2016

Cenp-A And Friends: A Collaborative Model For The Epigenetic Transmission Of Centromere Identity

Lucie Y. Guo
University of Pennsylvania, lucieguo@gmail.com

Cenp-A And Friends: A Collaborative Model For The Epigenetic Transmission Of Centromere Identity

Abstract

The accurate segregation of chromosomes during mitosis is essential for the survival and development of all eukaryotes, and this process requires the attachment of mitotic spindle microtubules to the kinetochore, which is assembled on the chromosomal centromere. Defections in centromere or kinetochore function can lead to the loss of genetic material during cell division, which can result in development defects or disease. Therefore, elucidating how the centromere is stably and accurately propagated across cell and organismal generations is crucial to our understanding of how genetic information is accurately inherited. Centromere location is specified epigenetically by the histone H3 variant termed centromere protein A (CENP-A). CENP-A molecules at the centromere are known to possess a remarkably stability, exhibiting almost no detectable turnover. This stability is crucial for maintenance of centromere identity, but the molecular basis for this stability is unclear. Additionally, new CENP-A molecules must be assembled onto centromeric chromatin at every cell cycle, or else this epigenetic mark will be diluted and inevitably lost over time. The process of CENP-A assembly is exquisitely regulated, but is poorly understood. In this thesis, we identified the role of an essential centromeric protein, CENP-C, in not only binding to CENP-A nucleosomes, but also reshaping and stabilizing it at centromeres. We then pinpoint the mechanism by which CENP-C stabilizes CENP-A nucleosomes through a critical arginine anchor, which drives the structural transition of the CENP-A nucleosome. We then assemble a core centromeric nucleosome complex (CCNC) containing the CENP-A nucleosome bound to the nucleosome-binding domains of both CENP-C and another centromeric protein, CENP-N, and provide the first biophysical insight into how both proteins collaborate to rigidify CENP-A nucleosomes in vitro. Additionally, using gene-editing and rapid protein degradation approaches, we demonstrate that CENP-C and CENP-N are both crucial in the maintenance of CENP-A nucleosomes in cells. And finally, we report the landscape of Cdk regulation on the CENP-A-specific chaperone, HJURP, which provide insight into the mechanism of nascent CENP-A assembly. Taken together, these findings advance our understanding in how centromere location is specified and maintained to ensure faithful inheritance of genetic information across cell divisions.

Degree Type

Dissertation

Degree Name

Doctor of Philosophy (PhD)

Graduate Group

Biochemistry & Molecular Biophysics

First Advisor

Ben E. Black

Subject Categories

Biochemistry | Biology | Biophysics

CENP-A AND FRIENDS: A COLLABORATIVE MODEL FOR THE EPIGENETIC TRANSMISSION OF CENTROMERE IDENTITY

Lucie Y. Guo

A DISSERTATION

In

Biochemistry and Biophysics
Presented to the Faculties of the University of Pennsylvania

In Partial Fulfillment of the Requirements for the
Degree of Doctor of Philosophy
2018

_____

Ben E. Black, Ph.D.
Associate Professor of Biochemistry and Biophysics
Supervisor of Dissertation

_____

Kim A. Sharp, Ph.D.
Associate Professor of Biochemistry and Biophysics
Graduate Group Chairperson

Dissertation committee:
Jim Shorter, Ph.D. Associate Professor of Biochemistry and Biophysics
Benjamin Garcia, Ph.D. Presidential Professor of Biochemistry and Biophysics
Roger A. Greenberg, M.D., Ph.D. Associate Professor of Cancer Biology
Shelley L. Berger, Ph.D. Daniel S. Och University Professor
Andrew J. Andrews, Ph.D. Assistant Professor, Fox Chase Cancer Center

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

I owe the following people (and many others) unwavering gratitude for their support during my PhD years:

First and foremost, to my advisor, **Ben Black**, who has been an incredible mentor and role model. Thank you for teaching me to think critically and rigorously, for continually pushing me beyond what I assume to be my limits, and always holding my work to the highest standards. Thank you for having invested so much time and energy into my scientific training. These years of working in your lab have undoubtedly been the most challenging, stimulating, and intellectually rewarding years of my life, and it has completely transformed me— not only into a better scientist, but also as a more diligent, resilient, and confident person. Your enthusiasm about science is infectious, and these years of working in your lab have surely persuaded me to try to continue making biomedical research a central part of my career down the road.

To the past and current members of the Black lab who I've had the privilege of working with: **Glennis Logsdon, Samantha Falk, Nikolina Sekulic, Tanya Panchenko, Morgan Gerace, Jennine Dawicki McKenna, Evan Smoak, Levani Zandarashvili, Praveen Kumar Allu, Hollis De Laney, Emily Bassett, Jamie DeNizio, Arunika Das, Hollis De Laney, Evan Selzer, Krystal Haislop, Kelly Karch**. You all are some of the smartest, kindest, most hard-working people I have ever met. I feel so grateful to have been part of a team with people who are just as passionate about the science as I have been. And a special thank-you to Nikolina, Samantha, Tanya, and Kevan— for having all the time you invested in training me when I first joined the lab. Also, to Glennis, who has been one of my best friends both in lab and in life. And to Kevan and Morgan, who have been amazing baymates.

Thank you to my outside collaborators for the project outlined in Chapter 3: **Kara McKinley** at the **Iain Cheeseman** lab (MIT), **Dani Fachinetti** (Institut Curie), **Don Cleveland** (UCSD).

Thank you to my collaborators for the project outlined in Chapter 4: **Xing-Jun Cao** at the **Ben Garcia** lab, and **Ana Stankovic** from the **Lars Jansen** lab (Instituto Gulbenkian).

Thank you to my thesis committee: **Jim Shorter, Ben Garcia, Roger Greenberg,** and **Shelley Berger** (as well as formerly **Walter Englander** and **Mitch Weiss**). Thank you so much for all your time and your valuable scientific feedback. Thank

you to **Andy Andrews** for kindly agreeing to read my thesis and attend my thesis defense.

Thank you to the **Walter Englander** lab for all your help with HXMS: **Leland Mayne, Zhong-yuan Kan, Wenbing Hu, Scott Ye.** My HXMS experiments would have been impossible without your guidance.

Thank you to other labs in the Penn community who have always given me valuable feedback during my research presentations: ChromoClub (**Mike Lampson** and **Katya Grishchuk** labs), the Crystal Talks community (**Ronen Marmorstein** lab and other labs), the GFBB group (**Ben Garcia**, **Huaying Fan**, **Roberto Bonasio** labs). Thank you for all the guidance you've given me over the years.

Thank you to the BMB Graduate Group (**Kim Sharp, Kate Ferguson, Rahul Kohli, Ruth Keris, Kelli McKenna**), and well as the Penn Medical Scientist Training Program (**Skip Brass, Maggie Krall**, **Maureen Kirsch, David Bittner**). Thank you for having provided me with such a supportive community in which to train as scientist.

Lastly— thank you to my family—my mom, dad, and my Irene, for your unwavering encouragement during all my many years of school. Thank you to my son Aiden, for having been such an easy-going and pleasant baby so far… otherwise it would have been a lot harder to get work done in lab! And thank you to my husband Ryan, without whom none of this would have been possible.

ABSTRACT

CENP-A AND FRIENDS: A COLLABORATIVE MODEL FOR THE EPIGENETIC TRANSMISSION OF CENTROMERE IDENTITY

Lucie Guo

Ben Black

The accurate segregation of chromosomes during mitosis is essential for the survival and development of all eukaryotes, and this process requires the attachment of mitotic spindle microtubules to the kinetochore, which is assembled on the chromosomal centromere. Defections in centromere or kinetochore function can lead to the loss of genetic material during cell division, which can result in development defects or disease. Therefore, elucidating how the centromere is stably and accurately propagated across cell and organismal generations is crucial to our understanding of how genetic information is accurately inherited. Centromere location is specified epigenetically by the histone H3 variant termed centromere protein A (CENP-A). CENP-A molecules at the centromere are known to possess a remarkably stability, exhibiting almost no detectable turnover. This stability is crucial for maintenance of centromere identity, but the molecular basis for this stability is unclear. Additionally, new CENP-A molecules must be assembled onto centromeric chromatin at every cell cycle, or else this epigenetic mark will be diluted and inevitably lost over time. The process of CENP-A assembly is exquisitely regulated, but is poorly understood. In this thesis, we identified the role of an essential centromeric protein, CENP-C, in not only binding to CENP-A nucleosomes, but also reshaping and stabilizing it at centromeres. We then pinpoint the mechanism by which CENP-C stabilizes CENP-A nucleosomes through a critical arginine anchor, which drives the structural transition of the CENP-A nucleosome. We then assemble a core centromeric nucleosome complex (CCNC) containing the CENP-A nucleosome bound to the nucleosome-binding domains of both CENP-C and another centromeric protein, CENP-N, and provide the first biophysical insight into how both proteins collaborate to rigidify CENP-A nucleosomes *in vitro*. Additionally, using gene-editing and rapid protein degradation approaches, we demonstrate that

CENP-C and CENP-N are both crucial in the maintenance of CENP-A nucleosomes in cells. And finally, we report the landscape of Cdk regulation on the CENP-A-specific chaperone, HJURP, which provide insight into the mechanism of nascent CENP-A assembly. Taken together, these findings advance our understanding in how centromere location is specified and maintained to ensure faithful inheritance of genetic information across cell divisions.

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: THE EPIGENETIC BASIS FOR CENTROMERE IDENTITY

1.1. INTRODUCTION

All eukaryotes require a highly regulated machinery to faithfully propagate genetic material during cell division. The centromere is a specialized region on each chromosome that provides the locus for forming a protein complex known as the kinetochore. During mitosis, the kinetochore connects chromosomes to spindle microtubules, which segregate chromosomes between daughter cells (**Figure 1**). Maintaining centromere identity is crucial for the survival and well-being of all organisms. Loss of a single centromere causes loss of that chromosome in one daughter cell and an extra copy in the other. On the other hand, presence of multiple functional centromeres on a single chromosome will render that chromosome vulnerable to breakage by opposing forces of spindle microtubules. Centromere defects during meiotic divisions cause aneuploidy (defined as gain or loss in chromosome number) in the gametes, which leads to spontaneous abortion or developmental defects in the embryo (Hassold and Hunt, 2001). Also, aneuploidy has been consistently observed in virtually all human cancers (Kops et al., 2005; Rajagopalan and Lengauer, 2004), and has been proposed to contribute to carcinogenesis by causing genomic instability (Sheltzer et al., 2011; Solomon et al., 2011).

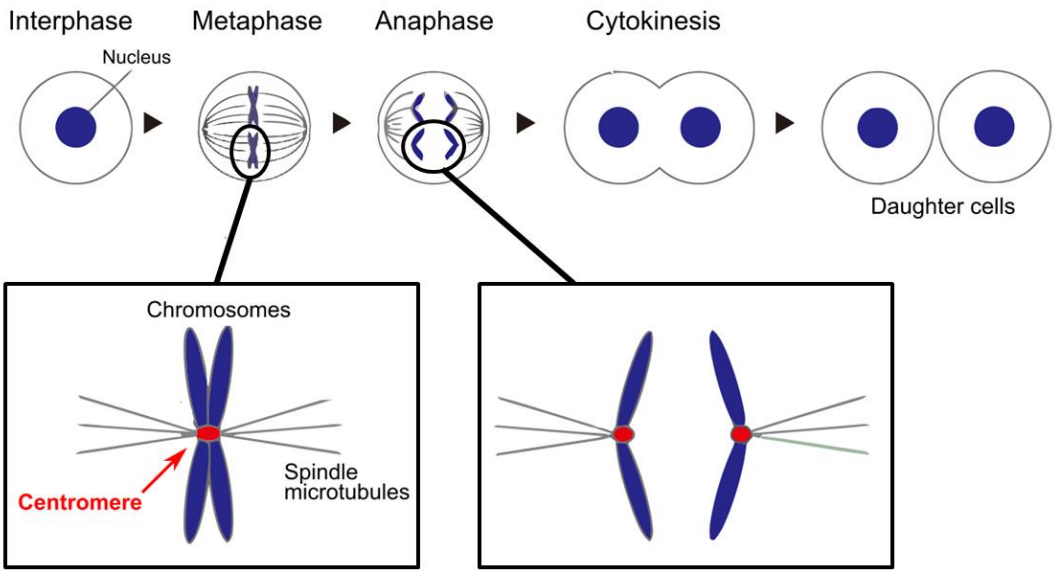


Figure 1. Schematic of the centromere in cell division.
(Adapted from http://www.nig.ac.jp)

In this chapter, I begin with a historical overview of the centromere and its epigenetic nature of inheritance. I then hone in on the epigenetic marker of centromere location, centromere protein A (CENP-A), and summarize its initial discovery and its unique structural features that play a role in specifying centromere identity. Then, since CENP-A does not exist in isolation at the centromere, but rather bound to a host of constitutively associated centromere proteins, I summarize these "friends of CENP-A" at the centromere, and our current understanding of their roles in eukaryotic centromere architecture. Lastly, I explain the key players and regulatory mechanisms ensuring the nascent assembly of CENP-A with each new cell divisions. Therefore, with this chapter, I aim to set the historical context for the two main focuses of my graduate work: elucidating the contributions from CENP-A's closest binding partners at the centromere, CENP-C and CENP-N, in endowing CENP-A with its remarkable ability to be stably retained at existing centromeres without any detectable turnover (Chapters 2 and 3), and regulation of the CENP-A-specific histone chaperone, HJURP, in the assembly of nascent CENP-A molecules at the centromere (Chapter 4).

1.2. CHROMOSOMES AND CENTROMERES

Our understanding of heredity first stems from work by Darwin, whose seminal 1859 book *On the Origin of Species* set forth the theory of evolution, which eventually became the unifying theory for all of the life sciences. But at the time of its publication, despite its compelling evidence for evolution, one of its major problems was lack of an underlying mechanism for heredity. During the 18th century, some scientists, including Dutch microbiologist Antonie van Leeuwenhoek, were influenced by prior thoughts in preformationism— the idea that living being exist as miniature versions of themselves prior to their development— and speculated the existence of a "homunculus" (little man) inside each sperm. Of course, our understanding of how inherited traits are transmitted from across generations stems from principles first proposed by Gregor Mendel, whose studies in peas were published in 1866 as "*Experiments in Plant Hybridisation*" which postulated what eventually became known as the Mendelian laws of inheritance, which included the principles of independent assortment (in which specific traits operate independently) and independence segregation (in which each pair of alleles segregate in germ cells and recombine during reproduction). And in 1882, Walther Flemming, a German anatomist, studied cell division in the fins and gills of salamanders, and discovered a structure that strongly stained with aniline dyes, which he named "chromatin". He saw that during cell division, chromatin separates into threadlike strings, which were later named "chromosomes" (mean "colored bodies")(**Figure 2**). Then, in the early 1900s, Sutton and Boveri independently postulated that genetic information is carried by chromosomes. Sutton examined the *Brachystola magna* (a grasshopper), and showed that there exist matched pairs of maternal and paternal chromosomes that separate during meiosis (Sutton, 1902, 1903). Boveri studied sea urchins and showed that proper embryonic development required the presence of all chromosomes. With that, the work by Sutton and Boveri, which was collectively called the chromosome theory of inheritance, became the fundamental unifying theory of genetics, and finally provided the underlying mechanism for the laws of Mendelian inheritance.

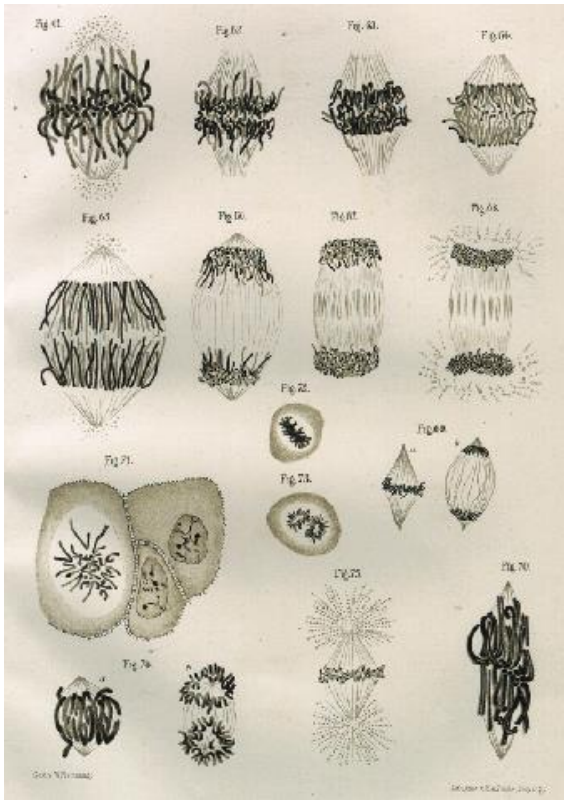


Figure 2. Drawing of mitosis by Walther Flemming in 1882 (O'Connor and Miko, 2008)

The architecture of chromatin remained elusive for many decades. Some proposed that each chromosome was a bundle of various strands of DNA wound around each other, and it wasn't until 1984 when Schwartz and Cantor separated intact chromosomal DNA by a pulsed-field gradient gel, that it was shown that each chromosome is one linear DNA molecule (Schwartz and Cantor, 1984). In 1936, the site where chromosomes associate with the spindle during cell division was given the name "centromere" (Darlington, 1936). Centromeres were originally define as regions of suppressed meiotic recombination (Beadle, 1932), and later recognized by cytogenetic techniques as the most constricted region of mitotic chromosomes, often residing within dark-staining heterochromatin by C-banding. The centromere was first cloned in *S. cerevisiae* (budding yeast) in 1980 by Clarke and Carbon, who identified the genes that persistently segregated with the centromere of chromosome III, and showed these DNA sequences were sufficient to direct

segregation of exogenous DNA (Clarke and Carbon, 1980). The centromere in budding yeast contains a small stretch of DNA spanning 125 bp that is sufficient for directing accurate chromosome segregation and contains three DNA elements: CDEI, CDEII, and CDEIII that are common to all chromosomes, a feature that is not present in higher eukaryotes (Bloom and Carbon, 1982; Saunders et al., 1988). Since budding yeast centromeres are small and possess a single microtubule attachment per chromosome, which is now known as "point centromeres". In contrast to the "point centromere" of the budding yeast, fission yeast and higher eukaryotes possess "regional centromeres" that span kilo- to mega-bases of highly repetitive DNA organized into hierarchy arrays of satellite repeats. The notable exception is *C. elegans* and other nematodes, as well as some insects and plants, which do not possess point nor regional centromeres, but rather holocentric centromeres, which diffusely span the entire length of the chromosome (Guerra et al., 2010; Hughes-Schrader and Ris, 1941).

1.3. CENP-A AS THE EPIGENETIC MARKER OF CENTROMERE LOCATION

Even though the budding yeast centromere relies on DNA for specifying its location, in all other eukaryotic species, there is a growing consensus that the location of the centromere is specified "epigenetically" by the protein that associate with centromeric DNA, particularly by centromeric protein A (CENP-A), which is known as Cse4 in *S. cerevisiae* (Stoler et al., 1995), HCP-3 in *C. elegans* (Buchwitz et al., 1999), Cnp1 in *S. pombe* (Takahashi et al., 2000), and CID in *Drosophila* (Henikoff et al., 2000). The first report of centromeric proteins came from an examination of serum samples from a patient with scleroderma, with a syndrome consisting of calcinosis, Raynaud's phenomenon, esophageal dismotility, sclerodactyly and telangiectasia (thus "CREST syndrome), who developed anti-centromere proteins as recognized by immunofluorescence (Earnshaw and Rothfield, 1985). These proteins were designated as CENP (CENtromere Protein)-A, B, and C, in order of smallest to

largest molecular weight as they appeared on the gel. And the first glimpse that centromeres could be specified not by any specific DNA sequences, but rather by the proteins, came from a study involving patient samples, in which the authors observed that on the inactive centromere of a dicentric chromosome, while the centromere DNA structures remained intact (as seen by traditional banding), CENP-A, B, and C were all undetectable (Earnshaw and Migeon, 1985). Using sera from CREST patients, the CENP-A protein was crudely extracted from nuclei of HeLa cells and shown to elute from a cation exchange column under denaturing conditions at nearly the identical NaCl concentration as H3 and H4; with these limited data, Palmer et al. boldly speculated that "CENP-A functions as a centromere-specific core histone, possibly substituting for H3 or H4" (Palmer et al., 1987). A few years later, taking advantage of the fact that CENP-A is quantitatively retained in bull spermatozoa (unlike the canonical histones), Palmer et al. purified CENP-A and obtained limited sequencing data suggesting that it contained regions of high similarity to histone H3, and some regions that are dissimilar, which they again precociously proposed to be "involved in loading CENP-A to centromeric DNA or in centromere-specific functions of CENP-A"(Palmer et al., 1991). CENP-A was cloned for the first time in 1994, and it was shown that CENP-A shares ~60% homology with histone H3, and the expression of epitope-tagged derivative of CENP-A was sufficient for targeting to centromeres (Sullivan et al., 1994). From these collective lines of evidence, there emerged the provocative finding that what makes the centromere unique from the rest of the chromatin is *"at the most fundamental level of chromatin structure, the nucleosome"* (Sullivan et al., 1994).

There exist multiple compelling lines of evidence that the presence of endogenous centromeric repeats is neither necessary nor sufficient for centromere identity. In *S. pombe*, plasmids containing minimal centromeric DNA can stochastically establish functional centromeres, albeit inefficiently, but can be faithfully propagated if the centromere becomes functional (Steiner and Clarke, 1994). Additionally in *S. pombe*, if the centromere were conditionally deleted, the survivors will gain centromeres at other novel chromosomal locations (Ishii et al., 2008). Indeed, one of most striking pieces of evidence that human centromeres are

inherited epigenetically came from the discovery of *neocentromeres*, which is a new centromere that forms on a normally non-centromeric region of chromatin, usually following the disruption of the natural centromere. The best-characterized neocentromere is on a chromosome 10 fragment, called mar del(10), which is mitotically stable and lacks detectable α-satellite sequences, which are the major repeat sequences underlying all human natural centromeres (Voullaire et al., 1993). Additionally, in pseudodicentric chromosomes, which contain two regions with α-satellite sequence (arisen from DNA translocation or duplication), only one of the two α-satellite regions remains an active centromere. Neocentromeres and active centromeres of pseudodicenteric chromosomes, despite being dissimilar in their underlying DNA sequence, maintain their ability to recruit CENP-A (Bassett et al., 2010; Lo et al., 2001; Voullaire et al., 1993). Furthermore, it has been shown that artificial targeting of Drosophila CENP-A (a.k.a. CID) fused to LacI to a stably integrated lac operator is able to generate a ectopic centromere that assembles a functional kinetochore be stably transmitted for several cell generations, even after the CID-LacI fusion protein is eliminated (Mendiburo et al., 2011). Novel centromere formation by artificial targeting proteins has also succeeded in human cells, by targeting fusion proteins of LacI with CENP-A (Logsdon et al., 2015) or the CENP-A-specific chaperone, HJURP (Barnhart et al., 2011). Taken together, there is growing consensus in the field that it is CENP-A nucleosomes, and not the underlying sequence, that specifies centromere location in higher eukaryotes.

1.4. STRUCTURAL FEATURES OF CENP-A NUCLEOSOMES

CENP-A nucleosomes have distinct structural properties compared to canonical H3 nucleosomes (summarized in **Figure 3**). Experiments in hydrogen-deuterium exchange coupled to mass spectrometry showed that the centromere-targeting domain (CATD) of CENP-A confers a unique structural rigidity, resulting in tetramers with H4 that are more conformationally rigid than the H3/H4 tetramer (Black et al., 2004). And when CENP-A/H4 tetramers are assembled into nucleosomes, despite the conformational constraints required for nucleosome

formation, nucleosomes assembled with CENP-A are still conformationally rigid than nucleosomes assembled with canonical histone H3 (Black et al., 2007a).

Crystal structure of the CENP-A/H4 tetramer revealed that the interface between the CENP-A molecules on either side of the nucleosome are rotated compared to the H3-H3 interface, as well as unique patches of hydrophobic residues on the CENP-A α2 helix and Loop1 (so called "hydrophobic stitches") that mediate the conformational rigidity that CENP-A confers to nucleosomes (Sekulic et al., 2010). Additionally, CENP-A possesses a bulged loop L1 containing side chains that generate a positive surface charge, in contrast to the negative surface charge at this location on canonical nucleosomes (Luger et al., 1997; Sekulic and Black, 2012; Sekulic et al., 2010).

The crystal structure of the CENP-A nucleosome claimed that the CENP-A/CENP-A interface to be rotated to a conventional degree, so that the $(CENP\text{-}A/H4)_2$ tetramer and the nucleosome maintain a shape similar to that of canonical H3 nucleosomes (Tachiwana et al., 2011). However, since a crystal structure is a snapshot of the nucleosome shape most conducible to crystallization, it remained a possibility that the CENP-A nucleosome could sample multiple conformations in solution, and that as the $(CENP\text{-}A/H4)_2$ tetramer is assembled into a nucleosome, the unique rotation at the CENP-A/CENP-A interface would generate a smaller radius of curvature of DNA wrapping near the dyad of the nucleosome, as well as wider spacing of the H2A/H2B heterodimers (Sekulic and Black, 2012). Indeed, in-solution FRET studies (which are further described in this thesis) have revealed that the H2B distances for CENP-A nucleosomes are ~5Å farther apart than expected from the crystal structure of the CENP-A nucleosome, whereas H2B distances in canonical nucleosomes are shorter and match the prediction from their crystal structure (Falk et al., 2015). Specifically, this nucleosome shape change directed by CENP-A is dominated by lateral passing of the two DNA gyres (Falk et al., 2016). Additionally, careful overlay of the structures of the $(CENP\text{-}A/H4)_2$ tetramer, the CENP-A nucleosome and the canonical H3 nucleosome reveal that the CENP-A nucleosome crystal structure has a CENP-A–CENP-A interface that exhibits rotation

but not compaction, and therefore is an intermediate structure between the (CENP-A/H4)$_2$ tetramer and the H3 nucleosome (Falk et al., 2016).

Another important distinguishing feature of CENP-A nucleosomes, which has emerged from multiple lines of evidence, is the unwrapping of its terminal DNA. Early reconstitutions of CENP-A and H3 nucleosomes and comparison of their electrophoretic mobility revealed that CENP-A nucleosomes exhibit unwrapping at the entry-exit site relative to canonical nucleosomes, which hinders the binding of the linker histone H1 (Conde e Silva et al., 2007). The αN helix, which would stabilize DNA interactions at the termini, is seen to be shorter in the crystal structure of the CENP-A nucleosome compared to that of the H3 nucleosome (Tachiwana et al., 2011). Hydrogen-deuterium exchange coupled to mass spectrometry (HXMS) of CENP-A vs. H3 polynucleosomes arrays has revealed that the αN helix of CENP-A is less protected than that of H3 upon folding of the nucleosome arrays (Panchenko et al., 2011). Chip-Seq experiments examining CENP-A at human centromeres and neocentromeres have also revealed that CENP-A nucleosome possess looser DNA termini than that of conventional H3 nucleosomes (Hasson et al., 2013). Recent experiments with electron cryomicroscopy (ECM) have again confirmed that CENP-A nucleosomal ends exhibit a high degree of flexibility, and that the αN helix of H3 plays a role in stabilizing the termini DNA in the conventional nucleosome (Roulland et al., 2016).

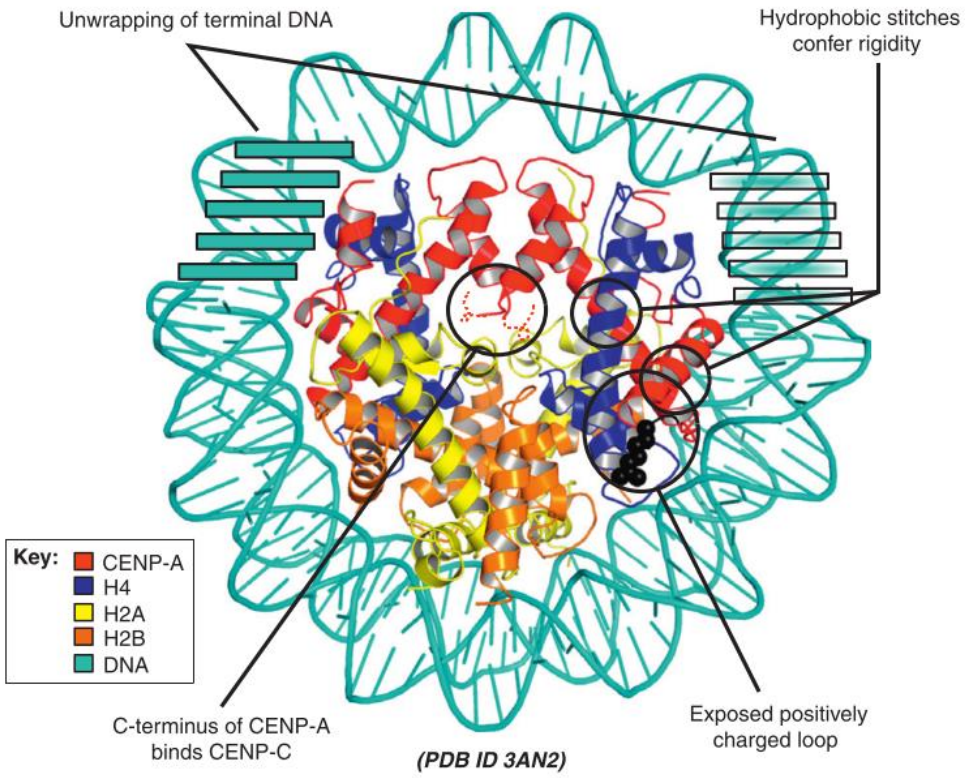


Figure 3. Unique physical properties of the mammalian CENP-A-containing nucleosome.
Distinguished physical properties of CENP-A nucleosome are highlighted in black circles. Clockwise from the top left, these include transient unwrapping/flexibility of the final helical turn of DNA at each nucleosome terminus; hydrophobic stitches that rigidify the CENP-A/H4 interface; a bulged loop L1 that is of opposite charge as on H3 in the conventional nucleosome; and the unstructured C-terminus that mediates recognition of CENP-A nucleosomes by CENP-C. Reproduced from (Sekulic and Black, 2012)

There has been some proposals that CENP-A resides in strange conformations, such as tetrasomes (Williams et al., 2009), hexasomes (Mizuguchi et al., 2007), and especially hemisomes with right-handed DNA wrapping (Bui et al., 2012; Furuyama et al., 2013; Shivaraju et al., 2012). However, the recent growing consensus in the field is that CENP-A nucleosomes exist as an octamer in vivo with conventional left-handed DNA wrapping (Bassett et al., 2012; Hasson et al., 2013; Padeganeh et al., 2013; Sekulic et al., 2010; Tachiwana et al., 2011), similar to the conventional H3 nucleosome.

1.5. RECOGNITION OF CENP-A NUCLEOSOMES BY CENTROMERE PROTEINS

The key function of the centromere is to recruit the kinetochore, which then mediates the attachment of sister chromatids to spindle microtubules. At the centromere, CENP-A is bound to a large complex of inner kinetochore proteins collectively referred to as the constitutive centromere-associated network (CCAN). To date, the CCAN is a group of 16 proteins that are known to form individually sub-complexes: CENP-C (Carroll et al., 2009; Kato et al., 2013), the CENP-L/N complex (Carroll et al., 2009; Fang et al., 2015; McKinley et al., 2015), the CENP-H/I/K/M complex (Basilico et al., 2014; Klare et al., 2015a; McKinley and Cheeseman, 2016), the CENP-O/P/Q/U/R complex, ,and the CENP-T/W/S/X complex (Nishino et al., 2012), and there are multivalent interactions between many members of the CCAN. Additionally, CENP-B is known to specifically bind to mammalian centromeric DNA, and recent evidence suggest that CENP-B could directly interact with CENP-A (Fachinetti et al., 2013). Many of these proteins had been identified in the CENP-A interactome (Foltz et al., 2006; Izuta et al., 2006). In this section, I summarize the literature on each of the sub-complexes within the CCAN, and the extensive network of known interactions between these sub-complexes.

1.5.1 CENP-C

1.5.1.1 CENP-C AS THE PLATFORM FOR KINETOCHORE ASSEMBLY

CENP-C was the first protein to be identified as a component of the human kinetochore plate (Saitoh et al., 1992). Among members of the CCAN, CENP-C has emerged as a key "founder" of kinetochore assembly. Human CENP-C is 943 amino acids in length, and sequence analysis predicts CENP-C to be almost entirely disordered (Westermann and Schleiffer, 2013), with the exception of a cupin fold at the C-terminus known to be required for dimerization (Cohen et al., 2008; Sugimoto

et al., 1997; Trazzi et al., 2009; Westermann and Schleiffer, 2013). CENP-C likely acts as a linear platform, with distinct modules that attach to downstream components (summarized in **Figure 4**). A fragment encompassing a.a. 200-400 on CENP-C has been shown to interact with CENP-H/K of the CENP-H/I/K/M subcomplex (Klare et al., 2015a). At its N-terminus, CENP-C contains a binding site for the Mis12 subcomplex of the KMN network (Przewloka et al., 2011; Screpanti et al., 2011), thus CENP-C functions as a primary link between the inner and outer kinetochore. While some studies have proposed two parallel and non-redundant pathways of kinetochore recruitment driven CENP-C and CENP-T (see "Section 1.5.3 The CENP-T/W/S/X and CENP-H/I/K/M subcomplex"), there is growing evidence that is actually a single pathway of kinetochore assembly, involving CENP-C as the founder, with CENP-H/I/K/M and CENP-T/W as followers that are dependent on CENP-C (Basilico et al., 2014; Carroll et al., 2010; Gascoigne et al., 2011a; Logsdon et al., 2015; Milks et al., 2009).

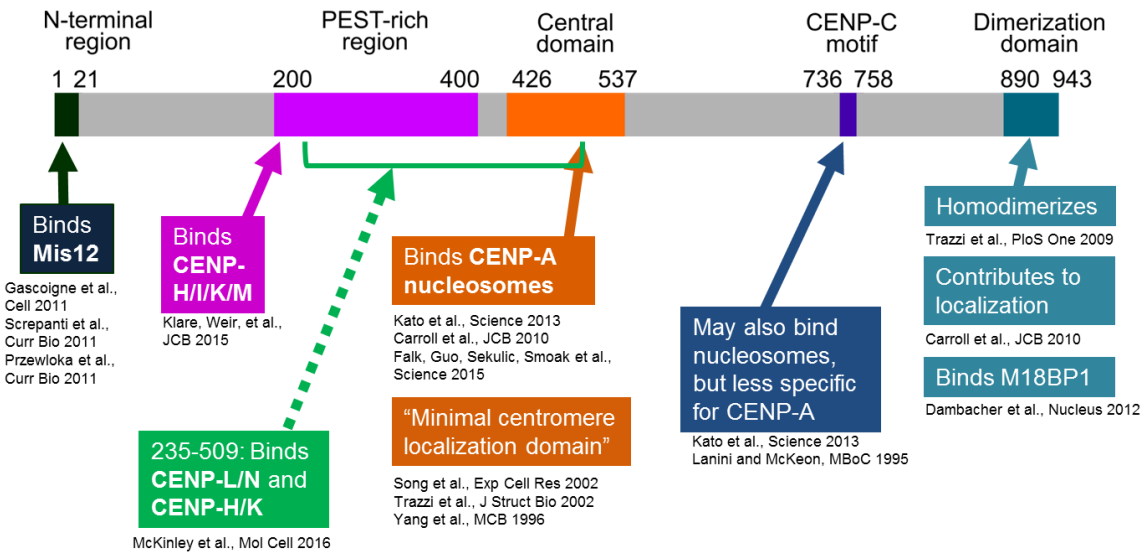


Figure 4. Summary of the known functions of each domain of the CENP-C protein at the outset of the projects described in this thesis.

1.5.1.2 Nucleosome recognition by CENP-C

For binding to the nucleosome, CENP-C contains two putative nucleosome-binding domains: the central domain [CENP-C^{CD} a.a. 426-537](Carroll et al., 2010) and the CENP-C motif [CENP-C^{CM} a.a. 736-758](Kato et al., 2013), which are proposed to engage the CENP-A nucleosome through similar histone contact points. CENP-C^{CD} is conserved in mammals (Kato et al., 2013), was mapped initially as the primary CENP-A nucleosome contact site, and has high specificity for CENP-A nucleosomes versus its counterparts with canonical H3 (Carroll et al., 2010). CENP-C^{CM}, on the other hand, is conserved from yeast to humans, and represents the only identified nucleosome-binding domain in species lacking a conserved CENP-C^{CD} (Kato et al., 2013). There currently exists an NMR model of the CENP-C^{CD} bound to a canonical nucleosome in which the 6 a.a. C-terminal tail of CENP-A replaces the corresponding region of histone H3, as well as a crystal structure of CENP-C^{CM} bound to this nucleosome (Kato et al., 2013). These structural information suggest that both CENP-C domains could interact with CENP-A nucleosomes in a similar manner, using multiple contact points, specifically the 6 amino-acid C-terminal tail ("LEEGLG") tail of CENP-A, the acidic patch of H2A, and the α2 helix of histone H4 (Kato et al., 2013). Both CENP-C^{CD} and CENP-C^{CM}, across multiple species, possess hydrophobic residues (WW or YW) that are proposed to interact with the "LEEGLG" tail, as well as a critical arginine residue (R522 in humans or R742 in Xenopus) that is proposed to interact with the acidic patch of H2A. This mechanism of nucleosome recognition by a critical arginine residue is a shared feature of a diverse set of nucleosome binding proteins studied to date (Armache et al., 2011; Barbera et al., 2006; Makde et al., 2010; McGinty et al., 2014; Morgan et al., 2016), and has been called the "arginine anchor" as an emerging paradigm for nucleosome recognition (**Figure 5**)(McGinty and Tan, 2016).

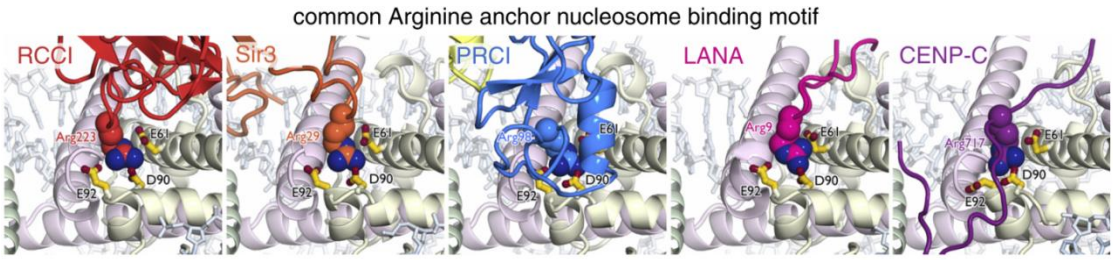


Figure 5. The arginine anchor motif.
Side chains of arginine anchor residues from RCC1, Sir3, PRC1 (Ring1B), LANA, and CENP-C shown in surface representation with side chains from acid triad of H2A in stick representation. Nucleosomal DNA shown as sticks. Reproduced from (McGinty and Tan, 2016)

Prior reports have suggested that either or both of the nucleosome binding domains of CENP-C could be important for its own localization to centromeres (Carroll et al., 2010; Kato et al., 2013; Lanini and McKeon, 1995; Milks et al., 2009; Trazzi et al., 2002; Yang et al., 1996). At the outset of this thesis, it had been unclear whether one or both of these domains are required for CENP-A nucleosome recognition.

1.5.2 The CENP-L/N subcomplex

1.5.2.1 Nucleosome recognition by CENP-N

Other than CENP-C, the only other member of the CCAN known to directly bind to CENP-A is the CENP-N subunit of the CENP-L/N subcomplex. CENP-N was first identified as a "reader" of epigenetic marks in CENP-A-containing chromatin, reported to directly bind the centromere targeting domain (CATD) of CENP-A, and also play a role in nascent CENP-A assembly (Carroll et al., 2009). The CATD provides exposed side-chains within its Loop 1 that are specific to the CENP-A nucleosome, which was predicted to be the sites for recognition by CENP-N (Sekulic and Black, 2009), and was shown to be important for the initial recruitment of CENP-N (Logsdon et al., 2015). Indeed, two residues on the Loop1 of CENP-A (R80 and G81, termed the "RG loop") was shown to be the key residues for recognizing CENP-N, since mononucleosomes assembled with CENP-A^{R80A,G81A} no longer binds to the N-terminal domain of CENP-A (Fang et al., 2015). For CENP-N, its nucleosome-

binding region was narrowed down to its N-terminal domain (a.a. 1-289), while deletion of its more conserved C-terminal domain (a.a. 290-339) did not affect nucleosome binding but is required for its partner CENP-L (Carroll et al., 2009). It is not known which exact CENP-N residues are required for the interaction interface between CENP-N and the CENP-A nucleosome. Mutation of two residues far apart in primary sequence (CENP-N^{R11A} and CENP-N^{R196A}) resulted in reduced CENP-A nucleosome binding, indicating that both residues in CENP-N contribute to recognition of CENP-A nucleosomes (Carroll et al., 2009). Furthermore, the recombinant CENP-L/N subcomplex has shown to be capable of directly binding CENP-C (McKinley et al., 2015; Nagpal et al., 2015), specifically via a.a. 235-509 of CENP-C (McKinley et al., 2015).

Even though CENP-N and CENP-C are both known to directly bind to CENP-A nucleosomes, at the outset of the projects described in this thesis, it had not been clearly shown with recombinant components whether CENP-N and CENP-C can bind to the same or different CENP-A nucleosomes. It was first proposed that CENP-C and CENP-N can bind to distinct sites on the same nucleosome, since it was reported that [^{35}S]methionine-labeled CENP-N could still bind to reconstituted nucleosomes incubated with CENP-C^{CD} from rabbit reticulocytes, but the band on the native gel that was indicated to be the complex was not further isolated to confirm that both binding partners are present (Carroll et al., 2010).

1.5.2.2 Cell-cycle dependence of CENP-N recruitment

It has been proposed that CENP-N recruitment to CENP-A chromatin is cell-cycle dependent. The centromeric localization of CENP-N has been shown to be increased at early S phase, peaking at middle/late S and G2 phase, and decreased during mitosis and G1 phase (Fang et al., 2015; Hellwig et al., 2011). Furthermore, recombinant CENP-N(1-289) binds more weakly to CENP-A polynucleosomes in the presence of $MgCl_2$, suggesting that chromatin compaction prevents CENP-N recruitment, possibly due to decreased accessibility of the RG loop (Fang et al., 2015).

1.5.3 THE CENP-T/W/S/X AND CENP-H/I/K/M SUBCOMPLEXES

CENP-T/W/S/X and CENP-H/I/K/M are both important heterotetramers within the CCAN. CENP-T-W and CENP-S-X each possess histone folds, and co-assemble to form a CENP-T/W/S/X heterotetramer that resembles a centromeric nucleosome-like structure, which binds to and induces positive supercoils within the DNA (Nishino et al., 2012; Takeuchi et al., 2014). CENP-H/I/K/M binds to the PEST-rich region (a.a. 200-400) of CENP-C (Klare et al., 2015a), and also directly interacts with CENP-T/W/S/X (Basilico et al., 2014). CENP-M resembles a small GTPase (structurally and evolutionarily), but cannot perform GTP-binding and conformational switching (Basilico et al., 2014). CENP-H/I has been reported to play a role in nascent CENP-A deposition (Okada et al., 2006).

It has been suggested that CENP-C and CENP-T form parallel but non-redundant pathways that recruit the outer kinetochore: targeting CENP-C and CENP-T independently to an ectopic chromosome locus revealed that CENP-C interacts with KNL1 and the Mis12 complex, whereas CENP-T directly interacts with Ndc80 (which in turn promotes recruitment of KNL1/Mis12), suggesting that these two pathways are not duplications, but rather distinct mechanisms (Rago et al., 2015). These two pathways have been reported to be susceptible to distinct modes of regulation, with Aurora B promoting KMN recruitment to CENP-C and Cdk promoting KMN recruitment to CENP-T (Rago et al., 2015). Furthermore, the targeting of CENP-T or CENP-C to ectopic chromosomal loci can result in CENP-A-independent kinetochore assembly (Gascoigne et al., 2011a; Hori et al., 2013). However, this model of CENP-T/W as a founder of an independent axis of kinetochore assembly has been called to question by various lines of evidence. Recruitment of CENP-T to ectopic centromeres during initial centromere establishment has been shown to require CENP-C, CENP-N, as well as the N-terminal tail of CENP-A (Logsdon et al., 2015). Furthermore, A point mutant affecting the interaction between CENP-M and CENP-I prevents kinetochore recruitment of the CENP-T/W complex (Basilico et al., 2014).

1.5.4 CENP-B

CENP-B is known to bind specifically to a 17bp DNA sequence known as the "CENP-B" box (Masumoto et al., 1989). CENP-B knockout mice are viable and largely developmentally and reproductively normal, with a few phenotypic defects: CENP-B-null mice have normal weight at birth, but subsequently lag behind. Additionally, the testes of CENP-null mice have significantly decreased weight and sperm content (Hudson et al., 1998; Perez-Castro et al., 1998). This is in stark contrast to the essential protein members of the CCAN, and is still perplexing to this field. However, CENP-B could play a role in centromere formation, since the *de novo* formation of human artificial chromosomes requires presence of a-satellite DNA containing CENP-B boxes (Okada et al., 2007), but paradoxically, human neocentromeres can exist in the absence of CENP-B (Voullaire et al., 1993), and the human Y chromosome does not contain CENP-B boxes (Masumoto et al., 1989). Nonetheless, the presence of CENP-B has been reported to help protect against chromosome missegregation: chromosomes without CENP-B (i.e., neocentromeres, the Y chromosome) are shown to mis-segregate at greater frequency (Fachinetti et al., 2015), and this dependence on CENP-B is heightened upon artificially reduced CENP-C levels at the centromere (Fachinetti et al., 2015). So perhaps, the functions of CENP-B have been made largely redundant by the multitude of players at the mammalian centromere.

Nonetheless, there is growing evidence that the presence of CENP-B on the CENP-B boxes can influence nearby nucleosomes. The crystal structure of the complex of CENP-B bound to DNA showed that the DNA-binding domain of CENP-B induces a 60° kink in the CENP-B box DNA (Tanaka et al., 2001) and induces translational positioning of nucleosomes (Tanaka et al., 2005; Yoda et al., 1998). Recently, ChiP-seq analyses of CENP-A-containing particles from human centromeres on a-sat DNA and naturally occurring neocentromeres has suggested that CENP-B plays a role in CENP-A-nucleosome phasing, and suggests that CENP-B

binding to the CENP-B box promotes asymmetric unwrapping of CENP-A-nucleosome terminal DNA (Hasson et al., 2013). Also, CENP-B may directly interact with CENP-A via the amino-terminal tail of CENP-A (Fachinetti et al., 2015).

1.6. PROPAGATION OF CENTROMERE LOCATION

Maintenance of genomic integrity requires accurate propagation of centromere number and position with each cell cycle. Failure to accurately propagate the centromere would cause the chromosome to be unable to attach to the mitotic spindle, leading to catastrophic consequences for cell division. Therefore, the machinery to ensure accurate inheritance of the epigenetic centromere marker, CENP-A, is of utmost importance for preserving genomic stability across cell and organismal generations.

Central to the accurate propagation of CENP-A is the extraordinary stability of the CENP-A molecule, which does not exchange upon assembly into a centromere (further explained in Chapter 2 and 3 of this thesis)(Bodor et al., 2013; Falk et al., 2015). Indeed, this stability has been measured out through the entire fertile lifespan of female mice (>1 year): the pool of CENP-A nucleosomes assembled before birth in the mouse oocyte is stably transmitted to embryos (Smoak et al., 2016). The CENP-A nucleosome itself contains multiple unique physical features (see "Section 1.4. Structural features of CENP-A nucleosomes") that likely contribute to its extraordinary stability. In addition to the "intrinsic" features of CENP-A nucleosomes, there are "extrinsic" factors, specifically its binding partners at the centromere, that play a key role in ensuring the stability of the CENP-A nucleosome. Chapters 2 and 3 describe in detail our efforts to attribute key roles to CENP-C and CENP-N in maintaining centromere identity.

Beyond the stable maintenance of existing CENP-A molecules at the centromere, the deposition of newly synthesized CENP-A molecules into centromeric chromatin, which occurs once in each cell cycle, is a crucial event for the propagation of centromere location. This chapter summarizes the machinery

responsible for depositing CENP-A, as well as the intricate regulatory mechanisms to ensure accurate deposition.

1.6.1 THE MACHINERY FOR DEPOSITING CENP-A

Whereas canonical H3.1-containing nucleosomes is assembled by chromatin assembly factor-1 (CAF-1), CENP-A assembly uses a CENP-A-selective histone chaperone, the Holliday junction recognition protein (HJURP), which was first identified from isolation of the soluble, yet-to-be loaded pool of CENP-A/H4 (the "prenucleosomal pool"); HJURP directly interacts with CENP/H4 in the pre-nucleosomal pool, and deposits CENP-A/H4 into centromeric nucleosomes in a cell cycle-dependent manner (Dunleavy et al., 2009; Foltz et al., 2009). HJURP is sufficient for assembling CENP-A nucleosomes onto plasmid DNA *in vitro* (Barnhart et al., 2011). HJURP is related to the yeast centromeric protein Scm3, which assembles Cse4 (the yeast CENP-A homologue) into centromeric chromatin (Camahort et al., 2007; Cho and Harrison, 2011; Mizuguchi et al., 2007; Stoler et al., 2007). A stretch of approximately 80 amino acids at the N-terminus of HJURP, a region that shares the most homology with the yeast Scm3 protein, is necessary and sufficient for binding CENP-A, and is called the CENP-A binding domain (CBD) (Shuaib et al., 2010). Specifically, HJURP interacts with the CENP-A targeting domain (Foltz et al., 2009). A crystal structure of HJURP CBD with CENP-A/H4 revealed that the C-terminal β-sheet domain of HJURP CBD caps the DNA-binding region of CENP-A/H4 and prevents it from spontaneous association with DNA (Hu et al., 2011). Furthermore, while the surface-exposed residues within the CATD determine specificity for HJURP recognition, and upon binding to HJURP, the contact points adjacent to the CATD serve to transmit stability throughout the histone fold domains of CENP-A and H4, suggesting that HJURP not only acts to shield CENP-A/H4 from nonspecific aggregation with nucleic acids prior to deposition, but also stabilizes the folded state of CENP-A/H4 (Bassett et al., 2012). Upon CENP-A assembly, HJURP forms an octameric CENP-A nucleosome from individual CENP-A/H4 heterodimers, and it has been reported that HJURP exists as a homodimer

through its C-terminal domain in the soluble prenucleosomal complex, as well as at chromatin as new CENP-A is assembled, thus providing a mechanism for how the assembly machinery assembles an octameric CENP-A nucleosome (Zasadzińska et al., 2013).

In addition to the CENP-A chaperone, HJURP, deposition of CENP-A in vertebrates requires Mis18 proteins. Whereas *S. pombe* has one Mis18 protein, vertebrates have two Mis18 paralogues, Mis18α and Mis18β, and additionally, humans possess the Mis18 binding protein (called Mis18BP1, also KNL2), which is required for recruiting Mis18 to centromeres (Fujita et al., 2007; Maddox et al., 2007). Mis18 has been suggested to alter histone modifications and the methylation status of centromeric chromatin to "prime" the centromere for CENP-A recruitment (Fujita et al., 2007; Kim et al., 2012). Mis18 is known to be required for the recruitment of HJURP to centromeres for CENP-A assembly (Moree et al., 2011; Perpelescu et al., 2015). Specifically, the Mis18α and β form a heterotetramer through their C-terminal coiled-coil domains, and these domains interact with HJURP directly to recruit it to centromeres; as HJURP binds, it disrupts the Mis18α-β heterotetramer and removes Mis18α from centromeres, which likely removes Mis18 from centromeres (Nardi et al., 2016). Therefore, Mis18 and HJURP work in a concerted fashion to assemble CENP-A at the centromere (summarized in **Figure 6**).

In order for centromere location to be accurately propagated from one cell division to the next, assembly of new CENP-A nucleosomes must occur at locations of existing CENP-A nucleosomes on the chromatin. It has been reported that Mis18 plays a direct role in specifying the location of nascent CENP-A assembly, by directly associating with the critical member of the CCAN, CENP-C (Dambacher et al., 2012; Moree et al., 2011). Specifically, the Mis18 complex contains two CENP-C binding domains, one on Mis18β and one on Mis18BP1, which are combinatorially required to generate robust centromeric localization of the Mis18 complex (Stellfox et al., 2016).

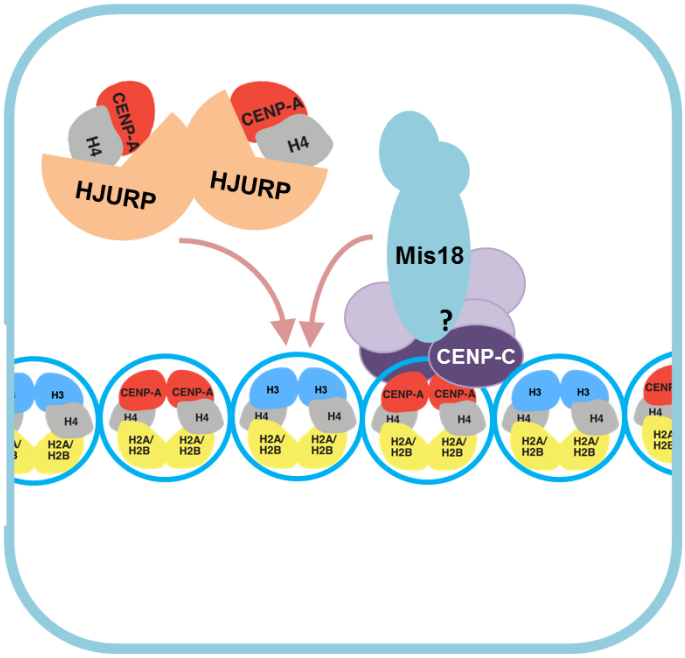


Figure 6. Summary of the players in the assembly of nascent CENP-A nucleosomes at the centromere.
HJURP exist as a dimer with the CENP-A/H4 heterotetramer in the "prenucleosomal complex", and the Mis18 complex recruits HJURP to centromeres and also directly binds to CENP-C.

1.6.2 Regulation of CENP-A assembly

Unlike canonical H3.1, whose assembly is coupled to DNA replication in S phase, replication of centromeric chromatin is uncoupled from centromeric DNA replication, and in metazoans is restricted to mitosis/early G1 phase (Jansen et al., 2007; Shelby et al., 2000) (Summarized in **Figure 7**). CENP-A assembly depends on passage through mitosis (Jansen et al., 2007). This cell-cycle coupled assembly of CENP-A is regulated by cyclin-dependent kinase (Cdk) activity: inhibiting Cdk1/2 in any phase of the cell cycle is sufficient to induce precocious CENP-A assembly (Silva et al., 2012). Specifically, CENP-A assembly machinery is present and poised for CENP-A assembly prior to mitosis, but sequestered away from centromeric chromatin until mitotic exit, and only allowed to assemble CENP-A upon decline of Cdk1/2 activity in telophase/G1 phase (Silva et al., 2012). The targets of Cdk regulation within the CENP-A assembly machinery includes the Mis18 complex,

since upon mutating all 24 of the putative Cdk consensus sites with Mis18BP1, Mis18BP1 is targeted prematurely to centromeres (Silva et al., 2012). However, CENP-A is not assembled prematurely with this Mis18BP1 mutant, indicating that there are likely additional targets of Cdk regulation among the CENP-A assembly machinery. Indeed, it has been reported that three putative Cdk consensus sites within the centromere-targeting domain of HJURP are susceptible to Cdk regulation (Müller et al., 2014; Wang et al., 2014), and mutation of all three sites to alanine results in premature CENP-A assembly, but only in a minority of cells (Müller et al., 2014). Therefore, it is likely that the full set of Cdk targets on the CENP-A assembly machinery has not been uncovered, and Chapter 4 of this thesis centers on using a quantitative phosphoproteomics approach to comprehensive map the landscape of Cdk regulation on the CENP-A-specific chaperone, HJURP.

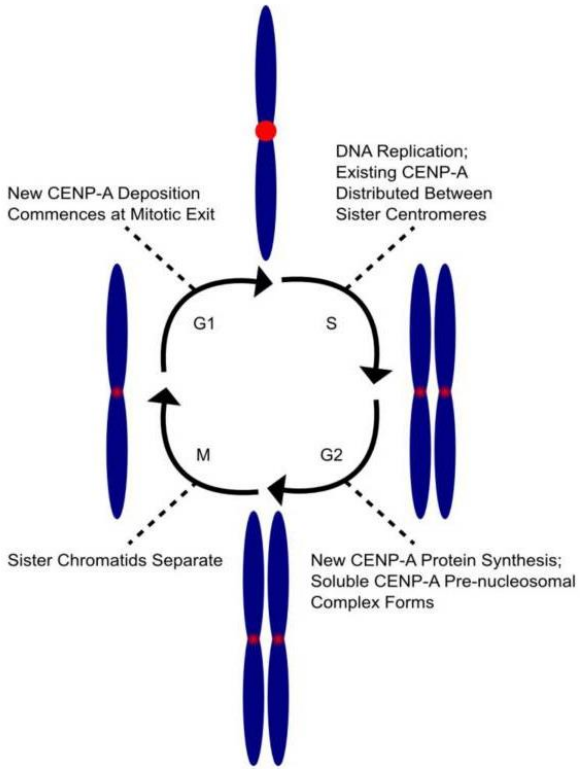


Figure 7. Model of CENP-A-containing chromatin throughout the cell cycle. Centromeric CENP-A levels fluctuate with the cell cycle. Prior to S phase, CENP-A is fully loaded at the centromere, but upon replication, the number of CENP-A molecules present on each daughter strand are reduced to half per centromere, since no new CENP-A is added. During G2, new CENP-A is synthesized and

assembles in a soluble prenucleosomal complex with its binding partner H4 and its chaperone HJURP, but is not loaded onto centromeres until G1. Therefore, cells progress with half-loaded centromeres through mitosis, until late anaphase/telophase when CENP-A is deposited by HJURP to restore CENP-A levels. Reproduced from (Falk and Black, 2012)

In addition to temporal regulation by Cdk1/2, initiation of CENP-A assembly has also been shown to depend on polo-like kinase 1 (Plk1), which acts as a licensing kinase to bind to and phosphorylate Mis18, and thereby promotes its centromeric localization (McKinley and Cheeseman, 2014). Plk1 activity is required for CENP-A assembly, more specifically for the localization of Mis18 to G1 centromeres (McKinley and Cheeseman, 2014). Therefore, regulation of CENP-A assembly by Cdk1/2 and Plk1 constitute a two-step paradigm.

Although it has been convincingly shown by multiple lines of evidence that CENP-A assembly is temporally restricted to the telophase/G1 phase of the cell cycle, it is still unclear why such regulation is necessary—in other words, what (if any) consequences there are for the cell if CENP-A assembly were to be forced to occur outside of telophase/G1. Even though one report claimed that CENP-A assembly could be forced to occur precociously, in G2 (Müller et al., 2014), it offered no insight into the mitotic consequences for those cells. Another report constitutively tethered the Mis18α subunit to the centromere by fusing it to the C-terminal domain of CENP-C, thereby bypassing both the Cdk regulation of the Mis18 complex and Plk1 licensing, and found that this resulted in severe mitotic defects, including misaligned chromosomes and multipolar spindles (McKinley and Cheeseman, 2014). However, it is unclear whether the severe mitotic defects could be derived from another outcome of the constitutive centromeric presence of the CENP-C-Mis18α fusion protein.

Since new CENP-A nucleosomes are not deposited until telophase/G1, centromeric chromatin could contain “gaps” after DNA replication has occurred, but before the full complement of CENP-A nucleosomes are re-acquired at the next telophase/G1. It has been proposed with evidence from stretched chromatin fibers and labeled H3.3 that H3.3 could be deposited in those "gaps", thus acting as a

"placeholder" after the dilution of CENP-A in S phase (Dunleavy et al., 2011). Indeed, there exists a crystal structure of a heterotypic nucleosome, containing one copy of CENP-A and one copy of H3.3, which still maintains CENP-C binding on only the CENP-A side of the nucleosome (Arimura et al., 2014). It is still unclear whether the loss of CENP-C from one side of the CENP-A/H3.3 heterotypic nucleosome could be relevant for S phase, and furthermore, it is unknown how H3.3 could be evicted from centromeric chromatin upon deposition of CENP-A in the next telophase/G1 phase.

Furthermore, CENP-A assembly is not only temporally restricted by the aforementioned mechanisms, but could also be susceptible to mechanisms of spatial restriction. The typical human centromere contains ~400 molecules of CENP-A, which is only ~4% of all centromeric nucleosomes (Bodor et al., 2014). Even this low concentration of centromeric CENP-A at physiologic centromeres is likely to be in excess of what is necessary for chromosome segregation, since cells with as little as 1% of endogenous CENP-A levels can still partially recruit kinetochore proteins and direct chromosome segregation, and disruption of kinetochore nucleation requires almost complete loss of CENP-A (Fachinetti et al., 2013). Therefore, the CENP-A nucleosomes present at the centromere is likely an ample pool, of which only a subset is required to nucleate centromere and kinetochore complexes (Bodor et al., 2014). There are likely regulatory mechanisms in place to limit the total protein level of CENP-A, since exogenous CENP-A expression in human cells often leads to downregulation of endogenous CENP-A (Falk et al., 2015; Jansen et al., 2007). A massive overexpression of CENP-A can result in ectopic deposition of CENP-A into chromosome arms (Falk et al., 2015; Gascoigne et al., 2011a; Van Hooser et al., 2001), possibly through a DAXX-mediated mechanism (Lacoste et al., 2014). Additionally, overexpression of HJURP in DT40 chicken cells results in centromere expansion surround natural centromeres, and ectopic HJURP localization onto artificial centromeres also results in centromere expansion (Perpelescu et al., 2015). Therefore, there may exist mechanisms to tightly tune the total protein levels of CENP-A and its chaperone, HJURP, to prevent ectopic CENP-A deposition to non-centromeric regions of chromatin.

1.7. SUMMARY

Kinetochores are assembled upon chromatin defined by the presence of CENP-A-containing nucleosomes, which are the epigenetic hallmark of centromere location. Understanding the mechanism by which CENP-A specifies centromere identity is paramount to the accurate propagation of genetic material over cell and organisms generations. Much has been learned since centromeres were first isolated in budding yeast in 1980, but there are crucial gaps in our understanding of the centromere. For example, although it is known that CENP-A molecules are remarkably stable, a characteristic that is critical for the stable transmission of centromere location, it is unknown how such stability is achieved. Chapter 2 focuses on the role of CENP-C, an essential centromeric protein and binding partner of CENP-A, in ensuring centromere stability by modulating physical properties of the CENP-A nucleosome. Chapter 3 builds upon the insights from Chapter 3, and focuses on our efforts in building a core centromeric nucleosome complex (CCNC) and uncovering its biophysical properties, and through doing so, unveils a definitive model for the stability of centromeric CENP-A nucleosome.

CHAPTER 2: CENP-C RESHAPES AND STABILIZES CENP-A NUCLEOSOMES AT THE CENTROMERE

Chapter 2 is based on the following publication:

> Falk, S.J.*, Guo, L.Y.*, Sekulic, N.*, Smoak, E.M.*, Mani, T., Logsdon, G.A., Gupta, K., Jansen, L.E.T., Van Duyne, G.D., Vinogradov, S.A., Lampson, M.A., and Black, B.E. (2015). CENP-C reshapes and stabilizes CENP-A nucleosomes at the centromere. *Science* 348, 699–704.
>
> *Co-first authors

2.1. ABSTRACT

Inheritance of each chromosome depends upon its centromere. A histone H3 variant, centromere protein A (CENP-A), is essential for epigenetically marking centromere location.We find that CENP-A is quantitatively retained at the centromere upon which it is initially assembled. CENP-C binds to CENP-A nucleosomes and is a prime candidate to stabilize centromeric chromatin. Using purified components, we find that CENP-C reshapes the octameric histone core of CENP-A nucleosomes, rigidifies both surface and internal nucleosome structure, and modulates terminal DNA to match the loose wrap that is found on native CENP-A nucleosomes at functional human centromeres. Thus, CENP-C affects nucleosome shape and dynamics in a manner analogous to allosteric regulation of enzymes. CENP-C depletion leads to rapid removal of CENP-A from centromeres, indicating their collaboration in maintaining centromere identity.

2.2. INTRODUCTION

Centromeres direct chromosome inheritance at cell division, and nucleosomes containing a histone H3 variant, centromere protein A (CENP-A), are central to current models of an epigenetic program for specifying centromere location (Black and Cleveland, 2011). The centromere inheritance model in

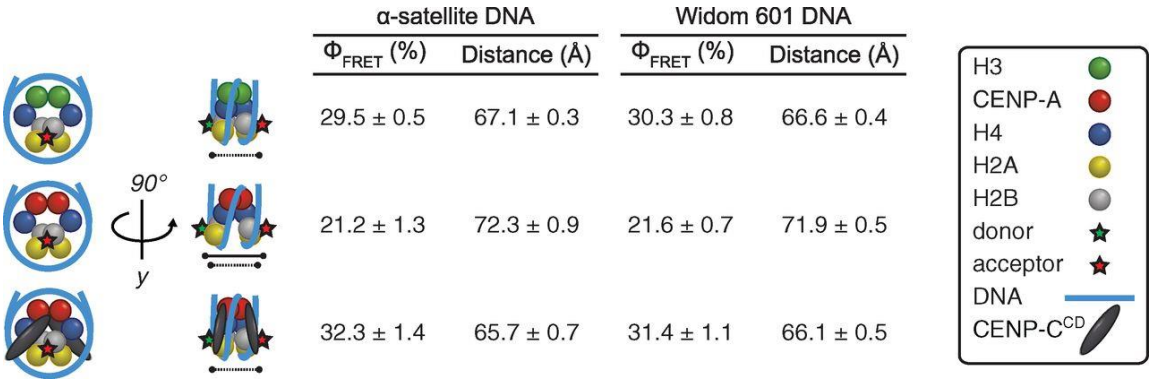
metazoans suggests that the high local concentration of preexisting CENP-A nucleosomes at the centromere guides the assembly of nascent CENP-A, which occurs once per cell cycle after mitotic exit. This model predicts that after initial assembly into centromeric chromatin, CENP-A must be stably retained at that centromere; otherwise, centromere identity would be lost before the next opportunity for new loading in the next cell cycle. Here, we use biochemical reconstitution to measure the shape and physical properties of CENP-A nucleosomes with and without its close binding partner, CENP-C, and combine these studies with functional tests that reveal the mechanisms underlying the high stability of centromeric chromatin.

2.3. RESULTS

2.3.1 CENP-C BINDING ALTERS THE SHAPE OF THE CENP-A NUCLEOSOME

CENP-C recognizes CENP-A nucleosomes via a region termed its central domain (amino acids 426 to 537; CENP-C^{CD})(Carroll et al., 2010; Kato et al., 2013). We first considered how CENP-C^{CD} may affect the overall shape of the CENP-A–containing nucleosome using an intranucleosomal fluorescence resonance energy transfer (FRET)–based approach. We designed an experiment to measure FRET efficiency, ΦFRET, between two fluorophores on defined positions on the H2B subunits of CENP-A nucleosomes in the absence or presence of CENP-C^{CD} and then used these measurements to calculate intranucleosomal distances **(Figure 8,** and *Fig. S1 in Falk et al., 2015*). The H2B distances for CENP-A nucleosomes in the absence of CENP-CCD are ~5 Å farther apart than expected from their crystal structure (PDB ID 3AN2)(Tachiwana et al., 2011), indicating that CENP-A–containing nucleosomes in solution prefer a histone octamer configuration not captured in the crystal structure. It is likely that CENP-A nucleosomes sample both conformations in solution, with crystal contacts stabilizing the form that was reported (Tachiwana et al., 2011). In contrast to CENP-A nucleosomes, conventional

nucleosomes have smaller H2B distances in solution (**Figure 8**) that are consistent with their crystal structure (Luger et al., 1997). Separation of H2A/H2B dimers from each other is consistent with a nucleosome model based on rotation of the CENP-A/CENP-A′ interface in $(CENP\text{-}A/H4)_2$ heterotetramers (Sekulic et al., 2010). Upon binding of CENP-C^{CD}, with the known stoichiometry of two CENP-C^{CD} molecules per nucleosome (Kato et al., 2013), the H2A/H2B distances shorten to ones that are nearly identical to those in conventional nucleosomes (**Figure 8**). The differences we observed between H3 nucleosomes, CENP-A nucleosomes, and CENP-A nucleosomes in a complex with CENP-C^{CD} are found using either the human α-satellite DNA sequence that corresponds to the most heavily occupied site at centromeres (Hasson et al., 2013) or the completely synthetic "601" nucleosome positioning sequence (Lowary and Widom, 1998) (**Figure 8**).



	α-satellite DNA		Widom 601 DNA	
	Φ_{FRET} (%)	Distance (Å)	Φ_{FRET} (%)	Distance (Å)
	29.5 ± 0.5	67.1 ± 0.3	30.3 ± 0.8	66.6 ± 0.4
	21.2 ± 1.3	72.3 ± 0.9	21.6 ± 0.7	71.9 ± 0.5
	32.3 ± 1.4	65.7 ± 0.7	31.4 ± 1.1	66.1 ± 0.5

Figure 8. CENP-A nucleosomes have a conventional shape only upon CENP-C^{CD} binding.
Calculated FRET efficiencies (Φ_{FRET}) and distances between donor and acceptor fluorophores on H2B S123C for the indicated nucleosomes on either α-satellite or Widom 601 DNA. Data are shown as the mean ± SEM of three independent nucleosome reconstitutions.

2.3.2 CENP-C BINDING RIGIDIFIES SECONDARY STRUCTURE AT THE INTERIOR OF THE CENP-A NUCLEOSOME

The shape change that we measure within the nucleosome upon CENP-C^{CD} binding most likely occurs through rotation at the four-helix bundles between histone dimer pairs within the octameric core, with interhistone contacts being stabilized or destabilized depending on the preference for rotational state. We tested this

prediction using hydrogen/deuterium exchange-mass spectrometry (HXMS). Strong protection of CENP-A nucleosomes (**Figure 9**, **Figure 10D**, and **Figure 11**) is conferred by CENP-C^{CD} binding on peptides spanning helices that are predicted (Kato et al., 2013) to contact it (i.e., the α3 helix and C-terminal residues of CENP-A, the α2 helices of both H4 and H2A, and regions of H2A encompassing its acidic patch residues). In addition to the surface changes induced by CENP-CCD, there are internal changes to the nucleosome that we measure by HX (**Figure 9A,B,** *and movie S1* in *Falk et al., 2015*) that are consistent with the change in nucleosome shape that we observed by FRET (**Figure 8**). The separation of H2A/H2B dimers in CENP-A nucleosomes lacking CENP-C^{CD} (**Figure 8**) is predicted to weaken an internal, intermolecular β sheet that serves as the physical connection between the H2A subunit on one face of the nucleosome and the H4 subunit on the opposite face. When CENP-C^{CD} binds to the CENP-A nucleosome, peptides spanning the corresponding β-sheet residues of both H2A and H4 exhibit extra protection from HX by 1 to 2 deuterons, where the same level of HX takes 5 to 10 times as long to occur than in CENP-A nucleosomes lacking CENP-C^{CD} (**Figure 9, Figure 12**, **Figure 13**)

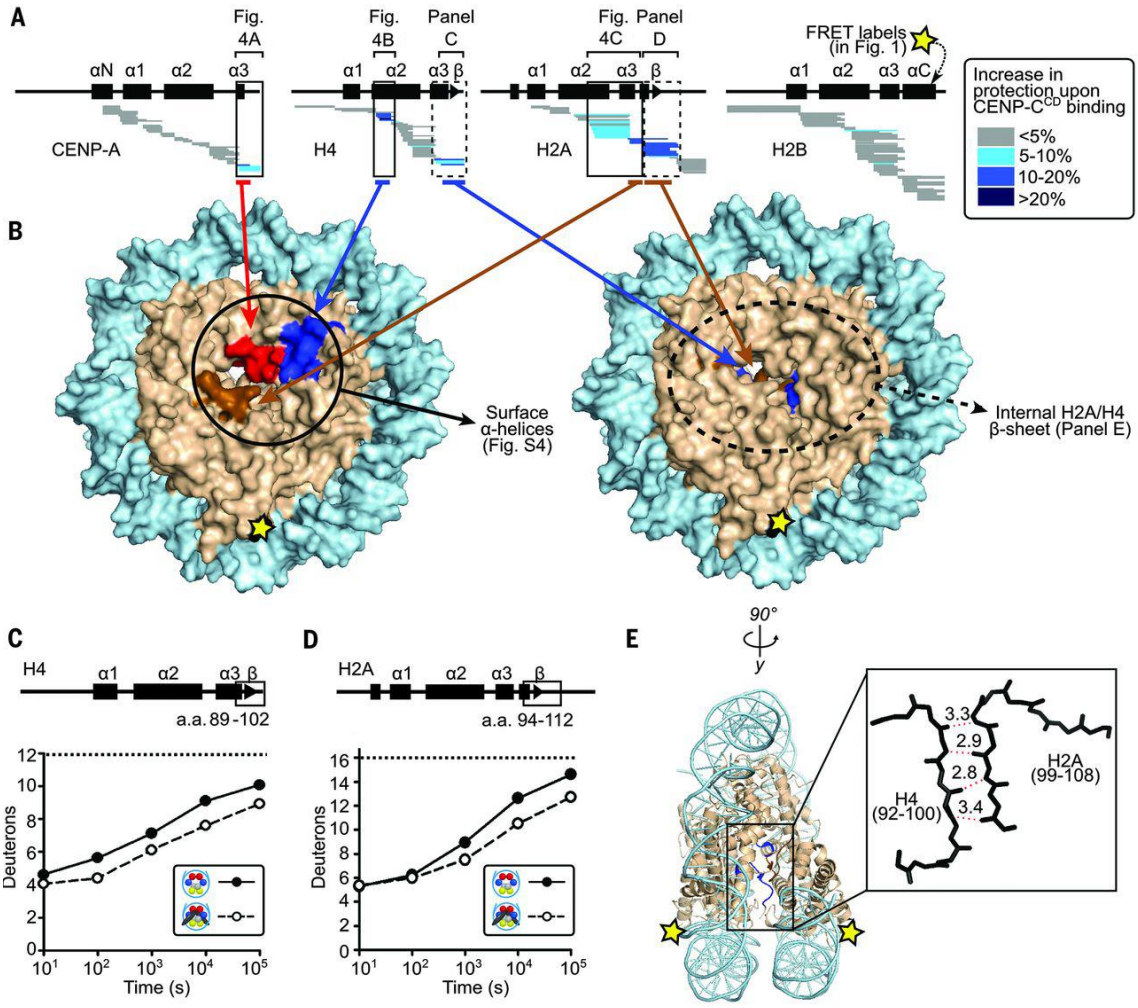


Figure 9. CENP-C^{CD} rigidifies CENP-A nucleosomes.
(A) HXMS of all histone subunits of the CENP-A nucleosome from a single time point (10^4 s; see all time points in
Figure **10**). Each horizontal bar represents an individual peptide, and peptides are placed beneath schematics of secondary structural elements. **(B)** Regions showing substantial protection from HX mapped onto the structure of the CENP-A nucleosome (PDB ID 3AN2). **(C and D)** Comparison of representative peptides spanning the β-sheet region in histone H4 and histone H2A over the time course. The maximum number of deuterons possible to measure by HXMS for each peptide is shown by the dotted line. **(E)** The internal H4/H2A interface mapped (see **Figure 12**) onto the canonical nucleosome crystal structure (PDB ID 1KX5).

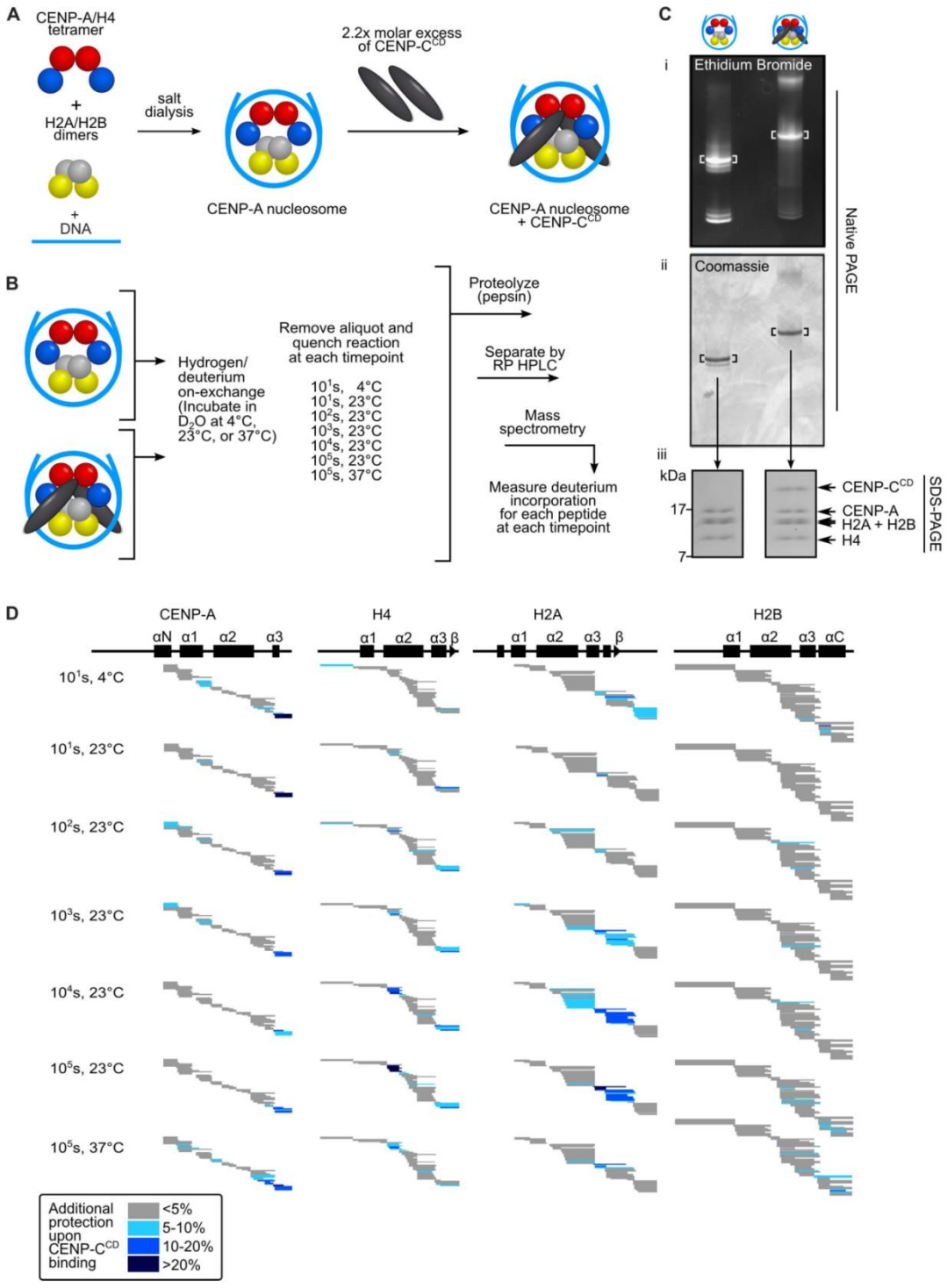


Figure 10. CENP-C^{CD} binding induces additional protection from HX at multiple regions within the CENP-A nucleosome.

(A,B) Experimental scheme for HXMS. HXMS is an approach that measures the exchange of
amide protons along the polypeptide backbone of the histones with deuterons from heavy water (D_2O) in the exchange buffer (Englander, 2006). Amide protons engaged in hydrogen bonds inside α-helices and the interior of β-sheets are protected from HX, exchanging only upon transient unfolding of structure, and thus this technique detects stabilization/destabilization of these structures. **(C)** CENP-A nucleosome alone and in complex with CENP-C^{CD} as assessed by native gel stained with ethidium bromide (i) or Coomassie Blue (ii). Bands from native gel were excised and assessed by SDS-PAGE (iii). **(D)** Regions of CENP-A nucleosomes where CENP-C^{CD} binding leads to additional protection from HX. The level of protection added by the presence of CENP-C^{CD} is determined by subtracting the level of HX of the CENP-A nucleosome alone from that of CENP-A nucleosomes bound by CENP-C^{CD}. Each horizontal bar represents an individual peptide, placed beneath schematics of secondary structural elements of the CENP-A nucleosome. Peptides that exhibit additional protection from HX upon CENP-C^{CD} binding are colored in blue, with shading according to the legend. It is notable that prior to nucleosome assembly, CENPA/H4 is protected from HX by HJURP (Bassett et al., 2012), and after chromatin assembly CENP-A/H4, as well as the nucleosome subunit, H2A, is protected from HX by CENP-C^{CD} (this study). HJURP is thought to protect CENP-A/H4 during the long cell cycle window between when CENP-A is expressed in late S-phase (Shelby et al., 2000) to when it is assembled at centromeres in the subsequent G1 (Jansen et al., 2007). Our studies strongly suggest that CENP-C is important for stabilizing CENP-A nucleosomes at the centromere to maintain centromere identity.

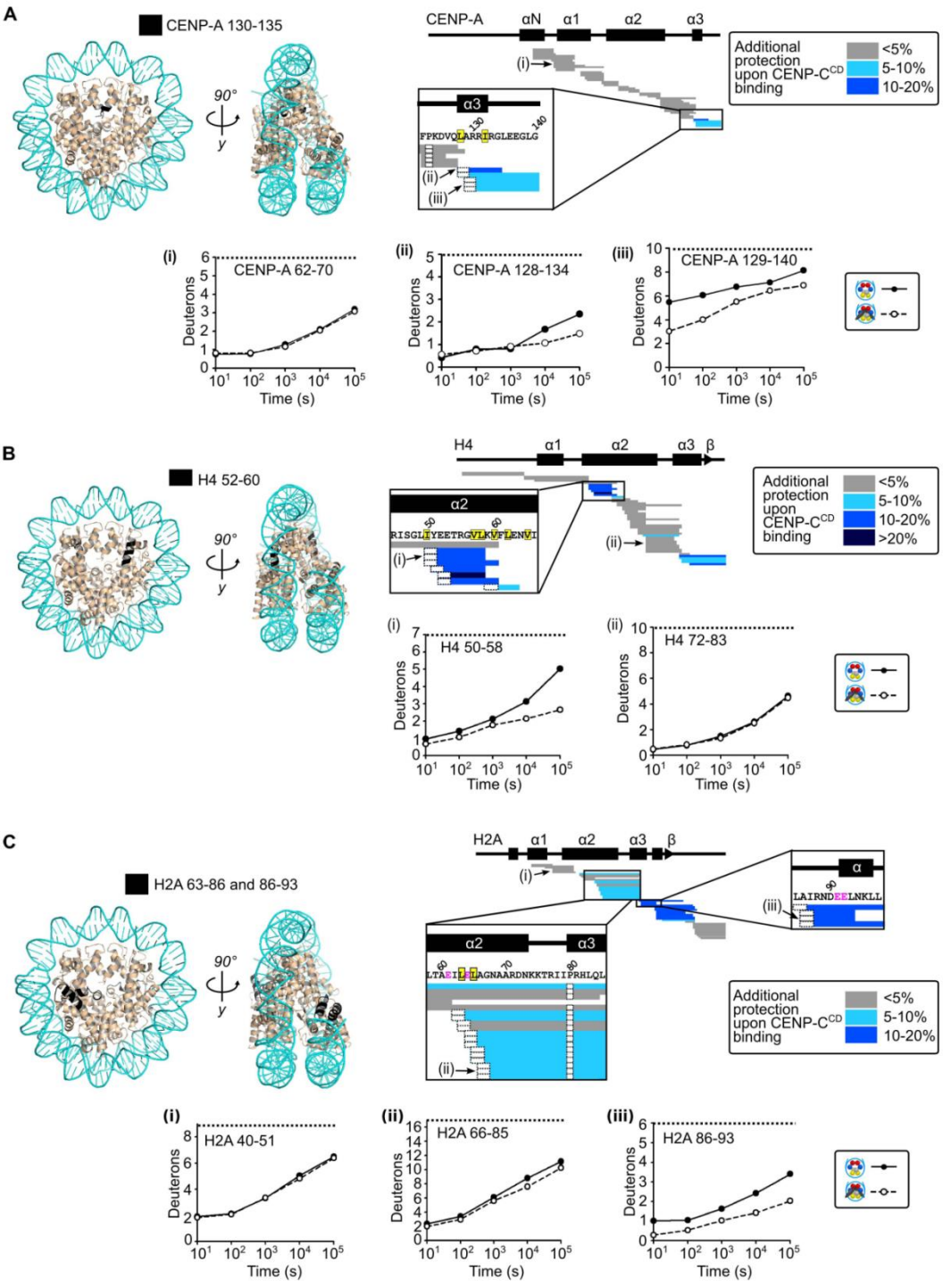


Figure 11. Regions on the surface of the nucleosome that exhibit additional protection from HX upon CENP-C^{CD} binding.
(A) CENP-A 130-140, **(B)** H4 52-60, and **(C)** H2A 63-93, mapped in black onto the structure of the CENP-A nucleosome (PDB 3AN2). Horizontal blocks represent peptides from CENP-A containing nucleosomes, placed beneath schematics showing

locations of α-helices and β-sheets of each histone. Multiple charge states were detected for a subset of peptides, each represented by its own block. Representative peptides from these regions exhibiting protection or from flanking regions exhibiting no extra protection are plotted as the number of deuterons exchanged at each time point. The maximum number of deuterons for each peptide possible to measure by HXMS is shown by the dotted line. Residues highlighted in yellow are those that exhibited methyl chemical shift perturbation in the NMR model of $H3_{1-132}$LEEGLG nucleosome bound to CENP-C^{CD} (Kato et al., 2013). The glutamate residues in pink on H2A are acidic patch residues. Note that PDB 3AN2 does not contain CENP-A residues 136-140, thus only 130-135 are colored.

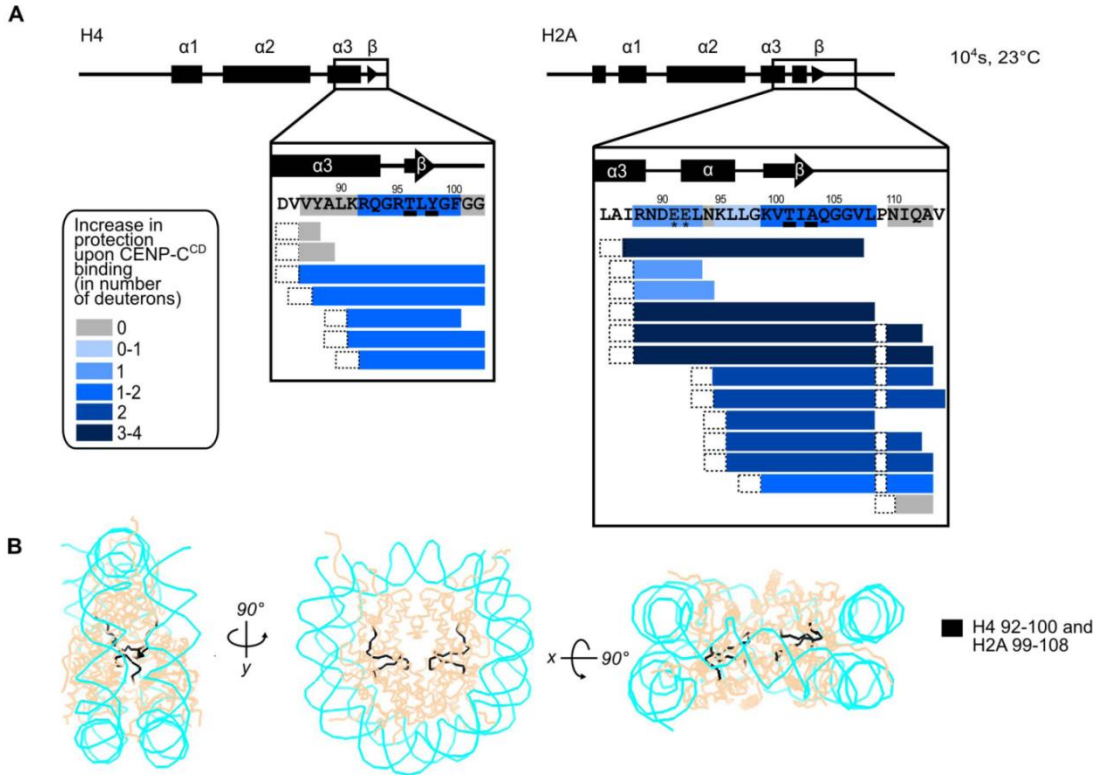


Figure 12. Mapping HX protection in the interior of the nucleosome when CENP-C^{CD} binds to the surface of CENP-A nucleosomes.
(A) All peptides encompassing the β-sheet interface between H4 and H2A (the 10^4 s time point is shown) from which the minimal regions that exhibit protection is deduced to be H4 92-100 and H2A 99-108 based on partially overlapping peptide HX data. Note that protection of H2A residues encompassing the β-sheet can be distinguished from those of surface contacts (i.e., acidic patch residues marked with asterisks). Underlined residues contain the backbone amide hydrogens that engage in hydrogen bonding within the β-sheet. The first two residues of each peptide and prolines are boxed in dashed black lines because exchange of the first two backbone amide protons cannot be measured (42) and prolines lack amide protons. **(B)** These minimal regions are colored in black in the nucleosome structure (PDB 1KX5) and are shown in various orientations. The H2A/H4 β-sheet is buried in the interior of

the nucleosome and is inaccessible to CENP-C^{CD} at the surface of the nucleosome (**Figure 9B, Figure 11**, and *movie S1 in Falk et al., 2015*), so we conclude that it is stabilized concomitantly with the shape change that brings H2A/H2B dimers together.

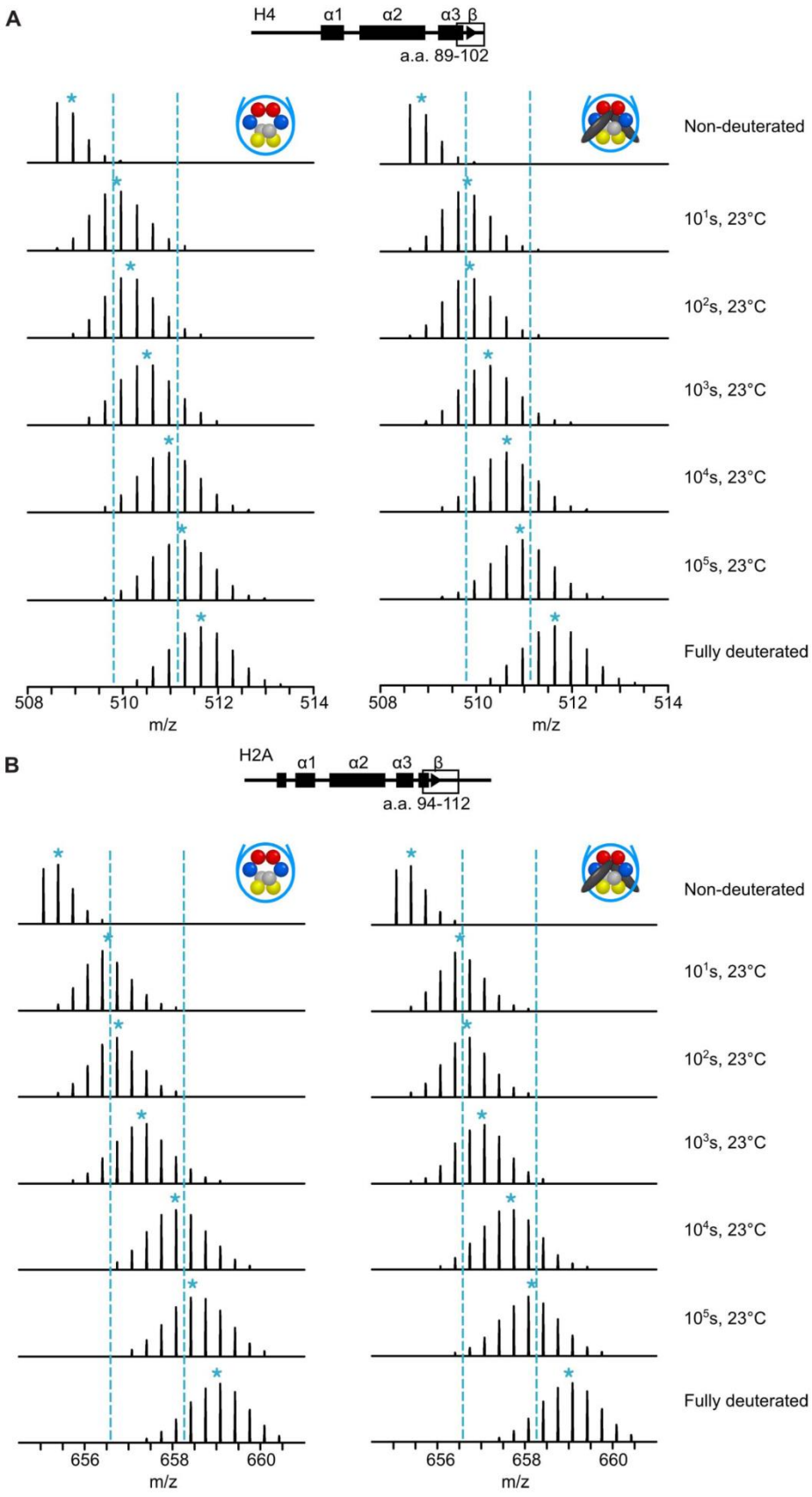

A
H4
α1
α2
α3
β
a.a. 89-102
Non-deuterated
10¹s, 23°C
10²s, 23°C
10³s, 23°C
10⁴s, 23°C
10⁵s, 23°C
Fully deuterated
508
510
512
514
m/z
B
H2A
α1
α2
α3
β
a.a. 94-112
Non-deuterated
10¹s, 23°C
10²s, 23°C
10³s, 23°C
10⁴s, 23°C
10⁵s, 23°C
Fully deuterated
656
658
660
m/z

Figure 13. HXMS of peptides spanning the β-sheet at the interface in H4 **(A)** and H2A **(B).**
In each panel, the peptide is shown from a CENP-A nucleosome (left) and a CENP-A nucleosome when in complex with CENP-C^{CD} (right). Dotted blue lines serve as guideposts to highlight the differences in m/z shifts between the two samples. A blue asterisk denotes the centroid location of each peptide envelope, and the numerical value in blue indicates the centroid mass of the peptide envelope as determined by the ExMS data analysis software (Kan et al., 2011).

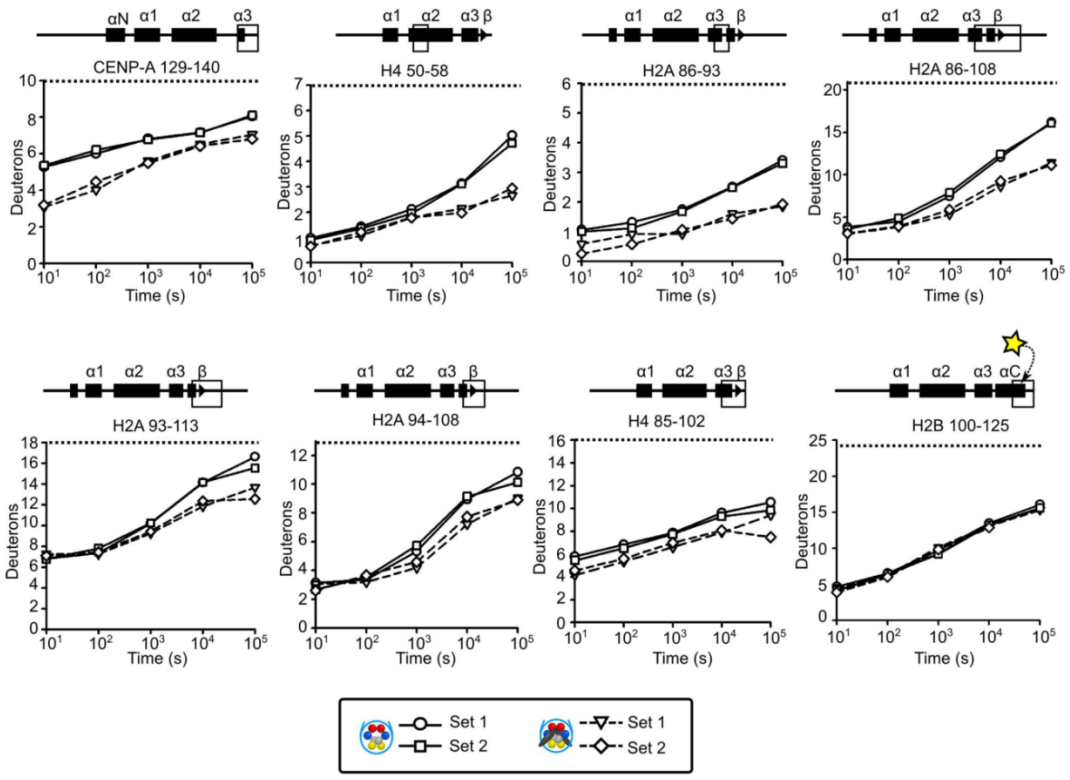


Figure 14. Faithful detection of 1-4 deuteron differences in HXMS of CENP-A nucleosomes with and without CENP-C^{CD} bound.
The indicated peptides are compared between two replicate datasets (set 1 and 2), which represent two independent CENP-C^{CD} purifications and two entirely independent nucleosome reconstitutions. Changes of 1-2 deuterons are well within the resolution of HXMS (Englander, 2006). The star symbol in the H2B diagram indicates the position of fluorophores used in FRET experiments.

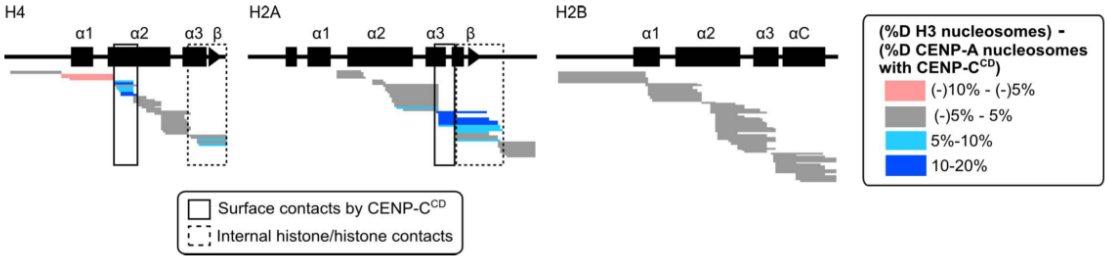


Figure 15. Comparison of HX behavior of canonical nucleosomes to CENP-A nucleosomes
bound by CENP-C^{CD}.
HXMS data with H3 nucleosomes at the 10^4 s time point showing the three histone subunits (H4, H2A, and H2B) common to the two types of nucleosomes. This is the same time point shown for CENP-A nucleosome comparisons in **Figure 9A**. Peptides spanning the H4/H2A interface show protection in CENP-A nucleosomes bound to CENP-C^{CD} as compared to H3 nucleosomes, but the protection is less pronounced than the additional HX protection to CENP-A nucleosomes conferred upon binding CENP-C^{CD} (see **Figure 9A**).

2.3.3 CENP-C ALTERS NUCLEOSOME TERMINAL DNA

Because CENP-C might also affect the extent that DNA wraps the nucleosomes, we reconstituted CENP-A nucleosomes using a 195–base pair (bp) DNA sequence from α-satellite DNA (Harp et al., 1996) that contains a contiguous sequence spanning the major binding site it occupies on human centromeres (Hasson et al., 2013) (**Figure 16A**). We first overdigested CENP-A nucleosomes and found very strong protection of 100 bp (*Fig. S9 in Falk et al., 2015*). Using a subsequent restriction digest of the 100-bp digestion product, we found that they were uniquely positioned, with their dyad precisely where the same-sized fragment previously mapped with native centromeric particles (Hasson et al., 2013)(*Fig. S9 in Falk et al., 2015*). CENP-A–containing nucleosomes have many discrete intermediate digestion products before the strongly protected 100-bp fragment is generated (**Figure 16A,B**, and *Fig. S10 in Falk et al., 2015*). When CENP-C^{CD} is bound, digestion products larger than a nucleosome core particle [e.g., >145 bp, where DNA strands could cross at ~165 bp for conventional nucleosomes (Kornberg, 1977)] are missing

at early time points (**Figure 16B**). This suggests that when CENP-C^{CD} binds to the nucleosome, the DNA above the dyad rarely crosses, as it would normally cross for conventional nucleosomes. Second, digestion to the 100-bp final fragment proceeds more quickly (**Figure 16B**). Thus, transient unwrapping of two helical turns (i.e., ~20 bp) from each terminus of the nucleosome is enhanced when CENP-C^{CD} is bound.

To determine whether CENP-C^{CD} binding leads to a steady-state structural change of nucleosomal DNA, we used small-angle neutron scattering (SANS) with contrast variation. When CENP-C^{CD} binds to CENP-A nucleosome core particles, the distance distribution profiles reflecting the shape in solution substantially redistribute for both the protein- and DNA-dominated measurements (**Figure 16C**, *Fig. S11 and Table S3 in Falk et al., 2015*). The increase in larger interatomic vectors for the protein component is expected to accompany an additional component (CENP-C^{CD}). The pronounced redistribution of vectors to both smaller and larger distances in DNA-dominated scattering when CENP-C^{CD} is bound is attributed to compaction of the nucleosome core (smaller vectors) and opening of the nucleosome terminal DNA when CENP-C^{CD} is bound (larger vectors).

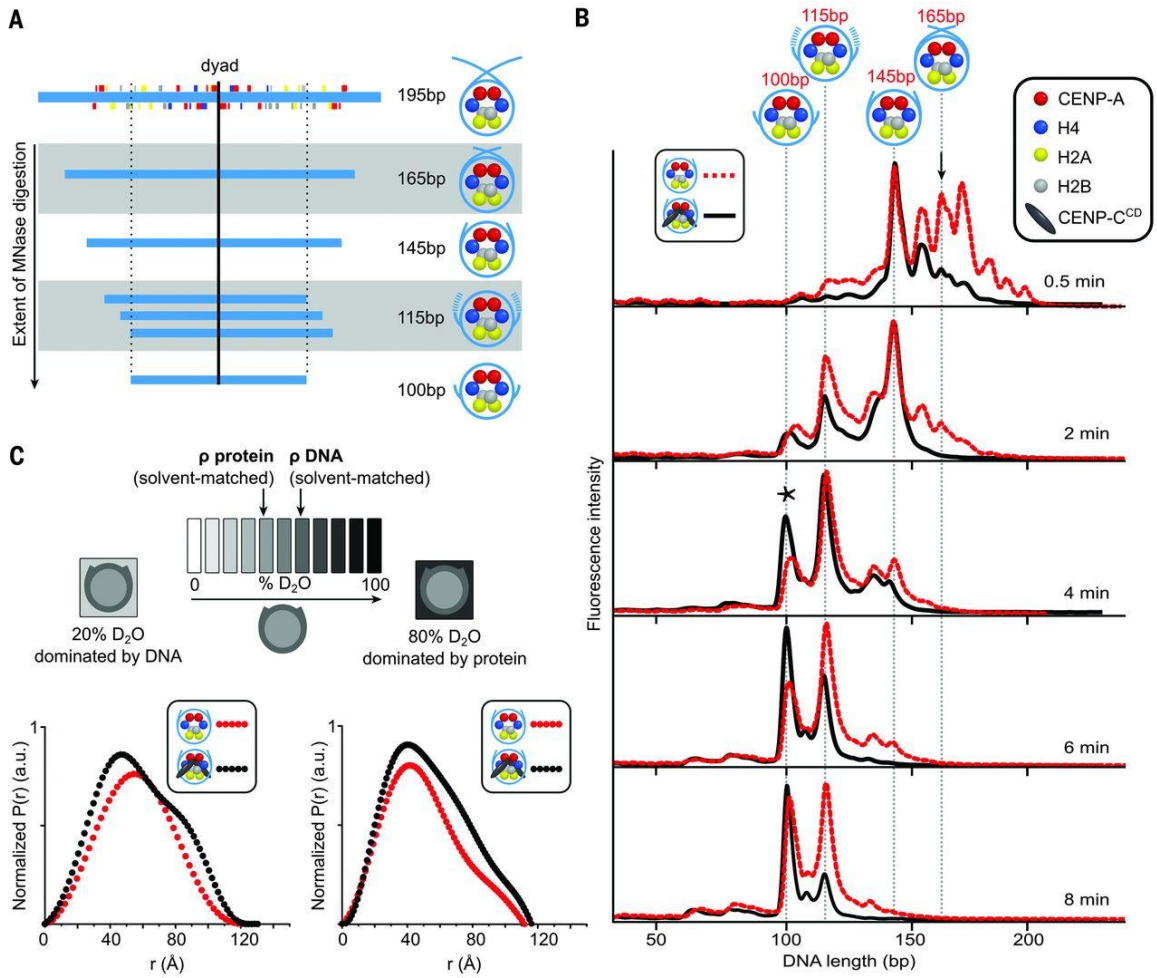


Figure 16. Alterations in the nucleosome terminal DNA upon CENP-C^{CD} binding. **(A)** Major micrococcal nuclease (MNase)–digested DNA fragments observed for CENP-A nucleosomes assembled on its native centromere sequence. **(B)** MNase digestion profiles of CENP-A nucleosomes in the absence (red) and presence (black) of CENP-C^{CD}. The black arrow (0.5 min) points to the 165-bp peak (DNA crossed at the dyad). The asterisk (4 min) denotes the final 100-bp peak. **(C)** Scheme of SANS contrast variation experiment together with paired distance distribution curves for CENP-A nucleosomes alone (red) and bound by CENP-C^{CD} (black) in the indicated SANS contrast variation conditions.

2.3.4 CENP-C STABILIZES CENP-A AT CENTROMERES

We took two complementary approaches in cells to determine whether CENP-A is stably retained at the centromere upon which it is initially deposited (see the legend for *Fig. S12 from Falk et al., 2015* that describes the motivation for these experiments). First, we used cell cycle–synchronized fluorescence pulse labeling of

CENP-A in “donor” cells and subsequent cell fusion with an “acceptor” cell line. The donor cells express SNAP-tagged CENP-A that has been pulse labeled with tetramethylrhodamine-Star (TMR*) to irreversibly label CENP-A (Jansen et al., 2007) before cell fusion. The acceptor cells express yellow fluorescent protein (YFP)–tagged CENP-A that is loaded at all centromeres, continuing even after fusion. At time points through the subsequent cell cycle (*Fig. S12 in Falk et al., 2015*) until the second mitosis (**Figure 17A**), we observed no detectable exchange of the TMR*-labeled donor CENP-A to the acceptor centromeres in a shared nucleoplasm. Quantitation of the fluorescence at each centromere in these heterokaryons yields a bimodal distribution. The donor centromere group with high TMR* and low YFP (**Figure 17B**, “x” symbols) has an average TMR* signal of 0.538 ± 0.005 (normalized arbitrary units where the maximal measured TMR* signal in each heterokaryon equals 1) (**Figure 17C**), whereas the acceptor centromere group with high YFP and low TMR* (**Figure 17B**, triangle symbols) has an average TMR* signal of 0.055 ± 0.005 (**Figure 17C**). These data indicate that once assembled at a centromere, an individual CENP-A molecule is stably maintained at that particular centromere.

As a complementary approach to test CENP-A stability at individual centromeres, we used a photoactivatable version of CENP-A (CENP-A-PAGFP). We induced expression of CENP-A to the extent that it is present at locations throughout the nucleus, but with clear enrichment at centromeres, and then activated a defined region of each cell nucleus (**Figure 17D**, 0 hours postphotoactivation). CENP-A-PAGFP signal is quantitatively retained at the activated centromeres and does not accumulate at unactivated centromeres (**Figure 17D,E**), indicating that there is negligible exchange between centromeres, consistent with our cell fusion results. In contrast, CENP-A-PAGFP signal in bulk chromatin decays, with about half of the protein removed by 8 hours after photoactivation.

To investigate whether CENP-C stabilizes CENP-A at centromeres, we combined SNAP labeling of CENP-A with CENP-C depletion (**Figure 17**), for which we generated a cell line with a chromosomally integrated, doxycycline-inducible CENP-C short hairpin RNA cassette. In our SNAP system, CENP-C depletion leads to a dramatic decrease in the retention over 24 hours of the existing pool of CENP-A at

centromeres (**Figure 17G,H**). Without CENP-C depletion, the average retention of CENP-A is slightly >100% (112% ± 63% SD), an increase that is explained by having a small pool of prenucleosomal CENP-A in the cell population that is labeled by the TMR* pulse and subsequently incorporated into centromeres. Nascent CENP-A deposition is also decreased when CENP-C is depleted (fig. **Figure 18C**)—consistent with its proposed role in the CENP-A assembly reaction (Erhardt et al., 2008; Moree et al., 2011)—but this would only affect incorporation of the small prenucleosomal pool in the CENP-A retention measurements (**Figure 17G,H**). Thus, our findings implicate CENP-C in stabilizing CENP-A nucleosomes at centromeres. We cannot rule out the possibility that removal of CENP-C in turn removes another centromere component that stabilizes CENP-A nucleosomes, but we favor the idea that CENP-C is the key molecule for stabilizing CENP-A nucleosomes based on its direct binding to it.

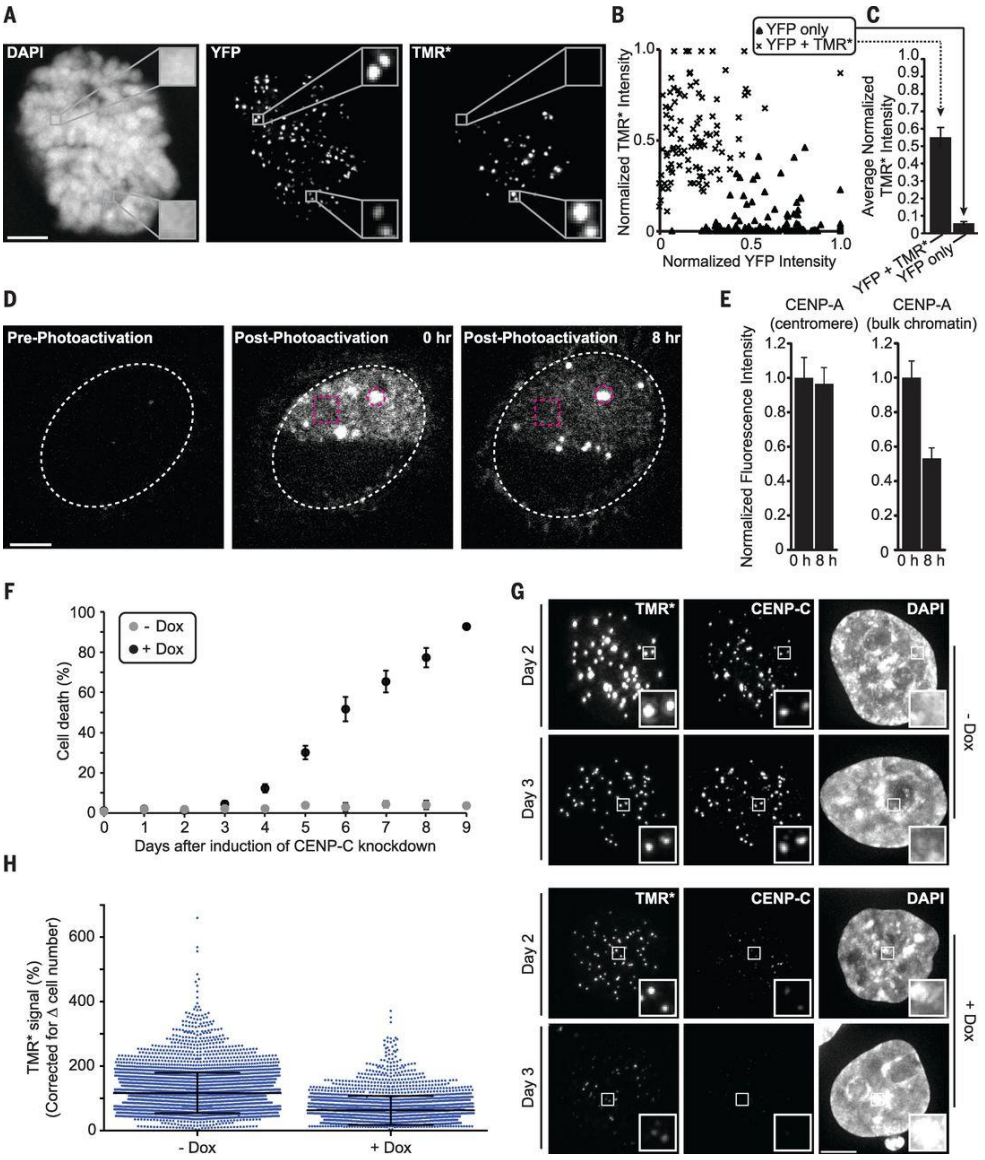


Figure 17. Depletion of CENP-C reduces the high stability of CENP-A at centromeres.

(A to C) Cells expressing SNAP-tagged CENP-A were pulse-labeled with TMR*, then fused with cells expressing YFP-tagged CENP-A. Representative images (A) show a cell in the second mitosis after fusion; insets show 3x magnification. X-means clustering was used to classify YFP only (triangles) or YFP and TMR* ("x" marks) centromeres (B), and mean (± SEM) TMR* intensity was calculated for each group (C). **(D and E)** Cells expressing high levels of CENP-A-PAGFP were photoactivated in bulk (box) and centromeric (circle) chromatin. Representative images (D) show a subset of centromeres in a single z section. Fluorescence intensity was quantified at 0 and 8 hours after photoactivation [(E), mean ± SEM]. **(F)** CENP-C knockdown begins causing cell death 4 days post-induction (mean ± SD). **(G and H)** Cells with (+ Dox) and without (– Dox) CENP-C depletion were pulse-labeled with TMR* (day 2), and the relative CENP-A-SNAP signals were analyzed (day 3). Quantification shows

CENP-A-SNAP signal retained at day 3 (>2500 centromeres plotted with mean ± SD). Scale bars, 5 μm.

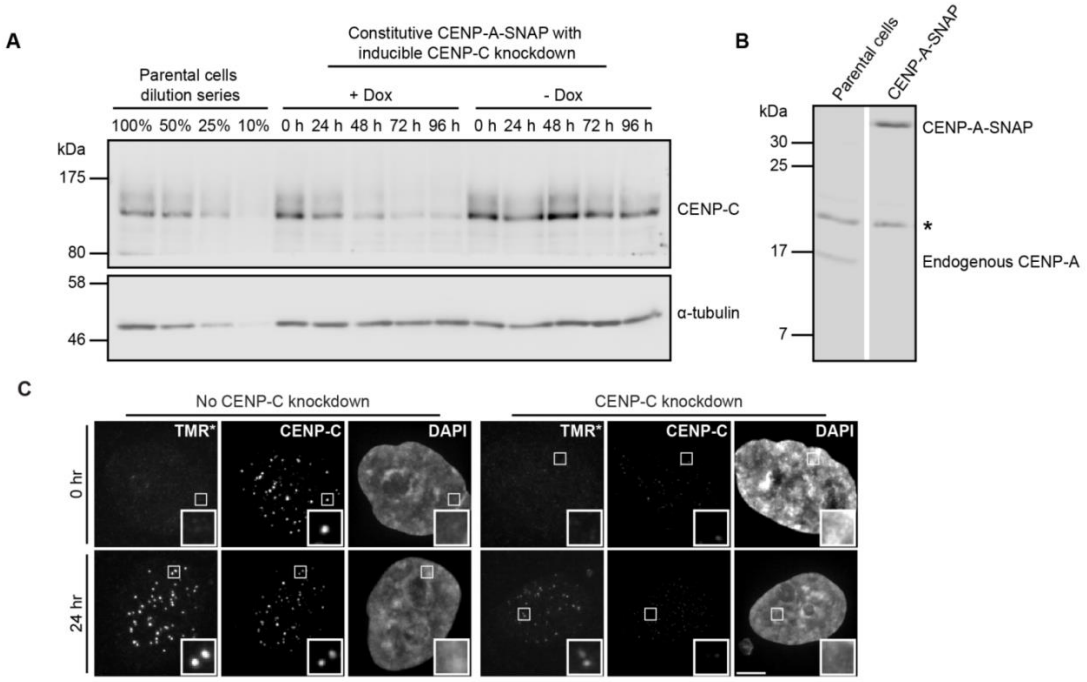


Figure 18. CENP-C knockdown effects on retention and assembly of CENP-A at the centromere.
(A) Immunoblot of CENP-C levels in inducible CENP-C knockdown cells at indicated timepoints after Dox addition, compared to the cells without Dox. Whole cell lysate dilutions from parental cells were used to measure the extent of CENP-C knockdown. α-tubulin levels were used as a loading control. CENP-C depletion requires several days prior to when cell death occurs **(Figure 17F)** following mitotic kinetochore failure (Fukagawa and Brown, 1997), so there is an experimental window of time in which we can test if CENP-A nucleosome retention persists after the majority of CENP-C protein has been depleted (**Figure 17F,G**). SNAP labeling of the existing pool of CENP-A (**Figure 17G** [Day 2]) combined with monitoring cell number allows one to account for the entire pool of CENP-A in the dividing cell population during the course of the experiment (Bodor et al., 2013). This approach also overcomes the limitation of the CENP-A-PAGFP approach (**Figure 17D**) where measurements beyond ~8 hr become problematic due to cell divisions. **(B)** Immunoblot of CENP-A levels in parental and CENP-A-SNAP cell lines. Asterisk denotes non-specific band. **(C)** We performed a quench-chase-pulse experiment where cells were synchronized using a double-thymidine block, pre-existing CENP-A was quenched with a non-fluorescent label, and the nascent pool of CENP-A was pulse-labeled with TMR* 6.5 hours post-release just prior to loading. Left, representative maximum projected immunofluorescence images of CENP-ASNAP csells. Right, representative images of CENP-A-SNAP + CENP-C knockdown cells. Insets are a 3x magnification of selected representative centromeres. Scale bar = 5 μm.

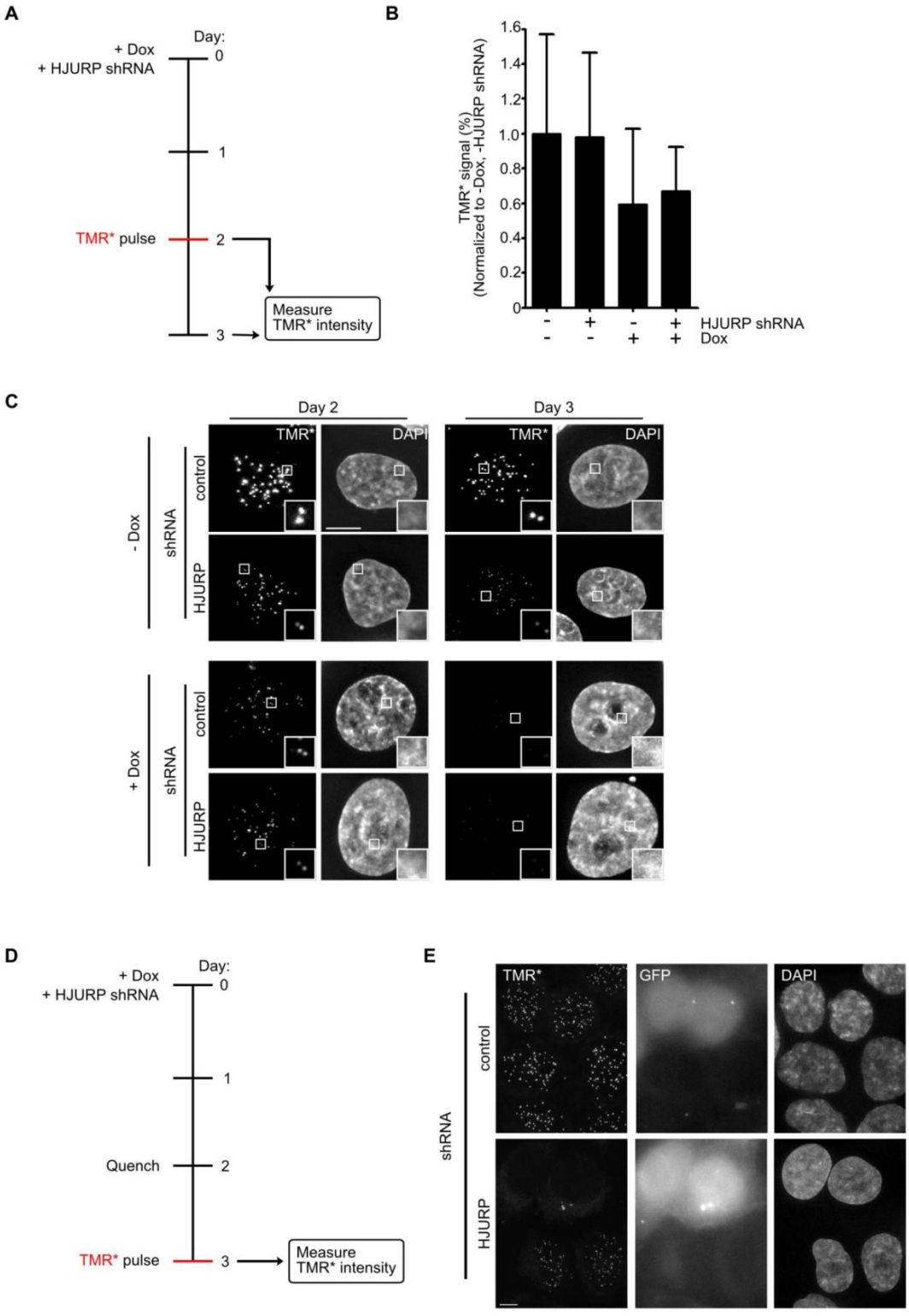


Figure 19. The reduction of CENP-A retention upon CENP-C knockdown is independent of new CENP-A chromatin assembly.

To determine whether CENP-C knockdown could reduce centromeric CENP-A levels in the absence of new CENP-A assembly, we quantified CENP-A retention upon co-depletion of CENP-C and HJURP**. (A)** Schematic of pulse-chase experiment to track CENP-A levels, combined with knockdown of both CENP-C and HJURP. Cells with (+ Dox) and without (-Dox) CENP-C depletion were co-transfected with shRNA-encoding plasmid targeting nucleotides 1288-1306 of HJURP (pSRP-HJURP) (Foltz et al., 2009) or the control plasmid (pSRP). A plasmid encoding GFP was co-transfected at a 9:1 (shRNA:GFP) ratio to mark transfected cells. The cells were pulse-labeled with TMR* (Day 2) and the CENP-A-SNAP signals were analyzed at day 2 and 3. **(B)** Quantification of CENP-A-SNAP signal retained at day 3. For each condition, >700 centromeres in GFP-positive cells were quantitated using the Centromere Recognition and Quantitation macro (Bodor et al., 2012) in ImageJ, and each TMR* value at day 3 was divided by its average TMR* value at day 2, then normalized to the - Dox, mock-transfected condition and plotted as mean ± s.d. **(C)** Representative images of GFP positive cells. CENP-C knockdown upon dox treatment was confirmed separately. Insets are 3x magnification. **(D)** We performed a quenchchase-pulse experiment on the same transfected cells used for panels A-C to assess whether or not the HJURP knockdown compromised new CENP-A assembly. By quenching pre-existing CENP-A with a non-fluorescent substrate and performing the TMR* labeling after 24 hours, we exclusively examine the pool of CENP-A-SNAP that had been newly synthesized and loaded onto centromeres. We TMR* labeled immediately after quenching and confirmed the lack of detectable TMR* signal in all conditions. **(E)** Representative images showing that transfection of HJURP shRNA abolishes new CENP-A loading. Note the assembly of CENP-A in the untransfected cells (GFP negative) adjacent to those that received HJURP shRNAs (GFP positive). Scale bars = 5 μm.

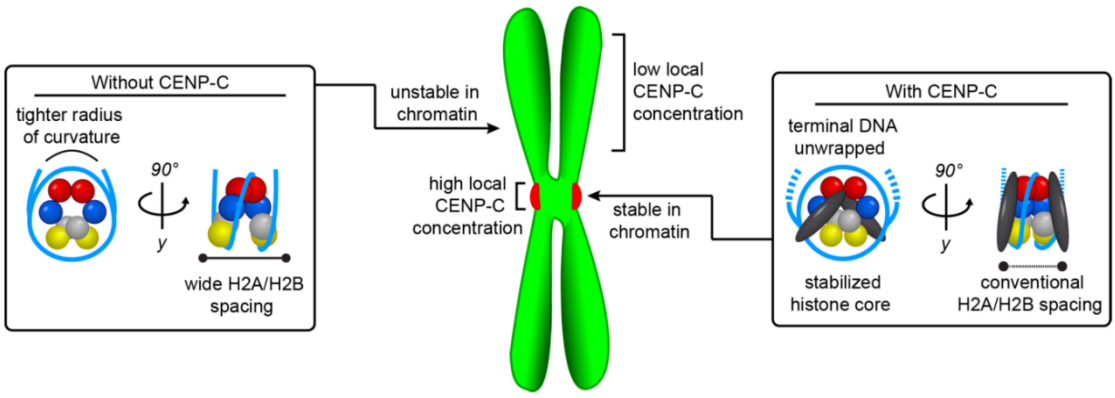


Figure 20. Summary model for collaboration of CENP-C with CENP-A nucleosomes in specifying centromere location.
At the centromere, there is a high local concentration of CENP-A, which results in a high local concentration of CENP-C. Together, CENP-A and CENP-C collaborate to form a stable complex that maintains the epigenetic mark of the centromere. In the chromatin arms, CENP-A levels do not reach a sufficient threshold to recruit CENP-C and CENP-A is quickly turned over. Our experiments support the idea that CENP-A-containing nucleosomes prefer an atypical shape in the absence of CENP-C, but

adopt a conventional overall histone octamer shape when CENP-C binds. In addition, our reconstitutions on native centromere DNA of octameric CENP-A nucleosomes very closely match the DNA wrapping properties of CENP-A nucleosomes isolated from functional human centromeres (Hasson et al., 2013), especially when CENP-C is bound (**Figure 16B**). This is in stark contrast to half-nucleosomes (termed hemisomes; i.e. one copy each of CENP-A, H4, H2A, and H2B) that wrap 65 bp of DNA (Furuyama et al., 2013) and have been proposed by others to be the major form at centromeres (Bui et al., 2012; Henikoff et al., 2014). Importantly, until now CENP-C has been considered primarily as a protein that recognizes CENP-A and bridges centromeric chromatin to other proteins important for centromere and kinetochore function (Carroll et al., 2010; Erhardt et al., 2008; Guse et al., 2011; Kato et al., 2013; Przewloka et al., 2011; Screpanti et al., 2011; Tomkiel et al., 1994) and helping target new CENP-A chromatin assembly at the centromere each cell cycle (Erhardt et al., 2008; Moree et al., 2011), but our findings that its binding directs changes to the shape and dynamics of the nucleosome suggest that it could also play a role in the special stability of CENP-A at centromeres in a manner analogous to allosteric regulation of enzymes. This has potential implications for chromatin regulation at diverse chromosome locations, as such a feature has not been reported for some other non-catalytic nucleosome binding proteins studied to date, like RCC1 and Sir3 (Armache et al., 2011; Makde et al., 2010), but now is worth considering for these and other nucleosome binding proteins. Directing a structural change upon binding of one component to a macromolecular complex to alter its behavior is a general strategy in biology, and our work with CENP-C importantly illustrates that a nucleosome—in this case, the special type at the centromere—is no exception.

2.4. DISCUSSION

CENP-A nucleosomes are highly stable at the centromeres upon which they are initially assembled. This stability is possible through collaboration with CENP-C. Along with the intranucleosomal rigidity of CENP-A and histone H4, where the key interfacial amino acids are important for accumulation at centromeres (Bassett et al., 2012; Black et al., 2004; Sekulic et al., 2010), the physical changes imposed by CENP-C combine to make CENP-A nucleosomes at centromeres very long-lived (**Figure 20**). Our data support a model of a steady-state octameric histone core where H2A/H2B dimers can exchange from either terminus of the CENP-A nucleosome. At the center, there is an essentially immobile $(CENP\text{-}A/H4)_2$ heterotetramer (Bodor et al., 2013)(**Figure 17** and *Figs. S12 to S14 in Falk et al., 2015*). Thus, the physical properties related to CENP-A nucleosome stability at

centromeres are tied to the intrinsic properties of the $(CENP\text{-}A/H4)_2$ heterotetramer (Bassett et al., 2012; Black et al., 2004; Sekulic et al., 2010) and the extrinsic properties imposed by CENP-C (**Figure 8, Figure 9, Figure 16**, **Figure 20**).

2.5. METHODS

2.5.1. FRET EXPERIMENTS (BY SAMANTHA FALK)

Recombinant human H2B was mutated using QuikChange (Stratagene) to contain a single cysteine (K120C or S123C) and then purified as described for the wildtype H2B (Sekulic et al., 2010). Lyophilized protein was dissolved in unfolding buffer (6 M Gnd-HCl, 10 mM Tris-HCl pH 7.5 at 20ºC, 0.4 mM TCEP) for 1 hr at RT and a 30-molar excess of either maleimido coumarin 343 (C343) or maleimido rhodamine B (RhB) dissolved in DMF was added dropwise to the protein. The reaction proceeded overnight shielded from light and was quenched with 10 mM DTT and run over a PD-10 column (GE Healthcare) to separate out free dye. Labeled H2B was then mixed with equimolar amounts of H2A for dimer reconstitution and purification using previously established methods (Dyer et al., 2004; Sekulic et al., 2010). Labeling efficiencies, E, ranged from 45-90% and were calculated by spectroscopy using the Beer-Lambert law using the following equation (Lakowicz, 2006): $E = [(A_{280} - (CF\ A_{max})/\varepsilon_{protein}l]/(A_{max}/\varepsilon_{fluorophore}l)$ (1) where A280 is the absorbance of protein at 280 nm, Amax is the absorbance of fluorophore at its maximum wavelength, εprotein and εfluorophore are the molar extinction coefficients for protein and fluorophore, respectively, l is the pathlength, and CF is the correction factor for contribution to the protein A280 from the fluorophore. Labeling efficiency was further confirmed by SDS-PAGE (Coomassie Blue staining) and mass spectrometry (*Figs. S1 and S2 in Falk et al., 2015*). α-satellite DNA derived from a sequence described by Harp, et al. (Harp et al., 1996) or the 601 DNA sequence described by Lowary and Widom (Lowary and Widom, 1998) were used in nucleosome assembly reactions. Briefly, a 145 bp region derived from a human α-satellite sequence with 25 bp of flanking DNA on each side was cloned into the pUC19 plasmid using EcoRI and XbaI restriction sites. The α-satellite DNA monomer

was then amplified from the plasmid by PCR using primers specific to the flanking DNA regions. The complete α-satellite sequence is: 5'-CGTATCGCCTCCCTCGCGCCATCAGATCAATATCCACCTGCAGATTCTACCAAAAGTGTATTTGGAAACTGCTCCATCAAAAGGCATGTTCAGCTCTGTGAGTGAAACTCCATCATCACAAAGAATATTCTGAGAATGCTTCCGTTTGCCTTTTATATGAACTTCCTGATCTGAGCGGGCTGGCAAGGCGCATAG- 3', with the 145 bp α-satellite region underlined. Typically, DNA from multiple 96-well PCR reactions were pooled, ethanol precipitated, resuspended in TE buffer and purified by anion- exchange chromatography. Widom 601 DNA was purified as described (Hasson et al., 2013). Nucleosomes were assembled on either DNA sequence and uniquely positioned using the gradual salt dialysis method followed by thermal shifting for 2 hr at 55°C (Dyer et al., 2004). Assembly was assessed by native PAGE (ethidium bromide and Coomassie Blue staining) and by SDS-PAGE (Coomassie Blue staining). As mentioned above, the fluorophores for FRET measurements were C343, serving asan energy donor (D), and RhB, serving as an acceptor (A). C343 and RhB were selected because their calculated R0 (Förster radius; distance at which energy transfer efficiency is 50%) is 58 Å, which is within the range of predicted dimer distances where energy transfer would be most sensitive to changes in FRET efficiency. For synthesis of fluorophores, all solvents and reagents were obtained from standard commercial sources and used as received. Selecto silica gel (Fisher Scientific, particle size 32-63 μm) was used for column chromatography. 1H NMR spectra were recorded on a Varian Unity 400 MHz spectrometer. Mass spectra were obtained on a MALDI- TOF MS Microflex LRF instrument (Bruker Daltonics), using α-cyano-4-hydroxycinnamic acid as a matrix. The compound maleimido C343 was synthesized by a CDMT-assisted peptide coupling of C343 and 1-(2-Aminoethyl)pyrrol-2,5-dione. 1-(2-Aminoethyl)pyrrol-2,5-dione was synthesized as described (Richter et al., 2012). C343 was dissolved in DMF at 0°C, 2-chloro-4,6-dimethoxy- 1,3,5-triazine and N-methylmorpholine (NMM) were added, and the mixture was stirred for 1 hr. 1-(2-Aminoethyl)pyrrol-2,5-dione and NMM were dissolved separately in DMF and added dropwise to the C343 mixture. The reaction was stirred at 0°C for 2 hr and then warmed to room temperature and stirred for 12 hr. The solvent was removed

under vacuum and the residue was purified by column chromatography (silica gel, DCM). The fraction containing maleimido C343 was collected, the solvent was evaporated, and the product was dried under vacuum. ^{1}H NMR (CDCl3, δ): 8.94 (s, 1H), 8.57 (s, 1H), 6.99 (d, 2H, 3J = 3.5 Hz), 6.70 (s, 2H), 3.80 (t, 1H, 3J = 5.8 Hz), 3.65 (m, 2H), 3.34 (m, 4H), 2.88 (t, 2H, 3J = 6.3 Hz), 2.77 (t, 2H, 3J = 6.1 Hz), 1.97 (m, 4H). For MALDI-TOF, the m/z (mass-to-charge ratio) calculated for C22H21N3O5 was 407.15; the following species were found; 407.102 [M]+ and 429.767 [M+Na]+. Maleimido RhB was synthesized by a HBTU-assisted peptide coupling of RhB piperazine amide (Nguyen and Francis, 2003) and N- maleimidoglycine (Kassianidis et al., 2006). ^{1}H NMR (CDCl3, δ): 7.78-7.72 (m, 3H), 7.53-7.51 (m, 1H), 7.29 (m, 2H), 7.10-7.05 (br s, 2H), 6.72 (s, 2H), 6.70 (s, 2H), 4.38 (s, 2H), 3.66-3.55 (m, 8H), 3.49-3.41 (m, 8H) 1.33 (t, 12H, ^{3}J = 7 Hz). For MALDI-TOF, the m/z calculated for $C_{38}H_{42}N_5O_5$+ was 648.32; the following species were found; 648.358. Steady-state emission measurements were performed on a FS900 spectrofluorometer (Edinburgh Instruments), equipped with a photon-counting R2658P PMT (Hamamatsu). Samples were excited at 450 nm, the wavelength at which the absorbance of an equimolar mixture of C343 and RhB is dominated by C343 (>99%), and measurements were performed using dilute solutions (ODmax< 0.1) in a Spectrosil quartz cuvette (1 cm optical path length, Starna Cells). As a result, only negligible RhB emission is observed under these conditions in the absence of FRET. Emission spectra were corrected by the detector quantum yield and normalized by the incident light intensity at the excitation wavelength. The final emission spectra used in quantum yield calculations (see below) are expressed in counts (photons) per second (CPS). Absorbance measurements were performed using a LAMBDA 35 UV/Vis spectrophotometer (PerkinElmer). FRET efficiency was calculated based on donor quenching in the presence of an acceptor fluorophore (Forster, 1946; Lorenz et al., 1999; Stryer and Haugland, 1967)). The quantum yield of fluorescence was calculated using the following equation (Crosby and Demas, 1971): $\Phi S = \Phi R[(A_R(\lambda_R)/A_S(\lambda_S)][n_S^2/n_R^2][D_S/D_R]$ (2) where Φ is quantum yield, A(λ) is the absorbance value at the designated excitation wavelength, n is the refractive index of the solution (nS = 1.333 and nR = 1.361), and D is the integrated emission

spectrum. The subscripts S and R refer to the sample and reference solutions, respectively. Rhodamine 6G in 100% ethanol was used as a reference actinometer ($\Phi R = 0.95$) (Kubin and Fletcher, 1982).

Because of the nature of nucleosome reconstitutions, nucleosomes reconstituted with both C343- and RhB-labeled dimers (i.e. our FRET samples) contain some percentage of C343-only nucleosomes. Both C343- and RhB-labeled dimers exhibited ~90% labeling efficiency, meaning that ~10% of the dimers used in a reconstitution reaction are unlabeled. This leads to a mixture of nucleosomes characterized by the following equation:

UU + DU + DD + DA + AA + AU = 1 (3) where UU represents the subset of nucleosomes that contain two unlabeled dimers, DU represents the subset of nucleosomes that contain one C343-labeled dimer and one unlabeled dimer, DD represents the subset of nucleosomes with two C343-labeled dimers, DA represents the subset of nucleosomes with one C343-labeled dimer and one RhB-labeled dimer, AA represents the subset of nucleosomes with two RhB-labeled dimers, and AU represents the subset of nucleosomes with one RhB-labeled dimer and one unlabeled dimer. Because we are using donor quenching to calculate FRET efficiency, we only consider C343-containing species, so equation 3 is simplified to the following equation: a + b = 1 (4) where a represents the normalized population of DU and DD nucleosomes and b represents the normalized population of DA nucleosomes in the FRET sample. Both a and b can be calculated using the known labeling efficiencies of both donor-labeled and acceptor-labeled dimers determined from spectroscopy and mass spectrometry analysis.

In order to account for the subset of DU and DD nucleosomes present in our FRET samples when measuring donor quenching, a separate control sample of C343-only nucleosomes are reconstituted and measured alongside every experimental sample. The following equation is then used to calculate the quantum yield of C343 in nucleosomes containing both C343 and RhB dimers: $\Phi DA = [\Phi T - a(\Phi DD)]/b$ (5), where ΦDA is the quantum yield of C343-RhB nucleosomes (DA), ΦT is the total quantum yield of all C343-containing nucleosomes (DU + DD + DA), ΦDD is the quantum yield of C343-only nucleosomes (DU + DD), and a and b

represent the fraction of C343-only nucleosomes and C343-Rhb nucleosomes in a sample, respectively, determined as described above. ΦT and ΦDD are calculated from the FRET sample and the C343-only sample, respectively, using equation 2 above.

FRET efficiency, ΦFRET, is then determined based on the following equation (Lakowicz, 2006): ΦFRET = 1-(ΦDA/ΦDD) (6). The distance, r, between the two fluorophores is then calculated using the following equation (Lakowicz, 2006): r = R0[(1/ΦFRET)-1]1/6 (7) where R0 is the Förster radius. For the C343/RhB pair, the R0 was calculated to be 58 Å, using the following equation (Lakowicz, 2006): R0 = 9790(Jκ2ΦDDn-4)1/6 Å (8) where J is the spectral overlap integral for C343/RhB pair, ΦDD is the quantum yield of C343, n is the refractive index of the solvent (n=1.333), and κ2=2/3 is the orientation factor for freely rotating fluorophores (Lakowicz, 2006). Our assumption of orientational averaging as in the case of freely rotating transition dipole moments was confirmed by our anisotropy measurements (see below). The measured anisotropy for the fluorophore pair was found to be less than 0.2 (*fig. S1 and Table S1 in Falk et al., 2015*), confirming that usage of formula (7) was appropriate for estimation of interchromophoric distances. Steady-state fluorescence anisotropy measurements were performed on a QuantaMaster spectrophotometer (PTI). Samples were diluted to 0.5-1.0 μM in 150 mM NaCl, 20 mM Tris-HCl pH 7.5 at 4°C, 1 mM EDTA, 1 mM DTT and excited at 450 nm for C343 and 567 nm for RhB. Anisotropy, r, was calculated in FeliX32 software using the following equation (Lakowicz, 2006): r = (IVV – G IVH)/(IVV + 2G IVH) (9) where IVV is the parallel polarized fluorescence intensity, IVH is the perpendicular polarized fluorescence intensity, and G is the correction factor for the setup. Lifetime measurements, τ, were performed using a FluoroLog fluorometer (Horiba Scientific). The excitation source was an LED (NanoLED), lmax=441 nm with an average repetition rate of 1 MHz. Samples were in 150 mM NaCl, 20 mM Tris-HCl pH 7.5 at 4°C, 1 mM EDTA, 1 mM DTT at 0.5-1.0 μM. Emission was measured at 491 nm using a bandpass filter (5 nm). Lifetimes were fitted exponentially using DAS6 software (Horiba Scientific).

2.5.2. HXMS

CENP-A mononucleosomes were reconstituted with the same 195 bp α-satellite DNA described above in the FRET studies and concentrated to 0.9 mg/ml with Centricon concentrators (Millipore, Billerica, MA). Recombinant human CENP-C^{CD} consisting of the central domain only (a.a. 426-537, the plasmid for recombinant human CENP-C^{CD} expression was a generous gift from A. Straight, Stanford, USA) was GST-tagged and purified over a GST column followed by PreScission protease cleavage (GE Healthcare) and ion-exchange chromatography and prepared in a buffer containing 20 mM Tris pH 7.5, 200 mM NaCl, 0.5 mM EDTA, 1 mM DTT. To form complexes with CENP-C^{CD}, 2.2 moles of recombinant CENP-C^{CD} were added per mole of CENP-A nucleosomes. To the nucleosome-only sample the buffer used for CENP-C^{CD} preparation was added so that the chemical composition of the buffers were identical in all cases. Deuterium on-exchange was carried out by adding 5 µL of each sample (containing approximately 4 µg of nucleosomes or complex) to 15 µL of deuterium on-exchange buffer (10 mM Tris, pD 7.5, 0.5 mM EDTA, in D2O) so that the final D2O content was 75%. Reactions were quenched at the indicated time points by withdrawing 20 µL of the reaction volume, mixing in 30 µL ice cold quench buffer (2.5 M GdHCl, 0.8% formic acid, 10% glycerol), and rapidly freezing in liquid nitrogen prior to proteolysis and LC-MS steps. HX samples were individually melted at 0°C then injected (50 µl) and pumped through an immobilized pepsin (Sigma) column at initial flow rate of 50 µl/min for 2 min followed by 150 µl/min for another 2 min. Pepsin was immobilized by coupling to Poros 20 AL support (Applied Biosystems) and packed into column housings of 2 mm x 2 cm (64 µL) (Upchurch). Protease-generated fragments were collected onto a C18 HPLC trap column (800 µm x 2 mm, Dionex) and eluted through an analytical C18 HPLC column (0.3 x 75 mm, Agilent) by a linear 12-55% buffer B gradient at 6 µl/min (Buffer A: 0.1% formic acid; Buffer B: 0.1% formic acid, 99.9% acetonitrile). The effluent was electrosprayed into the mass spectrometer (LTQ Orbitrap XL, Thermo Fisher Scientific). The SEQUEST (Bioworks) software program was used to identify the likely sequence of parent peptides using nondeuterated samples via tandem MS.

MATLAB based MS data analysis tool, ExMS, was used for data processing (Kan et al., 2011).

2.5.3. MNase digestions (by Nikolina Sekulic)

Nucleosomes were assembled using the same 195 bp α-satellite DNA sequence used in FRET studies using the same assembly approach described above. Nucleosomes were digested for various times with 2 U/µg of MNase (Roche) at room temperature (22°C). Reactions were terminated with the addition of guanidine thiocyanate and EGTA. The DNA was isolated using a MinElute PCR purification kit (Qiagen) and analyzed on an Agilent 2100 Bioanalyzer.

2.5.4. SANS (by Nikolina Sekulic)

Nucleosome core particles were assembled on the α-satellite 145 bp sequence described above. The sequence was cloned in tandem copies separated by EcoRV sites in pUC57. The 145 bp fragments were released by EcoRV digestion and purified away from the backbone by anion exchange chromatography. Following nucleosome reconstitutions, performed as described above, the nucleosomes were purified by preparative electrophoresis (Prep Cell, BioRad) using a 5% native gel to separate free DNA and any other non-nucleosomal species (Dyer et al., 2004). SANS experiments were performed at the National Institutes of Standards and Technology Center for Neutron Research NG-3. Samples were prepared by dialysis at 4°C against matching buffers containing 20% or 80% D2O for a minimum of 3 hr using a 6-8 kDa cutoff D-tube dialyzer (Novagen). Samples were centrifuged at 10,000 X g for 5 min at 4°C and then loaded into Hellma quartz cylindrical cells (outside diameter of 22 mm) with 1 mm path lengths and maintained at 6ºC during the experiment. Sample concentrations were determined by Bradford analysis and optical absorbance at 260 nm. Scattered neutrons were detected with a 64 cm × 64 cm two-dimensional position-sensitive detector with 128 × 128 pixels at a resolution of 0.5 cm/pixel. Data reduction was performed using the NCNR Igor Pro macro package (Kline, 2006). Raw counts were normalized to a common monitor count and

corrected for empty cell counts, ambient room background counts and non-uniform detector response. Data were placed on an absolute scale by normalizing the scattered intensity to the incident beam flux. Finally, the data were radially-averaged to produce scattered intensity, I(q), versus q curves. The scattered intensities from the samples were further corrected for buffer scattering and incoherent scattering from hydrogen in the samples. Data collection times varied from 0.5-2 hr, depending on the instrument configuration, sample concentration and buffer conditions. Sample-to-detector distances of 11 m (q-range 0.006-0.043 Å-1, where $q = 4\pi\sin(\theta)/\lambda$, where λ is the neutron wavelength and 2θ is the scattering angle), 5 m (q-range 0.011–0.094 Å–1), and 1.5 m (detector offset by 20.00 cm, q-range 0.03–0.4 Å–1) at a wavelength of 6 Å and a wavelength spread of 0.15 were collected for each contrast point. We observed good agreement between Rg and I(0) values determined from either inverse Fourier analysis using GNOM or from Guinier analysis. The program MuLCH (Whitten et al., 2008) was used to calculate theoretical contrast and to analyze contrast variation data. Distance distribution curves were normalized for total molecular mass for the complex.

2.5.5. SNAP LABELING EXPERIMENTS AND CELL FUSIONS (BY EVAN SMOAK AND SAMANTHA FALK)

CENP-A-SNAP HeLa cells for fusion experiments were labeled with TMR* (NEB) as described previously and subjected to a double thymidine block with a final thymidine concentration of 2 mM (Bodor et al., 2013; Jansen et al., 2007). YFP-CENP-A HeLa cells (Black et al., 2007b), CENP-A-SNAP Hela cells (Jansen et al., 2007), and SNAP-tagged core histone (H3.1, H3.3, H4, and H2B)-expressing HeLa cells (Black et al., 2007b; Bodor et al., 2013) are all established lines. After labeling with TMR*, CENP- A-SNAP HeLa cells were trypsinized, counted, and co-seeded onto poly-D-lysine (Sigma-Aldrich) treated coverslips along with an equivalent number of HeLa cells constitutively expressing YFP-CENP-A. Cells were arrested in growth medium (DMEM supplemented with 10% fetal bovine serum (FBS), 100 U/mL penicillin, and 100 μg/mL streptomycin) containing 2 mM thymidine for 17 hr. Cells

were then washed 3x with PBS, fused with 50% PEG-1500 (Roche) for 30 s and subsequently washed in PBS and placed in media containing 24 μM deoxycytidine to release from thymidine block. After 9 hr, cells were blocked again with media containing thymidine for 17 hr. Cells were released from thymidine with DMEM media containing 24 μM deoxycytidine and nocodazole was added 7 hr post-release at a final concentration of 400 ng/mL. Coverslips were fixed and processed for immunofluorescence at the timepoints outlined in *Fig. S12A in Falk et al, 2015*. HeLa-based cell lines for inducible CENP-A-SNAP with and without shRNAs, and constitutive CENP-A-SNAP with inducible shRNAs directed against CENP-C were generated by recombinase-mediated cassette exchange (RMCE) using the HILO RMCE system (a generous gift from E.V. Makeyev, Nanyang Technological University, Singapore (Khandelia et al., 2011)). pEM784 was used to express nuclear-localized Cre recombinase. pEM791 was modified for inducible expression of CENP-A-SNAP-HA3, CENP-A-SNAP-HA3 plus 2 shRNAs against CENP-C (5'-tgctgttgactttctaccttgaaggagttttggccgctgactgactccttcaatagaaagtcaa-3' and 5'-tgctgacaagtttgttcttggactcagttttggccactgactgactgagtccaaacaaacttgt-3'), constitutive CENP- A-SNAP-HA3 driven by the EF1α promoter plus 2 shRNAs against CENP-C, and CENP-A- PAGFP respectively. CENP-C knockdown was induced in constitutive CENP-A-SNAP cell lines by treating for 48 hr with 2 μg/mL doxycycline prior to TMR* labeling for pulse-chase experiments to measure the retention of CENP-A protein at centromeres. Cells were fixed either immediately after labeling or again 24 hr later. Cell number was also determined at these time points, so that the total level of CENP-A turnover could be calculated, as described (Bodor et al., 2013). For experiments to measure the amount of new CENP-A assembly with or without CENP-C knockdown, cells were treated with 50 ng/mL of doxycycline during a double thymidine block procedure that spanned 48 hr. Following release from the double thymidine block, CENP-A- SNAP was quenched with SNAP-Cell Block (NEB) then released for 6.5 hr to allow for new synthesis of CENP-A-SNAP protein. The nascent pool of CENP-A-SNAP protein was then pulse-labeled with TMR*, and the cells were cultured for an additional 17.5 hr prior to fixation and processing for immunofluorescence. A separate sample was labeled with TMR* immediately after

the quench step to confirm successful quenching of 'old' CENP-A. For immunofluorescence, cells were fixed in 4% formaldehyde for 10 min at room temperature followed by permeabilization using PBS + 0.5% Triton X-100. Samples were stained with DAPI before mounting with Vectashield medium (Vector laboratories). The following primary antibodies were used: mouse mAb anti-CENP-A (1:1000 Enzo), rabbit pAb anti-CENP-C (1:2000) (Bassett et al., 2010), and mouse mAb anti-HA.II antibody (1:1000, Covance). AlexaFluor488- and AlexaFluor647-conjugated secondary antibodies were obtained from Invitrogen and were used at 1:1000. Images were captured at 23°C using software (LAF; Leica) by a charge-coupled device camera (ORCA AG; Hamamatsu Photonics) mounted on an inverted microscope (DMI6000B; Leica) with a 100x 1.4 NA objective. For each sample, images were collected at either 0.2 μm z- sections (**Figure 17A**, **Figure 19**, and *S12-15 in Falk et al., 2015*) or 0.49 μm z-sections (**Figure 17D,G** and Fig. *S16 in Falk et al., 2015*) that were subsequently deconvolved using identical parameters. The z-stacks were projected as single two- dimensional images and assembled using PhotoShop (version 13.0; Adobe), ImageJ (1.48v) (Schneider et al., 2012), and Illustrator (version 16.0; Adobe). To quantify fluorescence intensity in cell fusions, individual centromeres from non-deconvolved maximum projections were selected and the intensity of both TMR* and YFP signal were determined after subtracting the background fluorescence measured from adjacent regions of the cell using ImageJ. For each unique fusion, the levels of fluorescence for both channels were normalized to the highest measured value in that channel, leading to normalized values for YFP intensity and TMR* intensity for each centromere in the fused cell. Thus, each centromere is a data point that has an associated TMR* and YFP value assigned to it, which were then run through the machine learning x-means clustering algorithm of Weka (Hall et al., 2009; Pelleg and Moore, 2000), which partitions the data points into n clusters based on their closeness to an assigned mean value. This generated the two groups of data points (YFP only and YFP + TMR*) in the plot seen in **Figure 17B**. To quantify fluorescence intensity in experiments with CENP-C knockdown, the Centromere Recognition and Quantification (CRaQ) macro (Bodor et al., 2012) was run in ImageJ with standard

settings using a reference channel and DAPI. Total CENP-A staining was used as the reference channel to define ROIs for quantification of TMR* intensity. CENP-A fluorescence intensity values at the final time point were normalized to reflect the total pool of labeled CENP-A by accounting for the increase in cell number in the dividing cell populations following TMR* pulse. 2800-4200 centromeres from >70 cells were analyzed for each time point.

2.5.6. PAGFP EXPERIMENTS (BY EVAN SMOAK)

CENP-A-PAGFP cells were generated with the RMCE system (Khandelia et al., 2011), as described above, and expression was induced with 1 μg/mL doxycycline 2 days prior to photoactivation and continued for the duration of the experiment. Cells were cultured in growth medium at 37°C in a humidified atmosphere with 5% CO2. For live imaging, cells were plated on 22 x 22 mm glass coverslips (#1.5; Thermo Fisher Scientific) coated with poly-D-lysine (Sigma-Aldrich). Coverslips were mounted in magnetic chambers (Chamlide CM-S22-1, LCI) using growth medium without phenol red (Invitrogen). Temperature was maintained at 37°C with 5% CO2 using an environmental chamber (Incubator BL; PeCon GmbH). Evaporation of media was prevented by applying a thin layer of mineral oil over the media within the magnetic chamber. Prior to photoactivation, a single plane image of the unactivated nucleus was acquired to be used for background subtraction. Cells were subsequently photoactivated by defining an ROI surrounding ~half of the nucleus and then activated using a pointable 405 nm laser (CrystaLaser) set to 10% power and one repetition using iLAS2 software run through MetaMorph, followed by acquisition of an image of a single z-plane. Cells were then followed by DIC for 8 hr, at which point a final single plane image was acquired. Images were acquired with a confocal microscope (DM4000; Leica) with a 100x 1.4 NA objective lens, an XY Piezo-Z stage (Applied Scientific Instrumentation), a spinning disk (Yokogawa Corporation of America), an electron multiplier charge-coupled device camera (ImageEM; Hamamatsu Photonics), and a laser merge module equipped with 488 nm and 593 nm lasers (LMM5; Spectral Applied Research) controlled by MetaMorph

software (Molecular Devices). To quantify the retention of CENP-A-PAGFP in bulk chromatin, a 25 x 25 pixel region-of-interest (ROI) was drawn in ImageJ on a region of photoactivated bulk chromatin and the average fluorescence was recorded. Recorded fluorescence measurements were corrected for background by subtracting the fluorescence value of the pre-photoactivated ROI. This corrected average fluorescence of the ROI was then multiplied by the area of the ROI in order to calculate the average fluorescence of the area. These area fluorescence measurements of the bulk chromatin ROI for each cell were averaged to generate the final numbers for comparison. Upon overexpression, CENP-A initially assembles into nucleosomes at centromeres and at locations throughout the genome (Heun et al., 2006; Lacoste et al., 2014). Functional centromeres do not spread throughout chromosomes under these conditions. We note that there is a small soluble pool of CENP-A in bulk chromatin that is mobile throughout the nucleus, but this does not significantly contribute to quantification and does not diffuse to the unactivated portion of the nucleus in earlier time points. To quantify the retention of CENP-A-PAGFP in centromeric chromatin, single planes were thresholded to create ROIs around all visibly photoactivated centromeres and the average fluorescence as well as the total centromeric area were both recorded. Fluorescence measurements were corrected for background by subtracting the fluorescence value of the pre-photoactivated ROIs. This corrected average fluorescence of the ROI was then multiplied by the area of the ROIs in order to calculate the average fluorescence of the area. These area fluorescence measurements of the centromeric chromatin ROIs for each cell were averaged to generate the final numbers for comparison.

2.5.7. CELL LETHALITY ASSAY

Constitutive CENP-A-SNAP HeLa cells with doxycycline-inducible shRNAs directed against CENP-C were seeded in 6-well plates in triplicate at 8.4 x 104 cells

per well and with daily introduction of 2 μg/mL dox. Cells were collected and stained with 0.4% Trypan Blue (CellGro) and counted on a hemocytometer to calculate the percentage of cell death based on trypan blue uptake.

2.5.8. IMMUNOBLOTTING

Samples derived from whole cell lysates were separated by SDS-PAGE and transferred to a nitrocellulose membrane for immunoblotting. Blots were probed using the following antibodies: human ACA (2 μg/mL, Antibodies Incorporated), rabbit anti-CENP-C (1.7 μg/mL) and mouse mAb anti-α-tubulin (1:4000, Sigma-Aldrich). Antibodies were detected using a horseradish peroxidase-conjugated secondary antibody at 1:10,000 (human) and 1:2000 (rabbit or mouse) (Jackson ImmunoResearch Laboratories) and enhanced chemiluminescence (Thermo Scientific).

CHAPTER 3: CENTROMERES ARE MAINTAINED BY FASTENING CENP-A TO DNA AND DIRECTING AN ARGININE ANCHOR-DEPENDENT NUCLEOSOME TRANSITION

Chapter 3 is based on the following publication:

> Guo, L.Y., Allu P.K., Zandarashvili, L., McKinley, K.L., Sekulic, N., Fachinetti, D., Logsdon G.A., Jamiolkowski, R.M., Cleveland, D.W., Cheeseman, I.M., and Black, B.E. Centromeres are maintained by fastening CENP-A to DNA and directing an arginine anchor-dependent nucleosome transition. *Nature Communications,* **8,** 15775 doi: 10.1038/ncomms15775 (2017).

Chapter 3 also includes my contribution in the following manuscript, which is shown in Figure 27 in this chapter:

> McKinley, K.L., Sekulic, N., Guo, L.Y., Tsinman, T., Black, B.E., and Cheeseman, I.M. (2015). The CENP-L-N complex forms a critical node in an integrated meshwork of interactions at the centromere-kinetochore interface. *Molecular Cell* 2015 Dec 17;60(6):886-98.

3.1. ABSTRACT

Maintaining centromere identity relies upon the persistence of the epigenetic mark provided by the histone H3 variant, CENP-A, but the molecular mechanisms that underlie its remarkable stability remain unclear. Here, we define the contributions of each of the three candidate CENP-A nucleosome-binding domains (two on CENP-C and one on CENP-N) to CENP-A stability using gene replacement and rapid protein degradation. Surprisingly, the most conserved domain, the CENP-C motif, is dispensable. Instead, the stability is conferred by the unfolded central domain of CENP-C and the folded N-terminal domain of CENP-N that becomes rigidified 1000-fold upon crossbridging CENP-A and its adjacent nucleosomal DNA. Disrupting the 'arginine anchor' on CENP-C for the nucleosomal acidic patch disrupts the CENP-A nucleosome structural transition and removes CENP-A

nucleosomes from centromeres. CENP-A nucleosome retention at centromeres requires a core centromeric nucleosome complex where CENP-C clamps down a stable nucleosome conformation and CENP-N fastens CENP-A to the DNA.

3.2. INTRODUCTION

The centromere is the specialized region of chromatin that directs accurate chromosome segregation in cell division (McKinley and Cheeseman, 2016; Westhorpe and Straight, 2015). The centromere recruits the proteinaceous kinetochore, which attaches to spindle microtubules during mitosis or meiosis. A model for the epigenetic specification of centromere identity has emerged wherein preexisting nucleosomes with a histone H3 variant named centromere protein A (CENP-A)(Earnshaw and Rothfield, 1985; Palmer and Margolis, 1985) direct the local assembly of newly synthesized CENP-A (Black and Cleveland, 2011; Fachinetti et al., 2013), with CENP-A deposition occurring once per cell cycle following completion of mitosis (Jansen et al., 2007; Schuh et al., 2007). Critically, this model relies on the stable maintenance of CENP-A nucleosomes at a single site on each chromosome throughout the remainder of the cell cycle.

Indeed, relative to the other H3 variants (i.e. H3.1 and H3.3) that turnover in chromatin (Bodor et al., 2013; Falk et al., 2015; Nashun et al., 2015), CENP-A experiences essentially no detectable turnover once assembled at a centromere (Bodor et al., 2013; Fachinetti et al., 2013; Falk et al., 2015; Jansen et al., 2007), and the stability has been measured out to >1 year where it preserves centromere identity in oocytes that are arrested in a prophase-like state during the entire fertile lifespan of female mice (Smoak et al., 2016). Particularly in the female germline or any somatic cell types that do not undergo very rapid divisions, maintaining centromere identity between rounds of CENP-A nucleosome assembly is critical for faithful chromosome inheritance. Thus, defining the molecular processes that confer the extraordinary stability of CENP-A nucleosomes is of outstanding interest in chromosome biology.

To date, both intrinsic features (i.e. those encoded in the sequence of CENP-A, itself) and extrinsic factors (i.e. constitutive centromere components that bind directly to CENP-A nucleosomes) have been considered as candidates that contribute to this distinctive stability. Residues that rigidify the interface between CENP-A and its partner histone, H4, are necessary but not sufficient for this stability (Bassett et al., 2012; Black et al., 2004; Falk et al., 2015; Sekulic et al., 2010), so extrinsic factors must be considered. The only two proteins of the constitutive centromere associated network (CCAN) known to make specific contacts with CENP-A nucleosomes on all functional mammalian centromeres are CENP-C and the CENP-N subunit of the CENP-L-N complex (Carroll et al., 2009, 2010; Kato et al., 2013; McKinley et al., 2015). Between these components of the CCAN, there are a total of three nucleosome-binding domains: two on CENP-C (the central domain [CENP-C^{CD} a.a. 426-537](Carroll et al., 2010) and the CENP-C motif [CENP-C^{CM} a.a. 736-758](Kato et al., 2013)) and one comprised of the N-terminal portion of CENP-N (CENP-N^{NT} a.a. 1-240)(Carroll et al., 2009; McKinley et al., 2015).

For the two nucleosome binding domains of CENP-C, CENP-C^{CD} and CENP-C^{CM} each are proposed to engage the CENP-A nucleosome through similar histone contact points and without any local secondary structure of their own (Kato et al., 2013). CENP-C^{CD} is conserved in mammals (Kato et al., 2013), was mapped initially as the primary CENP-A nucleosome contact site, and has high specificity for CENP-A nucleosomes versus its counterparts with canonical H3 (Carroll et al., 2010). CENP-C^{CD} also directs a structural transition of the CENP-A nucleosome that changes the shape of the octameric histone core, slides the gyres of the nucleosomal DNA past one another, and generates both surface and internal rigidity to the histone subunits (Falk et al., 2015, 2016). CENP-C^{CM}, on the other hand, is conserved from yeast to humans, and represents the only identified nucleosome-binding domain in species lacking a conserved CENP-C^{CD} (Carroll et al., 2010). CENP-C^{CM} is the only CENP-A nucleosome binding domain for which there exists atomic-level structural information, with a crystal structure of it bound to a canonical nucleosome in which

the 6 a.a. C-terminal tail of CENP-A replaces the corresponding region of histone H3 (Kato et al., 2013). This structure revealed that CENP-C^{CM} uses a so-called 'arginine anchor' to recognize the acidic patch on the H2A-H2B dimer (Kato et al., 2013). An arginine anchor is the shared feature of a diverse set of nucleosome binding proteins studied to date (Armache et al., 2011; Barbera et al., 2006; Makde et al., 2010; McGinty et al., 2014; Morgan et al., 2016), establishing an emerging paradigm for nucleosome recognition (McGinty and Tan, 2016).

Prior reports have suggested that either or both of the nucleosome binding domains of CENP-C could be important for its own localization to centromeres (Carroll et al., 2010; Kato et al., 2013; Lanini and McKeon, 1995; Milks et al., 2009; Trazzi et al., 2002; Yang et al., 1996). CENP-N^{NT} recognizes the CENP-A nucleosome via the CENP-A Targeting Domain (CATD(Black et al., 2004))(Carroll et al., 2009; Fang et al., 2015; McKinley et al., 2015), but it is not known whether its binding site on the nucleosome extends to other histones in a similar manner as the CENP-C nucleosome binding domains (Kato et al., 2013). While a prior study using labeled CENP-N and CENP-C expressed in reticulocyte extracts suggested that they can co-exist on the same nucleosome, there existed a need to use purified components to resolve proposals for CENP-C and CENP-N to bind to the same (Carroll et al., 2010) or different (Fang et al., 2015; Nagpal et al., 2015) CENP-A nucleosomes, and to study the nature of such a combined complex. Depletion of CENP-C reduces CENP-A nucleosome stability (Falk et al., 2015), but this finding does not delineate between a role for CENP-C^{CD} or CENP-C^{CM}. Also, CENP-C depletion leads to partial removal of the CENP-L-N complex (McKinley et al., 2015), so it remains possible that CENP-N^{NT} is responsible for CENP-A nucleosome retention. Thus, it is currently unclear which of the CENP-C or CENP-N domains is important for maintaining centromere identity and the extent to which they may cooperate to stabilize centromeric chromatin.

Here, we define the contributions of each of the three nucleosome binding domains present within the CCAN for maintaining centromere identity. To do this, we use a combination of gene editing, rapid inducible degradation of centromere

components, biochemical reconstitution, hydroxyl radical footprinting, and hydrogen/deuterium exchange coupled to mass spectrometry (HXMS). Our data establish an essential Core Centromeric Nucleosome Complex (CCNC) that is critical for CENP-A stability and maintenance of centromere integrity.

3.3. RESULTS

3.3.1 CENP-C^{CD} CONFERS STABILITY TO CENP-A NUCLEOSOMES

Upon embarking on our effort to define the molecular processes that confer stability to CENP-A nucleosomes, we first turned our attention to CENP-C. We reasoned that if either or both of the nucleosome binding domains of CENP-C were indeed required for its localization to centromeres (Carroll et al., 2010; Kato et al., 2013; Lanini and McKeon, 1995; Milks et al., 2009; Trazzi et al., 2002; Yang et al., 1996), then we could not accurately define which of the domains may confer stability to CENP-A. To define the requirements for these domains in the absence of endogenous CENP-C, we employed a human DLD-1 cell line in which both alleles of CENP-C are tagged with an auxin-inducible degron (AID)(Holland et al., 2012; Nishimura et al., 2009) and EYFP tags (Fachinetti et al., 2015) (**Figure 21a**). In this background, we introduced untagged versions of either wild type or mutant CENP-C proteins, constitutively expressed from a unique genomic locus (**Figure 21b**). The AID-EYFP-tagged CENP-C with an otherwise wildtype protein coding sequence is degraded to below the level of detection within 30 min of addition of the synthetic auxin, indole-3-acetic acid (IAA)(**Figure 22a-d**). This allowed us to exclusively detect the rescue constructs with an antibody directed against CENP-C (**Figure 21c-e**). Since all constructs are expressed at roughly equal levels as the AID-EYFP-tagged version prior to IAA treatment, there is an expected drop in the amount that is

detectable at centromeres after IAA treatment even with the wildtype full-length version [CENP-C(FL); **Figure 21d,e**, and **Figure 22e,f**]. The removal of the CD led to partially diminished CENP-C localization, whereas removal of the CM had little effect, even when removed in combination with the CD (**Figure 21d,e**). Thus, the CD and CM are not strictly necessary for CENP-C localization, consistent with the fact that CENP-C makes multiple other direct contacts within the meshwork of the CCAN (Klare et al., 2015b; McKinley et al., 2015; Nagpal et al., 2015).

Previously, we found that slow reduction of CENP-C (via shRNA treatment) causes a marked decrease in the retention of SNAP-tagged and tetramethylrhodamine-Star (TMR*) pulse-labeled CENP-A at centromeres (Falk et al., 2015). This strategy allows us to monitor the pool of CENP-A nucleosomes existing prior to the pulse labeling (Bodor et al., 2013; Falk et al., 2015; Jansen et al., 2007), so that any effects of CENP-C depletion on nascent CENP-A assembly (Carroll et al., 2010; Erhardt et al., 2008; Moree et al., 2011) do not complicate our analysis. Therefore, we next tested whether these domains of CENP-C are required for the retention of CENP-A at centromeres. We first added the SNAP tag to endogenous CENP-A using CRISPR-Cas9-mediated genome editing (**Figure 21f**, and **Figure 22g-j**), and confirmed that rapid removal of CENP-C-AID-EYFP with no rescue causes a dramatic decrease over 24 h of the existing pool of CENP-A at centromeres (**Figure 22k-m**). CENP-C(FL) rescued CENP-A retention, whereas replacement with CENP-C(ΔCD) resulted in a marked decrease in CENP-A retention (**Figure 21g-i**). In contrast, CENP-C(ΔCM) did not diminish CENP-A retention (**Figure 21g-i**). Therefore, CENP-C^{CD} is required for the retention of CENP-A at centromeres.

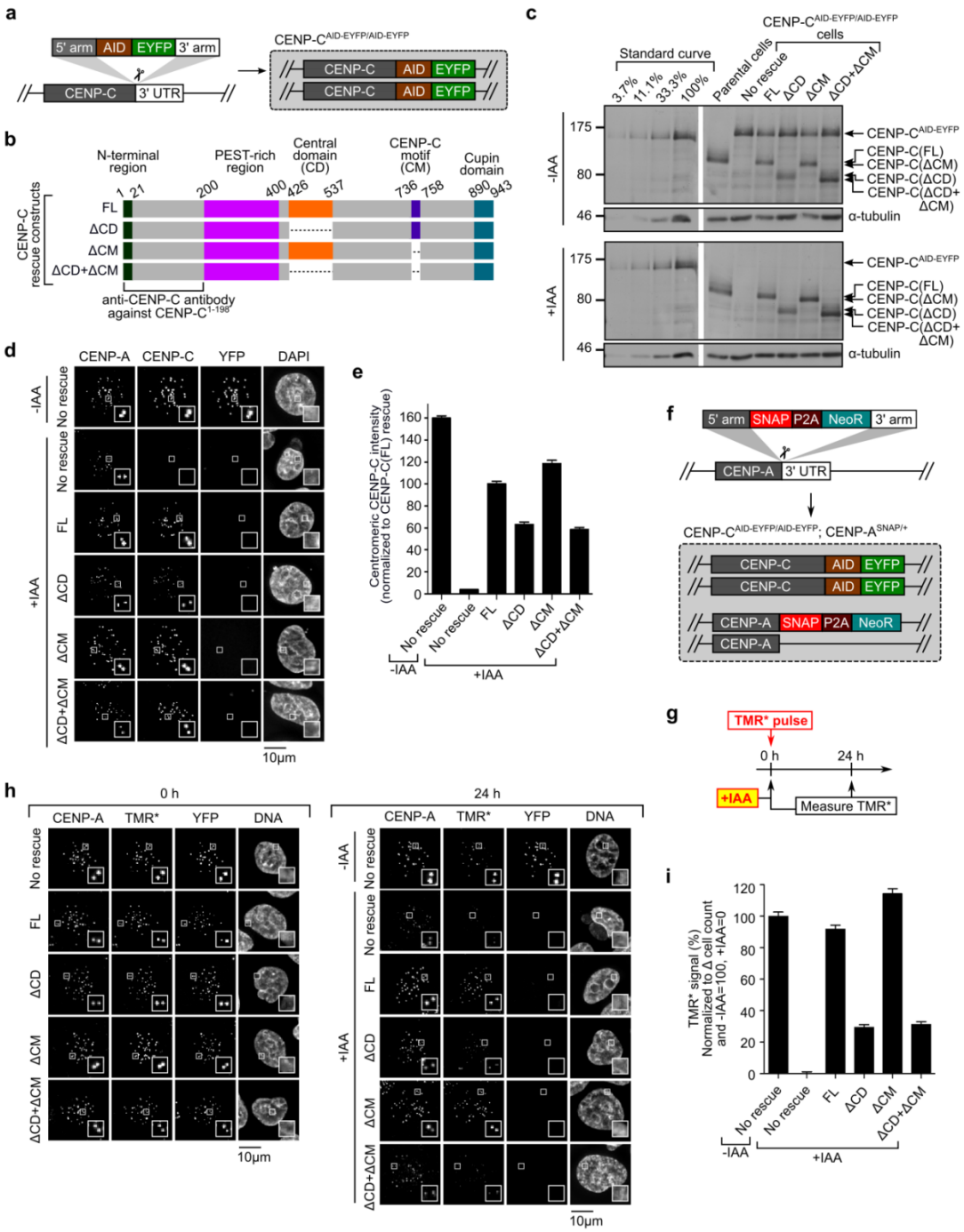


Figure 21. CENP-C^{CD} is the only nucleosome binding domain of CENP-C required for retention of CENP-A nucleosomes.
(a) Schematic of CENP-C$^{AID\text{-}EYFP/AID\text{-}EYFP}$ cells. **(b)** Rescue constructs constitutively expressed at unique FRT site. FL=full length. **(c)** Immunoblot of CENP-C$^{AID\text{-}EYFP/AID\text{-}EYFP}$ cells (with and without 4 h of auxin-induced CENP-C depletion), using an antibody generated against CENP-C (a.a. 1-198). **(d)** Representative images, in which the loss of YFP signal verifies depletion of CENP-C-AID-EYFP after 24 h of IAA, and CENP-C antibody then exclusively detects rescue constructs. **(e)** Quantitation of

d. **(f)** Schematic for SNAP-tagging CENP-A at its endogenous locus. **(g)** Schematic for pulse-chase experiment in which CENP-C$^{AID\text{-}EYFP/AID\text{-}EYFP}$ cells expressing rescue constructs of either CENP-C(FL) or CENP-C domain deletion mutants were pulse-labeled with TMR* and assessed for retention of the existing pool of CENP-A molecules. **(h)** Representative images from experiment diagrammed in g. **(i)** Quantitation of h. See also **Figure 22k-m**. All graphs are shown as mean ±95% confidence interval (n>2000 centromeres in all cases).

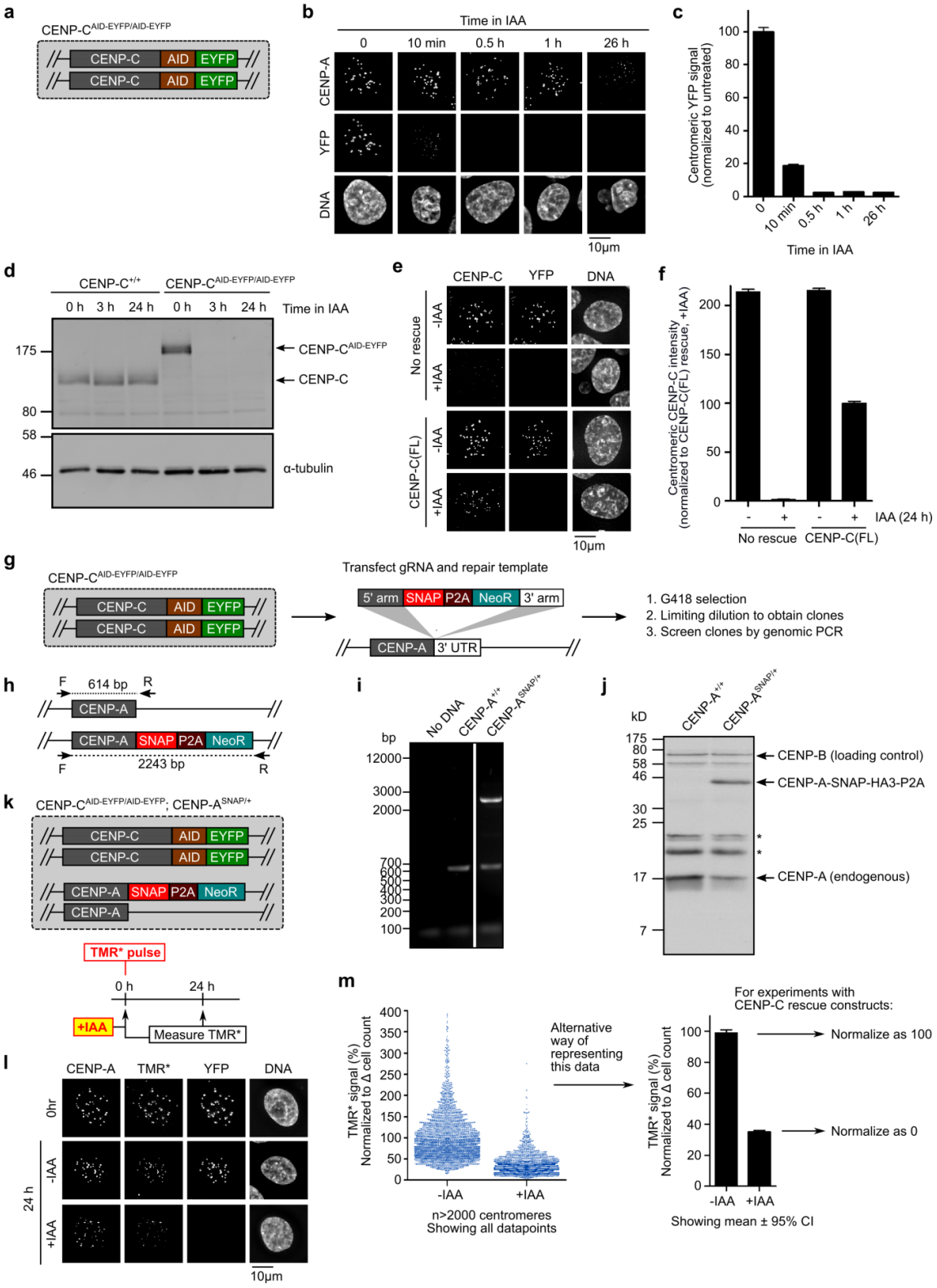


Figure 22. Compromised retention of centromeric CENP-A nucleosomes upon rapid auxin-induced degradation of CENP-C.

(a) Schematic of CENP-C$^{AID\text{-}EYFP/AID\text{-}EYFP}$ cells. **(b)** Timecourse of CENP-C-AID-EYFP signal after various lengths of IAA treatment. **(c)** Quantitation of b. Mean ± 95% confidence interval (n>2000 centromeres in all cases). **(d)** Immunoblot of CENP-C$^{AID\text{-}EYFP/AID\text{-}EYFP}$ cells using anti-CENP-C and anti-α-tubulin after various lengths of IAA treatment. **(e)** CENP-C$^{AID\text{-}EYFP/AID\text{-}EYFP}$ cells expressing full-length, untagged CENP-C rescue construct at the unique FRT site. Prior to IAA treatment, anti-CENP-C antibody detects both the rescue construct and the CENP-C-AID-EYFP at the endogenous gene locus. After IAA treatment, the AID-tagged CENP-C is depleted (as verified by the loss of YFP signal), and anti-CENP-C antibody exclusively detects the rescue construct. **(f)** Quantitation of e. Since both CENP-C-AID-EYFP and the rescue construct are present in the cell prior to IAA addition, there is an expected drop (~2-fold) in centromeric CENP-C signal upon rapid removal of the AID-tagged CENP-C. Mean ± 95% confidence interval (n>2000 centromeres in all cases). **(g)** Schematic for SNAP-tagging CENP-A at its endogenous locus by CRISPR/Cas, in CENP-C$^{AID\text{-}EYFP/AID\text{-}EYFP}$ cells. **(h)** Schematic for screening clones by genomic PCR. Incorporation of the repair template containing SNAP is expected to result in a 2.2 kb PCR product. **(i)** PCR with genomic DNA extracted from parental cells and a heterozygous clone in which one allele in which CENP-A is SNAP-tagged by CRISPR. **(j)** Verification of presence of SNAP-tagged CENP-A by immunoblot with anti-centromere antibodies (ACA). In addition to CENP-A, ACA also recognizes CENP-B, which here serves as a loading control. The heterozygous clone has both endogenous CENP-A as well as SNAP-tagged CENP-A. **(k)** Schematic of the clone used for all TMR* experiments in this study (CENP-C$^{AID\text{-}EYFP/AID\text{-}EYFP}$ cells with CENP-A that is SNAP-tagged at its endogenous locus), and schematic of pulse-chase experiment to measure CENP-A retention at the centromere after complete depletion of CENP-C. **(l)** Representative images from experiment diagrammed in k. **(m)** Left: Quantitation of panel l, shown with all datapoints (n>2000 centromeres in all cases), and displayed as in our prior study (Falk et al., 2015). Right: Plotting these same data as a bar graph showing mean ± 95% confidence interval (Cumming et al., 2007). For TMR* experiments with CENP-C rescue constructs (**Figure 21g-i, Figure 25d,e,** and **Figure 35d-f**), the value of the -IAA condition is normalized as 100%, and the value of the +IAA condition is normalized as 0%.

3.3.2 THE ARGININE ANCHOR OF CENP-CCD STABILIZES CENP-A

We next determined whether or not the stability CENP-C^{CD} imparts to CENP-A nucleosomes is attributable to the CENP-A nucleosome structural transition (Falk et al., 2015, 2016). We reasoned that by reducing the contact points of the CENP-C^{CD} with the H2A-H2B dimer (Kato et al., 2013) that we could generate a version of

CENP-C that could bind to CENP-A nucleosomes but not drive the nucleosome structural transition (Falk et al., 2015, 2016) that we predicted would be central to stabilizing CENP-A at centromeres. We chose two adjacent arginines (R521 and R522) within the CENP-C^{CD} that are proposed to contact an acidic patch of the nucleosome on the surface of the H2A-H2B dimer (Kato et al., 2013), and performed quantitative binding studies on CENP-A nucleosomes in which one histone subunit (H2B) is fluorescently labeled at a site (at the position corresponding to K120) distal to the binding surface of CENP-C^{CD} (**Figure 23a**) We reasoned it was likely that existing isothermal calorimetry data (Kato et al., 2013) (that used canonical nucleosomes with the C-terminal 6 a.a. of CENP-A appended to conventional H3) does not clearly distinguish *altered* binding from a *complete loss* of binding; since the complex binding surface for CENP-C^{CD} on the nucleosome involves 3 different histone subunits and a nucleosome structural transition for *bona fide* CENP-A nucleosomes (Falk et al., 2015, 2016), complicating the interpretation of thermodynamic measurements. Instead, our native PAGE analyses would clearly distinguish unbound, bound, and higher-order aggregates that can form with the WT protein at very high concentrations. We find that while the binding affinity of CENP-C^{CD} R521A is close to that of WT CENP-C^{CD}, CENP-C^{CD} R522A exhibits a 2 to 3-fold decrease in binding affinity (**Figure 24a,b**). The discrete mobility on native PAGE of the CENP-A nucleosome core particle (NCP)-CENP-C^{CD} complex is lost in R522A, but preserved in R521As (**Figure 24a,c**). Therefore, both mutations preserve the ability to bind to the NCP, but that the R522A may perturb the highly ordered nature of the complex that CENP-C^{CD} forms with the NCP. Meanwhile, mutation of one of the hydrophobic residues (W530) proposed to contact the hydrophobic tail of CENP-A abolishes binding to CENP-A nucleosomes (**Figure 23b-e**).

We then measured the effects of these mutations using HXMS, an approach that revealed HX protection that maps unambiguously to the buried center of the nucleosome to a region encapsulating a β-sheet that forms between H2A and H4,

coinciding with the CENP-A nucleosome structural transition conferred by WT CENP-C^{CD} (Falk et al., 2015). HX measures the rate of amide proton exchange along the polypeptide backbone of proteins, and protection occurs through stabilization of H-bonds within secondary structure (Englander, 2006) (i.e., within histone α-helices or between the β-strands that form with loop L1/L2 contacts between histone pairs like CENP-A and H4 (Black et al., 2004; Falk et al., 2015)) or via direct backbone interactions (Englander, 2006). Reconstitution of complexes at the concentrations and high nucleosome saturation required for the clear interpretation of HXMS experiments (see Methods) were achieved with WT, R521A, and R522A versions of CENP-C^{CD} (**Figure 24c**). HXMS analysis revealed that CENP-C^{CD}(R521A) forms a similar complex as wildtype CENP-C^{CD}, with protection of the surface helices of the CENP-A NCP as well as the interior H2A-H4 interface (**Figure 24d**)(Falk et al., 2015). On the other hand, although CENP-C^{CD}(R522A) still binds to CENP-A NCPs (**Figure 24a-c**), its mode of binding is grossly perturbed: it still contacts and stabilizes the H4 α2 helix on the surface of the NCP, but the HX protection is reduced at the other surface helices (one each on CENP-A and H2A)(**Figure 24d**). We interpret these results to mean that the reduced HX protection on the surface of H2A is directly due to the removal of the CENP-C^{CD} arginine anchor. The altered dynamics in HX are extended to reduce protection at a contact site with CENP-A (near CENP-C a.a. 530) that lies between the CENP-C^{CD} N-terminal (CENP-C a.a. 522) nucleosome contact point (on H2A) and its C-terminal (near CENP-C a.a. 535) contact point (on H4)(Kato et al., 2013) (**Figure 24d-f, h**). Most importantly, removal of the arginine anchor by the R522A mutation leads to loss of HX protection at the H2A-H4 interface at the interior of the nucleosome (**Figure 24d, g, i**). Since CENP-C(R522A) binding fails to stabilize the internal H4/H2A interface similar to CENP-C(Δ519-533), which is missing the entire region required for binding to CENP-A nucleosomes, we predicted that both mutants would also fail to confer CENP-A stability. After generating the respective cell lines (**Figure 25a**), and adding IAA to remove CENP-C-AID-YFP, all mutants localize to centromeres to a level

equivalent to the wild type protein (**Figure 25b,c**). We measured retention of TMR*-labeled CENP-A, and found that both CENP-C(Δ519-533) and CENP-C(R522A) were markedly reduced in their ability to retain CENP-A at centromeres, whereas the CENP-C(R521A) mutation had no effect (**Figure 25d,e**). The R522A result is particularly striking, indicating that R522 is the key arginine anchor of CENP-C^{CD} and providing a prime example of how an arginine anchor on a nucleosome binding protein can be a lynchpin for a central biological process such as maintaining centromere identity. Together with our HXMS results, these data strongly indicate that stabilization of the interior of the CENP-A nucleosome requires the H2A-H2B contacts via R522 of CENP-C and that the stability of CENP-A nucleosomes due to CENP-C at functional centromeres can be attributed exclusively to the CENP-C^{CD}.

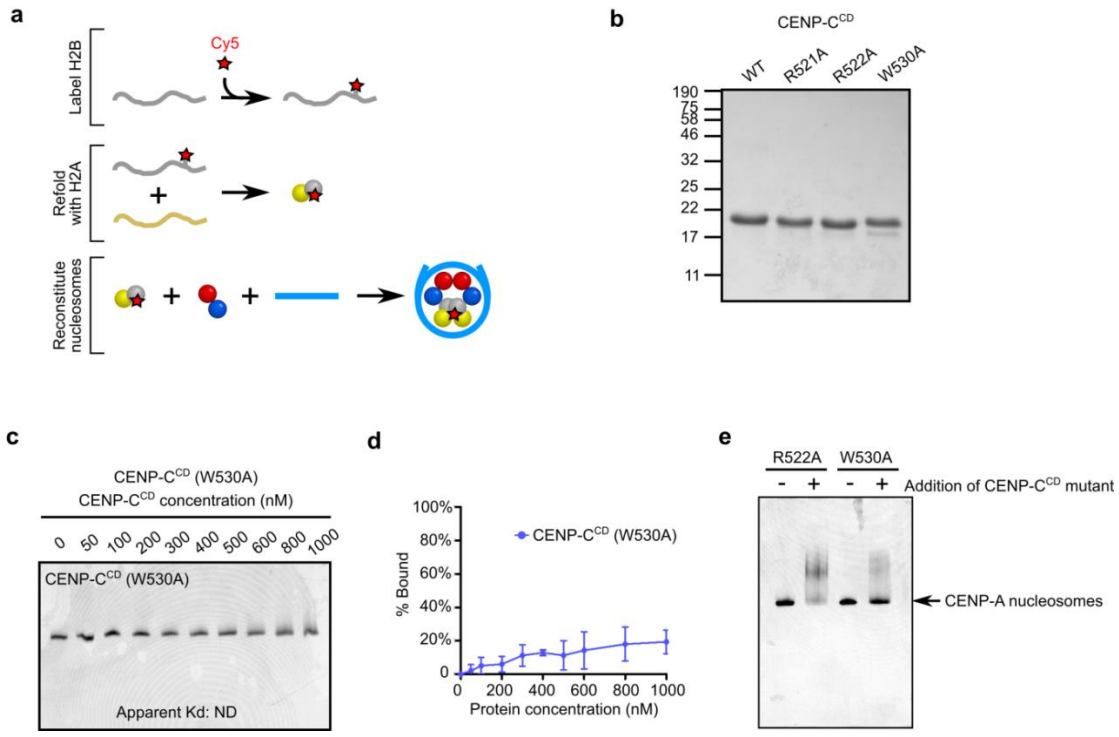


Figure 23. CENP-C^{CD}(W530A) fails to bind to CENP-A nucleosomes.
(a) Schematic for labeling H2B K120C with Cy5, subsequent refolding with H2A to form histone dimers, and reconstitution into nucleosomes. **(b)** SDS-PAGE gels stained with Coomassie Blue of CENP-C^{CD} WT and mutant proteins used for binding assays shown in this figure and **Figure 24**. **(c)** Representative native PAGE analysis of CENP-A NCPs harboring Cy5-labeled histone H2B that have been incubated with

the indicated concentrations of CENP-C^{CD}(W530A). Only low level of binding to the nucleosome was observed for the W530A mutant. **(d)** Quantitation of three independent experiments (values shown are mean ± SD) performed as in panel c. Note that for some data points, the error bars are too small to be visible in the graph. The apparent K_d for W530A was not determined (ND) because of insufficient binding. (e) The W530A mutant does not assemble with CENP-A nucleosomes to an extent that would make an HXMS experiment useful or interpretable. CENP-C^{CD} mutants incubated with CENP-A nucleosomes containing Cy5-H2B. CENP-C^{CD} (R522A) can form a complex with CENP-A nucleosomes (under the same conditions as those used to assemble nucleosomes for the HXMS experiments in **Figure 24**), while at this concentration, CENP-C^{CD} (W530A) displays very little binding to CENP-A nucleosomes (note that the majority of the nucleosomes in the presence of W530A are unbound from the mutant CENP-C protein). The high proportion of unbound nucleosomes would dominate the HXMS experiment and would not be capable for making a useful comparison to wild type or R522A versions of CENP-C.

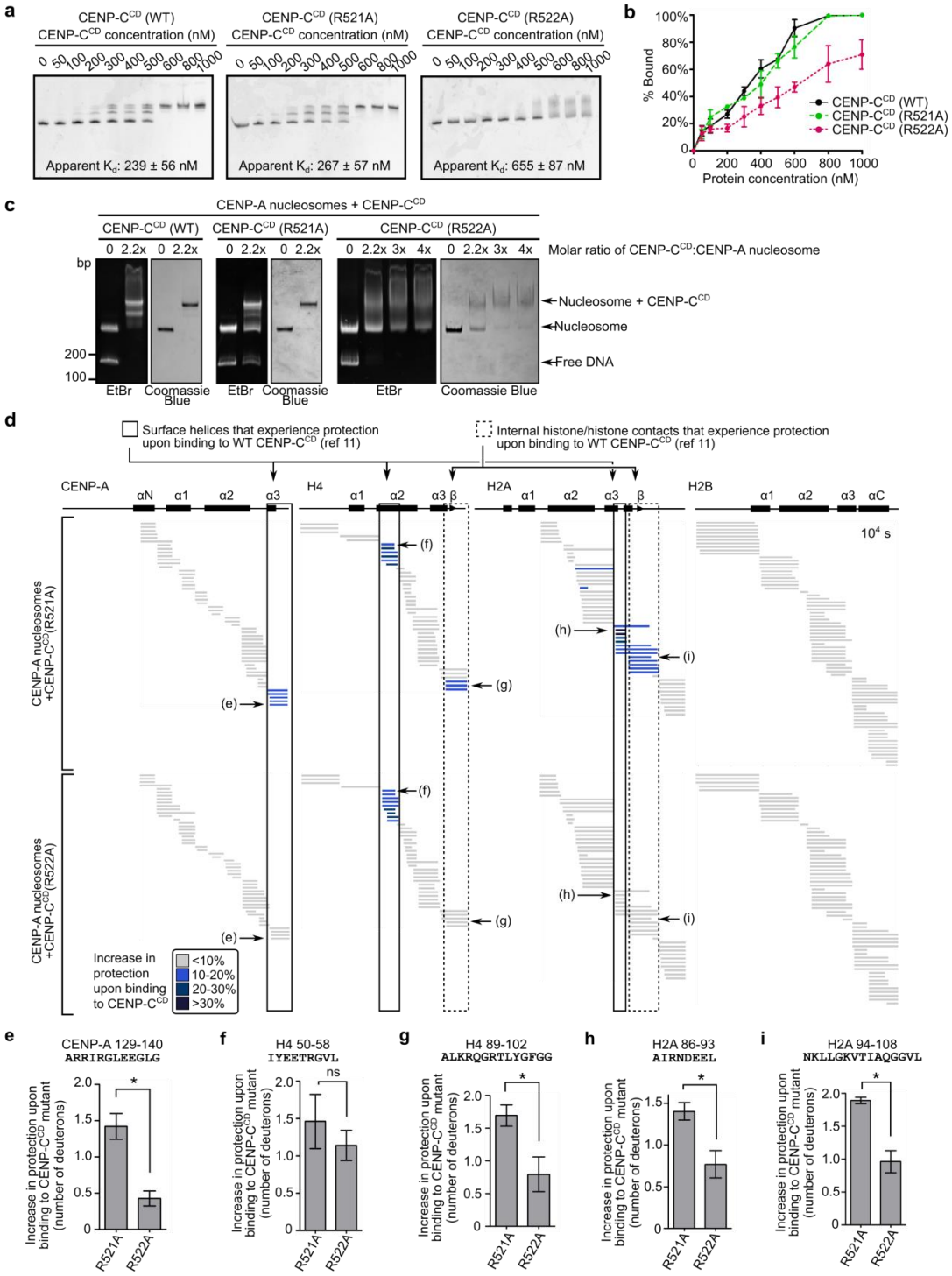


Figure 24. The arginine anchor of CENP-C^{CD} is critical for the CENP-A nucleosome structural transition.

(a) Representative native PAGE analysis of CENP-A NCPs harboring Cy5-labeled histone H2B that have been incubated with the indicated concentrations of CENP-C^{CD} (WT or the indicated mutants). Each reaction contains 200 nM nucleosomes. Cy5

fluorescence was detected on a Typhoon Phosphorimager, and CENP-C binding retards the mobility. Both WT and R521A show crisp shifts to bands with 1 or 2 copies of CENP-C bound to the nucleosome. R522A exhibits a more smeary appearance when bound to the CENP-A nucleosome (see also panel c), and the species with a single molecule of CENP-C^{CD}(R522A) was not clearly resolved. Listed on the graphs are apparent K_d values for these binding experiments (values shown are mean ± SD; n = 3). **(b)** Quantitation of three independent experiments (values shown are mean ± SD) performed as in panel a. Note that for some data points, the error bars are too small to be visible in the graph. **(c)** CENP-A NCPs in complex with WT or mutant CENP-C^{CD}, as assessed by native PAGE stained with ethidium bromide (EtBr) and then Coomassie Blue. **(d)** HXMS of all histone subunits of the CENP-A NCP from a single timepoint (10^4 s), showing regions that exhibit additional protection from HX upon binding of CENP-C^{CD}(R521A) or CENP-C^{CD}(R522A). Each horizontal bar represents an individual peptide, placed beneath schematics of secondary structural elements of the CENP-A nucleosome. When available, we present the data from all measurable charge states of each of the unique peptides (here and in the similarly formatted plots in the experiments presented in **Figure 29** and **Figure 32**). **(e-i)** Representative peptides from various histone regions, comparing protection from exchange when the nucleosome is bound to CENP-C^{CD} R521A versus R522A, showing faithful detection of differences between the two mutants across multiple replicate experiments (plotted as the mean ± SD; n = 3). Asterisks denotes differences that are statistically significant ($p<0.05$; Student's t-test).

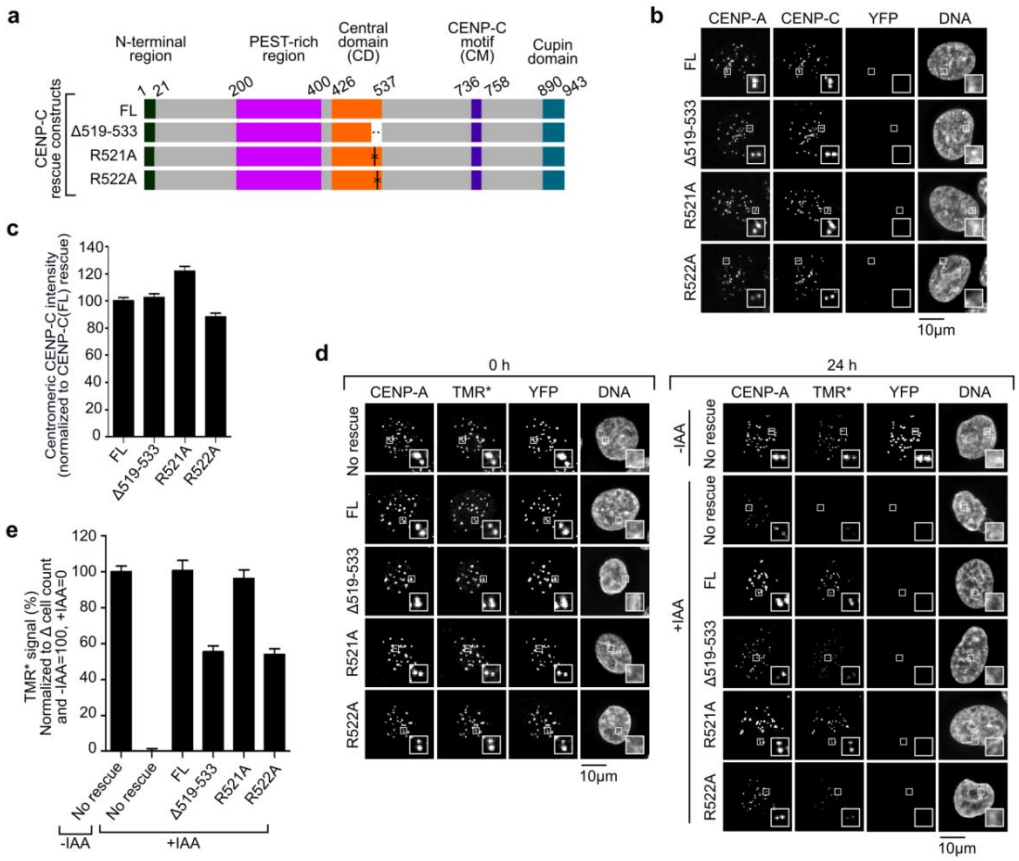


Figure 25. The arginine anchor of CENP-C^{CD} is required for CENP-A nucleosome stability at centromeres.
(a) Rescue constructs constitutively expressed at the unique FRT site in CENP-C$^{AID\text{-}EYFP/AID\text{-}EYFP}$ cells. FL=full length. **(b)** Representative images showing localization of CENP-C rescue constructs in CENP-C$^{AID\text{-}EYFP/AID\text{-}EYFP}$ cells after 24 h of IAA treatment. Scale bar indicates 10 μm. **(c)** Quantitation of b. **(d)** Representative images showing CENP-A retention as measured by TMR* assay in cells, similar to schematic in **Figure 21g**. Scale bar indicates 10 μm. **(e)** Quantitation of d. All graphs are shown as mean ± 95% confidence interval (n>2000 centromeres in all cases).

3.3.3 CENP-N^{NT} FASTENS CENP-A TO NUCLEOSOMAL DNA

Although a substantial component (~50%) of CENP-A retention at centromeres is attributable to the CENP-A nucleosome structural transition conferred by CENP-C^{CD} (**Figure 25d,e**), complete removal of CENP-C leads to a more pronounced defect (**Figure 25d,e**, and **Figure 22k-m**), implying that an interacting partner outside of the CENP-C^{CD} also contributes to CENP-A retention. We next considered the CENP-L-N complex because CENP-N^{NT} directly contacts the CENP-A nucleosome (Carroll et al., 2009; McKinley et al., 2015), but additionally requires

CENP-C for its centromere localization (McKinley et al., 2015). Prior work found that the interaction of CENP-L-N with CENP-C occurs in a region (CENP-C a.a. 235-509)(McKinley et al., 2015) overlapping with the CD (CENP-C a.a. 426-537) but outside of the histone contact residues (CENP-C a.a. 519-533). We found that CENP-$C^{235-425}$ binds to CENP-L-N at similar levels to CENP-$C^{235-509}$ (**Figure 26a**), and that CENP-$C^{235-352}$ was sufficient for this interaction (**Figure 28a**). Consistent with this, CENP-C(Δ519-533) almost completely rescues CENP-L-N localization (**Figure 26b,c** and **Figure 28b,c**). Taken together, our findings suggest that CENP-C and CENP-N bind to each other with interaction interfaces that are distinct from their nucleosome interaction interfaces. Thus, in principle, CENP-N^{NT} and CENP-C^{CD} could both contribute to the stability of CENP-A nucleosomes at human centromeres.

CENP-N was previously known to directly bind to the CENP-A nucleosome through its N-terminal domain (a.a. 1-289) (Carroll et al., 2009), and we further narrowed down the minimal interaction down to a.a. 1-240, and CENP-N(1-240) is heretofore referred to as CENP-N^{NT} (**Figure 27**)(McKinley et al., 2015).

Prior work found that the interaction of CENP-L/N with CENP-C occurs in a region (a.a. 235-509)(McKinley et al., 2015) overlapping with the CD (a.a. 426-537) but outside of the histone contact residues (a.a. 519-533). We found that CENP-$C^{235-425}$ binds to CENP-L/N at similar levels to CENP-$C^{235-509}$ (**Figure 26a**), and that CENP-$C^{235-352}$ was sufficient for this interaction (**Figure 28a**). Consistent with this, CENP-C(Δ519-533) almost completely rescues CENP-L/N localization (**Figure 26b,c**, and **Figure 28b,c**). Taken together, our findings suggest that CENP-C and CENP-N bind to each other with interaction interfaces that are distinct from their nucleosome interaction interfaces. Thus, in principle, CENP-N^{NT} and CENP-C^{CD} could both contribute to the stability of CENP-A nucleosomes at human centromeres.

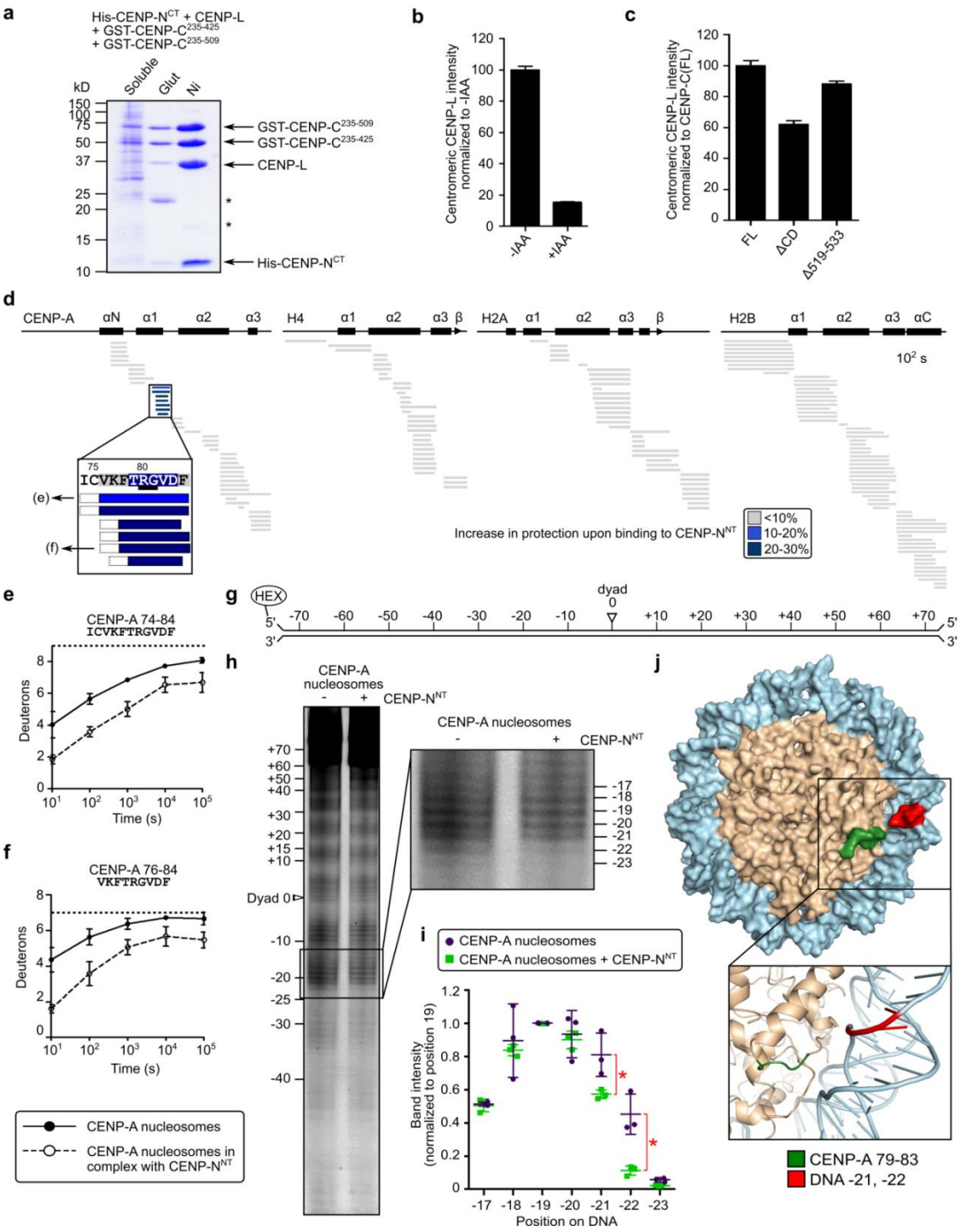


Figure 26. CENP-N^{NT} crossbridges CENP-A to DNA.
(a) Coomassie Blue-stained SDS-PAGE of co-purification with described protocol(McKinley et al., 2015) of CENP-L/His-CENP-N^{CT} with GST-CENP-C$^{235\text{-}509}$ and GST-CENP-C$^{235\text{-}425}$ by glutathione-agarose (Glut) or Nickel-NTA-agarose (Ni). **(b)** Localization of CENP-L-N in CENP-C$^{AID\text{-}EYFP/AID\text{-}EYFP}$ cells before and after 24 Eh of IAA treatment, assessed using anti-CENP-L(McKinley et al., 2015) (See Supplementary Fig. 3b for images). **(c)** Localization of CENP-L-N in CENP-C$^{AID\text{-}EYFP/AID\text{-}EYFP}$ cells

constitutively expressing the rescue constructs CENP-C(FL), CENP-C(ΔCD), or CENP-C(Δ519-533), after 24 h of IAA treatment. (See Supplementary Fig. 3c for images) All graphs are shown as mean ± 95% confidence interval (n>2000 centromeres in all cases). **(d)** HXMS of all histone subunits of the CENP-A NCP from a single timepoint (10^2 s), showing protection at CENP-A(79-83) upon binding to CENP-N^{NT}. The first two residues of each peptide are boxed in dashed black lines because exchange of the first two backbone amide protons cannot be measured(Bai et al., 1993). **(e,f)** Representative peptides spanning the CENP-A surface bulge over the time course. The maximum number of deuterons possible to measure by HXMS for each peptide is shown by the dotted line. All peptides are plotted at every timepoint as mean ± SD from triplicate experiments. Note that for some data points, the error bars are too small to be visible in the graph. (g) Schematic of the 5'-fluorescently labeled 147 bp α-satellite DNA sequence used in footprinting experiments. (h) Representative hydroxyl radical footprinting experiment of CENP-A nucleosomes vs. CENP-A nucleosomes in complex with CENP-N^{NT}, with inset showing magnification of positions -17 to -23. (i) Quantitation of band intensities from 3 independent experiments, shown as mean ± SD normalized to DNA position -19 (this position was chosen because it was expected to be very exposed for hydroxyl radical-mediated cleavage with and without CENP-N^{NT}). Asterisks denotes differences that are statistically significant ($p<0.05$; Student's t-test). (j) A molecular model of the CENP-A nucleosome (PDB 3AN2)(Tachiwana et al., 2011) in which the DNA sequence was modified(Falk et al., 2016) to that used in the footprinting experiment: CENP-A a.a. 79-83 is labeled in green, and DNA positions -21 and -22 are labeled in red.

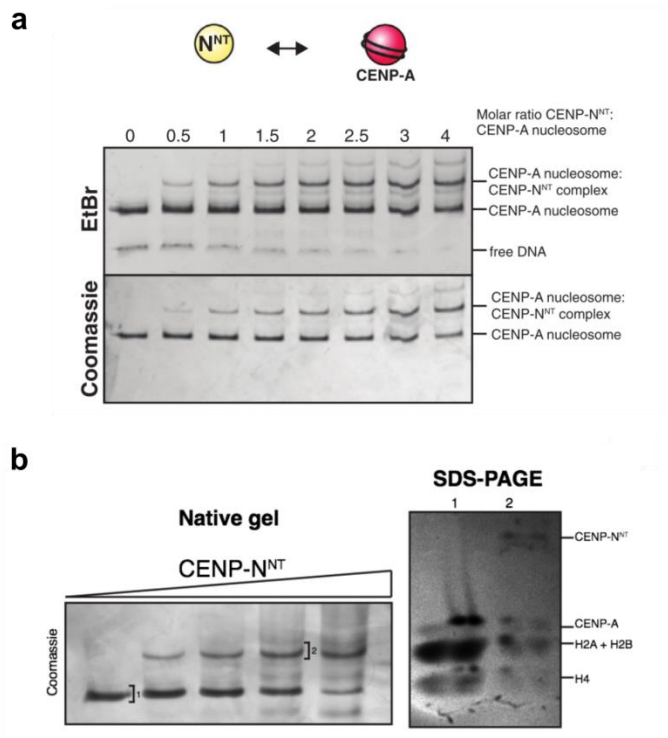


Figure 27. CENP-N(1-240) is sufficient to bind to the CENP-A nucleosome.

(a) Native gel showing binding of CENP-N(1-240)-His (referred to as CENP-N^{NT}) to CENP-A nucleosomes **(b)** 2-dimensional gel analysis of binding of CENP-N^{NT} to CENP-A nucleosomes. Left: native gel of CENP- A nucleosomes with increasing amounts of CENP-N^{NT}. The bands marked 1 and 2 were excised and run on an SDS-PAGE gel (right) to confirm the species (histones and/or CENP-N^{NT}) in the native gel band.
[This figure consists of my contribution to McKinley et al., 2015]

To measure the location and magnitude of the stability conferred by CENP-N^{NT} to the CENP-A NCP, we performed HXMS on the assembled complex (**Figure 26d**, and **Figure 28d**). The only region on the entire NCP where we detected additional protection from HX in the presence of CENP-N^{NT} is within the CENP-A targeting domain (CATD)(Black et al., 2004), at a discrete portion that spans the C-terminal region of the α1 helix and the N-terminal portion of L1 (**Figure 26d-g**, and **Figure 28e-j**). This location corresponds to a major surface structural feature unique to CENP-A nucleosomes: a bulge of opposite charge as the same site of canonical nucleosomes containing H3 (**Figure 26e**)(Luger et al., 1997; Sekulic et al., 2010; Tachiwana et al., 2011). The HXMS results are consistent with previous work demonstrating that CENP-N recognizes the CATD (Carroll et al., 2009; McKinley et al., 2015). The region of HX protection is centered around residues Arg80 and Gly81 on CENP-A (**Figure 26d**, inset), where mutations disrupt CENP-N binding (Fang et al., 2015).

We next considered how CENP-N specifically recognizes CENP-A when it is in a histone complex wrapped with DNA and how this might contribute to its function at centromere. We employed a well-established approach for nucleosomes (Hayes et al., 1990; Syed et al., 2010; Tullius, 1988), recently extended to CENP-A nucleosomes assembled with a synthetic positioning sequence (Roulland et al., 2016) that employs hydroxyl radical mediated cleavage of DNA. We used the same natural CENP-A nucleosome positioning sequence from human centromeres (Falk et al., 2015; Hasson et al., 2013) used in our HXMS experiments, but where it is end-labeled (Falk et al., 2016) for hydroxyl radical footprinting (**Figure 26g**). CENP-A nucleosome positioning is strong enough to readily detect the expected ~10 bp periodicity of protection from hydroxyl radical cleavage caused by each superhelical

turn of the DNA on the surface of the histone octamer (**Figure 26h**). CENP-N^{NT} binding does not alter the phasing, but there is very strong added protection at -21 and -22 nt from the dyad axis of the nucleosome (**Figure 26h,i**). This location is immediately adjacent to the bulged L1 of CENP-A that is protected from HX (**Figure 26j**). Thus, we envision a continuous binding surface that spans and crossbridges CENP-A and nucleosomal DNA.

This raised the questions of whether the nucleosome binding surface of CENP-N^{NT} is an extended, unstructured segment, as in CENP-C^{CD}, or a well-folded domain. Fortunately, our HXMS experiments on the CENP-A nucleosome complex with CENP-N^{NT} yielded near complete coverage of both the histone fold domains of each nucleosome subunit (**Figure 26d**) and CENP-N^{NT}, itself (**Figure 29**), for which there is little or nothing known regarding its structure and dynamics. For CENP-N^{NT}, we found substantial HX protection for the CENP-N^{NT} molecule alone (**Figure 29**), suggesting it is a folded domain. This protection was markedly increased—taking 100-1000 times as long to reach the same level of HX—upon binding to CENP-A NCPs (**Figure 29**). The dramatic protection from HX on CENP-N^{NT} upon binding to the CENP-A NCP extended through its entire N-terminal ~200 a.a (**Figure 29**, and **Figure 30a-g**). Residues ~209-240 were disordered both before and after binding to CENP-A NCPs (**Figure 29a** and **Figure 30h-i**), indicating that this region is unlikely to be involved in binding. Indeed, a further truncation of CENP-N^{1-205} retained the ability to bind to CENP-A nucleosomes (**Figure 30j**). Since there are key residues for this interaction (R11 and R196)(Carroll et al., 2009) at each end of this domain, it is likely that the folded nature of the CENP-N^{NT} brings together key residues that form the binding surface with CENP-A NCPs. Indeed, important residues for CENP-A nucleosome binding are found at various locations across this region of CENP-N (Carroll et al., 2009).

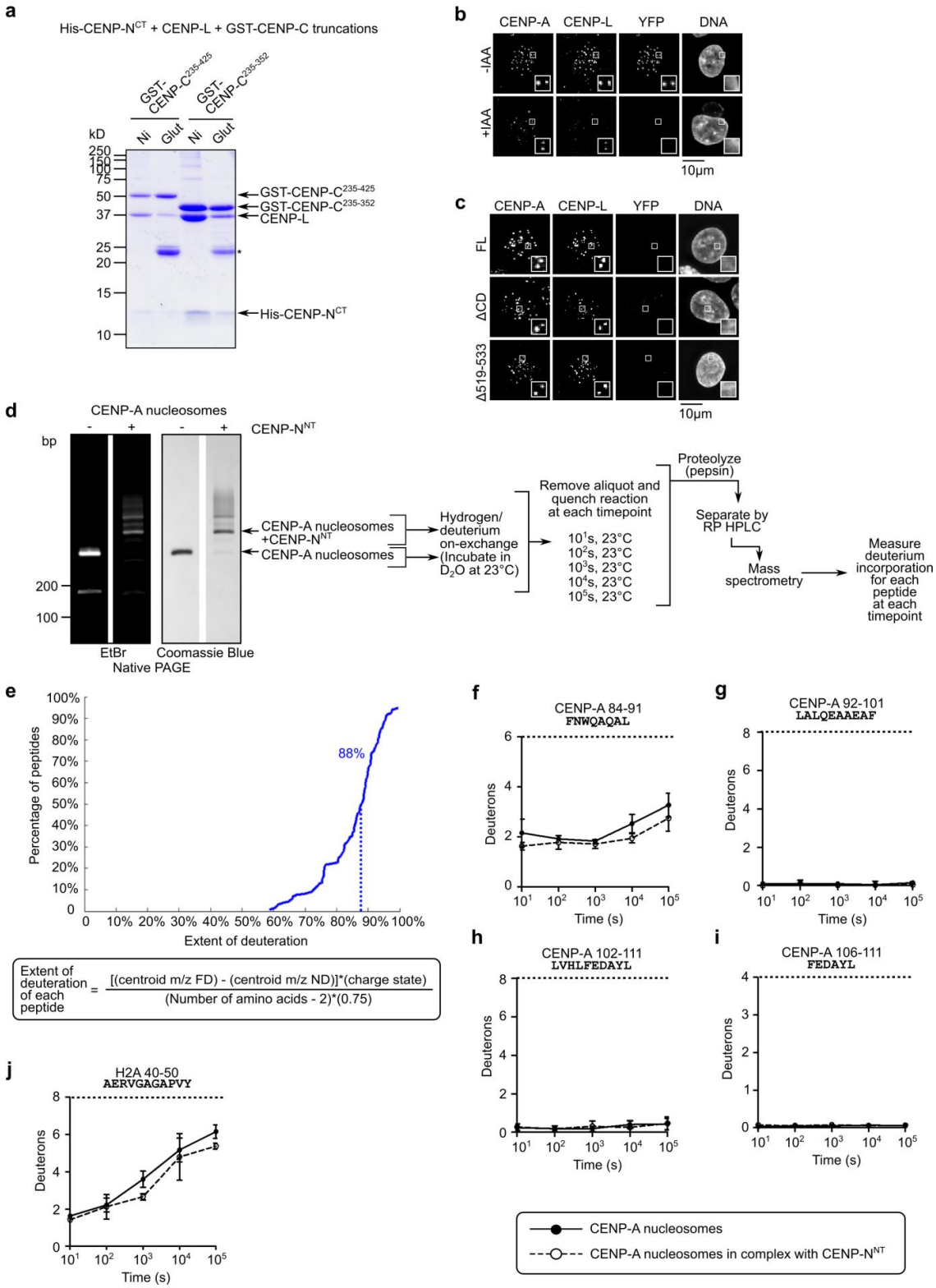


Figure 28. CENP-L/N^{CT} binds CENP-C$^{235-352}$, and CENP-N^{NT} binds the CENP-A nucleosome surface bulge.

(a) SDS-PAGE showing co-purification performed as described(McKinley et al., 2015) of fragments of GST-CENP-C with CENP-L/His-CENP-N^{CT}, demonstrating that

CENP-C$^{235\text{-}352}$ is sufficient for the interaction with CENP-L/N^{CT}. **(b)** Representative images showing localization of CENP-L/N in CENP-C$^{AID\text{-}EYFP/AID\text{-}EYFP}$ cells before and after 24 h of IAA treatment, assessed using anti-CENP-L(McKinley et al., 2015). See quantitation in Fig. 4b. **(c)** Representative images showing localization of CENP-L/N in CENP-C$^{AID\text{-}EYFP/AID\text{-}EYFP}$ cells constitutively expressing the rescue constructs CENP-C(FL), CENP-C(ΔCD), or CENP-C(Δ519-533), after 24 h of IAA treatment, assessed using anti-CENP-L(McKinley et al., 2015) . See quantitation in **Figure 26c**. **(d)** CENP-A NCPs alone and in complex with CENP-N^{NT} as assessed by native PAGE stained with EtBr or Coomassie Blue, and schematic for HXMS experiment. **(e)** Evidence that our HXMS experiments have minimal back-exchange. Cumulative distribution curve of a representative fully deuterated (FD) sample (see the Methods for a description of how FD samples are prepared), showing the extent of deuteration of all peptides compared to the theoretical maximum amount of deuteration of each peptide (i.e., if every amide proton were exchanged for a deuteron). The median deuteration was ~88% for the FD sample, therefore the back-exchange after the quench step was only ~12%, which is well within the optimal range (better than most published HXMS experiments(Walters et al., 2012)). **(f-i)** Representative peptides spanning the α2 helix of CENP-A. We note that our data do not exclude that the binding site might extend to adjacent surface residues on the N-terminal portion of α2 helix(Sekulic and Black, 2009; Tachiwana et al., 2011) that also lies within the CATD. Even prior to CENP-N^{NT} binding, this particular region (CENP-A a.a. ~85-111) of the α2 helix of CENP-A undergoes HX too slowly (>12 days(Black et al., 2007a)) for us to detect. All peptides are plotted at every time point as mean ± SD from triplicate experiments. At every other location of the folded core of the octameric CENP-A nucleosome, however, HX is fast enough that we could have measured changes if imparted by CENP-N^{NT}, but none were observed (See Fig. 4d). Note that for some data points, the error bars are too small to be visible in the graph. **(j)** A peptide of H2A a.a. 40-50, showing mild protection upon binding to CENP-N.

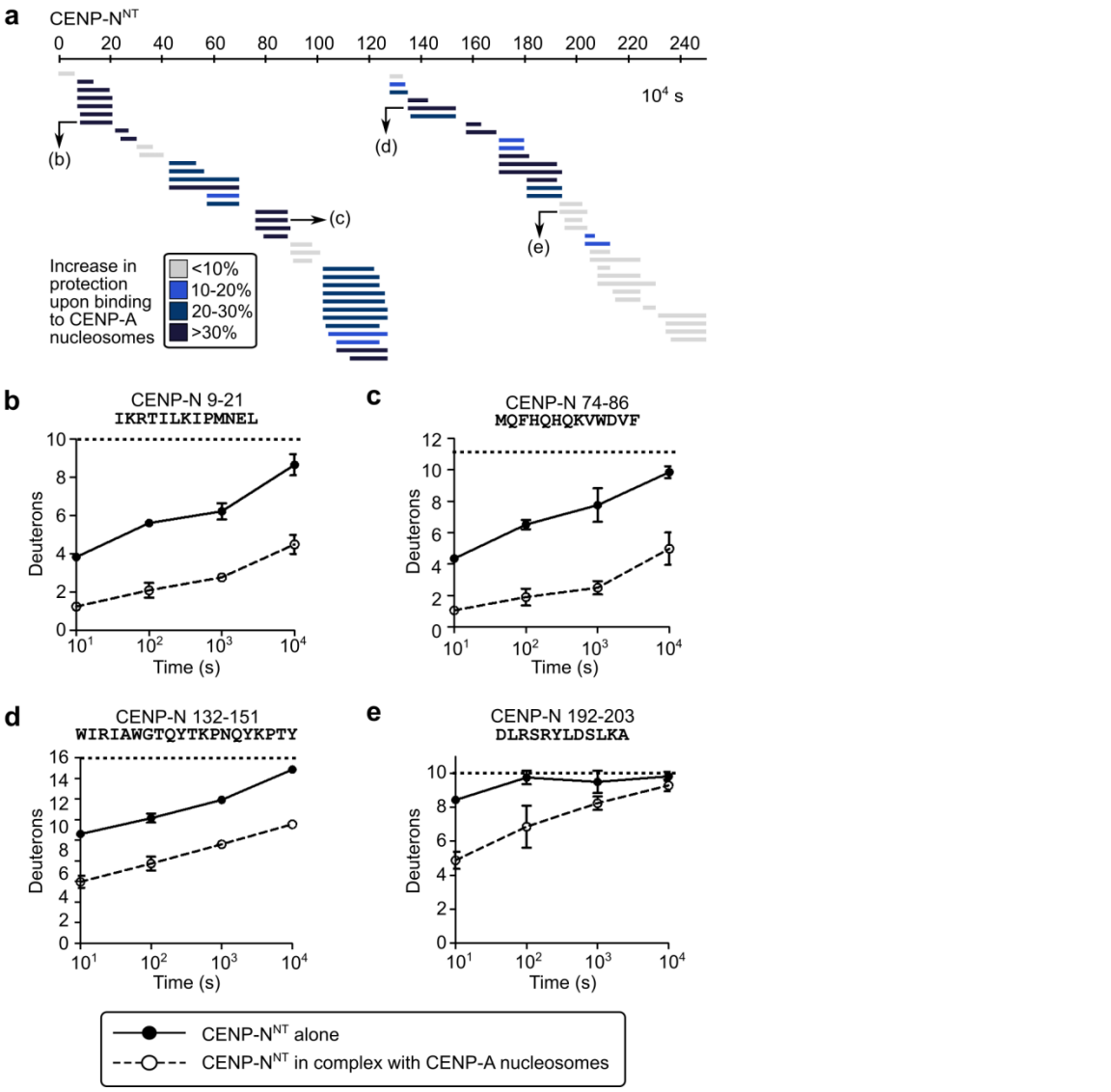


Figure 29. CENP-N^{NT} undergoes global stabilization upon binding to the CENP-A nucleosome.
(a) HXMS of CENP-N^{NT} from a single timepoint (10^4 s), showing substantial protection from HX spanning the ~200 a.a. domain upon binding to CENP-A NCP.
(b-e) Representative peptides spanning CENP-N^{NT} over the timecourse. All peptides are plotted at every timepoint as mean ± SD from triplicate experiments. Note that for some data points, the error bars are too small to be visible in the graph.

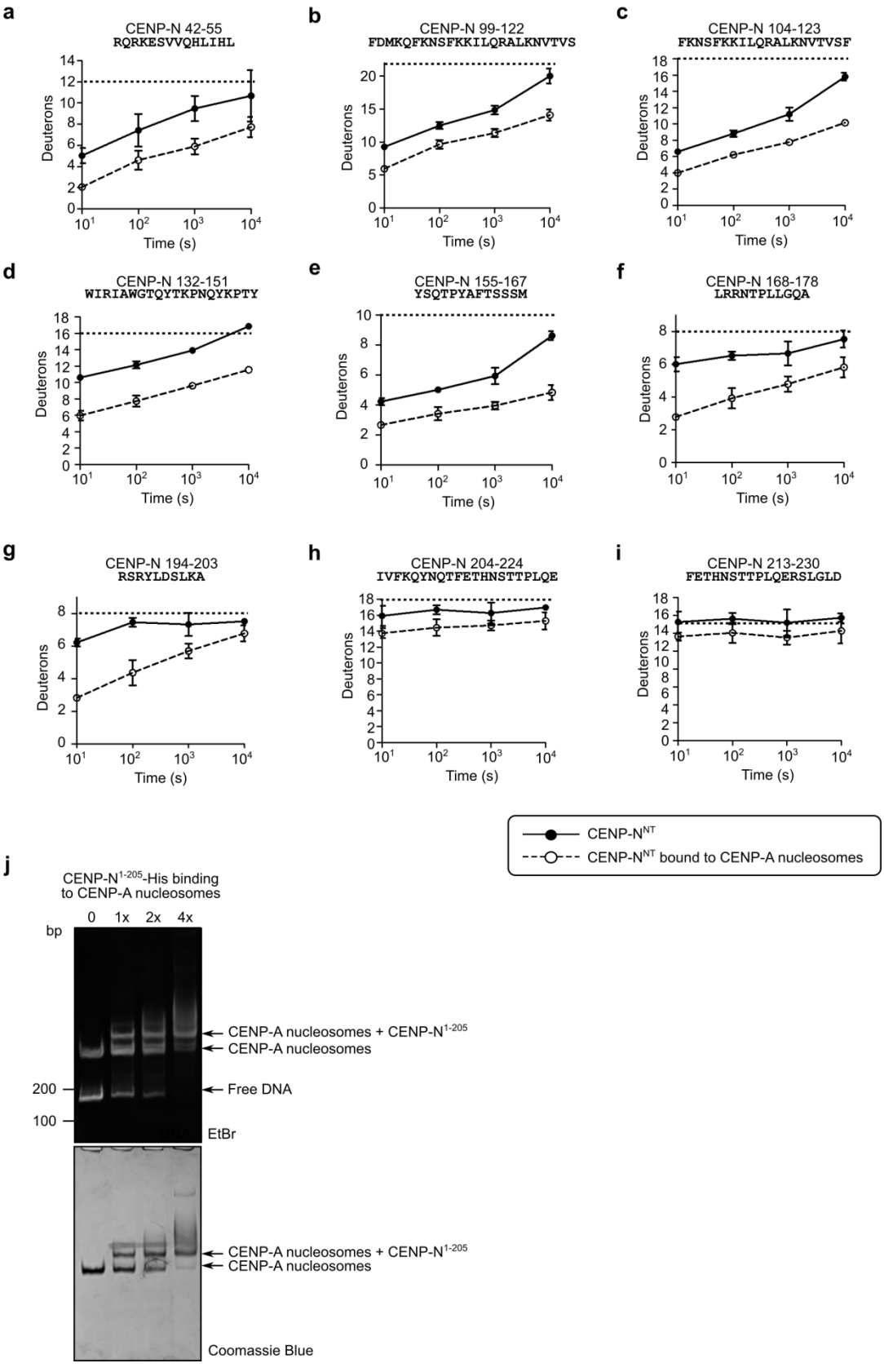

a
CENP-N 42-55
RQRKESVVQHLIHL
b
CENP-N 99-122
FDMKQFKNSFKKILQRALKNVTVS
c
CENP-N 104-123
FKNSFKKILQRALKNVTVSF
d
CENP-N 132-151
WIRIAWGTQYTKPNQYKPTY
e
CENP-N 155-167
YSQTPYAFTSSSM
f
CENP-N 168-178
LRRNTPLLGQA
g
CENP-N 194-203
RSRYLDSLKA
h
CENP-N 204-224
IVFKQYNQTFETHNSTTPLQE
i
CENP-N 213-230
FETHNSTTPLQERSLGLD
Deuterons
Time (s)
CENP-N^NT
CENP-N^NT bound to CENP-A nucleosomes
j
CENP-N^1-205-His binding
to CENP-A nucleosomes
0 1x 2x 4x
bp
200
100
CENP-A nucleosomes + CENP-N^1-205
CENP-A nucleosomes
Free DNA
EtBr
CENP-A nucleosomes + CENP-N^1-205
CENP-A nucleosomes
Coomassie Blue

Figure 30. The N-terminal 205 amino acids of CENP-N constitute its minimal nucleosome-binding domain.
(a-i) Representative peptides spanning various regions of CENP-N^{NT}. CENP-N^{NT} experiences substantial protection from HX across its entire first ~200 amino acids (a-g), indicating that it consists of a folded domain that becomes globally rigidified upon binding to CENP-A NCPs. The residues 206-230 (h,i) reach full deuteration even at the earliest timepoint, and show no difference upon binding to CENP-A NCPs, which indicates that this is a disordered region potentially dispensable for binding to CENP-A NCPs. The maximum number of deuterons possible to measure by HXMS for each peptide is shown by the dotted line. All peptides are plotted at every time point as mean ± SD from triplicate experiments. Note that for some data points, the error bars are too small to be visible in the graph.
(j) Guided by HXMS data (a-i), we further truncated CENP-N^{NT} into just its first 205 amino acids. Native PAGE shows CENP-A NCPs with increasing molar ratios of CENP-N^{1-205}, stained either with EtBr or Coomassie Blue. As expected, CENP-N^{1-205} is sufficient for binding to CENP-A NCPs, supporting the notion that this is the minimal nucleosome-binding domain on CENP-N.

3.3.4 The Core Centromeric Nucleosome Complex (CCNC)

Since the HX protection on CENP-A NCPs from CENP-N^{NT} is discrete (**Figure 26d**) at a nucleosomal surface contact point that remains accessible after CENP-C^{CD} binding (Falk et al., 2015; Kato et al., 2013), it seemed reasonable to reconstitute nucleosome complexes with both domains bound simultaneously. Using established conditions that generate a complex with one copy of CENP-C^{CD} bound to each face of the nucleosome (Falk et al., 2015; Kato et al., 2013), we added increasing amounts of CENP-N^{NT} (**Figure 31a**). We observed a concentration-dependent and stepwise formation of complexes where one and two copies of CENP-N^{NT} bound to the CENP-A NCP complex containing two copies of CENP-C^{CD} (**Figure 31a,b**). The complexes were stable through native PAGE analysis, and the dominant species contained equimolar amounts (i.e. two copies each) of each core histone (CENP-A, H4, H2A, and H2B), CENP-C^{CD}, and CENP-N^{NT} (**Figure 31b**). We term this complex the Core Centromeric Nucleosome Complex (CCNC)(**Figure 31c**, **Figure 32a,c**). The complex was purified by preparative native PAGE (**Figure 32c**) and is also stable through

sucrose gradient (**Figure 32d,e**), indicating that the CCNC is stable throughout the lengthy (several hours) separation, even with no gel matrices involved whatsoever.

CENP-A nucleosomes within the CCNC experience protection from HX at multiple sites (**Figure 31e,f**, and **Figure 32b,f-i**) corresponding to the additive contributions of CENP-C^{CD} (Falk et al., 2015) and CENP-N^{NT} (**Figure 31d-f**). Furthermore, the rigidity conferred to CENP-N is measured out to 100,000 s of exchange, and exhibits clear EX2 behavior at all timepoints—without any evidence of bimodal peaks or any other fast exchanging species that could have corresponded to a substantially populated unbound, unprotected state—thus providing unambiguous evidence that the complex is stable in solution even on timescales of ~28 h (**Figure 31d**). The CCNC exhibits surface protection on CENP-A, H4, and H2A and protection at the internal interhistone H2A-H4 β-sheet that are all conferred by CENP-C^{CD} (Falk et al., 2015), as well as the surface bulge protection at the α1-helix and L1 conferred by CENP-N^{NT} (**Figure 31e,f** and **Figure 32g-i**). The discrete HX protection pattern emphasizes the specific nature of CCNC assembly in solution. CENP-C^{CD}, itself, undergoes rapid HX, consistent with CENP-C^{CD} existing as a primarily linear polypeptide lacking defined secondary structure(Kato et al., 2013), although there is reduced HX at the earliest timepoints within ~a.a. 515-537 when bound to CENP-A nucleosomes (**Figure 33**). Within the CCNC, CENP-N^{NT} still experiences massive slowing of HX across *most* of its folded nucleosome-binding domain, with exception of its a.a. 99-122 region, suggesting that its mode of nucleosome binding could be altered at that specific location by the co-presence of CENP-C^{CD} (**Figure 34**). Thus, in the context of the CCNC, each non-histone subunit acts in a complementary way to stabilize the particle: CENP-C^{CD} binding directs a structural transition of the nucleosome that stabilizes the interior of the octameric histone core and stabilizes surfaces helices on three of the histone subunits (**Figure 31e,f**), whereas CENP-N^{NT} stabilizes the CENP-A surface bulge on the nucleosome surface (**Figure 31e,f**) and crossbridges it to the adjacent DNA (**Figure 31g-i**).

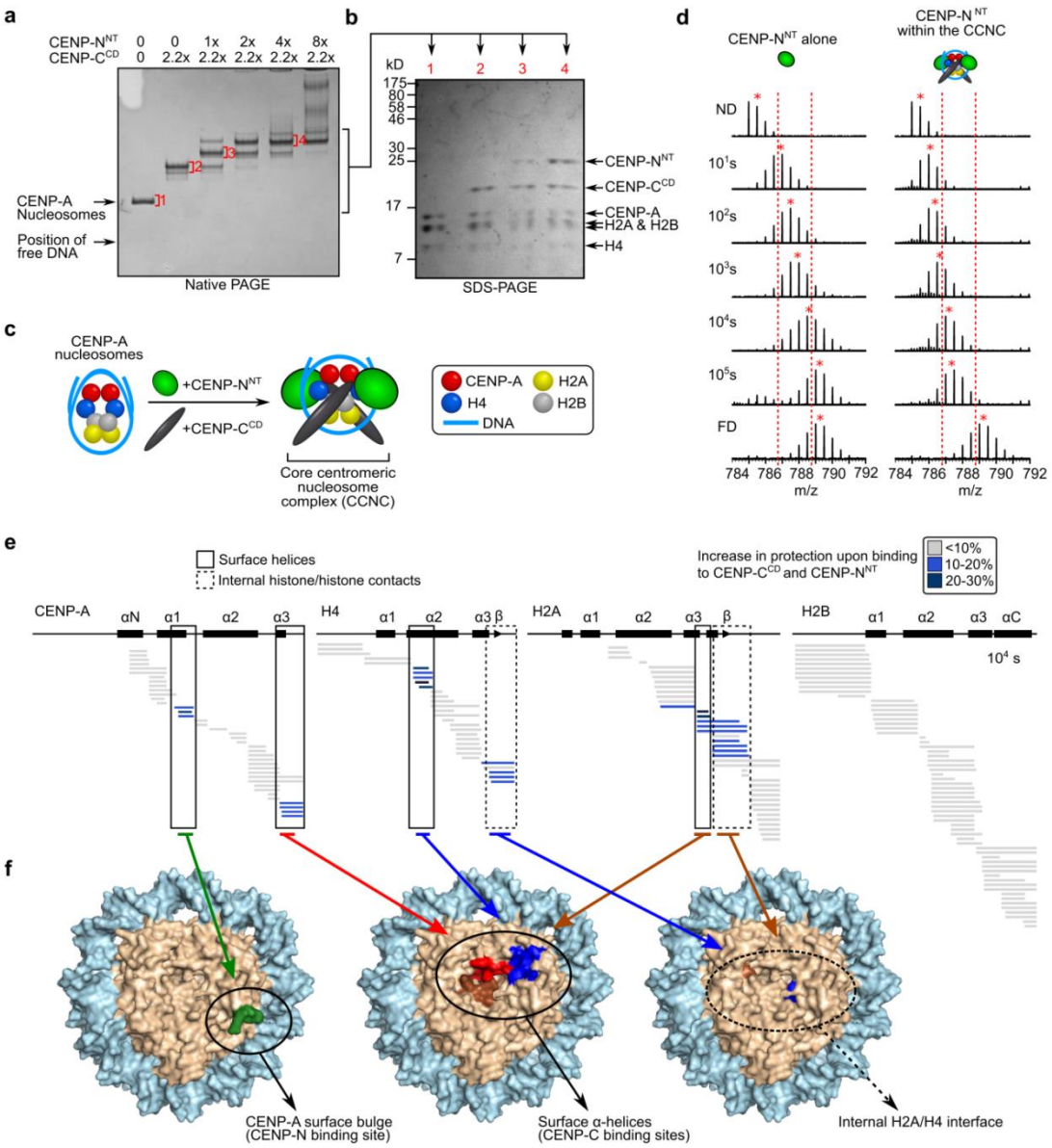


Figure 31. CENP-C^{CD} and CENP-N^{NT} simultaneously bind to the same CENP-A NCP and generate internal and surface stability.
(a) Coomassie Blue-stained native PAGE of binding reactions with CENP-N^{NT} and CENP-C^{CD} and CENP-A NCPs. **(b)** Indicated bands from native PAGE excised and run on SDS-PAGE. **(c)** Schematic of formation of the core centromeric nucleosome complex (CCNC). **(d)** A representative peptide of CENP-N^{NT} (a.a. 9-21) that shows substantial protection upon binding to CENP-A nucleosomes. The peptide is shown from CENP-N^{NT} alone (left) vs. as part of the core centromeric nucleosome complex (CCNC) (right). Dotted red lines serve as guideposts to highlight the differences in m/z shifts between the two samples. A red asterisk denotes the centroid location of each peptide envelope, and the numerical value in blue indicates the centroid mass of the peptide envelope. It is important to note that this peptide exhibits clear EX2 behavior at all timepoints when part of the CCNC (without any evidence of bimodal

peaks), indicating that this complex is stable in solution even on timescales of 100,000 s (~28 hr). **(e)** HXMS of all histone subunits of the CENP-A NCP from a single timepoint (10^4 s). **(f)** Regions showing substantial protection from HX mapped onto the structure of the CENP-A NCP (PDB ID 3AN2). Left: the exposed CENP-A bulge, to which CENP-N binds. Middle: the surface helices to which CENP-C binds. Right: internal histone-histone contacts that undergo stability upon CENP-C binding.

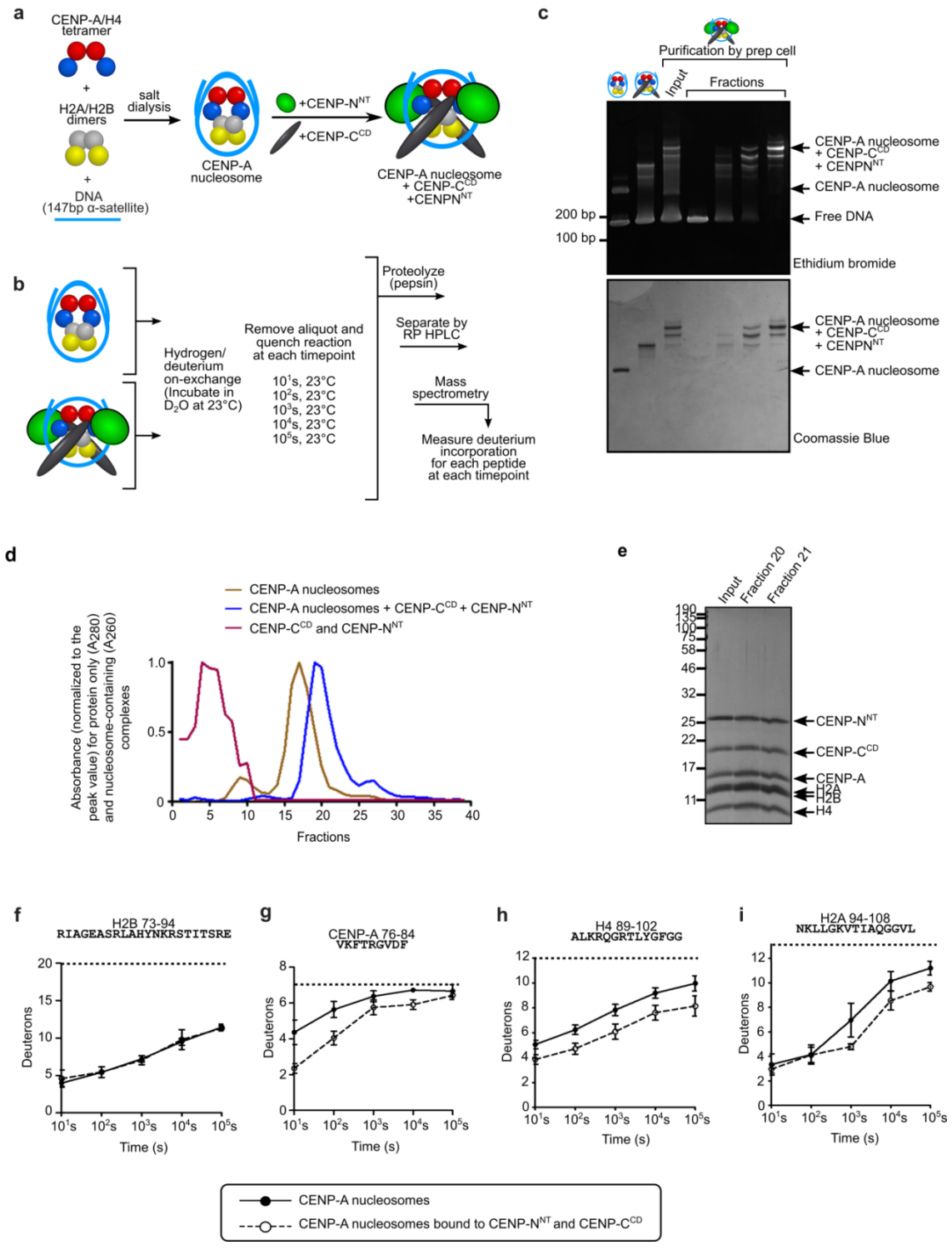


Figure 32. CENP-A NCPs in complex with both CENP-N^{NT} and CENP-C^{CD} experience additive HX protection.
(a,b) Experimental scheme for HXMS of the CENP-A NCP in complex with both CENP-N^{NT} and CENP-C^{CD}. **(c)** Purification of the complex of CENP-A nucleosomes bound to CENP-C^{CD} and CENP-N^{NT} (the CCNC) by preparative native PAGE ("Prep

Cell"). Fractions are collected as they are eluted out of the bottom of the gel. The native PAGE is stained both by ethidium bromide (For DNA) and Coomassie Blue (for protein). **(d)** Sucrose gradient elution profiles of CENP-A nucleosomes, CENP-A nucleosomes in complex with CENP-C^{CD} and CENP-N^{NT}, and the CENP-C^{CD} and CENP-N^{NT} proteins alone. Complexes were subject to a linear 5-30 % sucrose gradient with centrifugation at 35,000 rpm for 13 hr at 4°C. The samples were fractionated from top to bottom, and each fraction was analyzed for absorbance at 280 nm (for nucleosome and complex) or 260 nm (for CENP-C and CENP-N proteins alone). The absorbance values are plotted, with the highest value in each run normalized to 1.0. **(e)** SDS-PAGE gels stained with Coomassie-Blue, showing input and peak fractions of CENP-A nucleosomes bound to CENP-C^{CD} and CENP-N^{NT} (the CCNC). **(f-i)** Representative peptides within the CCNC. (f) A histone peptide that shows no difference between CENP-A nucleosomes vs. the CCNC. (g-i) Histone peptides spanning the CENP-A surface bulge (g), and the β-sheet region between histone H4 and H2A (h,i) over the timecourse. The maximum number of deuterons possible to measure by HXMS for each peptide is shown by the dotted line. All peptides are plotted at every time point as mean ± SD from triplicate experiments. Note that for some data points, the error bars are too small to be visible in the graph.

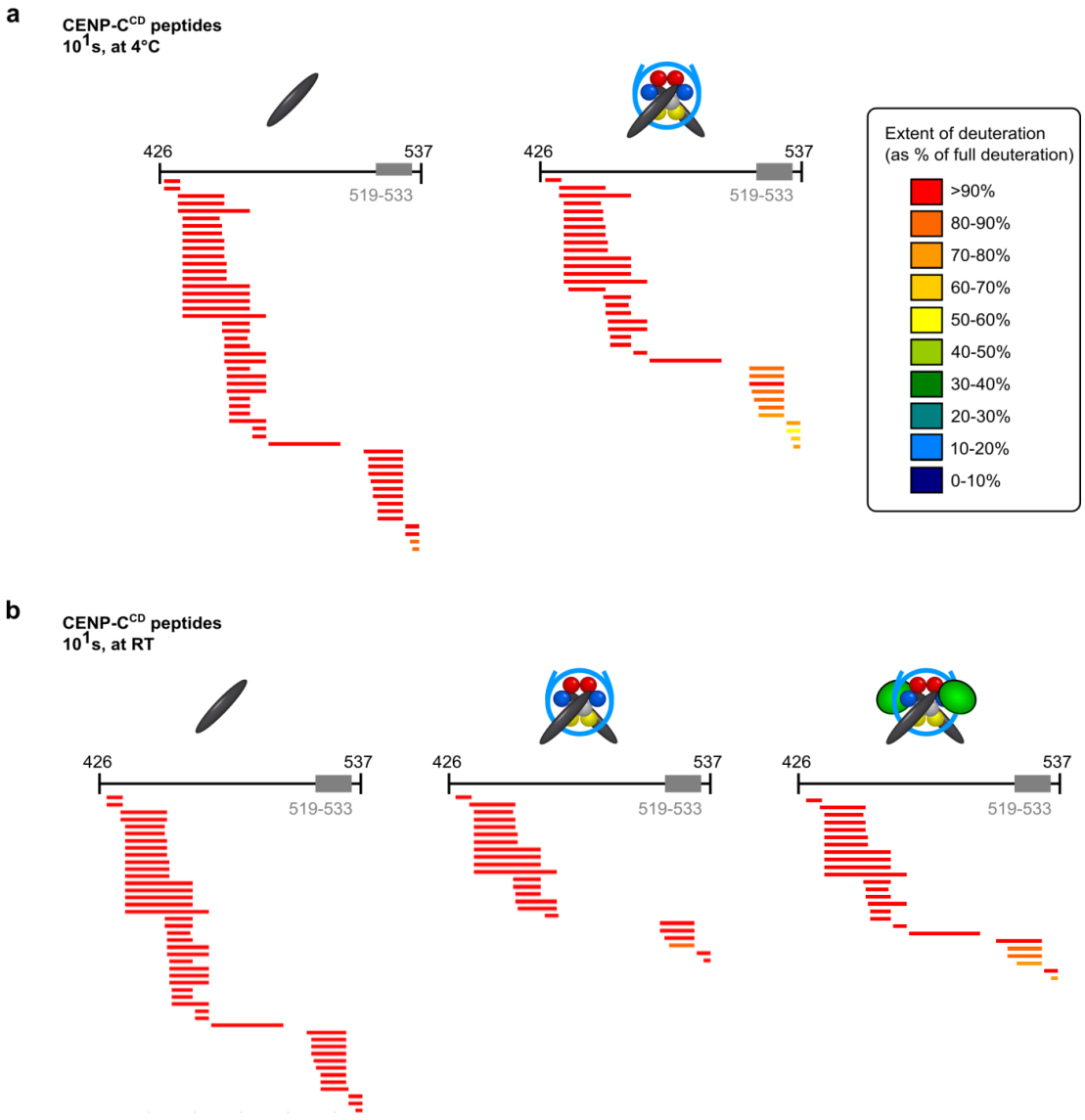


Figure 33. CENP-C^{CD} lacks detectable secondary structure and binds the histone surface of CENP-A nucleosomes with residues ~515-537.
(a) CENP-C^{CD} peptides alone vs. when in complex with CENP-A nucleosomes, at the 10^1 s, 4°C timepoint (which is the earliest timepoint we can test, and is equivalent to 10^0 s at room temperature [RT]). This shows that CENP-C^{CD}, when alone, is disordered and is mostly completely exchanged even at this earliest timepoint. CENP-C^{CD} when in complex with CENP-A nucleosomes shows some mild protection in the a.a. 513-537 region, which encompasses the region that contacts CENP-A nucleosomes (~a.a. 519-533). Each horizontal bar represents an individual peptide from CENP-C^{CD} alone or in complex with CENP-A nucleosomes and is color-coded for percent deuteration. **(b)** CENP-C^{CD} peptides alone versus when in complex with CENP-A nucleosomes versus when part of the CCNC, at the 10^1 s (RT) timepoint. Again, some mild protection is seen in the a.a. 513-537 region when CENP-C^{CD} is bound to CENP-A nucleosomes or part of the CCNC, but CENP-C^{CD} still overall lacks

secondary structure in three instances. Thus, because protection is so minor at such an early timepoint (10 s), the presence of CENP-N does not cause the gain of secondary structure or any other change that would otherwise lead to strong protection from HX.

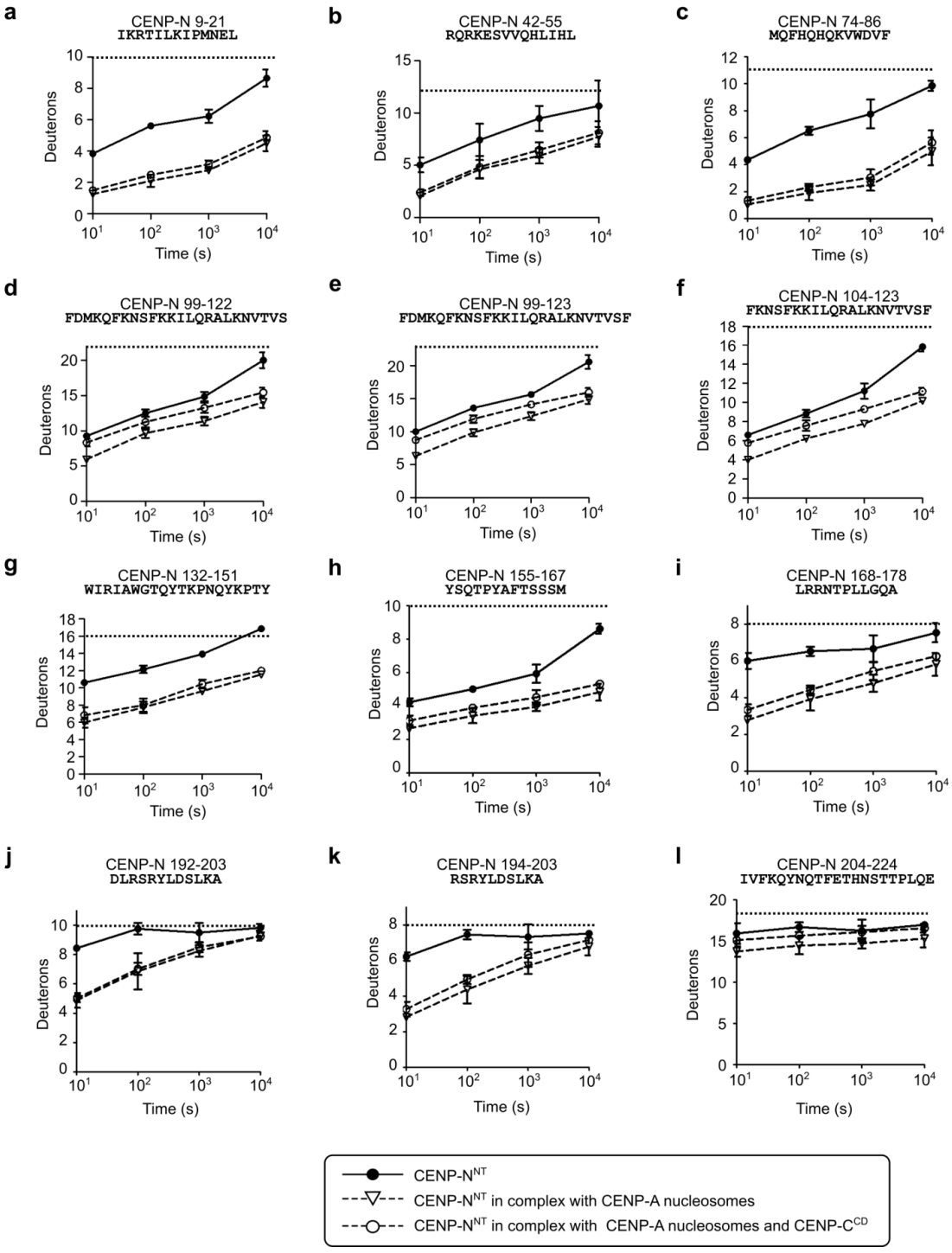


Figure 34. The HXMS behavior of CENP-N^{NT} in the presence and absence of CENP-C^{CD}.

(a-l) Representative peptides of CENP-N^{NT} across the timecourse, either unbound or bound to CENP-A NCPs (in the presence or absence of CENP-C^{CD}). The maximum number of deuterons possible to measure by HXMS for each peptide is shown by the dotted line. All peptides are plotted at every time point as mean ± SD from triplicate experiments. Note that for some data points, the error bars are too small to be visible in the graph. Across *most* of the folded nucleosome-binding domain of CENP-N^{NT}, the massive slowing of HX is similar in the presence or absence of CENP-C^{CD} (a-c, g-k). Within a region that maps unambiguously to a.a. 99-122, however, CENP-N only exhibits substantial HX protection when bound alone to the CENP-A NCP (d-f). This suggests that a very local region (i.e. between a.a. 99-122) has a structural change in the presence of CENP-C that leads to the change in HX behavior we observe. The region C-terminal of a.a. 206 is disordered regardless of whether CENP-N^{NT} is bound to the NCP (l, also see **Figure 30**).

3.3.5 CENP-A NUCLEOSOME STABILITY REQUIRES BOTH CENP-C AND CENP-N

The finding that a stable CCNC can be assembled from its component parts (**Figure 31**) supports the notion that CENP-N provides stability to centromeric chromatin that cannot be attributed to CENP-C^{CD}. To test this notion, we focused our analysis back on our cell lines where the levels of the two components can be modulated. CENP-C and CENP-N display partially interdependent localization to centromeres in human cells (McKinley et al., 2015), complicating the analysis of their interactions with CENP-A. During the 24 h timescale in which we measure CENP-A retention at centromeres, CENP-C removal also leads to loss of most but not all CENP-L-N (**Figure 26b**, and **Figure 28b**). CENP-N removal, using a similar AID-tagging approach of both CENP-N alleles(McKinley et al., 2015), reduces CENP-C levels at centromeres by ~half (**Figure 36a-c**). Interestingly, CENP-N-AID removal produces a pronounced defect in CENP-A retention (**Figure 35a-c**), suggesting a direct role of CENP-N in CENP-A retention, since reducing CENP-C levels by half is unlikely to be responsible for the full magnitude of this effect. Importantly, combined removal of CENP-N (by siRNA treatment) and CENP-C (by IAA treatment) had an additive effect, severely compromising CENP-A retention (**Figure 35d-f**, and **Figure 36d,e**). This indicates that the low level of CENP-L-N remaining after 24 h of CENP-C depletion (**Figure 26b**, and **Figure 28b**) accounts for the residual stability

to CENP-A nucleosomes not accounted for by the CENP-C^{CD} alone (**Figure 24d,e**). More importantly, this supports a model wherein CENP-C and CENP-N are roughly equal partners necessary to form the CCNC and maintain CENP-A nucleosome levels at centromeres (**Figure 37**).

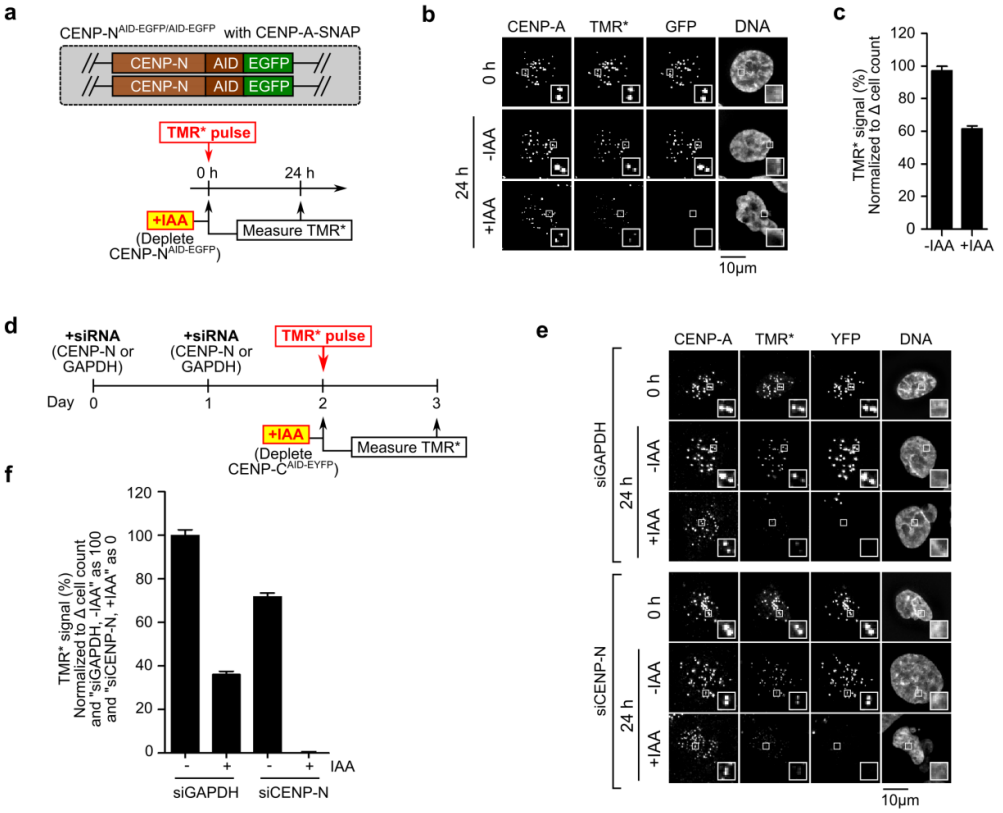


Figure 35. CENP-C and CENP-N collaborate to maintain CENP-A nucleosomes at centromeres.
(a) Schematic for experiment in which CENP-N$^{AID-EGFP/AID-EGFP}$ cells expressing CENP-A-SNAP at a unique FRT site were pulse-labeled with TMR* and assessed for retention of the existing pool of CENP-A molecules. **(b)** Representative images from experiment diagrammed in a. Scale bar indicates 10 μm. **(c)** Quantitation of b. **(d)** Schematic of experiment, in which CENP-C$^{AID-EYFP/AID-EYFP}$ cells were treated with siCENP-N or siGAPDH and pulse-labeled with TMR*, and the relative CENP-A-SNAP signals were analyzed after 24 h (with or without CENP-C depletion by IAA treatment). **(e)** Representative images from experiment described in d. Scale bar indicates 10 μm. **(f)** Quantitation of e. The value of the "siGAPDH, -IAA" condition is normalized as 100%, and the value of the "siCENP-N, +IAA" condition is normalized as 0%. All graphs are shown as mean ± 95% confidence interval (n>2000 centromeres in all cases).

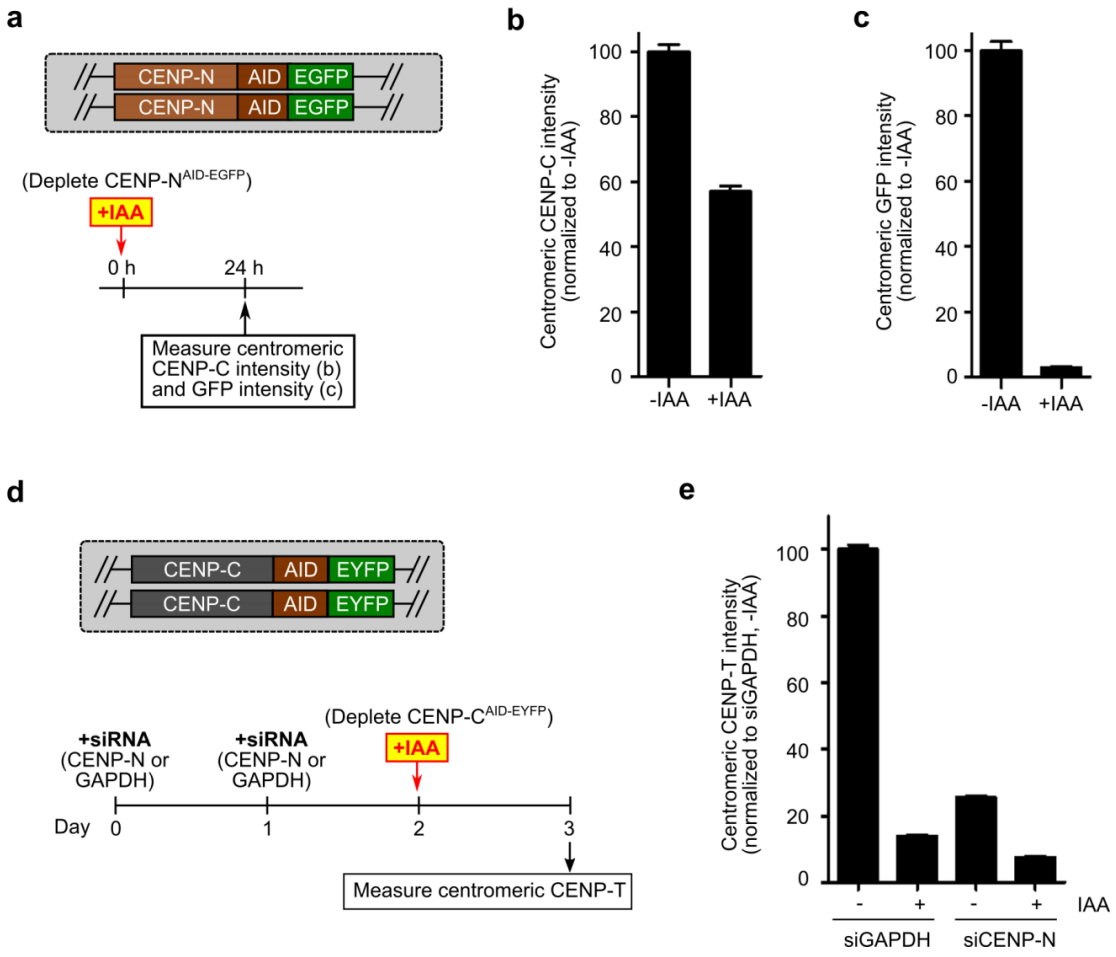


Figure 36. Effect of CENP-N depletion on centromeric CENP-C and CENP-T levels. **(a)** AID-tagged CENP-N was depleted by 24 h of IAA in CENP-N$^{AID\text{-}EGFP/AID\text{-}EGFP}$ cells, then assessed for centromeric CENP-C localization, and disappearance of CENP-N-AID-EGFP. **(b)** Quantitation of centromeric CENP-C intensity before and after 24 h IAA treatment. Mean ± 95% confidence interval (n>2000 centromeres in all cases). **(c)** Quantitation of centromeric CENP-N-AID-EGFP intensity before and after 24 h IAA treatment. Mean ± 95% confidence interval (n>2000 centromeres in all cases). **(d)** CENP-C$^{AID\text{-}EYFP/AID\text{-}EYFP}$ cells were subject to CENP-N depletion via siRNA, as in **Figure 35d**, and assessed for centromeric CENP-T localization. **(e)** Quantitation of centromeric CENP-T intensity. As expected, CENP-T levels are markedly reduced in the presence of CENP-N depletion (Logsdon et al., 2015; McKinley et al., 2015; Samejima et al., 2015), which indicates the expected impact of substantial CENP-N depletion by the siRNA approach used here and in the experiment in **Figure 35d-f.** Mean ± 95% confidence interval (n>2000 centromeres in all cases).

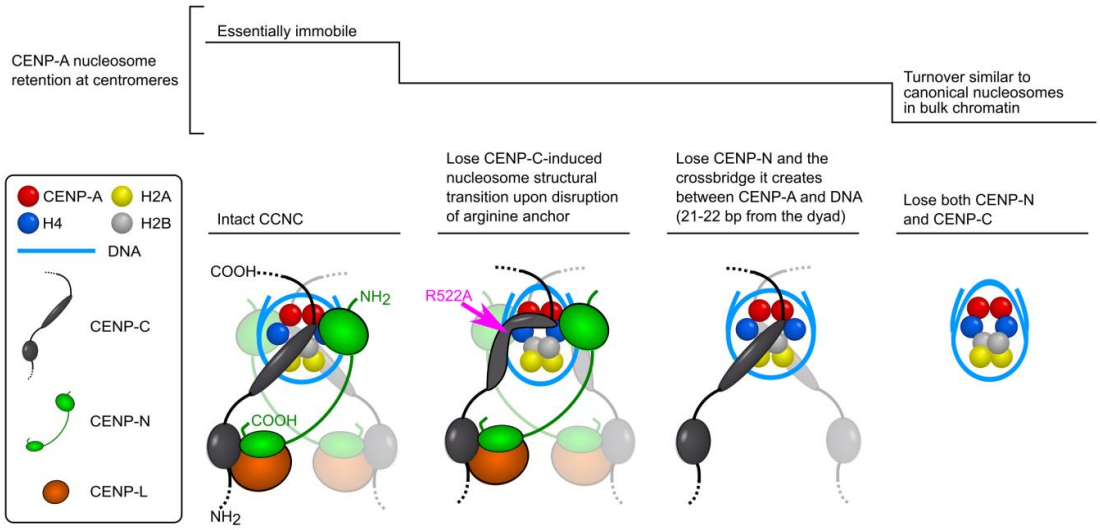


Figure 37. Model of the physical basis for the stability of CENP-A nucleosomes within the CCNC.
See text for the details of our model of centromere maintenance. We note that CENP-C^{CD} is shown as an elongated oval that represents a structured loop that has no conventional secondary structure, despite having been historically called a "domain". Also, a flexible linker is shown between CENP-C^{CD} and the CENP-C contact point with the CENP-L-N complex, in line with proposals that CENP-C largely exists as an extended and unfolded protein that may span >100 nm at mitotic kinetochores (Screpanti et al., 2011; Wan et al., 2009; Ye et al., 2016). Additionally, it is also not known if there is a fixed or variable distance at centromeres from the CENP-C-L-N contact point to the CENP-A nucleosome. It is also unclear if this contact point on CENP-C with CENP-L-N is a folded domain or if it contacts one or both subunits of CENP-L-N.

3.4. Discussion

Our physical studies of CENP-A nucleosome complexes combined with gene replacement and rapid depletion of the non-histone CCAN proteins, CENP-C and CENP-N, provide the molecular basis for the extraordinary stability of CENP-A nucleosomes that is at the heart of the epigenetic mechanism that maintains the identity of centromere location on every chromosome. At steady-state, we envision that the relevant nucleosome required to maintain centromere identity has ~147 bp of DNA wrapped around an octameric histone core containing two copies each of CENP-A, H4, H2A, H2B, and two copies each of CENP-C and CENP-N. CENP-C binding

confers the most pronounced physical changes in the CENP-A nucleosome structural transition that alters nucleosome shape, enhances the tendency of CENP-A nucleosomes to sample states with 20 bp of DNA unwrapping at each nucleosome terminus, and confers internal and surface rigidity to the histone core(Falk et al., 2015, 2016). CENP-N, while having a more discrete impact on the histone core of the NCP (**Figure 26**), has a binding site that crossbridges a key DNA contact point on the NCP to the (CENP-A/H4)$_2$ heterotetramer (**Figure 26g-j**). We expect that normally these are the direct chromatin contacts at the interface with the kinetochore, including those recently reconstituted with purified components (Weir et al., 2016). Though, since contacts between CENP-A and the other CCNC components can be bypassed in mitosis (Hoffmann et al., 2016), CCNC function may be more relevant to maintaining centromere identity during the remainder of the cell cycle.

Independent recognition of a single nucleosome by two different chromatin components, as we find occurs within the CCNC, has not been well studied in any chromatin context. The small but growing list of physical studies of nucleosome recognition proteins (Armache et al., 2011; Barbera et al., 2006; Makde et al., 2010; McGinty et al., 2014; Morgan et al., 2016) uniformly involves a key contact point between an arginine anchor with the nucleosomal acidic patch (McGinty and Tan, 2016). Three previous studies had claimed that mutation of R522 (or its corresponding position in Xenopus CENP-C) disrupts centromere targeting of CENP-C, but none provided a definitive answer for mammalian CENP-C: one study was done in *Xenopus* extracts (Carroll et al., 2010), and the two studies in human cells used truncated CENP-C transgenes that were overexpressed (Kato et al., 2013; Song et al., 2002). Our study advances the field in part because it interrogates the nucleosome-binding domains of CENP-C in a gene replacement system, using one in which the endogenous CENP-C is rapidly and completely removed, and the replacement CENP-C constructs are untagged, full-length, and expressed at near endogenous levels (**Figure 21c**). More broadly, our findings show a remarkable role

for an arginine anchor beyond their established role in nucleosome recognition (McGinty and Tan, 2016) to a role in altering nucleosome shape and function. R522A preserves the ability of CENP-C to *bind* to CENP-A nucleosomes (**Figure 24a-c**) and accumulate at centromeres (**Figure 26b,c**), but we pinpoint a role for R522 for CENP-A maintenance at the centromere (**Figure 26d,e**), driving the nucleosome structural transition that stabilizes the interior of the CENP-A nucleosome (**Figure 24d,e**). It is possible that mutation of another residue within CENP-C^{CD} could also retain binding while compromising the CENP-A nucleosome structural transition, but disruption of the R522 arginine anchor in our gene replacement systems indicates that this common feature in diverse nucleosome binding proteins can play an important functional role, beyond the role of molecular recognition.

CENP-C^{CD} is particularly remarkable because its high specificity for CENP-A nucleosomes is mediated by a very small feature (the 6 a.a. C-terminal tail)(Carroll et al., 2009), but its binding confers stabilization that spreads throughout much of the octameric core of the nucleosome as well as to the position of the DNA gyres (Falk et al., 2015, 2016). Remarkably, CENP-C^{CD} does this without having any defined secondary structure of its own.

Most aspects of the mechanism used by CENP-N^{NT} contrast starkly with that used by CENP-C^{CD}. The only notable similarity is that CENP-N^{NT} uses a small feature on the surface of the CENP-A nucleosome to achieve its high specificity of binding to CENP-A nucleosomes. In contrast to the widespread HX protection conferred to the NCP by CENP-C^{CD}, the only HX protection we observed with CENP-N^{NT} maps to loop L1 (**Figure 26d-f**). Our findings provide clear biophysical evidence that CENP-N^{NT} recognizes NCPs without accessing the acidic patch on H2A-H2B at all, making it unique relative to other nucleosome recognition domains studied to date (Armache et al., 2011; Barbera et al., 2006; Makde et al., 2010; McGinty et al., 2014; Morgan et al., 2016) and leaving open that site on the NCP for CENP-C^{CD} to bind. CENP-N^{NT} itself is a folded domain (again, in contrast to CENP-C^{CD}) even prior to engaging the CENP-A NCP (**Figure 29**, **Figure 30**, **Figure 34**). Its discrete contact points on Loop

1 of CENP-A and the adjacent nucleosomal DNA 21-22 bp from the dyad axis of symmetry stabilizes its own secondary structure (**Figure 26**, **Figure 30**, **Figure 34**). Therefore, the combination of previous work (Carroll et al., 2009, 2010, Falk et al., 2015, 2016; Fang et al., 2015; Kato et al., 2013) and the work presented here shows that CENP-C^{CD} and CENP-N^{NT} defy expectations in that the *unfolded* one (CENP-C^{CD}) generates substantial structural changes in the NCP (**Figure 31e-f** and (Falk et al., 2015; Kato et al., 2013)), whereas the *folded* one (CENP-N^{NT}) changes core histone dynamics only very locally at the points of contact with CENP-A and its adjacent nucleosomal DNA (**Figure 26d-j**). Furthermore, our combined HXMS and hydroxyl radical footprinting shows that CENP-N fastens CENP-A to its adjacent DNA, providing an example of a crossbridging mechanism for maintaining nucleosome-encoded epigenetic information that perhaps represent a more general mode for maintaining nucleosome-encoded epigenetic information involving other histone variants (or post-translationally modified canonical histones).

A steady-state CCNC complex required to faithfully maintain CENP-A retention at centromeres does not necessitate that all components exhibit matched turnover rates, themselves. H2A-H2B dimers can come on and off through partial disassembly of nucleosomes, as with canonical nucleosomes. CENP-C and CENP-N could similarly exchange, and indeed both proteins display dynamic behaviors at centromeres (Hellwig et al., 2011; Hemmerich et al., 2008; Smoak et al., 2016), with CENP-N varying in quantity at the centromere depending on the cell cycle stage (Fang et al., 2015; Hellwig et al., 2011; McKinley et al., 2015). We note that H2A-H2B are not nearly as stable at centromeres as CENP-A and H4 (Bodor et al., 2013; Falk et al., 2015). The binding mode for CENP-N^{NT} suggests an explanation for this: CENP-C^{CD} binding protects all of the core histones from dissociating from DNA, but CENP-N^{NT} would only protect CENP-A and H4. The hydrophobic stitches between CENP-A and H4 themselves provide yet another required feature to rigidify the particle and maintain it at centromeres (Bassett et al., 2012; Black et al., 2004; Sekulic et al., 2010).

In conclusion, our findings demonstrate that the individual CENP-C^{CD} and CENP-N^{NT} subunits independently bind to the nucleosome with non-overlapping effects on the stability and/or shape of the nucleosome (**Figure 37**). When both proteins are present, they impart additive effects on the physical properties of CENP-A nucleosomes. At any given time, there are multiple molecules of CENP-C and CENP-N present at the centromere—both directly bound to the CENP-A nucleosomes *and* directly bound to each other—in a manner that locks in centromere location. By tying faithful inheritance of chromosomes to an epigenetic mark in which the CCNC acts as the fundamental repeating unit, mammals have evolved a remarkably resilient form of chromatin.

3.5. METHODS

3.5.1. GENERATION OF CELL LINES

Using DLD-1 Flp-In T-Rex cells stably expressing Tir1(Holland et al., 2012) with CENP-$C^{AID\text{-}EYFP/AID\text{-}EYFP}$ (Hoffmann et al., 2016) as a starting point, endogenous CENP-A was tagged with C-terminal SNAP using CRISPR/Cas9-mediated genome engineering. The sgRNA was designed to target the 3'UTR of CENP-A. The oligonucleotides (5'-CACCGCTGACAGAAACACTGGGTGC-3' and 5'-AAACGCACCCAGTGTTTCTGTCAGC-3') were annealed and inserted into pX330 which already contains Cas9 (Ran et al., 2013). To generate the repair template, the SNAP-3xHA sequence (Jansen et al., 2007) followed by a viral 2A peptide (Kim et al., 2011) and the neomycin resistance gene was synthesized as a gBlock (Integrated DNA Technologies), and 5' and 3' homology arms of ~800 bp each were amplified from DLD-1 genomic DNA by PCR. All three pieces were inserted into a pUC19 backbone using HiFi DNA Assembly (NEB). The repair template and pX330 were co-transfected with Lipofectamine 2000 (Invitrogen) in 9:1 ratio, and selected after 5 days using 750 µg/ml G418. To isolate monoclonal cell lines, cells were subject to limiting dilution after G418 selection. To screen clones, PCR of genomic DNA was performed for every clone (using primers 5'-CCTTCCCCACTCCTTCACAGGC-3' and

5'-CCTGTGAAAGAGGATGAGCTTACC-3'); insertion of the SNAP tag results in a PCR product of 2243 bp (whereas the PCR product is 614 bp if the allele is unmodified). Clones containing SNAP-tagged CENP-A were further validated by immunoblotting and TMR* visualization. Stable cell lines constitutively expressing CENP-C rescue constructs were generated by Flp/FRT recombination. Domain deletions and point mutants of CENP-C were generated by PCR site-directed mutagenesis, and the sequences of all constructs (WT and mutant versions) were validated by DNA sequencing. CENP-C constructs were inserted into a pcDNA5/FRT vector and co-transfected with pOG44 (Invitrogen), a plasmid expressing the Flp recombinase, into cells with Lipofectamine 2000 (Invitrogen) according to manufacturer's instructions. Following selection in 400 μg/ml Hygromycin B, colonies were pooled into polyclonal cell lines. CENP-N$^{AID\text{-}EGFP/AID\text{-}EGFP}$ cells expressing CENP-A-SNAP were also generated by Flp/FRT recombination: CENP-A-SNAP was inserted into a pcDNA5/FRT vector and co-transfected with pOG44 into CENP-N$^{AID\text{-}EGFP/AID\text{-}EGFP}$ cells (McKinley et al., 2015) and selected with Hygromycin B as described above.

3.5.2. CELL CULTURE

The indicated DLD-1 derivatives described above were cultured in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% fetal bovine serum (FBS), 100 U ml^{-1} penicillin and 100 μg ml^{-1} streptomycin. All cell lines were maintained with 2 μg ml^{-1} puromycin (Sigma). Cell lines in which CENP-A is SNAP-tagged by CRISPR/Cas9-mediated genome editing were maintained with 750 μg ml^{-1} G418. Cell lines containing CENP-C rescue constructs introduced by Flp/FRT recombination were maintained with 400 μg ml^{-1} Hygromycin B. CENP-N$^{AID\text{-}EGFP/AID\text{-}EGFP}$ cells with CENP-A-SNAP at the FRT site were maintained in 300 μg ml^{-1} G418 and 400 μg ml^{-1} Hygromycin B. To induce degradation of AID-tagged CENP-C or CENP-N, indole-3-acetic acid (IAA; Sigma) was prepared in water and added to cells at 500 μM for the indicated amounts of time.

3.5.3. IMMUNOBLOTTING

Samples derived from whole cell lysates were separated by SDS-PAGE and transferred to a nitrocellulose membrane for immunoblotting. Blots were probed using the following primary antibodies: rabbit anti-CENP-C (1.7 μg ml^{-1})(Bassett et al., 2010), mouse mAb anti-α-tubulin (1:4000, Sigma-Aldrich #T9026), or human Anti-Centromere Antibodies (ACA) (2 μg ml^{-1}, Antibodies Incorporated #15-235). The blots were subsequently probed using the following HRP-conjugated secondary antibodies: Donkey Anti-Human IgG (1:10,000, Jackson ImmunoResearch Laboratories #709-035-149), Amersham ECL Mouse IgG (1:2,000, GE Life Sciences #NA931), Amersham ECL Rabbit IgG (1:2,000, GE Life Sciences #NA934V). Antibodies were detected by enhanced chemiluminescence (Thermo Scientific).

3.5.4. SNAP LABELING EXPERIMENTS

DLD-1 cells were pulse-labeled with 2 μM TMR* (NEB) in complete medium for 15 min at 37°C, washed with PBS and incubated in the culture medium for 2 h to allow excess TMR* to diffuse out of cells. Cells were then either fixed immediately (for the "0 h" timepoint), or cultured for another 24 h in the presence or absence of 500 μM IAA to induce degradation of the AID-tagged CCAN protein (for the "24 h" timepoints). Cell number was also determined at these timepoints using a hemocytometer, so that the total level of CENP-A turnover could be calculated, as described(Bodor et al., 2013; Falk et al., 2015): CENP-A turnover was calculated as [(TMR* intensity at 24 h)/(Avg TMR* intensity at 0 h)]*(Change in cell number). SiRNA knockdown of CENP-N was performed as described(Logsdon et al., 2015). Briefly, cells were treated with 20 μM CENP-N siRNAs (siGENOME SMARTpool; Dharmacon, GE Life Sciences #M-015872-02-0005), or GAPDH siRNAs (ON-TARGETplus GAPD Control; Dharmacon, GE Life Sciences #D-001830-01-05).

3.5.5. IMMUNOFLUORESCENCE AND MICROSCOPY

For experiments involving CENP-A, CENP-C, or CENP-T immunofluorescence, DLD-1 cells were fixed in 4% formaldehyde for 10 min at room temperature and quenched with 100 mM Tris (pH 7.5) for 5 min, followed by permeabilization using PBS containing 0.1% Triton X-100. For experiments involving CENP-L immunofluorescence, DLD-1 cells were pre-extracted with PBS containing 0.1% Triton X-100 for 30 s, fixed with 4% formaldehyde for 10 min and quenched with 100 mM Tris (pH 7.5) for 5 min. All coverslips were then blocked in PBS supplemented with 2% FBS, 2% BSA, and 0.1% Tween prior to antibody incubations. The following primary antibodies were used: mouse mAb anti-CENP-A (1:1000, Enzo Life Sciences #ADI-KAM-CC006-E), rabbit pAb anti-CENP-C (1.7 μg ml^{-1})(Bassett et al., 2010), rabbit pAb anti-CENP-T (1 μg ml^{-1})(Gascoigne et al., 2011b), and rabbit pAb anti-CENP-L (1:1000)(McKinley et al., 2015). Secondary antibodies conjugated to fluorophores were used: Cy3 Goat anti-Rabbit (1:200, Jackson ImmunoResearch Laboratories #111-165-144) and Cy5 Donkey anti-Mouse (1:200, Jackson ImmunoResearch Laboratories #715-175-151). Samples were stained with DAPI before mounting with VectaShield medium (Vector Laboratories). Images were captured at room temperature on an inverted fluorescence microscope (DMI6000 B; Leica) equipped with a charge-coupled device camera (ORCA AG; Hamamatsu Photonics) and a 40x oil immersion objective. Images were collected at 0.59 μm z-sections and subsequently deconvolved using identical parameters. To display as figures, the z stacks were projected as single two-dimensional images and assembled using ImageJ (NIH). To quantify fluorescence intensity of centromeres, the CraQ macro(Bodor et al., 2012) was run in ImageJ with standard settings using DAPI and total CENP-A staining as the reference channel to define ROIs for quantification of TMR* intensity. One representative experiment is displayed from 2 or more independent experiments. At least 2000 centromeres were analyzed for each timepoint.

3.5.6. RECOMBINANT PROTEIN PURIFICATION

Human histones and CENP-A were prepared as described (Sekulic and Black, 2016a; Sekulic et al., 2010). Briefly, histones H2A and H2B are expressed as monomers in inclusion bodies and purified under denaturing conditions, then refolded into H2A-H2B dimers. (CENP-A/H4)2 is expressed off of a bicistronic construct as a soluble heterotetramer and purified by hydroxyapatite column followed by cation exchange. Recombinant human CENP-C^{CD} consisting of the central domain (a.a. 426-537) was expressed from a plasmid kindly provided by A. Straight (Stanford)(Carroll et al., 2010; Falk et al., 2015). CENP-C is expressed as a GST fusion protein and affinity-purified on a glutathione column. GST is then cleaved by PreScission protease and separated from CENP-C by cation exchange(Carroll et al., 2010; Falk et al., 2015; Sekulic and Black, 2016a). sPCR site-directed mutagenesis was performed to generate CENP-C^{CD}(R521A) and CENP-C^{CD}(R522A), and they were expressed and purified using the same protocol as wildtype CENP-C^{CD}. Recombinant human CENP-N^{NT}-His was purified with a protocol adapted from a previous study(McKinley et al., 2015): CENP-N^{NT}-His was grown in BL21(DE3)pLysS cells for 6 h at 18°C, and purified on a 1 ml HisTrap FF column (GE Healthcare) via FPLC, with elution buffer of 50 mM sodium phosphate pH 8.0, 500 mM NaCl, 250 mM imidazole, 1 mM βME, and 50% glycerol. PCR-directed mutagenesis was performed to generate the further truncated construct, CENP-N^{1-205}-His, and it was purified with the same protocol as CENP-N^{NT}-His. Sequential purifications of complexes co-expressing GST- and His-tagged subunits were performed as described(McKinley et al., 2015). Briefly, complexes were first purified on Ni-agarose, and the elution was bound to glutathione agarose, washed three times, and eluted.

3.5.7. Assembly of NCPs and complexes

Six identical repeats of a 147 bp DNA sequence derived from an α-satellite sequence from the human X chromosome (Yang et al., 1982) was cloned into a pUC57 backbone, with each repeat separated by an EcoRV site. The sequence of each repeat is 5'-

ATCAAATATCCACCTGCAGATTCTACCAAAAGTGTATTTGGAAACTGCTCCATCAAAAGG
CATGTTCAGCTCTGTGAGTGAAACTCCATCATCACAAAGAATATTCTGAGAATGCTTCCG
TTTGCCTTTTATATGAACTTCCTCGAT-3'. This sequence corresponds to the major binding site that the CENP-A nucleosome occupies on human centromeres (Hasson et al., 2013). Preparation of DNA for NCP assembly was performed as described (Sekulic and Black, 2016a). Briefly, the plasmid described above was grown, isolated, and subjected to EcoRV digestion followed by separation of plasmid and insert by anion chromatography using Source 15Q resin (GE Healthcare). With the purified DNA, CENP-A NCPs were assembled and uniquely positioned using gradual salt dialysis followed by thermal shifting for 2 hr at 55°C (Dyer et al., 2004; Sekulic and Black, 2016b). Formation of complexes with CENP-C^{CD} was performed as described (Falk et al., 2015), in which 2.2 moles of CENP-C^{CD} were added per mole of CENP-A NCPs. To form the complex with CENP-N^{NT}, 4 moles of recombinant CENP-N^{NT}-His were added per mole of CENP-A NCPs. To form the complex with both CENP-N^{NT} and CENP-C^{CD}, 4 moles of CENP-N^{NT}-His and 2.2 moles of CENP-C^{CD} were added per mole of CENP-A NCPs. Complexes were analyzed by 5% native PAGE, stained with ethidium bromide to visualize DNA and Coomassie Brilliant Blue to visualize protein components. Following formation of complexes (or NCPs, in the case of the nucleosome-alone sample), samples were purified by preparative electrophoresis (Prep Cell, BioRad) using a 5% native gel to isolate the relevant complex from other species, such as free DNA (Dyer et al., 2004).

3.5.8. Binding assays (with Praveen Kumar Allu)

Recombinant human H2B K120C was purified as described for wildtype H2B (Falk et al., 2015) from inclusion bodies. Lyophilized protein was dissolved in unfolding buffer (7 M urea, 10 mM Tris-HCl pH 7.5 at 20°C, 0.4 mM TCEP) for 1 hr at RT and a 15-fold molar excess of sulfo-Cy5-maleimide (Lumiprobe) was dissolved in DMSO and added dropwise to the protein. The reaction proceeded overnight shielded from light and was quenched with 5 mM sodium 2-sulfanylethanesulfonate (MESNA) and run over a PD-10 column (GE Healthcare) to separate out free dye. Labeled H2B was

then mixed with equimolar amounts of H2A for dimer reconstitution and purification using established methods (Dyer et al., 2004; Sekulic et al., 2010), but using SDS-PAGE gels to determine concentrations of H2A and labeled-H2B monomers for refolding. Three independent assays were performed for calculating apparent K_d values for CENP-C WT and mutant proteins using CENP-A nucleosomes with labelled Cy5-H2B prepared on 147 bp DNA by gradient dialysis. Briefly, 200 nM of nucleosomes were incubated with increasing concentration of CENP-C WT or CENP-C mutants in TCS buffer (20 mM Tris-Cl pH 7.5, 1 mM EDTA, and 1 mM DTT) and incubated on ice for 1 hr before separating by 5 % native PAGE. After electrophoresis, gels were analyzed in a Typhoon 9200 imager (GE Healthcare), and the percentage of unbound vs. unbound nucleosomes were quantified using ImageJ. The apparent K_d values were calculated from the binding curve fitted from three independent experiments.

3.5.9. HXMS

Deuterium on-exchange was carried out by adding 5 µL of each sample (containing approximately 4 µg of NCPs or the indicated complex) to 15 µL of deuterium on-exchange buffer (10 mM Tris, pD 7.5, 0.5 mM EDTA, in D_2O) so that the final D_2O content was 75%. Reactions were quenched at the indicated timepoints by withdrawing 20 µL of the reaction volume, mixing in 30 µL ice cold quench buffer (2.5 M GdHCl, 0.8% formic acid, 10% glycerol), and rapidly freezing in liquid nitrogen prior to proteolysis and LC-MS steps. HX samples were individually melted at 0°C then injected (50 µl) and pumped through an immobilized pepsin (Sigma) column at initial flow rate of 50 µl min^{-1} for 2 min followed by 150 µl min^{-1} for another 2 min. Pepsin was immobilized by coupling to Poros 20 AL support (Applied Biosystems) and packed into column housings of 2 mm x 2 cm (64 µL) (Upchurch). Protease-generated fragments were collected onto a TARGA C8 5 µm Piccolo HPLC column (1.0 x 5.0 mm, Higgins Analytical) and eluted through an analytical C18 HPLC column (0.3 x 75 mm, Agilent) by a linear 12-55% buffer B gradient at 6 µl min^{-1} (Buffer A: 0.1% formic acid; Buffer B: 0.1% formic acid, 99.9% acetonitrile).

The effluent was electrosprayed into the mass spectrometer (LTQ Orbitrap XL, Thermo Fisher Scientific).

3.5.10. HXMS DATA ANALYSIS

The SEQUEST (Bioworks) software program was used to identify the likely sequence of parent peptides using non-deuterated samples via tandem MS. MATLAB-based MS data analysis tool, ExMS, was used for data processing (Kan et al., 2011). For all peptides found by SEQUEST, ExMS first analyses the non-deuterated sample to identify the peptide envelope centroid values as well as the chromatographic elution time ranges of each parental non-deuterated peptide. ExMS then uses the information from the non-deuterated analyses to identify deuterated peptides in each sample of the HXMS timecourse. Each individual deuterated peptide is corrected for loss of deuterium label during HXMS data collection (i.e., back exchange after quench) by normalizing to the maximal deuteration level of that peptide, which we measure in a "fully deuterated" (FD) reference sample. The FD sample are prepared in 75% deuterium just as is done in the on-exchange experiment, but under acidic denaturing conditions (0.5% formic acid), and incubated overnight so that each amide proton undergoes full exchange. The extent of back-exchange is calculated by comparing the extent of full deuteration as measured in the FD sample to the theoretical maximal deuteration (i.e., if no back-exchange occurs), which takes into account the 75% deuterium content of the samples. The median extent of back-exchange in our datasets is ~12% (**Figure 28e**), which is within the range for the lowest amount of deuterium loss ever reported for bottom-up HXMS (10% ± 5%)(Walters et al., 2012). For comparing two different HXMS datasets, we can plot the percent difference of each peptide, which is calculated by subtracting the percent deuteration of one sample from that of another, and plotted according to the color legend in 10% increments (as in **Figure 24d**, **Figure 26d**, **Figure 29a, Figure 31e**). We can also calculate the number of deuterons within each peptide that are exchanged at each timepoint, and plotted as in **Figure 26e-f**, **Figure 29b-e**, and **Figure 28f-j, Figure 30a-i, Figure**

32f-i, and **Figure 34**. These plots include data from 3 separate exchange reactions, with each data point shown as mean ± SD.

3.5.10. HYDROXYL RADICAL FOOTPRINTING (BY PRAVEEN KUMAR ALLU)

CENP-A nucleosomes assembled with HEX-labeled 147 bp α-satellite DNA (Falk et al., 2016) were reconstituted and then purified using a sucrose gradient. 4 µg of HEX-labeled CENP-A nucleosomes alone or complexed to CENP-N^{NT} were used in each reaction. The hydroxyl radical cleavage reaction was initiated by addition of 5 µl of 40 mM $FeAmSO_4$/80 mM EDTA, 2 M ascorbate, and 2.4% H_2O_2 to a 30 µl reaction mixture. Each reaction was carried out for 5min at room temperature, and terminated with 200 µl of stop solution (0.1% SDS, 25 mM EDTA, 1% glycerol, and 100 mM Tris, pH 7.4). Further phenol/chloroform extraction and ethanol precipitation was carried out to extract DNA fragments. Samples were separated by denaturing PAGE (10% polyacrylamide, 7 M urea, 88 mM Tris–borate, and 2 mM EDTA, pH 8.3)(Falk et al., 2016). Gels were imaged on a Typhoon 9200 imager (GE Healthcare). Band intensities were quantified from ImageJ from three independent experiments.

3.5.11. SUCROSE GRADIENT SEDIMENTATION (BY PRAVEEN KUMAR ALLU)

100 µg of CENP-A nucleosomes or the CCNC complex were subjected to 5-30% sucrose gradient centrifugation at 165,000g on a SW60 rotor (Beckman Coulter) for 13 h at 4°C. The samples were fractionated from top to bottom, and each fraction was analyzed for absorbance at 260 nm (for nucleosome and complex) or 280 nm (for CENP-C and CENP-N proteins alone).

CHAPTER 4: IDENTIFICATION OF CDK-DEPENDENT HJURP PHOSPHOSITES

Chapter 4 consists of my contributions to the following manuscript, and is the result of our collaboration with the Ben Garcia lab at the University of Pennsylvania (with postdoctoral researcher Xing-Jun Cao):

Stankovic, A., Guo, L.Y., Bodor, D.L., Mata, J.F., Cao, X., Bailey, A.O., Shabanowitz, J., Hunt, D.F., Garcia, B.A., Black, B.E., and Jansen, L.E.T. A dual inhibitory mechanism restricts centromeric chromatin assembly to G1 phase. *Molecular Cell,* 2017 Jan 19;65(2):231-246.

4.1. ABSTRACT

Maintenance of genomic stability requires accurate propagation of centromere location with every cell division. Replication of centromeric DNA in S phase must be followed by loading of new CENP-A molecules into centromeric chromatin, to prevent loss of CENP-A nucleosomes from successive dilutions. But unlike canonical histones, CENP-A loading is not coupled to DNA replication, but rather occurs exclusively after mitotic exit. This unique timing is controlled by cyclin-dependent kinases 1 and 2 (Cdk1/2), whose activity prevents CENP-A loading until the decline of Cdk1/2 activity in telophase/G1, but the targets of Cdk regulation are unknown. By combining the powerful methods of stable isotope labeling by amino acids in cell culture (SILAC), affinity-purification of the CENP-A-containing chromatin assembly complex with HJURP and its partner histone H4, titanium dioxide enrichment of phosphopeptides, as well as high-resolution mass spectrometry, we quantitatively measured the changes of multiple phosphorylation sites on CENP-A-containing complexes upon Cdk inhibition. We will present the results of these experiments, including findings that putative Cdk sites on HJURP are heavily phosphorylated in early mitosis, and are rapidly dephosphorylated upon Cdk inhibition. In contrast, putative Cdk sites on the CENP-A N-terminal tail remain phosphorylated even after Cdk inhibition. All of these post-translational

modifications are strong candidates to participate in the regulation of the CENP-A chromatin assembly pathway.

4.2. INTRODUCTION

Centromere location is specified epigenetically by the histone H3 variant termed centromere-protein A (CENP-A). Epigenetic inheritance of centromere location requires newly expressed CENP-A to be loaded into centromeric chromatin at each cell cycle. Contrary to canonical histones, CENP-A loading does not occur at the same time as DNA replication, but instead occurs following mitotic exit (Jansen et al., 2007). This unique loading mechanism is driven by cyclin-dependent kinases, whose activity during S, G2, and M phases sequesters CENP-A in a "prenucleosomal complex" where it is bound to its chaperone HJURP and its histone partner H4, and this inhibition is relieved as Cdk activity declines during mitotic exit (Silva et al., 2012). We sought to identify the Cdk targets that control this unique timing of CENP-A.

Stable isotope labeling with amino acids in cell culture (SILAC) is a robust tool of quantitative phosphoproteomics. SILAC entails in vivo metabolic incorporation of "heavy" 13C- or 15N-labeled amino acids into proteins during normal cell growth and division. Light and heavy peptides can be distinguished by predictable mass differences. Since there is no chemical difference (other than isotopic composition) between light and heavy amino acids, comparative cell populations exhibit identical biochemical and cellular properties (Mann, 2006).

4.3. RESULTS

4.3.1 IDENTIFICATION OF CDK-DEPENDENT PHOSPHORYLATION SITES ON THE CENP-A CHROMATIN ASSEMBLY COMPLEX BY QUANTITATIVE PHOSPHOPROTEOMICS

HJURP, the CENP-A specific chaperone is a phospho-protein and features several putative Cdk sites, making it a prime candidate for cell cycle control of CENP-A assembly. To precisely determine which residues are phosphorylated in the relevant cell cycle window, we used stable isotope labeling by amino acids in cell culture (SILAC) coupled to mass spectrometry. The SILAC approach allows for direct and accurate quantitation of changes in phosphorylation levels of residues. We grew two parallel cultures of cells stably expressing CENP-A-GFP (Bailey et al., 2013), one in normal media ("light") and one in media in which the essential amino acids lysine and arginine are replaced with their heavy isotopes ("heavy"). Both cultures were arrested in mitosis, one of which was subsequently released into early G1 phase by Roscovitine treatment (**Figure 38**, left, and

Figure 39). We harvested and combined equal numbers of the heavy/light cells after 30 min of Roscovitine (or DMSO only) treatment. At this stage HJURP is partially dephosphorylated (based on shift in SDS-PAGE mobility of the phosphorylated HJURP pool) (

Figure 39b,c), but has not completed centromere chromatin assembly (Jansen et al., 2007; Silva et al., 2012). We then isolated GFP-tagged CENP-A from the soluble fraction, which co-purified the relevant endogenous pool of HJURP. We detected 6 phosphorylated residues corresponding to putative Cdk consensus sites within HJURP, all of which were dephosphorylated to varying degrees upon mitotic exit, ranging from 25-70% decrease relative to mitotic values (**Figure 38b**, right,

Figure 39, and **Table 1**). In contrast, unphosphorylated peptides and previously reported phospho-sites on the CENP-A N-terminal tail (Bailey et al., 2013), which reside within Cdk consensus motifs, remain unaffected by forced mitotic exit, indicating that protein levels of CENP-A and HJURP remain unaffected and that HJURP is specifically targeted for dephosphorylation upon G1 entry. Three of the identified residues correspond to recently reported phospho-sites (Müller et al.,

2014; Wang et al., 2014), and our analysis shows that these (S412, S448, S473) are neither the sole nor the most responsive sites to the silencing of mitotic Cdk activity.

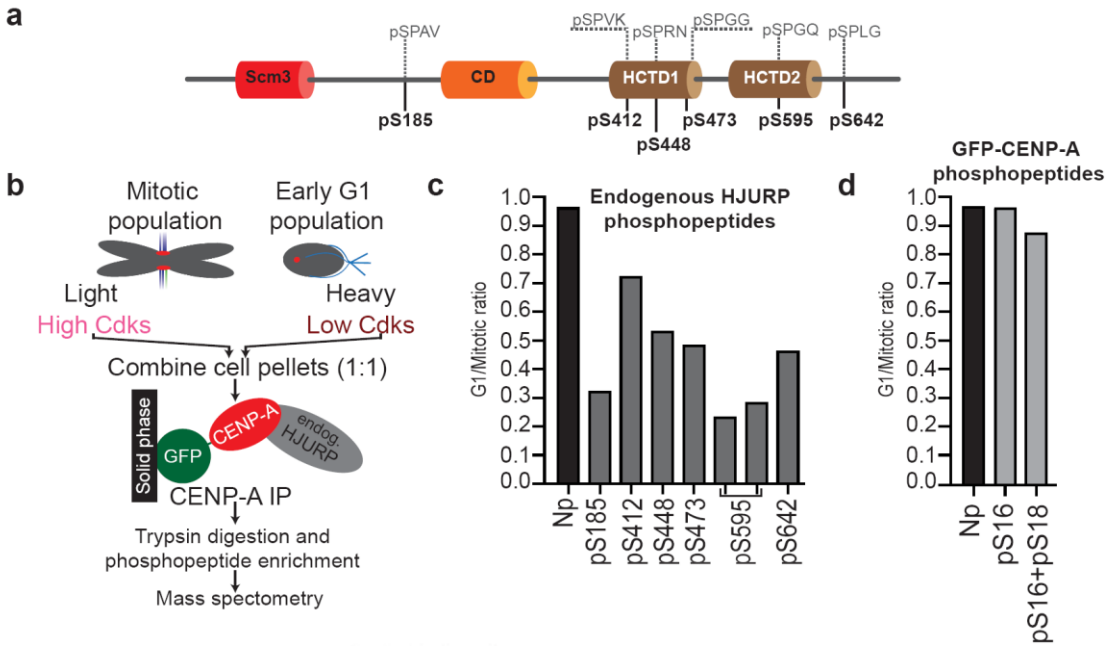
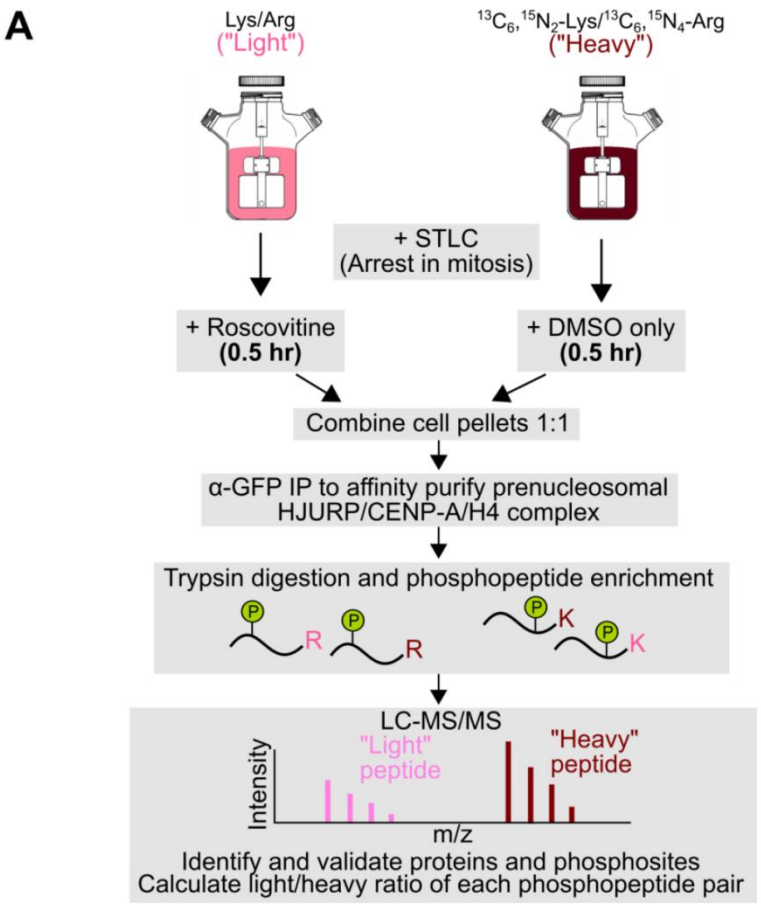


Figure 38. HJURP and is dephosphorylated upon mitotic exit.
(a) (Top) Schematic representation of HJURP protein, along with previously recognized domains (CENP-A binding domain (Scm3), conserved domain (CD), HJURP C-terminal domain 1 and 2 (HCTD1 and 2)) and the position of identified phospho-sites by SILAC in b. Amino acid sequences flanking phospho-sites are annotated. **(b)** Schematics of SILAC experiment. Cells stably expressing GFP-CENP-A were grown to equilibrium in either light or heavy medium and arrested in mitosis by STLC treatment. The light cells were released into G1 by Roscovitine treatment for 30 min. At this stage HJURP is partially dephosphorylated (based on shift in SDS-PAGE mobility of the phosphorylated HJURP pool) (

Figure 39b,c), but has not completed centromere chromatin assembly. Heavy and light cells were combined in equal numbers, followed by GFP-pull down of the prenucleosomal CENP-A complex containing the relevant endogenous HJURP pool, trypsin digestion, phosphopeptide enrichment and analysis by LC-MS/MS. (See more detailed schematic in

Figure 39a). **(c)** Phosphorylated Cdk sites are listed. The L/H ratio of a representative non-phosphorylated peptide is shown as internal control, **(d)** L/H ratios of Cdk consensus sites within the N-terminal tail of CENP-A. (See **Table 1** for data from two additional replicate experiments)


A
Lys/Arg
("Light")
$^{13}C_6$,$^{15}N_2$-Lys/$^{13}C_6$,$^{15}N_4$-Arg
("Heavy")
+ STLC
(Arrest in mitosis)
+ Roscovitine
(0.5 hr)
+ DMSO only
(0.5 hr)
Combine cell pellets 1:1
α-GFP IP to affinity purify prenucleosomal
HJURP/CENP-A/H4 complex
Trypsin digestion and phosphopeptide enrichment
LC-MS/MS
"Light"
peptide
"Heavy"
peptide
Intensity
m/z
Identify and validate proteins and phosphosites
Calculate light/heavy ratio of each phosphopeptide pair

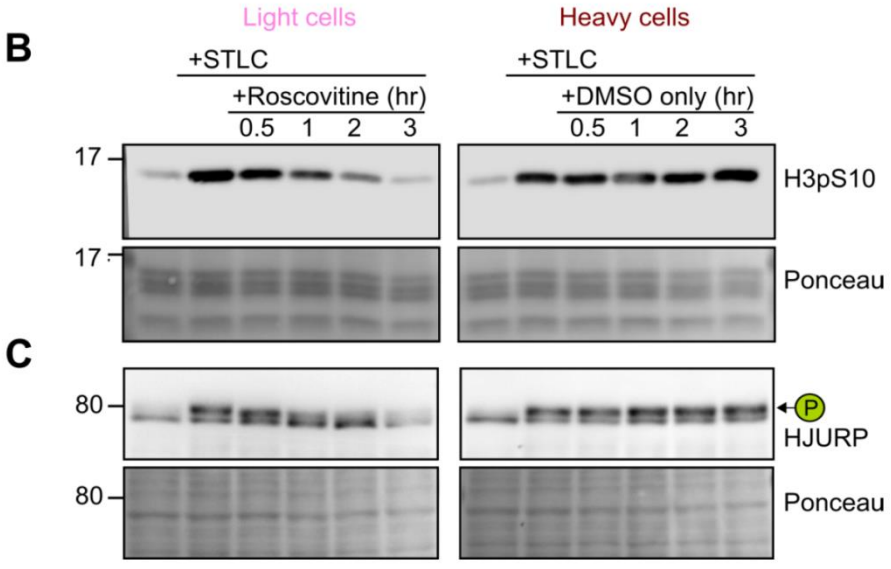

B
Light cells
Heavy cells
+STLC
+Roscovitine (hr)
+STLC
+DMSO only (hr)
0.5 1 2 3
0.5 1 2 3
17
H3pS10
17
Ponceau
C
80
HJURP
80
Ponceau

Figure 39. Strategy for SILAC experiment with Roscovitine-induced mitotic exit. **(a)** Schematic of SILAC experiment. **(b)** Western blots for the mitotic marker H3pS10 indicating cell cycle position of HeLa S3 cells. Cell were arrested in mitosis with the Eg5 inhibitor STLC followed by treatment with DMSO control or Roscovitine (light cells) to force mitotic exit caused by Cdk inhibition. (c) Western blots for HJURP (isolated from soluble fraction) showing dephosphorylation (as seen by shift in SDS-PAGE mobility of phosphorylated HJURP) upon Roscovitine treatment of "light" cells. Based on this, we harvested cells after 30 minutes of Roscovitine (or DMSO) treatment, balancing between HJURP dephosphorylation and completion of HJURP-mediated centromeric chromatin assembly.

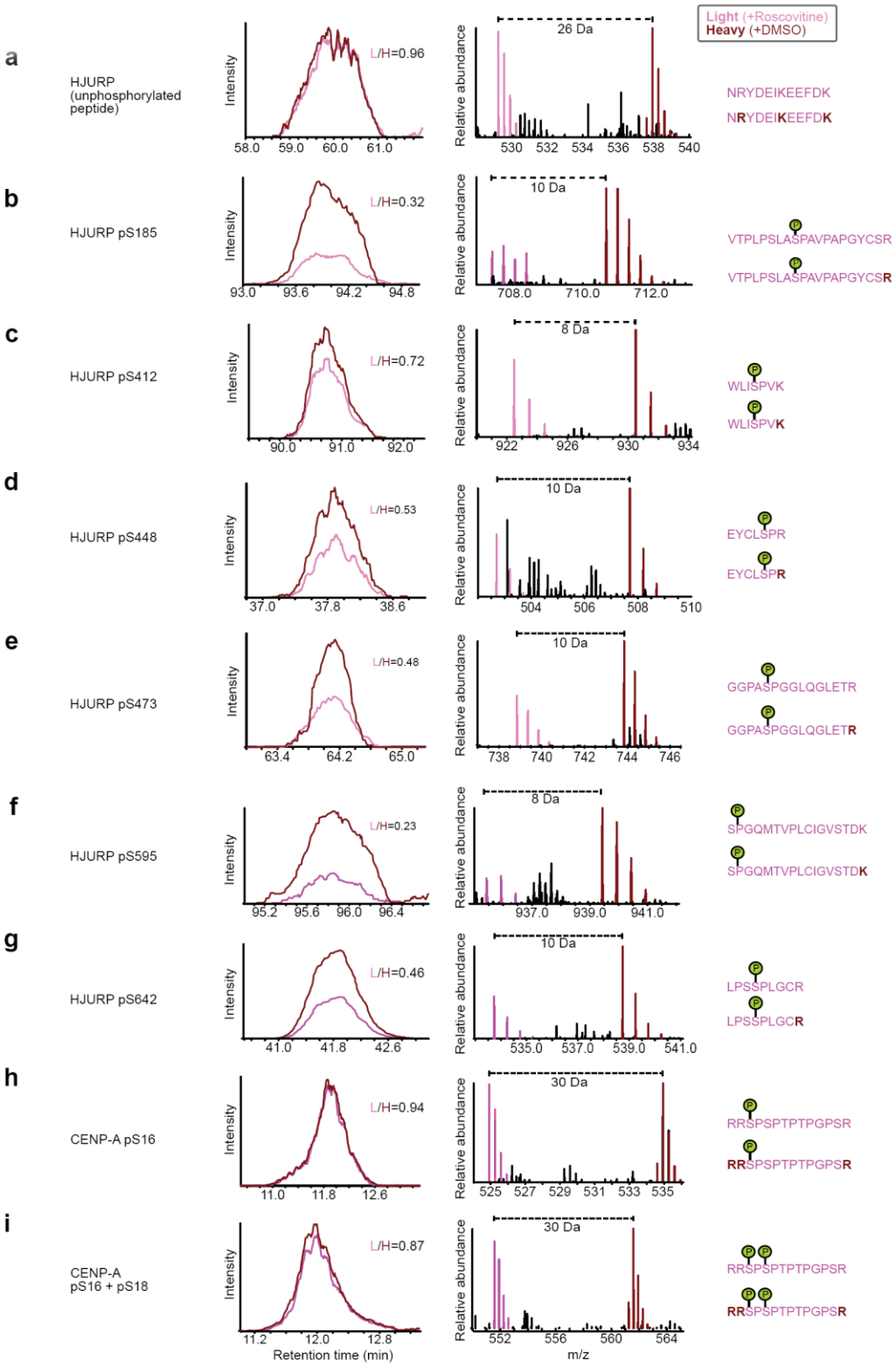

Light (+Roscovitine)
Heavy (+DMSO)
a
HJURP (unphosphorylated peptide)
L/H=0.96
26 Da
NRYDEIKEEFDK
NRYDEIKEEFDK
b
HJURP pS185
L/H=0.32
10 Da
VTPLPSLASPAVPAPGYCSR
VTPLPSLASPAVPAPGYCSR
c
HJURP pS412
L/H=0.72
8 Da
WLISPVK
WLISPVK
d
HJURP pS448
L/H=0.53
10 Da
EYCLSPR
EYCLSPR
e
HJURP pS473
L/H=0.48
10 Da
GGPASPGGLQGLETR
GGPASPGGLQGLETR
f
HJURP pS595
L/H=0.23
8 Da
SPGQMTVPLCIGVSTDK
SPGQMTVPLCIGVSTDK
g
HJURP pS642
L/H=0.46
10 Da
LPSSPLGCR
LPSSPLGCR
h
CENP-A pS16
L/H=0.94
30 Da
RRSPSPTPTPGPSR
RRSPSPTPTPGPSR
i
CENP-A pS16 + pS18
L/H=0.87
30 Da
RRSPSPTPTPGPSR
RRSPSPTPTPGPSR
Intensity
Relative abundance
Retention time (min)
m/z

Figure 40. Representative chromatograms and spectra of HJURP and CENP-A phosphopeptides in SILAC
Extracted ion chromatograms of all phosphopeptides from **Figure 38**, showing co-elution of phosphopeptide pairs (left panels), and representative mass spectra showing isotopic envelopes of light vs. heavy peptides (right panels). Phosphopeptides from "light" (Roscovitine-treated) cells are colored in pink, while the phosphopeptides from "heavy" (mock-treated) cells are colored in dark red. Each peptide pair is separated by a predictable mass difference (calculated from the number of lysines and arginines in the peptide), which is labeled with dotted lines between the monoisotopic peaks of light and heavy peptides. **(a)** A representative unphosphorylated HJURP peptide as internal control. **(b-g)** HJURP phosphopeptides. (h-i) CENP-A phosphopeptides.

		(Roscovitine-treated)/(Mock-treated) Ratio		
Phospho-site(s) covered	Phosphopeptide	Experiment #1 (Forward labeling)	Experiment #2 (Reverse labeling)	Experiment #3 (Reverse labeling)
HJURP pS185	VTPLPSLA**pS**P**S**AVPAPGYCSR	**0.32**	nd	**0.30**
HJURP pS412	WLI**pS**P**S**V**K**	**0.72**	nd	**0.62**
HJURP pS448	EYCL**pS**P**S**R	**0.53**	nd	nd
HJURP pS473	GGPA**pS**P**S**GGLQGLETR	**0.48**	**0.37**	**0.42**
HJURP pS595	**pS**P**S**GQMTVPLCIGVSTDK	**0.23**	nd	**0.25**
	YCLK**pS**P**S**GQMTVPLCIGVSTDK	**0.28**	nd	**0.30**
HJURP pS642	LPS**pS**P**S**LGCR	**0.46**	**0.48**	**0.39**
CENP-A pS16+pS18	**pS**P**pS**P**S**TPTPGPSR	nd	**0.93**	nd
	R**pS**P**pS**P**S**TPTPGPSR	nd	**1.01**	nd
	RR**pS**P**pS**P**S**TPTPGPSR	**0.87**	**0.98**	nd

CENP-A pS16	RR**pS****P**SPTPTPGPSR	**0.94**	**1.13**	nd

Table 1. G1/Mitotic Ratios of HJURP and CENP-A phosphopeptides are reproducible.
Data from three independent SILAC experiments (the data described in **Figure 38**,

Figure 39, and

Figure 40 are displayed in the column labeled "Experiment #1"). For the forward labeling experiment, the "light" cells are treated with Roscovitine while the "heavy" cells are mock-treated. For reverse labeling experiments, the "light" cells are mock-treated while the "heavy" cells are treated with Roscovitine. Red residues are the sites where phosphate groups were unambiguously mapped. Bolded residues are the Cdk consensus motifs in each peptide. (nd=no data)

4.4. DISCUSSION

Maintenance of genomic stability requires accurate propagation of centromere location across cell divisions. Cyclin-dependent kinases (Cdks) are a critical controller of the unique, exquisitely regulated mechanism of CENP-A replenishment at the centromere. Our SILAC effort enables us to quantitatively map the landscape of Cdk-regulation on the CENP-A prenucleosomal complex, including the endogenous HJURP protein. We detected 6 Cdk consensus sites within HJURP, all of which declined in the level of phosphorylation after Cdk inhibition, to varying degrees. In contrast, Cdk consensus sites on the CENP-A N-terminal tail that are known to be heavily phosphorylated4 remain phosphorylated even after Cdk inhibition. Total protein levels of both HJURP and CENP-A were nearly identical between the light and heavy cells, since their unphosphorylated peptides had light/heavy ratios of approximately one. These results are consistent with recent reports that the three HJURP phosphosites within its HCTD1 are Cdk targets (Müller et al., 2014; Wang et al., 2014) but our present data reveal that those particular sites are not the sole targets of Cdk regulation of HJURP, nor are they the most responsive to the silencing of Cdk1 and Cdk2 that recapitulates what occurs during mitotic exit.

4.5. METHODS

4.5.1. SILAC AND AFFINITY PURIFICATION OF PRENUCLEOSOMAL HJURP/CENP-A/H4 COMPLEX

SILAC labelling medium (MEM Eagle Joklik Modification) deficient in lysine and arginine was reconstituted according to manufacturer's instructions (Sigma-Aldrich), and supplemented with normal lysine and arginine (Sigma-Aldrich) for "light" medium, and 50 mg/ L $^{13}C_6$,$^{15}N_2$-lysine and 50 mg/L $^{13}C_6$, $^{15}N_4$-arginine (Silantes) for "heavy" medium. Both media were supplemented with 10% dialyzed FBS (Gemini), GlutaMax (Gibco), 1 mM HEPES, 1% Pen/Strep, MEM non-essential amino acids (Gibco), and 120 mg/L proline to prevent arginine-to-proline conversion. Two parallel cultures of previously characterized HeLaS3 cells stably expressing localization and purification (LAP)-tagged CENP-A (Bailey et al., 2013) were cultured in spinner flasks for at least 6 cell doublings to allow full incorporation of the stable isotope-containing amino acids. Heavy isotope labeling efficiency of ~98% was confirmed by mass spectrometry after trypsin digestion of proteins extracted from heavy-labeled cells. To enrich for mitotic cells, both cultures were treated with 50 μM S-trityl-L-cysteine for 17 h. Subsequently, the "light" cells were treated with 100μM R-Roscovitine (AdipoGen) for 30 min while the "heavy" cells were mock-treated with DMSO. Cell cycle status and HJURP phospho-status was monitored by immunoblotting for H3pS10 (Upstate) and an anti-HJURP antibody generated against a C-terminal fragment (1 μg/ml)(Bassett et al., 2012), respectively. Cell pellets from 1.4 x 10^9 of "light" and "heavy" cells were combined in 1:1 ratio. Affinity purification of the prenucleosomal HJURP/CENP-A/H4 complex was performed as previously described (Bailey et al., 2013) except that protein elution was performed with 2% SDS and heating at 95°C.

4.5.2. MASS SPECTROMETRY AND DATA ANALYSIS

Purified CENP-A and associated proteins were precipitated using pre-chilled acetone (4 X volume) followed by successive washing. Dried protein pellets were

reconstituted with 0.1% RapiGest SF Surfactant (Waters) in 100 mM NH4HCO3, pH 8.0. Resuspended proteins were reduced using DTT, alkylated with iodoacetamide, and digested using trypsin. Since trypsin cleaves only after lysines and arginines, this ensures that every resulting peptide will contain at least one lysine or arginine, so that the all heavy peptides are distinguishable from their corresponding light peptides by predictable mass differences. Rapidgest was removed by adding 0.5% TFA and incubation for 30min at 37°C. The sample was thenThe peptides were desalted with StageTips (Thermo), followed by phosphoenrichment phosphopeptide enrichment by TiO2 prior to analysis by Q-Exactive Hybrid Quadrupole-Orbitrap mass spectrometer (Thermo Fisher Scientific). The pFind search engine was used to search the UniProt human protein database to identify peptides (Wang et al., 2007). Quantification was done using extracted-ion chromatograms (XICs) of each light and heavy peptide pair, and L/H ratio represents the ratio of total area under each elution peak.

CHAPTER 5: CONCLUSION

5.1. SUMMARY

The centromere is a critical region in chromatin upon which assembles the proteinaceous kinetochore complex, and it is known that the histone variant H3 is the epigenetic marker of centromere location. How centromere identity is propagated across cell and organismal generations is a crucial question for biology, since any loss or spurious duplication of centromere location can have disastrous consequences for the cell. CENP-A is known to be remarkably stable at the centromere, and the underlying mechanism for its stability had been poorly understood. In Chapter 2, we identify the role of an essential binding partner, CENP-C, in reshaping CENP-A nucleosomes and stabilizing it at the centromere. We show by multiple structural methods that CENP-C not only binds to CENP-A nucleosomes, but also alters its physical properties. Not only is this a breakthrough for the centromere field that attributes the extraordinary stability of CENP-A to its essential binding partner, but it presented the first example of how a non-catalytic nucleosome binding protein can alter nucleosome conformation.

Although Chapter 2 contains exciting findings, we did not rule out the possibility that the effect of CENP-C on CENP-A stability is indirect, and does not depend on its direct binding to CENP-A nucleosome—since after all, the depletion of CENP-C is a bit of a "sledgehammer approach" that removes many other proteins from the centromere. To show this directly, we would need to take a more surgical approach, to make perturbations in the interaction between CENP-A and CENP-C without derailing the entire CCAN network. Therefore, in Chapter 3, we pinpointed a mechanism by which CENP-C drives such a structural transition in the CENP-A nucleosome via a critical arginine anchor, and we use gene-editing approaches to confirm this mechanism in a CENP-A maintenance assay. By making point mutations in this gene replacement system, we conducted the cleanest set of CENP-C structure-function studies to date, sand correlate the loss of the biophysical structural transition to the decrease of CENP-A retention in cells. We also show that another "friend" of CENP-A at the centromere, CENP-N, also plays a role in stabilizing CENP-

A nucleosomes. We assemble a core centromeric nucleosome complex with the nucleosome-binding domains of CENP-C as well as CENP-N, another critical binding partner of CENP-A at the centromere. We demonstrate that CENP-C and CENP-N can simultaneously engage and rigidify the CENP-A nucleosome at the centromere, and that they both play a role in stabilizing the CENP-A nucleosome in cells.

In Chapter 4, we shed insight into the exquisitely regulated mechanism of nascent CENP-A assembly, which is restricted to a specific phase of the cell cycle. Using the powerful, quantitative SILAC approach, we identify multiple Cdk-dependent phosphorylation sites on the CENP-A chaperone, HJURP, and we measure their exact G1/mitotic ratios. These sites are likely candidates for the regulation of CENP-A assembly.

5.2. FUTURE DIRECTIONS FOR CHAPTERS 2 AND 3

In Chapters 2 and 3, we have elucidated the role of two critical binding partners, CENP-C and CENP-N, in stabilizing CENP-A at the centromere in cells, and we have also reconstituted and examined biophysically the core centromeric nucleosome complex, consisting of the CENP-A nucleosome bound to CENP-C^{CD} and CENP-N^{NT}. Our work opens up many new questions that are just waiting to be investigated.

5.2.1. TOWARD A STRUCTURE FOR THE CCNC

Our finding that as CENP-N^{NT} binds to the CENP-A nucleosome, it makes only a small footprint on CENP-A (at its exposed Loop1) by HXMS, but almost the entire CENP-N N-terminal domain (~200 amino acids) is globally and dramatically stabilized upon binding to the CENP-A nucleosome. We are first to show that CENP-N^{NT}, even before binding to the nucleosome, is a folded domain. And upon engaging with the CENP-A nucleosome, it experiences this striking stabilization of its secondary structure throughout the already-folded domain. For a long time in the field, CENP-N has been notoriously difficult to work with biochemically, since it is known to have low solubility in solution and a tendency to form aggregates. While

we have improved purification of CENP-N^{NT} which greatly extended its "shelf life", our finding provides an explanation for why CENP-N^{NT} tends to be better behaved when in complex with CENP-A nucleosomes: the complex stays stable for weeks in solution, whereas CENP-N on its own (in the absence of glycerol) will precipitate out of solution in a matter of days. Since the complex is well-behaved, obtaining a structural model of the CCNC, either by crystallography or by cryoelectron microscopy (cryo-EM), would obviously be large leaps for the centromere field. Our HXMS indicate that CENP-C^{CD} and CENP-N^{NT} bind simultaneously to different regions on the CENP-A nucleosome, that CENP-C^{CD} is linear and disordered while CENP-N has inherent secondary structure (e.g., helices and sheets). A structure of the CCNC will likely unveil the secondary structure of CENP-N, and provide a valuable model for how both proteins engage with the CENP-A nucleosome.

5.2.2. Structure-function studies of CENP-N

We have shown that rapid auxin-induced degradation of the CENP-N protein results in a decrease in CENP-A stability, and that the domain of CENP-N responsible for this stability is likely its N-terminal domain, which directly interacts with the CENP-A nucleosome. Although previous reports have proposed R11 and R196 of CENP-N as critical residues for this interaction (Carroll et al., 2009), this is unlikely to be the complete picture, since CENP-N likely makes multiple contacts with the nucleosome. We cannot use our HXMS to pinpoint the crucial residues in CENP-N, since the entire N-terminal domain undergoes dramatic stabilization in secondary structure.

The first strategy for structure-function studies might be to make smaller domain deletions with the N-terminal domain of full-length CENP-N. Our HXMS data shows that CENP-N 205-240 are disordered before and after binding to the CENP-A nucleosome, so that is unlikely to be a region involved in the interaction with the CENP-A nucleosome. Within CENP-N 1-205, the a.a. 87-98 region also seems more disordered than its surrounding regions, suggesting that it is possible that CENP-N^{NT} exists as two folded domains (e.g., a.a. 8-86 and a.a. 99-193), with a.a 87-98 as a

flexible linker between the two domains. (However, it is also possible that this region is simply a long loop). One strategy would be to start with CENP-N(Δ8-86) and CENP-N(Δ99-193) to test in a CENP-A-SNAP assay, to see whether one or both of these sub-domains are required for stabilizing CENP-A at the centromere, and then hone in on point mutants. It would then be interesting to test the phenotypically interesting CENP-N mutants as recombinant proteins, by first testing whether they can still make complexes with CENP-A nucleosomes, and if so, testing such complexes in HXMS to examine the effects of these mutants on the protection of CENP-A, as well as the protection of the CENP-N protein itself.

5.2.3. TOWARD BIOPHYSICAL ELUCIDATION OF A LARGER CCNC

In our study, we reconstituted the minimal CCNC that includes the CENP-A nucleosome bound to CENP-C(426-537) and CENP-N(1-240). An obvious extension of this study would be to attempt to reconstitute and biophysically characterize a larger complex. In **Figure 37**, we summarize that not only does CENP-C and CENP-N individually bind the same CENP-A nucleosome, CENP-C also interacts with the CENP-L/N subcomplex (via CENP-C a.a. 235-352, which we mapped in **Figure 28**). A recent study assembled a seven-submit CCAN subcomplex (Weir et al., 2016), but there is much structural insight yet to be learned about the these subunits work together to establish and maintain centromeric chromatin.

5.2.4. CELL CYCLE DEPENDENCE OF THE ROLES OF CENP-C AND CENP-N IN STABILIZING THE CENP-A NUCLEOSOME

We have shown that CENP-C and CENP-N both serve critical roles in stabilizing the CENP-A nucleosome at centromeres, but it is still unknown whether their protective roles are most crucial in a specific phase of the cell cycle, or if they are providing stability to CENP-A nucleosome throughout the entire cell cycle. S-phase is especially intriguing, since the CENP-A nucleosome could be especially vulnerable in S phase, since chromatin undergoes some disruption and subsequent

restoration as it passes through the replication fork (Probst et al., 2009). For CENP-A to be stably transmitted in each cell division, CENP-A nucleosomes need to correctly reassembled behind the replication fork. This process is a crucial, yet poorly understood aspect of centromere inheritance, and it is possible that its binding partners at the centromere are important at the step. While the proteins of the CCAN are constitutively present at the centromere, their localization is dynamic (Fang et al., 2015; Hellwig et al., 2011; Hemmerich et al., 2008). CENP-N exchanges quickly throughout the cell cycle but slows down during late S-phase, at which time its levels at the centromere is increased, and CENP-N is loaded into centromeric chromatin in late S phase (Fang et al., 2015; Hellwig et al., 2011; Hemmerich et al., 2008). CENP-C, on the other hand, is known to be dynamically exchanging across most of the cell cycle, but relatively more stably bound during S phase (Hemmerich et al., 2008).

Therefore, based on the literature, and from our own data, we speculate that CENP-N may play a role for stabilizing (possibly through reassembling) CENP-A nucleosomes after they are disrupted by the replication fork. As CENP-N levels rise in late S-phase, it could mark the centromeres that have completed replication (Stellfox et al., 2013), and lock CENP-A onto the adjacent DNA (by binding to both DNA as well as CENP-A).

5.2.5 Single molecule FRET analyses of the CCNC

Past FRET experiments have shown that CENP-C locks the CENP-A nucleosome into a more compact conformation (Falk et al., 2015, 2016). However, we do not yet have dynamic information about nucleosome fluctuations: i.e., whether the CENP-A nucleosome stably resides in one state (without fluctuating), or whether they fluctuate between two or more discrete states. It is already known that CENP-A nucleosomes are less compact than H3 nucleosomes (Falk et al., 2015, 2016), but it would be interesting to test whether fluctuations (if they exist) differ between CENP-A nucleosomes vs. H3 nucleosomes. Could both types of nucleosomes exist in open vs. closed states, but just that CENP-A spends more time

occupying the open state? (Or is it that only CENP-A nucleosomes can access this open state?) We already expect CENP-C to affect CENP-A nucleosome conformation. But does it do so through dynamics? And of course, how does CENP-C affect CENP-A nucleosome dynamics? If nucleosomes indeed fluctuate between two or more discrete states, CENP-C could change the nucleosome in two possible ways: either a) CENP-C binding *drives* the equilibrium and partition of those states towards the closed state, or b) CENP-C preferentially binds the closed state and blocks fluctuations back to the open state (like a ratchet). Or, if nucleosomes stably reside in an open state (without fluctuating) until being bound by CENP-C, we would expect it to change the conformation to a closed state that is also stable. And of course, does CENP-N change dynamics of the CENP-A nucleosome?

5.3. FUTURE DIRECTIONS FOR CHAPTER 4

In Chapter 4, we identified the Cdk-dependent phosphorylation sites on the CENP-A chaperone, HJURP, using a quantitative phosphoproteomics approach. These sites could be critical sites of the regulation of CENP-A assembly. Previous attempts to characterize HJURP mutants in cells have involved massively overexpressing HJURP, which complicates the analyses, especially for claims for gain-of-function HJURP mutants (Müller et al., 2014). Instead, we will examine HJURP phosphomutants by HJURP replacement in diploid cells, to aim to arrive at unambiguous answers for whether (and how) HJURP phosphomutants affect HJURP function. One method is by adding AID tags onto HJURP in our DLD-1 Flp-In TRex cells, which would allow us to rapidly deplete HJURP protein, a method that was successful for CENP-C (in Chapters 2 and 3). We can subsequently introduce exogenous HJURP rescue mutants into the unique FRT site, which can be either constitutive or inducible (if containing a Tet operator). If the rescue construct is made inducible, we can modulate the level of expression by varying the amount of doxycycline to add to the culture media, which would allow us to arrive at near-endogenous levels of HJURP expression. Additionally, generation of phospho-specific HJURP antibodies would be helpful in characterizing the phosphomutants.

5.4. FINAL THOUGHTS

In the human body, about one hundred billion cells divide over the course of a single day. The process that ensures faithful transmission of genetic material across cell and organismal generations is an exquisitely regulated process, and is undoubtedly one of the most important processes in all of eukaryotic life. Our work has shed insight into the epigenetic maintenance of centromere identity, which is fundamental in ensuring accurate segregation of chromosomes during cell division. The location of the centromere, encoded by CENP-A, is a crucial piece of information for each chromosome—if that information is lost or duplicated, it can result in chromosome breakage, aneuploidy, and cell death.

The inheritance of CENP-A can be viewed as one of the purest examples of "epigenetics", a term first coined by the English embryologist Conrad Waddington in the 1940s to mean "above genetics". Ever since the human genome was sequenced, "epigenetics" has come to mean heritable changes that are not caused by alteration of the genetic code itself—in other words, the hope that human beings are more than just the sum of our genes. It is quite an astonishing result of evolution— that this crucial task of long-term inheritance of centromeric location has fallen on the shoulders not of any DNA sequence, but rather a histone variant. It makes sense, then, that the centromere is quite a sacred place in the chromosome, built with multiple mechanisms for protecting the robustness of this epigenetic mark. The faithful perpetuation of centromere identity relies on the extraordinary long-term stability of the CENP-A molecule, and this thesis provides evidence that CENP-A does not do it alone, but rather with the help of friends at the centromere.

APPENDIX A: PROTOCOLS FOR CHAPTERS 2 AND 3

A1. GENE-EDITING OF DLD-1 CELLS BY CRISPR/CAS

A1A. EXTRACTING GENOMIC DNA

Day 1:

1. Make 1ml DNA extraction buffer:

Recipe	Stock	Amount to add from stock
100mM Tris HCl pH 7.4	1M Tris HCl pH 7.5 at 4C	100µl (1:10)
200mM NaCl	5M NaCl	40µl (1:25)
5mM EDTA	0.5M EDTA	10µl (1:100)
0.2% SDS	20% SDS	10µl (1:100)
1mg/ml Proteinase K	2.5mg/ml stock	400µl
	ddH2O	440µl
		Total: 1ml

2. Lyse pellet in 600µl DNA extraction buffer
   - I've tried either lysing cells directly off of a dish after PBS wash, using a cell scraper to collect cells, or trypsinizing the cells and washing the pellet before adding the DNA extraction buffer-- either should work.
   - Also, you can adjust the volume of buffer based on how many cells you are lysing: I've used 600ul for a 10cm dish which works fine if the cells aren't too confluent. I've also used ~150-200ul for a well for a 12-well plate.
3. Incubate for 30min at 37C (Note: had increased from 0.1mg/ml, which is for doing an overnight incubation at 55C). Lysate should be goopy after incubation.
4. In Eppendorf tube, add equal volume phenol:chloroform:isoamyl alcohol (25:24:1) vortex to mix. (Remember to dispose of phenol waste in special container)

5. Centrifuge 14,000rpm for 5min at RT.
6. Prepare a fresh tube with 0.1V of 3M sodium acetate, pH 5.5 (or 4M ammonium acetate)
7. Transfer upper phase to the fresh tube (should see a white layer between upper and lower phases).
8. Add 2V ice cold 100% EtOH. DNA will become visible.
9. Vortex to mix. Leave overnight at -20C.

Day 2:

1. Centrifuge 14,000rpm for 15min at 4C to pellet DNA. Aspirate supernatant.
2. Wash pellet with 70% ice cold EtOH to remove excess salt.
3. Centrifuge 14,000rpm, 15min at 4C to pellet DNA. Aspirate supernatant.
4. Allow pellet to air-dry.
5. Resuspend pellet in 10mM Tris-HCl, pH 8.0.
6. Measure concentration by Nanodrop.

A1B. DESIGNING GRNAS FOR CRISPR

The protocol described below uses the CENP-A C-terminus as an example, and is adapted from Ran et al, 2013.

1. Obtain genomic sequence for gene of interest (CENP-A, in my case) from the public domain. Use this sequence to design primers for genomic DNA extraction from the specific cell lines to be used for targeting (DLD-1 cells, in my case), to PCR up a region of <900bp (which is the limit at the sequencing core) that spans the location of desired double-stranded break. In this case, the desired location of the double-stranded break would be at the end of the exon 4 (…LEEGLG), right before the TGA stop codon. Alternatively, one could just use the genomic sequence from the public domain to design the gRNAs. However, if there were polymorphisms between that sequence and the cell

lines to be used that result in a sequence mismatch within the bases to which the gRNA is expected to bind, then the gRNAs may not work.

2. Perform genomic DNA extraction on the cell lines (see "Protocol for genomic DNA extraction"), and genomic PCR with the aforementioned primers. Send for sequencing.
3. Using http://crispr.mit.edu, paste 200 bp of sequence (~100bp upstream and downstream of the site of desired DSB) into the window. The site will generate a list of possible gRNAs, each with a "quality score", which represents the inverse likelihood of off-target binding (see **Figure 41**).
    a. Locate each gRNA on the sequence of the gene and determine the location of the DSB (see **Figure 42**). To maximize efficiency of cleavage, choose gRNAs that would cleave within ~30bp of ideal DSB site. The shorter this distance, the better. If this distance is >100bp, the gRNA is unlikely to work.
    b. Among the gRNAs with a distance from ideal site of DSB of <30bp, choose the gRNAs with the highest quality scores. I usually try 3 gRNAs concurrently, and continue with the one that is most efficient.
4. The pX330 plasmid (containing Cas9) contains BbsI sites for insertion of the gRNA. If using this plasmid, order each gRNA as a pair of 20 bp oligos with BbsI site appended (see **Table 2**).

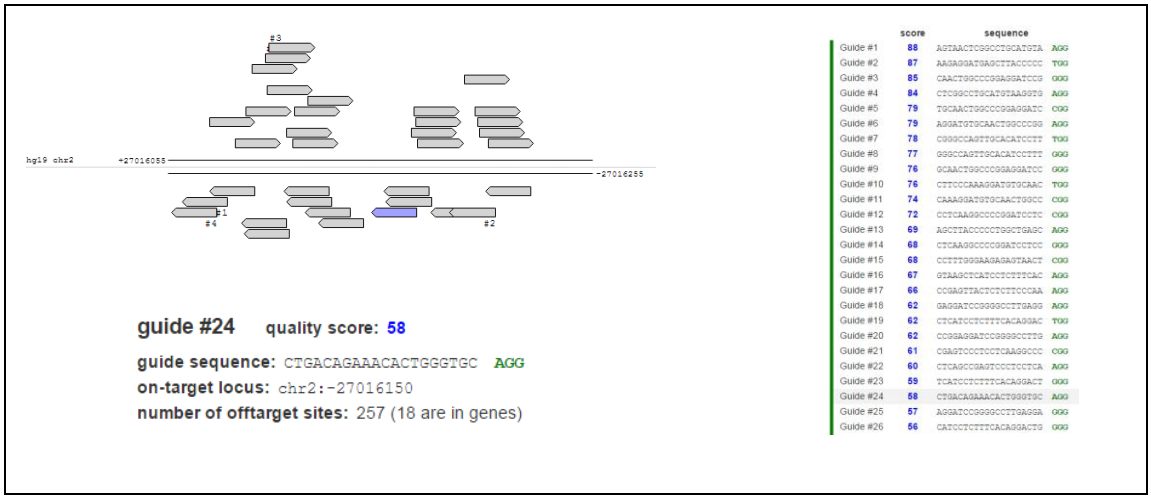


Figure 41. List of gRNAs as generated for CENP-A-SNAP

CRISPR List of possible gRNAs generated by http://crispr.mit.edu, including the quality score (inverse likelihood of off-target binding).

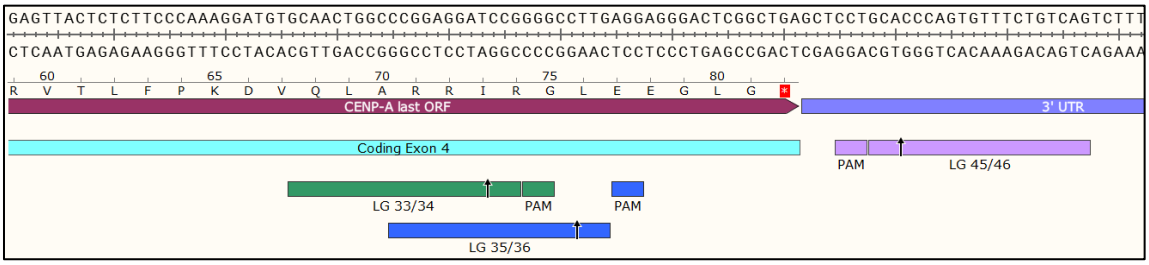


Figure 42. Alignment of gRNAs with CENP-A sequence

Snapshot from SnapGene Viewer showing the locations of the gRNAs (labeled by their oligo number) and their associated protospacer adjacent motif (PAM). Each double-stranded break will be directed to 3 base pairs away from the PAM, and is indicated by a cleavage arrow.

Location of DSB	Sequence of oligo	Oligo number	Distance from ideal site of DSB	Quality score
CENP-A 3' UTR	CACCGCTGACAGAAACACTGGGTGC	LG-045	12 bp	58
	AAACGCACCCAGTGTTTCTGTCAGC	LG-046		
CENP-A Exon 4	CACCGCAACTGGCCCGGAGGATCCG	LG-033	25 bp	85
	AAACCGGATCCTCCGGGCCAGTTGC	LG-034		
CENP-A Exon 4	CACCGCCGGAGGATCCGGGGCCTTG	LG-035	17 bp	62
	AAACCAAGGCCCCGGATCCTCCGGC	LG-036		

Table 2. List of the 3 successful gRNAs that cuts the CENP-A C-terminus.

Within the sequence of each oligo, the 20 bp sequence that anneals to the DNA are bolded (and the unbolded, blue sequence are the BbsI site appended onto each oligo, including an extra guanine).

5. Anneal each pair of oligos, using a protocol such as the one below:
    a. Combine 1µl of oligo 1 (100µM), 1µl of oligo 2 (100µM), 1µl of NEB Buffer 2.1, and 7µl of ddH_2O.

b. Run annealing program on thermocycler: start with 95°C for 1min, then decrease by 5°C every minute, until temperature reaches room temperature (i.e., 95°C for 1min, 90°C for 1min, 85°C for 1min… until 22°C)

6. Linearize pX330 backbone with BbsI
    a. Important note: we have had poor experience with BbsI from NEB. Instead, FastDigest BpiI (ThermoFisher #FD1014) is an isoschizomer is a much more reliable alternate.
    b. Protocol: Combine 2µg of pX330 with 2µl of BpiI, 4µl of 10X FastDigest Buffer, and 26µl of ddH2O. Incubate at 37°C for 1hr. Verify on agarose gel that backbone has been successfully linearized. Then gel-extract (e.g., using QIAquick Gel Extraction Kit, from Qiagen).
7. Ligation: insert gRNA into pX330 backbone:
    a. For each reaction, use 50ng of backbone (Bpil-cleaved pX330), and calculate amount of insert to add:

$$50\text{ng backbone} \times \left(\frac{\text{length of insert}}{\text{length of backbone}}\right) \times 5 = \text{amount of insert to use (ng)}$$

    b. Set up ligation reactions (including backbone-only control).

	Backbone-only control	Backbone + insert
pX330 (BpiI-cut, gel-extracted)	50ng	50ng
Insert (annealed oligos, diluted 1:100 in ddH2O)	0	Amount calculated from previous section
TaKaRa Solution I	5µl	5µl
ddH2O	Add to 10µl total	Add to 10µl total

c. Incubate reactions at 14°C for 1hr. Then transform 1µl of each reaction into competent cells. Grow up cultures and extract DNA by Qiagen Miniprep kit.
d. Verify sequences by sequencing, using a sequencing primer that binds to the U6 promoter (LG-053: GAGGGCCTATTTCCCATGATTCC).

A1C. DESIGNING REPAIR TEMPLATES FOR CRISPR

(using example of adding SNAP-NeoR to the CENP-A C-terminus)

1. Our repair templates have been designed in a pUC19 backbone, containing the synthetic region for insertion flanked by ~800bp of homology on either side (**Figure 43**). The repair template is thus put together by HiFi Assembly (NEB) using 4 pieces: the linearized pUC19 backbone (by XbaI, in this case), the 3' and 5' homology arms, and the synthetic region for insertion. For SNAP-NeoR, the region for insertion was synthesized as a gBlock (IDT), while the homology asrms were derived from PCR of genomic DNA from DLD-1 cells (see "Protocol for genomic DNA extraction"). The primers used for amplifying each region are listed below **(Table 3)**. I try to have overlapping ends of 20-30 bp for every fragment.

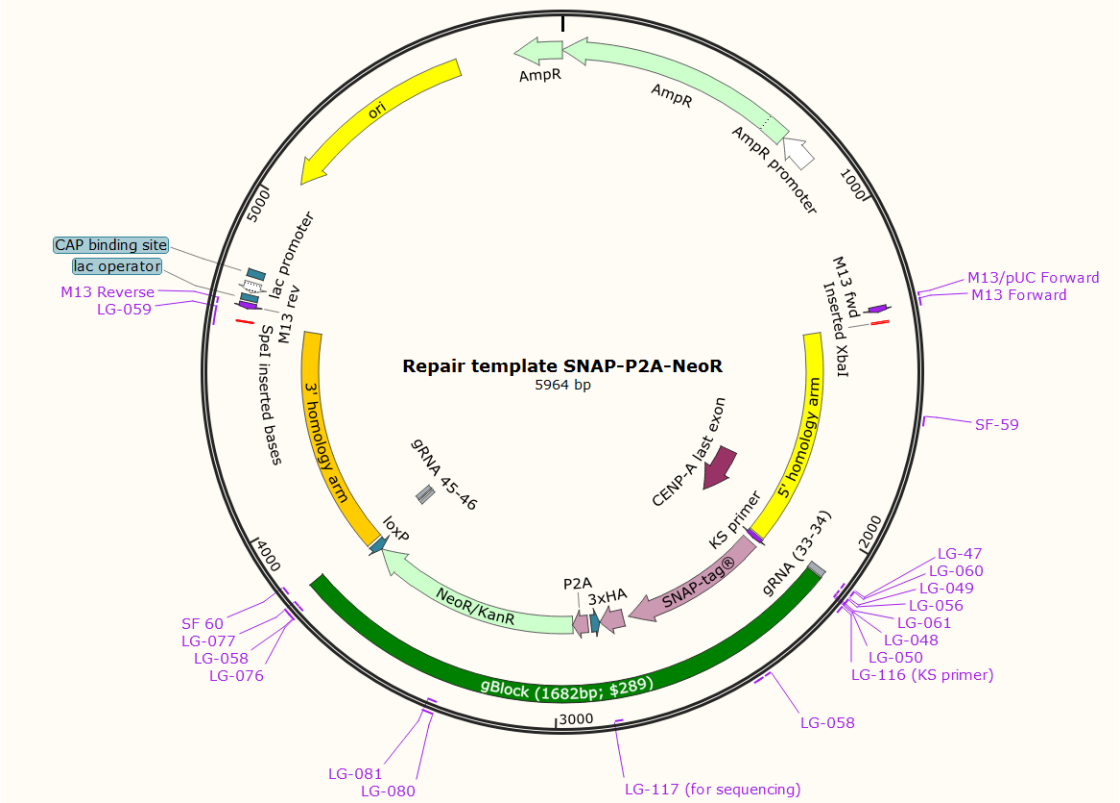


Figure 43. Plasmid map of repair template.

Section of repair template		Sequence of oligos used for PCR	Number of oligo
For 5' homology arm from DLD-1 genomic DNA	F	GTGAATTCGAGCTCGGTACCCGGGGATCC TCTAGAGGAACTCTCTCGTTTGTCCAC	LG-054
	R	CCTCGAGAAGGCCGAGTCCCTCCTCAAG	LG-055
For synthetic gBlock (containing SNAP-NeoR)	F	GCCTTGAGGAGGGACTCGG	LG-056
	R	AAGACTGACAGAAACACTGGGTG	LG-057
	R	ACGCCAAGCTTGCATGCCTGCAGGTCGAC TAGTGCCTTTTCTCCCATACCACAG	LG-059

Table 3. Oligos used to generate each section of the repair template.

2. Set up reaction for HiFi Assembly (NEB), using NEBuilder (online tool) to calculate amount of each DNA piece to add. I have had most success if all 4

pieces for the HiFi Assembly are all gel-extracted (using QiaQuick kit). See **Table 4** as example for setting up reaction. Add ddH2O to 10µl, and add 10ul of NEBuilder HiFi DNA Assembly Master Mix (NEB). Incubate reaction at 50°C for 1hr. Then transform 2µl of this reaction into 50µl competent cells, and plate the entire mix onto LB plates with appropriate antibiotic(s).

Gel-extracted DNA fragment	bp	moles (for 1:1 ratio)	ng to add
pUC19 XbaI-cut	2686bp	0.100 pmol	166 ng
5' arm	846bp	0.100 pmol	52.28 ng
SNAP-NeoR	1682bp	0.100 pmol	104.0 ng
3'arm	884bp	0.100 pmol	54.63 ng

Table 4. Example for setting up HiFi Assembly .

3. Grow up a few clones and obtain DNA by Miniprep. Screen clones by diagnostic digest (500ng DNA is sufficient).
    a. Expected sizes of fragments from NdeI + HindIII diagnostic digest of a successful HiFi Assembly: 2423bp, 1824bp, 1070bp, 647bp. (See **Figure 44** for diagnostic digest of 8 clones, 6 of which were correct).

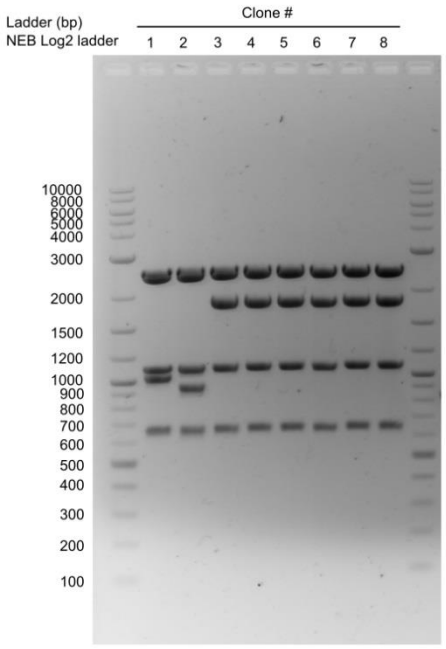


Figure 44. Example of diagnostic digest to screen for correct assembly of repair template

Digest with NdeI and HindIII for 8 different clones obtained from NEB HiFi Assembly. Clones 3-8 matched the predicted sizes of fragments, indicating correct assembly.

4. Verify by sequencing (note: need multiple primers to sequence the entire region of interest spanning both homology arms and the synthetic region for insertion).
5. Important! Make sure to mutagenize the repair template so that the gRNA will not cut it. The easiest way would be to mutate the PAM sequence (i.e., mutate one of both of the Gs in NGG), if the PAM is in a non-coding region or if the mutation can occur at a wobble base. If this is not possible, then mutate ~4 bases in the 20bp gRNA sequence.

A1D. TRANSFECTION AND SELECTION FOR CLONES

1. Transfect DLD-1 cells using Lipofectamine 2000 according to manufacturer's instructions. I transfect DLD-1 cells in 6-well plates at 80-90% confluency. For each well, I have optimized to use 100ng of gRNA plasmid and 900ng of repair template, with 10.5µl of Lipofectamine 2000.
2. Homology-directed repair will occur exclusively in S-phase. I usually wait ~5 days after transfection before beginning selection for clones (by G418

selection, this case, or by FACS if the repair template contains a fluorescent protein).

3. Before beginning selection, verify that the CRISPR was successful in the transfected cells. In this case, I performed immunofluorescence with anti-HA antibody and looked for HA signal at centromeres. All three gRNAs that I attempted (see **Table 2)** yielded HA-positive cells, with LG-045/046 having had the highest efficiency (~1%).
4. Carry out selection for clones. In this case, I treated with 750µg/ml G418 (determined from G418 kill-curve in DLD-1 cells) for 7 days, combined the surviving cells, and performed limiting dilution in 96-well plates (1 cell per 2-3 wells).
5. Expand clones that grow out of the 96-well plates. Once they get to the 12-well stage, can begin extracting DNA to screen by genomic PCR. I have found that DNA from half of a confluent 12-well is sufficient for genomic DNA extraction.
6. Perform genomic PCR on clones.
    a. In this case, I used two primers (SF59: CCTTCCCCACTCCTTCACAGGC and SF60: CCTGTGAAAGAGGATGAGCTTACC) that span the location of the DSB and insertion.
7. Validate clones by immunofluorescence and western blot.
    a. In this case, I obtained both homozygous and heterozygous clones, and this was evident by genomic PCR as well as by western blot (**Figure 45**).

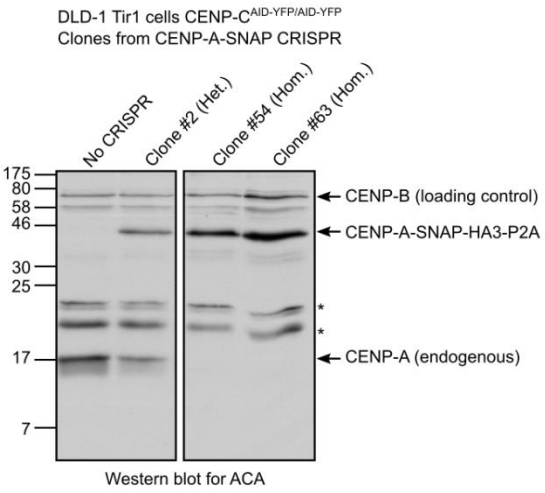


Figure 45. Western blot of clones from CRISPR/Cas-mediated gene editing in DLD-1 cells to insert SNAP-tag at the C-terminus of CENP-A.

ACA is used as primary antibody. From this blot, it can be determined that Clone #2 is tagged on one allele (heterozygous), while Clones #54 and #63 are tagged on both alleles (homozygous).

A2. Pulse-chase SNAP experiments to assess CENP-A

MAINTENANCE

This protocol was adapted for DLD-1 cells from the methods used in Falk, Guo, Sekulic, Smoak et al 2015. Specifically, it was used for the CENP-C$^{AID\text{-}EYFP/AID\text{-}EYFP}$ and CENP-N$^{AID\text{-}EGFP/AID\text{-}EGFP}$ cells, to measure CENP-A-SNAP retention after 24 hours of IAA treatment. This protocol can be adapted to assess additional cell lines (e.g., cell lines with CENP-C rescue constructs at the FRT site), and also for different lengths of IAA treatment.

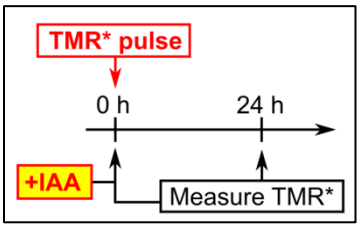


Figure 46. Schematic for TMR*-labeling experiment as described in this protocol.

Day 1:

1. Seed DLD-1 cells expressing CENP-A SNAP.
    a. Seed 1 well in a 24-well plate (labeled “0 hr”), and another 2 wells in a second 24-well plate (labeled “24 hr”). Each well should contain a circular 13mm polylysine-coated coverslip. Seed these wells at 1 x 10^5

cells/well, with ~500µl of media in each well, containing no antibiotics.

b. Seed 1 well in a 6-well plate (labeled "0 hr"), and another 2 wells in a second 6-well plate (labeled "24 hr"). These wells will be used only for cell counting, therefore do not need coverslips in the wells. Seed these wells at 5 x 10^5 cells/well, with ~2.5ml of media in each well, containing no antibiotics (because the surface area of a well in a 6-well plate is approximately 5 times that of a well in a 24-well plate).

c. Note: I have found it easier to seed the 0hr and 24hr cells on separate plates (as described above), instead of seeding them on a same plate. This way, I can directly fix the 0hr coverslip by adding 4% formaldehyde directly to the well (followed by quench, and storage in PBS) then store the entire plate at 4°C overnight (see detailed protocol on immunofluorescence, on Day 3). I have found that this is less stressful than trying to transfer the small 13mm coverslips out of the plate prior to fixation, especially when working with multiple cell lines in parallel.

Day 2

1. Pulse label cells with 2µM TMR* for 15min. Remember to turn off light in the tissue culture hood when working with TMR*, since it is light-sensitive.
    a. To reconstitute SNAP-Cell® TMR* (NEB S9105S), add 30µl tissue culture-grade DMSO to the 30nmol of lyophilized TMR*. Rotate at room temperature, in the dark, for 30min to mix.
    b. Make TMR* solution in media:
        i. Dilute TMR* 1:500 in media, and prepare to use 250µl of this for each well in 24-well plate. Make a master mix to be used for all the wells. Pipet up and down ~10 times after TMR* addition to mix well. Then, spin clarify and transfer to new tube. Keep this warm and in the dark until ready to add to cells.

ii. Similarly, treat the 6-wells (for cell counting) similarly, but do NOT need to TMR*-label these wells (since TMR* is expensive). Instead, just dilute tissue-culture grade DMSO 1:500 in media, and prepare to use 800µl for each well in 6-well plate. Keep this warm and in the dark until ready to add to cells.

c. From the 24-well and 6-well plates, aspirate existing media out of wells.

d. Wash each well carefully 2x with 1x PBS.

i. For this and all subsequent steps for handling cells, add solutions to wells by aiming toward the sides of wells for pipetting (instead of directly on top of cells), and letting the solution slowly drip into the bottom of each well. This is surely gentler on the cells.

ii. Note: others in the lab have noticed problems with cells sloughing off the coverslips, and have found that using media instead of PBS for washes can prevent this. But I have not noticed a problem, so I have consistently stuck with using PBS for washes.

e. Add 250µl of TMR* (1:500 mix) to each well in 24-well plate. (Similarly, add 800µl of DMSO (1:500 mix) to each well in 6-well plate). Put plates back in the incubators. Incubate for 15min exactly.

2. After 15min incubation, aspirate solutions out of wells, and carefully wash 2x with PBS.

3. Add media into each well (500µl for 24-well plates, 2.5ml for 6-well plates), with no antibiotics or drugs, and put plates back in incubators. Incubate for 2 hours.

a. This is a critical step, which serves to allow any unbound TMR* diffuse out of cells. Without this step, the background TMR* signal would be too high in the 0 hr timepoints, which would make analysis difficult. Previously in HeLa cells (as in Falk*, Guo*, Sekulic*, Smoak* et al. 2015), this step had been only 30 min, but I have found that the TMR*

background at the 0hr timepoint in DLD-1 cells is still too high after 30 min, and that extending this to 2hr greatly reduced this background.

4. After the 2 hour incubation, the different plates will be treated differently:
    a. For 24-well plates:
        i. The “0 hr” coverslip:
            1. Aspirate out the media
            2. Add in 500µl of 4% formaldehyde (right into the well). Incubate for 10min. Remember to keep plate in the dark.
            3. Remove formaldehyde, and quench with 500µl of 100mM Tris. Incubate for 5min.
            4. Wash 2x with 1X PBS. For last wash, use large volume (e.g., 1ml), and leave in the well. Wrap the entire plate in foil and store at 4°C until tomorrow. Tomorrow, resume immunofluorescence of this plate along with the “24 hr” timepoints.
        ii. The “24 hr” coverslips:
            1. Aspirate out the media.
            2. Carefully add in fresh media (500µl per well), remembering to add 500µM IAA into the wells designated for IAA treatment.
                a. Write down the time of IAA addition, since this is the beginning of the 24-hr IAA treatment period.
    b. For 6-well plates:
        i. The “0 hr” well:
            1. Aspirate out media
            2. Wash well 1x with 1X PBS
            3. Add 500µl of trypsin, let cells detach from well. Then add 500µl of media.

4. Carefully transfer the entire 1ml into an Eppendorf tube, making sure to monodisperse cells as best as possible. Then count the cells. (Note: it is OK to not count cells immediately, in order to align the "24 hr" wells as best as possible, so that the time of IAA addition is the same between the 24-well and the 6-well plates. But if cells are not counted immediately, do plate the tube on ice so that cells do not keep growing, and count within ~1hr).
5. To count cells, carefully invert tube to mix wells (and/or mix carefully with a transfer pipet), and load 10µl onto the hemocytometer. The hemocytometer has 9 equally sized squares. For my experiments, I always count 5 squares, calculate the average, then load a second time, and count another 5 squares. If the averages between the two loads differ by more than 15%, I count a third load.
    a. To calculate the number of total cells:
        i. (Avg number of cells per square)*10000 = Number of cells per ml.
    b. Note: the hemocytometer is only considered accurate if the number of cells per square is between ~50 and ~200. Therefore, for this protocol, in which I am seeding 5E5 cells per well, I always add 500µl trypsin with 500µl media so that I get between 50-100 cells per square (which equals 5E5 to 1E6 cells total).
    c. The cell counting is undoubtedly the most tedious aspect of the entire protocol, but getting accurate cell counts is absolutely crucial to make the results interpretable. I have attempted to

automate this process with the Countess automated cell counter from ThermoFisher, but I have found this unhelpful. (I got a large degree of variability between loads, especially since DLD-1 cells are not as perfectly circular as the fluorescent beads that they use for demo. Therefore, it did not save me time to use the Countess).

ii. The "24 hr" well:
    1. Aspirate out the media.
    2. Carefully add in fresh media (2.5ml per well), remembering to add 500μM IAA into the wells designated for IAA treatment. Try to do this step as close in time as possible to the IAA addition in the 24-well plate.

Day 3

1. After 24 hours of IAA addition:
    a. For the "24 hr" timepoints in the 24-well plate:
        i. Aspirate out media
        ii. Fix by adding 500μl 4% formaldehyde for 10min.
        iii. Remove formaldehyde and quench with 500μl 100mM Tris.
        iv. Wash 2X with 1x PBS. At this point, pull the foil-wrapped 24-well plate containing the "0 hr" timepoint coverslip out of the 4°C, and add this coverslip to all subsequent steps.
        v. Remove 1x PBS and permeabilize coverslips with 500μl of 0.1% Triton in PBS. Incubate for 5min.
        vi. Wash 2X with 1x PBS.
        vii. Carefully transfer coverslips into a humidified chamber. For me, I use a 15cm tissue culture dish, and place a piece of parafilm on the bottom of the plate, upon which I place my

coverslips (face up). I then place a few rolled-up sheets of wet Kimwipes inside the plate to keep the chamber humidified.

viii. Add 75µl of IF Block (PBS supplemented with 2% FBS, 2% BSA, and 0.1% Tween) to each coverslip, making sure the IF Block spreads across the entire coverslip. Incubate for 20min at room temperature. During incubation step, cover the dish with foil.

ix. Wash coverslips 2X with 1x PBS.

x. Add 75µl of primary antibody mix to each coverslip. Incubate for 1hr at room temperature. During incubation step, cover the dish with foil.

1. For this protocol, my primary antibody mix is mouse mAb anti-CENP-A (Enzo), diluted 1:1000 in IF Block.

xi. Wash coverslips 2X with 1x PBS.

xii. Add 75µl of secondary antibody mix to each coverslip. Incubate for 1hr at room temperature. During incubation step, cover the dish with foil.

1. For this protocol, my secondary antibody mix is Cy5 Donkey anti-Mouse (Jackson Labs), diluted 1:200 in IF Block.

xiii. Wash coverslips 2X with 1x PBS.

xiv. Add 75µl of DAPI (Diluted 1:10,000 in PBST) to coverslips. Incubate for 10min.

xv. Wash coverslips 2X with 2x PBST.

xvi. Mount coverslips to slides with 3.5µl of VectaShield.

xvii. Store slides in the dark at 4°C until ready to be imaged.

b. For the "24 hr" timepoints in the 6-well plate:

i. Count cells as described for Day 2.

Notes about imaging and analysis:

1. Capture images at room temperature on an inverted fluorescence microscope (DMI6000 B; Leica) equipped with a charge-coupled device camera (ORCA AG; Hamamatsu Photonics) and a 40x oil immersion objective. Use 0.59 μm z-sections (which I found to be sufficient for asynchronous cells).
    a. Note: I have tried this experiment with both 40x objective as well as the 100x objective, and have gotten the same results. Using the default parameters, CraQ will pick up more centromeres per cell in the 100x images (~20 centromeres/cell for 100x, and ~9 centromeres per cell for 40x). But because 40x images will contain many more cells per field, I still obtain a greater n with 40x compared to 100x, so I have stuck with imaging in 40x for these SNAP experiments.
2. I perform analysis on deconvolved images, since the deconvolved images will be ultimately used for making figures. But I have found that using non-deconvolved images for quantitation does NOT change the trends of the results.
3. To quantify fluorescence intensity of centromeres, run the CraQ macro on ImageJ (Bodor et al., 2012) with standard settings using DAPI and total CENP-A staining as reference channels to define ROIs for quantification of TMR* intensity.
4. To calculate CENP-A turnover:

$$\text{CENP-A turnover} = \frac{(\text{TMR} * \text{ intensity at 24hr}) - (\text{Avg TMR} * \text{ intensity at 0hr})}{(\text{Cell number at 24hr})/(\text{Cell number at 0hr})}$$

5. The CENP-A turnover for untreated cells should be equal to 100%.

A3. Purification of CENP-N^{NT}-His

Protocol is adapted from McKinley et al., Mol Cell 2015.

Expression:

1. Transform construct into pLysS and grow overnight at 37°C.
2. Grow starter culture overnight at 37°C.
3. Inoculate large flasks of LB + antibiotics with overnight culture.
4. Grow to OD 0.7-1 at 37°C.
5. Turn temperature down to 18°C, induce with 1mM IPTG, and grow for 6 hours. See **Figure 47** for induction bands.
6. Freeze in lysis buffer at -80°C
    a. Lysis buffer: 50mM Na phosphate pH 8.0, 300mM NaCl, 10mM imidazole, 0.1% Tween, 1mM PMSF

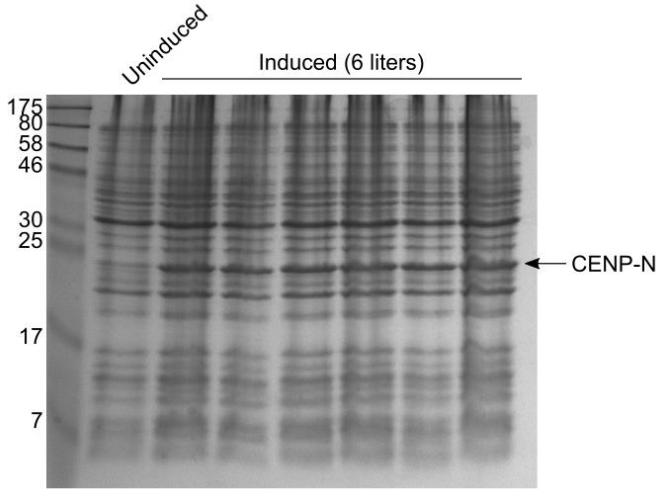


Figure 47. Expression of CENP-N^{NT} in pLysS cells.

Lysis:

1. Thaw pellet in 37°C water bath and place on ice. Add PMSF, aprotinin, and leupeptin/pepstatin. Incubate 15 minutes with occasional stirring.
2. Dounce with 10 plunges using the manual homogenizer.
3. Sonicate: 30 pulses on at 80%, keep on ice for 3min. Repeat for a total of 3 times.
4. Spin at 14700rpm at 4°C in SS-34 rotor for 30min. Take the supernatant, and spin for another 30min.
    a. Note: My experience has been that spinning at 30min just once still leaves the lysate very viscous (which can result in clogging the column). Adding another spin (or alternatively, making the spin longer, such as 50min) has been helpful to reduce viscosity.

Purification with 1ml HiTrap HP column (max pressure 0.3 MPa)

1. Make buffers.
    a. Lysis buffer: 50mM Na phosphate pH 8.0, 300mM NaCl, 10mM imidazole, 0.1% Tween
    b. Wash buffer: 50 mM Na phosphate pH 8.0, 500 mM NaCl, 40 mM imidazole, 0.1% Tween, BME, 20% glycerol
    c. Elution buffer: 50 mM Na phosphate pH 8.0, 500 mM NaCl, 250 mM imidazole, BME , 50% glycerol. Make sure to adjust pH to 8.0
2. Load clarified lysate (from ~2L of cells) onto column, no faster than 1 ml/min. Once loaded, start wash with Wash Buffer at 2 ml/min. Then elute with Elution Buffer at 1 ml/min.
3. Note: Storage of CENP-N protein in 50% glycerol is crucial for long-term stability. I have found that the fractions containing CENP-N are usually so concentrated (~4mg/ml) that further concentration is not necessary. By directly eluting in buffer containing 50% glycerol, further dilution is not necessary. See **Figure 48** for an example gel of final product of the CENP-N purification. It is beneficial to perform an additional gel filtration step, which gets rid of minor contaminants (**Figure 48**).

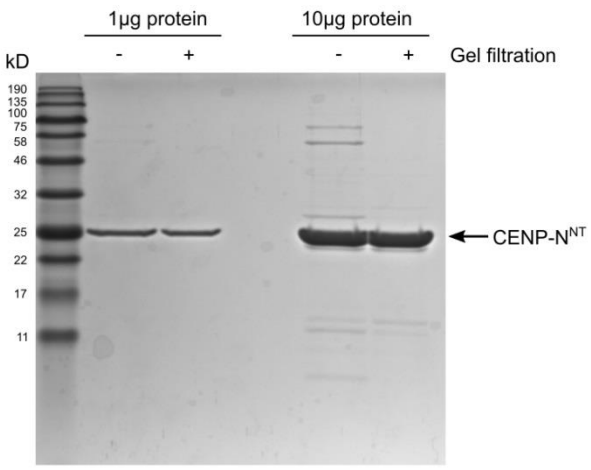


Figure 48. SDS-PAGE gel of final product of CENP-N^{NT} purification.

The gel includes samples from before and after the additional gel filtration step.

A4. Reconstitution of CENP-A nucleosomes and assembly of the Core Centromeric Nucleosome Complex (CCNC)

This protocol of assembling CENP-A nucleosomes by stepwise dialysis is similar to protocols prior described (Dyer et al., 2004; Luger et al., 1999; Sekulic and Black, 2016a), with a few modifications throughout.

1. First, measure concentrations of (CENP-A/H4)2 heterotetramer, H2A/H2B heterodimer, and DNA.
2. Calculate volumes to add for each of the aforementioned components, so that histones are present at 1 heterotetramer:2 heterodimers molar ratio. Components should be concentrated enough that the final concentration of DNA in the dialyzer can be ~0.7mg/ml. And calculate how much KCl stock to add so that the starting reconstitution mix can be 2M KCl (to match the KCl concentration of the Rb-High buffer).
3. Make buffers for reconstitution. Make sure all buffers are pre-cooled.
    a. "RB-High" (high salt buffer): 10 mM Tris–HCl pH 7.5 at 4°C, 2M KCl, 1 mM EDTA, 1 mM DTT. Make 500ml of this.
    b. "RB-Low" (low salt buffer): 10 mM Tris–HCl pH 7.5 at 4°C, 0.25M KCl, 1 mM EDTA, 1 mM DTT. Make 2L of this.

c. “TCS-0” (buffer with no salt): 10 mM Tris–HCl pH 7.5 at 4°C, 1 mM EDTA, 1 mM DTT. Make 1.5L of this.
d. Note: In the Rb-High and Rb-Low buffers, NaCl can be substituted for KCl.

4. Assemble the reconstitution mixture by adding components in this order: DNA, 4M KCl, CENP-A/H4, H2A/H2B, and Rb-High buffer.
5. Incubate mixture on ice for 30min and transfer to dialyzer of appropriate size (after pre-wetting the dialyzer in Rb-High buffer)
6. Begin dialysis with flow rate of 1.6ml/min overnight (~16hr) using a two-channel peristaltic pump as described (Dyer et al., 2004; Luger et al., 1999; Sekulic and Black, 2016a).
7. After 16 hr of dialysis, there should be ~500ml of Rb-Low still left. Move the dialyzer into a fresh beaker containing the ~500ml of Rb-Low. If precipitate is visible, take reaction out of the dialyzer and spin-clarify, then transfer to fresh dialyzer. Dialyze for a few hours in Rb-Low.
8. Dialyze in TCS-0 buffer, 500ml each (x3, a few hours each, the last dialysis usually goes overnight).
9. Check nucleosome quality with native PAGE.
    a. Pour 5% native gel (with 1.5mm combs) with BioRad Mini protean tetra cell casting module. Use ~10ml per gel:

	10ml	20ml
40% acrylamide	1.25 ml	2.5 ml
2% Bis	0.416 ml	0.832 ml
5x TBE	0.8 ml	1.6 ml
H2O	7.87 ml	15.74 ml
TEMED	5 ul	10 ul
APS	50 ul	100 ul

    b. Pre-run native gel at 150V for 1hr, at 4°C. Use 0.2X TBE as running buffer.
    c. Prepare gel samples. Each sample should contain no more than 1-3 µg of nucleosomes (as measured by DNA concentration). Loading 1µg

will result in a very crisp band. If planning to excise the band from the native gel to resolve on denaturing SDS-PAGE, then load 3µg per lane, to allow for more protein per band. Loading more than 3µg will more easily result in smeary bands. Add equal volume 10% sucrose. Load no more than 10µl of volume per lane, otherwise bands are likely to be smeary (because the native gel has no stacking layer, unlike SDS-PAGE gels). But loading less than ~5µl is also suboptimal since the solution may not spread evenly within the well. So loading 5µl-10µl is optimal.

d. After gel is finished running, stain with EtBr, then with Coomassie Blue (see **Figure 49** for example native gel).

10. If the nucleosome prep results in multiple bands (as is common for many α-satellite sequences), perform thermal shifting of nucleosomes, as described (Dyer et al., 2004; Luger et al., 1999; Sekulic and Black, 2016a), to allow histone octamer to slide along the DNA to the most favorable position on DNA. Heat-shift nucleosomes by incubating them at 55°C f or 2hr, then running a native PAGE to verify successful heat-shifting (**Figure 49**).

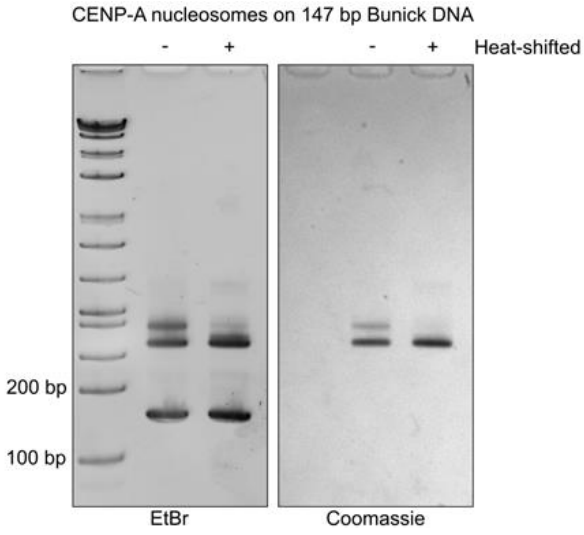


Figure 49. CENP-A nucleosomes assembled 147bp Bunick DNA, with and without thermal shifting. The gel is stained with EtBr as well as with Coomassie Blue.

11. To determine protein components within a band on the native gel, it is helpful to excise the band from the native gel and resolve its components on denaturing SDS-PAGE gel (i.e., 2D gel):

    a. It is imperative to run the denaturing gel (or, at least make the gel samples) as soon as possible after the native gel has destained.
    b. Place the native gel (Coomassie-stained and destained) on a piece of Saran wrap. Use a clean razorblade to cut out the bands of interest.
    c. If the native gel is 1.5mm (as opposed to 1.0mm), it is necessary to cut the band further to fit it within the wells of the 1.5mm SDS-PAGE gel. To do so, I find it easiest to bisect the band along its coronal axis, then placing the two identical pieces vertically in the well of the SDS-PAGE gel, then pipetting in ~20µl of SDS sample buffer into the lane.

For examples 2D gels, see

    d. **Figure 10** and **Figure 31**

12. To assemble complexes of CENP-A nucleosome with CENP-C^{CD} and/or CENP-N^{NT}, incubate with 2.2 molecules of CENP-C for *each* molecule of CENP-A nucleosome, and incubate with 4 molecules of CENP-N for each molecule of CENP-A nucleosomes. See **Figure** 31 for example of these complexes run on native gel. Complexes can be purified by preparative native PAGE as described (Dyer et al., 2004; Luger et al., 1999; Sekulic and Black, 2016a).

13. For making HXMS reactions from these complexes (i.e., comparing nucleosome alone vs. nucleosome bound to protein), it is imperative to add equal volume of Blank buffer to the nucleosome-alone sample, so that the final buffer content can be identical between the two samples to be compared.

A5. HXMS OF NUCLEOSOME AND COMPLEXES

A5A. MAKING BUFFERS FOR HXMS REACTIONS

Keep in mind that pD = pH + 0.4

Make stocks:

- 1M Tris in 10ml D_2O
    - (0.010mol)(121.14g/mol)=1.21g of Trizma Base
    - pH to ~7.1 at room temp with NaOD (so that pD ~7.5)
- 0.25M EDTA in 10ml D_2O
    - (0.0025mol)(372.24g/mol)=0.93g of EDTA
    - pH to ~7.6 at room temp with DCl (so that pD ~8). Note: EDTA will not dissolve until pD approaches ~8.

On-exchange Buffer for timecourse (10mM Tris, pD ~7.5 at RT; 0.5mM EDTA, pD ~8 at RT):

- In 10ml of D_2O, add:
    - 100ul of 1M Tris in D2O (pD ~7.5 at RT)
    - 20ul of 0.25M EDTA in D2O (pD ~8 at RT)
    - No need to further pH

On-exchange Buffer for Fully Deuterated (FD) samples (0.s5% formic acid in D2O):

- In 10ml D_2O, add 57ul of 88% formic acid

Quench Buffer for timecourse (2.5M GdHCl, 0.8% formic acid, 10% glycerol):

- 2.388g of GdHCl
- 91ul of 88% formic acid
- 1ml of 100% glycerol
- Add to 10ml with autoclaved milliQ water
- Sterile-filter through 0.22um filter.
- Adjusting pH of this quench buffer is very important. See below for instructions.

Quench Buffer for fully deuterated (FD) samples (2.5M GdHCl, 0.5% formic acid, 10% glycerol):

- 2.388g of GdHCl
- 57ul of 88% formic acid
- 1ml of 100% glycerol
- Add to 10ml with autoclaved milliQ water
- Adjusting pH of this quench buffer is very important. See below for instructions.

Steps for adjusting pH of Quench Buffers:

1. Objective is to make sure that the pH of our quench buffer + on exchange buffer will be 2.4-2.5. This is the pH at which back-exchange will be minimized).
2. Make all buffers.
3. Then, make separate "test reactions" of 600µl total (since that is the minimum volume required by the pH meter):
   - 60ul TCS buffer
   - 180ul exchange buffer
   - 360ul quench buffer
   - (Thus all components are proportional to each actual reaction, which would contain 5ul TCS buffer with sample + 15ul on-exchange buffer + 30ul quench = 50µl total)
   - (Note: for the FD buffers, make sure to use the FD Quench Buffer + On-exchange buffer for FD samples)
     a. Measure pH of the "test reaction". Then adjust pH of Quench Buffer, and make another 600µl of "test" reaction to measure pH, and continue adjusting the pH of quench buffer until the "test" reaction reaches pH 2.4-2.5.

Store all buffers at 4°C in the dark, and make sure to wrap any solutions with D_2O with parafilm.

A5B. SET UP HXMS REACTIONS

Non-deuterated (ND) samples

- These samples are used for collecting a peptide pool, and also for the all-H control at the beginning of each data collection day.
- For each reaction:
    - 5µl of sample at ~1mg/ml (if samples are nucleosomes or complex, use DNA concentration)
    - 15µl of buffer (e.g., TCS-0 buffer)
- For each reaction, pipet 20µl of this reaction into 30µl of cold Quench Buffer, then flash freeze in liquid nitrogen.
- Note: For nucleosome samples, it is OK to use a histone mix instead of nucleosomes to make the non-deuterated samples (e.g., 1:1 mix of CENP-A/H4 and H2A/H2B), if the peptides in the histone mix have the same retention times as the peptides in nucleosomes (this can be verified empirically).
- In general, I make >20 ND samples for each timecourse.

Fully-deuterated (FD) samples

- Set up a tube containing multiple reactions (I generally make >10 FD samples for each timecourse) of 5µl of sample (at ~1mg/ml) with 15µl of FD On-exchange Buffer. Wrap this in parafilm and foil during incubation.
- When starting to work with a new protein, it is necessary to try various temperatures and incubation times (e.g., 24hr at RT, 24hr at 37°C, 48hr at RT…) and determine which condition allows for the fullest exchange. (See section on "Quality checking FD sample"

On-exchange timecourse samples:

- Prepare tubes by adding 30µl of cold Quench Buffer to each tube, and labeling them exactly according to the planned timecourse (see example timecourse in **Table 5**). Keep all tubes on ice.
- Prepare samples for on-exchange:

    - As in example timecourse, prepare 6 tubes total, 3 of which will contain CENP-A nucleosomes (CA-A, CA-B, and CA-C), and 3 of which will contain CENP-A nucleosomes + CENP-C (CAC-A, CAC-B, CAC-C).
    - Add 27.5µl of sample (at ~1mg/ml) to each tube. (Since each reaction will require 5µl of sample, this calculates for 5.5 reactions).
- To begin the exchange:
    - Pipet 82.5µl of On-Exchange Buffer to each tube (containing 27.5µl of sample). Then immediately start the timer (counting up). At each timepoint, pipet 20µl from the on-exchange reaction into the appropriate tube containing 30µl of cold Quench Buffer. Pipet up and down quickly (especially to mix the glycerol) and drop into liquid nitrogen.
    - Store all HXMS samples at -80°C until ready to run on mass spec.

Time	CENP-A nucs			CENP-A nucs + CENP-C		
10^1s (10 sec)	CA-A1	CA-B1	CA-C1	CAC-A1	CAC-B1	CAC-C1
10^2s (1min 40s)	CA-A2	CA-B2	CA-C2	CAC-A2	CAC-B2	CAC-C2
10^3s (16min 40s)	CA-A3	CA-B3	CA-C3	CAC-A3	CAC-B3	CAC-C3
10^4 (2hr 46min 40s)	CA-A4	CA-B4	CA-C4	CAC-A4	CAC-B4	CAC-C4
10^5s (27hr 46min 40s)	CA-A5	CA-B5	CA-C5	CAC-A5	CAC-B5	CAC-C5

Table 5. Example of timecourse for HXMS, with two samples, 5 timepoints performed in triplicate.

A5C. CROSSLINKING PEPSIN TO POROS RESIN

Make buffers

- Buffer A (500ml): 50mM sodium citrate, pH 4.4. Filter.
- Buffer B (25ml): 1.5M sodium sulphate, pH 4.4. Filter

    - Note: A previous protocol says: the *sodium sulphate should be dissolved in hot (>80C) water in a volume close to the final buffer volume (~5.35g in 25ml). Allow solution to slowly cool at room temperature. When temp is close to 30C, adjust pH to 4.4 (with concentrated HCl) using as little volume as possible to prevent buffer from cooling too rapidly. Allow buffer to cool to 25-30C before use.* However, we have found so far that the high temperature is not necessary for dissolving sodium sulphate.
- Buffer C (1ml)
    - 100mM sodium phosphate
    - 0.1M ethanolamine (6.26µl of 16.6M stock)
    - 5µl of 1M sodium cyanoborohydride
        - This must be hydrated *fresh* with milliQ water and used *immediately*. Beware that this reducing agents very toxic, so wear mask, goggles, and lab coat when weighing out. The dry powder should be kept on a desiccator.
    - Note: the pH will be ~9 after ethanolamine addition. This is above pH 4.4-4.5, But the volume added is small relative to the total volume, so do not need to adjust pH.
- Buffer D (50ml): 50mM sodium citrate, 1M NaCl, pH 4.4. Filter
- Buffer E (500ml): 0.1% formic acid. Filter and de-gas.

Important note: Buffers A, B, D need to be at pH 4.4-4.5, where proteases have low activity. At pH 6 and above the protease is irreversibly inactivated.

Day 1 (makes enough for 2-3 columns (UpChurch, C-130B)

1. Crack open a PD-10 column (GE, 8.3ml bed volume, 5cm bed length) and pour out storage buffer. (Cut open the ridge with razorblade, place needle at the end of column). Equilibrate with 25ml of Buffer A
2. Dissolve protease at 100mg/ml (50mg in 0.5ml of Buffer A)
    - Pepsin: Sigma P6887

- Fungal protease: Sigma P2143

3. Load protease onto column, and chase with 500µl of Buffer A, then 2ml of Buffer A. Collect flowthrough in 500µl gractions (thus we have collected 5 flowthrough fractions).
4. Add 200µl of Buffer A at a time, collecting 200µl fractions. Repeat 15-20 times. Then plug the column and Ponceau-stain the fractions to see where the protein is (using 2µl of sample per dot on membrane). Combine the fractions that contain protease in a 15ml conical.
5. Calculate "x" which is just a scaling factor:

   (total µl of eluted protease)/285=x
6. To the eluted protease, add:

   50(x)µl of 1M sodium cyanoborohydride

   For example, if x=5, then add 250µl
7. Slowly add 160(x)µl of Buffer B
8. Weigh out 100mg of POROS-20AL dry resin (Applied Biosystems #1-6028-02) into an Eppendorf tube. Add enough of the protease mixture to dissolve the resin, and transfer the suspension to the 15ml conical. Gently agitate to form a homogenous solution.
9. Add 330(x)µl of Buffer B over 2 hours (by adding the same amount every 5 minutes. Keep the solution homogenous by intermittent swirling.
   - For example, if the total volume is 1.62ml, we can add 67.5µl every 5 minutes, for 24 times total.
10. Tumble gently overnight at RT.

Day 2

1. Next morning, quench the reaction with 75µl of Buffer
2. Tumble gently for 5 more hours at room temperature
3. Using a fine sintered glass funnel, wash the resin with: 6ml Buffer A, then 5ml Buffer D, then 6ml Buffer A
4. Gently use spatula to scrape resin from funnel and place in 15ml conical
5. Resuspend resin with Buffer A to make 50:50 solution

6. Store in a 15ml conical at 4°C

   Note: unused cross-linked proteases usually last several months. Can still be used after 6 months, but may lose some efficiency.

A5D. SETTING UP AND CALIBRATING THE ORBITRAP

Preparing system to run samples

1. Prepare "Pre-injection Mix" of 3:2 Quench Buffer: milliQ water. This will be at 1M GdHCl, and is the solution to be used for all washes.
2. Turn on the power switch at the bottom of the cart.
3. Connect the HXMS Box to start cooling. Make sure the arrow points toward the blue dot (otherwise the box will be heated instead). Place a temperature probe inside the box, and monitor temperature of the box throughout the day.
4. Turn on the Waters pump.
5. Take knobs off of the tubing in the Box.
6. Connect the tubing to the Waters pump, but do so under flow: set Waters pump to 0.050ml/min and press "Run".
7. Flip on the air switch for the Eksigent (on the wall), and turn on the Eksigent (the switch is on the side of the box). The green light should come on on the other side.
8. On the computer, open XCalibur software. Wait for "Eksigent LC Channel 1" to say "Ready to Download".
9. Open Eksigent software on computer. Go to "System" → "Direct Control". Set to 12%, at 6µl/min. Then click Start.
10. Now connect the tubing from the Eksigent to the HXMS Box.
11. Open ThermoTunePlus on the computer. Then open "LMtune.LTQtune."
12. Begin calibration
    a. Load the syringe attached to the Orbitrap with Calibration Mix (CalMix). Press the Pump button on the Orbi and wait for masses to

appear on the computer. Check to make sure they are relatively stable. The NL should be in the 10^7 range.

b. In ThermoTunePlus, go to "Control" → "Calibrate"
c. First check to see if Orbi is already calibrated. Go to the "Check" tab, and check "Mass Calibration" under "Positive Ion Mode". Document the Universal and Low RMSD values in the Excel spreadsheet on the desktop. If values are >1.5, then calibrate. Otherwise, can proceed without calibration.
d. To perform the actual calibration, go to the "Semi-Automatic" tab, and check "Mass Calibration" under "Positive Ion Mode".
e. After calibration, perform another check: Go to the "Check" tab, and check "Mass Calibration" under "Positive Ion Mode". Document the Universal and Low RMSD values in the Excel spreadsheet on the desktop.

13. After calibration, make sure to connect the tubing from the Orbi to the HXMS Box (instead of to the syringe for CalMix).
14. Now ready to start the day of runs. Start with a pre-clean run.
    a. Choose a name for the run. Specify Path and Inst Method in XCaliber. For pre-clean and washes in between runs, Inst Meth should be "C:\Xcalibur\methods\Lucie\3-29-11 yeast cleanup_tp".
    b. Check the method. Right click → "Open File" to make sure everything is correct, especially the pre-flush times.
    c. Click "Run Sample" (not Run Sequence), and do NOT click "OK" yet.
    d. Make sure both valves are in P1.
    e. Draw up 50ul of Pre-Injection Mix into the Hamilton syringe. Inject. Then quickly change V1 to P2, and click "OK" on the computer to start running the method. Prepare to change the water pump from 0.050ml/min to 0.150ml/min, but do NOT click "Menu" yet.
    f. After 2min, click "Menu" to change the water pump to 0.150ml/min

g. After another 2min (4 min total), flip V2 to P2. Then change water pump back to 0.050ml/min.
h. Wait for 15min clean-up method to finish.

15. After the pre-clean run is finished:
    a. Set both valves to P1
    b. Make sure the water pump is on 0.050 ml/min
    c. Check temperature probe in the box. Temperature should be close to 0C.
    d. Wait for pressures stabilize.
    e. Rinse the syringe and tubing:
        1. Rinse the Hamilton syringe 3 times with 1% formic acid + TFA pH 2.20 (the buffer for the waters pump).
        2. Rinse the injection tubing: With V1 in P1, fill up the Hamilton syringe with 1% formic acid + TFA pH 2.20 and inject into V1. There should be liquid oozing out of the bypass tubing. Catch with Kim wipes to prevent a puddle.

A5E. RUNNING ND SAMPLES FOR GENERATING PEPTIDE POOL

1. Keep all HXMS samples on dry ice before injecting into the system.
2. After the pre-clean (see previous section), start an ND run.
3. For ND runs, specify "C:\Xcalibur\methods\Lucie\20111205_forNDs_4minpreflush" for Inst Meth.
    a. The method consists of a linear 12-55% buffer B gradient at 6 µl/min (Buffer A: 0.1% formic acid; Buffer B: 0.1% formic acid, 99.9% acetonitrile). The method for ND will collect MS/MS data.
4. Click "Run Sample" (not Run Sequence), and do NOT click "OK" yet.
5. Make sure both valves are in P1. Take sample out from dry ice and let thaw slowly at 0°C. Draw sample (50ul) into the Hamilton syringe. Inject. Then quickly change V1 to P2, and click "OK" on the computer to start running the

method. Prepare to change the water pump from 0.050ml/min to 0.150ml/min, but do NOT click "Menu" yet.

6. After 2min, click "Menu" to change the water pump to 0.150ml/min
7. After another 2min (4 min total), Change V2 to P2. Then change water pump back to 0.050ml/min. Then wait for the 25min run to finish.
8. In between each run, perform a wash step.
9. After the ND run (and after each subsequent run), do a wash run (same method as the pre-clean).
10. To continue building the peptide pool:
    a. Use SEQUEST to generate a list of peptides from this ND run. Then use ExMS to generate an exclusion list with these peptides, and load this exclusion list into the Method file of the next ND run. See section "A5g. Generating exclusion list and final peptide pool" for more details.
    b. For the 3rd ND run, generate an exclusion list containing peptides from both of the first two runs, and load that into the Method file.
    c. You should see that the number of peptides within the Ppep threshold (e.g., <0.90) will decrease with every iterative ND run. Depending on the length of your protein, usually after ~4 runs, we start getting diminishing returns.

A5F. PROTEIN SEQUENCES

CENP-A/H4/H2A/H2B histone mix

Histone CENP-A (MS index 1-140):

MGPRRRSRKPEAPRRRSPSPTPTPGPSRRGPSLGASSHQHSRRRQGWLKE
IRKLQKSTHLLIRKLPFSRLAREICVKFTRGVDFNWQAQALLALQEAAEAF
LVHLFEDAYLLTLHAGRVTLFPKDVQLARRIRGLEEGLG

Histone H4 (MS index 141-243):

MSGRGKGGKGLGKGGAKRHRKVLRDNIQGITKPAIRRLARRGGVKRISGLI
YEETRGVLKVFLENVIRDAVTYTEHAKRKTVTAMDVVYALKRQGRTLYG
FGG

Histone H2A (with N-terminal GPLG tag remnant) (MS index 244-377):

GPLGMSGRGKQGGKARAKAKSRSSRAGLQFPVGRVHRLLRKGNYAERVG
AGAPVYMAAVLEYLTAEILELAGNAARDNKKTRIIPRHLQLAIRNDEELN
KLLGKVTIAQGGVLPNIQAVLLPKKTESHHKAKGK

Histone H2B (MS index 378-503):

MPEPAKSAPAPKKGSKKAVTKAQKKDGKKRKRSRKESYSVYVYKVLKQV
HPDTGISSKAMGIMNSFVNDIFERIAGEASRLAHYNKRSTITSREIQTAVRL
LLPGELAKHAVSEGTKAVTKYTSSK

CENP-C^{CD}

MAKPAEEQLDVGQSKDENIHTSHITQDEFQRNSDRNMEEHEEMGNDCVSKKQM
PPVGSKKSSTRKDKEESKKKRFSSESKNKLVPEEVTSTVTKSRRISRRPSDWWVV
KSEE

CENP-N^{NT}-His

MDETVAEFIKRTILKIPMNELTTILKAWDFLSENQLQTVNFRQRKESVVQHLIHLC
EEKRASISDAALLDIIYMQFHQHQKVWDVFQMSKGPGEDVDLFDMKQFKNSFKK
ILQRALKNVTVSFRETEENAVWIRIAWGTQYTKPNQYKPTYVVYYSQTPYAFTSSS
MLRRNTPLLGQALTIASKHHQIVKMDLRSRYLDSLKAIVFKQYNQTFETHNSTTP
LQERSLGLDINMDSRIIHEGSSHHHHHHH

A5G. GENERATING EXCLUSION LIST AND FINAL PEPTIDE POOL

- Summary: Depending on protein length, running ~4 NDs should provide a complete peptide pool. Following the first run, use ExMS to generate an

exclusion list from the SEQUEST results. For the next run, upload the exclusion list into the Xcalibur method file.

- To create FASTA database:
    - The database should include the sequences of the proteins in the sample (see protein sequences of my samples, in previous section).
    - Also, make sure to include ALL the proteins used by the lab as well as the protease (pepsin or fungal protease) used. Human keratins, MS standards, and the proteome of the organism of expression are also good ideas. You can also include the reverse protein sequences.
- To generate Excel file of SEQUEST results from RAW file:
    - Open Bioworks Browser software
    - Drag the run (.RAW file) into the window
    - Go to Actions → SEQUEST Search (to match against a database)
    - Specify .params file (with appropriate FASTA database)
    - Use peptide tolerance of 8 ppm (default is 4 ppm), fragment tolerance of 0.1 (LM) AMU
    - Right click → Display → Peptides → make sure Retention times is checked (not Scan)
    - Right click → Export → Excel
    - This will generate an excel file containing all the peptides found in the run, including their Ppep scores.
- Generating an exclusion list in ExMS:
    - Select 5: MS/MS related utilities
    - Select 2: To make MS/MS exclusion (reject) list
    - Select 1. Thermo Bioworks/SEQUEST (the result table saved as .xls file)
    - Designate P score (I use 0.90)
    - Xcorr score: [1,1,1,1,1]
    - Number of MS/MS experiments: 1 on first run, 2 on second, etc.
    - RT range of MS/MS experiments: [0 25] for a 25 minute run
    - RT window for exclusion list: 1 s

    - This should generate a .txt file. In the next run, import this exclusion list into the Method file of the run.
- To generate peptide pool in ExMS from several iterative ND runs:
    - To begin analyzing ND sample (which will ask to generate a peptide pool as the first step), start ExMS and choose 1: Start auto-processing HX MS dataset
        - Alternatively, to just generate the pool without analyzing a run, choose 5: MS/MS related utilities → 1: To make experimental/theory peptide pools
    - Input name of protein and sequence. For my sequences, see "A5f. Protein sequences".
    - Choose 1: Thermo Bioworks/SEQUEST (the result table saved as .xls file)
    - Input the score system to be used (1=XCorr; 2=Ppep): Ppep.
    - Input Ppep score threshold for filtering peptides: 90
    - How many MS/MS experiments (result tables of peptide search) to establish the "peptidesPool"?: Enter how many ND runs from which to build the peptide pool.
    - Then, choose each .xls file to be analyzed.
    - ExMS will then generate the peptide pool file ("ExMS_preload") with an accompanying .fig file (see examples in **Figure 50** and **Figure 51**).

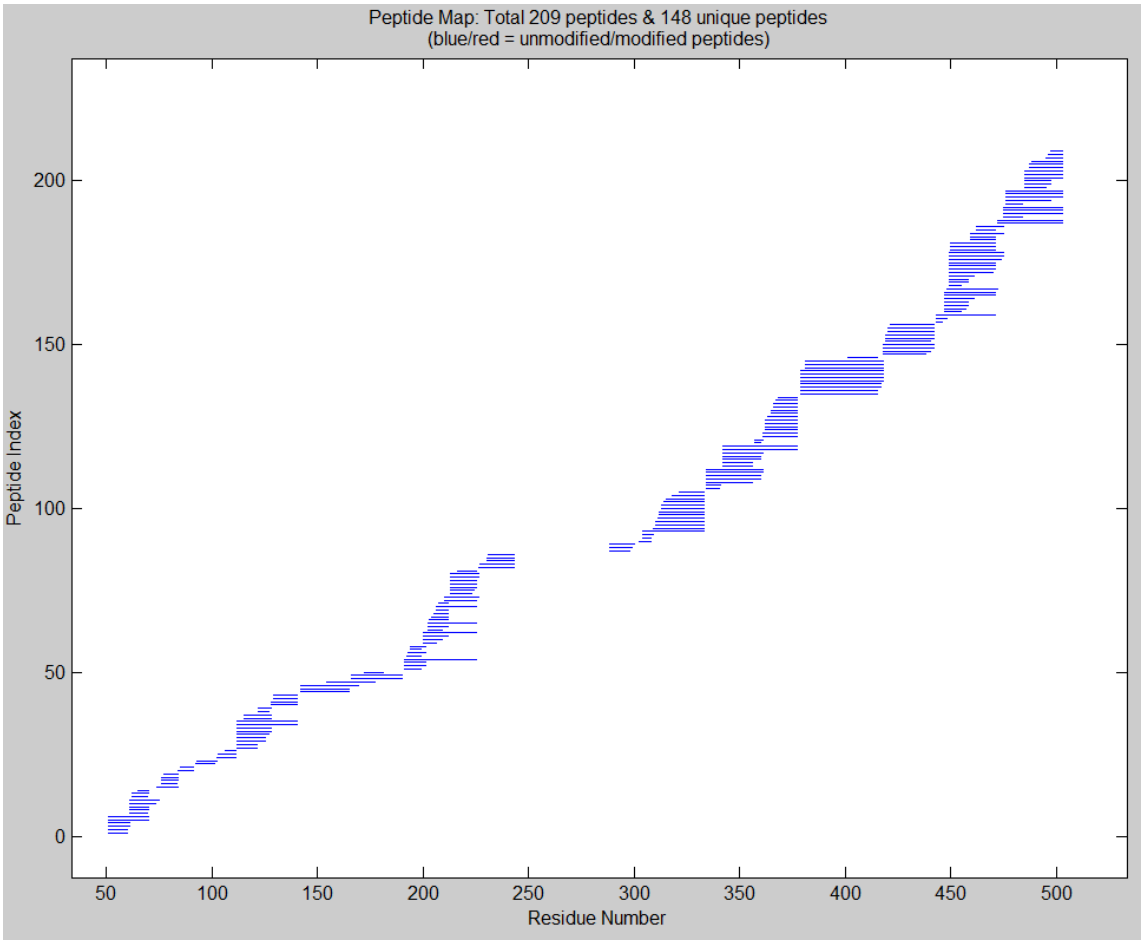


Figure 50. Peptide pool showing CENP-A/H4/H2A/H2B from several iterative ND runs.

See "A5f. Protein sequences" for how the residue number lines up with each protein.

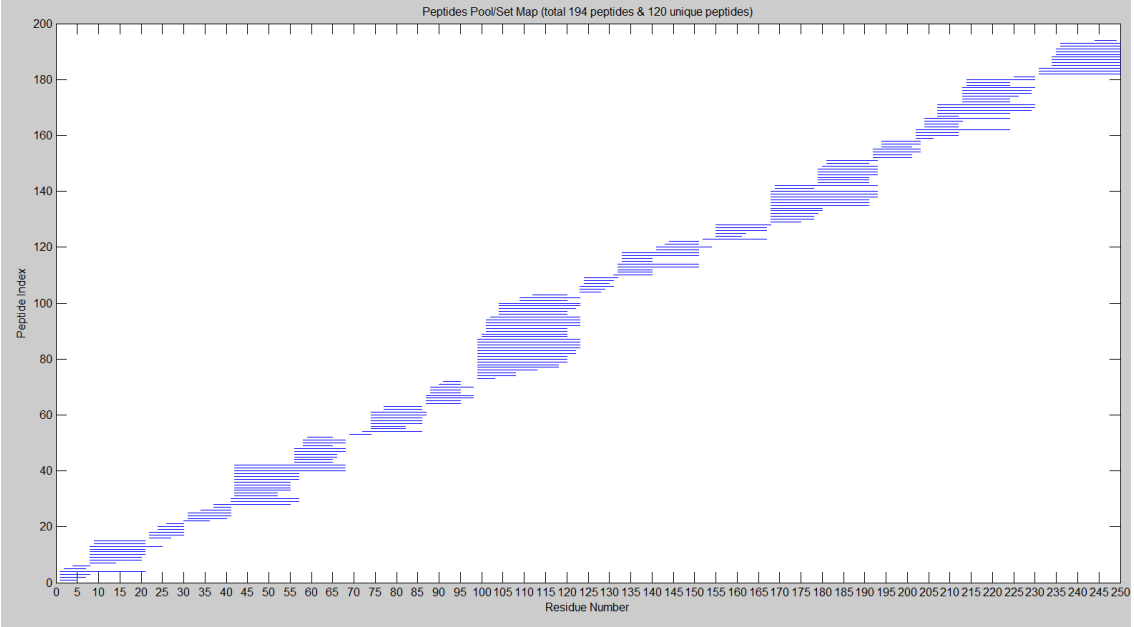


Figure 51. Peptide pool showing CENP-N^{NT}, from several iterative ND runs.

A5H. RUNNING THE DEUTERATED TIMECOURSE SAMPLES

1. Keep all HXMS samples on dry ice before injecting into the system. Remember to let each sample thaw slowly at 0°C on ice.
2. Start with a pre-clean,,then start the day with an All-H run (with an ND sample). See "A5e. Running ND samples for generating peptide pool" for protocol and method file. Then make sure that the retention times of the peptides in this run align well (within ~half a minute) to those in the peptide pool.
3. I usually run the FD sample as the second run of the day. For the FD run (and for all deuterated runs), specify "Deuterium_run_MS1_12to55_tp" as the Method file. Just like the file for the ND runs, this file consists of a linear 12-55% buffer B gradient at 6 µl/min (Buffer A: 0.1% formic acid; Buffer B: 0.1% formic acid, 99.9% acetonitrile). But it does not collect MS/MS data.
1. After the FD, start running timecourse samples, with a wash run between each run. For my first dataset with CENP-A nucleosomes +/- CENP-C, I also ran an empty gradient (a 25min run, but without injecting sample) after each wash, to make sure that no peptides have been carried over from the

previous run. I never saw any carryover of peptides, and when I repeated the dataset, I saw no difference in the results. So for subsequent datasets, I no longer did the empty gradients.

4. At the end of day, before disconnecting the system, remember to store the analytical C18 HPLC column in 90% acetonitrile: bump up to 90% Buffer B stepwise (e..g., start with 20%, then up to 30%, then up to 40%... to 90%).

Notes:

- I usually run the same timepoint of two samples back-to-back, to minimize the variability in data collection between the two samples (and randomly select which one to run first). For example, for the timecourse shown in **Table 5**, I can run CA-A1, then CAC-A1, then CAC-A3, then CA-A3).
- I have found by experience that the data looks tighter if the entire timecourse is run over a single day, or over two consecutive days. (Note: if doing runs continuously, make sure to re-calibrate the Orbi every 12 hours).
- Make sure that all plumbing is running well before starting the timecourse, because changing any component of the plumbing (especially the C18 analytical column) can dramatically alter the retention times of peptides, which will make analysis very problematic.

A5I. ANALYSIS OF DEUTERATED SAMPLES BY EXMS

1. Convert RAW files to mzXML files (to be used by ExMS)
    a. Open TPP Web Tools (purple flower icon). Username: guest. Then password should auto-fill (should also be guest)
    b. On the desktop there is a link to "TPP data directory". (C:\Inetpub\wwwroot\ISB\data). Draw .RAW files into this folder
    c. Then on the web interface: Analysis Pipeline
        i. Specify RAW Input File(s) to convert to mzXML (select our files)

ii. Conversion Options: check “Profile” (not Centroid)
iii. Convert

d. Now you should see .mzXML files appearing in the TPP data directory folder. Remove them from this public folder immediately after transferring the files.

2. Use ExMS to auto-process the mzXML files by selecting “1: Start auto-processing HX MS dataset”, and following the steps as asked by the program. After preprocessing, ExMS will generate an “ExMS_wholeResults” file.
    a. We can perform auto-check of the results by selecting “2: Check auto-processed results or re-check” and selecting “1: To just use this auto-determined good peptide set (Will need no manual check)”. This will generate an “ExMS_wholeResults_afterCheck” file.
3. Quality check the fully deuterated (FD) sample
    a. To check whether the FD sample will serve as an adequate control for full deuteration, run the FracDeut script (written by J. Dawicki-McKenna). The code requires the “ExMS_wholeResults_afterCheck” files for both the ND and the FD. The code will calculate the theoretical full deuteration of each peptide and compare it to the experimental degree of deuteration, then plot a histogram of fractional deuteration for all peptides (**Figure 52**). For my FD samples of histone mixes, I usually get a median fractional deuteration of >85%.

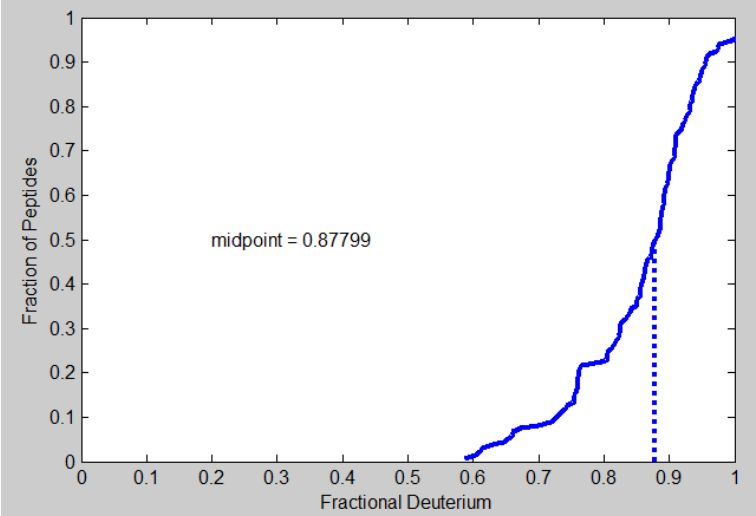


Figure 52. Example FracDeut analysis of a fully deuterated (FD) histone mix sample.

4. Manually check peptides
    a. After letting ExMS perform autocheck of the data, you can manually check each peptide or selected peptides. To perform manual checking, select "2: Check auto-processed results or re-check". Then, you can manually check every peptide by choosing "3: To check all peptides no matter auto-determined good or not (Maximum manual check)", or just check the peptides that ExMS could not confidently identify by choosing "2: To check those peptides outside the good set (Partial manual check)". For guidelines on how to perform manual checking, see published literature (Kan et al., 2011).
5. Plotting data
    a. Data can be plotted into figures by using Matlab scripts written by T. Panchenko and M. Salman. This code first prompts the user for input, then performs necessary calculations and creates HDX Ribbon plots, Difference plots where applicable as well as XY plots for all data sets.

APPENDIX B: PROTOCOLS FOR CHAPTER 4

B1. SILAC EXPERIMENTS

B1A. MAKING MEDIA FOR SILAC

Overview: this protocol makes 4L of SILAC media, which emulates the Joklik formula (from Sigma). My SILAC experiments were always 2L light + 2L heavy, since this was determined to be the minimum scale needed to detect the phosphosites of interest. This media is made completely from scratch, by adding nutrient powder into milliQ water. Lysine and Arginine are added later.

1. First measure out 3L milliQ water in a 4L flask, keep this stirring at room temperature.
2. Weigh out each individual component in table below. Note that many of these components require some time to dissolve, so for best results, keep the water stirring as the components are added.
    1. Note: for vitamins, since so little quantity is required, I found it easiest to weigh out the smallest quantity that can be comfortably measured and then dissolve that in 1ml ddH2O. Then, add the volume that would correspond to the appropriate final quantity.

Component	mg/L	For 4L
Salts and sugars		
Magnesium chloride x 6H2O (Sigma M9272)	200	800
Potassium chloride (Fisher BP366)	400	1600
Sodium chloride (Sigma S9888)	6500	26000
Sodium dihydrogen phosphate x H2O (Sigma S9638)	1327	5308
D(+)-Glucose anhydrous (Sigma D9434)	2000	8000

Amino acids (NOT including Lys, Arg, and Pro, which are to be added later)		
L-Cystine (Sigma C7352)	24	96
L-Histidine base (Sigma H6034)	31	124
L-Isoleucine (I7403)	52	208
L-Leucine (Sigma L8912)	52	208
L-Methionine (Sigma M5308)	15	60
L-Phenylalanine (Sigma P5482)	32	128
L-Threonine (Sigma T8441)	48	192
L-Tryptophan (Sigma T8941)	10	40
L-Tyrosine (Sigma T8566)	32.6	130.4
L-Valine (Sigma V0513)	46	184
Vitamins		
D-Calcium pantothenate (Sigma P5155, stored at 4°C)	1	4
Choline chloride (Sigma C7527)	1	4
Folic acid (Sigma F8758)	1	4
Myo-inositol (Sigma I7508)	2	8
Nicotinamide (Sigma N0636)	1	4
Pyridoxal-HCl (Sigma P6155, stored at -20°C)	1	4
Riboflavin (Sigma R4500)	0.1	0.4
Thiamine-HCl (Sigma T1270)	1	4

3. After adding all components, wait to make sure there are no precipitates.
4. Then add phenol red to media (10mg/L, so add 40mg for 4L)

5. At this point, the media will look yellow (acidic). Add NaOH until it looks more red.
6. Measure pH, and add more NaOH until pH reaches ~7.
7. Add 8g of sodium bicarbonate (Sigma S9638)
8. Pour in water to ~3.6L.
9. Measure pH again. Adjust pH to ~7.2-7.3.
10. Sterile-filter into 1L bottles using 0.22μm filter, 900ml per bottle (thus leaving 100ml for adding serum). This is now a master stock of Joklik media, and can be stored at 4°C for several weeks. Note: the media should go through the 0.22μm filter easily. If it is failing to go through easily, it might indicate that there is precipitation in the media, and that it is unable to be used since it could be missing key nutrients.
11. When beginning a SILAC experiment, make light and heavy media separately, from this master stock.
12. Make separate stocks of lysine, arginine, and proline to be added individually. Store these hydrated stocks at -20°C until ready to use. Can refreeze and thaw as many times as needed.
    a. For lysine (make 50mg/ml stock in ddH2O, which Is 1000x)
        1. Light: L-Lysine (Sigma L5501)
        2. Heavy: $^{13}C^{15}N$ labeled Lysine-HCl, 100mg (Silantes 211603902)
    b. For arginine (make 50mg/ml stock in ddH2O, which Is 1000x)
        1. Light: L-Arginine (Sigma L5006)
        2. Heavy: $^{13}C^{15}N$ labeled Arginine-HCl, 100mg (Silantes 201603902)
    c. For proline (Sigma P0380), make 120mg/ml stock in ddH2O, which is 1000x
13. To make 500ml of light (or heavy) media:
    a. Add 450ml of Joklik media (from master stock)
    b. 50ml of dialyzed FBS (Gemini # 100108). It is important to use dialyzed FBS, since normal FBS contain trace amounts of lysine and arginine.

c. 5ml of GlutaMax (Gibco 35050-061, which is at 100X). I find this easier to work with compared to L-Glutamine, especially since this stock is stored at room temperature.
d. 5ml of Pen/Strep
e. 5ml of 1M HEPES
f. 5ml of 100X MEM Non-Essential Amino Acids Solution (Gibco 11140-050, containing Gly, Ala, Asn, Asp, Glu, Pro, and Ser).
g. 500µl of 1000x Lysine (50mg/ml) – light or heavy
h. 500µl of 1000x Arginine (50mg/ml) – light or heavy
i. 500µl of 1000x Proline (120mg/ml)

B2B. GROWING HELAS3 SPINNER CELLS IN SILAC MEDIA

1. Make sure spinner flasks are autoclaved, once in ddH2O, and once dry.
2. HeLaS3 cells expressing CENP-A-LAP (Bailey et al., 2013) are frozen in liquid nitrogen at 2 x 10^7 cells/vial.
3. Make sure to pre-warm media in 37°C water bath prior to use.
4. Combine two vials into some volume of normal media (e.g., 50ml), then split into two equal 50ml conicals (e.g., 25ml each). Then spin down gently (800rpm, 3min), and carefully remove the media. Then, to one cell pellet, add 50ml of Light media. To the other conical, add 50ml of Heavy media. Carefully and thoroughly mix each tube of cells using a transfer pipet. Monodispersing the cells is crucial for optimal growth, since clumped cells do not grow as well.
    a. Thus, both the Light and Heavy flasks are starting from equivalent cells, and the starting concentration of cells in each flask is ~4x10^5 cells/ml.
5. Gently pour each conical of cells into a pre-warmed 100ml spinner flask. Make sure to label which flask is Light, and which is Heavy.
6. Check cell density and viability daily. Do not let the cells overgrow (>1x106 cells/ml), and do not split them to lower than ~3x105 cells/ml. If cells are

doubly daily, and percent of dead cells is <10%, then they are growing well. Every day, add more pre-warmed media (light or heavy, according to which flask).

7. Once the capacity of the flask is maxed out, transfer to the next biggest size of spinner flask (and make sure it is autoclaved twice-- first in water and next in dry-- and pre-warmed to 37C).

APPENDIX C: LIST OF PLASMIDS

Plasmids containing guide RNAs (gRNA), for transfection in mammalian cells (for CRISPR/Cas-mediated gene editing). [All are in the pX330 (BB901) backbone]

BB #	Description	Vector	Cloning notes
957	pX330 with gRNA to CENP-A C-terminus (LG-033 + LG-034)	pX330 (BB901)	Annealed oligos LG-035 and LG-036 and ligated with pX330 (BB901) that had been digested with FastDigest BpiI (Thermo).
958	pX330 with gRNA to CENP-A C-terminus (LG-035 + LG-036	pX330 (BB901)	Annealed oligos LG-035 and LG-036 and ligated with pX330 (BB901) that had been digested with FastDigest BpiI (Thermo).
963	pX330 with gRNA to CENP-A C-terminus (LG-045 + LG-046)	pX330 (BB901)	Annealed oligos LG-045 and LG-046 and ligated with pX330 (BB901) that had been digested with FastDigest BpiI (Thermo).

Plasmids containing repair templates, for transfection in mammalian cells (for CRISPR/Cas-mediated gene editing):

BB #	**Description**	**Vector**	**Notes on cloning, source, or usage**
965	SNAP-P2A-NeoR	pUC19	Repair template for CENP-A-SNAP CRISPR, containing ~800bp homology arms on either end of the exogenous gene

Plasmids for transfection in mammalian DLD-1 Flp-In T-Rex cells, to insert construct into unique FRT site by Flp/FRT recombination:

BB #	**Description**	**Vector**	**Notes on cloning, source, or usage**
959	Flp recombinase		This is a commercial plasmid from Invitrogen (pOG44).
961	pcDNA5/FRT with CENP-A-SNAP	pcDNA/FRT (BB955)	PCR'ed up BB921 with LG-037 and LG-038 (used KAPA HiFi with GC-rich buffer!); inserted into BB955 backbone by KpnI and PspXI digest.

962	pcDNA5/FRT-CENP-C (untagged, full-length)	pcDNA/FRT (BB955)	PCR'ed up BB891 with LG-039 and LG-040; inserted into BB955 backbone by NheI and PspXI digest.
1007	pcDNA5-FRT-CENP-C(Δ426-537)	pcDNA/FRT (BB955)	PCR BB892 with LG-039 and LG-040. Ligation with BB962 after NheI + PspXI digestion.
1008	pcDNA5-FRT-CENP-C(Δ736-758)	pcDNA/FRT (BB955)	PCR BB903 with LG-039 and LG-040. Ligation with BB962 after NheI + PspXI digestion.
1009	pcDNA5-FRT-CENP-C(Δ426-537 + Δ736-758)	pcDNA/FRT (BB955)	PCR BB904 with LG-039 and LG-040. Ligation with BB962 after NheI + PspXI digestion.
1135	pcDNA5-FRT-CENP-C(Δ519-533)	pcDNA/FRT (BB955)	Mutagenesis of BB962 with LG 104 and LG 105
1136	pcDNA-FRT-CENP-C(R521A)	pcDNA/FRT (BB955)	Mutagenesis of BB962 with LG 106 and LG 107
1137	pcDNA-FRT-CENP-C(R522A)	pcDNA/FRT (BB955)	Mutagenesis of BB962 with LG 108 and LG 109

Plasmids for expression in bacterial cells:

BB #	**Description**	**Vector**	**Notes on cloning, source, or usage**
675	pGEX-GST-CENP-C(426-537)	pGEX	Gift from Aaron Straight (used in Carroll et al., 2010, and Falk et al., 2015)
1015	CENP-N(1-240)-HIS		His-tagged CENP-N(1-240). Made by Kara McKinley (first used in McKinley et al., 2015)
1072	CENP-N(1-205)-His		His-tagged CENP-N(1-205)
1145	GST-CENP-C(R521A)	pGEX	CENP-C^{CD}, R521A
1086	GST-CENP-C(R522A)	pGEX	CENP-C^{CD}, R522A

APPENDIX D: LIST OF OLIGOS

LG#	Sequence	Purpose	Notes
LG-033	CACCGCAACTGGCCCGGAGGATCCG	CENP-A-SNAP CRISPR: gRNA #1	
LG-034	AAACCGGATCCTCCGGGCCAGTTGC	CENP-A-SNAP CRISPR: gRNA #1	
LG-035	CACCGCCGGAGGATCCGGGGCCTTG	CENP-A-SNAP CRISPR: gRNA #2	
LG-036	AAACCAAGGCCCCGGATCCTCCGGC	CENP-A-SNAP CRISPR: gRNA #2	
LG-037	GACTGGTACCGACGTCACCGGTCCGGCCGGATCTATGGG	To insert CENP-A-SNAP from BB921 into pcDNA/FRT (N-terminal KpnI-AgeI sites)	Tm=59C; 71% GC; 17bp annealed
LG-038	GATCATGCTCGAGTCTCAAGCGTAATCTGGAACGTCATATG	To insert CENP-A-SNAP from BB921 into pcDNA/FRT (C-terminal PspXI site)	Tm=60C; 43% GC; 28bp annealed
LG-039	GACTGCTAGCATGGCTGCGTCCGGTCTG	To insert CENP-C from BB891 into pcDNA/FRT (N-terminal NheI site and ATG)	Tm=60-61C, 75% GC; 16 annealed bases
LG-040	CCCTCTAGACTCGAGTCATCTTTTTATCTG	To insert CENP-C from BB891 into pcDNA/FRT (already has C-terminal PspXI; keeping the stop codon)	Tm=59C; 43% GC; 30bp annealed
LG-045	CACCGCTGACAGAAACACTGGGTGC	CENP-A-SNAP CRISPR: gRNA #3	
LG-046	AAACGCACCCAGTGTTTCTGTCAGC	CENP-A-SNAP CRISPR: gRNA #3	
LG-049	CCGGGGCCTTGAAGAGGGACTCGGC	F mutagenesis primer for repair template for gRNA #035-036 (for both BB956 and with NeoR)	Tm=67C
LG-050	GCCGAGTCCCTCTTCAAGGCCCCGG	R mutagenesis primer for repair template for gRNA #035-036 (for both BB956 and with NeoR)	Tm=67C
LG-051	AGATTACGCTTGAGCTCTTGCACCCAGTGTTTCTG	F mutagenesis primer for repair template for gRNA #045-046 (for	Tm=66C

		BB956 ONLY)	
LG-052	CAGAAACACTGGGTGCAAGAGCTCAAGCGTAATCT	R mutagenesis primer for repair template for gRNA #045-046 (for BB956 ONLY)	Tm=64C
LG-053	GAGGGCCTATTTCCCATGATTCC	U6 F promoter (Recommended by Ran et al 2013) for sequencing gRNAs in pX330	Tm=58C, 52% GC
LG-054	GTGAATTCGAGCTCGGTACCCGGGGATCCTCTAGAGGAACTCTCTCGTTTGTCCAC	CENP-A-SNAP CRISPR (with NeoR): F for 5' homology arm gPCR	Tm=57C, 52% GC
LG-055	CCTCGAGAAGGCCGAGTCCCTCCTCAAG	CENP-A-SNAP CRISPR (with NeoR): R for 5' homology arm gPCR	Tm=58C, 67% GC
LG-056	GCCTTGAGGAGGGACTCGG	CENP-A-SNAP CRISPR (with NeoR): F for gBlock	Tm=60C, 68% GC
LG-057	AAGACTGACAGAAACACTGGGTG	CENP-A-SNAP CRISPR (with NeoR): R for gBlock	Tm=58C, 48% GC
LG-058	ATGACTCGAGGCTCCTGCACCCAGTGTT	CENP-A-SNAP CRISPR (with NeoR): F for 3' homology arm gPCR	Tm=59C, 61% GC
LG-059	ACGCCAAGCTTGCATGCCTGCAGGTCGACTAGTGCCTTTTCTCCCATACCACAG	CENP-A-SNAP CRISPR (with NeoR): R for 3' homology arm gPCR	Tm=57C, 52% GC
LG-060	GTGCAACTGGCCCGAAGAATTCGCGGCCTTGAGGAGGG	F mutagenesis primer for repair template for gRNA #033-034	Tm=63C, 63% GC
LG-061	CCCTCCTCAAGGCCGCGAATTCTTCGGGCCAGTTGCAC	R mutagenesis primer for repair template for gRNA #033-034	Tm=67C, 63% GC
LG-076	GTTATGACTCGAGGCTCTTGCACCCAGTGTTTCTG	F mutagenesis primer for repair template for gRNA #045-046 (for the one with NeoR ONLY)	Tm=65C, annealed 34bp (51% GC)
LG-077	CAGAAACACTGGGTGCAAGAGCCTCGAGTCATAAC	R mutagenesis primer for repair template for gRNA #045-046 (for the one with NeoR ONLY)	Tm=64C, annealed 34bp (51% GC)
LG-080	GCTACCTGCCCATTCGACCACCAAGCG	F mutagenesis for L-->F in CRISPR repair template with SNAP-NeoR	Tm-64C
LG-081	CGCTTGGTGGTCGAATGGGCAGGTAGC	R mutagenesis for L-->F in CRISPR repair template with SNAP-NeoR	Tm-64C
LG-104	GTGACTTCAACTGTCACGAAATCAGAGGAGAGTCCT	CENP-C CD mutagenesis: F d519-533	Tm=52C

LG-105	AGGACTCTCCTCTGATTTCGTGACAGTTGAAGTCAC	CENP-C CD mutagenesis: R d519-533	Tm=55C
LG-106	AGAAGTGACTTCAACTGTCACGAAAAGTGCCAGAATTTCCAGGCGT	CENP-C CD mutagenesis: F R521A	Tm=63C
LG-107	ACGCCTGGAAATTCTGGCACTTTTCGTGACAGTTGAAGTCACTTCT	CENP-C CD mutagenesis: R R521A	Tm=63C
LG-108	TCAACTGTCACGAAAAGTCGAGCCATTTCCAGGCGTCCATCTGAT	CENP-C CD mutagenesis: F R522A	Tm=63C
LG-109	ATCAGATGGACGCCTGGAAATGGCTCGACTTTTCGTGACAGTTGA	CENP-C CD mutagenesis: R R522A	Tm=63C
LG-116	CGAGGTCGACGGTATCG	For sequencing (this is the common "KS primer")	Tm=55C, 65% GC
LG-117	GCGTAATCTGGAACGTCATATGG	For sequencing (R primer that anneals to "HA")	Tm=57C, 48% GC
LG-191	GGAGATATACATATGGATGAGACTGTTG	F primer for CENP-N(205)-His gBlock	Tm=55C, 28 annealed bases, 39% GC
LG-192	CCAAGCTTAGATCTGGATCCTC	R primer for CENP-N(205)-His gBlock	Tm=55C, 22 annealed bases, 50% GC

References

Arimura, Y., Shirayama, K., Horikoshi, N., Fujita, R., Taguchi, H., Kagawa, W., Fukagawa, T., Almouzni, G., and Kurumizaka, H. (2014). Crystal structure and stable property of the cancer-associated heterotypic nucleosome containing CENP-A and H3.3. Sci. Rep. *4*, 7115.

Armache, K., Garlick, J.D., Canzio, D., Narlikar, G.J., and Kingston, R.E. (2011). Structural basis of silencing: Sir3 BAH domain in complex with a nucleosome at 3.0 Å resolution. Science *334*, 977–982.

Bai, Y., Milne, J.S., Mayne, L., and Englander, S.W. (1993). Primary structure effects on peptide group hydrogen exchange. Proteins *17*, 75–86.

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

Barbera, A.J., Chodaparambil, J. V, Kelley-Clarke, B., Joukov, V., Walter, J.C., Luger, K., and Kaye, K.M. (2006). The nucleosomal surface as a docking station for Kaposi's sarcoma herpesvirus LANA. Science *311*, 856–861.

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

Basilico, F., Maffini, S., Weir, J.R., Prumbaum, D., Rojas, A.M., Zimniak, T., De Antoni, A., Jeganathan, S., Voss, B., van Gerwen, S., et al. (2014). The pseudo GTPase CENP-M drives human kinetochore assembly. Elife *3*, e02978.

Bassett, E.A., Wood, S., Salimian, K.J., Ajith, S., Foltz, D.R., and Black, B.E. (2010). Epigenetic centromere specification directs aurora B accumulation but is insufficient to efficiently correct mitotic errors. J. Cell Biol. *190*, 177–185.

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

Beadle, G.W. (1932). A Possible Influence of the Spindle Fibre on Crossing-Over in Drosophila. Proc. Natl. Acad. Sci. U. S. A. *18*, 160–165.

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

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

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

Black, B.E., Jansen, L.E.T., Maddox, P.S., Foltz, D.R., Desai, A.B., Shah, J. V, and Cleveland, D.W. (2007b). Centromere identity maintained by nucleosomes assembled with histone H3 containing the CENP-A targeting domain. Mol. Cell *25*, 309–322.

Bloom, K.S., and Carbon, J. (1982). Yeast centromere DNA is in a unique and highly ordered structure in chromosomes and small circular minichromosomes. Cell *29*, 305–317.

Bodor, D.L., Rodríguez, M.G., Moreno, N., and Jansen, L.E.T. (2012). Analysis of protein turnover by quantitative SNAP-based pulse-chase imaging. Curr. Protoc. Cell Biol. *Chapter 8*.

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

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

Buchwitz, B.J., Ahmad, K., Moore, L.L., Roth, M.B., and Henikoff, S. (1999). A histone-H3-like protein in C . elegans. Nature *401*, 547–548.

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

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

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

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

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

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

Cohen, R.L., Espelin, C.W., De Wulf, P., Sorger, P.K., Harrison, S.C., and Simons, K.T. (2008). Structural and functional dissection of Mif2p, a conserved DNA-binding kinetochore protein. Mol. Biol. Cell *19*, 4480–4491.

Conde e Silva, N., Black, B.E., Sivolob, A., Filipski, J., Cleveland, D.W., and Prunell, A. (2007). CENP-A-containing Nucleosomes: Easier Disassembly versus Exclusive Centromeric Localization. J. Mol. Biol. *370*, 555–573.

Crosby, G.A., and Demas, J.N. (1971). Measurement of photoluminescence quantum yields. Review. J. Phys. Chem. *75*, 991–1024.

Cumming, G., Fidler, F., and Vaux, D.L. (2007). Error bars in experimental biology. J. Cell Biol. *177*, 7–11.

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

Darlington, C.D. (1936). The external mechanics of the chromosomes. Proc. R. Soc. Lond. B *121*, 264–273.

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

Dunleavy, E.M., Almouzni, G., and Karpen, G.H. (2011). H3.3 is deposited at centromeres in S phase as a placeholder for newly assembled CENP-A in G_1 phase. Nucleus *2*, 146–157.

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

Earnshaw, W.C., and Migeon, B.R. (1985). Three related centromere proteins are absent from the inactive centromere of a stable isodicentric chromosome. Chromosoma *92*, 290–296.

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

Englander, S.W. (2006). Hydrogen exchange and mass spectrometry: A historical perspective. J. Am. Soc. Mass Spectrom. *17*, 1481–1489.

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

Fachinetti, D., Diego Folco, H., Nechemia-Arbely, Y., Valente, L.P., Nguyen, K., Wong,

A.J., Zhu, Q., Holland, A.J., Desai, A., Jansen, L.E.T., et al. (2013). A two-step mechanism for epigenetic specification of centromere identity and function. Nat. Cell Biol. *15*, 1056–1066.

Fachinetti, D., Han, J.S., McMahon, M.A., Ly, P., Abdullah, A., Wong, A.J., and Cleveland, D.W. (2015). DNA sequence-specific binding of CENP-B enhances the fidelity of human centromere function. Dev. Cell *33*, 314–327.

Falk, S.J., and Black, B.E. (2012). Biochimica et Biophysica Acta Centromeric chromatin and the pathway that drives its propagation. Biochim. Biophys. Acta *1819*, 313–321.

Falk, S.J., Guo, L.Y., Sekulic, N., Smoak, E.M., Mani, T., Logsdon, G.A., Gupta, K., Jansen, L.E.T., Van Duyne, G.D., Vinogradov, S.A., et al. (2015). CENP-C reshapes and stabilizes CENP-A nucleosomes at the centromere. Science *348*, 699–704.

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

Fang, J., Liu, Y., Wei, Y., Deng, W., Yu, Z., Huang, L., Teng, Y., Yao, T., You, Q., Ruan, H., et al. (2015). Structural transitions of centromeric chromatin regulate the cell cycle-dependent recruitment of CENP-N. Genes Dev. *29*, 1058–1073.

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

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

Forster, T. (1946). Energiewanderung und Fluoreszenz. Naturwissenschaften *33*, 166–175.

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

Fukagawa, T., and Brown, W.R. (1997). Efficient conditional mutation of the vertebrate CENP-C gene. Hum. Mol. Genet. *6*, 2301–2308.

Furuyama, T., Codomo, C.A., and Henikoff, S. (2013). Reconstitution of hemisomes on budding yeast centromeric DNA. Nucleic Acids Res. *41*, 5769–5783.

Gascoigne, K.E., Takeuchi, K., Suzuki, A., Hori, T., Fukagawa, T., and Cheeseman, I.M. (2011a). Induced ectopic kinetochore assembly bypasses the requirement for CENP-A nucleosomes. Cell *145*, 410–422.

Gascoigne, K.E., Takeuchi, K., Suzuki, A., Hori, T., Fukagawa, T., and Cheeseman, I.M. (2011b). Induced ectopic kinetochore assembly bypasses the requirement for CENP-A nucleosomes. Cell *145*, 410–422.

Guerra, M., Cabral, G., Cuacos, M., Gonz??lez-Garc??a, M., Gonz??lez-S??nchez, M., Vega, J., and Puertas, M.J. (2010). Neocentrics and holokinetics (holocentrics): Chromosomes out of the centromeric rules. Cytogenet. Genome Res. *129*, 82–96.

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

Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., and Witten, I. (2009). The WEKA data mining software: An update. SIGKDD Explor. *11*, 10–18.

Harp, J.M., Uberbacher, E.C., Roberson, A.E., Palmer, E.L., Gewiess, A., and Bunick, G.J. (1996). X-ray Diffraction Analysis of Crystals Containing Twofold Symmetric Nucleosome Core Particles. Acta Crystallogr. Sect. D, Biol. Crystallogr. *52*, 283–288.

Hassold, T., and Hunt, P. (2001). To err (meiotically) is human: the genesis of human aneuploidy. Nat. Rev. Genet. *2*, 280–291.

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

Hayes, J.J., Tullius, T.D., and Wolffe, a P. (1990). The structure of DNA in a nucleosome. Proc. Natl. Acad. Sci. U. S. A. *87*, 7405–7409.

Hellwig, D., Emmerth, S., Ulbricht, T., Döring, V., Hoischen, C., Martin, R., Samora, C.P., McAinsh, A.D., Carroll, C.W., Straight, A.F., et al. (2011). Dynamics of CENP-N kinetochore binding during the cell cycle. J. Cell Sci. *124*, 3871–3883.

Hemmerich, P., Weidtkamp-Peters, S., Hoischen, C., Schmiedeberg, L., Erliandri, I., and Diekmann, S. (2008). Dynamics of inner kinetochore assembly and maintenance in living cells. J. Cell Biol. *180*, 1101–1114.

Henikoff, S., Ahmad, K., Platero, J.S., and van Steensel, B. (2000). Heterochromatic deposition of centromeric histone H3-like proteins. Proc. Natl. Acad. Sci. U. S. A. *97*, 716–721.

Henikoff, S., Ramachandran, S., Krassovsky, K., Bryson, T.D., Codomo, C.A., Brogaard, K., Widom, J., Wang, J.P., and Henikoff, J.G. (2014). The budding yeast centromere DNA element II wraps a stable Cse4 hemisome in either orientation in vivo. Elife *2014*, 1–23.

Heun, P., Erhardt, S., Blower, M.D., Weiss, S., Skora, A.D., and Karpen, G.H. (2006). Mislocalization of the drosophila centromere-specific histone CID promotes formation of functional ectopic kinetochores. Dev. Cell *10*, 303–315.

Hoffmann, S., Dumont, M., Barra, V., Ly, P., Nechemia-Arbely, Y., McMahon, M.A., Hervé, S., Cleveland, D.W., and Fachinetti, D. (2016). CENP-A is dispensable for mitotic centromere function after initial centromere/kinetochore assembly. Cell Rep. *17*, 2394–2404.

Holland, A.J., Fachinetti, D., Han, J.S., and Cleveland, D.W. (2012). Inducible, reversible system for the rapid and complete degradation of proteins in mammalian cells. Proc. Natl. Acad. Sci. U. S. A. *109*, E3350–E3357.

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

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

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.

Hudson, D.F., Fowler, K.J., Earle, E., Saffery, R., Kalitsis, P., Trowell, H., Hill, J., Wreford, N.G., De Kretser, D.M., Cancilla, M.R., et al. (1998). Centromere protein B null mice are mitotically and meiotically normal but have lower body and testis weights. J. Cell Biol. *141*, 309–319.

Hughes-Schrader, S., and Ris, H. (1941). The diffuse spindle attachment of coccids, verified by the mitotic behavior of induced chromosome fragments. J. Exp. Zool. *87*, 429–456.

Ishii, K., Ogiyama, Y., Chikashige, Y., Soejima, S., Masuda, F., Kakuma, T., Hiraoka, Y., and Takahashi, K. (2008). Heterochromatin integrity affects chromosome reorganization after centromere dysfunction. Science *321*, 1088–1091.

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

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

Kan, Z.-Y., Mayne, L., Chetty, P.S., and Englander, S.W. (2011). ExMS: data analysis for HX-MS experiments. J. Am. Soc. Mass Spectrom. *22*, 1906–1915.

Kassianidis, E., Pearson, R.J., and Philp, D. (2006). Probing structural effects on replication efficiency through comparative analyses of families of potential self-replicators. Chem. - A Eur. J. *12*, 8798–8812.

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

Khandelia, P., Yap, K., and Makeyev, E. V (2011). Streamlined platform for short hairpin RNA interference and transgenesis in cultured mammalian cells. PNAS *108*, 12799–12804.

Kim, I.S., Lee, M., Park, K.C., Jeon, Y., Park, J.H., Hwang, E.J., Jeon, T.I., Ko, S., Lee, H., Baek, S.H., et al. (2012). Roles of Mis18α in epigenetic regulation of centromeric chromatin and CENP-A loading. Mol. Cell *46*, 260–273.

Kim, J.H., Lee, S.-R., Li, L.-H., Park, H.-J., Park, J.-H., Lee, K.Y., Kim, M.-K., Shin, B.A., and Choi, S.-Y. (2011). High cleavage efficiency of a 2A peptide derived from porcine teschovirus-1 in human cell lines, zebrafish and mice. PLoS One *6*, e18556.

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

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

Kline, S.R. (2006). Reduction and analysis of SANS and USANS data using IGOR Pro. J. Appl. Crystallogr. *39*, 895–900.

Kops, G.J., Weaver, B.A., and Cleveland, D.W. (2005). On the road to cancer: aneuploidy and the mitotic checkpoint. Nat. Rev. Cancer *5*, 773–785.

Kornberg, R.D. (1977). Structure of chromatin. Annu. Rev. Biochem. *46*, 931–954.

Kubin, R.F., and Fletcher, A.N. (1982). Fluorescence quantum yields of some rhodamine dyes. J. Lumin. *27*, 455–462.

Lacoste, N., Woolfe, A., Tachiwana, H., Garea, A.V., Barth, T., Cantaloube, S., Kurumizaka, H., Imhof, A., and Almouzni, G. (2014). Mislocalization of the centromeric histone variant CenH3/CENP-A in human cells depends on the chaperone DAXX. Mol. Cell *53*, 631–644.

Lakowicz, J.R. (2006). Principles of Fluorescence Spectroscopy (Boston, MA: Springer US).

Lanini, L., and McKeon, F. (1995). Domains required for CENP-C assembly at the kinetochore. Mol. Biol. Cell *6*, 1049–1059.

Lo, A.W.I., Craig, J.M., Saffery, R., Kalitsis, P., Irvine, D. V., Earle, E., Magliano, D.J., and Choo, K.H.A. (2001). A 330 kb CENP-A binding domain and altered replication timing at a human neocentromere. EMBO J. *20*, 2087–2096.

Logsdon, G.A., Barrey, E.J., Bassett, E. a, DeNizio, J.E., Guo, L.Y., Panchenko, T., Dawicki-McKenna, J.M., Heun, P., and Black, B.E. (2015). Both tails and the centromere targeting domain of CENP-A are required for centromere establishment. J. Cell Biol. *208*, 521–531.

Lorenz, M., Hillisch, a, Goodman, S.D., and Diekmann, S. (1999). Global structure similarities of intact and nicked DNA complexed with IHF measured in solution by fluorescence resonance energy transfer. Nucleic Acids Res. *27*, 4619–4625.

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

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

Luger, K., Rechsteiner, T.J., and Richmond, T.J. (1999). Expression and purification of recombinant histones and nucleosome reconstitution. Methods Mol. Biol. *119*, 1–16.

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

Makde, R.D., England, J.R., Yennawar, H.P., and Tan, S. (2010). Structure of RCC1 chromatin factor bound to the nucleosome core particle. Nature *467*, 562–566.

Mann, M. (2006). Functional and quantitative proteomics using SILAC. Nat. Rev. Mol. Cell Biol. *7*, 952–958.

Masumoto, H., Masukata, H., Muro, Y., Nozaki, N., and Okazaki, T. (1989). A human centromere antigen (CENP-B) interacts with a short specific sequence in alphoid DNA, a human centromeric satellite. J. Cell Biol. *109*, 1963–1973.

McGinty, R.K., and Tan, S. (2016). Recognition of the nucleosome by chromatin factors and enzymes. Curr. Opin. Struct. Biol. *37*, 54–61.

McGinty, R.K., Henrici, R.C., and Tan, S. (2014). Crystal structure of the PRC1 ubiquitylation module bound to the nucleosome. Nature *514*, 591–596.

McKinley, K.L., and Cheeseman, I.M. (2014). Polo-like Kinase 1 Licenses CENP-A Deposition at Centromeres. Cell *158*, 397–411.

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

McKinley, K.L., Sekulic, N., Guo, L.Y., Tsinman, T., Black, B.E., and Cheeseman, I.M. (2015). The CENP-L-N complex forms a critical node in an integrated meshwork of interactions at the centromere-kinetochore interface. Mol. Cell *60*, 886–898.

Mendiburo, M.J., Padeken, J., Fülöp, S., Schepers, A., and Heun, P. (2011). Drosophila CENH3 is sufficient for centromere formation. Science *334*, 686–690.

Milks, K.J., Moree, B., and Straight, A.F. (2009). Dissection of CENP-C-directed centromere and kinetochore assembly. Mol. Biol. Cell *20*, 4246–4255.

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

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

Morgan, M.T., Haj-Yahya, M., Ringel, A.E., Bandi, P., Brik, A., and Wolberger, C. (2016). Structural basis for histone H2B deubiquitination by the SAGA DUB module. Science *351*, 725–728.

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

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

Nardi, I.K., Zasadzińska, E., Stellfox, M.E., Knippler, C.M., and Foltz, D.R. (2016). Licensing of Centromeric Chromatin Assembly through the Mis18α-Mis18β Heterotetramer. Mol. Cell *61*, 774–787.

Nashun, B., Hill, P.W.S., Smallwood, S.A., Dharmalingam, G., Amouroux, R., Clark, S.J., Sharma, V., Ndjetehe, E., Pelczar, P., Festenstein, R.J., et al. (2015). Continuous histone replacement by Hira is essential for normal transcriptional regulation and de novo DNA methylation during mouse oogenesis. Mol. Cell *60*, 611–625.

Nguyen, T., and Francis, M.B. (2003). Practical synthetic route to functionalized rhodamine dyes. Org. Lett. *5*, 3245–3248.

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

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

O'Connor, C., and Miko, I. (2008). Developing the chromosome theory. Nat. Educcation *1*, 44.

Okada, M., Cheeseman, I.M., Hori, T., Okawa, K., McLeod, I.X., Yates, J.R., Desai, A., and

Fukagawa, T. (2006). The CENP-H-I complex is required for the efficient incorporation of newly synthesized CENP-A into centromeres. Nat. Cell Biol. *8*, 446–457.

Okada, T., Ohzeki, J., Nakano, M., Yoda, K., Brinkley, W.R., Larionov, V., and Masumoto, H. (2007). CENP-B controls centromere formation depending on the chromatin context. Cell *131*, 1287–1300.

Padeganeh, A., Ryan, J., Boisvert, J., Ladouceur, A.-M., Dorn, J.F., and Maddox, P.S. (2013). Octameric CENP-A nucleosomes are present at human centromeres throughout the cell cycle. Curr. Biol. *23*, 764–769.

Palmer, D.K., and Margolis, R.L. (1985). Kinetochore components recognized by human autoantibodies are present on mononucleosomes. Mol. Cell. Biol. *5*, 173–186.

Palmer, D.K., O'Day, K., Wener, M.H., Andrews, B.S., and Margolis, R.L. (1987). A 17-kD centromere protein (CENP-A) copurifies with nucleosome core particles and with histones. J. Cell Biol. *104*, 805–815.

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

Panchenko, T., Sorensen, T.C., Woodcock, C.L., Kan, Z., Wood, S., Resch, M.G., Luger, K., Englander, S.W., Hansen, J.C., and Black, B.E. (2011). Replacement of histone H3 with CENP-A directs global nucleosome array condensation and loosening of nucleosome superhelical termini. Proc. Natl. Acad. Sci. U. S. A. *108*, 16588–16593.

Pelleg, D., and Moore, A. (2000). Proceedings of the 17th International Conf. on Machine Learning (San Francisco).

Perez-Castro, a V, Shamanski, F.L., Meneses, J.J., Lovato, T.L., Vogel, K.G., Moyzis, R.K., and Pedersen, R. (1998). Centromeric protein B null mice are viable with no apparent abnormalities. Dev. Biol. *201*, 135–143.

Perpelescu, M., Hori, T., Toyoda, A., Misu, S., Monma, N., Ikeo, K., Obuse, C., Fujiyama, A., and Fukagawa, T. (2015). HJURP is involved in the expansion of centromeric chromatin. Mol. Biol. Cell *26*, 2742–2754.

Probst, A. V, Dunleavy, E., and Almouzni, G. (2009). Epigenetic inheritance during the cell cycle. Nat Rev Mol Cell Biol *10*, 192–206.

Przewloka, M.R., Venkei, Z., Bolanos-Garcia, V.M., Debski, J., Dadlez, M., and Glover, D.M. (2011). CENP-C is a structural platform for kinetochore assembly. Curr. Biol. *21*, 399–405.

Rago, F., Gascoigne, K.E., and Cheeseman, I.M. (2015). Distinct organization and regulation of the outer kinetochore KMN network downstream of CENP-C and CENP-T. Curr. Biol. *25*, 671–677.

Rajagopalan, H., and Lengauer, C. (2004). Aneuploidy and cancer. Nature *432*, 338–341.

Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. Nat. Protoc. *8*, 2281–2308.

Richter, M., Chakrabarti, A., Ruttekolk, I.R., Wiesner, B., Beyermann, M., Brock, R., and Rademann, J. (2012). Multivalent design of apoptosis-inducing bid-BH3 peptide-oligosaccharides boosts the intracellular activity at identical overall peptide concentrations. Chemistry *18*, 16708–16715.

Roulland, Y., Ouararhni, K., Naidenov, M., Ramos, L., Shuaib, M., Syed, S.H., Lone, I.N., Boopathi, R., Fontaine, E., Papai, G., et al. (2016). The flexible ends of CENP-A nucleosome are required for mitotic fidelity. Mol. Cell *63*, 674–685.

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

Samejima, I., Spanos, C., Alves, F.D.L., Hori, T., Perpelescu, M., Zou, J., Rappsilber, J., Fukagawa, T., and Earnshaw, W.C. (2015). Whole-proteome genetic analysis of dependencies in assembly of a vertebrate kinetochore. J. Cell Biol. *211*, 1141–1156.

Saunders, M., Fitzgerald-Hayes, M., and Bloom, K. (1988). Chromatin structure of altered yeast centromeres. Proc. Natl. Acad. Sci. U. S. A. *85*, 175–179.

Schneider, C. a, Rasband, W.S., and Eliceiri, K.W. (2012). NIH Image to ImageJ: 25 years of image analysis. Nat. Methods *9*, 671–675.

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

Schwartz, D.C., and Cantor, C.R. (1984). Separation of yeast chromosome-sized DNAs by pulsed field gradient gel electrophoresis. Cell *37*, 67–75.

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

Sekulic, N., and Black, B.E. (2009). A reader for centromeric chromatin. Nat. Cell Biol. *11*, 793–795.

Sekulic, N., and Black, B.E. (2012). Molecular underpinnings of centromere identity and maintenance. Trends Biochem. Sci. *37*, 220–229.

Sekulic, N., and Black, B.E. (2016a). Preparation of recombinant centromeric nucleosomes and formation of complexes with nonhistone centromere proteins. Methods Enzymol. *573*, 67–96.

Sekulic, N., and Black, B.E. (2016b). Preparation of recombinant centromeric nucleosomes and formation of complexes with nonhistone centromere proteins. Methods Enzymol. Online ahead of Print.

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

Shelby, R.D., Monier, K., and Sullivan, K.F. (2000). Chromatin assembly at kinetochores is uncoupled from DNA replication. J. Cell Biol. *151*, 1113–1118.

Sheltzer, J.M., Blank, H.M., Pfau, S.J., Tange, Y., George, B.M., Humpton, T.J., Brito, I.L., Hiraoka, Y., Niwa, O., and Amon, A. (2011). Aneuploidy drives genomic instability in yeast. Science *333*, 1026–1030.

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

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

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

Smoak, E.M., Stein, P., Schultz, R.M., Lampson, M.A., and Black, B.E. (2016). Long-term retention of CENP-A nucleosomes in mammalian oocytes underpins transgenerational inheritance of centromere identity. Curr. Biol. *26*, 1110–1116.

Solomon, D.A., Kim, T., Diaz-Martinez, L.A., Fair, J., Elkahloun, A.G., Harris, B.T., Toretsky, J.A., Rosenberg, S.A., Shukla, N., Ladanyi, M., et al. (2011). Mutational inactivation of STAG2 causes aneuploidy in human cancer. Science *333*, 1039–1043.

Song, K., Gronemeyer, B., Lu, W., Eugster, E., and Tomkiel, J.E. (2002). Mutational analysis of the central centromere targeting domain of human centromere protein C, (CENP-C). Exp. Cell Res. *275*, 81–91.

Steiner, N.C., and Clarke, L. (1994). A novel epigenetic effect can alter centromere function in fission yeast. Cell *79*, 865–874.

Stellfox, M.E., Bailey, A.O., and Foltz, D.R. (2013). Putting CENP-A in its place. Cell. Mol. Life Sci. *70*, 387–406.

Stellfox, M.E., Nardi, I.K., Knippler, C.M., and Foltz, D.R. (2016). Differential Binding Partners of the Mis18α/β YIPPEE Domains Regulate Mis18 Complex Recruitment to Centromeres. Cell Rep. *15*, 2127–2135.

Stoler, S., Keith, K.C., Kurnick, K.F.E., and 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., and Baker, R.E. (2007). Scm3, an essential Saccharomyces cerevisiae centromere protein required for G2/M progression and Cse4 localization. Proc. Natl. Acad. Sci. U. S. A. *104*, 10571–10576.

Stryer, L., and Haugland, R.P. (1967). Energy transfer: a spectroscopic ruler. Proc. Natl. Acad. Sci. U. S. A. *58*, 719–726.

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

Sullivan, K.F., Hechenberger, M., and Masri, K. (1994). Human CENP-A contains a histone H3 related histone fold domain that is required for targeting to the centromere. J. Cell Biol. *127*, 581–592.

Sutton, W.S. (1902). On the morphology of the chromosome group in Brachystola magna. Biol. Bull. *14*, 24–39.

Sutton, W.S. (1903). The Chromosomes in Heredity. Biol. Bull. *4*, 231.

Syed, S.H., Goutte-Gattat, D., Becker, N., Meyer, S., Shukla, M.S., Hayes, J.J., Everaers, R., Angelov, D., Bednar, J., and Dimitrov, S. (2010). Single-base resolution mapping of H1-nucleosome interactions and 3D organization of the nucleosome. Proc. Natl. Acad. Sci. U. S. A. *107*, 9620–9625.

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

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

Takeuchi, K., Nishino, T., Mayanagi, K., Horikoshi, N., Osakabe, A., Tachiwana, H., Hori, T., Kurumizaka, H., and Fukagawa, T. (2014). The centromeric nucleosome-like CENP-T-W-S-X complex induces positive supercoils into DNA. Nucleic Acids Res. *42*, 1644–1655.

Tanaka, Y., Nureki, O., Kurumizaka, H., Fukai, S., Kawaguchi, S., Ikuta, M., Iwahara, J., Okazaki, T., and Yokoyama, S. (2001). Crystal structure of the CENP-B protein-DNA complex: The DNA-binding domains of CENP-B induce kinks in the CENP-B box DNA. EMBO J. *20*, 6612–6618.

Tanaka, Y., Tachiwana, H., Yoda, K., Masumoto, H., Okazaki, T., Kurumizaka, H., and Yokoyama, S. (2005). Human centromere protein B induces translational positioning of nucleosomes on alpha-satellite sequences. J. Biol. Chem. *280*, 41609–41618.

Tomkiel, J., Cooke, C.A., Saitoh, H., Bernat, R.L., and Earnshaw, W.C. (1994). CENP-C is required for maintaining proper kinetochore size and for a timely transition to anaphase. J. Cell Biol. *125*, 531–545.

Trazzi, S., Bernardoni, R., Diolaiti, D., Politi, V., Earnshaw, W.C., Perini, G., and Della Valle, G. (2002). In vivo functional dissection of human inner kinetochore protein CENP-C. J. Struct. Biol. *140*, 39–48.

Trazzi, S., Perini, G., Bernardoni, R., Zoli, M., Reese, J.C., Musacchio, A., and Della Valle, G. (2009). The C-terminal domain of CENP-C displays multiple and critical functions for mammalian centromere formation. PLoS One *4*, e5832.

Tullius, T.D. (1988). DNA footprinting with hydroxyl radical. Nature *332*, 663–664.

Voullaire, L.E., Slater, H.R., Petrovic, V., and Choo, K.H. (1993). A functional marker centromere with no detectable alpha-satellite, satellite III, or CENP-B protein: activation of a latent centromere? Am. J. Hum. Genet. *52*, 1153–1163.

Walters, B.T., Ricciuti, A., Mayne, L., and Englander, S.W. (2012). Minimizing back exchange in the hydrogen exchange-mass spectrometry experiment. J. Am. Soc. Mass Spectrom. *23*, 2132–2139.

Wan, X., O'Quinn, R.P., Pierce, H.L., Joglekar, A.P., Gall, W.E., DeLuca, J.G., Carroll, C.W., Liu, S.T., Yen, T.J., McEwen, B.F., 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 Mis18β interacts with and specifies the centromeric assembly of molecular chaperone holliday junction recognition protein (HJURP). J. Biol. Chem. *289*, 8326–8336.

Wang, L.-H., Li, D.-Q., Fu, Y., Wang, H.-P., Zhang, J.-F., Yuan, Z.-F., Sun, R.-X., Zeng, R., He, S.-M., and Gao, W. (2007). pFind 2.0: a software package for peptide and protein identification via tandem mass spectrometry. Rapid Commun. Mass Spectrom. *21*, 2985–2991.

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

Westermann, S., and Schleiffer, A. (2013). Family matters: structural and functional conservation of centromere-associated proteins from yeast to humans. Trends Cell Biol. *23*, 260–269.

Westhorpe, F.G., and Straight, A.F. (2015). The centromere: epigenetic control of chromosome segregation during mitosis. Cold Spring Harb. Perspect. Biol. *7*, 1–26.

Whitten, A.E., Cai, S., and Trewhella, J. (2008). MULCh: Modules for the analysis of small-angle neutron contrast variation data from biomolecular assemblies. J. Appl. Crystallogr. *41*, 222–226.

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

Yang, C.H., Tomkiel, J., Saitoh, H., Johnson, D.H., and Earnshaw, W.C. (1996). Identification of overlapping DNA-binding and centromere-targeting domains in the human kinetochore protein CENP-C. Mol. Cell. Biol. *16*, 3576–3586.

Yang, T.P., Hansen, S.K., Oishi, K.K., Ryder, O.A., and Hamkalo, B.A. (1982). Characterization of a cloned repetitive DNA sequence concentrated on the human X chromosome. Proc. Natl. Acad. Sci. U. S. A. *79*, 6593–6597.

Ye, A.A., Cane, S., and Maresca, T.J. (2016). Chromosome biorientation produces hundreds of piconewtons at a metazoan kinetochore. Nat. Commun. *7*, 13221.

Yoda, K., Ando, S., Okuda, A., Kikuchi, A., and Okazaki, T. (1998). In vitro assembly of the CENP-B/alpha-satellite DNA/core histone complex: CENP-B causes nucleosome positioning. Genes Cells *3*, 533–548.

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