembly centomeric nucleosomes using recombinant Scm3 and histones. As of this writing it is not known whether the Mis18 complex can do this as well or if it plays any type of active role in the CENP-A deposition machinery along with its function of recruiting HJURP-CENP-A/H4 to the centromere[12,70]. Although the Scm3 domain of hHJURP is sufficient to assemble CENP-A/H4 into nucleosomes *in vitro*, full-length HJURP has not been shown, at least as of this writing, to have the same activity. One hypothesis is that the Scm3 domain alone is constitutively active and hyper-deposits CENP-A/H4 into nucleosomes when not in the presence of the rest of HJURP, which could be having regulatory effects on Scm3's

ability to deposit CENP-A. Furthermore, although the Scm3 domain alone is sufficient to assemble CENP-A chromatin *in vitro*, the *quality* of the chromatin cannot really be assayed other than to say nucleosomes are present, whether there is proper spacing, DNA wrapping on a nucleosome-by-nucleosome basis has not been thoroughly researched. It could be full-length HJURP regulates the Scm3 domain's ability to deposit CENP-A into chromatin and other factors are necessary and needed along with full length HJURP to get proper CENP-A deposition *in vitro* and in cells. A possible candidate for a new member of the CENP-A deposition machinery could be the Mis18 complex, and we will go over some data in this thesis that suggests this complex might very well be necessary and/or sufficient to assemble CENP-A chromatin in conjunction with HJURP.

Centromere Priming Components: Mis18

Mis18 was originally found in fission yeast and exists as a single protein that when knocked down causes massive chromosome mis-segregation[83]. In humans and higher eukaryotes there are two Mis18 homologs, Mis18α and Mis18β, along with a third partner that has no official homolog in fission yeast as of now, Mis18BP1. Depletion of any one of the complex components prevents the whole complex from localizing to centromeres[84]. Mis18α and Mis18β exist in a pre-chromatin complex that is independent of Mis18BP1 and all three complex members come together on chromatin (Stellfox et al. Under Review 2014)[84]. Mis16 (RbAP46/48 in humans) was also found to interact with Mis18 in fission yeast and its knockdown also caused massive chromosome mis-segregation defects. Depletion of Mis18 or Mis16 in fission yeast resulted in depletion of CENP-A/Cnp1 at centromeres[83]. This result has been recapitulated in our lab in human cells whereby knockdown of any Mis18 component abolishes CENP-A loading at

centromeres[12]. Taken together, these results suggest Mis18 is involved in the CENP-A deposition pathway in some way that is conserved across species. Some species, such as *C. elegans* (which surprisingly has a Mis18BP1 homolog) and *Drosophila*, completely lack any known Mis18α and Mis18β homologs[85]. Despite this, for the vast majority of species Mis18 is highly conserved and absolutely required for proper CENP-A deposition every cell cycle.

In *Xenopus* and humans, the Mis18 complex localizes to centromeres just after the cell exits mitosis, immediately prior to CENP-A deposition[67,84]. This is in contrast to fission yeast where Mis18 occupies the centromere throughout the cell cycle except during a short window in mitosis[80].

Research done in our own lab has shown Mis18 localizes to centromeres immediately prior to HJURP recruitment and subsequent CENP-A deposition. Knockdown of any of the three Mis18 components not only abolishes CENP-A deposition but also HJURP localization at centromeres[12]. This suggests that Mis18 is involved in the recruitment process of HJURP and other CENP-A deposition machinery; as of this writing how it is doing this is unknown. The centromere community has suggested the Mis18 complex epigenetically "primes" the centromere to make it compliant for HJURP recruitment and CENP-A deposition. Temperature sensitive Mis18 mutants in fission yeast showed increased acetylation at centromeres[83]. This is in contrast to human cells where HDAC inhibitors actually rescues the Mis18 depletion phenotype and restores CENP-A at centromeres[84]. Recent work has shown Mis18α interacts directly with methyl-transferases DNMT3A and DNMT3B that somehow require Mis18α for their methylation activity at centromeres. Depletion of Mis18α led to lack of methylation at centromeres and

decreased CENP-A levels[86]. Finally, The interaction of the Mis18 proteins with Mis16/RbAP46/48 discussed above suggests they may play a role in chromatin remodeling at the centromere. Although the Mis18 complex is clearly helping establish the epigenetic landscape at centromeres, how it is doing this and what the significance of this is in reference to HJURP recruitment and CENP-A deposition has not be fully fleshed out.

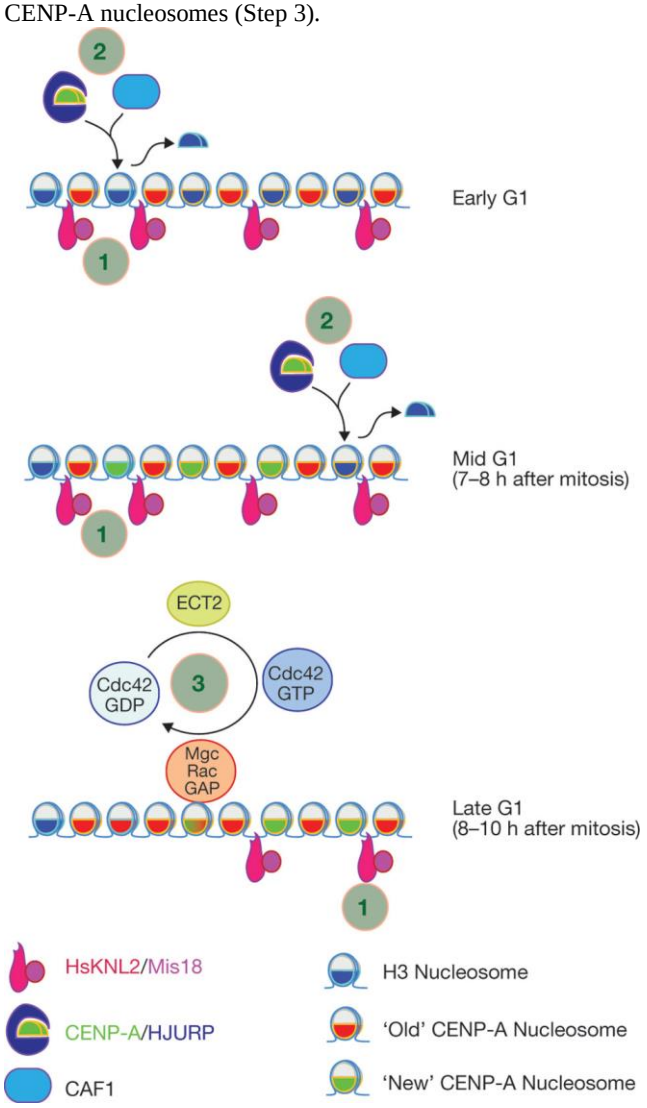
Some have hypothesized that the Mis18 complex could be directly recruiting HJURP through a physical interaction. In fission yeast it was revealed that Scm3 could directly binding to Mis18 *in vitro*[78]. Our lab, along with others[60], have established that the Mis18 complex does indeed have a direct interaction with HJURP *in vitro* and *in vivo.* These findings will be explained more in the second chapter of this work. There has not been a robust interaction detected *in vivo* between HJURP/Scm3 and the Mis18 complex, this suggests the interactions are being highly regulated and are short-lived in the cell on a temporal scale. In line with this, Mis18 and HJURP only colocalize to centromeres during early G1-phase in cells [71]. Research in this dissertation (particularly in chapter 2) will be elucidating a mechanism that can explain why the HJURP-Mis18 interaction is so short lived and how this could be involved in proper CENP-A deposition and regulation during G1-phase of the cell cycle.

It is clear that the highly conserved Mis18 complex, in organisms that contain it, is absolutely indispensable for proper HJURP recruitment and CENP-A deposition. Exactly how the Mis18 complex is acting on centromeric chromatin to facilitate CENP-A deposition is not fully understood. Furthermore, the consequence of Mis18 interacting directly with HJURP and how it pertains to CENP-A deposition has not been thoroughly

researched as of this writing. On an even more basic level, the stoichiometry, interaction domains, and overall structure of the Mis18 complex has not been looked into at all. Chapter 2 of this dissertation will be going over some of these basic questions and thereby shedding more light on the highly important, yet under-studied Mis18 complex. An overall process of how Mis18 “primes” centromeric chromatin can be seen in Figure 1-4.

Figure 1-4: Centromere specification is modulated by the Mis18 complex and associated factors.

During G1, the centromere is specified in a three-step process. After licensing (1) and loading (2)MgcRacGAP and ECT2 cycle Cdc42 GTPase activity to modify newly incorporated CENP-A nucleosomes to make them molecularly identical to pre-existing CENP-A nucleosomes (Step 3).



DNA Replication Licensing

Replication origins are licensed by the formation of a stable pre-replicative complex (pre-RC), which binds throughout G1 phase until replication is initiated in S-phase. Similar to the control of Mis18 recruitment, the assembly of the pre-RC is inhibited by Cdk activity[87–89]. After mitosis, the drop in Cdk activity allows both Mis18 and the RC complexes to bind to their respective sites on chromatin. As of now a mechanism for centromeric licensing, similar to that of DNA replication licensing, has not been put forth in the field. Cdt1 and Cdc6 interact directly with the MCM components and are required to load them onto DNA[90]. Activation of the MCM2–7 complex to initiate origin firing and DNA replication occurs during S-phase[91]. Following activation of the origin, several mechanisms inhibit re-association of Cdt1 including binding to Geminin and degradation by ubiquitylation, which therefore inhibits re-licensing until the following M/G1 phase[92,93].

Licensing of DNA to be replicated once and only once during the cell cycle is of utmost importance to the cell in order to maintain its genetic stability. When cells have unreplicated or damaged DNA, checkpoint control mechanisms prevents the initiation of M-phase by blocking the activation of the mitotic Cdc2 protein kinase until the DNA is completely replicated or repaired. Blockage of re-replicating DNA is further controlled by preventing origin re-initiation and firing. Fusion experiments in mammalian cells were performed where G1-phase cells were fused with S-phase cells. The G1 cells prematurely pushed into replictation[94]. The model for preventing re-replication of DNA put forth a hypothesis that cells contained “licensing factors” that would access chromatin only at

the end of M-phase and would be degraded or inactivated after the initiation of replication[95].

A very important group of licensing factors is a protein complex known as the Origin Recognition Complex (ORC) that binds to replication origins. In *S. cerevisiae* replication origins have been highly characterized. Generally they are a 100-200bp DNA sequence that contain an 11bp domain called the Autonomously Replicating Sequences Consensus Sequence. Fractionation of protein factors that bind this DNA sequence were performed and Orc1 and Orc6 were purified and characerized[96]. The mutation or deletion of ORC genes showed that all ORC components are essential. ORC proteins were identified in several eukaryotes, including *S. pombe*, *Drosophila melanogaster*, *Xenopus laevis* and human[97].

Cdc6 was found in budding yeast and it's homolog, cdc18, was later found in fission yeast[98,99]. Both of these proteins accumulate at the end of mitosis and a mutant of cdc18 in *S. pombe* cannot initiate replication but can still enter mitosis resulting a lethal phenotype[100]. Additionally, ScCdc6 has been implicated in inactivating mitotic Cdk's during the exit from mitosis[101]. Mammalian homologs of Cdc6/18 have been isolated as well[102] and are hypothesized to be carrying out similar function as their yeast counterparts.

The MCM (mini-chromosome maintenance) complex were originally isolated as genes from a screening for mutations that cause high rates of mini-chromosome loss in budding yeast[103]. Biochemical studies have revealed that the MCM complex works as a replicative helicase in a trimeric complex consisting of Mcm-4, -6, and -7. Weak ATPase and helicase activity has been seen *in vitro* for this complex[104].

Taken together, these findings highlight how DNA replication is controlled to happen only once during the cell cycle. Indeed, if this process becomes unregulated aberrant replication occurs usually followed by cell death. Similar to DNA replication, centromeres are propagated on chromatin once and only once and at a specific window in the cell cycle. Both of these processes are temporally regulated by Cdk activity showing even more similarities between the them. As of now, no mechanisms for if and how centromeres are licensed has been put forward in the field. A brief overview of DNA replication licensing can be seen in Figure 1-5.

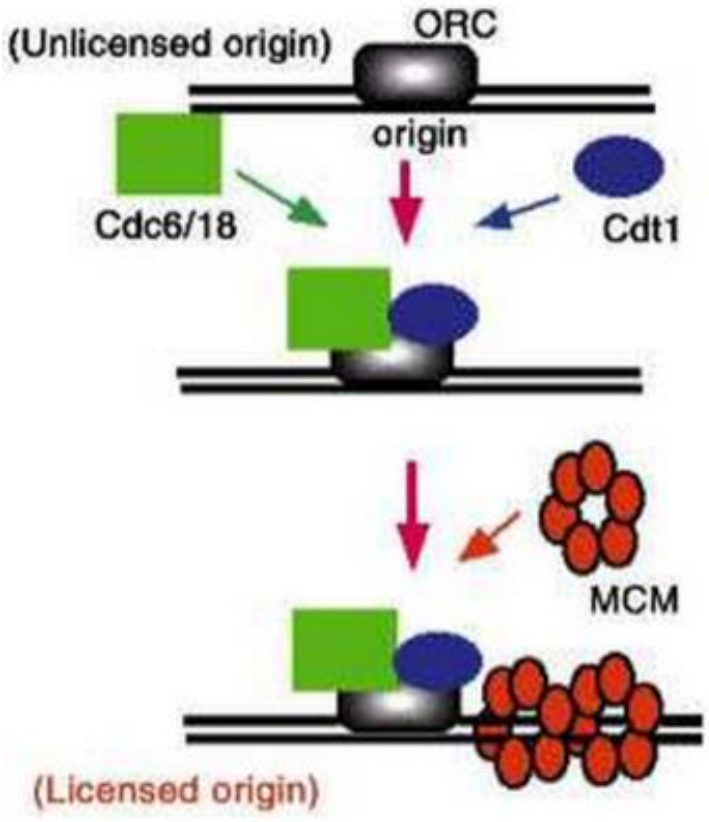


Figure 1-5: Establishment of DNA replication licensing

DNA replication licensing is established by a step-wise assembly process of initiator proteins that bind to the origin DNA. The ORC acts as a base for cdc16/8, which then loads the MCM complex onto chromatin. This process establishes DNA replication licensing.

Figure used with permission from Nishitani et al., 2002.

CHAPTER 2: Licensing of centromeric chromatin assembly through the Mis18α-Mis18β heterotetramer

(This chapter contains content and data from the manuscript of the same title in final review at the journal of Molecular Cell)

Abstract:

Centromeres are specialized chromatin domains that are specified by the centromere-specific CENP-A nucleosome. The stable inheritance of centromeres is an epigenetic process and requires the deposition of new CENP-A nucleosomes by HJURP. We show that HJURP is recruited to centromeres through a direct interaction between the HJURP centromere targeting domain and the C-terminal coiled-coil domain of the Mis18α–β complex. We demonstrate that Mis18α and Mis18β form a heterotetramer through the conserved coiled-coil domains in the C-termini of the Mis18 paralogs. *S.pombe* contains a single Mis18 isoform and forms a homotetramer; therefore, the tetrameric form of Mis18 is a conserved feature from humans to fission yeast. Fluorescence recovery after photobleaching demonstrated Mis18α is stably bound to centromeres, with little turnover. HJURP binding disrupts the Mis18α–β heterotetramer and removes Mis18α from centromeres. We propose that stable binding of Mis18 to the centromere in telophase licenses the site for CENP-A deposition. Binding of HJURP deposits CENP-A at the site and facilitates the removal Mis18, thus restricting CENP-A deposition at the site to a single event per cell cycle.

Introduction:

The centromere is an epigenetically specified chromosomal locus in humans that orchestrates the segregation of chromosomes during cell division. The centromere is defined by the presence of the histone H3 variant, CENP-A (Centromere Protein-A) (Cleveland et al. 2003; Hori et al. 2008; Wan et al. 2009). The presence of the CENP-A nucleosome is sufficient to recruit the constitutive centromere associated network (CCAN) and mitotic kinetochore proteins that are required for proper chromosome segregation (Barnhart et al. 2011; Guse et al. 2011; Mendiburo et al. 2011).

New CENP-A nucleosomes must be deposited at the existing centromere in every cell cycle to maintain centromere identity. CENP-A uses a conserved chromatin assembly factor HJURP (Holliday Junction Recognition Protein), or Scm3 (suppressor of chromosome segregation) in yeast, to distinguish it from other histone H3 variants and facilitates its deposition into centromeric chromatin (Dunleavy et al. 2009; Foltz et al. 2009; Shuaib et al. 2010; Barnhart et al. 2011; Bernad et al. 2011);Camahort et al. 2007; Mizuguchi et al. 2007; Stoler et al. 2007; Pidoux et al. 2009; Williams et al. 2009; Dechassa et al. 2011). HJURP binds the CENP-A/H4 heterodimer through the N-terminal Scm3 homology domain (Sanchez-Pulido et al. 2009; Shuaib et al. 2010; Hu et al. 2011) and contains a centromere targeting domain (CenTD) within the first C-terminal repeat that is absent in yeast Scm3 (Zasadzinska et al. 2013).

The Mis18 proteins are required for HJURP/Scm3 recruitment and therefore CENP-A/Cnp1 deposition in diverse eukaryotes with regional centromeres, but absent from yeast with point centromeres (Stoler et al. 1995; Hayashi et al. 2004; Camahort et al. 2007; Fujita et al. 2007; Mizuguchi et al. 2007; Pidoux et al. 2009; Williams et al. 2009;

Moree et al. 2011). *S. pombe* contains a single Mis18 protein; however vertebrates possess two Mis18 paralogs, Mis18α and Mis18β that share ~27% amino acid identity and each display 30% identity with *S.p.*Mis18. Humans and some higher eukaryotes possess a Mis18 associated protein Mis18BP1/Knl-2 (Fujita et al. 2007; Maddox et al. 2007). The human Mis18 proteins are recruited to centromeres during late telophase and remain associated with the centromere during early G1-phase when new CENP-A is deposited. Recruitment of the Mis18 complex is regulated by Cdk1 and Plk1 phosphorylation (Silva et al. 2012; Barnhart-Dailey and Foltz 2014; McKinley and Cheeseman 2014). Mis18α and Mis18β have been previously shown to affect histone modifications and the methylation status of underlying chromatin (Fujita et al. 2007; Kim et al. 2012). Eliminating histone methylation within the alpha satellite DNA repeats of a human artificial chromosome alters the recruitment of HJURP (Bergmann et al. 2011).

We demonstrate that human Mis18α and Mis18β associate through C-terminal coiled-coils to form a heterotetramer. The tetrameric stoichiometry of the Mis18 complex is conserved in fission yeast. We find that tetramer formation is essential for Mis18 centromere recruitment. Using FRAP we observe that Mis18α is highly stable at centromeres during G1-phase. The Mis18α-β heterotetramer is disrupted upon binding of the HJURP CenTD to the Mis18α-β coiled-coils, which leads to the removal of Mis18α–β from centromeres. We propose that these dynamics ensure each Mis18 complex bound per cell cycle directs the deposition of a single CENP-A nucleosome, and thereby limits the amount of new CENP-A recruited.

Results:

Mis18 proteins form a conserved tetramer.

We performed size exclusion chromatography (S.E.C.) and glycerol gradient sedimentation using recombinant human Mis18α-β complex (Figure 1A-D, Figure S1A-D). Mis18α or Mis18β alone migrated with a Stokes radius of 5.1 nm and 4.3 nm, respectively (Figure 1A-B, F) and sedimented with S-values of 2.7S and 3.2S (Figure 1A-B, F). Based on this analysis, Mis18α and Mis18β alone exhibit native molecular weights twice the size of their predicted monomeric molecular weights, consistent with formation of a homodimer (Figure 1F). The calculated molecular weight of the combined Mis18α-β complex, formed by incubating equivalent amounts of Mis18α and Mis18β shows that this complex exists as a heterotetramer (Figure 1C, F). *S. pombe*, as well as most non-vertebrate eukaryotes contains a single Mis18 homologue. Recombinant *S.p.*Mis18 formed a tetramer out of the single Mis18 protein (Figure 1D,S1A). Taken together, these data show that Mis18 proteins form an evolutionally conserved tetramer.

To determine whether the human Mis18α–β forms a tetramer *in vivo,* we determined the native molecular weight of the Mis18 complex from mitotic HeLa cells stably expressing GFP-tagged Mis18α. Chromatin free extracts (CFE) from nocodazole-blocked cells were analyzed by S.E.C. and glycerol gradient sedimentation (Figure 1E, F, Figure S1B-C). The native molecular weight of the cell-derived Mis18 complex was 154kD, close to the 152kD predicted molecular weight of a Mis18 tetramer containing two GFP-Mis18α (51kD) and two Mis18β molecules.

Mis18 proteins interact through conserved coiled-coils.

Sequence alignment of Mis18 proteins from a wide variety of eukaryotes identified the previously recognized YIPPEE domain and a second conserved region within the Mis18 C-termini (Figure 2A, S2A). The newly identified C-terminal conserved domains span amino acids 198-227 of human Mis18α and 189-219 of Mis18β and are predicted to form coiled-coils based on the MARCOIL program (Figure S2B,C) (Delorenzi and Speed 2002). To determine the contribution of the coiled-coils to Mis18 homodimer and heterotetramer formation we purified recombinant full length (FL), coiled-coil deletion (ΔCC), and coiled-coil alone (CC) fragments of Mis18α and Mis18β (Figure 2B, S1A). MBP-Mis18α^{CC} interacted with HisNusA-Mis18α^{FL} and HisNusA-Mis18α^{CC} fragments, but not with the YIPPEE/Mis18 domain alone (HisNusA-Mis18α$^{\Delta CC}$) (Figure 2C). Likewise, MBP-Mis18β^{CC} interacted with the Mis18β^{FL} and the Mis18β^{CC} but not the YIPPEE/Mis18 domain alone (Figure 2D). Therefore, Mis18α and Mis18β form homodimers through the coiled-coil domain.

To assess the role of the coiled-coil in mediating heterotetramer formation we determined if the C-terminal domains of Mis18α could bind Mis18β. We found that MBP-Mis18α^{CC} associated with HisNusA-Mis18β^{FL} and HisNusA-Mis18β^{CC} but not HisNusA-Mis18β$^{\Delta CC}$ (Figure 2E). The same held true for the reciprocal experiment when we used the coiled coil domain of Mis18β (MBP-Mis18β^{CC}) to pulldown HisNusA-Mis18α (Figure 2F). No interaction was detected between the coiled-coil and YIPPEE/Mis18 domains. Together these results show the Mis18 C-terminal coiled-coils contribute to

both the homodimerization and the formation of the heterotetramer, potentially forming a four-helix coiled-coil.

To assess whether the coiled-coils mediate homodimer and heterotetramer formation of Mis18α and Mis18β *in vivo* we targeted Mis18α to a non-centromeric site in U2OS cells containing a LacO array by fusing Mis18α to mCherry and the LacI repressor (mCLI) (Figure 2G) (Janicki and Spector 2003). mCLI-Mis18α was co-expressed with GFP-Mis18α or GFP-Mis18β and the recruitment of the GFP-tagged protein to the LacO array was analyzed. Mis18α recruited Mis18β to the array in 100% of transfected cells (Figure 2H). Consistent with our *in vitro* experiments the minimal coiled-coil domains were equally efficient and mediating the Mis18α-β interaction at the array (Figure 2H). Likewise, we observed efficient formation of the Mis18 or Mis18 homodimer at the array by targeting either paralog and expressing the GFP version of the same protein (Figure 2I, Figure S4E). Mutation of the conserved cysteine residues or other highly conserved amino acids within the YIPPEE/Mis18 domain (Mis18αY176A, Mis18βY172A) did not alter the ability of Mis18α and Mis18β to interact *in vitro* or *in vivo* (Figure S3A-D). Therefore, the coiled coil domains were sufficient to mediate the homo- as well as the heteromultimerization.

Mis18 localization requires formation of the heterotetramer via the coiled-coil domain.

Hydrophobic amino acids contribute to formation of the coiled-coils (Figure 2A, Figure S2A-C)(Grigoryan and Keating 2008). We made two double mutants of highly conserved hydrophobic residues at the 'a' and 'd' positions within the Mis18α coiled-coil (I201G/L205G, and L215G/L219G) (Figure S2C) and tested their effect on Mis18 complex formation. Both mutants abolished the predicted coiled-coil domain in

Mis18α but are not predicted to affect the alpha-helix (Figure 3A, S1A). The effect of these mutations on Mis18 homodimer and tetramer formation were tested using the *in-vitro* pulldown. HisNusA tagged mutant Mis18 proteins along with their wild-type counterpart were pre-assembled with either MBP-Mis18α^{CC} or MBP-Mis18β^{CC}. The Mis18α coiled-coil mutants were not able to pulldown either Mis18α or Mis18β wild-type or the coiled-coil domains (Figure 3B-C). The coiled-coil domain severely reduced the ability of Mis18α and Mis18β to homo- or hetero-multimerize *in vivo* when assayed at the lacO array (Figure 3D). These data demonstrate that the coiled-coil motif is responsible for the interaction between C-terminal alpha helixes of Mis18α and Mis18β.

In contrast, we found that the single amino acid mutation I201G selectively eliminated the ability of Mis18α to form the homodimer, but did not alter its ability to bind Mis18β. At the LacO array mCLI-Mis18α^{I201G} was able to recruit GFP-Mis18β, but failed to recruit GFP-Mis18α (Figure 3D). Likewise, Mis18α^{I201G} co-immunoprecipitated Mis18β but not Mis18α (□□□□□□□□E).□Therefore, the Mis18α^{I201G} mutant would be expected to form a Mis18α−β heterodimer.

The interaction of Mis18α and Mis18β with Mis18BP1 is critical for recruitment of the complex to centromeres (Fujita et al. 2007). We tested whether the association of the Mis18 complex with Mis18BP1 was affected by heterotetramer formation using the LacO array (Figure 3F). mCLI-tagged Mis18BP1 was tethered to the LacO array and the recruitment of FLAG-Mis18β and GFP-tagged Mis18α WT or coiled-coil mutants was examined. As expected, Mis18BP1 recruited Mis18αWT and Mis18β. In contrast, mutations that completely disrupted the Mis18 heterotetramer (Mis18α$^{I201G/L205G}$, Mis18α$^{L215G/L219G}$) eliminated the ability of Mis18BP1 to recruit Mis18α and Mis18β.

Likewise, the Mis18α^{I201G} mutant, which is expected to bind Mis18β and form a heterodimer, was unable to be interact with mCLI-Mis18BP1 at the array.

GFP-tagged Mis18α$^{I201G/L205G}$, Mis18α$^{L215G/L219G}$ or Mis18^{I201G} were transfected into U2OS cells to examine the role of heterotetramer formation in centromere localization. Localization of the coiled-coil mutants to centromeres was completely abolished (Figure 3G,S4F). GFP-tagged deletions containing either the YIPPEE/Mis18 or coiled-coil domains of Mis18α or Mis18β did not localize to centromeres; therefore, the coiled-coil domains are not sufficient to determine centromere localization (Figure S4A-D). Mutation of the conserved cysteine residues (Fujita et al. 2007) or other highly conserved amino acids within the YIPPEE domain also results in a loss of centromere localization (Figure S3E). Together these data show the functional YIPPEE/Mis18 domain and heterotetramer formation through the coiled-coil are both required for the complex to associate with centromeres. The Mis18α–β heterodimer is not sufficient to recruit Mis18 to centromeres or to interact with Mis18BP1. This suggests that multiple YIPPEE/Mis18α motifs are required for the Mis18 complex to stably bind centromeres.

Mis18 recruits HJURP through a direct interaction with the coiled-coil domain.

Mis18 precedes HJURP to centromeres and is required for HJURP centromere recruitment (Fujita et al. 2007; Barnhart et al. 2011). To determine if Mis18α-β are sufficient to recruit HJURP and deposit CENP-A independently of other centromere proteins we re-localized Mis18α-β to the non-centromeric LacO array (Figure 4A). Cells were co-transfected with mCLI-Mis18α, mCLI-Mis18β, HA-HJURP and HA-CENP-A. Under these condition CENP-A was robustly recruited the mCLI-Mis18 bound arrays (Figure 4A, left panel). CENP-A was retained at the LacO array following IPTG

treatment, which disrupts binding of LacI, consistent with stable CENP-A nucleosome assembly at the array (Figure 4A, right panel). We showed previously that mCLI-HJURP was able to assemble CENP-A nucleosome and these domains facilitated the formation of active centromeres and kinetochores during mitosis (Barnhart et al. 2011; Bassett et al. 2012). CENP-A was retained in a similar percentage of IPTG treated mCLI-Mis18α–β expressing cells as cells expressing mCLI-HJURP.

To determine if the recruitment of HJURP to centromeres by Mis18α–β is due to a direct interaction, we expressed and purified recombinant full-length MBP-HJURP (Figure 4B, S1A). MBP-HJURP was incubated with differentially-tagged Mis18α and Mis18β separately or with the pre-formed Mis18α–β heterotetramer. *In vitro* MPB-pulldowns showed that both Mis18 proteins were able to bind HJURP alone (Figure 4B). Binding was more robust (~2-fold) between HJURP and the Mis18α–β heterotetramer.

The Mis18 coiled-coils recruit HJURP.

In order to determine the domain of Mis18α–β required to recruit HJURP, mCLI and GFP-tagged Mis18 truncation mutants were co-expressed along with HA-HJURP in the LacO-containing U2OS cells. Targeting of the Mis18α–β coiled-coils together to the array by co-expressing mCLI-Mis18α^{CC} and GFP-Mis18β^{CC} or vice versa was sufficient to recruit HJURP to the array and as efficient as full-length Mis18 (Figure 4C). Targeting the YIPPEE domain alone (Mis18α/$\beta^{\Delta CC}$) was not sufficient to recruit HJURP. Mis18α/β mutants that disrupt the YIPPEE domains (but not the coiled-coil) and abolish Mis18 centromere recruitment (Figure S3A,E) were tethered to the array and recruited HJURP similarly to their WT counterparts (Figure S5A). This again shows the coiled-coils are the

minimal domain sufficient to recruit HJURP to chromatin. Each coiled-coil alone also recruited HJURP to the array but to a lesser degree than when Mis18α and Mis18β were present. This recapitulates our *in vitro* results showing that Mis18α or Mis18β alone both bound HJURP, albeit less well when compared to the complete Mis18α–β complex (Figure 4B).

HJURP binding disrupts the Mis18 heterotetramer.

Previously we identified a centromere targeting domain (CenTD) of HJURP (a.a.348-482) (Zasadzinska et al. 2013). *In vitro* MBP pulldowns revealed that the Mis18 complex bound exclusively to the R1 fragment containing the HJURP CenTD (Figure 4D,E, S1A). This demonstrates that the HJURP CenTD functions as a direct binding site for the Mis18 complex consistent with recent a recent report from Wang et al (2014).

To assess the stoichiometry of the Mis18 complex bound to the HJURP CenTD we performed S.E.C. and glycerol gradient sedimentation on the purified Mis18α-β complex bound to the MBP-HJURP^{R1} fragment (Figure 5A-C, S1A). The MBP-HJURP^{R1} fragment alone ran as a monomer (Figure 5A, S1D,F-G) consistent with previously reported data (Zasadzinska et al. 2013). To recapitulate the endogenous assembly of the Mis18-HJURP complex, we combined Mis18α and Mis18β prior to the addition of MBP-HJURP^{R1}. We observed the Mis18α-β complex shifted its migration to the same fraction as MBP-HJURP^{R1} following fractionation by S.E.C, consistent with a direct interaction between these proteins. However, the HJURP bound Mis18 complex displayed a smaller stokes radius than the Mis18α-β heterotetramer (4.7 nm versus 5.5 nm) (Figure 5B,C). Based on the S.E.C. and glycerol gradient sedimentation we calculated the size of the complex formed by Mis18α, Mis18β, and MBP-HJURP^{R1} to be 114 kD which is

consistent with a three-protein complex containing a single copy of Mis18α, Mis18β and MBP-HJURP^{R1} (Figure 5C, S1E-I). An *in vitro* pulldown was performed from the peak size exclusion fraction containing the Mis18-HJURP^{R1} complex (Figure 5D) showed that both Mis18α and Mis18β are bound to HJURP and bound to one another. Taken together, these results suggest the Mis18 heterotetramer is reduced to a Mis18 dimer upon binding to HJURP. Although we cannot completely eliminate the possibility that subcomplexes containing only one Mis18 paralog are also present.

To confirm these *in vitro* results *in vivo*, HEK293T cells were transfected with HA-Mis18α/β with or without GFP-HJURP^{R1}. Chromatin free extracts were analyzed by S.E.C. Similar to the *in vitro* results, when HJURP^{R1} was present there was a characteristic decrease in stokes radius of the Mis18 complex suggesting the Mis18 complex was being disrupted by HJURP^{R1} binding (Figure S5B).

To determine if HJURP interaction *in vivo* disrupts the hetero or homo-dimerization of Mis18 we used the LacO array in U2OS cells. mCLI Mis18α was tethered to the array and cotransfected with GFP-Mis18α and FLAG-Mis18β with or without HA-HJURP^{R1}. Cells transfected with mCLI-Mis18α at the array recruited both GFP-Mis18α and FLAG-Mis18β consistent with the formation of the heteromeric complex (Figure 5E-left panel,F). If the heterotetramer is disrupted into heterodimers by HJURP^{R1} we expect a loss of GFP-Mis18α fluorescence at the array and reduction of FLAG-Mis18β when HJURP^{R1} is transfected (Figure 5E- right panel). As predicted, transfection of HJURP^{R1} almost completely abolished the GFP-Mis18α signal and reduced the FLAG-Mis18β fluorescence by 50% when compared to cells with no HJURP^{R1} transfected (Figure 5F). These results are consistent with disruption of the Mis18 complex into

heterodimers and recapitulate our *in-vitro* pulldown results (Figure 5C). These data show that the Mis18α-β complex recruits HJURP to the centromere through direct binding of the CenTD and that HJURP disrupts the Mis18α–β heterotetramer into heterodimers upon binding.

Mis18α is stable and does not turnover at the centromere during early G1.

We hypothesized that for the Mis18α-β complex license centromeres, following its recruitment in late telophase new Mis18 proteins should not be recruited to centromeres. To determine if the Mis18α-β heterotetramer is stably bound to centromeres or if it cycles on and off during early G1 we performed fluorescence recovery assay using a HeLa cell line stably expressing GFP-Mis18α. We selectively photobleached Mis18α positive centromeres and allowed them to recover for 5 min (Figure 6A). GFP Mis18α was only able to recover 20.8% of the pre-bleach level (Figure 6B) showing that once the Mis18α-β complex is bound to the centromere in late telophase it does not actively turnover. This shows the Mis18α-β complex is not replaced at centromeres after binding of the complex in late mitosis.

HJURP facilitates removal of Mis18 from centromeres.

We showed that HJURP binding disrupted the Mis18 heterotetramer into heterodimers, and Mis18 mutants that favor heterodimer formation cannot bind to centromeres. We hypothesize that HJURP binding and disruption of the Mis18α-β complex is the final step in centromere licensing that promotes removal of the Mis18 complex from the centromere. Therefore, we predicted that in the absence of HJURP the Mis18 complex should persist at centromeres. To assess this hypothesis we performed HJURP or control

siRNA depletion in an asynchronously dividing GFP-Mis18α HeLa cell line (Figure 6C-E). Twenty-four hrs post depletion, cells progressing from metaphase to G1 were imaged for 6 hrs (Figure 6C-D). Consistent with our hypothesis, GFP-Mis18α persisted for longer time periods into G1 when HJURP was suppressed (Figure 6C). At 3 hrs post metaphase, 80% of cells treated with HJURP siRNA showed strong GFP-Mis18α puncta compared to control cells in which GFP-Mis18α persisted at centromeres for less than 2 hrs following metaphase (Figure 6C-D,S5C). The rate of GFP-Mis18α loss in the HJURP siRNA condition was 5-times slower than in controls. In the siRNA treated cells small amounts of HJURP may still be present despite significant suppression of HJURP and this may contribute to removal Mis18α over time. Alternatively, a secondary pathway may exist for the slow removal of Mis18α from centromeres that is independent of HJURP. Taken together, these data show that HJURP is required for dissociation of the Mis18 complex from the centromere in G1-phase

CENP-A levels at centromeres are controlled by Mis18.

Based on our model we predict that CENP-A levels should remains relatively constant across generations. To test this we performed live-cell fluorescence imaging of YFP-CENP-A across three cellular generations. Cells were imaged at the G1/S boundary in the first (1 cell), second (2 cell) and third (4 cell) generations (Figure S6A-C). We chose to measure YFP-CENP-A intensity at the G1/S boundary to ensure that centromere assembly was completed, and all cells contain a 2N DNA content. The mean CENP-A intensity between cells varied less than the distribution of centromeres within the same cell. There was no significant difference in the YFP-CENP-A intensities between generation I, II and III within the same lineage (Figure S6C). This is consistent with a

mechanism that pegs the amount of new CENP-A deposited in each cell cycle to the amount of existing CENP-A.

To test if the availability of Mis18 determines how much CENP-A is loaded we overexpressed GFP-Mis18α/β in U2OS cells and assessed the amount of CENP-A present at centromeres after 48 hrs (Figure 6F-G Figure S6D). There was a clear and significant increase in total CENP-A fluorescence at centromeres in cells that had the Mis18 complex overexpressed. In contrast, when HJURP was overexpressed the CENP-A levels were not significantly different from that of the control. This indicates that although HJURP is the CENP-A chaperone responsible for shuttling it to the centromere, it is the amount of Mis18 present that controls how much CENP-A is deposited every cell cycle.

Discussion

Here we demonstrate that human Mis18α and Mis18β form a Mis18α–β heterotetramer (HT) through a conserved coiled-coil domain within their C-termini (Figure 7). Using *S.pombe* Mis18 we show that tetramer formation is an evolutionarily conserved feature of the Mis18 complex across divergent eukaryotes. We find that once Mis18α is bound to centromere is early G1 it is highly stable. Mutations that completely eliminate coiled-coil formation, or favor Mis18 heterodimer formation ($Mis18^{HetD}$), show that $Mis18^{HT}$ formation is required for the complex to recognize Mis18BP1 and stably bind to centromeres. We observe that HJUPR binding dissociated the $Mis18^{HT}$ into heterodimers. We propose a model where binding of HJURP to Mis18α–β complex disrupts the heterotetramer into the $Mis18^{HetD}$ and eliminates the ability of the Mis18 complex to be retained at the centromere. The stable association of Mis18 at centromeres and its removal by HJURP binding limits CENP-A deposition to a single round and couples the amount of new CENP-A to the pre-existing size of the centromere.

We and others have shown that mutations within the Mis18α or Mis18β YIPPEE/Mis18 domain, including the CXXC motifs, eliminate their ability to bind to centromeres. The YIPPEE/Mis18 domains must be required for recognition of the components involved in recruiting Mis18 to centromeres. These interacting proteins are likely to include Mis18BP1, as well as CCAN proteins such as CENP-C that are known to be involved in Mis18 centromere localization in humans, and the Eic1 (a.k.a. Mis19) protein in *S.pombe* which bridges the recognition of the spMis18 to the CCAN (Hayashi et al. 2004; Fujita et al. 2007; Maddox et al. 2007; Moree et al. 2011; Dambacher et al. 2012; Subramanian et

al. 2014). In addition, we demonstrate that the YIPPEE/Mis18 domain must be organized into the tetramer by the C-terminal coiled-coil for centromere localization to occur, as point mutations of the residues that contribute to coiled-coil formation eliminate alter Mis18BP1 recognition and recruitment of the complex to centromeres.

The formation of the $Mis18^{HT}$ and the requirement for the coiled-coil domain for hetero and homotypic interactions suggest that the C-termini of Mis18α and Mis18β form a 4-helix coiled-coil. Based on the observation that both Mis18 paralogs can form homodimers in isolation we predict that the Mis18α:Mis18β stoichiometry of the complex is 2:2. However, we cannot eliminate the possibility that the endogenous complex contains a 3:1 ratio. The exact arrangement of the subunits within the $Mis18^{HT}$ is not known. Coiled-coils can form parallel or anti parallel (Grigoryan and Keating 2008), so it conceivable that YIPPEE/Mis18 domains of the alpha subunits may be closely juxtaposed or at opposite ends of the coiled coil structure. Given that Mis18α and Mis18β form both homo and hetero dimers, we favor a pattern in which each coiled-coils make homotypic and heterotypic interactions within the structure (Figure S2). In favor of this model we see that mutations in Mis18α at the ‘a‘ and ‘d’ position affect both homdimer and heterodimer formation; however, a single mutation at position ‘d’ position ($Mis18\alpha^{I201G}$) disrupts only homodimer formation, suggesting that homodimer and heterodimer interactions may take place on different "faces" of the coiled-coil. In depth structural analysis of the coiled-coil of Mis18α and Mis18β will be required to fully address these questions. HJURP and Scm3 directly interact with Mis18 (Pidoux et al. 2009; Wang et al. 2014) (Figure 4). However, Scm3 lacks the CenTD required for HJURP localization. This suggests that HJURP and Scm3 use different domains to

recognize a conserved feature of the Mis18, and we propose that the conserved of Mis18 is the C-terminal Coiled-coil that we described here.

We have shown that total CENP-A levels at centromeres do not vary across multiple generations (Figure S6A-C). This suggests that the process of new CENP-A deposition is restricted in order to ensure that amount of new CENP-A deposited in G1 is equal to the amount of CENP-A already present in the centromere. This can be achieved through a mechanism whereby a single CENP-A nucleosome determines the deposition of a new CENP-A nucleosome once and only once. This process is reminiscent of replication licensing, which allows each replication origin to be activated only once per cell cycle (Machida et al. 2005).

Here we provide direct evidence that Mis18 acts as a licensing factor for centromere deposition, similar to regulation of DNA replication by the pre-replication complex (pre-RC). Replication origins are licensed by the formation of a stable pre-replicative complex (pre-RC), which binds throughout G1 phase until replication is initiated in S-phase. Similar to the Pre-RC, we show that the $Mis18^{HT}$ complex is stably recruited to the centromeres prior to new CENP-A deposition. Temporal control of Mis18 recruitment and the assembly of the pre-RC are inhibited by Cdk activity (Dahmann et al. 1995; Hua et al. 1997; Noton and Diffley 2000; Silva et al. 2012). The drop in Cdk activity upon exit from mitosis allows both the pre-RC and Mis18 complexes to bind their respective sites within chromatin.

In vitro and in cells, the binding of HJURP disrupts the Mis18 heterotetramer, forming a $Mis18^{HetD}$ and eliminates the ability of Mis18 to continue to recongnize Mis18BP1 and bind the centromere. Likewise, following activation of the MCM2–7 complex to initiate

origin firing and DNA replication, several mechanisms inhibit re-association of the MCM loaders Cdt1 and Cdc6, including binding to Geminin and degradation by ubiquitylation, to inhibit re-licensing until the following M/G1 phase (Wohlschlegel et al. 2000; Diffley and Labib 2002; Aladjem 2007; Truong and Wu 2011). Following the disruption of the $Mis18^{HT}$, additional interactions with the centromere may contribute to association of HJURP and the final deposition of CENP-A. Recently, Muller and colleagues (Muller et al. 2014) identified DNA binding activity of the "conserved" domain of HJURP that may contribute to the association of HJURP with the centromere once Mis18 has recruited it.

Consistent with our observation in human cells, artificially decreasing the amount of CENP-A (a.k.a. CID) levels in Drosophila sperm causes a heritable reduction in the amount of CENP-A at centromeres (Raychaudhuri et al. 2012). This demonstrates that the mechanism which pegs the amount of new CENP-A to the amount already present at the centromeres despite the divergent centromere position pathway involving the Cal1 chromatin assembly factor (Erhardt et al. 2008; Mellone et al. 2011; Chen et al. 2014). Chronic overexpression of CENP-A can lead to increased CENP-A at centromeres, suggesting that over time the high levels of CENP-A can push this system toward over-deposition (Bodor et al. 2014). However, the same experiment showed no increase in CENP-C with increased CENP-A, leaving open the possibility that a core set of CENP-A nucleosomes is differentially recognized and subject to licensing as we describe here.

The highest degree of Mis18α recruitment occurs in late-telophase and the amount of Mis18 present steadily decreases through early G1 (Figure 6A,C), suggesting that Mis18 is loaded onto chromatin in a single event and is gradually removed from centromere as HJURP binds Mis18 and new CENP-A is deposited. In the HJURP siRNA treated

condition Mis18α was still lost from centromeres but at a much lower rate. This may be due to incomplete removal of HJURP, or through additional mechanisms to inhibit Mis18 rebinding to centromeres such as ubiquitylation-mediated degradation. Ubiquitylation by βTrCP, has been shown previously to regulate Mis18β protein levels (Kim et al. 2014). HJURP binding may eliminate the Mis18 heterotetramer and directly affect the ability of Mis18 to recognize the centromere, but it may also facilitate the proteasome-mediated degradation of Mis18. In addition, βTrCP mediated ubiquitylation may also degrade Mis18 that is not bound to chromatin following the initial binding phase in late telophase. We observed that overexpression of Mis18 resulted in increased CENP-A deposition within one or two cell cycles. Increasing Mis18 levels (~9-fold, Figure S6D) may swamp the ubiquitylation system and result in more CENP-A assembly because Mis18 that would usually be degraded is available to re-bind an already assembled site.

Centromeric licensing through the Mis18α–β heterotetramer may cooperate with other mechanisms to restrict CENP-A to pre-existing centromeres. The ubiquitylation and degradation of CENP-A by Psh1 in yeast and Ppa in Drosophila at non-centromeric sites avoids the accumulation of CENP-A outside of the pre-existing centromere and ensures the propagation of CENP-A chromatin exclusively at existing centromeres (Hewawasam et al. 2010; Ranjitkar et al. 2010; Moreno-Moreno et al. 2011). In addition, Histone H3K9 dimethylation makes heterochromatin regions refractory to centromere formation and helps to spatially restrict CENP-A nucleosome deposition to the centromere (Lam et al. 2006). Together these mechanisms ensure the stable inheritance of CENP-A containing chromatin.

Materials and Methods:

Cell transfections and immunocytochemistry. Parental U2OS and LacO-TRE (Janicki et al. 2004) cells were transfected with Lipofectamine-2000 (Invitrogen) according to the manufacturers protocol. Cells were processed for immunocytochemistry 48 hrs after transfection. Immunocytochemistry and immunoprecipitation experiments were conducted as described previously (Zasadzinska et al. 2013). Metaphase spreads were prepared as described in the supplemental materials and methods. Images of fixed cells were collected using a 100× oil immersion objective lens on a DeltaVision deconvolution microscope (Applied Precision) using SoftWoRX acquisition software. Images were deconvolved and presented as stacked images. Images within experiments were collected with identical exposure times and scaled equally.

In-vitro pulldowns. Recombinant proteins were purified as described in the supplemental materials and methods. Proteins were combined for 3 hrs at room temperature at 1:1 molar ratio in binding buffer: 50mM Tris-HCl pH 7.5, 250mM NaCl, 20mM $MgCl_2$, 0.5% NP-40, 10% glycerol, and 5mM BME. Pulldowns between Mis18 proteins were conducted using 160nM recombinant protein in each condition. For pulldowns involving HJURP, 125 nM Mis18α dimer, 125nM Mis18β dimer or 125nM Mis18 α-β tetramer were pre-incubated for 3 hrs to form the complex prior to adding 62.5 nM HJURP. Affinity matrices were pre-incubated in the binding buffer with 0.2 mg/ml BSA for 1 hr and added to pre-formed complexes for 40 min. Complexes were washed 6-times in 1ml of binding buffer, for Ni-NTA 40mM imidazole was added. Bound complexes were eluted in SDS sample buffer and boiled. Western blots were performed using antibodies

against the 6X His (Santa-Cruz), MBP (New England Biolabs), Mis18β (Bethyl Labs), and the HA epitope (Covance).

Hydrodynamics. Stokes radii were determined by S.E.C. on Superdex 200 10/300 column (GE-Healthcare) in buffer containing 50mM Tris-HCl pH 7.5, 350mM NaCl, 2.5% glycerol, 0.05% NP-40, and 5mM BME. Proteins were pre-mixed with one another at a 1:1 molar ratio at a final concentration of 7.4 uM for all conditions. Fractions (1 ml) were analyzed by western blot. Gradient sedimentation was performed on a 12ml 10-20% glycerol gradient made using a BioCOMP gradient station in buffer containing 50mM Tris-HCl pH 7.5, 350mM NaCl, 0.05% NP-40. Gradients were centrifuged at 40,000 rpm, for 36 hrs at 4ºC in a SW41Ti rotor and 500ul fractions collected. Native molecular weights of the Mis18 proteins were calculated based on the measured S-values and stokes radii (Siegel and Monty 1966).

Live-cell imaging and siRNA depletion GFP-Mis18α expressing HeLa cells were plated into 8 well coverglass (Thermo) ($1.0x10^4$ cells/well) and transfected with 20nM HJURP (Ambion, #S30814) or control siRNA (Ambion, #4390846) using RNAiMax (Thermo) Cells were imaged 24 hrs later at 37ºC in 5% CO_2 using a 60× oil immersion objective lens (numerical aperture = 1.40; Olympus) on a deconvolution microscope (DeltaVision) using a camera (CoolSNAP HQ2; Photometrics). Images were collected beginning in metaphase at 20 min intervals for 6 hr.

Fluorescence Recovery after Photobleaching Stable HeLa GFP-Mis18α expressing cells were grown on glass-bottomed culture dishes (MatTek Corporation). Prior to imaging, growth media was replaced with Leibovitz's L-15 medium without phenol red (Gibco) with 10% FBS (Optima, Atlanta Biologicals). Photobleaching was conducted using a

Zeiss LSM 510 UV Confocal Microscope. Two pre-bleach images were collected. Individual centromeres were bleached with 70 iterations of the 488 nm laser and the fluorescence recovery at the centromere was assessed at 10-second intervals. Fluorescence recovery at photobleached centromeres was analyzed using ImageJ, normalized to account for sample bleaching (Phair et al. 2004), and average fluorescence recovery data (GFP-Mis18α, n=18) was fit to a single exponent curve $A*(1-e^{-kt})$.

Acknowledgements:

We thank I. Cheeseman and D. Cleveland for reagents. We thank D. Burke and P.T. Stukenberg for helpful comments on the manuscript. D.R.F was supported by an American Cancer Society Research Scholar Award, March of Dimes Basil O'Conner Award and NIH R01GM111907. M.E.S. and I.K.N. were supported by NIH T32CA00910937. C.M.K. was supported by a Harrison Award from the University of Virginia.

References:

Aladjem MI. 2007. Replication in context: dynamic regulation of DNA replication patterns in metazoans. *Nat Rev Genet* **8**: 588-600.

Barnhart MC, Kuich PH, Stellfox ME, Ward JA, Bassett EA, Black BE, Foltz DR. 2011. HJURP is a CENP-A chromatin assembly factor sufficient to form a functional de novo kinetochore. *J Cell Biol* **194**: 229-243.

Barnhart-Dailey MC, Foltz DR. 2014. Centromere licensing: Mis18 is required to Polo-ver. *Curr Biol* **24**: R808-810.

Bassett EA, DeNizio J, Barnhart-Dailey MC, Panchenko T, Sekulic N, Rogers DJ, Foltz DR, Black BE. 2012. HJURP uses distinct CENP-A surfaces to recognize and to stabilize CENP-A/histone H4 for centromere assembly. *Dev Cell* **22**: 749-762.

Bergmann JH, Rodriguez MG, Martins NM, Kimura H, Kelly DA, Masumoto H, Larionov V, Jansen LE, Earnshaw WC. 2011. Epigenetic engineering shows H3K4me2 is required for HJURP targeting and CENP-A assembly on a synthetic human kinetochore. *EMBO J* **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. *J Cell Biol* **192**: 569-582.

Bodor DL, Mata JF, Sergeev M, David AF, Salimian KJ, Panchenko T, Cleveland DW, Black BE, Shah JV, Jansen LE. 2014. The quantitative architecture of centromeric chromatin. *eLife* **3**: e02137.

Camahort R, Li B, Florens L, Swanson SK, Washburn MP, Gerton JL. 2007. Scm3 is essential to recruit the histone h3 variant cse4 to centromeres and to maintain a functional kinetochore. *Mol Cell* **26**: 853-865.

Chen CC, Dechassa ML, Bettini E, Ledoux MB, Belisario C, Heun P, Luger K, Mellone BG. 2014. CAL1 is the Drosophila CENP-A assembly factor. *J Cell Biol* **204**: 313-329.

Cleveland DW, Mao Y, Sullivan KF. 2003. Centromeres and kinetochores: from epigenetics to mitotic checkpoint signaling. *Cell* **112**: 407-421.

Dahmann C, Diffley JF, Nasmyth KA. 1995. S-phase-promoting cyclin-dependent kinases prevent re-replication by inhibiting the transition of replication origins to a pre-replicative state. *Curr Biol* **5**: 1257-1269.

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

Dechassa ML, Wyns K, Li M, Hall MA, Wang MD, Luger K. 2011. Structure and Scm3-mediated assembly of budding yeast centromeric nucleosomes. *Nat Commun* **2**: 313.

Delorenzi M, Speed T. 2002. An HMM model for coiled-coil domains and a comparison with PSSM-based predictions. *Bioinformatics* **18**: 617-625.

Diffley JF, Labib K. 2002. The chromosome replication cycle. *J Cell Sci* **115**: 869-872.

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

Erhardt S, Mellone BG, Betts CM, Zhang W, Karpen GH, Straight AF. 2008. Genome-wide analysis reveals a cell cycle-dependent mechanism controlling centromere propagation. *J Cell Biol* **183**: 805-818.
Foltz DR, Jansen LE, Bailey AO, Yates JR, 3rd, Bassett EA, Wood S, Black BE, Cleveland DW. 2009. Centromere-specific assembly of CENP-a nucleosomes is mediated by HJURP. *Cell* **137**: 472-484.
Fujita Y, Hayashi T, Kiyomitsu T, Toyoda Y, Kokubu A, Obuse C, Yanagida M. 2007. Priming of centromere for CENP-A recruitment by human hMis18alpha, hMis18beta, and M18BP1. *Dev Cell* **12**: 17-30.
Grigoryan G, Keating AE. 2008. Structural specificity in coiled-coil interactions. *Current opinion in structural biology* **18**: 477-483.
Guse A, Carroll CW, Moree B, Fuller CJ, Straight AF. 2011. In vitro centromere and kinetochore assembly on defined chromatin templates. *Nature* **477**: 354-358.
Hayashi T, Fujita Y, Iwasaki O, Adachi Y, Takahashi K, 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 JL, Gerton JL. 2010. Psh1 is an E3 ubiquitin ligase that targets the centromeric histone variant Cse4. *Mol Cell* **40**: 444-454.
Hori T, Amano M, Suzuki A, Backer CB, Welburn JP, Dong Y, McEwen BF, Shang WH, Suzuki E, Okawa K et al. 2008. CCAN makes multiple contacts with centromeric DNA to provide distinct pathways to the outer kinetochore. *Cell* **135**: 1039-1052.
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.
Hua XH, Yan H, Newport J. 1997. A role for Cdk2 kinase in negatively regulating DNA replication during S phase of the cell cycle. *J Cell Biol* **137**: 183-192.
Janicki SM, Spector DL. 2003. Nuclear choreography: interpretations from living cells. *Curr Opin Cell Biol* **15**: 149-157.
Janicki SM, Tsukamoto T, Salghetti SE, Tansey WP, Sachidanandam R, Prasanth KV, Ried T, Shav-Tal Y, Bertrand E, Singer RH et al. 2004. From silencing to gene expression: real-time analysis in single cells. *Cell* **116**: 683-698.
Kim IS, Lee M, Park JH, Jeon R, Baek SH, Kim KI. 2014. betaTrCP-mediated ubiquitylation regulates protein stability of Mis18beta in a cell cycle-dependent manner. *Biochem Biophys Res Commun* **443**: 62-67.
Kim IS, Lee M, Park KC, Jeon Y, Park JH, Hwang EJ, Jeon TI, Ko S, Lee H, Baek SH et al. 2012. Roles of Mis18alpha in Epigenetic Regulation of Centromeric Chromatin and CENP-A Loading. *Mol Cell*.
Lam AL, Boivin CD, Bonney CF, Rudd MK, Sullivan BA. 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.
Machida YJ, Hamlin JL, Dutta A. 2005. Right place, right time, and only once: replication initiation in metazoans. *Cell* **123**: 13-24.
Maddox PS, Hyndman F, Monen J, Oegema K, Desai A. 2007. Functional genomics identifies a Myb domain-containing protein family required for assembly of CENP-A chromatin. *J Cell Biol* **176**: 757-763.

McKinley KL, Cheeseman IM. 2014. Polo-like kinase 1 licenses CENP-A deposition at centromeres. *Cell* **158**: 397-411.
Mellone BG, Grive KJ, Shteyn V, Bowers SR, Oderberg I, Karpen GH. 2011. Assembly of Drosophila centromeric chromatin proteins during mitosis. *PLoS Genet* **7**: e1002068.
Mendiburo MJ, Padeken J, Fulop S, Schepers A, Heun P. 2011. Drosophila CENH3 is sufficient for centromere formation. *Science* **334**: 686-690.
Mizuguchi G, Xiao H, Wisniewski J, Smith MM, Wu C. 2007. Nonhistone Scm3 and histones CenH3-H4 assemble the core of centromere-specific nucleosomes. *Cell* **129**: 1153-1164.
Moree B, Meyer CB, Fuller CJ, Straight AF. 2011. CENP-C recruits M18BP1 to centromeres to promote CENP-A chromatin assembly. *J Cell Biol* **194**: 855-871.
Moreno-Moreno O, Medina-Giro S, Torras-Llort M, Azorin F. 2011. The F Box Protein Partner of Paired Regulates Stability of Drosophila Centromeric Histone H3, CenH3(CID). *Curr Biol* **21**: 1488-1493.
Muller S, Montes de Oca R, Lacoste N, Dingli F, Loew D, Almouzni G. 2014. Phosphorylation and DNA binding of HJURP determine its centromeric recruitment and function in CenH3(CENP-A) loading. *Cell reports* **8**: 190-203.
Noton E, Diffley JF. 2000. CDK inactivation is the only essential function of the APC/C and the mitotic exit network proteins for origin resetting during mitosis. *Mol Cell* **5**: 85-95.
Phair RD, Gorski SA, Misteli T. 2004. Measurement of dynamic protein binding to chromatin in vivo, using photobleaching microscopy. *Methods Enzymol* **375**: 393-414.
Pidoux AL, Choi ES, Abbott JK, Liu X, Kagansky A, Castillo AG, Hamilton GL, 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.
Ranjitkar P, Press MO, Yi X, Baker R, MacCoss MJ, 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.
Raychaudhuri N, Dubruille R, Orsi GA, Bagheri HC, Loppin B, Lehner CF. 2012. Transgenerational propagation and quantitative maintenance of paternal centromeres depends on Cid/Cenp-A presence in Drosophila sperm. *PLoS Biol* **10**: e1001434.
Sanchez-Pulido L, Pidoux AL, Ponting CP, Allshire RC. 2009. Common ancestry of the CENP-A chaperones Scm3 and HJURP. *Cell* **137**: 1173-1174.
Shuaib M, Ouararhni K, Dimitrov S, 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 LM, Monty KJ. 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 MC, Bodor DL, Stellfox ME, Martins NM, Hochegger H, Foltz DR, Jansen LE. 2012. Cdk activity couples epigenetic centromere inheritance to cell cycle progression. *Dev Cell* **22**: 52-63.
Stoler S, Keith KC, Curnick KE, Fitzgerald-Hayes M. 1995. A mutation in CSE4, an essential gene encoding a novel chromatin-associated protein in yeast, causes chromosome nondisjunction and cell cycle arrest at mitosis. *Genes Dev* **9**: 573-586.

Stoler S, Rogers K, Weitze S, Morey L, Fitzgerald-Hayes M, Baker RE. 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.
Subramanian L, Toda NR, Rappsilber J, Allshire RC. 2014. Eic1 links Mis18 with the CCAN/Mis6/Ctf19 complex to promote CENP-A assembly. *Open biology* **4**: 140043.
Truong LN, Wu X. 2011. Prevention of DNA re-replication in eukaryotic cells. *Journal of molecular cell biology* **3**: 13-22.
Wan X, O'Quinn RP, Pierce HL, Joglekar AP, Gall WE, DeLuca JG, Carroll CW, Liu ST, Yen TJ, McEwen BF et al. 2009. Protein architecture of the human kinetochore microtubule attachment site. *Cell* **137**: 672-684.
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). *J Biol Chem* **289**: 8326-8336.
Williams JS, Hayashi T, Yanagida M, Russell P. 2009. Fission yeast Scm3 mediates stable assembly of Cnp1/CENP-A into centromeric chromatin. *Mol Cell* **33**: 287-298.
Wohlschlegel JA, Dwyer BT, Dhar SK, Cvetic C, Walter JC, Dutta A. 2000. Inhibition of eukaryotic DNA replication by geminin binding to Cdt1. *Science* **290**: 2309-2312.
Zasadzinska E, Barnhart-Dailey MC, Kuich PH, Foltz DR. 2013. Dimerization of the CENP-A assembly factor HJURP is required for centromeric nucleosome deposition. *EMBO J* **32**: 2113-2124.

Figure Legends

Figure 1. Mis18α and Mis18β form a conserved tetramer. Glycerol gradient sedimentation (left) and S.E.C. (right) performed using recombinant (A) His-tagged Mis18α, (B) Strep-tagged Mis18β, (C) Mis18α and Mis18β combined and (D) S.pombe Mis18. Arrows at the top indicate the sedimentation-values (S) and Stokes radii (nm) of standards. Proteins were detected by immunoblot and the antibodies used are shown to the right. Dots signify the peak fractions and the proteins represent: black, Mis18α; grey, Mis18β; open circles, *S.p.*Mis18. E. Chromatin free extracts (CFE) from mitotically arrested cells stably expressing GFP-tagged Mis18α were separated on glycerol gradient (left) and S.E.C S200 column (right) and immunoblotted using anti-GFP and anti-Mis18β. F. Table summarizing the hydrodynamic analyses in A-E. Caclulated molecular weights were calculated from Stokes radii and sedimentation coefficients (Siegel and Monty 1966). Stokes radii (Rs) and sedimentation values were calculated from linear standard curves (Figure S1A-D). All values are the average of two replicates. Stoichiometry of the individual complexes was calculated by dividing the calculated molecular weight of the complex by the theoretical molecular weights of the monomeric proteins, determined based on amino acid content.

Figure 2. Mis18 proteins multimerize through the conserved coiled-coil domains. A. Diagram showing the conservation of the predicted C-terminal coiled-coils. The positions of the characteristic hydrophobic (H) and charged (C) residues are shown. Shading indicates the degree of amino acid similarity. B. Schematic of the Mis18 fragments used (CC, coiled-coil). C. *In vitro* pulldowns using differentially tagged recombinant Mis18α fragments showing that Mis18α homomultimerize through its predicted C-terminal

coiled-coil. Filled ovals or squares indicate proteins present. Proteins or affinity tags were detected by immunoblot. D. *In vitro* pulldowns of differentially tagged recombinant Mis18β fragments showing that Mis18β can homomultimerize through its predicted C-terminal coiled-coil. E,F. *In vitro* pulldown between the Mis18α and Mis18β coiled-coil domains alone showing this domain can mediate Mis18α and Mis18β heterotypic interactions. Proteins or affinity tags were detected by immunoblot. G. Schematic of the LacO array system in U2OS cells. GFP-tagged prey proteins were assayed for recruitment to the non-centromeric LacO array by the mCherry LacI-fused (mCLI) bait protein. H,I. U2OS-LacO cells co-transfected with the indicated mCLI-Mis18α/β constructs and GFP-Mis18α/$\beta^{CC/FL}$. Cells were stained with an anti-CENP-C antibody to mark centromeres. Percentages of cells with recruitment of the prey proteins to the array are shown. Scale bars, 5μm

Figure 3. Mis18 coiled-coils are required for proper centromere localization. A. MARCOIL coiled coil prediction for Mis18α with indicated mutations in the conserved hydrophobic amino acids. B-C. *In vitro* pulldowns using differentially tagged recombinant Mis18α WT and coiled-coil mutants. Filled ovals and squares indicate proteins present in the input and pulldowns. Proteins or affinity tags were detected by immunoblot. D. U2OS-LacO cells were co-transfected with mCLI-Mis18α/β and GFP-Mis18$\alpha^{WT/I201G/I201G}$ or Mis18α^{L205}. Percentages of cells with recruitment of the prey proteins to the array are shown. Scale bars, 5μm. E. Anti-GFP immunoprecipitation conducted from U2OS cells transfected with GFP-Mis18α (WT or Mis18αI205G mutant), HA-Mis18α (WT), and FLAG-Mis18β (WT). F. U2OS-LacO cells were co-transfected with mCLI-Mis18BP1, GFP-Mis18$\alpha^{WT/I201G/I201G}$ or Mis18α^{L205}, and FLAG-

Mis18β^{WT}. Percentages of cells with recruitment of the prey proteins to the array are shown. Scale bars, 5μm. G. U2OS cells were co-transfected with GFP-Mis18α WT or mutant constructs. RFP-histone H2B was used as a transfection marker. Centromeres were identified using anti-CENP-T antibody. The percentage of cells in which GFP-Mis18 is recruited to centromeres is shown ± S.D. Scale bars, 5μm.

Figure 4. The Mis18α/β complex promotes CENP-A deposition through a direct physical interaction with HJURP. A. U2OS-LacO cells were co-transfected with the indicated mCLI, GFP-TET, HA-CENP-A, and HA-HJURP constructs with/without 10mM IPTG treatment to disrupt the LacI interaction with the LacO array. Cells were stained with anti-CENP-A antibody. The LacO array was marked by GFP-TET. Recruitment of CENP-A to the LacO array can be seen in the magnified boxed regions to the right. Percentages of cells with CENP-A recruitment to the array are shown. Scale bars, 5μm. B. I*n vitro* pulldown of recombinant HJURP^{R1} and differentially tagged Mis18α and Mis18β. Filled ovals and squares indicate proteins present in the input and pulldowns. Proteins or affinity tags were detected by immunoblot. C. U2OS-LacO cells were co-transfected with mCLI-Mis18α/β^{CC}, GFP-Mis18α/β^{CC}, and HA-HJURP. Cells were stained with an anti-HA antibody. Percentages of cells with recruitment of the prey proteins to the array are shown. Scale bars, 5μm. D. Schematic of the domain structure of HJURP and the recombinant fragments used in this experiment. CenTD, centromere-targeting domain. E. *In vitro* pulldown between recombinant HJURP fragments, Mis18α and Mis18β. Proteins or affinity tags were detected by immunoblot.

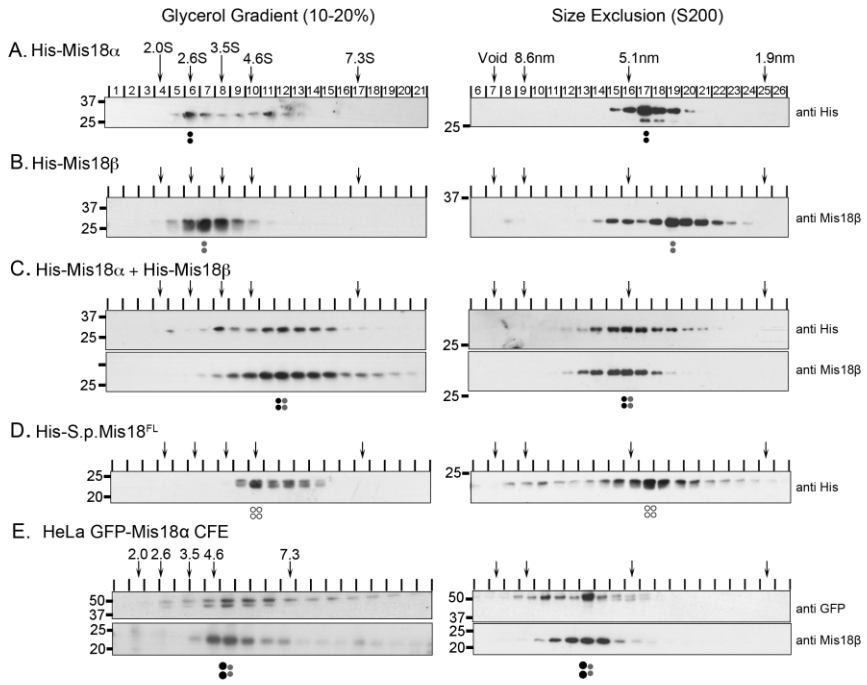
Figure 5. HJURP disrupts the Mis18α/β heterotetramer. A-C. S.E.C. performed on recombinant HJURP fragment (a.a. 348-555) MBP-HJURP^{R1}, Mis18α/β and MBP-

HJURP^{R1} + Mis18α/β. Stokes radii of standards are indicated by arrows. Proteins were detected by immunoblot. Shapes under the blots identify the peak fractions represent and the proteins present: black, Mis18α; grey, Mis18β; open oval, HJURP^{R1}. D. *In-vitro* pulldowns of the peak S.E.C. fraction on either amylose or strep-tactin beads showing that the peak complex contains Mis18α, Mis18β and HJURP^{R1}. E. Predicted model of Mis18α/β recruitment and heterotetramer disruption in the presence of HJURP^{R1} at the LacO array. F. U2OS-LacO cells were co-transfected with mCLI-Mis18α, GFP-Mis18α, FLAG-Mis18β, and HA-HJURP^{R1}. Cells were stained with an anti-FLAG antibody to determine Mis18β recruitment. The LacO arrays are magnified in the boxed regions to right with either GFP-Mis18α (left) or FLAG-Mis18β (right). Scale bars, 5μm. G. Fluorescence quantitation of the mCherry (mCLI-Mis18α), GFP (GFP-Mis18α), and FLAG (FLAG-Mis18β) channels with no HJURP or HJURP^{R1} transfected into the cells. A student's two-tailed t-test was performed comparing the two conditions in each channel. P-values are noted above and asterisks.

Figure 6. The stable pool Mis18α is removed from centromeres following HJURP binding. A. Representative images of fluorescence recovery after photobleaching (FRAP) of GFP-Mis18α. B. Averaged fluorescence recovery values were fitted to an exponential recovery curve, ± S.E.M. C. Live-cell imaging of GFP Mis18α HeLa cells treated with either control or HJURP siRNA. Cells were imaged starting in metaphase when GFP-Mis18α is absent. D. Quantitation of the persistence of GFP-Mis18α in centromeric foci in control and HJURP siRNA treated cells. Cells were examined for Mis18α foci at 20-minute intervals following metaphase. Bars indicate the time points during which Mis18 was visible. E. Western blot of cell populations from control siRNA

and HJURP siRNA treated cells to assess the efficiency of HJURP suppression. F. Parental U2OS cells were transfected with the indicated GFP constructs along with RFP-H2B as a transfection marker. Cells were stained with anti-CENP-T and CENP-A antibodies. Scale bars, 5μm. G. Bar graph of total CENP-A fluorescence between the different conditions, ±S.D. P-values obtained from a two-tailed t-test with a significance level >0.05.

Figure 7. CENP-A levels at centromeres are controlled by Mis18. A model of centromeric licensing by the Mis18α-β heterotetramer. Mis18α-β is recruited to centromeres during late telophase at which time it becomes stably bound to the centromere. The pre-nucleosomal complex of CENP-A and HJURP recognizes the centromere through a direct interaction between the Mis18α-β coiled-coil domains and the CenTD domain of HJURP. Binding of HJURP disrupts the Mis18α-β heterotetramer and reduces Mis18 binding to Mis18BP1 and the centromere. Degradation of Mis18α and Mis18β may further limit re-recruitment of the complex to centromeres.



F.

	Monomer Theoretical MW (kD)	Rs (nm)	Sedimentation-Value	Calculated MW (kD)	Ratio (Calculated MW/Monomer MW)
H.s. Mis18α	28.1	5.1	2.7	57.9	2.1 (1α:1α)
H.s. Mis18β	29.5	4.3	3.2	56.9	2.0 (1β:1β)
H.s. Mis18α/β	28.1 (α), 29.5(β)	5.5	5.3	122.6	4.3 (2α:2β)
S.p. Mis18 FL	24.7	5.1	4.4	94.7	3.8
S.p. Mis18 ΔCC	18.5	2.2	N/A	N/A	N/A
HeLa GFP-Mis18α CFE	51.3 (α), 24.7 (β)	7.2	5.1	154.4	4.1 (2α:2β)

Figure 1 Nardi et al.

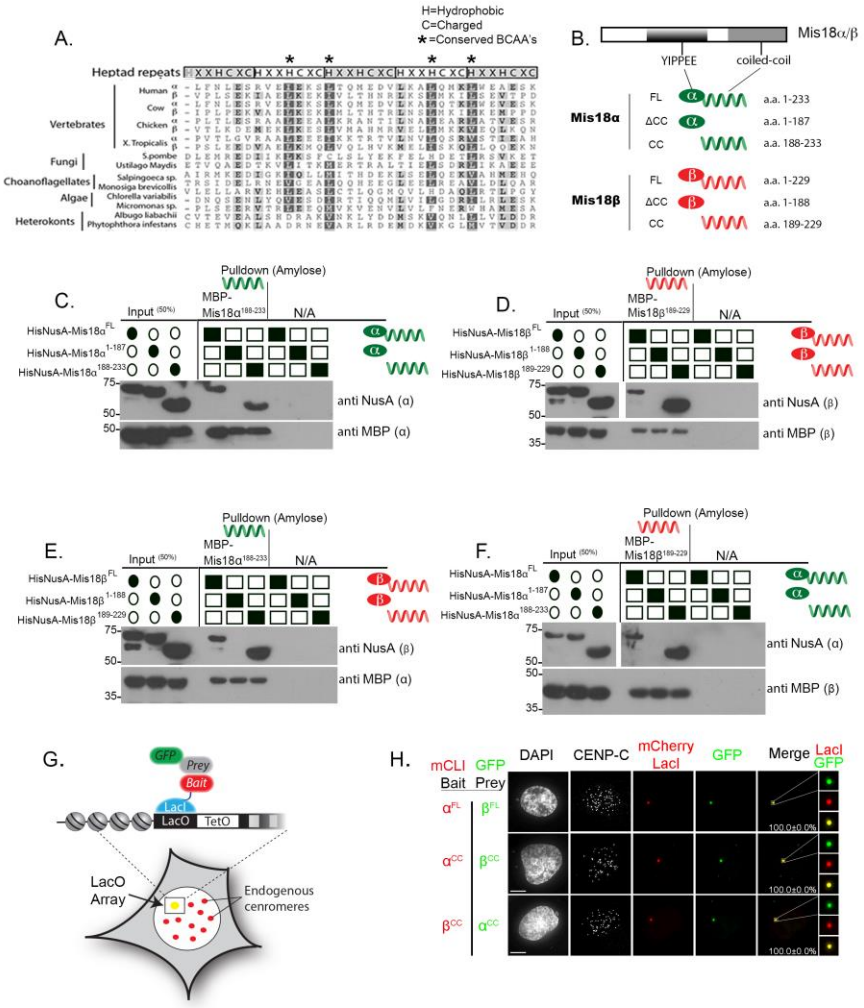


Figure 2. Nardi et al.

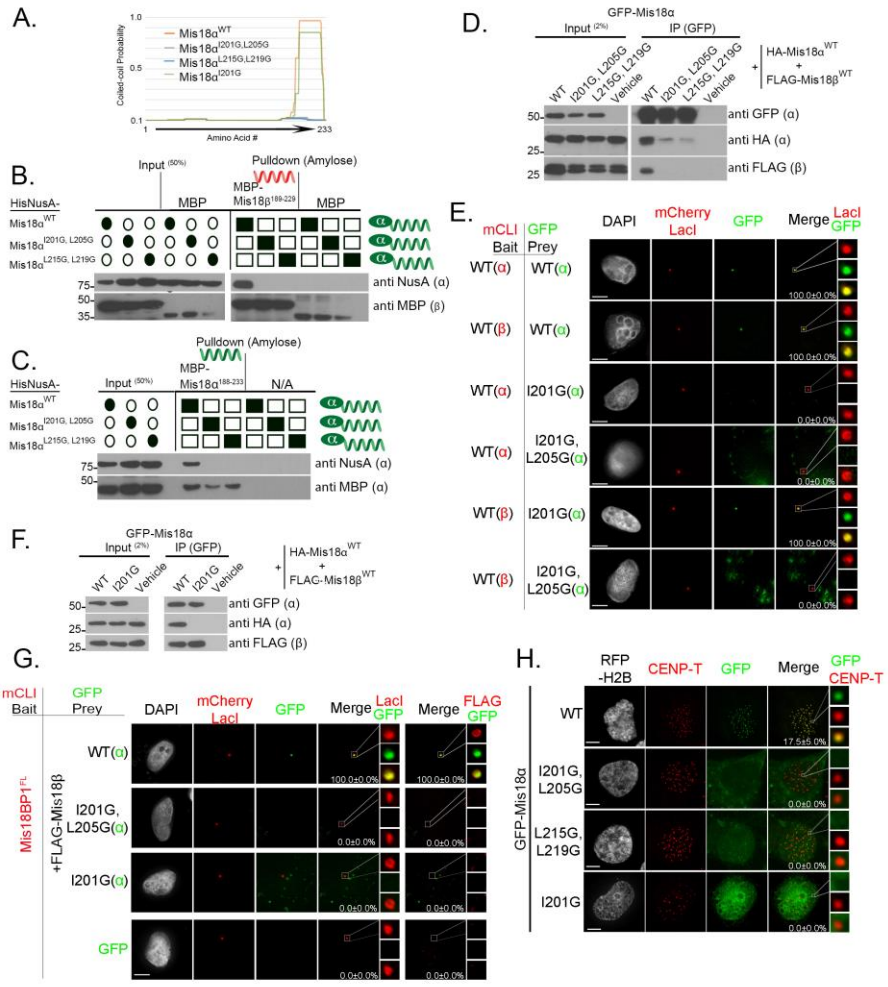


Figure 3. Nardi et al.

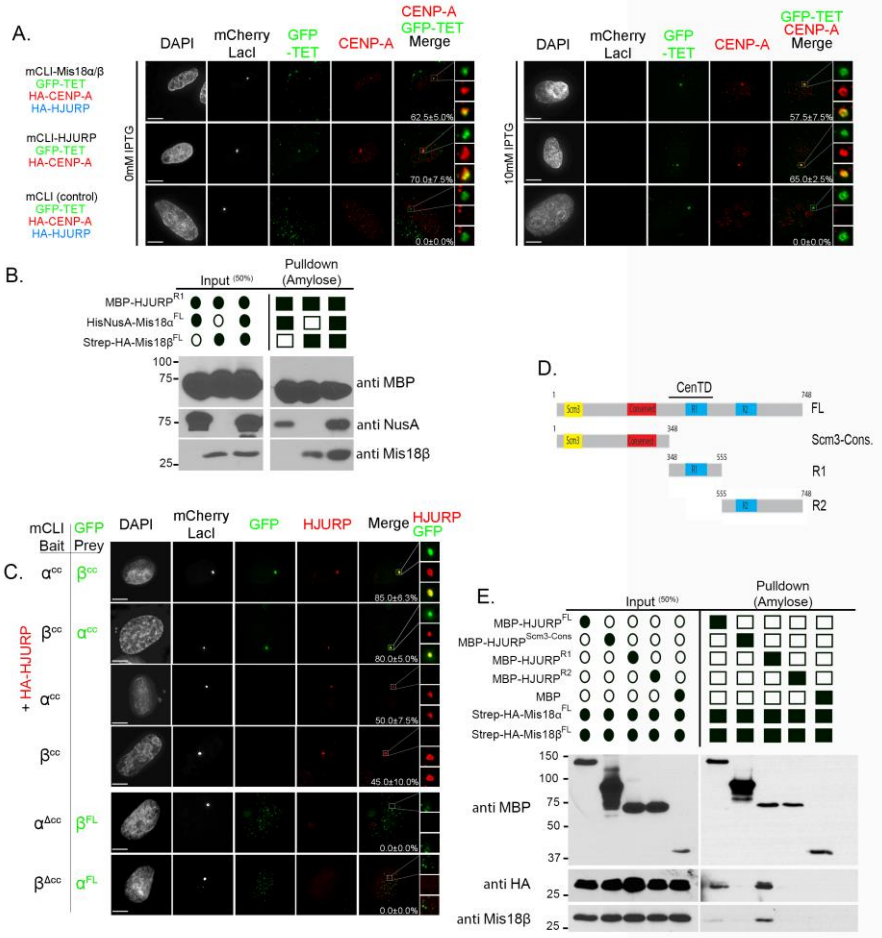


Figure 4. Nardi et al.

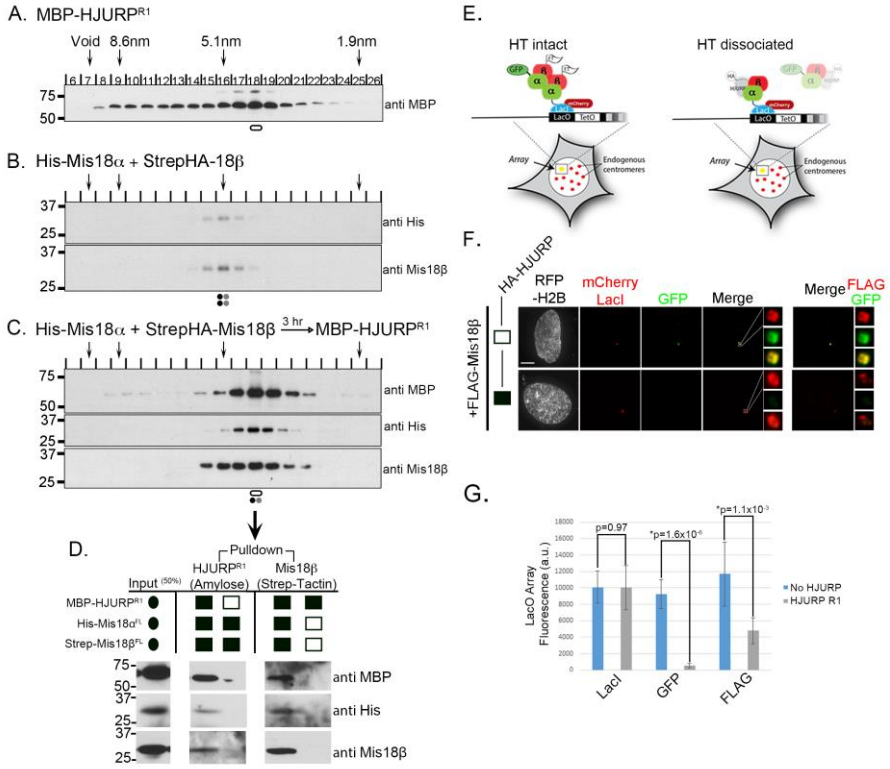


Figure 5. Nardi et al.

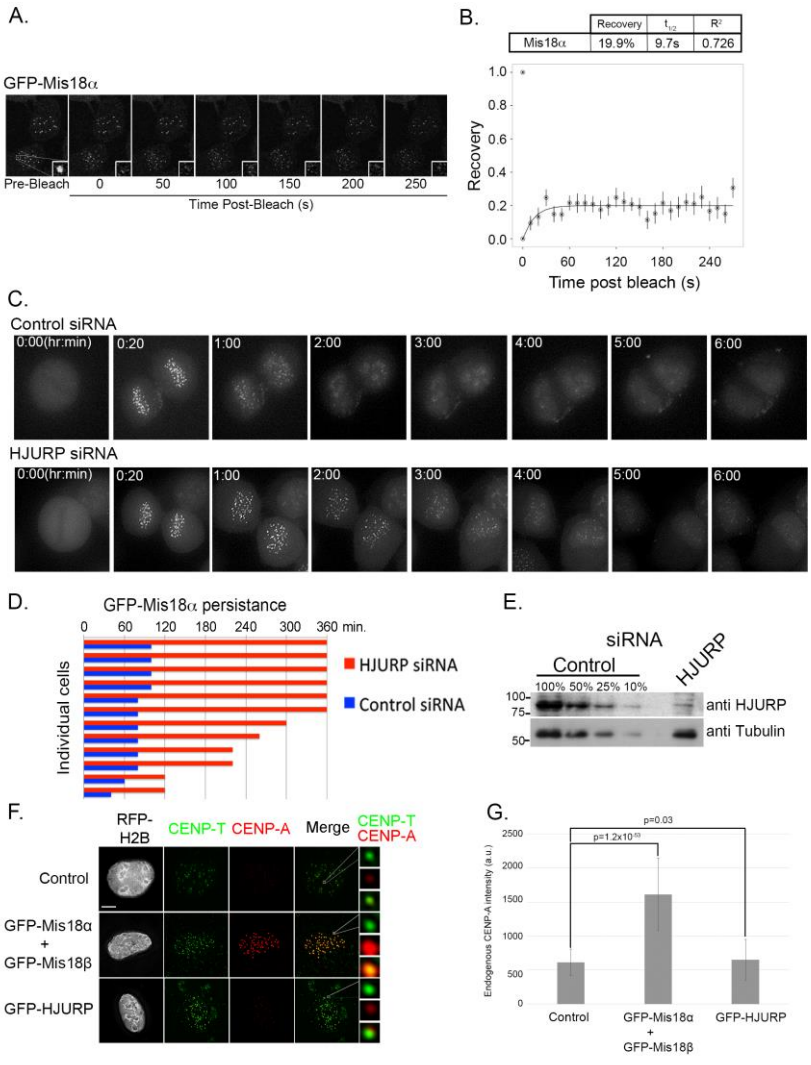


Figure 6, Nardi et. al

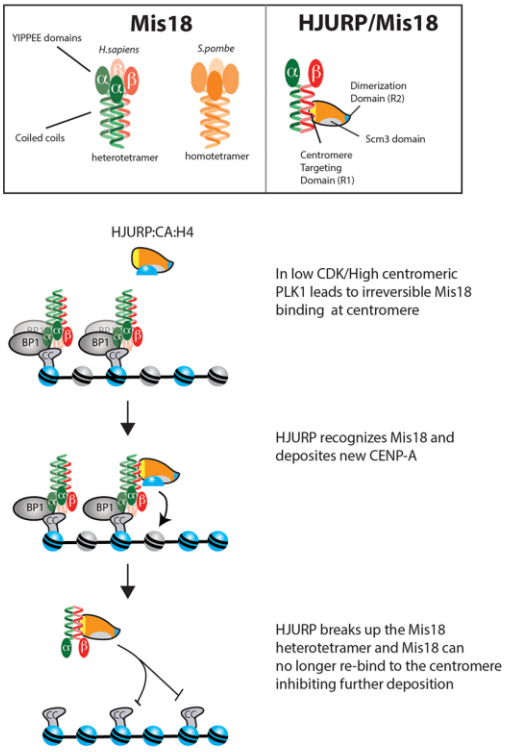
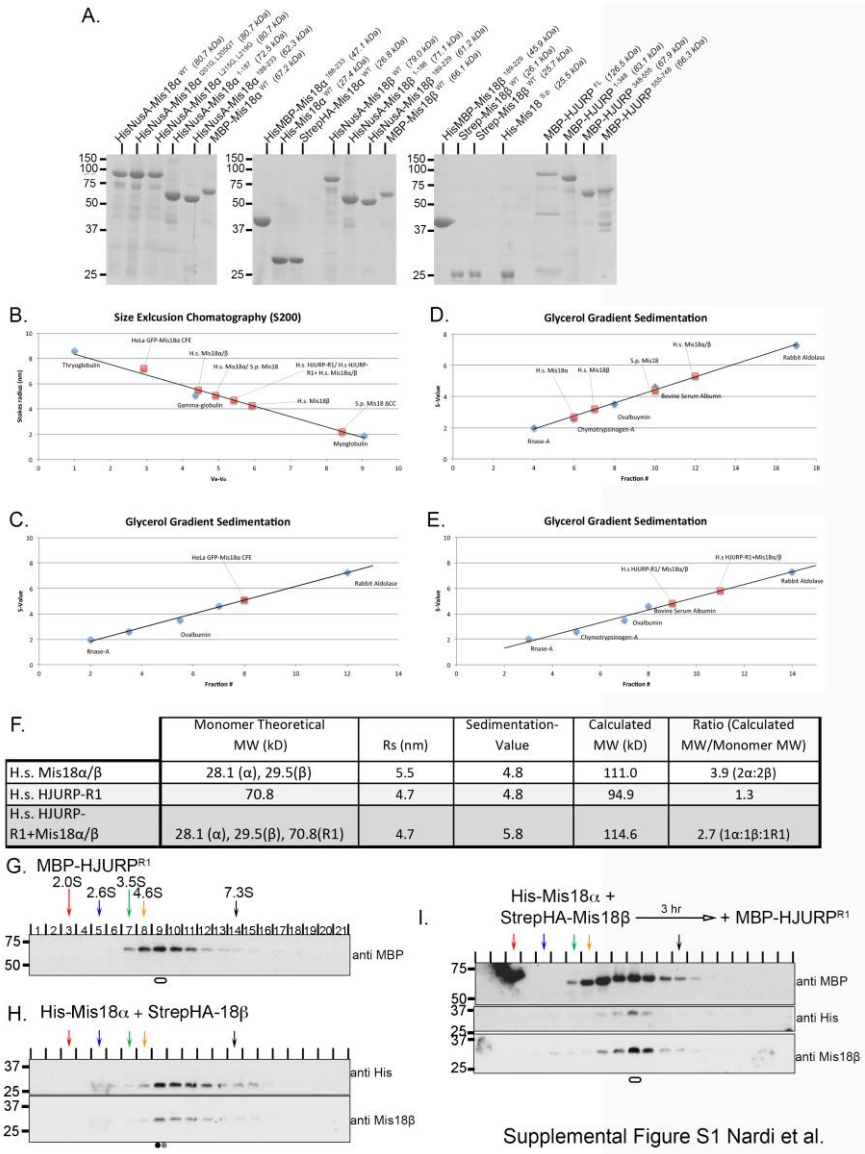


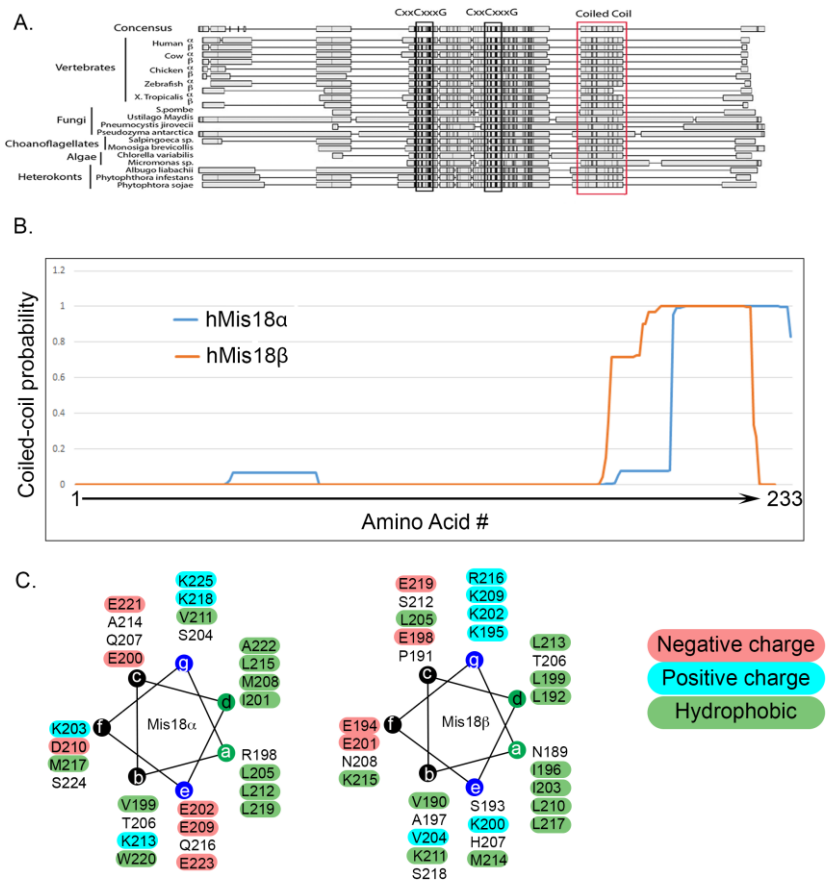
Figure 7, Nardi et. al

Supplemental Figures

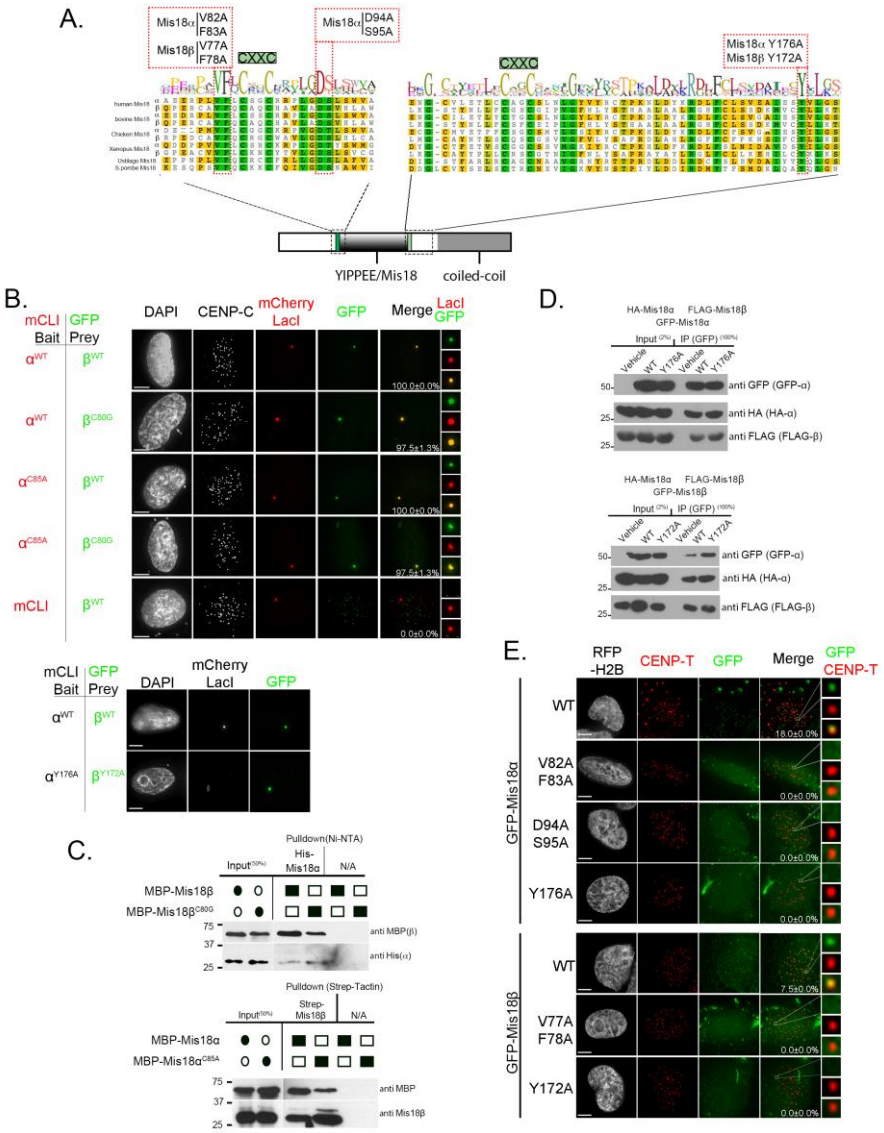


	Monomer Theoretical MW (kD)	Rs (nm)	Sedimentation-Value	Calculated MW (kD)	Ratio (Calculated MW/Monomer MW)
H.s. Mis18α/β	28.1 (α), 29.5(β)	5.5	4.8	111.0	3.9 (2α:2β)
H.s. HJURP-R1	70.8	4.7	4.8	94.9	1.3
H.s. HJURP-R1+Mis18α/β	28.1 (α), 29.5(β), 70.8(R1)	4.7	5.8	114.6	2.7 (1α:1β:1R1)

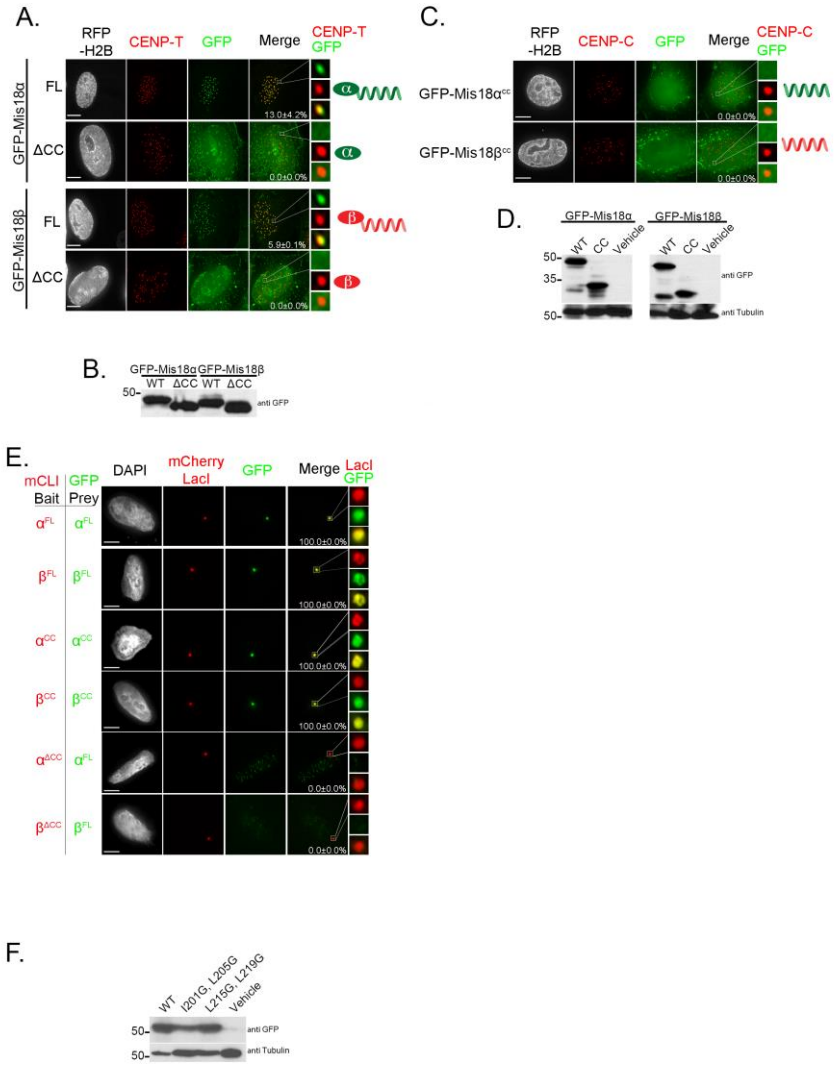
Supplemental Figure S1 Nardi et al.



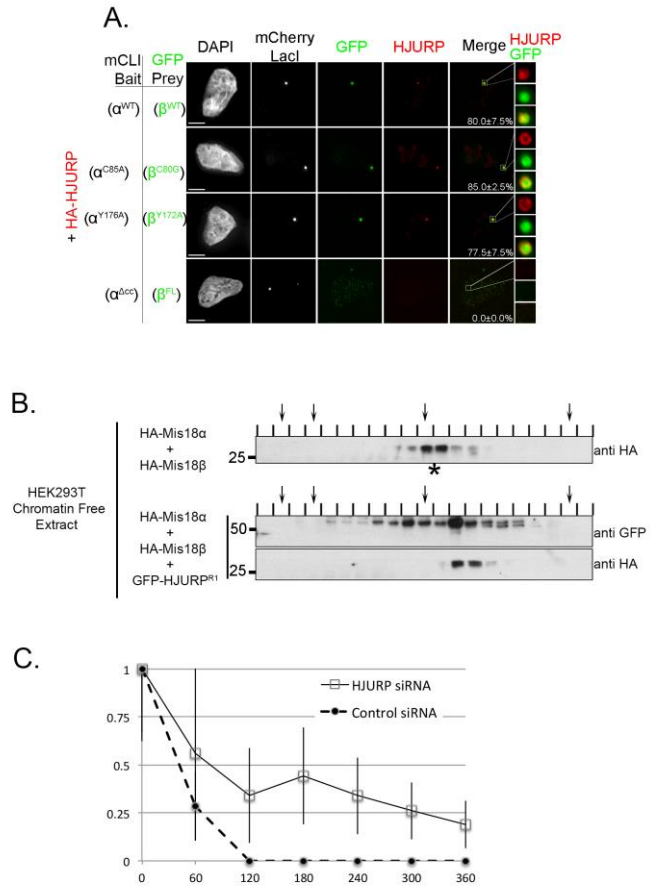
Supplemental Figure S2 Nardi et al.



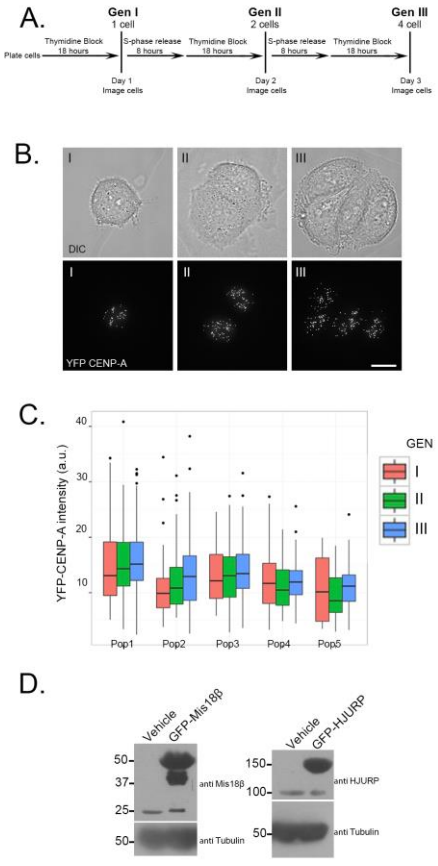
Supplemental Figure S3 Nardi et al.



Supplemental Figure S4 Nardi et al.



Supplemental Figure S5 Nardi et al.



Supplemental Figure S6 Nardi et al.

Supplemental Figure Legends

Figure S1. A. Coomassie gel of recombinant proteins used for *in-vitro* analysis. Predicted Molecular weights are indicated for each protein in parentheses next to each name. Molecular weight markers used on the gel are indicated on the left sides (kDa). B. Linear curve with line of best fit for size exclusion standards thyroglobulin (Rs=8.6), Gamma globulin (Rs=5.1), and myoglobulin (Rs=1.9) (Biorad, blue dots, unboxed labels). The X-axis is calculated by subtracting the elution volume (Ve) by the void volume of the column (Vo). The y-axis is the Stokes radii of each standard and calculated protein. The proteins of interest were (boxed labels) overlaid onto the line of best fit independently and ther Rs values were calculated from the three original standards. C-E Linear curves with a line of best fit for glycerol gradient sedimentation. Standards used were Rnase-A (2.0S), Chymotrypsinogen-A (2.6S), Olvalbumin from hen egg (3.5S), Bovine Serum Albumin (4.6S), and Rabbit Aldolase (7.3S) (Blue dots, unboxed labels). From these standards a line of best fit was made and used to calculate the S-values of the proteins of interest (Red dots, boxed labels). The X-axis the fraction # of the peak elution for each standard and protein of interest while the Y-axis corresponds the S-value. The proteins of interest were overlaid on the line of best fit independently. Graphs S1C was used to calculate the S-values of the recombinant Mis18 proteins (Figure 1A-D,F). Graph S1D was used to calculated the S-value of the HeLa GFP-Mis18α CFE (Figure 1E,F). Graph S1E was used to calculate the S-values of the HJURP-R1 and the Mis18 complex alone and in complex together (Figure 5A-C,E). F. Table summarizing hydrodynamic analysis of $HJURP^{R1}$ and the Mis18α/β separately and mixed together from Figure 5 A-C. Stokes radii and sedimentation values were calculated from linear curves (Figure S1B,E)

were used to calculate the native molecular weights (Siegel and Monty 1966). Values are the average of two replicates. Stoichiometry of the individual complexes was calculated by dividing the calculated predicted molecular weights of the complex by the calculated molecular weights of the monomeric proteins. G-I. Glycerol gradient sedimentation performed on recombinant HJURP fragment (aa 348-555) MBP-HJURP^{R1}, Mis18α-β and MBP-HJURP^{R1}. Sedimentation-values of standards are indicated by arrows. Proteins were detected by Western blot. Shapes under the blots identify the peak fractions represent and the proteins present. This analysis was done in conjunction with S.E.C analysis to determine stokes radii of the complexes as in Figure 5A-C.

Figure S2. Mis18 coiled-coil conserved across species. A. ClustalW alignment was conducted of Mis18 sequences from a broad range of phylogeny. The coiled-coil domains are boxed in red. Shading indicates the degree of amino acid similarity, with black indicating 100% similarity. B. Mis18 proteins were run through the MARCOIL prediction program showing they both contain C-terminal predicted coiled-coil domains. C. Alpha-helical representation of the amino acids present within the Mis18α and Mis18β coiled coils. a-g indicate the position with the heptad repeat. D. Possible configuration of a two-component four helical coiled-coil. Lines indicate the axes of separation that would yield a heterodimer.

Figure S3. Conserved residues in the YIPPEE domain of Mis18 necessary for centromere recruitment A. Sequence alignment of Mis18 across species. Green indicates residues that are highly conserved and yellow indicates residues that are less conserved across species. B. U2OS LacO cells were transfected with Mis18 wild-type or YIPPEE mutants as bait (mCLI) and prey (GFP). Boxed regions are magnified to the right.

Percentages of cells with GFP-Mis18 co-localized with the array are indicated; ± S.D. Scale bars, 5μm. C. *In vitro* pull-downs using recombinant Mis18WT and cysteine mutants. D. Immunoprecipitation conducted from U2OS cells transfected with wild-type or distal YIPPEE domain mutants of Mis18α or Mis18β. E. Localization of transiently transfected GFP-tagged wild-type Mis18 or the indicated mutants in parental U2OS cells. Cells were co-transfected with RFP-H2B to identify transfected cells and stained with anti-CENP-T to mark centromeres. Boxed regions are magnified to the right. Percentages of cells with Mis18 localized to centromeres are indicated ± S.D. Scale bars, 5μm.

Figure S4. Mis18 heterotetramer must form for proper centromere recruitment A. Parental U2OS cells were transfected with GFP-tagged wild-type Mis18 or ΔCC mutants. Cells were co-transfected with RFP-H2B to identify transfected cells and stained with anti-CENP-T to mark centromeres. Boxed regions are magnified to the right. Percentages of cells with Mis18 localized to centromeres are indicated ± S.D. Scale bars, 5μm. B. Western blot of U2OS cells transiently transfected with the Mis18 Full-length or YIPPEE domain to show expression. Anti-tubulin is a loading control. C. Parental U2OS cells transfected with the indicated Mis18 coiled-coil domain. Cells were co-transfected with RFP-H2B to identify transfected cells and stained with anti-CENP-T to mark centromeres. Boxed regions are magnified to the right. Percentages of cells with Mis18 localized to centromeres are indicated ± S.D. D. Western blot of U2OS cells transiently transfected with the Mis18 Full-length or coiled-coil domain to show expression. Anti-tubulin is a loading control. E. U2OS LacO cells were transfected with the indicated Mis18 wild-type and truncation mutants (mCLI) and prey (GFP). Boxed regions are magnified to the right. Percentages of cells with GFP co-localized with the array are

indicated; ± S.D. Scale bars, 5μm. F. Western blot of Mis18 WT and mutants transfected U2OS cells. Anti-tubulin was used as a loading control.

Figure S5. A. U2OS LacO cells were transfected with Mis18 wild-type or YIPPEE mutants as bait (mCLI) and prey (GFP). Boxed regions are magnified to the right. Percentages of cells with GFP-Mis18 co-localized with the array are indicated; ± S.D. Scale bars, 5μm. B. S.E.C. of HEK293T cells transfected with HA-Mis18α/β +/- GFP-HJURP^{R1}. Chromatin free extracts were made and run over an S200 column. Proteins were detected by western blot. Asterisk's under blots represent peak fractions and the Stokes radii of the standards and void are noted above the fractions. C. Averaged centromeric intensities for the stably expressing GFP-Mis18α HeLa cells with either HJURP or control siRNA from Figure 6C.

Figure S6. A. Flow chart of the protocol used to plate, block, release, and image cells at successive G1/S boundaries B. Representative images of cells at each generation (generation I, II, and III). Cells were imaged to assess the changes in YFP-CENP-A intensity across cell generations C. Vertical box plots of integrated YFP-CENP-A fluorescence across generations for individual cells and their progeny. Bars indicate standard deviation while horizontal lines in the boxes indicate the mean. D. Immunoblot of overexpressed GFP-tagged Mis18β and HJURP from figure 6F.

Supplemental Materials & Methods

DNA constructs. DNA constructs created using gateway cloning and the vectors used are listed in supplemental table S1. Mis18α was created by PCR amplification with a forward primer that added a 6X-His tag flanked with an Nde1 cut site and a reverse primer that was flanked by a Kpn1 cut site. Mis18β was amplified with a forward primer harboring an Nde1 cut site with no affinity tag built on and a reverse primer harboring a BamH1 cut site. Both PCR products were cut with the respective restriction enzymes and ligated into pST50trc1 (Mis18α) and pst39 (Mis18β) expression plasmids. Two complimentary oligos harboring flanking Nde1 compatible overhangs that encoded for a Strep-TEV tag were ligated together and ligated into the pST39 harboring Mis18β that had been cut with Nde1.

Cell culture. Parental U2OS and LacO-TRE were transfected with 1 µg of GFP containing Mis18 plasmid and 0.02 µg RFP-H2B (transfection marker) using Lipofectamine-2000 (Invitrogen) in serum-free Optimem Media (Gibco). U2OS-LacO-TRE cells were transfected with 0.2µg mCherry-LacI construct, 0.4 µg GFP construct, and 0.4 µg HA-HJURP construct (1:2:2). For immunocytochemistry U2OS-LacO-TRE cells were pre-extracted with 0.1% Triton X-100 in PBS for 3 min, fixed with 4% formaldehyde in PBS for 10 min, and quenched with 100 mM Tris, pH 7.5, for 10 min at room temperature. Coverslips were blocked in 2% FBS, 2% BSA and 0.1% Triton X-100 in PBS. Parental U2OS cells were not pre-extracted. Coverslips were mounted with ProLong (Invitrogen).

Antibodies. Cells were stained with rabbit anti–CENP-T (1:3000, D. Cleveland, UCSD) mouse anti-CENP-C (1:3000) and mouse anti-HA monoclonal antibody (1:1000,

Covance), and a mouse monoclonal CENP-A antibody (1:1000, Sigma). Secondary antibodies used were donkey anti–rabbit and goat anti-mouse Cy5-conjugated (1:6000, Jackson ImmunoResearch). DNA was stained with 0.2 ug/ml DAPI.

Protein expression and purification. Recombinant proteins were expressed in the Rosetta (DE3) pLysS bacterial strain. Bacteria were grown in LB media to an OD of 0.6 at 37ºC and induced at 18ºC with 0.1mM IPTG for 16 hr. His-tagged HJURP fragments were purified on Ni-NTA (Qiagen). Bacteria were lysed using a steel Wheaton-dounce in buffer contained 50mM Tris-HCl pH 7.5, 250mM NaCl, 20mM MgCl2, 0.5mM CaCl2, 10% glycerol, 0.1% NP-40, 5mM BME, LPC, 1mM PMSF, 20mM imidazole, and 0.15mg/ml RNase-A. DnaseI was added and lysates were centrifuged at 22,000x*g* for 20 min. Supernatants were collected and pellets were re-extracted with a second round of lysis. Cleared lysates were incubated with Ni-NTA agarose (Qiagen) for 1 hour and washed twice in lysis buffer with 40mM imidazole. Proteins were eluted in lysis buffer containing 250mM imidazole and no protease inhibitors. Full-length HJURP was further purified on a Superdex 200 10/300 column (G.E. Healthcare) and re-concentrated on Ni-NTA agarose (Qiagen). Mis18 proteins were purified as described in lysis buffer containing: 50mM Tris-HCl, 350mM NaCl, 0.5mM CaCl2, 10% glycerol, 0.1% NP-40, 5mM BME, LPC, and 1mM PMSF, and imidazole. MBP tagged proteins were purified using maltose agarose (Qiagen) and eluted in buffer containing 10mM maltose. Strep–tagged proteins were purified using Strep-Tactin Super-flow plus (Qiagen) and eluted in buffer containing 2.5mM d-desthiobiotin.

Live cell imaging. YFP-CENP-A expressing HeLa cells were plated onto gridded coverslips (MatTek) $1.0x10^4$ cells per well and treated 24 hrs later with 01.ug/ml

thymidine for 18 hrs. Thymidine was washed out and new media containing deoxycytidine at 24uM was added. After this, thymidine was again added so that the single cells would progress through mitosis and then be blocked at the G1/S boundary again for imaging. This process was repeated for a total of three cell generations, and the same cells were imaged at each generation during each G1/S block. Fluorescence intensities were measured from raw-imaged files and images were then deconvolved and presented as stacked images.

Immunoprecipitations. Cells were lysed 24 h post transfection in RIPA buffer (150 mM NaCl, 1% NP-40, 0.3% deoxycholate, 0.15% SDS, 50 mM Tris HCl pH 7.5, 1 mM EDTA, 10% glycerol, Protease Inhibitors (Roche), 200 μM NaV, 0.5 mM PMSF, 5 mM NaF, 50 mM β-mercaptoethanol, 5 μM microcystin) on ice for 15 min with occasional vortexing. Extracts were DNAseI (1:200, NEB Biolabs) treated and sonicated where indicated. Lysates were centrifuged at 18 000 *g* for 10 min at 4°C and pre-cleared with Protein A agarose (Biorad) for 2 h 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 min, 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.

CHAPTER 3: Mis18 as a component of the CENP-A deposition machinery

Abstract:

Centromeres are specialized chromatin domains specified by the centromere specific CENP-A nucleosome. The stable inheritance of centromeres is an epigenetic process that requires the deposition of new CENP-A nucleosomes by HJURP and other associated factors such as the evolutionarily conserved Mis18α-β complex. We show the Mis18 complex directly and specifically binds to pre-nucleosomal CENP-A *in vitro*. This interaction occurs between the CENP-A CATD and the conserved N-terminal YIPPEE domains of Mis18α-β. Histidine 104 located within the CATD of CENP-A is necessary for the CENP-A interaction with the Mis18 complex. Mutation of this residue leads to loss of CENP-A at centromeres, which we show is due to the inability of CENP-A to interact with the Mis18 complex at chromatin. We propose the direct interaction of the Mis18 complex with CENP-A at chromatin is an integral step for proper CENP-A deposition into nucleosomes at centromeres every cell cycle.

Introduction:

Centromere specification involves the deposition of the centromere specific histone CENP-A (Centromere Protein-A). Similar to canonical histones that are deposited on chromatin by a series of chaperones and assembly factors, CENP-A is loaded at centromeres by its own set of unique assembly factors. A key player is HJURP (Holliday Junction Recognition Protein), which was demonstrated to act as the CENP-A specific chaperone through direct binding to CENP-A and is required for its chromatin assembly at centromeres[1,2]. Although HJURP is integral for CENP-A assembly into chromatin other proteins have been shown to be required for CENP-A recruitment and deposition, such as the Mis18 complex.

The Mis18 complex is composed of three subunits in higher eukaryotes, Mis18α, Mis18β, and Mis18BP1. The Mis18 proteins are required for HJURP recruitment and therefore CENP-A deposition in a diverse array of eukaryotes harboring regional centromeres[3–9]. Recent work by our

group and others has shown the Mis18 complex recruits HJURP through a direct interaction[10,11]. During CENP-A deposition, this puts the Mis18 complex and CENP-A in close proximity to one another but an interaction, direct or indirect, between the Mis18 complex and CENP-A has not been shown. Although HJURP is associated with several pre-chromatin factors and chromatin bound factors (such as Mis18), the active enzymatic steps involved in CENP-A offloading and deposition into chromatin have not been elucidated. This is in stark contrast to canonical histone deposition pathways such as that of Histone H3.1 (H3).

H3 (of which CENP-A is a variant), is assembled into chromatin by multiple subunits collectively known as the CAF complex that include CAF-1, Asf1, NAP-1, and RCAF. Together, these subunits form a complex that binds and assembles H3 into nucleosomes during DNA replication and even during DNA repair[12,13]. The process of shuttling H3/Histone H4 (H4) into chromatin involves transport by Asf1 from the cytosol to the nucleus followed by acetylation of H3 lysine 56[14,15] which denotes newly synthesized H3. Asf1 is proposed to hand-off H3/H4 to CAF-1[16]. The CAF/H3/H4 complex interacts with PCNA, a protein involved in the DNA polymerase machinery. This interaction is believed to facilitate offloading of H3/H4 into nucleosomes from the CAF complex at sites of DNA replication[17,18]. In regards to CENP-A deposition, how HJURP offloads CENP-A into chromatin, and if other factors are involved in the process, has not been well studied. Given the Mis18 complex directly binds HJURP this puts the complex in close proximity to HJURP bound CENP-A making it a likely factor involved in CENP-A assembly into chromatin.

Unlike H3 deposition, which occurs during S-phase when DNA is being unwound to be replicated, CENP-A deposition takes places in G1-phase in a replication independent manner[19,20]. In this stage the centromeric DNA is still tightly wound around the pre-existing nucleosomes. In order to shuttle new CENP-A into the chromatin the DNA needs to be unwound then rewound properly around the new CENP-A nucleosomes. This process is facilitated by several DNA

interacting subunits in H3 nucleosome deposition such as PCNA mentioned above, but no mechanism or factors have been put forth in the context of CENP-A deposition at centromeres. Crystallization of a fragment of fission yeast Mis18, of which Mis18α and Mis18β are paralogs, found a conserved YIPPEE domain in the N-terminus of Mis18[21]. The YIPPEE protein itself was originally discovered in a yeast interaction trap screen. After subsequent cloning, YIPPEE was found to be a conserved gene family of proteins present in a diverse range of eukaryotic organisms from fungi to humans[22]. YIPPEE domains are putative zinc-binding/DNA binding domains that facilitate protein-DNA interactions. We hypothesize the YIPPEE domains within Mis18α and Mis18β are playing an active role in deposition of new CENP-A at centromeric chromatin.

Here we demonstrate the human Mis18 complex can directly and specifically bind to CENP-A/H4 *in vitro*. Furthermore, we show the YIPPEE domains of Mis18 are necessary for CENP-A retention at chromatin and disruption of this domain by deletion or mutation leads to loss of CENP-A from chromatin at an exogenous locus in humans cells. We have found that mutation of histidine 104 within the CATD of CENP-A abolishes CENP-A centromere localization and Mis18 binding *in vivo*, but retains its ability to bind HJURP, demonstrating that Mis18 binding is required for CENP-A deposition. We propose this unique and direct interaction between Mis18 and CENP-A is necessary for CENP-A deposition into chromatin and denotes a role for Mis18 in the CENP-A deposition machinery.

Results:

The Mis18 complex directly binds to the CENP-A CATD- Previous work has shown that HJURP directly binds to CENP-A/H4 as its pre-nucleosomal chaperone[1,2]. Recent findings by our lab and others have also determined Mis18 directly interacts with HJURP[10,11]. We hypothesized Mis18 could be in close contact with and even bind directly to CENP-A at centromeric chromatin. To

test if CENP-A and Mis18 directly interact, we purified recombinant MBP-Mis18α, HisNusA-Mis18β, and untagged CENP-A/H4 heterotetramer. Individual canonical histones were purified as well and dialyzed to form H3/H4 tetramer and H2A/H2B dimer (Figure 1A). MBP-Mis18α was pre-incubated with HisNusA-Mis18β to form the Mis18 complex. After incubation, CENP-A/H4 heterotetramer or H3/H4 heterotetramer were added in with the Mis18 complex or MBP alone as a control (Figure 1B). A pulldown on amylose beads was performed and proteins were visualized by coomassie stain. Strikingly, the Mis18 complex bound directly and specifically with CENP-A/H4 and not H3/H4. As a negative control, neither CENP-A/H4 nor H3/H4 bound with the MBP tag alone. This is the first time a direct and specific interaction between the Mis18 complex and CENP-A/H4 has been observed.

The CATD in CENP-A is necessary for its recruitment to centromeres and direct binding to the Scm3 domain of HJURP. Indeed, swapping this domain into canonical H3.1 to make a chimeric H3-CATD protein is sufficient to target the chimera to centromeres[1,2,23,24]. We hypothesized Mis18 could be binding to CENP-A/H4 through direct recognition of the CATD. To test this, we performed another *in vitro* pulldown between the Mis18 complex and either CENP-A/H4, H3/H4, or H3-CATD/H4 (a generous gift from Ben Black) (Figure 1D-E). Pulldowns were performed as in figure 1B. Interestingly, the Mis18 complex bound to CENP-A/H4 and also H3-CATD/H4, but not H3/H4.

Lastly, a pulldown was performed between CENA-A/H4 along with individual MBP-Mis18α or Mis18β comparing their binding to CENP-A to that of the pre-formed Mis18α-β complex. Pulldowns were performed as in figure 1B (Figure 1C). Proteins were again visualized by coomassie. From these initial results it suggests that the pre-formed Mis18 heterotetramer is no better at binding to CENP-A/H4 than the individual Mis18 homodimers, except the pulldown was 3:1 molar excess of Mis18 dimer to CENP-A/H4 tetramer. The overabundance of Mis18 could be a reason for the similarity in binding robustness of the homodimers versus the

heterotretramer. A follow up experiment should be performed with differing ratios of Mis18 to CENP-A/H4 to ascertain whether binding is different between the Mis18 homodimers and the heterotetramer. Taken together, these results show a direct interaction between the Mis18 complex and CENP-A/H4 through the CENP-A CATD.

Specific residues within CATD are required to target CENP-A to centromeres- Prior work has recognized key amino acids within the CATD of CENP-A that are necessary for both centromere localization and binding of HJURP[25,26]. The previous work used a set of mutations which substituted in non-conserved H3 amino acids to the CATD of CENP-A and tested them for HJURP binding (Figure 2A)[19,27]. From that work we focused on the CENP-A CATD mutants that still bound HJURP at the LacO array[27] to test if they were able to interact with Mis18 at chromatin.

GFP versions of the CENP-A CATD mutants were made and named α-helix 2.1, 2.2, 2.3, and L1 (Figure 2A-C) as was done in the original work[27]. Centromere localization of these CENP-A mutants was assayed along with GFP-CENP-A and GFP-H3-CATD as positive controls (Figure 2C). Cells were stained with CENP-T as a marker for centromeres and RFP-H2B was used as a marker for transfection. Although all three CENP-$A^{\alpha 2}$ (α2.1-2.3) mutants showed significantly reduced centromere localization when compared to CENP-A^{WT}, CENP-$A^{\alpha 2.2}$ showed the most pronounced phenotype with only 10% of the cells having centromere localization. The α2.1 and α2.3 mutants gave a more intermediate phenotype showing more GFP nuclear background suggesting that they might be localizing to other chromosomal loci outside of the centromere. This suggests multiple residues within the CATD may be involved in centromere localization with some mutations giving a more gradual phenotype than others.

Mis18 is required for CENP-A retention at chromatin- Previous work has shown HJURP binds CENP-A/H4 as a heterodimer[28] through recognition of key amino acids within the CENP-A CATD [27,28] (Figure 3A). Given the α2.1-α2.3 mutants are contained within the CATD of CENP-A

we wanted to verify the previous works results showing these CENP-A CATD mutants could still bind HJURP[27]. To test this we used an exogenous LacO array in human U2OS cells[29] (Figure 3B-C). mCherry-LacI-HJURP^{FL} (mCLI-HJURP^{FL}) was tethered at the array and the same GFP versions of the CENP-A mutants were transfected in to assay recruitment frequency to the array. All of the mutants were able to be recruited to the array by HJURP as efficiently as CENP-A^{WT}. This suggests the lack of centromere localization of CENP-A$^{\alpha 2.2}$ (Figure 2C) is not due to its inability to bind HJURP (Figure 3C).

Having ascertained Mis18 directly and specifically binds CENP-A through the CATD (Figure 1) we hypothesized the lack of CENP-A$^{\alpha 2.2}$ recruitment to centromeres could be due to a disruption of the Mis18 interaction with CENP-A. To assess this, we again used the LacO array system (Figure 3D). Mis18α and Mis18β were co-tethered to the array and GFP-HJURP^{FL} along with HA versions of the CENP-A (Figure 2A) were transfected in as well to the assay recruitment efficiency of the CENP-A mutants to the array (Figure 3D-E). CENP-A^{L1}, CENP-A$^{\alpha 2.1}$, CENP-A$^{\alpha 2.3}$, and H3^{CATD} were all recruited to the Mis18 array as efficiently as CENP-A^{WT}, while CENP-A$^{\alpha 2.2}$ was not significantly recruitment to the array by the Mis18 complex. Taken together, these results show the Mis18 complex is absolutely required for CENP-A recruitment and retention at chromatin through recognition of specific residues within the CENP-A CATD.

Mis18 recognizes histidine 104 within the CATD of CENP-A- CENP-A/H4 exists as a heterotetramer in solution with two copies of CENP-A along with two copies of H4[25]. HJURP/Scm3 binding occludes the CENP-A/CENP-A interface to bind CENP-A/H4 as a heterodimer[28,30,31]. The CENP-A/H4 interface is kept intact even upon HJURP/Scm3 binding. The crystal structure of the HJURP bound to CENP-A/H4 dimer reveals six residues within the CATD that contact H4 to form hydrophobic stitches[28,32]. Previous research has shown these six residues within the CATD, when mutated to their H3 counterparts (Figure 4A, green residues, B, yellow residues) can still bind HJURP but lose centromeric targeting[27]. This suggested to us that the lack

of centromere targeting and Mis18 interaction of CENP-A$^{\alpha 2.2}$ could be due to this mutant's inability to form an intermediate tetramer before being assembled into nucleosomes at centromeres. Based on the crystal structure (Figure 4B)[32], the only residue within CENP-A$^{\alpha 2.2}$ that is not required for an interaction with H4 is the histidine 104 residue with the CATD (Figure 4A-B, purple residue).

We mutated this single residue to its H3 counterpart, Glycine (Figure 4A) to make a YFP-CENP-A^{H104G} mutant and tested it for centromere localization in U2OS cells in comparison to YFP-CENP-A^{WT} (Figure 4C-D). Fluorescence intensity of individual centromeres showed localization of the CENP-A^{H104G} mutant was completely abolished and western blot revealed it was expressing just as well as its CENP-A^{WT} counterpart (Figure 4E).

From the crystal structure of CENP-A/H4 in complex with HJURP (Figure 4F)[28], The H104 side-chain within the CATD has an aromatic interaction with F29 within the Scm3 domain of HJURP. We tested whether this residue is necessary for CENP-A/HJURP interaction using the CENP-A^{H104G} mutant and assaying its ability to recruit to the array in LacO-U2OS cells with mCLI-HJURP^{FL} tethered there (Figure 4G). YFP-CENP-A^{H104G} was recruited to the array by HJURP in similar frequency to its WT counterpart. These data show the lack of CENP-A^{H104G} recruitment to centromeres is not due to a lack of HJURP binding.

Given that CENP-A^{H104G} cannot be recruited to centromeres, yet can still interact with HJURP, we hypothesized its lack of centromere recruitment and retention at chromatin could be due to the disruption of the Mis18 interaction. To test this, Mis18α and Mis18β were co-tethered to the array in LacO-U2OS and HA-HJURP^{FL} along with YFP-CENP-A$^{WT/H104G}$ were transfected in as well (Figure 4H). Strikingly, YFP-CENP-A^{H104G} was unable to be recruited to the Mis18 arrays at all which was in stark contrast to CENP-A^{WT} which showed robust recruitment efficiency. Taken together, this shows a specific interaction between the Mis18 complex and H104 within the

CATD of CENP-A that is necessary for CENP-A recruitment and retention at centomeric chromatin.

Mis18 YIPPEE domains required for CENP-A retention at chromatin- Recent work has found the crystal structure of the YIPPEE domain in fission yeast Mis18[21], of which both Mis18α and Mis18β are paralogs. This conserved YIPPEE domain is also present in Mis18α and Mis18β (Figure 5A). Previous work by our lab showed the Mis18 coiled-coil domains are required and necessary for HJURP binding[10]. To determine if the coiled-coils were also regulating CENP-A binding and retention at chromatin or if this was a function of the YIPPEE domains, we tethered the coiled-coil domains or full-length versions of Mis18α/β to the array and transfected in YFP-CENP-A^{WT} in the presence and absence of HA-HJURP^{FL} (Figure 5C). We then assayed YFP-CENP-A^{WT} recruitment efficiency to the array. Surprisingly, YFP-CENP-A was not found at any of the arrays harboring the minimal Mis18 coiled-coils even though HA-HJURP^{FL} was still recruited as efficiently as the full-length Mis18 proteins. This is in contrast to the full-length Mis18 proteins at the array, which retained CENP-A at chromatin. These results suggest that there may be a hand off mechanism of CENP-A from HJURP to Mis18 into chromatin, and the YIPPEE domains of Mis18 are required for this process.

From these results it was clear the Mis18 YIPPEE domains are required for CENP-A retention at chromatin. Crystallization of fission yeast Mis18 YIPPEE domain confirmed that the previously reported key amino acids are not only highly conserved but integral for the YIPPEE domain formation[21]. These amino acids include the previously characterized CXXC motif[4], along with the newly found VF and DS motifs taken from the crystal structure (Figure 5B, asterisk's above residues). We tested if these conserved amino acids within the YIPPEE domain were regulating the Mis18/CENP-A interaction at chromatin. We tethered WT and mutant versions of Mis18α/β (Mis18α$^{V82A,\ F83A/C85A}$ and/or Mis18β$^{V77A,\ F78A/C80G}$) in different combinations to the LacO array and co-transfected in YFP-CENP-A^{WT} and HA-HJURP^{FL} (Figure 5B,D). None of the Mis18

YIPPEE mutants, even when co-tethered to the array with a WT counterpart, could retain CENP-A at chromatin despite still being able to recruit HA-HJURP^{FL}. Although these results are promising our previous work has shown these mutants are unable to recruit to centromeres[10]. This suggests the YIPPEE domains are required for multiple functions besides CENP-A retention.

From these initial Mis18 YIPPEE mutants (Figure 5B-C) we hypothesized we could scissor down a Mis18 separation of function mutant that could still recruit to centromeres while being unable to retain CENP-A at chromatin. Four single point mutants were created in Mis18α (GFP-Mis18α$^{V82A/F83A/D94A/S95A}$) and tested for centromere localization (Figure 5E). Strikingly, Mis18α^{V82A} retained its ability to recruit to centromeres as robustly as Mis18α^{WT}. Testing this single point mutant (and making the analogous point mutation in Mis18β^{V77A}) at the LacO array for its ability to retain CENP-A at chromatin is the next logical experiment. Despite the exciting result with this mutant, the Valine residue in the Mis18 proteins might very well be dispensable for CENP-A binding but should be tested nonetheless.

Taken together, these results show the conserved YIPPEE domain in the Mis18 proteins are necessary for CENP-A recruitment and overall retention at chromatin. This suggests a new function for the Mis18 complex as a part of the CENP-A deposition machinery at centomeric chromatin. Further study will be needed to elucidate this potential new role for the Mis18 complex in CENP-A deposition.

Discussion:

In this study we have shown the Mis18α-β complex directly interacts with CENP-A/H4 *in vitro*, and this interaction is taking place between a conserved H104 residue within the CATD of CENP-A. Moreover, the conserved YIPPEE domains of the Mis18 proteins are necessary for CENP-A binding and retention at chromatin, and without them CENP-A is not stably incorporated into centromeric chromatin. We propose a hand-off model of pre-nucleosomal

CENP-A from HJURP to the Mis18 complex upon HJURP recruitment to centromeres through a direct interaction with the Mis18 coiled-coils[10] (Figure 6). We observe the Mis18 conserved YIPPEE domains are necessary for CENP-A retention at chromatin and believe they are involved in the direct interaction we show *in vitro* between the Mis18 complex and CENP-A/H4. An alternative hypothesis is the Mis18 complex could be recognizing CENP-A within a nucleosome. We have previously established new CENP-A deposition at the centromere is dependent on the Mis18 complex being present in early G1-phase to recruit HJURP bound with pre-nucleosomal CENP-A[10]. Although we are assuming the Mis18 complex is interacting with pre-nucleosomal CENP-A, we cannot preclude the possibility of the Mis18 complex interacting with pre-existing CENP-A nucleosomes at the centromere. If the Mis18 YIPPEE domains are mutated to make them dysfunctional this could be abolishing a Mis18/nucleosomal CENP-A interaction that is necessary for Mis18 recognition of the centromere. Without Mis18 at centromeric chromatin no new CENP-A can be deposited. An initial testing can be done *in vitro* to assess if the Mis18 complex can recognize CENP-A within the context of a nucleosome using our recombinant proteins we have already generated.

Commented [DF1]: Discuss the alternative hypothesis, that Mis18 is recognizing CENP-A within the nucleosome.

We and others have shown the mutations within the Mis18α or Mis18β YIPPEE domain, including the CXXC motifs, eliminate their ability to bind to centromeres[10]. The YIPPEE domain then must be required for recognition of the centromere through its interactions with other factors such as Mis18BP1, as well as CCAN proteins such as CENP-C that is known to be involved in Mis18 centromere localization in humans[33,34]. Although the YIPPEE domain is clearly involved in Mis18 centromere recruitment and recognition we propose a dual role for the domain in stabilizing and retaining CENP-A at centromeric chromatin. We and others have made multiple mutants within the YIPPEE domain but have not gone in and systematically mutated residues on a point by point basis to assess if a Mis18 separation of function mutant exists that can recruit to centromeres but not bind and retain CENP-A at chromatin. Our Mis18α^{V82A} mutant showed

promise in that it is still able to localize to centromeres as robustly as WT. Testing this mutant, and making the analogous Mis18β^{V77A}, for CENP-A binding *in vitro* and at the array will be the next step we will take to assess if this is a valid mutant for our hypothesis.

The crystal structure of HJURP bound to CENP-A/H4 shows it binds the histones as a dimer[28] which is in contrast to the natural state of CENP-A/H4 in solution which exists as a stable heterotetramer[32]. This suggests that somehow during the deposition process of pre-nucleosomal CENP-A into nucleosomes it has to reform a tetramer, at least intermediately, before its stable incorporation into chromatin. This would be similar to the deposition dynamics of H3.1 which is chaperoned initially by the Asf1 and exists as a characteristic heterodimer with H4[35], just like CENP-A/H4 bound to HJURP. During DNA replication, H3/H4 dimers are thought to reform into tetramers through an interaction with the CAF complex and other proteins that target them to the replication forks, such as PCNA[18,36]. Unlike H3/H4 tetramer reformation, CENP-A/H4 tetramer reformation prior to nucleosome assembly has not been fully elucidated. We have shown that partially disrupting the CENP-A/H4 interaction through H3 residue substitution with the CENP-A$^{\alpha 2.2}$ mutant abolishes CENP-A recruitment to centromeres and hinders Mis18 from retaining CENP-A at chromatin. Furthermore, the *in vitro* binding of the Mis18 complex to CENP-A takes place with CENP-A/H4 that is initially purified as a tetramer before it is added to the binding reaction. From this data we propose the Mis18 complex recognizes and binds to CENP-A/H4 as a heterotetramer and could be integral in heterotetramer reformation from the initial HJURP bound CENP-A/H4 heterodimers. To assay this, two populations of recombinant CENP-A/H4 need to be made harboring different fluorophores. These pools should be mixed and the baseline FRET can be assessed between the two populations. Mis18 should be added in and FRET should again be recorded. If there is no change in FRET then this suggests Mis18 is binding CENP-A/H4 as a tetramer. Scm3 should be added in with the two populations of CENP-A/H4 separately from Mis18. We know Scm3 binds CENP-A/H4 as a dimer so there should be a loss of FRET signal

upon its addition. To show if Mis18 is involved in bringing the CENP-A/H4 dimers back into tetramers, Scm3 should be initially mixed with the two populations of CENP-A/H4. There should be a loss of FRET signal as the tetramers are bound as dimers. Increasing amounts of Mis18 should then be titrated into the reaction, if there is then an increasing FRET signal this shows Mis18 is not only binding CENP-A/H4 as a tetramer but it is helping it reform into a tetramer after the initial Scm3 binding. If not, then Mis18 is binding CENP-A/H4 as a dimer from HJURP and there are other factors that are helping CENP-A/H4 reform into a tetramer before nucleosome deposition.

Taken together, these results suggest a new role for the Mis18α–β complex in CENP-A deposition at centromeres. It is clear from this work that without Mis18, CENP-A is unable to be retained at chromatin signifying the involvement of the Mis18 complex as a new component in the CENP-A deposition machinery.

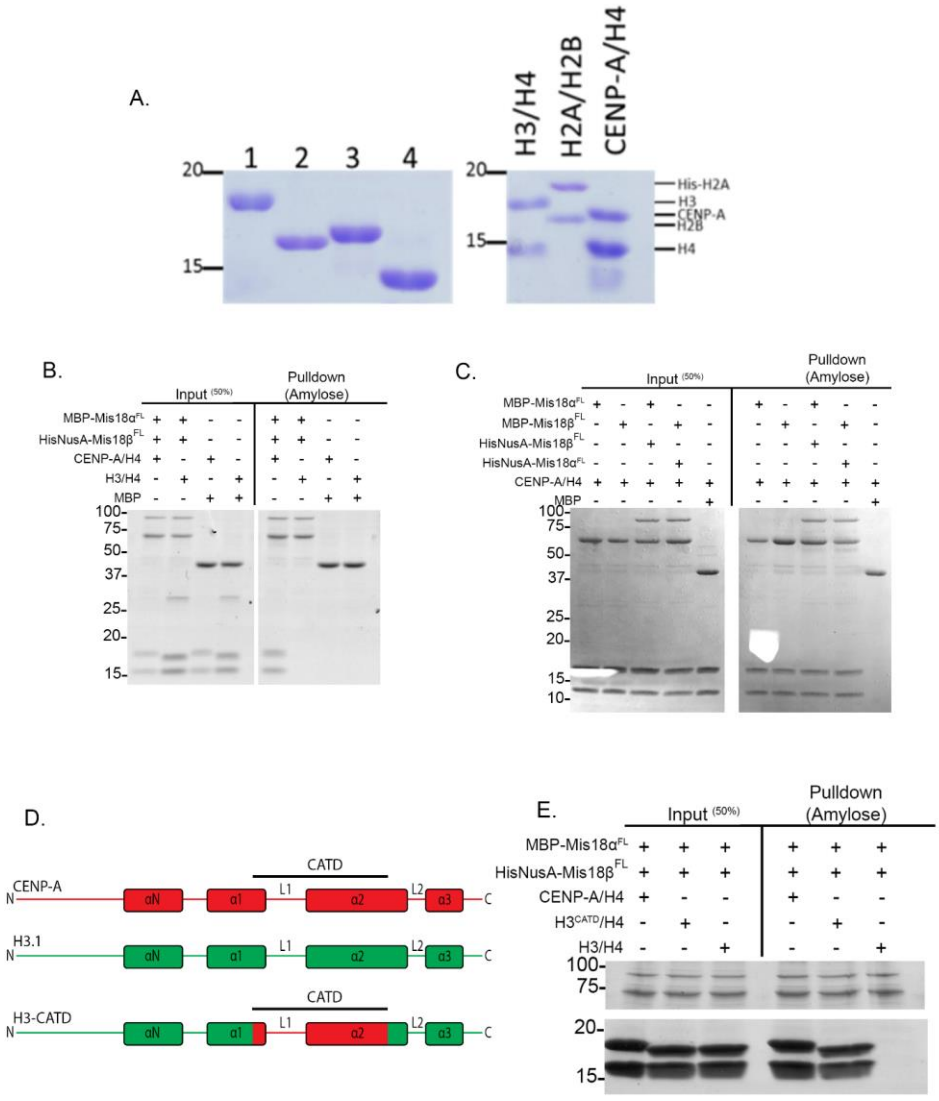

A.
1
2
3
4
20
15
H3/H4
H2A/H2B
CENP-A/H4
His-H2A
H3
CENP-A
H2B
H4
B.
Input (50%)
Pulldown (Amylose)
MBP-Mis18α^FL
HisNusA-Mis18β^FL
CENP-A/H4
H3/H4
MBP
100
75
50
37
25
20
15
C.
Input (50%)
Pulldown (Amylose)
MBP-Mis18α^FL
MBP-Mis18β^FL
HisNusA-Mis18β^FL
HisNusA-Mis18α^FL
CENP-A/H4
MBP
100
75
50
37
25
20
15
10
D.
CENP-A
H3.1
H3-CATD
N
C
αN
α1
L1
α2
L2
α3
CATD
E.
Input (50%)
Pulldown (Amylose)
MBP-Mis18α^FL
HisNusA-Mis18β^FL
CENP-A/H4
H3^CATD/H4
H3/H4
100
75
20
15

Figure 1

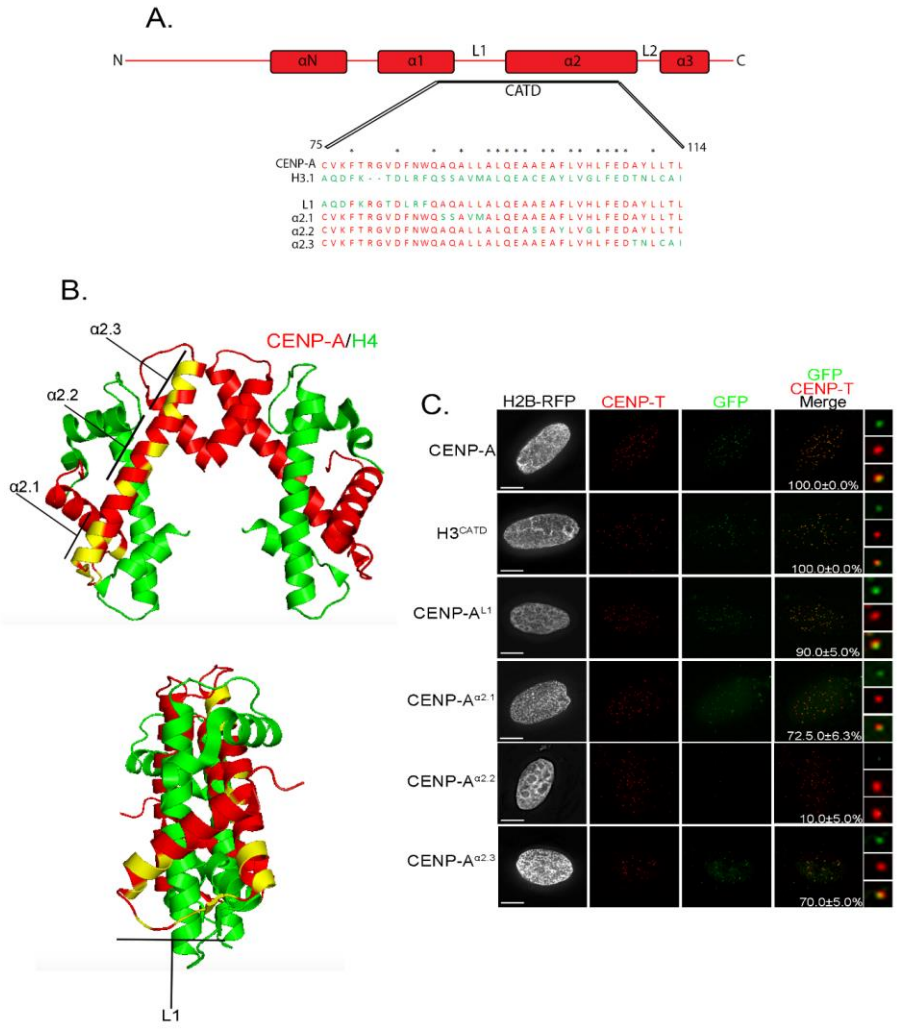

A.
N
αN
α1
L1
α2
L2
α3
C
CATD
75
114
CENP-A CVKFTRGVDFNWQAQALLALQEAAEAFLVHLFEDAYLLTL
H3.1 AQDFK--TDLRFQSSAVMALQEACEAYLVGLFEDTNLCAI
L1 AQDFKRGTDLRFQAQALLALQEAAEAFLVHLFEDAYLLTL
α2.1 CVKFTRGVDFNWQSSAVMALQEAAEAFLVHLFEDAYLLTL
α2.2 CVKFTRGVDFNWQAQALLALQEASEAYLVGLFEDAYLLTL
α2.3 CVKFTRGVDFNWQAQALLALQEAAEAFLVHLFEDTNLCAI
B.
α2.3
CENP-A/H4
α2.2
α2.1
L1
C.
H2B-RFP
CENP-T
GFP
GFP
CENP-T
Merge
CENP-A
100.0±0.0%
H3CATD
100.0±0.0%
CENP-AL1
90.0±5.0%
CENP-Aα2.1
72.5.0±6.3%
CENP-Aα2.2
10.0±5.0%
CENP-Aα2.3
70.0±5.0%

Figure 2

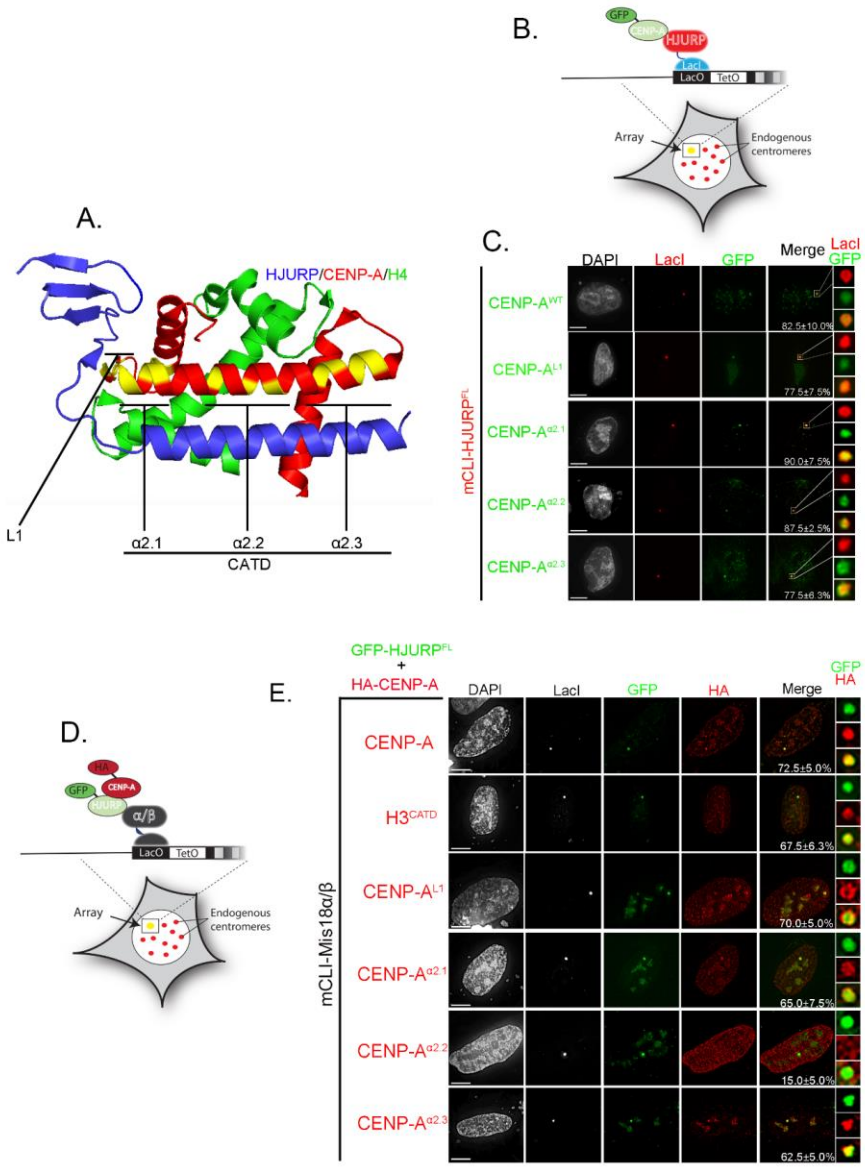

A.
HJURP/CENP-A/H4
L1
α2.1
α2.2
α2.3
CATD
B.
GFP
CENP-A
HJURP
LacI
LacO
TetO
Array
Endogenous centromeres
C.
DAPI
LacI
GFP
Merge
LacI GFP
mCLI-HJURP^FL
CENP-A^WT
CENP-A^L1
CENP-A^α2.1
CENP-A^α2.2
CENP-A^α2.3
82.5±10.0%
77.5±7.5%
90.0±7.5%
87.5±2.5%
77.5±6.3%
D.
HA
CENP-A
GFP
HJURP
α/β
LacO
TetO
Array
Endogenous centromeres
E.
GFP-HJURP^FL
+
HA-CENP-A
DAPI
LacI
GFP
HA
Merge
GFP HA
mCLI-Mis18α/β
CENP-A
H3^CATD
CENP-A^L1
CENP-A^α2.1
CENP-A^α2.2
CENP-A^α2.3
72.5±5.0%
67.5±6.3%
70.0±5.0%
65.0±7.5%
15.0±5.0%
62.5±5.0%

Figure 3

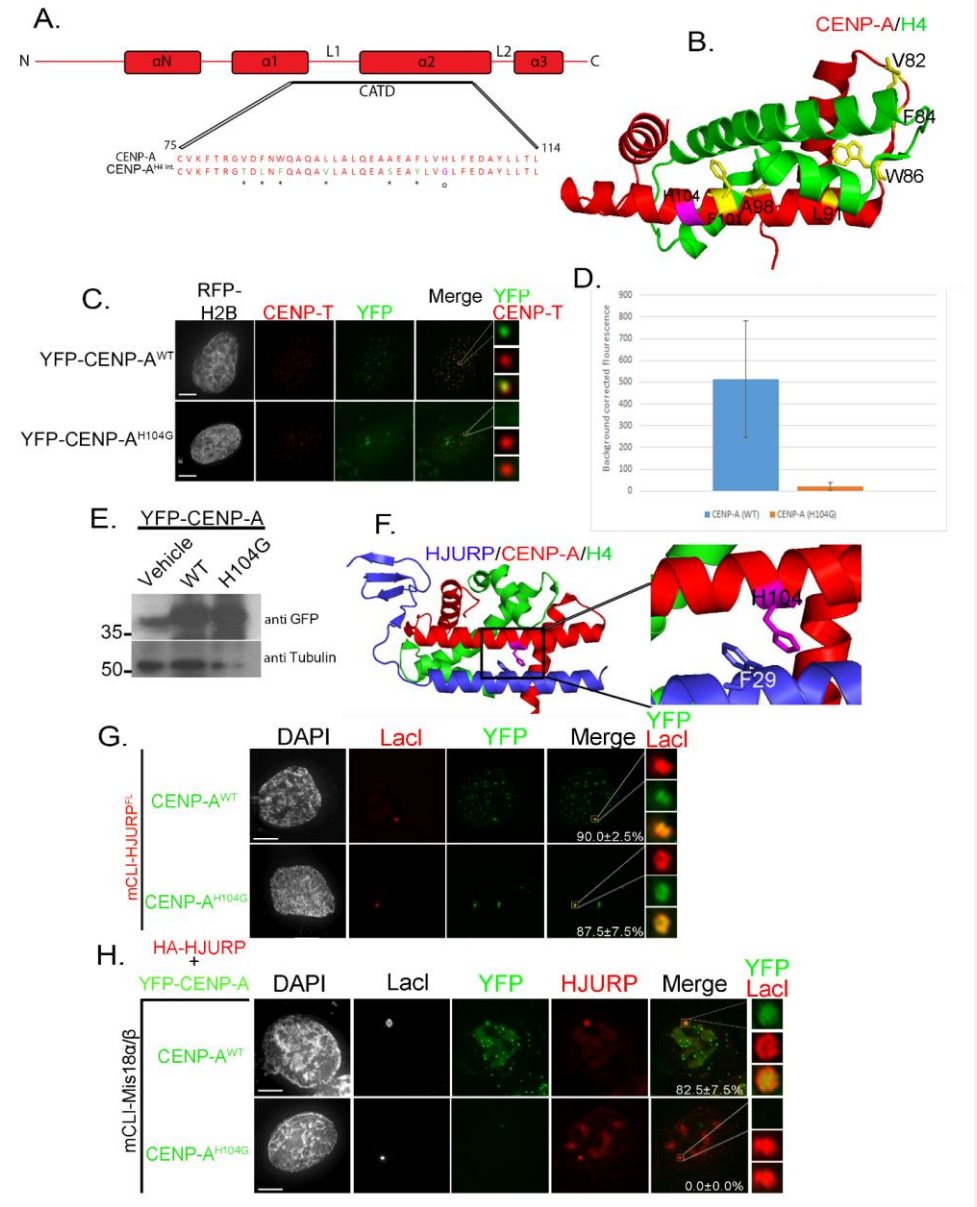

A.
N
αN
α1
L1
α2
L2
α3
C
CATD
75
114
CENP-A
CENP-A^H4 HL
B.
CENP-A/H4
V82
F84
W86
H104
A98
F101
L97
C.
RFP-H2B
CENP-T
YFP
Merge
YFP CENP-T
YFP-CENP-A^WT
YFP-CENP-A^H104G
D.
Background corrected flourescence
CENP-A (WT)
CENP-A (H104G)
E.
YFP-CENP-A
Vehicle
WT
H104G
35-
50-
anti GFP
anti Tubulin
F.
HJURP/CENP-A/H4
H104
F29
G.
DAPI
LacI
YFP
Merge
YFP LacI
mCLI-HJURP^FL
CENP-A^WT
CENP-A^H104G
90.0±2.5%
87.5±7.5%
H.
HA-HJURP
+
YFP-CENP-A
DAPI
LacI
YFP
HJURP
Merge
YFP LacI
mCLI-Mis18α/β
CENP-A^WT
CENP-A^H104G
82.5±7.5%
0.0±0.0%

Figure 4

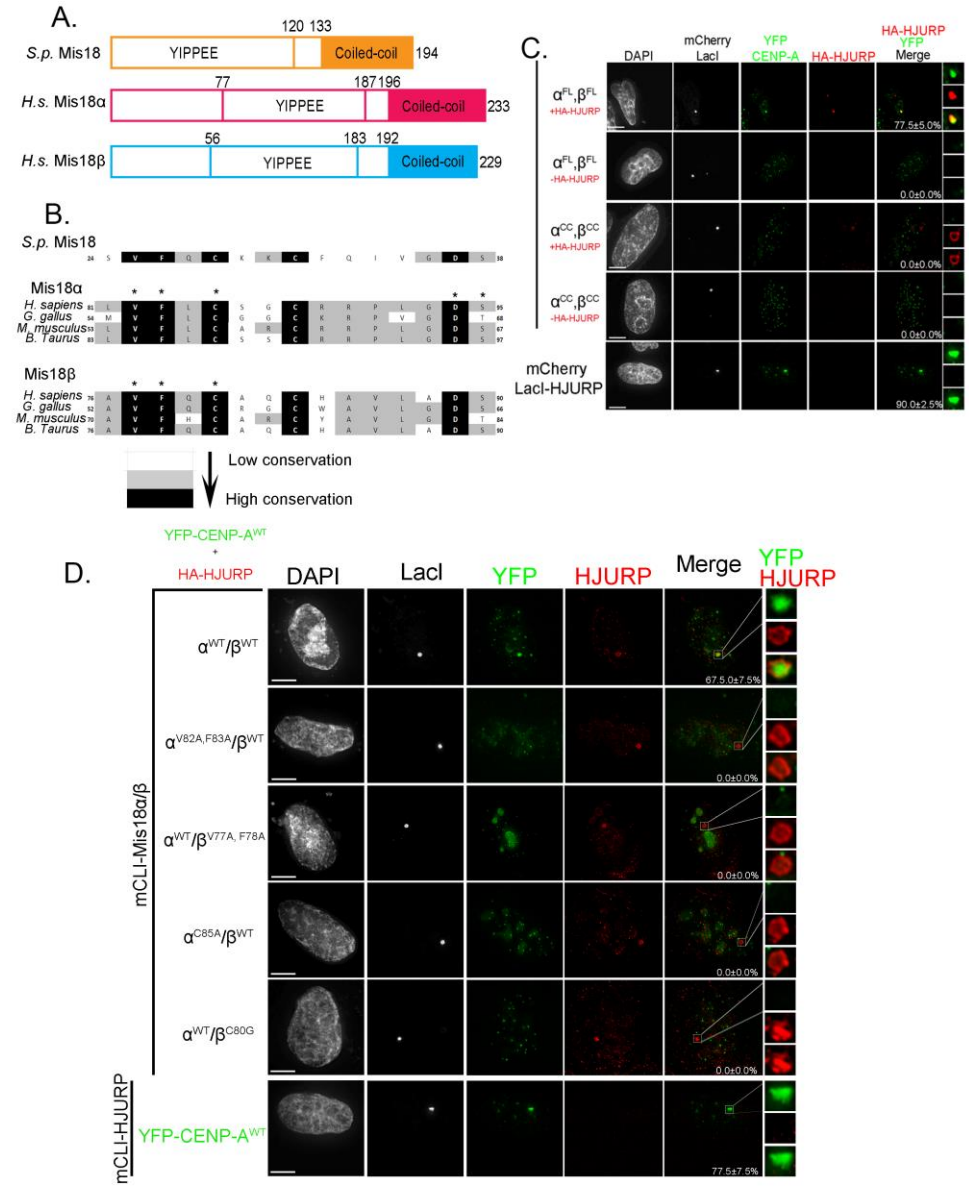

A.
S.p. Mis18
YIPPEE
Coiled-coil
H.s. Mis18α
H.s. Mis18β
B.
Mis18α
Mis18β
H. sapiens
G. gallus
M. musculus
B. Taurus
Low conservation
High conservation
C.
DAPI
mCherry LacI
YFP CENP-A
HA-HJURP
HA-HJURP YFP Merge
mCherry LacI-HJURP
77.5±5.0%
0.0±0.0%
90.0±2.5%
D.
YFP-CENP-AWT + HA-HJURP
DAPI
LacI
YFP
HJURP
Merge
YFP HJURP
mCLI-Mis18α/β
αWT/βWT
αV82A,F83A/βWT
αWT/βV77A, F78A
αC85A/βWT
αWT/βC80G
mCLI-HJURP
YFP-CENP-AWT
67.5±7.5%
77.5±7.5%

Figure 5

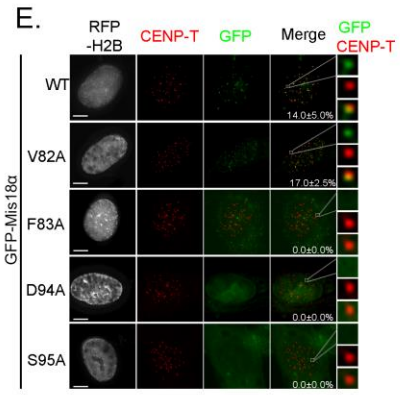

E.
RFP
-H2B
CENP-T
GFP
Merge
GFP
CENP-T
GFP-Mis18α
WT
V82A
F83A
D94A
S95A
14.0±5.0%
17.0±2.5%
0.0±0.0%
0.0±0.0%
0.0±0.0%

Figure 5

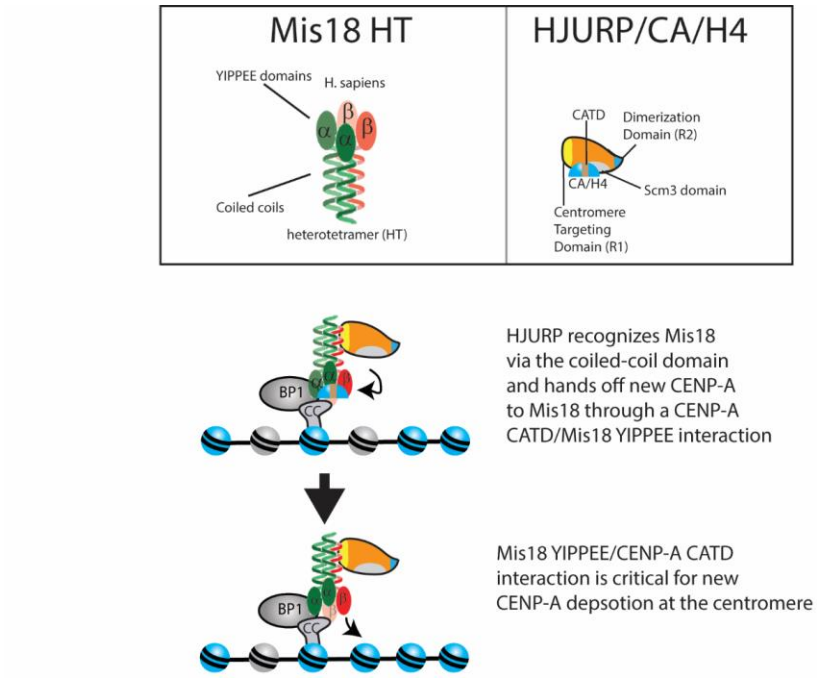

Mis18 HT
YIPPEE domains
H. sapiens
Coiled coils
heterotetramer (HT)
HJURP/CA/H4
CATD
Dimerization Domain (R2)
CA/H4
Scm3 domain
Centromere Targeting Domain (R1)
BP1
CC
HJURP recognizes Mis18 via the coiled-coil domain and hands off new CENP-A to Mis18 through a CENP-A CATD/Mis18 YIPPEE interaction
BP1
CC
Mis18 YIPPEE/CENP-A CATD interaction is critical for new CENP-A depsotion at the centromere

Figure 6

Figure Legends

Figure 1: The Mis18 complex directly and specifically binds to CENP-A/H4: A. (left) Purified recombinant histones His-H2A(1), H2B(2), H3.1(3), and H4(4). (right) Individual histones were mixed and dialysis was performed to make the individual H3/H4 tetramer and H2A/H2B dimer B. *In vitro* pulldown between the Mis18α-β complex and CENP-A/H4 or H3/H4 heterotetramer. MBP-Mis18α was used as the prey protein and pulled down with amylose beads. The Mis18 proteins were pre-incubated with each other to form the Mis18 complex prior to addition of CENP-A/H4 or H3/H4. MBP was used as a bait protein negative control to make sure CENP-A/H4 or H3/H4 were not specifically binding to the tag alone. C. *In vitro* pulldown as in B comparing the robustness of binding between Mis18α alone, Mis18β alone, or the pre-formed Mis18α-β complex. D. Diagram with domains of CENP-A, Histone H3.1, and the chimera H3-CATD used in the pulldown in E. *In vitro* pulldown as in B comparing the binding of CENP-A/H4, H3-CATD/H4, and H3/H4 tetramers with the pre-formed Mis18 complex.

Figure 2: Specific residues within CATD are required to target CENP-A to centromeres: A. Diagram of CENP-A and sequences of the CENP-A mutants used in the experiments. Asterisk's above residues indicate residues shared between H3 and CENP-A. B. Crystal structure of CENP-A/H4 heterotetramer[32] (CENP-A=red, H4=green) with mutated residues from 2A highlighted in yellow along with their mutant names. C. Parental U2OS cells transfected with GFP versions of the indicated CENP-A mutants. Cells were stained with CENP-T as an individual centromere marker and RFP-H2B was used as an independent marker for transfection. Percentages represent proportion of cells with centromere localization while images to the right are of an individual magnified centromere in the cells. Scale bars, 5um.

Figure 3: Mis18 is required for CENP-A retention at chromatin: A. Crystal structure of CENP-A/H4 dimer in complex with the Scm3 domain of HJURP[28] (HJURP=blue, CENP-A=red, H4=green). Mutated residues are highlighted in yellow along with their mutant names. B.

Schematic of the LacO array in U2OS cells showing the LacI-HJURP^{FL} tethered to the array and co-transfected in with different versions of GFP-CENP-A. This is the experimental setup used in C. U2OS-LacO cells were co-transfected with mCLI-HJURP^{FL}, GFP-CENP-A$^{L1/\alpha2.1/\alpha2.2/\alpha2.3}$. GFP-CENP-A^{WT} was used as a positive control. Percentage of cells with recruitment of the prey proteins (GFP-CENP-A) to the array are shown. Magnified arrays are shown to the right of the merged images. Scale bars, 5μm. D. Schematic of the LacO array in U2OS cells showing LacI-Mis18α/β tethered to the array and co-transfected in with GFP-HJURP^{FL} and different versions of HA-CENP-A. This is the experimental setup used in E. U2OS-LacO cells were co-transfected with mCLI-Mis18α/β, GFP-HJURP^{FL}, and HA-CENP-A$^{L1/\alpha2.1/\alpha2.2/\alpha2.3}$. HA-CENP-A^{WT} and H3^{CATD} were used as positive controls. Percentage of cells with recruitment of the prey proteins (HA-CENP-A/HA-H3^{CATD}) to the array are shown. Magnified arrays are shown to the right of the merged images. Scale bars, 5μm.

Figure 4: Mis18 recognizes histidine 104 within the CATD of CENP-A: A. Diagram of CENP-A and the mutated residues (astericks, green) noted in the previous study that abolish CENP-A/H4 tetramerization interactions[32]. H104 within the CATD, based on the crystal structure does not help confer a CENP-A/H4 interaction and was mutated to its H3 counterpart (Glycine, purple residue) to make CENP-A^{H104G}. B. Crystal structure of a CENP-A/H4 dimer (CENP-A=red, H4=green) showing the six CENP-A residues within the CATD that help confer a CENP-A/H4 interaction. C. Parental U2OS cells transfected with YFP versions of CENP-A^{WT} or CENP-A^{H104G}. Cells were stained with CENP-T as an individual centromere marker and RFP-H2B was used as an independent marker for transfection. Images to the right of the merged are of an individual magnified centromere in the cells. Scale bars, 5um. D. Background corrected fluorescence intensity of YFP-CENP-A$^{WT/H104G}$ at individual centromeres. Values are averages from all the centromeres pooled, +/- S.D. E. Western blot of YFP-CENP-A$^{WT/H104G}$ to assess expression levels in parental U2OS cells. Tubulin was used as a loading control. F. Crystal

structure of CENP-A/H4 dimer in complex with the Scm3 domain of HJURP[28] (HJURP=blue, CENP-A=red, H4=green). Magnified image to the right highlights the H104 residue (purple) within the CATD of CENP-A interaction with F29 contained in the Scm3 domain of HJURP. G. U2OS-LacO cells were co-transfected with mCLI-HJURP^{FL} and YFP-CENP-A$^{WT/H104G}$. YFP-CENP-A^{WT} was used as a positive control. Percentage of cells with recruitment of the prey proteins (YFP-CENP-A) to the array are shown. Magnified arrays are shown to the right of the merged images. Scale bars, 5μm. H. U2OS-LacO cells were co-transfected with mCLI-Mis18α/β, HA-HJURP^{FL}, and YFP-CENP-A$^{WT/H104G}$. YFP-CENP-A^{WT} was used as a positive control. Percentage of cells with recruitment of the prey proteins (YFP-CENP-A) to the array are shown. Magnified arrays are shown to the right of the merged images. Scale bars, 5μm.

Figure 5: Mis18 YIPPEE domains required for CENP-A retention at chromatin: A. Diagram showing the conserved YIPPEE and coiled-coil domains in *S.p.* Mis18 and *H.s.* Mis18α and Mis18β. B. Amino acid conservation of a portion of the Mis18 YIPPEE domain across eukaryotes. The alignments include orthologs from *S. pombe (S.p.), G. gallus, M. musculus, H. sapiens,* and *B. Taurus*. Amino acid residues mutated in this study are defined by asterick's. C. LacO-U2OS cells were used. MCH-LI-Mis18$^{FL/CC}$ were tethered at the array and YFP-CENP-A was co-transfected into the cells in the presence and absence of HJURP. Cells were stained with DAPI to visualize nuclei and an HA antibody (Covance) was used to visualize the HA-HJURP. To the right of the merged images the array is magnified. Percentages indicate the proportion of arrays that were positive for YFP-CENP-A +/-SD. MCH-LI-HJURP tethered to the array as a positive control for YFP-CENP-A recruitment. Scale bars, 5um. D. U2OS-LacO cells were co-transfected with mCLI-Mis18α$^{WT/V82A,F83A/C85A}$/β$^{WT/V77A,F78A/C80G}$, HA-HJURP^{FL}, and YFP-CENP-A^{WT}. Percentage of cells with recruitment of the prey proteins (YFP-CENP-A) to the array are shown. Magnified arrays are shown to the right of the merged images. Scale bars, 5μm. MCLI-HJURP^{FL} co-transfected with YFP-CENP-A^{WT} was used as a positive control for CENP-A

recruitment to the array. E. Parental U2OS cells transfected with GFP versions of the indicated Mis18α mutants. Mis18α^{WT} was used as a positive control. Cells were stained with CENP-T as an individual centromere marker and RFP-H2B was used as an independent marker for transfection. Percentages represent proportion of cells with centromere localization while images to the right of the merged are of an individual magnified centromere in the cells. Scale bars, 5um.

Figure 6: Mis18 depicted as a factor crucial for CENP-A retention at centromeric chromatin.

Materials and Methods:

DNA Constructs Gateway cloning was to create DNA constructs first by PCR amplification with flanking attB sites using Vent polymerase and cloning into the PDONR221 entry plasmid using BP clonase enzyme (Invitrogen). Entry clones were transferred into destination vectors as previously described using LR clonase[10]

Cell transfections and immunocytochemistry. Parental U2OS and LacO-TRE (Janicki et al. 2004) cells were transfected with Lipofectamine-2000 (Invitrogen) according to the manufacturers protocol. Cells were processed for immunocytochemistry 48 hrs after transfection. Immunocytochemistry and immunoprecipitation experiments were conducted as described previously (Zasadzinska et al. 2013). Metaphase spreads were prepared as described in the supplemental materials and methods. Images of fixed cells were collected using a 100× oil immersion objective lens on a DeltaVision deconvolution microscope (Applied Precision) using SoftWoRX acquisition software. Images were deconvolved and presented as stacked images. Images within experiments were collected with identical exposure times and scaled equally.

In-vitro pull-downs. Recombinant proteins were purified as described in the supplemental materials and methods. Proteins were combined for 3 hrs at room temperature at 1:1 molar ratio in binding buffer: 50mM Tris-HCl pH 7.5, 250mM NaCl, 20mM $MgCl_2$, 0.5% NP-40, 10% glycerol, and 5mM BME. For pull-downs involving HJURP, Mis18α and Mis18β were pre-

incubated for 3 hrs to form the complex prior to adding HJURP. Affinity matrices were pre-incubated in the binding buffer with 0.2 mg/ml BSA for 1 hr and added to pre-formed complexes for 40 min. Complexes were washed 6-times in 1ml of binding buffer, for Ni-NTA 40mM imidazole was added. Bound complexes were eluted in SDS sample buffer and boiled. Western blots were performed using antibodies against the 6X His (Santa-Cruz), MBP (New England Biolabs), Mis18β (Bethyl Labs), and the HA epitope (Covance).

Protein expression and purification. Recombinant proteins were expressed in the Rosetta (DE3) pLysS bacterial strain. Bacteria were grown in LB media to an OD of 0.6 at 37ºC and induced at 18ºC with 0.1mM IPTG for 16 hr. His-tagged HJURP fragments were purified on Ni-NTA (Qiagen). Bacteria were lysed using a steel Wheaton-dounce in buffer contained 50mM Tris-HCl pH 7.5, 250mM NaCl, 20mM MgCl2, 0.5mM CaCl2, 10% glycerol, 0.1% NP-40, 5mM BME, LPC, 1mM PMSF, 20mM imidazole, and 0.15mg/ml RNase-A. DnaseI was added and lysates were centrifuged at 22,000x*g* for 20 min. Supernatants were collected and pellets were re-extracted with a second round of lysis. Cleared lysates were incubated with Ni-NTA agarose (Qiagen) for 1 hour and washed twice in lysis buffer with 40mM imidazole. Proteins were eluted in lysis buffer containing 250mM imidazole and no protease inhibitors. Full-length HJURP was further purified on a Superdex 200 10/300 column (G.E. Healthcare) and re-concentrated on Ni-NTA agarose (Qiagen). Mis18 proteins were purified as described in lysis buffer containing: 50mM Tris-HCl, 350mM NaCl, 0.5mM CaCl2, 10% glycerol, 0.1% NP-40, 5mM BME, LPC, and 1mM PMSF, and imidazole. MBP tagged proteins were purified using maltose agarose (Qiagen) and eluted in buffer containing 10mM maltose. Strep–tagged proteins were purified using Strep-Tactin Super-flow plus (Qiagen) and eluted in buffer containing 2.5mM d-desthiobiotin.

CENP-A/H4 Purification:

1. Thaw bacterial resuspension. Homogenize each 10 ml resuspension (from 500 ml culture) in an ice-cold 15 ml capacity Wheaton dounce-type stainless steel homogenizer until homogeneous (~5 full passes). Spin at 26,000 X g for 15'. Discard supernatant containing soluble proteins (but not CENP-A/H4). Resuspend pellet in 10 ml hydroxyapatite start buffer (80 mM sodium phosphate (pH=6.9), 300 mM NaCl, leupeptin (2.5 ug/ml), pepstatin (2.5 ug/ml), aprotinin (3.15 ug/ml), 1 mM PMSF, 14 mM BME) by pipetting up and down. Spin as above. Discard supernatant. Add 10 ml hydroxyapatite start buffer.

2. Sonicate pellet for 30 s at ~80% capacity, constant pulse with microtip sonicator. Repeat four times with at least 2' between sonications to chill in a 50/50 ice/water slurry.

**note: CENP-A/H4 is much more efficiently released from the insoluble pellet when sonicated in small batches (10 ml in a 50 ml tube), as opposed to pooling prior to this step.

3. Spin at 26,000 X g for 15' and collect supernatant.

4. Pre-equilibrate an HT column with 2 column volumes 80 mM sodium phosphate, 300 mM NaCl, and finally 1 column volume hydroxyapatite start buffer prior to loading sample. Spin at 26,000 X g for 15' again to clarify, and then apply to a Biogel HT (BioRad) column. Use 2 ml packed resin per 1 ml sonicated extract (this ratio is based on the limited DNA-binding capacity of the resin and the high [DNA] in the sonicated extract).

5. Wash column with 1 column volume hydroxyapatite start buffer, and 5-6 column volumes of hydroxyapatite wash buffer (80 mM sodium phosphate (pH=6.9), 500 mM NaCl, leupeptin (2.5 ug/ml), pepstatin (2.5 ug/ml), aprotinin (3.15 ug/ml), 1 mM PMSF, 14 mM BME). Monitor DNA and protein contaminants by $A_{260/280}$ in the wash fractions.

6. Elute CENP-A/H4 with 80 mM sodium phosphate (pH=6.9), 2.5 M NaCl, 1 mM PMSF, 14 mM BME.

7. Pool CENP-A/H4 fractions and dialyze into 20 mM sodium phosphate, 450 mM NaCl, 1 mM PMSF, 14 mM BME. I did 35 ml sample into 2 L buffer to get the final [NaCl] = ~500 mM.

8. Spin at 26,000 X g for 15' again to clarify, and then apply sample to HiTrap SP FF (1 ml column). Buffer A=20 mM sodium phosphate (pH=6.9), 500 mM NaCl, 1 mM PMSF, 14 mM BME. Buffer B= 20 mM sodium phosphate (pH=6.9), 2 M NaCl, 1 mM PMSF, 14 mM BME. Wash with 5 column volumes Buffer A. Then do a 0-20% (A→B) gradient over 20 column volumes. Then step to 100% B to elute CENP-A/H4 (should be a very sharp peak).; do 10 column volumes.

Histone H3/H4 Purification: See protocol linked below

http://research.fhcrc.org/content/dam/stripe/tsukiyama/files/Protocols/expression.pdf

Histone H3/H4 tetramer and H2A/H2B dimer reconstitution: To reconstitute the H3.1/H4 tetramer, the histones (about 5 mg of each histone), purified under denaturing conditions noted above, were mixed in a 1:1 stoichiometry. Then, the mixture was dialyzed against 1 liter of 50 mM Tris–HCl buffer (pH 8.0) containing 10 mM DTT, 2 mM EDTA, 6 M urea, and 2 M NaCl for 1 h at room temperature, and the dialysis was continued overnight at 4 C. After overnight dialysis, the H2A/H2B mixture was dialyzed against 1 liter of 20 mM Tris–HCl buffer (pH 8.0) containing 5 mM DTT, 1 mM EDTA, 1 mM PMSF, 5% glycerol, and 2 M NaCl, to renature the H3/H4 tetramer. After an overnight dialysis against the buffer without urea, the impurities and excess histones were precipitated. The salt concentration was reduced by stepwise dialysis against the buffer containing 1 M NaCl for 4 h, 0.5 M NaCl for 4 h, and finally 0.1 M NaCl overnight at 4 C. After centrifugation at 27,000g for 20 min, the reconstituted H3/H4 tetramer was obtained in the soluble fraction and was analyzed by 15 SDS–PAGE[37].

References:

1. Dunleavy, E. M. *et al*. HJURP is a cell-cycle-dependent maintenance and deposition

factor of CENP-A at centromeres. *Cell* **137,** 485–97 (2009).

2. Foltz, D. R. *et al.* Centromere-specific assembly of CENP-a nucleosomes is mediated by HJURP. *Cell* **137,** 472–84 (2009).

3. Stoler, S. *et al.* Scm3, an essential Saccharomyces cerevisiae centromere protein required for G2/M progression and Cse4 localization. *Proc. Natl. Acad. Sci. U. S. A.* **104,** 10571–6 (2007).

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

5. Camahort, R. *et al*. Cse4 is Part of an Octameric Nucleosome in Budding Yeast. *Mol. Cell* **35,** 794–805 (2009).

6. Mizuguchi, G., Xiao, H., Wisniewski, J., Smith, M. M. & Wu, C. Nonhistone Scm3 and histones CenH3-H4 assemble the core of centromere-specific nucleosomes. *Cell* **129,** 1153–64 (2007).

7. Sanchez-Pulido, L., Pidoux, A. L., Ponting, C. P. & Allshire, R. C. Common ancestry of the CENP-A chaperones Scm3 and HJURP. *Cell* **137,** 1173–4 (2009).

8. Williams, J. S., Hayashi, T., Yanagida, M. & Russell, P. Fission yeast Scm3 mediates stable assembly of Cnp1/CENP-A into centromeric chromatin. *Mol. Cell* **33,** 287–98 (2009).

9. Moree, B., Meyer, C. B., Fuller, C. J. & Straight, A. F. CENP-C recruits M18BP1 to centromeres to promote CENP-A chromatin assembly. *J. Cell Biol.* **194,** 855–71 (2011).

10. Nardi, I. K., Zasadzińska, E., Stellfox, M. E., Knippler, C. M. & Foltz, D. R. Licensing of Centromeric Chromatin Assembly through the Mis18a-Mis18B Heterotetramer. *Mol. Cell* **61,** 774–787 (2016).

11. Wang, J. *et al.* Mitotic regulator Mis18β interacts with and specifies the centromeric

assembly of molecular chaperone HJURP. *J. Biol. Chem.* (2014). doi:10.1074/jbc.M113.529958

12. Verreault, A., Kaufman, P. D., Kobayashi, R. & Stillman, B. Nucleosome Assembly by a Complex of CAF-1 and Acetylated Histones H3/H4. *Cell* **87,** 95–104 (2015).

13. Green, E. M. *et al.* Replication-Independent Histone Deposition by the HIR Complex and Asf1. *Curr. Biol.* **15,** 2044 (2005).

14. Han, J. *et al.* Rtt109 Acetylates Histone H3 Lysine 56 and Functions in DNA Replication. *Science (80-.).* **315,** 653–655 (2007).

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

16. Tyler, J. K. *et al.* Interaction between the Drosophila CAF-1 and ASF1 Chromatin Assembly Factors. *Mol. Cell. Biol.* **21,** 6574–6584 (2001).

17. Moggs, J. G. *et al.* A CAF-1–PCNA-Mediated Chromatin Assembly Pathway Triggered by Sensing DNA Damage. *Mol. Cell. Biol.* **20,** 1206–1218 (2000).

18. Shibahara, K. & Stillman, B. Replication-Dependent Marking of DNA by PCNA Facilitates CAF-1-Coupled Inheritance of Chromatin. *Cell* **96,** 575–585 (1999).

19. Shelby, R. D., Vafa, O. & Sullivan, K. F. Assembly of CENP-A into Centromeric Chromatin Requires a Cooperative Array of Nucleosomal DNA Contact Sites. *Cell* **136,** 501–513 (1997).

20. Shelby, R. D., Monier, K. & Sullivan, K. F. Chromatin Assembly at Kinetochores Is Uncoupled from DNA Replication. *Reactions* **151,** 1113–1118 (2000).

21. Subramanian, L. *et al*. Centromere localization and function of Mis18 requires Yippee-like domain-mediated oligomerization. *EMBO Rep.* (2016). at

<http://embor.embopress.org/content/early/2016/03/02/embr.201541520.abstract>

22. Roxström-Lindquist, K. & Faye, I. The Drosophila gene Yippee reveals a novel family of putative zinc binding proteins highly conserved among eukaryotes. *Insect Mol. Biol.* **10,** 77–86 (2001).

23. Hori, T. *et al.* CCAN makes multiple contacts with centromeric DNA to provide distinct pathways to the outer kinetochore. *Cell* **135,** 1039–52 (2008).

24. Foltz, D. R. *et al.* The human CENP-A centromeric nucleosome-associated complex. *Nat. Cell Biol.* **8,** 458–69 (2006).

25. Black, B. E. *et al.* Structural determinants for generating centromeric chromatin. *Nature* **430,** 578–82 (2004).

26. Shelby, R. D., Vafa, O. & Sullivan, K. F. Assembly of CENP-A into Centromeric Chromatin Requires a Cooperative Array of Nucleosomal DNA Contact Sites. *J. Cell Biol.* **136,** 501–513 (1997).

27. Bassett, E. a *et al.* HJURP uses distinct CENP-A surfaces to recognize and to stabilize CENP-A/histone H4 for centromere assembly. *Dev. Cell* **22,** 749–62 (2012).

28. Hu, H. *et al.* Structure of a CENP-A – histone H4 heterodimer in complex with chaperone HJURP. 901–906 (2011). doi:10.1101/gad.2045111.GENES

29. Janicki, S. M. *et al.* From silencing to gene expression: real-time analysis in single cells. *Cell* **116,** 683–98 (2004).

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

31. Zhou, Z. *et al.* Structural basis for recognition of centromere histone variant CenH3 by

the chaperone Scm3. *Nature* **472,** 234–237 (2011).

32. Sekulic, N., Bassett, E. a, Rogers, D. J. & Black, B. E. The structure of (CENP-A-H4)(2) reveals physical features that mark centromeres. *Nature* **467,** 347–51 (2010).

33. Maddox, P. S., Hyndman, F., Monen, J., Oegema, K. & Desai, A. Functional genomics identifies a Myb domain-containing protein family required for assembly of CENP-A chromatin. *J. Cell Biol.* **176,** 757–63 (2007).

34. Moree, B., Meyer, C. B., Fuller, C. J. & Straight, A. F. CENP-C recruits M18BP1 to centromeres to promote CENP-A chromatin assembly. *J. Cell Biol.* (2011). doi:10.1083/jcb.201106079

35. Tagami, H., Ray-gallet, D. & Nakatani, Y. Mediate Nucleosome Assembly Pathways Dependent or Independent of DNA Synthesis. **116,** 51–61 (2004).

36. Ransom, M., Dennehey, B. K. & Tyler, J. K. Chaperoning Histones during DNA Replication and Repair. *Cell* **140,** 183–195 (2010).

37. Tanaka, Y. *et al*. Expression and purification of recombinant human histones. **33,** 3–11 (2004).

CHAPTER 4: The CENP T/W/S/X experiments: The long lost project/ HJURP and its involvement in DNA damage

Abstract:

CENP T/W/S/X were found to form a complex and be interdependent on one another for centromere localization and CCAN integrity[23,26,55,71]. The T/W, S/X, and T/W/S/X complex have been crystallized and found to form a "nucleosome-like" particle that looks strikingly like a canonical nucleosome[26]. At the start of my tenure in the Foltz lab I was interested in looking at how the CCAN was involved in specifying the centromere every cell cycle. I set out to elucidate the mechanism for how the T/W/S/X complex was involved in centromere specification. This chapter is devoted to the early work I did trying to purify the complex and get proper reagents ready to begin to answer questions such as how is the T/W/S/X complex wrapping DNA? Is it protecting the linker DNA between nucleosomes? Does it preferentially interact with H3 or CENP-A nucleosomes? Do CENP-T, W, S, and X have chaperones similar to canonical histones? Lastly, Do CENP-S and X localize the Fanconi Anemia complex to centromeres and what is the consequence of this? This chapter will go over the reagents I have in hand as well as future questions that can be taken on by a new member coming into the Foltz lab.

In part 2 of this chapter I will go over preliminary work on HJURPs potential involvement in the DNA damage and repair pathway. To date only one paper has been published on this topic and it was this paper that gave HJURP its name (Holliday-junction recognition protein) as it had an affinity for holliday-junction DNA *in vitro*[128]. This work used flow cytometry as a main technique, which has a steep learning curve, and my initial results were not able to be replicated easily. I will discuss the pitfalls and data that I have and possible troubleshooting steps one can take to successfully reopen this project.

Introduction:

CENP-S/X and the Fanconi Anemia complex

CENP-S was identified in CENP-M and CENP-U/50 affinity purifications[23,25]. CENP-S was found to be a binding partner to a protein that went by the name of Stra13, otherwise known as CENP-X. Both proteins have histone-fold domains and look strikingly like canonical histones[26]. Depletion of either CENP-S or CENP-X induces mitotic abnormalities and kinetochore instability. The presence of the S/X complex is necessary for recruitment of kinetochore proteins KNL1 and Ndc80/Hec1[25]. CENP-T was also detected as a CENP-S associated protein and together these CENPs form a T/W/S/X heterotetramer *in vitro*[26]. As of now the relationship between CENP-S/X and CENP-T *in vivo* remains unclear. CENP-S/CENP-X are also known as MHF1/MHF2 which are biding partners of FANCM, a member of the Fanconi anemia nuclear pore complex[27,129]. Whether or not FANCM or other members of the Fanconi anemia complex localize the centromere and what the function of this in reference to centromere propagation is still not known at this time.

Fanconi anemia (FA) is a complex and chronic disorder that takes place in the bone marrow and is usually diagnosed in early adulthood. A common presentation of FA is acute myeloid leukemia followed by subsequent bone marrow failure. Currently, sixteen FANC genes are associated with mutations in patients (FANCA-FANCQ). The gene products collaborate in a pathway of DNA repair of inter-stand crosslinks (ICLs) that can arise from exposure to chemicals such as cisplatin and aldehydes[130]. Because these patients cannot repair ICLs efficiently delivering chemotherapies and radiation

treatment to treat their leukemia's (which arise in a majority of patients with this disease) can be fatal.

On a mechanistic level the FA complex repairs ICLs by covalently linking the two strands of the DNA double helix and inhibiting replication and transcription until the site is repaired. The FANCM-FAAP24-MHF1-MHF2 sub-complex anchors the FA complex and initially recognized the ICL[131]. The key player in detecting ICLs to be repaired is FANCM. FANCM-MHF1-MHF2-FAAP24 is a subcomplex of the overall FA complex that initially localizes to ICLs when the FA complex is active in S-phase[132]. A crystal structure of FANCM in complex with MHF1/MHF2 reveals FANCM interactions with the whole MHF1/MHF2 (CENP-S/CENP-X) tetramer and that the MHF1/MHF2 complex is necessary for FANCM's DNA binding ability[129]. Specifically, MHF1/MHF2 interact with FANCM$^{661-800}$ that contains three helices and an irregular coil. These elements wrap around both MHF proteins to create a buried surface that greatly increases the DNA binding affinity of full-length FANCM, and also stimulate its DNA branch migration activity. The authors claim the MHF1/MHF2 complex localizes FANCM to centromeres but the data and IF images, in this authors opinion are not convincing, and more work needs to be done to show this[129]. As of this writing no other FA complex members have been found to localize to centromeres but it is highly possible these FA members could be integral in centromere integrity and repair throughout the cell cycle. Whether or not there are differing pools of MHF1/MHF2 with the FA complex and at the centromeres or if these pools dynamically intermix has also not been elucidated.

CENP-T/W

CENP-T is a component of the NAC[23]. Like CENP S and X, it too has a histone fold domain in its C-terminus with a long N-terminal tail that reaches out to the outer kinetochore[24]. CENP-W is much smaller in size and is entirely represented by a histone fold domain. The crystal structure of the T/W complex revealed they interact with one another like canonical histones in a dimer formation[26] which looks strikingly like the Histone H2A/H2B dimer. CENP T/W and S/X together form a T/W/S/X heterotetramer that has DNA binding activity. Mutations in either the DNA binding domains of the T/W or S/X sub-complexes abolish their ability to localize to centromeres suggesting the formation of these "nucleosome-like" particles is happening in cells and is necessary for proper function. Although DNA binding activity for the T/W/S/X complex was shown *in vitro*, it is this authors opinion that the data was not convincing as to whether this complex was wrapping DNA like canonical histones or just "globbing" onto the DNA due to electrostatic interactions with no higher order structure like that of canonical nucleosomes[26]. Furthermore, how and where the T/W/S/X complex binds to DNA in the context of nucleosomes that are already at the chromatin was also not explored and would be a good avenue to go down for anyone trying to pick up this project.

HJURP & DNA damage

Ataxia telangiectasia mutated (ATM) is a serine/threonine protein kinase that is recruited and activated by DNA double-strand breaks. It phosphorylates several key proteins that initiate activation of the DNA damage checkpoint, leading to cell cycle arrest, DNA repair or apoptosis. Several of these targets including p53 and CHK2

are tumor suppressors. Identification of the ATM gene by Shiloh and colleagues[133] revealed a very large gene that coded for a protein containing a phosphoinositide-3 (PI3) kinase-like sequence in its carboxyl terminus. Initially it was found that upon exposure of cells to ionizing radiation, ATM phosphorylation activity increased[134]. A key experiment in ATM null mice beautifully showed lack of p53 induction upon exposure to ionizing radation[135]. In the subsequent years ATM has been found to have a variety of phosphorylation targets and is required for DNA damage repair of double-stranded breaks (DSBs) through homologous recombination (HR).

Homologous recombination is a type of genetic recombination in which nucleotide sequences are exchanged between two similar or identical molecules of DNA. It is most widely used by cells to accurately repair harmful breaks that occur on both strands of DNA, known as double-strand breaks. A key protein that is actually involved in the physical repair of DSBs through HR is Rad51. Rad51 plays a major role in homologous recombination of DNA during double strand break repair. In this process, an ATP dependent DNA strand exchange takes place in which a template strand invades base-paired strands of homologous DNA molecules. Rad51 is involved in the search for homology and strand pairing stages of the process[136,137].

As of this writing there has only been a single paper that has implicated HJURP as a component of DNA repair through HR. The researchers showed HJURP was highly upregulated upon IR and colocalized with other DNA damage proteins at sites of DSBs. Furthermore, they showed HJURP bound holliday-junction DNA *in vitro* that supposedly mimics DNA strand invasion structures that occur during HR. Lastly, they showed HJURP localization and upregulation upon DNA damage was ATM dependent[128]. No

further work has been done on HJURP and its involvement in the DNA damage pathway. What I will present here is preliminary data that tried to recapitulate the findings of the first paper through cells harboring a locus that could be cut to induce a DSB and assayed for its repair by a GFP reporter gene that could only be expressed upon DSB repair[138]. The results were not easy to replicate and I will discuss why I think that is in the results and discussion section as well as future directions for the project.

Results:

S/X

The stumbling block for this project stopped at proper purification of the T/W and S/X complexes. Herein, I will discuss where I think I went wrong and what can be done to rectify the issues if someone were to try to get the reagents in hand. I began trying to purify CENP-S/X in a pst44 bicystronic vector with an MBP tag on CENP-X and a Streptactin tag on CENP-S. CENP-S had a TEV cleavage site while CENP-X had a Factor Xa cleavage site. Purification of the complex was done under 25mM Tris PH=7.5, 10% glycerol, 1X benzamidine, 1X LPC, 1mM PMSF, 250mM NaCl, 20mM $MgCl_2$, and 5mM BME. Purification methodology for the S/X complex is the same for that of Mis18 and can be found in the experimental methods in this section. Off of 1L of bacterial culture induced at 0.6 OD with 0.2mM IPTG for 18 hours at 18^{o}C, the purification was successful. In this gel, a one-step MBP purification was taken (Figure 4-1). Subsequent two-step purifications were done but as can be seen from gel 2 (Figure 4-2), the loss in the tandem affinity purification was about 5-fold just based on coomassie staining. Given that CENP-S and CENP-X have such a high affinity for one another testing the complex for the proportion of the tetramer after one-step purification should be done. Maybe one step is all you need and the complex is homogenous for the tetramer.

Subsequent size exclusion chromatography was performed on a one-step purification of CENP-S/X off amylose beads to see if the complex was running as a tetramer off the single affinity step. Unfortunately, the complex ran near the void (F10 is the void) of the S200 column, which suggested to us the bulky tags would need to be cleaved off to properly characterize the protein we had in hand (Figure 4-3). From this

analysis though the good news was that the complex was very homogenous with both CENP-S and CENP-X running together with one another, suggesting whatever we had in hand was a pure complex. Whether or not it was a tetramer was still not known.

Because the CENP-S/X complex was running near the void with the bulky MBP tag, we tried to cleave off the MBP tag from CENP-X that harbored the Factor-Xa cleavage site. From Figure 4-1. 125ul of elution F2 was taken. F2 was calculated to have 0.8mg/ml of protein as the concentration. 10 ul of protein was used in each reaction with an increasing amount of Factor Xa (Sigma) from 0.1, 0.2, 0.3, 0.4, 0.5, and 0.6ul. The base buffer the reaction was done in was 25mM Tris PH=7.5, 10% glycerol, 100mM NaCl, 20mM $MgCl_2$, 50mM Potassium Acetate, 1mM $CaCl_2$, and 5mM BME. Total reaction volumes were kept at 20ul and done overnight for 12 hours at 4^0C. It should be noted that the initial purifications of CENP-S/X had benzamidine in the buffer, which I have learned inhibits Factor Xa cleavage efficiency. Leaving this out of future purifications would be wise. As can be seen by the gel, the differing amounts of Factor Xa did not actually increase the cleavage efficiency, which might have been due in part to the benzamidine in the buffer. Also to note though, Sigma claims 1ul of Factor Xa can cleave 50ug of protein so the reaction may have been saturated for FXa under the current conditions and was not going to get any better. Unfortunately, cleavage did work, but CENP-X at least on the gel, could not be visualized. CENP-X is a small protein (about 7.8kDa) and the gel shown is a 15% SDS-PAGE gel. I have had poor resolution on gels of this percentage near the bottom. I would recommend in the future visualizing CENP-X, trying a gradient gel from 4-20%. Once cleaved, CENP-X might also have fallen out of solution (Figure 4-4). Factor Xa works best under salt conditions between 50-100mM,

but I purified the proteins in 250mM NaCl. The crystallization of CENP-X was done under 200mM NaCl[26] which might give us a hint that this protein requires higher salt to stay soluble. Tev-protease is said to work under higher salt conditions and may be a cleavage site that could be added in the future instead of Factor Xa.

T/W

The CENP-T/W complex was purified separately from S/X in a pst44 bicystronic vector. CENP-W was tagged with an N-terminal MBP tag that harbored a Factor Xa cleavage site while CENP-T^{HFD} was tag with a 6X-histidine tag with a Tev cleavage site. One-step affinity purifications using either the histidine tag or MBP tag were done on 1 liter of bacterial cultures inoculated at an OD of 0.6 with 0.2mM IPTG for 18 hours at 18^{O}C. Pulldowns were done with a base buffer of 25mM Tris PH=7.5, 250mM NaCl, 20mM $MgCl_2$, 0.1% NP-40, 5mM BME, and 10% glycerol with either amylose or imidazole added to elute the proteins off amylose or nickel beads respectively. Both coomassie gels shown had 6 fractions eluted, 250ul for each fraction. The gel shown has 10ul of each fraction was loaded (Figure 4-5). As can be seen the single 6X-his purification was much more efficient than the amylose pulldown. Bradford assay revealed the peak fraction (F2 for both) of the 6X-his purification contained 2mg/ml of protein while the amylose peak fraction contained 1mg/ml of protein (Figure 4-5).

Size exclusion chromatography was performed off F2 from the nickel purification (Figure 4-5, gel on right) over an S200 column. The column buffer was the same as the purification buffer but the glycerol was reduced to 5%. A total of 24 fractions were collected off of SEC at 0.5ml per fraction with the column running at 0.3ml/min.

Unfortunately, without the MBP tag cleaved off CENP-W, the T/W complex ran near the void (F9) at F11. Fortunately, the complex, even off of one-step affinity purification, ran as a homogenous population (Figure 4-6). As of this time no cleavage of MBP from CENP-W has been performed on the complex but would be the next step in making the CENP-T/W complex that matches the crystallized version[26].

HJURP & DNA damage

To look into HJURPs potential role in DNA damage repair we employed a system in U2OS cells that could induce a DSB at a single site on a chromosome. Briefly, this locus harbors an Sce1 (a restriction enzyme cleavage site) cut site that is found nowhere else in the human genome. This site contains a GFP gene that is truncated by an Sce1 cleavage site and does not express. Upon transfection of a plasmid harboring the Sce1 coding sequence, this site is cut creating a single DSB. If repair of this site takes place through HR, the GFP gene will be resected and a downstream DNA gene cassette containing the untruncated GFP gene that no longer harbors the Sce1 cut site will repair the truncation. Proper repair will lead to GFP being expressed denoting the locus was repaired through HR. Populations of cells harboring repair and subsequent GFP expression can be assayed and sorted by fluorescence activated cell sorting (flow cytometry or FACS) (Figure 4-7A). To try to recapitulate the results from the initial HJURP DNA damage paper[128] we devised a protocol to test whether HJURP knockdown impaired HR repair in this system.

From here on out for this section the materials and methods will be discussed in conjunction with the results so that I can go over the potential methodological pitfalls that

I believe held the project up. To start the U2OS cells were plated in a 6 well plate format with 100,000 cells in each well in DMEM (Qiagen) media with FBS and antibiotics. Culturing of the cells prior to this had the cells kept in the same media but with 100ug/ml Puromycin added to select for the DNA damage cassette so it would not be lost in subsequent passaging events. The following day, HJURP 5' siRNA was transfected into the cells with lipofectamine RNAiMax. The methods for this are as follows; Wells containing the cells (should be at about 30% confluency the next day) were washed in 1X PBS and 2ml of serum free OPTIMEM (Qiagen) was added to the wells. In separate sterile tubes, each tube containing 500ul of OPTIMEM, for each well one tube had the siRNA of choice aliquoted in. The siRNAs used were at a concentration of 20mM and were at a final concentration in the wells (which had a final volume of 3ml) of 20uM. In the adjacent tubes 6ul of RNAiMax was aliquoted in. The tubes containing the siRNA and the RNAiMax were combined (NEVER DO THIS IN BATCH, keep tubes at no more than 1ml) and let sit in the hood for 15 minutes. The mixture was then sprinkled on the cells with intermittent mixing of the media. Knockdown of the target gene was allowed to take place for 24 hours. The next day, a plasmid harboring a constitutively expressed coding sequence of Sce1 (pIN15) was added to the cells. The methods for this are as follows; Cells were washed with 1X PBS and 2ml of serum free OPTIMEM was added into each well. In two tubes, 250ul of OPTIMEM was added to each tube for each condition (two tubes per condition). In one of the tubes 2.5ul of lipfectamin-2000 was added while to the other tube 1ug of the Sce1 containing plasmid was added. The tubes were combined under the fume hood and let incubate for 15 minutes. Then, just like the RNAiMax, the mixture was sprinkled onto cells with intermittent mixing. The cells were

allowed to let sit in the incubator for 48 hours. Upon the day of analysis, cells were washed with 1X PBS, and then sloughed off with 1XPBS +3mM EDTA. They were resuspended with an additional 750ul of phenol red free DMEM (Qiagen), taken up for FACS analysis, and gated for GFP positive cells to denote repair of the locus via HR (Figure 4-7B). GAPDH and Rad51 siRNAs were used as negative and positive controls respectively while HJURP was knocked down with two different siRNAs, FLEG and HJURP 5'.

An example of how cells are gated can be seen in Figure 4-8A-C. Here the GAPDH cell population is gated initially for the single cell population (Figure 4-8A) then for cells fluorescing at 530nm and 580nm (Figure 4-8B-C, denoting GFP expression and locus repair). For this step I would recommend getting training on FACS equipment as well as on the software to be able to make the graphs and do the analysis for yourself, practice often and get comfortable with it. For the initial experiment, 5.4% of cells with GAPDH siRNA repaired the locus and fluoresced GFP. For Rad51, almost none of the cells were able to repair the locus, which is expected. Meanwhile, the two siRNAs against HJURP gave an intermediate phenotype, which suggests HJURP is involved in HR, but is not a nodal element like Rad51 (Figure 4-8D). What was not done in this experiment or the subsequent experiments was a western blot to assess knockdown efficiency of the target proteins. This NEEDS to be done in the future so that the results can be interpreted properly.

Next we wanted to see if we could rescue the DNA lack of repair phenotype in the HJURP knockdown conditions by constitutively transfecting in HJURP constructs resistant to the siRNA. Different fragments of HJURP were made by gateway cloning

that harbored a C-terminal Dsred tag (Figure 4-9). These constructs are IN 47-51 respectively, and can be found in the plasmid database. The destination plasmid harboring the DsRed tag is PDF408 which is a lentiviral vector, and these constructs were supposed to be used to be stably integrated into the U2OS cells so that no transfection of a rescue plasmid would be needed after siRNA knockdown. This was never done, and at first we simply did transient transfection of the FL-HJURP DsRed plasmid to see if we could rescue the HJURP siRNA lack of DNA repair phenotype. The same methodology was used as above (Figure 4-8B) but on top of the Sce1 transfection, cells were also transfected with vector DNA (pUC19) or IN47 (FL-HJURP harboring a C-terminal DsRed tag) at 0.5ug. An siRNA against the 3' end of HJURP was also used. When compared to Figure 4-8. A smaller population of cells were fluorescing GFP (denoting repair) upon GAPDH siRNA induction (5.39% to 3.0%). Furthermore, When the Sce1 plasmid alone was transfected into cells without siRNA treatment, 8.4% of cells were GFP positive. This suggested to us that even though GAPDH is supposedly a "house-keeping" gene the siRNA treatment itself might be making the cells "sick" and the decreased repair phenotype we were seeing was due to the transfection reagents and not the gene knockdown itself. The HJURP 5' siRNA again gave a similar phenotype as compare with Figure 4-8 but this was not the case for HJURP 3' siRNA. Again though, western blots need to be done for each experiment as the 3' siRNA might not have efficiently knocked down HJURP which could be the cause of the discrepancy. Unfortunately, transient transfection of FL-HJURP harboring a DsRed tag (IN47) did not rescue the lack of DNA repair phenotype in the HJURP knockdown conditions (Figure 4-10). That being said, this plasmid was never made for transient transfection instead it was

supposed to be used to make stable cell lines. Future experiments should be making stable cell lines harboring these constructs and assay whether they are expressing properly.

In summation the results for HJURP and its potential involvement in DNA damage were inconclusive for this assay. I believe there are a few methodological reasons for this. For one, cells were treated with siRNA for only 24 hours before the media was taken off to make way for DNA transfection. In most cases in our lab, siRNA containing media was left on cells were a minimum of 48 hours. The subtle phenotype we are seeing could be due to insufficient knockdown of the target genes. Moreover, treating cells back to back with siRNA then DNA could be making the cells very sick and unable to properly repair themselves due to cellular senescence etc. Anecdotally, Lipofectamine-2000 in high amounts is highly toxic to cells and can kill them if too much is put on at one time. Suffice to say the treatment protocol (Figure 4-7B) could be too harsh for the cells to properly recover from. I would recommend making stable cell lines harboring HJURP rescue constructs along with Sce1 stably integrated. These constructs should be doxycycline inducible so they can be turned on when needed. This way, siRNA will be the only treatment given to cells, this will simplify the protocol greatly and solve for transfection efficiency which can be finicky with lipofectamine-2000, as if the cells are too confluent the transfection efficiency is almost null.

Discussion:

T/W/S/X

The CENP-T/W/S/X complex project represented my first foray into recombinant protein purification and reagent making which was a huge stumbling block, but a wonderful learning experience for me. That being said, for this project to get off the ground, the whole T/W/S/X complex must be reconstituted *in vitro* to move forward. I hope that the initial results I have obtained help whoever works on this project next for potential pitfalls. I would recommend developing assays and questions before starting this project that are interesting and novel and that could be turned into a unique publication in the field. I mistakenly tried to chase the coat tails of the group who had crystallized the complex and was always playing catch up instead of trying to blaze a new trail for myself.

An interesting question to start with is to assay for whether the T/W/S/X components have unique chaperones that aid in their deposition at centromeric chromatin. Although these structures are forming a "nucleosome-like" complex and they look like canonical histones themselves, a chaperone mechanism has not been explored for this unique complex. We have LAP-CENP-S, X, T, and W HeLa cell lines that could absolutely be of use in answering this question.

Furthermore, how the T/W/S/X complex integrates into chromatin in the context of nucleosomes was not explored in the T/W/S/X crystallization manuscript. Although the complex can bind DNA, where is it binding when in the presence of nucleosomes? Does it bind to the linker region between nucleosomes? Does it help space nucleosomes evenly? These questions would be best answered using an *in vitro* nucleosome assembly

approach, but once the system is set up information gleaned from this process could be completely new and relevant to the field.

Lastly, answering if T/W/S/X heterotetramer exists in cells and the biological relevance of this in the context of centromere location and propagation would be a fresh path to go down. Although the complex exists as a heterotetramer *in vitro* whether this same complex exists at actual centromeres has not been established.

HJURP & DNA Damage

Although the initial results for HJURP hint at its role in DNA damage, the experimental protocol needs to be simplified with less steps for error which was discussed above. Going forward with the project one would need to carve out a unique experimental approach to take it further than just broadly finding that HJURP is somehow involved in the DNA damage pathway. More specifically, what is HJURP mechanistically doing to repair the DNA? Moreover, the Kato paper was published before HJURP was known to be the CENP-A specific chaperone and although the results were promising, knocking down HJURP for even a short period of time is going to cause centromere instability which could actually be the causative agent of the phenotype Kato and our work are observing.

To start, anyone picking up this project should try to elucidate where exactly HJURP fits into the DNA damage pathway. Based on the prior research it is ATM dependent and localizes to sites of DNA damage, doing ChIP for HJURP on the Sce1 locus after DNA damage and subsequent mass spectrometry could flesh out if HJURP is binding to the site of DNA damage and what its binding partners are. This analysis would

show what/if HJURP is in complex with any DNA damage pathway proteins and would also hint at potential targets that are upstream and downstream of HJURP in the pathway. Knockdown of these targets and ChIP for HJURP at the Sce1 locus would elucidate whether or not these components are necessary for HJURP localization to sites of DNA damage and one would be able to start putting HJURP in the DNA repair pathway.

Previous data has shown HJURP deposits CENP-A in early G1-phase just after mitosis. Unpublished work from Ewelina Zasadzinska in our lab shows HJURP is actually localizing to centromeres, albeit transiently in S-phase, which is when a majority of DNA repair complexes are active. If HJURP does indeed have a dual role in CENP-A deposition and DNA repair the next logical question would be; Are these functions of HJURP temporally regulated by the cell cycle? Tandem affinity purifications of our LAP-tagged HJURP cell line and mass-spec analysis of its binding partners in cells blocked in G1 and S-phase can be done to see if the binding partners of HJURP change depending on cell cycle.

Lastly, now that it is known HJURP is the CENP-A specific chaperone, does it need to be and is it bound to CENP-A when it helps repair DNA? Creating stable version of the U2OS DNA damage cell line harboring the TLTY box mutant version of HJURP that cannot bind CENP-A[73] and inducing a DSB in an HJURP knockdown background could assess whether HJURP needs CENP-A bound to repair the DNA. If so, then why? Does CENP-A transiently get deposited at sites of DNA damage? Does it help with DNA repair? What is the function of CENP-A deposition at these sites? The sky is the limit for this experiment if it works out and at this time no one is working on this question.

Figure Legends:

Figure 4-1: One step MBP purification of CENP-S/X: 1L purification with a coomassie gel. Lane 1 (I) is the input, lane 2 is the insoluble pellet, S=supernatant or what was being put over the amylose beads. F1-F8 are the elution fractions with BSA standards run to try to quantify the amount of protein present. CENP-S and CENP-X are pointed out on the gel. Elution fractions are 10ul/250ul per fraction.

Figure 4-2: Second step affinity purification of CENP-S/X over streptactin beads: 1L purification with a coomassie gel. Lane 1 (I) is the input, lane 2 is the insoluble pellet, S=supernatant or what was being put over the amylose beads. F1-F8 are the elution fractions with BSA standards run to try to quantify the amount of protein present. CENP-S and CENP-X are pointed out on the gel. As can be seen there is a sizeable decrease and loss of protein from the first affinity step, roughly speaking 65%. Elution fractions loaded are 10ul/250ul per fraction.

Figure 4-3: CENP-S/X run as a single complex over SEC: One step affinity purification of CENP-S/X pooled fractions from Figure 4-1. This was done over amylose beads pulling on the CENP-X. Fractions F1-F8 were pooled and run over an S200 column. I=Input, while F6-F15 represent fractions taken off SEC. F10=void. 500ul fractions were taken at each fraction and the column buffer was 25mM tris ph=7.5, 5% glycerol, 5mM BME, 250mM NaCl, 20mM $MgCl_2$, and 0.1% NP-40. Column was run at 0.3ml/min for a total volume of 25ml.

Figure 4-4: Factor Xa cleavage of MBP tag off CENP-X: From left to right input (lane 2), 3-8 10ul of MBP-CENP-X/S mixed with 0.1-0.6ul of Factor Xa (Sigma) o/n at 4^{o}C for

12 hours. Cleaved MBP and CENP-S can be visualized by the coomassie stain while CENP-X could not be made out.

Figure 4-5: CEP-T/W one-step affinity purifications: (left panel) affinity purification of CENP-T/W using one step amylose purification to pull on MBP-CENP-W. I=Input, S=supernatant, F=Flow, 1-6=elutions fractions. BSA standards are 0.5,1,2,3,4ug. (Right panel) affinity purification of CENP-T/W using one step nickel purification to pull on His-CENP-T^{HFD}. I=Input, S=supernatant, F=Flow, 1-6=elutions fractions. BSA standards are ,1, 2, 4ug

Figure 4-6: CENP-T/W run as a single complex over SEC: One step affinity purification of CENP-T/W pooled fractions from Figure 4-5 (right panel). This was done over nickel beads pulling on the CENP-T^{HFD}. Fractions F1-F6 were pooled and run over an S200 column. I=Input, while F6-F18 represent fractions taken off SEC. F9=void. 500ul fractions were taken at each fraction and the column buffer was 25mM tris ph=7.5, 5% glycerol, 5mM BME, 250mM NaCl, 20mM $MgCl_2$, and 0.1% NP-40. Column was run at 0.3ml/min for a total volume of 25ml.

Figure 4-7: Schematic for assaying HJURPs involvement in DNA damage repair: A. U2OS cell system-harboring locus with truncated GFP gene and Sce1 cut site with a downstream untrunctated GFP gene to be used for repair via HR. B.Protocol for knocking down HJURP to assess its potential role HR.

Figure 4-8: FACS Analysis of DR-GFP U2OS cells transfected with different siRNAs and SceI to induce a DSB: (A-C) An example of how cell populations are analyzed post FACS analysis. These cells were treated with GAPDH siRNA and gated first for single cells (A). Then the single cells were gated on their GFP fluorescence,

which emits at 530nm(B). The GFP positive population was hard to gate on so it was distinguished by gating the emission of the cells at 530nm versus 580nm (C). This shows the cells have a basal level of autoflourescence at these wavelengths and makes the GFP positive population more easily distinguishable at 580nm. D. Bar graph showing percentage of cells that were GFP positive (Y-axis) with indicated siRNA treatment (X-axis).

Figure 4-9: *Fragments of HJURP harboring different domains and a C-terminal DsRed tag.*

Figure 4-10: FACS Analysis of DR-GFP U2OS cells transfected with different siRNAs, rescue plasmids, SceI to induce a DSB: Bar graph showing percentage of cells that were GFP positive (Y-axis) with indicated siRNA treatment and subseqnent DNA transfection treatment, Vector=pUC19 dna and IN47=FL-HJURP harboring a C-terminal DsRed tag that was used to try to rescue the phenotype. GAPDH, HJ3', HJ5', and Rad51 were the siRNA's used while Sce1 (IN15) was the plasmid used to induce a DSB in the cells (X-axis).

Materials & Methods:

CENP-T/W/S/X

DNA constructs. DNA constructs were created by PCR amplification with a forward primer that added a 6X-His tag flanked with an Nde1 cut site for CENP-T^{HFD} and a reverse primer that was flanked by a Kpn1 cut site and was subsequently cloned into position 2 of a pst44 multicystronic vector. CENP-S was cloned the same way except the forward primer had a 1X Streptactin Tev tag added on. CENP-W and CENP-X were

taken from the human orfeome (v5.1) and cloned into a custom MBP vector (IN12-3) using LR clonase. They were both subsequently PCRed as MBP fusion constructs with a forward primer that added an Nde1 cut site and a reverse primer with a Bsrg1 overhang. From here the PCR products were cut with the necessary restriction enzymes and ligated into apst50trc1 cut with the same enzymes. They were subsequently cut out from these entry vectors and cloned into position 1 of a pst44 harboring Strep-CENP-S (MBP-CENP-X) or His-CENP-T^{HFD} (MBP-CENP-W).

Protein expression and purification. Recombinant proteins were expressed in the Rosetta (DE3) pLysS bacterial strain. Bacteria were grown in LB media to an OD of 0.6 at 37^{o}C and induced at 18^{o}C with 0.1mM IPTG for 16 hr. His-tagged CENP-T/W fragments were purified on Ni-NTA (Qiagen) while CENP-S/X were purified on amylose beads (Qiagen). Bacteria were lysed using a steel Wheaton-dounce in buffer contained 50mM Tris-HCl pH 7.5, 250mM NaCl, 20mM MgCl2, 0.5mM CaCl2, 10% glycerol, 0.1% NP-40, 5mM BME, LPC, 1mM PMSF, 20mM imidazole, and 0.15mg/ml RNase-A. DnaseI was added and lysates were centrifuged at 22,000x*g* for 20 min. Supernatants were collected and pellets were re-extracted with a second round of lysis. Cleared lysates were incubated with Ni-NTA agarose or amylose (Qiagen) for 1 hour and washed twice in lysis buffer with 40mM imidazole (only for T/W). Proteins were eluted in lysis buffer containing 250mM imidazole or 10mM maltose and no protease inhibitors.

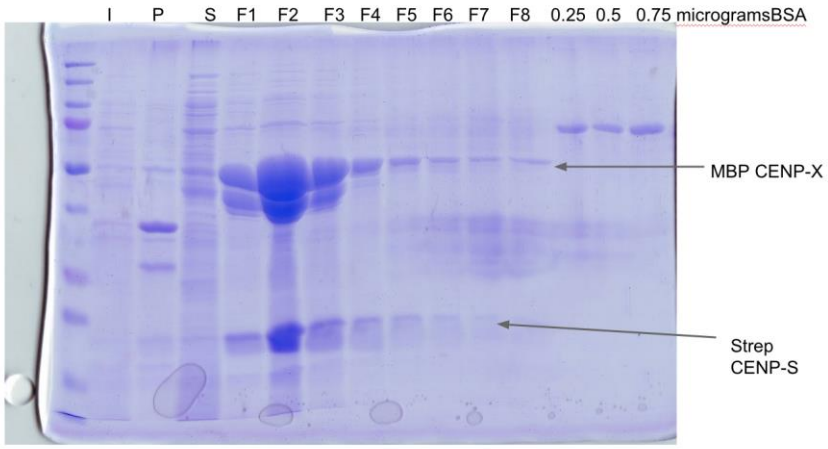

I P S F1 F2 F3 F4 F5 F6 F7 F8 0.25 0.5 0.75 micrograms BSA
MBP CENP-X
Strep
CENP-S

Figure 4-1

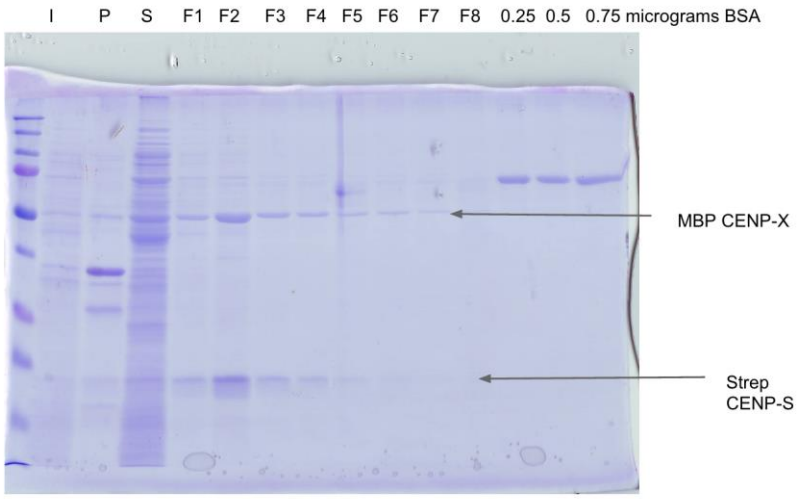

I P S F1 F2 F3 F4 F5 F6 F7 F8 0.25 0.5 0.75 micrograms BSA
MBP CENP-X
Strep
CENP-S

Figure 4-2

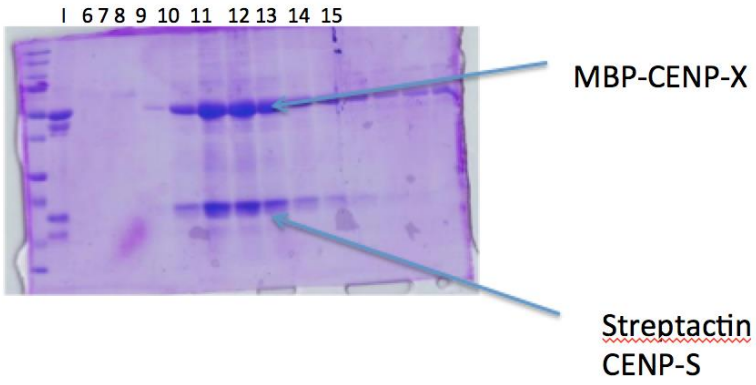


Figure 4-3

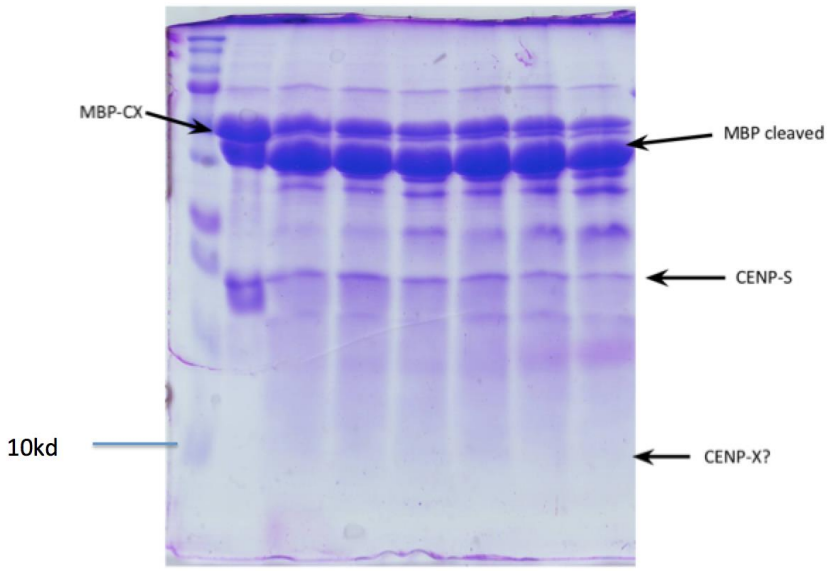


Figure 4-4

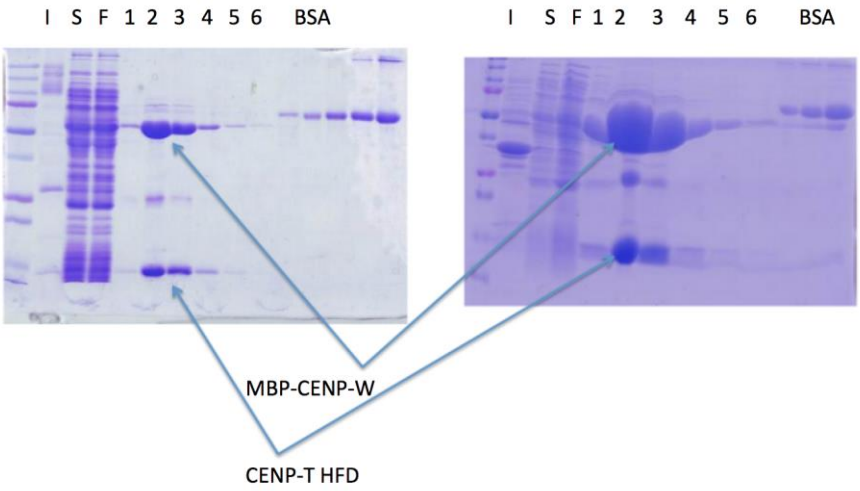


Figure 4-5

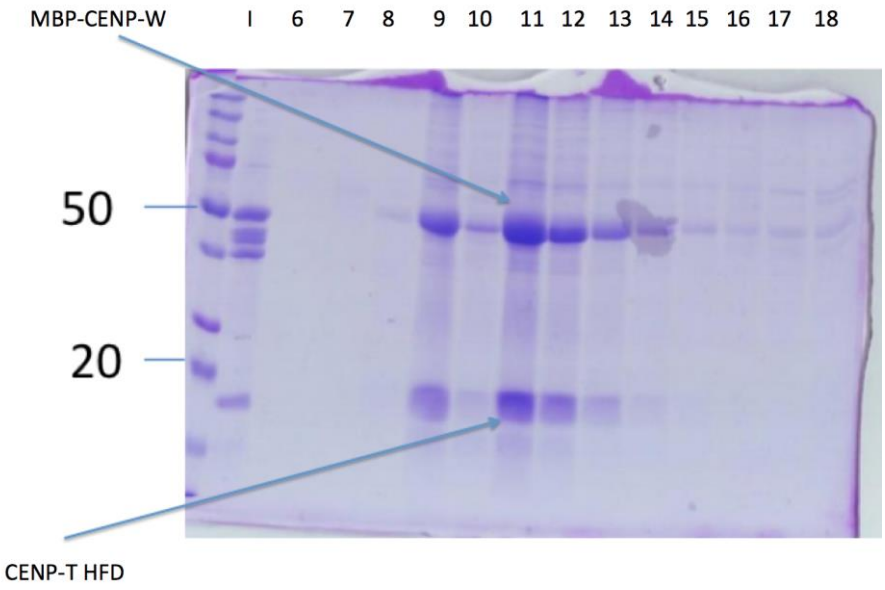
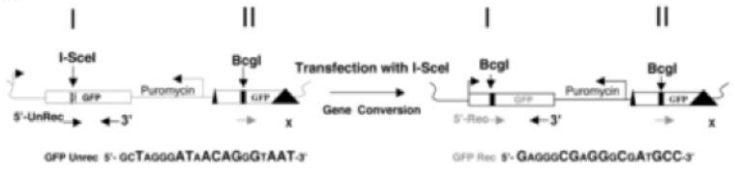


Figure 4-6

A.


I
II
I
II
I-SceI
BcgI
Transfection with I-SceI
BcgI
BcgI
Puromycin
Puromycin
GFP
GFP
GFP
5'-UnRec
3'
Gene Conversion
5'-Rec
3'
X
X
GFP Unrec 5'- GCTAGGGATAACAGGGTAAT-3'
GFP Rec 5'- GAGGGCGAGGGCGATGCC-3'

B.

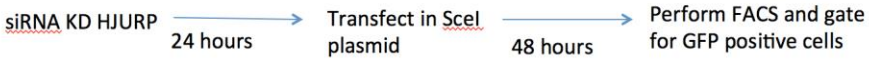

siRNA KD HJURP
24 hours
Transfect in SceI plasmid
48 hours
Perform FACS and gate for GFP positive cells

Figure 4-7

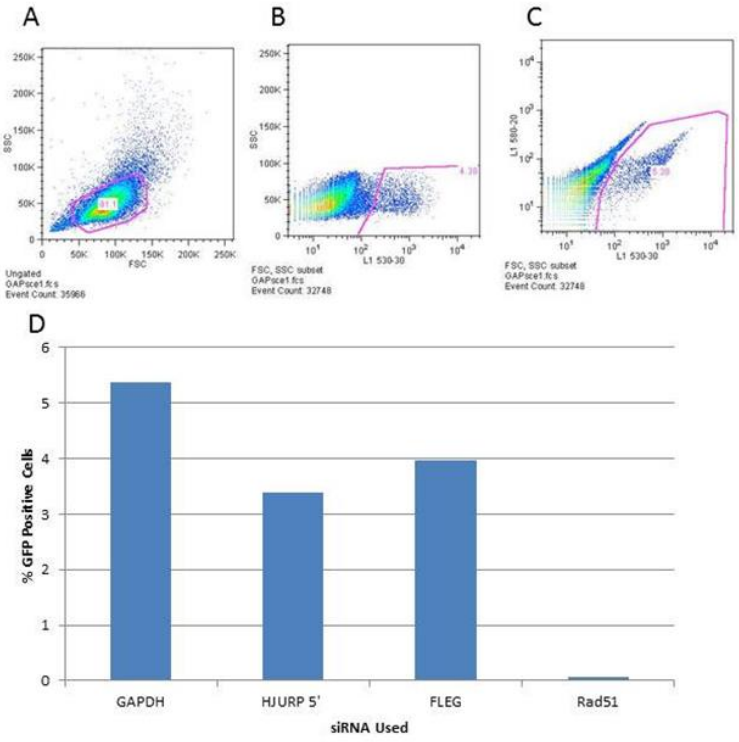

A
B
C
D
SSC
FSC
Ungated
GAPsce1.fcs
Event Count: 35966
L1 530-30
FSC, SSC subset
GAPsce1.fcs
Event Count: 32748
L1 580-20
% GFP Positive Cells
GAPDH
HJURP 5'
FLEG
Rad51
siRNA Used

Figure 4-8

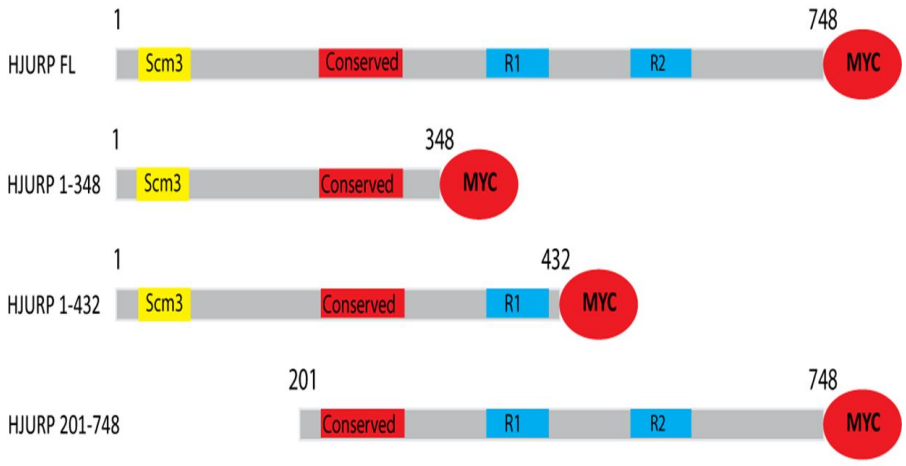

1
748
HJURP FL
Scm3
Conserved
R1
R2
MYC
1
348
HJURP 1-348
Scm3
Conserved
MYC
1
432
HJURP 1-432
Scm3
Conserved
R1
MYC
201
748
HJURP 201-748
Conserved
R1
R2
MYC

Figure 4-9

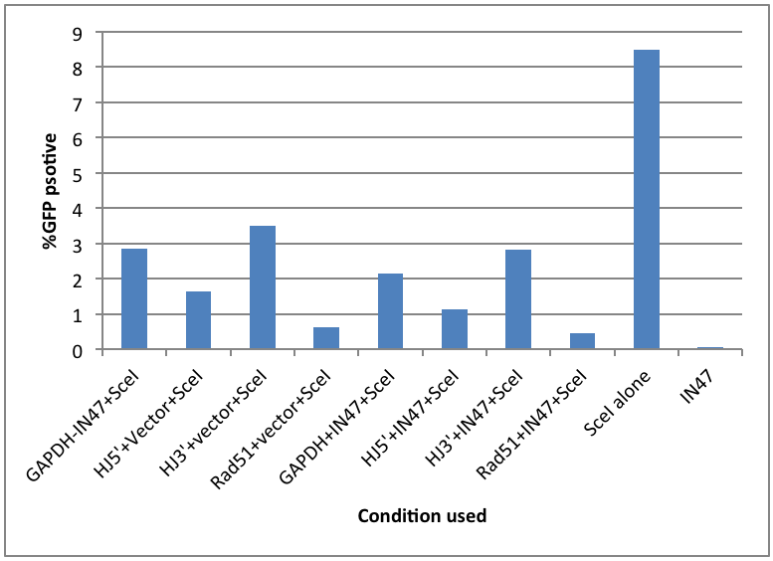

%GFP psotive
9
8
7
6
5
4
3
2
1
0
GAPDH-IN47+SceI
HJ5'+Vector+SceI
HJ3'+vector+SceI
Rad51+vector+SceI
GAPDH+IN47+SceI
HJ5'+IN47+SceI
HJ3'+IN47+SceI
Rad51+IN47+SceI
SceI alone
IN47
Condition used

Figure 4-10

References

1. Shang, W. *et al.* Chickens possess centromeres with both extended tandem repeats and short non-tandem-repetitive sequences. 1219–1228 (2010). doi:10.1101/gr.106245.110.20
2. Cleveland, D. W., Mao, Y. & Sullivan, K. F. Centromeres and kinetochores: from epigenetics to mitotic checkpoint signaling. *Cell* **112,** 407–21 (2003).
3. Sun, X., Wahlstrom, J. & Karpen, G. Molecular structure of a functional Drosophila centromere. *Cell* **91,** 1007–1019 (1997).
4. Zheng, D. *et al.* NIH Public Access. *Cell* **137,** 1–16 (2011).
5. Marshall, O. J., Chueh, A. C., Wong, L. H. & Choo, K. H. A. Neocentromeres : New Insights into Centromere Structure , Disease Development , and Karyotype Evolution. *J. Hum. Genet.* 261–282 (2008). doi:10.1016/j.ajhg.2007.11.009.
6. Bailey, A. O. *et al.* Posttranslational modification of CENP-A influences the conformation of centromeric chromatin. *Proc. Natl. Acad. Sci. U. S. A.* **110,** 11827–32 (2013).
7. Ribeiro, S. A. *et al.* A super-resolution map of the vertebrate kinetochore. *Proc. Natl. Acad. Sci. U. S. A.* **107,** 10484–10489 (2010).
8. Beth A Sullivan1,2 and Gary H Karpen1. Centromeric chromatin exhibits a histone modification pattern that is distinct from both euchromatin and heterochromatin. *Nat Struct Mol Biol.* **29,** 997–1003 (2005).
9. Saffery, R. *et al.* Human centromeres and neocentromeres show identical distribution patterns of >20 functionally important kinetochore-associated proteins. *Hum. Mol. Genet.* **9,** 175–185 (2000).
10. Ogiyama, Y., Ohno, Y., Kubota, Y. & Ishii, K. Epigenetically induced paucity of histone H2A.Z stabilizes fission-yeast ectopic centromeres. *Nat Struct Mol Biol* **20,** 1397–1406 (2013).
11. Williams, B. C., Murphy, T. D., Goldberg, M. L. & Karpen, G. H. Neocentromere activity of structurally acentric mini-chromosomes in Drosophila. *Nat Genet* **18,** 30–38 (1998).
12. Barnhart, M. C. *et al.* HJURP is a CENP-A chromatin assembly factor sufficient to form a functional de novo kinetochore. *J. Cell Biol.* **194,** 229–43 (2011).
13. Chen, C.-C. *et al.* CAL1 is the Drosophila CENP-A assembly factor. *J. Cell Biol.* **204 ,** 313–329 (2014).
14. Gascoigne, K. E. *et al.* Induced ectopic kinetochore assembly bypasses the requirement for CENP-A nucleosomes. *Cell* **145,** 410–22 (2011).
15. Mendiburo, M. J., Padeken, J., Fülöp, S., Schepers, A. & Heun, P. Drosophila CENH3 is sufficient for centromere formation. *Science* **334,** 686–90 (2011).
16. Perpelescu, M. & Fukagawa, T. The ABCs of CENPs. *Chromosoma* **120,** 425–46 (2011).

17. Carroll, C. W., Milks, K. J. & Straight, A. F. Dual recognition of CENP-A nucleosomes is required for centromere assembly. *J. Cell Biol.* **189,** 1143–55 (2010).

18. Carroll, C. W., Silva, M. C. C., Godek, K. M., Jansen, L. E. T. & Straight, A. F. Centromere assembly requires the direct recognition of CENP-A nucleosomes by CENP-N. *Nat. Cell Biol.* **11,** 896–902 (2009).

19. Ando, S., Yang, H., Nozaki, N., Okazaki, T. & Yoda, K. CENP-A, -B, and -C Chromatin Complex That Contains the I-Type α-Satellite Array Constitutes the Prekinetochore in HeLa Cells. *Mol. Cell. Biol.* **22,** 2229–2241 (2002).

20. Kwon, M.-S., Hori, T., Okada, M. & Fukagawa, T. CENP-C Is Involved in Chromosome Segregation, Mitotic Checkpoint Function, and Kinetochore Assembly. *Mol. Biol. Cell* **18,** 2155–2168 (2007).

21. Milks, K. J., Moree, B. & Straight, A. F. Dissection of CENP-C – directed Centromere and Kinetochore Assembly. *Mol. Biol. Cell* **20,** 4246–4255 (2009).

22. Liu, S.-T. *et al.* Human CENP-I specifies localization of CENP-F, MAD1 and MAD2 to kinetochores and is essential for mitosis. *Nat Cell Biol* **5,** 341–345 (2003).

23. Foltz, D. R. *et al.* The human CENP-A centromeric nucleosome-associated complex. *Nat. Cell Biol.* **8,** 458–69 (2006).

24. Hori, T. *et al.* CCAN makes multiple contacts with centromeric DNA to provide distinct pathways to the outer kinetochore. *Cell* **135,** 1039–52 (2008).

25. Amano, M. *et al.* The CENP-S complex is essential for the stable assembly of outer kinetochore structure. *J. Cell Biol.* **186,** 173–82 (2009).

26. Nishino, T. *et al.* CENP-T-W-S-X Forms a Unique Centromeric Chromatin Structure with a Histone-like Fold. *Cell* **148,** 487–501 (2012).

27. Yan, Z. *et al.* A histone-fold complex and FANCM form a conserved DNA-remodeling complex to maintain genome stability. *Mol. Cell* **37,** 865–78 (2010).

28. Okada, M., Okawa, K., Isobe, T. & Fukagawa, T. CENP-H – containing Complex Facilitates Centromere Deposition of CENP-A in Cooperation with FACT and CHD1. *Mol. Biol. Cell* **20,** 3986 –3995 (2009).

29. Amaro, A. C. *et al.* Molecular control of kinetochore-microtubule dynamics and chromosome oscillations. *Nat. Cell Biol.* **12,** 319–329 (2010).

30. Matson, D. R., Demirel, P. B., Stukenberg, P. T. & Burke, D. J. A conserved role for COMA / CENP-H / I / N kinetochore proteins in the spindle checkpoint. 542–547 (2012). doi:10.1101/gad.184184.111

31. Richmond, R. K., Sargent, D. F., Richmond, T. J., Luger, K. & Ma, A. W. Crystal structure of the nucleosome ˚ resolution core particle at 2 . 8 A. *Nature* **7,** (1997).

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

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

(2007).

34. Sekulic, N., Bassett, E. a, Rogers, D. J. & Black, B. E. The structure of (CENP-A-H4)(2) reveals physical features that mark centromeres. *Nature* **467,** 347–51 (2010).

35. Hasson, D. *et al.* The octamer is the major form of CENP-A nucleosomes at human centromeres. *Nat. Struct. Mol. Biol.* **20,** 687–695 (2013).

36. Miell, M. D. D. *et al.* CENP-A confers a reduction in height on octameric nucleosomes. *Nat. Struct. Mol. Biol.* **20,** 763–765 (2013).

37. Black, B. E. *et al.* Structural determinants for generating centromeric chromatin. *Nature* **430,** 578–82 (2004).

38. Shelby, R. D., Vafa, O. & Sullivan, K. F. Assembly of CENP-A into Centromeric Chromatin Requires a Cooperative Array of Nucleosomal DNA Contact Sites. *Cell* **136,** 501–513 (1997).

39. Morey, L., Barnes, K., Chen, Y., Fitzgerald-Hayes, M. & Baker, R. E. The Histone Fold Domain of Cse4 Is Sufficient for CEN Targeting and Propagation of Active Centromeres in Budding Yeast. *Eukaryot. Cell* **3,** 1533–1543 (2004).

40. Hu, H. *et al.* Structure of a CENP-A – histone H4 heterodimer in complex with chaperone HJURP. 901–906 (2011). doi:10.1101/gad.2045111.GENES

41. Dalal, Y., Furuyama, T., Vermaak, D. & Henikoff, S. Structure, dynamics, and evolution of centromeric nucleosomes. *Proc. Natl. Acad. Sci. U. S. A.* **104,** 15974–81 (2007).

42. Dimitriadis, E. K., Weber, C., Gill, R. K., Diekmann, S. & Dalal, Y. Tetrameric organization of vertebrate centromeric nucleosomes. *Proc. Natl. Acad. Sci. U. S. A.* **107,** 20317–20322 (2010).

43. Bassett, E. a *et al.* HJURP uses distinct CENP-A surfaces to recognize and to stabilize CENP-A/histone H4 for centromere assembly. *Dev. Cell* **22,** 749–62 (2012).

44. Camahort, R. *et al.* Cse4 is Part of an Octameric Nucleosome in Budding Yeast. *Mol. Cell* **35,** 794–805 (2009).

45. Wu, Q. *et al.* Expression and prognostic significance of centromere protein A in human lung adenocarcinoma. *Lung Cancer* (2012). doi:10.1016/j.lungcan.2012.04.007

46. Shelby, R. D., Monier, K. & Sullivan, K. F. Chromatin Assembly at Kinetochores Is Uncoupled from DNA Replication. *Reactions* **151,** 1113–1118 (2000).

47. Dunleavy, E. M., Almouzni, G. & Karpen, G. H. 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 (2011).

48. Mellone, B. G. *et al.* Assembly of Drosophila Centromeric Chromatin Proteins during Mitosis. *PLoS Genet.* **7,** e1002068 (2011).

49. Schuh, M., Lehner, C. F. & Heidmann, S. Incorporation of Drosophila CID/CENP-A and CENP-C into Centromeres during Early Embryonic Anaphase. *Curr. Biol.*

17, 237–243 (2015).

50. Pearson, C. G. *et al.* Stable Kinetochore-Microtubule Attachment Constrains Centromere Positioning in Metaphase. *Curr. Biol.* **14,** 1962–1967 (2015).
51. Takahashi, K., Takayama, Y., Masuda, F., Kobayashi, Y. & Saitoh, S. Two distinct pathways responsible for the loading of CENP-A to centromeres in the fission yeast cell cycle. *Philos. Trans. R. Soc. B Biol. Sci.* **360,** 595–607 (2005).
52. Takayama, Y. *et al.* Biphasic Incorporation of Centromeric Histone CENP-A in Fission Yeast. *Mol. Biol. Cell* **19,** 682–690 (2008).
53. Lermontova, I. *et al.* Loading of Arabidopsis Centromeric Histone CENH3 Occurs Mainly during G2 and Requires the Presence of the Histone Fold Domain. *Plant Cell* **18,** 2443–2451 (2006).
54. Dornblut, C. *et al.* A CENP-S/X complex assembles at the centromere in S and G2 phases of the human cell cycle. *Open Biol.* **4,** 130229 (2014).
55. Prendergast, L. *et al.* Premitotic Assembly of Human CENPs -T and -W Switches Centromeric Chromatin to a Mitotic State. *PLoS Biol.* **9,** e1001082 (2011).
56. Hellwig, D. *et al.* Dynamics of CENP-N kinetochore binding during the cell cycle. *J. Cell Sci.* **124,** 3871–3883 (2011).
57. Hemmerich, P. *et al.* Dynamics of inner kinetochore assembly and maintenance in living cells. *J. Cell Biol.* **180,** 1101–1114 (2008).
58. Stellfox, M. E., Bailey, A. O. & Foltz, D. R. Putting CENP-A in its place. *Cell. Mol. Life Sci.* **70,** 387–406 (2013).
59. Silva, M. C. C. *et al.* Cdk activity couples epigenetic centromere inheritance to cell cycle progression. *Dev. Cell* **22,** 52–63 (2012).
60. Wang, J. *et al.* Mitotic regulator Mis18β interacts with and specifies the centromeric assembly of molecular chaperone HJURP. *J. Biol. Chem.* (2014). doi:10.1074/jbc.M113.529958
61. McKinley, K. L. & Cheeseman, I. M. Polo-like Kinase 1 Licenses CENP-A Deposition at Centromeres. *Cell* **158,** 397–411 (2014).
62. Yang, C. H., Tomkiel, J., Saitoh, H., Johnson, D. H. & Earnshaw, W. C. Identification of overlapping DNA-binding and centromere-targeting domains in the human kinetochore protein CENP-C. *Mol. Cell. Biol.* **16,** 3576–86 (1996).
63. Saitoh, H. *et al.* CENP-C, an autoantigen in scleroderma, is a component of the human inner kinetochore plate. *Cell* **70,** 115–25 (1992).
64. Wan, X. *et al.* Protein architecture of the human kinetochore microtubule attachment site. *Cell* **137,** 672–84 (2009).
65. Dambacher, S. *et al.* CENP-C facilitates the recruitment of M18BP1 to centromeric chromatin Do not distribute . © 2012 Landes Bioscience . 101–110 (2012).
66. Heun, P. *et al.* Mislocalization of the Drosophila centromere-specific histone CID promotes formation of functional ectopic kinetochores. *Dev. Cell* **10,** 303–15

(2006).

67. Moree, B., Meyer, C. B., Fuller, C. J. & Straight, A. F. CENP-C recruits M18BP1 to centromeres to promote CENP-A chromatin assembly. *J. Cell Biol.* (2011). doi:10.1083/jcb.201106079

68. Erhardt, S. *et al.* Genome-wide analysis reveals a cell cycle–dependent mechanism controlling centromere propagation. *J. Cell Biol.* **183 ,** 805–818 (2008).

69. Kato, T. *et al.* Activation of Holliday junction recognizing protein involved in the chromosomal stability and immortality of cancer cells. *Cancer Res.* **67,** 8544–53 (2007).

70. Dunleavy, E. M. *et al.* HJURP is a cell-cycle-dependent maintenance and deposition factor of CENP-A at centromeres. *Cell* **137,** 485–97 (2009).

71. Foltz, D. R. *et al.* Centromere-specific assembly of CENP-a nucleosomes is mediated by HJURP. *Cell* **137,** 472–84 (2009).

72. Mizuguchi, G., Xiao, H., Wisniewski, J., Smith, M. M. & Wu, C. Nonhistone Scm3 and histones CenH3-H4 assemble the core of centromere-specific nucleosomes. *Cell* **129,** 1153–64 (2007).

73. Shuaib, M., Ouararhni, K., Dimitrov, S. & Hamiche, A. 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,** 1349–54 (2010).

74. Zasadzińska, E., Barnhart-Dailey, M. C., Kuich, P. H. J. L. & Foltz, D. R. Dimerization of the CENP-A assembly factor HJURP is required for centromeric nucleosome deposition. *EMBO J.* **32,** 2113–24 (2013).

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

76. Zhou, Z. *et al.* Structural basis for recognition of centromere histone variant CenH3 by the chaperone Scm3. *Nature* **472,** 234–237 (2011).

77. Camahort, R. *et al.* Scm3 is essential to recruit the histone h3 variant cse4 to centromeres and to maintain a functional kinetochore. *Mol. Cell* **26,** 853–65 (2007).

78. Sanchez-Pulido, L., Pidoux, A. L., Ponting, C. P. & Allshire, R. C. Common ancestry of the CENP-A chaperones Scm3 and HJURP. *Cell* **137,** 1173–4 (2009).

79. Stoler, S. *et al.* Scm3, an essential Saccharomyces cerevisiae centromere protein required for G2/M progression and Cse4 localization. *Proc. Natl. Acad. Sci. U. S. A.* **104,** 10571–6 (2007).

80. Williams, J. S., Hayashi, T., Yanagida, M. & Russell, P. Fission yeast Scm3 mediates stable assembly of Cnp1/CENP-A into centromeric chromatin. *Mol. Cell* **33,** 287–98 (2009).

81. Espelin, C. W., Kaplan, K. B. & Sorger, P. K. Probing the Architecture of a Simple Kinetochore Using DNA–Protein Crosslinking . *J. Cell Biol.* **139,** 1383–1396 (1997).

82. Xiao, H. *et al.* Nonhistone Scm3 binds to AT-rich DNA to organize atypical

centromeric nucleosome of budding yeast. *Mol. Cell* **43,** 369–80 (2011).

83. Hayashi, T. *et al.* Mis16 and Mis18 are required for CENP-A loading and histone deacetylation at centromeres. *Cell* **118,** 715–29 (2004).
84. Fujita, Y. *et al.* Priming of centromere for CENP-A recruitment by human hMis18alpha, hMis18beta, and M18BP1. *Dev. Cell* **12,** 17–30 (2007).
85. Maddox, P. S., Hyndman, F., Monen, J., Oegema, K. & Desai, A. Functional genomics identifies a Myb domain-containing protein family required for assembly of CENP-A chromatin. *J. Cell Biol.* **176,** 757–63 (2007).
86. Kim, I. S. *et al.* Roles of Mis18α in epigenetic regulation of centromeric chromatin and CENP-A loading. *Mol. Cell* **46,** 260–73 (2012).
87. Dahmann, C., Diffley, J. F. X. & Nasmyth, K. A. S-phase-promoting cyclin-dependent kinases prevent re-replication by inhibiting the transition of replication origins to a pre-replicative state. **5,** (1995).
88. Hua, X. H., Yan, H. & Newport, J. A Role for Cdk2 Kinase in Negatively Regulating DNA Replication during S Phase of the Cell Cycle. **137,** 183–192 (1997).
89. Bianco, J. N. *et al.* Analysis of DNA replication profiles in budding yeast and mammalian cells using DNA combing. *Methods* **57,** 149–157 (2012).
90. Truong, L. N. & Wu, X. Prevention of DNA re-replication in eukaryotic cells. 13–22 (2011).
91. Diffley, J. F. X. The chromosome replication cycle. **5,** 869–872 (2002).
92. Wohlschlegel, J. a *et al.* Inhibition of eukaryotic DNA replication by geminin binding to Cdt1. *Science* **290,** 2309–12 (2000).
93. Aladjem, M. I. Replication in context: dynamic regulation of DNA replication patterns in metazoans. *Nat Rev Genet* **8,** 588–600 (2007).
94. RAO, P. N. & JOHNSON, R. T. Mammalian Cell Fusion : Studies on the Regulation of DNA Synthesis and Mitosis. *Nature* **225,** 159–164 (1970).
95. Blow, J. J. & Laskey, R. A. A role for the nuclear envelope in controlling DNA replication within the cell cycle. *Nature* **332,** 546–548 (1988).
96. Bell, S. P. & Stillman, B. ATP-dependent recognition of eukaryotic origins of DNA replication by a multiprotein complex. *Nature* **357,** 128–134 (1992).
97. Gavin, K. A., Hidaka, M. & Stillman, B. Conserved Initiator Proteins in Eukaryotes. *Sci.* **270 ,** 1667–1671 (1995).
98. Hereford, L. M. & Hartwell, L. H. Sequential gene function in the initiation of Saccharomyces cerevisiae DNA synthesis. *J. Mol. Biol.* **84,** 445–461 (1974).
99. Nurse, P., Thuriaux, P. & Nasmyth, K. Genetic control of the cell division cycle in the fission yeast Schizosaccharomyces pombe. *Mol. Gen. Genet. MGG* **146,** 167–178 (1976).
100. Kelly, T. J. *et al.* The fission yeast cdc18+ gene product couples S phase to START and mitosis. *Cell* **74,** 371–382 (2015).

101. Calzada, A., Sacristan, M., Sanchez, E. & Bueno, A. Cdc6 cooperates with Sic1 and Hct1 to inactivate mitotic cyclin-dependent kinases. *Nature* **412,** 355–358 (2001).

102. Williams, R. S., Shohet, R. V. & Stillman, B. A human protein related to yeast Cdc6p. *Proc. Natl. Acad. Sci. U. S. A.* **94,** 142–147 (1997).

103. Maine, G. T., Sinha, P. & Tye, B.-K. Mutants of S. CEREVISIAE Defective in the Maintenance of Minichromosomes. *Genetics* **106,** 365–385 (1984).

104. Ishimi, Y., Komamura, Y., You, Z. & Kimura, H. Biochemical function of mouse minichromosome maintenance 2 protein. *J. Biol. Chem.* **273,** 8369–8375 (1998).

105. Barnhart, M. C. *et al.* HJURP is a CENP-A chromatin assembly factor sufficient to form a functional de novo kinetochore. *J. Cell Biol.* **194,** 229–43 (2011).

106. Green, E. M. *et al.* Replication-Independent Histone Deposition by the HIR Complex and Asf1. *Curr. Biol.* **15,** 2044 (2005).

107. Tyler, J. K. *et al.* Interaction between the Drosophila CAF-1 and ASF1 Chromatin Assembly Factors. *Mol. Cell. Biol.* **21,** 6574–6584 (2001).

108. Sharp, J. A., Fouts, E. T., Krawitz, D. C. & Kaufman, P. D. Yeast histone deposition protein Asf1p requires Hir proteins and PCNA for heterochromatic silencing. *Curr. Biol.* **11,** 463–473 (2015).

109. Lewis, P. W., Elsaesser, S. J., Noh, K.-M., Stadler, S. C. & Allis, C. D. 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 (2010).

110. Atp-dependent, T. HIRA and Daxx Constitute Two Independent Histone H3 . 3-Containing Predeposition Complexes. **LXXV,** (2010).

111. Verreault, A., Kaufman, P. D., Kobayashi, R. & Stillman, B. Nucleosome Assembly by a Complex of CAF-1 and Acetylated Histones H3/H4. *Cell* **87,** 95–104 (2015).

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

113. Moggs, J. G. *et al.* A CAF-1–PCNA-Mediated Chromatin Assembly Pathway Triggered by Sensing DNA Damage. *Mol. Cell. Biol.* **20,** 1206–1218 (2000).

114. Shibahara, K. & Stillman, B. Replication-Dependent Marking of DNA by PCNA Facilitates CAF-1-Coupled Inheritance of Chromatin. *Cell* **96,** 575–585 (1999).

115. Krude, T. Chromatin Assembly Factor 1 (CAF-1) Colocalizes with Replication Foci in HeLa Cell Nuclei. *Exp. Cell Res.* **220,** 304–311 (1995).

116. Green, C. M. & Almouzni, G. Local action of the chromatin assembly factor CAF-1 at sites of nucleotide excision repair in vivo. *EMBO J.* **22,** 5163–5174 (2003).

117. Kim, U. J., Han, M., Kayne, P. & Grunstein, M. Effects of histone H4 depletion on the cell cycle and transcription of Saccharomyces cerevisiae. *EMBO J.* **7,** 2211–2219 (1988).

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

119. Tchénio, T., Casella, J.-F. & Heidmann, T. A Truncated Form of the Human CAF-1 p150 Subunit Impairs the Maintenance of Transcriptional Gene Silencing in Mammalian Cells. *Mol. Cell. Biol.* **21,** 1953–1961 (2001).

120. Ray-Gallet, D. *et al.* HIRA Is Critical for a Nucleosome Assembly Pathway Independent of DNA Synthesis. *Mol. Cell* **9,** 1091–1100 (2002).

121. Osley, M. A. & Lycan, D. Trans-acting regulatory mutations that alter transcription of Saccharomyces cerevisiae histone genes. *Mol. Cell. Biol.* **7,** 4204–4210 (1987).

122. Xu, H., Kim, U. J., Schuster, T. & Grunstein, M. Identification of a new set of cell cycle-regulatory genes that regulate S-phase transcription of histone genes in Saccharomyces cerevisiae. *Mol. Cell. Biol.* **12,** 5249–5259 (1992).

123. Tyler, J. K. *et al.* The RCAF complex mediates chromatin assembly during DNA replication and repair. *Nature* **402,** 555–560 (1999).

124. Mello, J. A. *et al.* Human Asf1 and CAF-1 interact and synergize in a repair-coupled nucleosome assembly pathway. *EMBO Rep.* **3,** 329–334 (2002).

125. Sutton, A., Bucaria, J., Osley, M. A. & Sternglanz, R. Yeast ASF1 protein is required for cell cycle regulation of histone gene transcription. *Genetics* **158,** 587–596 (2001).

126. Saunders, A. *et al.* Tracking FACT and the RNA Polymerase II Elongation Complex Through Chromatin in Vivo. *Sci.* **301 ,** 1094–1096 (2003).

127. Tanaka, Y. *et al.* Expression and purification of recombinant human histones. **33,** 3–11 (2004).

128. Kato, T. *et al.* Activation of Holliday junction recognizing protein involved in the chromosomal stability and immortality of cancer cells. *Cancer Res.* **67,** 8544–53 (2007).

129. Tao, Y. *et al.* The structure of the FANCM-MHF complex reveals physical features for functional assembly. *Nat. Commun.* **3,** 782 (2012).

130. Kozekov, I. D. *et al.* DNA Interchain Cross-Links Formed by Acrolein and Crotonaldehyde. *J. Am. Chem. Soc.* **125,** 50–61 (2003).

131. Walden, H. & Deans, A. J. The Fanconi Anemia DNA Repair Pathway: Structural and Functional Insights into a Complex Disorder. *Annu. Rev. Biophys.* **43,** 257–278 (2014).

132. Niedernhofer, L. J. The Fanconi Anemia Signalosome Anchor. *Mol. Cell* **25,** 487–490 (2007).

133. Savitsky, K. *et al.* The complete sequence of the coding region of the ATM gene reveals similarity to cell cycle regulators in different species. *Hum. Mol. Genet.* **4 ,** 2025–2032 (1995).

134. Pandita, T. K. *et al.* Ionizing radiation activates the ATM kinase throughout the

cell cycle. 1386–1391 (2000).

135. Liao, M.-J., Yin, C., Barlow, C., Wynshaw-Boris, A. & van Dyke, T. Atm Is Dispensable for p53 Apoptosis and Tumor Suppression Triggered by Cell Cycle Dysfunction. *Mol. Cell. Biol.* **19,** 3095–3102 (1999).

136. Park, J.-Y. *et al.* Identification of a novel human Rad51 variant that promotes DNA strand exchange. *Nucleic Acids Res.* **36,** 3226–3234 (2008).

137. Baumann, P. & West, S. C. Role of the human RAD51 protein in homologous recombination and double-stranded-break repair. *Trends Biochem. Sci.* **23,** 247–251 (2015).

138. Nakanishi, K., Cavallo, F., Brunet, E. & Jasin, M. Homologous Recombination Assay for Interstrand Cross-Link Repair. *Methods Mol. Biol.* **745,** 283–291 (2011).

APPENDIX

Mis18-Alpha			
Plasmid Contents	Designation	Resistance (A=Ampicillin, K=Kanamycin)	Expression organism (H=human, B=bacteria)
His Mis18-alpha	pCK1	K	B
MBP-Mis18-alpha FL	IN59	A	B
MBP-Mis18 alpha (aa1-175)	IN53	A	B
MBP mis18 alpha (aa1-164)	IN74	A	B
GFP MIs18-alpha	pMS68.2	A	H
mCherry-LacI-Mis18-alpha	Same name in (pMB)	A	H
mCherry Laci Mis18 alpha (aa168-233)	IN146	A	H
GFP mis18 alpha (aa168-233)	IN146	A	H
HisMBP Mis18 Alpha aa 188-233	IN220	A	B
HNusA MIs18 Alpha aa 188-233	IN221	A	B
HisMBP MIs18 Alpha aa 1-187	IN222	A	B
HisNusA MIs18 Alpha aa 1-187	IN223	A	B
mCherry LacI Mis18 alpha (aa1-175)	pMS73.2	A	H
GFP Mis18 alpha (aa1-175)	pMS72	A	H
MBP Mis18 alpha C85A	IN130	A	B
mCherry lacI Mis18 Alpha C85A	IN132	A	H
GFP Mis18 alpha I201G L205G	IN226	A	H
GFP Mis18 alpha L215G L219G	IN227	A	H
Strep HA Mis18 alpha	IN137	A	B
GFP Mis18 alpha V82A F83A	IN165	A	H
GFP Mis18 alpha D94A S95A	IN176	A	H
GFP MIs18 alpha Y176A	IN215	A	H
HA Mis18 alpha C-terminal	pMS84.4	A	H
HisNusA Mis18 Alpha L215G L219G	IN229	A	B
HisMBP Mis18 Alpha I201G L205G	IN230	A	B
HisNusA MIs18 Alpha I201G L205G	IN231	A	B
HisMBP Mis18 Alpha L215G L219G	IN232	A	B

Mis18-Beta			
Strep Mis18 beta	pCK2	K	B
mCherry-LacI-Mis18-beta	pMS75.3	A	H
MBP-Mis18-beta FL	IN54	A	B
MBP mis18 beta (aa1-188)	IN55	A	B
HisNusA-Mis18-beta (aa164-229)	IN75-3	A	B
GFP Mis18-beta	pMS71.2	A	H
mCherry LacI Mis18 beta (aa164-229)	IN149	A	H
GFP mis18 beta (aa164-229)	IN147	A	H
mCherry LacI Mis18 beta (aa1-188)	pMS74.3	A	H
GFP Mis18 beta C80G	IN134	A	H
mCherry LacI	pDF287	A	H
GFP Mis18 beta (aa1-188)	pMS76.3	A	H
MBP Mis18 beta C80G	IN131	A	B
Strep HA mis18 beta	IN143	A	B
HA mis18 beta C-terminal	pMS85	A	H
3X-FLAG Mis18 beta	IN195	A	H
GFP Mis18 beta V77A F78A	IN173	A	H
GFP Mis18 beta Y172A	IN216	A	H
His NusA mis18 beta aa 1-188	IN217	A	B
HNusA Mis18 beta aa 189-229	IN218	A	B
HisMBP Mis18n beta aa 189-229	IN219	A	B
Mis18-S.p.			
His-Mis18 S. pombe	IN90-2	A	B
GFP MIs18 pombe	IN214	A	H
His-Fxa Mis18 S.p. aa1-141	IN91	A	B
HJURP			
HA HJURP FL	pDF241-1	A	H
MBP HJURP FL	IN89-6	A	B
MBP HJURP 1-348	IN123	A	B
MBP HJURP 348-555	IN120	A	B
MBP HJURP 555-748	IN127	A	B
GFP-HJURP R1 (aa348-555)	pEZ11	A	H
HJURP Scm3 Cons. DsRed PLVU	IN44	A	H
HJURP Scm3 Cons. R1 DsRed PLVU	IN45	A	H

HJURP Cons. R1 R2 DsRed PLVU	IN46	A	H
HJURP FL DsRed PLVU	IN47	A	H
CENP-A			
HA CENP-A	pDF88	A	H
YFP CENP-A H104G	IN203	A	H
YFP-CENP-A WT	pDF23	A	H
HA-H3-CATD	pDF500	A	H
HA-CENPA-L1	pDF501	A	H
HA-CENP-A H2.1	pDF503	A	H
HA-CENP-A H2.2	pDF504	A	H
HA-CENP-A H2.3	pDF505	A	H
T/W/S/X			
MBP-CENP-X, Strep-CENP-S bicystronic	IN30	A	B
MBP CENp-W, His CENP-T HFD bicystronic	IN50	A	B
Misc.			
MBP	IN42-3	A	B
GFP Tet	pDF318	A	H
Isce1-DR-GFP	IN16	A	B
mCherry-LacI	pDF287	A	H
RFP-H2B	pDF438	A	H

List of Abbreviations:

ATP – Adenosine triphosphate

CAL1 – Chromosome alignment defect 1

CATD – CENP-A targeting domain

CCAN – Constitutive centromere associated network

CD – Conserved domain of HJURP

CDK- Cyclin dependent kinase 1

CENP-A/Cnp1/CenH3/Cse4/CID – Centromere Protein-A

Centromere protein A CENP-ACAD – CENP-A distal complex

CENP-ANAC – CENP-A nucleosome associated complex

GAPDH -- Glyceraldehyde-3-phosphate dehydrogenase

GFP – green fluorescence protein

H2A – Histone H2A

H2B - Histone H2B

H3 – Histone H3

H4 - Histone H4

HJURP C-terminal domain HJURP/HJ – Holliday recognition protein

IgG -- Immunoglobulin G

IP -- Immunoprecipitation

IPTG -- Isopropyl 1-D-1-thiogalactopyranoside

kDa – Kilodalton

KNL2 – Kinetochore null protein 2

LacI – Lac repressor LacO/TRE – Lac operator / Tet response element

LAP – Localization and affinity purification tag

LPC – Leupeptin, pepstatin, chymotrypsin

MBP – Maltose binding protein

Mis18 – Missegregation of chromosomes phenotype mutant

Mis12- Missegregation of chromosomes phenotype mutant

Mis18BP1 – Mis18 binding protein

Npm1 – Nucleophosmin 1

PBS – Phosphate buffered saline

PIPES -- 1,4-Piperazinediethanesulfonic acid

PMSF -- Phenylmethanesulfonyl fluoride

PCR – Polymerase chain reaction

RbAp46/48 – Retinoblastoma binding protein 46/48 RIP

RNA – Ribonucleic acid

Scm3 – Suppressor of chromosome missegregation protein 3

FANC- Fanconi Anemia Complex

FAAP- Fanconi Anemia Associated Protein

CENP- Centromere Protein

CenTD – HJURP Centromere Targeting Domain

Rad51 – DNA Repair Protein Homolog 1

ATM - Ataxia telangiectasia mutated

NBS1 - Nibrin