

Localization of *Drosophila* CENP-A to non-centromeric sites depends on the NuRD complex

Engin Demirdizen, Matthias Spiller-Becker, Arion Förtsch, Alexander Wilhelm, Samuel Corless, Debora Bade, Andrea Bergner, Bernd Hessling and Sylvia Erhardt¹*

ZMBH, DKFZ-ZMBH-Alliance and CellNetworks - Cluster of Excellence, University of Heidelberg, Im Neuenheimer Feld 282, 69120 Heidelberg, Germany

Received February 21, 2019; Revised September 12, 2019; Editorial Decision October 09, 2019; Accepted October 24, 2019

ABSTRACT

Centromere function requires the presence of the histone H3 variant CENP-A in most eukaryotes. The precise localization and protein amount of CENP-A are crucial for correct chromosome segregation, and misregulation can lead to aneuploidy. To characterize the loading of CENP-A to non-centromeric chromatin, we utilized different truncation- and localization-deficient CENP-A mutant constructs in *Drosophila melanogaster* cultured cells, and show that the N-terminus of *Drosophila melanogaster* CENP-A is required for nuclear localization and protein stability, and that CENP-A associated proteins, rather than CENP-A itself, determine its localization. Co-expression of mutant CENP-A with its loading factor CAL1 leads to exclusive centromere loading of CENP-A whereas co-expression with the histone-binding protein RbAp48 leads to exclusive non-centromeric CENP-A incorporation. Mass spectrometry analysis of non-centromeric CENP-A interacting partners identified the RbAp48-containing NuRD chromatin remodeling complex. Further analysis confirmed that NuRD is required for ectopic CENP-A incorporation, and RbAp48 and MTA1-like subunits of NuRD together with the N-terminal tail of CENP-A mediate the interaction. In summary, our data show that *Drosophila* CENP-A has no intrinsic specificity for centromeric chromatin and utilizes separate loading mechanisms for its incorporation into centromeric and ectopic sites. This suggests that the specific association and availability of CENP-A inter-

acting factors are the major determinants of CENP-A loading specificity.

INTRODUCTION

Centromeres act as the assembly site of the multi-subunit kinetochore complex and spindle microtubules, facilitating chromosome segregation during mitosis and meiosis (1–3). In most eukaryotes, centromeres are identified by the H3 variant CENP-A, also known as CID in *Drosophila melanogaster*, which replaces H3 in a subset of centromeric nucleosomes (1,4,5). CENP-A overexpression and stable misincorporation can create neocentromeres, defined as active centromeres located at an ectopic site (6,7). Error correction mechanisms remove ectopic CENP-A during S-phase to prevent neocentromere formation in mammalian cells (8), a process that is regulated by ubiquitin-mediated proteolysis in yeasts and flies (9–13). If uncorrected, ectopic CENP-A could lead to abnormal kinetochore formation and segregation defects, ultimately contributing to genome instability (14).

Overexpressed and mislocalized CENP-A is proposed to be a prognostic and predictive marker for tumor malignancy (15–17), however, little is known about CENP-A misincorporation mechanism(s). In human cells, overexpression of CENP-A leads to the formation of H3.3-CENP-A heterotypic nucleosomes that require DAXX, an H3.3-specific chaperone, for chromatin incorporation (18,19). Importantly, tethering of the CENP-A-specific histone chaperone HJURP is sufficient for ectopic CENP-A localization and kinetochore assembly (20), and it has also been suggested that HJURP is the limiting factor for centromeric restriction of CENP-A (16,21). CAF-1, an H3.1 and H3.3-specific chaperone in *Saccharomyces cerevisiae*, and HIRA, an H3.3-specific chaperone, were also implicated in mislocalization of CENP-A in human colon can-

*To whom correspondence should be addressed. Tel: +49 6221 54 6898; Fax: +49 6221 54 5892; Email: s.erhardt@zmbh.uni-heidelberg.de
Present addresses:

Matthias Spiller-Becker, Active Motif, Inc. 1914 Palomar Oaks Way, Suite 150 Carlsbad, CA 92008, USA.

Andrea Bergner, Max Planck Institute for Medical Research, Jahnstraße 29, 69120 Heidelberg, Germany.

Alexander Wilhelm, Johann Wolfgang Goethe-Universität Frankfurt am Main, Max-von-Laue-Str. 9, 60323 Frankfurt, Germany.

Bernd Hessling, DKFZ, Im Neuenheimer Feld 580, 69120 Heidelberg, Germany.

Debora Bade, Expert Medical Events B.V., Biltstraat 106, 3572 BJ Utrecht, the Netherlands.

cer cells (22,23). It is yet unclear if other loading factors or complexes are also involved in mislocalizing CENP-A or which factors act in other organisms. In *Drosophila*, CAL1 has been identified as the CENP-A-specific loading factor (24,25).

Many chromatin remodeling complexes exist in eukaryotes, which use ATP-dependent DNA translocation activity to bind and move nucleosomes. Additional subunits in these complexes mediate the specific mechanism of chromatin remodeling (26). Some of the subunits can be found in several different complexes, for instance RbAp46/48 is a component of the ISWI family class remodeling complex NURF and also in the CHD family remodeler NuRD (27). RbAp46/48 is thought to function by interacting directly with histones, especially as a chaperone of H3-H4 at newly replicated DNA within the CAF-1 complex (28). RbAp48 has been reported to interact with CENP-A-containing chromatin in *Drosophila* cultured cells (29,30). A RbAp48-containing HAT1 complex mediates acetylation of histone H4 K5 and 12, which is important for CENP-A deposition to centromeric chromatin in human cells (31). This is the same acetyltransferase that is required for *Drosophila* CENP-A loading to centromeres (29).

CENP-A targeting to centromeres is functionally conserved and requires the CENP-A-targeting domain (CATD) within the histone fold domain (HFD) of CENP-A (32–35). CATD is required for CENP-A chaperone specificity, centromeric targeting, and altered nucleosome structure (32,33) reviewed in (36). The function of the N-terminal histone tail of CENP-A is less understood. CENP-A N-terminal tails are extremely divergent in length and sequence: The human CENP-A N-terminus is relatively short whereas the N-termini of, for example, *Drosophila*, *Arabidopsis*, and budding yeast CENP-A proteins are relatively long, 125 amino acids in the case of *Drosophila melanogaster* (34). So far only the tail of yeast Cse4 has been shown to be essential (37). Overexpression of tail-truncated mammalian CENP-A suggested a role of the tail in the recruitment of some kinetochore proteins, and the spindle assembly checkpoint kinase BubR1 was shown to be recruited to the tail of *Drosophila* CENP-A when tethered ectopically (38,39). Furthermore, EGFP-tagged deletion constructs revealed a function of the *Arabidopsis* HTR12 tail in meiosis (40). Previous work also showed that loss of the CENP-A tail has no major impact on centromere localization (11,35). However, whether the CENP-A N-terminal region has additional regulatory roles, for instance in the regulation of non-centromeric CENP-A is unclear. Importantly, there seems to be a base level of CENP-A outside of centromeres at least during some cell cycle stages (8,41,42) but little is known about how non-centromeric loading of CENP-A takes place, is prevented, reversed, or kept to a minimum. We, therefore, decided to revisit the functional significance of the CENP-A N-terminus in the context of CENP-A mislocalization in order to address potential functions for the tail at both centromeric and non-centromeric sites.

Here, we report that tail-truncated CENP-A (Δ N-CENP-A) cannot localize efficiently to centromeres, but accumulates in the cytoplasm and is targeted by the proteasome for degradation, suggesting that the CENP-A tail is

critical for regulating protein localization to the nucleus and, thereby, stability. Centromeric localization of Δ N-CENP-A, however, was achieved by overexpression of the CENP-A-specific chaperone CAL1 (25). Moreover, overexpression of the histone-binding factor RbAp48 caused the incorporation of CENP-A to non-centromeric chromatin without any preferences for centromeric regions. This localization of CENP-A to non-centromeric chromatin requires the RbAp48-containing chromatin remodeling complex NuRD that interacts with the N-terminal tail of CENP-A via its MTA1-like subunit. Our data suggest that CENP-A-associated proteins influence where CENP-A is loaded in chromatin and that non-centromeric loading is facilitated by a NuRD complex-dependent alternative mechanism.

MATERIALS AND METHODS

DNA constructs, cell lines and RNA interference

All methods concerning standard molecular biology techniques were essentially performed as described in (Sambrook and W Russell, 2001). Most proteins analyzed in this study were expressed from the CuSO₄-inducible pMT-V5-His vector (Life technologies). A detailed description of the vector cloning system can be found on the Life technologies/Invitrogen website (DES-TOPO TA expression kit). Alternatively, the cloning was also performed by restriction enzyme digestion approach. RbAp48^{MUT} construct (57) was kindly gifted by Prof. Ernest Laue and Dr. Wei Zhang and cloned into pMT vector. Mi-2 wild type and ATPase catalytic domain loss of function mutants (H1151R, R1161Q), ATPase motor brace region gain of function mutant (H1196Y) and CHD domain loss of function mutant (R552Q) were kindly gifted by Prof. Alexander Brehm (58). They were cloned into pMT vector. *Drosophila* S2 cells stably expressing pMT- and pLAP-constructs were obtained by Cellfectin II mediated transfection (life technologies™) using pHygromycin as a selection marker. The proteasome was inhibited with 20 μ M MG132 for 6–24 h. Induction of gene expression from the pMT promoter was induced with 0.05 (low) or 1 mM (high) CuSO₄ for 16–24 h. All DNA vector constructs and primer sequences used in this study are summarized in Supplemental Table S3.

Double-stranded RNA (dsRNA) was made using the Ambion MEGAscript kit according to the manufacturer's protocol. 1.5–2 million cells were cultivated in 6-well plates for 24 h and washed twice with serum-free medium. Then, 20–25 μ g of dsRNA was incubated with the cells for 1 h in serum-free medium. After increasing the final serum concentration to 10% by fresh medium, cells were incubated with dsRNA for 4 days. On day 4, knockdown cells were treated and analyzed.

Yeast two hybrid

The yeast two hybrid (Y2H) assay was performed as described previously (80). Briefly, competent cells of the yeast strain SGY37VIII, which carry the LacZ gene under the control of a LexA-operator, were prepared as described previously (81). 0.1–0.5 μ g of pMM5 constructs (fusion with LexA-DNA binding domain) and/or pMM6 constructs

(fusion with Gal4 transcription-activating domain) were transformed into 15 μ l of competent SGY37VIII cells. 100 μ l of LiPEG was added and cells were vortexed and incubated for 20 min at RT. Then, 10 μ l DMSO (Fisher Scientific, Loughborough, UK) were added and cells were incubated for 10 min at 42°C. Afterwards, cells were harvested at 3000 g/3 min, resuspended in 200 μ l of sterile YPD and plated onto SC-Leu/His plates. Plates were incubated for 2 days at 30°C and single colonies were resuspended in 50 μ l PBS, of which 5 μ l cell suspension was spotted onto SC-selective plates and incubated 1 day at 30°C. Cells were then tested for protein interaction by X-Gal overlay. Plates were incubated at 30°C for 12–24 h and then photographed.

Microscopy and image analysis

Indirect Immunofluorescence (IF) was performed essentially as described in Erhardt *et al.* (24). For mitotic chromosome preparations, cells were arrested with 3.3 μ g/ml Colcemid (PAA) for 1–2 h, incubated in 0.5% Na-citrate for 7 min and cytospun using a Shandon 4 (Thermo) at 900 rpm for 10 min. The antibodies used were rabbit α -CENP-A, 1:500 (Aaron Straight); chicken α -CID 1:500 (Erhardt Lab); guinea pig α -CENP-C, 1:1000 (Gary Karpen); guinea pig α -CENP-C, 1:5000 (Erhardt Lab); rabbit α -His 1:1000 (abcam); mouse α -V5, 1:1000 (life technologies™); α -lamin, 1:500 (Developmental Studies); mouse α -tubulin 1:1000 (Sigma); mouse α -GFP 1:500 (Roche); chicken α -GFP 1:200; rabbit α -Mi-2 1:500 (Alexander Brehm). Secondary antibodies were applied at a concentration of 1:500 (Alexa Fluor®).

Live cell imaging was performed in 8-well chambers (Ibidi) with 1 mM CuSO₄ (\pm 20 μ M MG132). Imaging was performed every 20 min for up to 16 h with the following settings: 20 μ m in Z, 0.5 μ m interval distance, 0.03 sec exposure for mCherry-tubulin, 0.04 s exposure for induced EGFP-tagged proteins. Living and fixed cells were imaged on a DeltaVision® Core, deconvolved, projected and background subtracted. mCherry-tubulin signals were equalized. All processing was performed using softWoRx suite and ImageJ. To quantify fluorescence of the Live Cell Imaging experiments in an unbiased way the contrast was adjusted equally for every picture and a rolling ball background subtraction was carried out. Subsequently the Raw-IntegratedDensity pixel value of the whole cell and the respective nucleus was measured over 8 h in a 30 min interval, while the measured areas were adjusted according to the movement of the cell, for every timepoint using ImageJ software. The data was subsequently analyzed and statistical tests were done using RStudio software. Time course graph was plotted using Excel.

To determine the penetrance of observed phenotypes, cells or chromosomes were counted for the presence or the absence of the respective phenotypes, and the percentages were calculated accordingly. Signal intensities of maximum projected cells and chromosomes were determined using the thresholding and measuring by ImageJ software tool. The bar graphs were plotted using the GraphPad Prism. Error bars represent standard error of the mean (SEM). Significance was determined by using the Student's *t*-test.

Co-immunoprecipitation of GFP-tagged proteins

GFP fusion proteins were precipitated with the GFP-binding protein (GBP) covalently bound to NHS-sepharose (82). S2 cells expressing GFP protein at low levels under pCopia promoter (GFP only) were used as control for IP. Cell lysis was performed in IP buffer (50 mM Tris-HCl pH 8, 150 mM NaCl, 1% NP-40, 0.5% Na-Doc, 10 mM NaF, 0.1% SDS, 40 mM NEM, 1 mM EDTA, 1 \times Roche complete protease inhibitor cocktail). The lysate was centrifuged for 30 min at high speed (13000 rpm). Then the supernatant was transferred to a fresh reaction tube carrying 50 μ l of NHS-GBP (equilibrated in lysis buffer). Co-IP was performed for 2–3 h at 4°C. Beads were collected by centrifugation (3000 rpm, 1 min) and washed six times in 1 ml IP buffer. Proteins were eluted in 2 volumes 2 \times SLB for 5 min at 95°C at vigorous shaking and separated by SDS-PAGE followed by western blot or mass spectrometry analysis.

Co-immunoprecipitation of V5-tagged proteins, endogenous Mi-2 and H3 from fly cultured S2 cells and early embryos

In early 0–6 h embryos, CENP-A-V5 expression was induced by crossing CENP-A-V5 flies with actin-Gal4 driver flies at 29°C. About 20 μ l dry volume of embryos were harvested and lysed in an IP buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 1% NP-40, 10 mM NaF, 2 mM PMSF, 20 mM NEM, 1 \times protease inhibitor cocktail-Roche) using a pestle. 1.0–2.0 \times 10⁸ S2 cells expressing V5-tagged proteins or endogenous Mi-2 and H3 were washed with cold PBS and then lysed in an IP buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 1% Triton X-100, 10 mM NaF, 2 mM PMSF, 20 mM NEM, 1 μ g/ml Aprotinin/leupeptin, 0.5 μ g/ml Pepstatin). After incubating at 4°C for 20 min, the lysate was centrifuged for 30 min at maximum speed. Subsequently, the supernatant was transferred to a precooled tube containing 1 μ g of V5/H3/Mi-2-antibody covalently coupled to 20 μ l dry volume of agarose G beads (Roche). The Co-IP was performed for 2–3 h at 4°C. Beads were then collected by centrifugation (3000 rpm, 1 min) and washed 6 \times in 20 volumes of lysis buffer. Proteins were eluted in 2 volumes 2 \times SLB for 5 min at 95°C at vigorous shaking and separated by SDS-PAGE followed by western blot.

Immunoblotting

Cells were lysed in RIPA buffer containing 50 mM Tris-HCl pH7.5, 500 mM NaCl, 1% NP-40, 0.5% sodium deoxycholate, 0.1% SDS, 2 mM PMSF, Roche complete, 250 u/ml Benzonase (Sigma), sonicated (Bioruptor, Diagenode) and supplemented with 4 \times SDS Laemmli buffer (SLB). Alternatively, after wash with PBS, cells were directly supplemented with 2 \times SLB (40 ul/1 million) and sonicated. For denaturation, the samples were boiled at 95°C for 5 min. For protein fractionation, cells were lysed in lysis buffer containing 50 mM Tris-HCl pH 7.5, 150 mM NaCl, 1% NP-40, 2 mM PMSF, Roche complete. The lysate was centrifuged at 11 300 rpm for 15 min, and the supernatant contained soluble fraction. Residual pellet was resuspended with the same lysis buffer supplemented with 250 u/ml benzonase, sonicated and centrifuged at 11 300

rpm for 15 min. The supernatant contained chromatin fraction. The protein samples were separated using SDS-PAGE gel electrophoresis and transferred onto a 0.45 μ m nitrocellulose membrane in Tris/glycine/methanol containing buffer. Primary antibodies used for blotting are rabbit α -CID 1:2000 (Active Motif); rabbit α -H3 1:5000 (Abcam); rabbit α -CAL1 1:2500 (Erhardt Lab); mouse α -actin 1:5000 (Millipore); mouse α -alpha tubulin 1:5000 (Sigma); mouse α -V5 1:5000 (Invitrogen); mouse α -GFP 1:2500 (Roche); mouse α -Lamin 1:5000 (Developmental Studies); rabbit α -RbAp48 1:20 000 (Alexander Brehm); rabbit α -Mi-2 1:8000 (Alexander Brehm); guinea pig α -MTA1-like 1:2500 (Alexander Brehm). Secondary antibodies used for blotting are HRP-conjugated α -rabbit 1:5000 (Abcam); α -guinea pig 1:5000 (Abcam); α -mouse 1:10 000 (Abcam). Signal detection was performed using ECL chemiluminescent solution (Thermo Fisher).

Crosslinking IP (Xlink-IP)

Xlink-IP protocol was performed based on (49). Briefly, 10^8 cells were washed in PBS and incubated in 10 ml 0.4% PFA at RT for 7 min with mild agitation. Then cells were centrifuged, and PFA was removed. 0.5 ml 1.25 mM ice cold glycine-PBS was added to quench the reaction. After centrifugation and removal of glycine-PBS, cells were lysed in 1 ml Xlink IP buffer (50 mM Tris-HCl pH 8, 150 mM NaCl, 1% NP-40, 0.5% Na-Doc, 10 mM NaF, 0.1% SDS, 40 mM NEM, 1 mM EDTA, 1 \times Roche complete protease inhibitor cocktail) for 30–60 min at 4°C on a rotator. Sonication for 50 cycles (10 s ON, 5 s OFF) was performed. The lysates were centrifuged at 13 000 rpm for 30 min and then transferred to the GBP beads (equilibrated). GBP-binder protocol was performed as described above. Elution was done with 50 μ l 4 \times laemmli at 65°C for 10 min to keep Xlinks or at 95°C for 10 min to reverse Xlinks. The working of the protocol was confirmed by WB and coomassie staining, and samples were analyzed by Mass spec.

Mass spectrometry of proteins and data analysis

Protein samples prepared for mass spectrometry (MS) were separated on SDS-PAGE. After coomassie staining gels have been sliced into equidistant pieces for fractionation, which have been cut out and digested with Trypsin using the Digest Pro Robotic System from Intavis. Peptides have been separated on an in-house packed C18 reversed-phase column of 25 cm length using a 120 min gradient from 3% to 36% ACN and directly injected to a Q-Exactive HF mass spectrometer.

Data analysis was carried out by MaxQuant (version 1.5.3.30) (83). In total, 12671 peptides and 1791 proteins could have been identified by MSMS based on an FDR cut-off of 0.01 on peptide and protein level. Match between runs option was enabled to transfer peptide identifications across Raw files based on accurate retention time and m/z. Quantification was done using a label free quantification approach based on the MaxLFQ algorithm. A minimum number of two quantified peptides was required for protein quantification, requantify option was activated, to enable quantification of proteins with very high ratios. In total

1787 proteins were quantified. MaxQuant raw output files have been filtered and visualized using in-house compiled R-scripts. For the Fisher Exact Test, a sub-group of data (e.g. 137 proteins having a more than two-fold higher abundance in B3A samples) were compared with total data (1791 proteins). Perseus software was used for the calculation of Fisher Exact Test. Fisher exact test was applied by using the Perseus software to identify protein categories (84), like GO terms, showing consistent abundance changes across their members (FDR filter of 0.02 using Benjamini–Hochberg corrected *P*-values).

The listed proteins in Supplementary Table S1 include enriched factors filtered according to \log_2 LFQ >1 over the control. Scatterplot of \log_2 LFQ-intensities for CENP-A^{FL} versus CENP-A^{B3(A)} is shown in Figure 3C, and the enriched factors are shown in Supplementary Table S2. The interesting candidate protein complexes were determined using String pathway analysis online tool.

Correlation test and protein alignment analysis

ICGC data (US donors with both RNA and SNV data) – International Cancer Genome Consortium- was analyzed using USCS XENA tool. Gene expression for CENP-A, SMARCA1, BPTF, MTA1, MTA2, MTA3, HDAC1, RBBP4 and RBBP7 was selected. Pearson's Rho comparison of CENPA with each other component's gene expression was performed and assembled in Inkscape.

Multiple sequence alignment of CENP-A amino acid sequences for five different *Drosophila* species and the phylogenetic tree comparison were performed using clustal omega. Multiple sequence alignment of CENP-A amino acid sequences for *Drosophila melanogaster* and human was performed using T-coffee.

RESULTS

The B3 domain of the N-terminal tail of CENP-A is required for protein stability and nuclear localization

To investigate if the N-terminal tail (N-tail) of CENP-A is required for non-centromeric localization, we created truncation constructs by removing the tail in 25 amino acid (aa) increments ($\Delta 25$, $\Delta 50$, $\Delta 75$, $\Delta 100$ and $\Delta 125 = \Delta$ N-CENP-A), which were expressed from a copper sulfate (CuSO₄)-inducible metallothionein (pMT) or constitutively-expressing pCopia promoter tagged with V5-His or EGFP (Figure 1A). Deletion of the first 100 residues had no obvious effect on protein localization and stability compared to full-length CENP-A-V5-His (CENP-A^{FL}; Figure 1B-C). Only the deletion of the entire 125 aa N-terminus (Δ N-CENP-A) resulted in a loss of protein levels, even though the constructs were all expressed at comparable levels after CuSO₄-induction as determined by RT-PCR (Supplementary Figure S1A). Proteasome inhibition did not prominently affect $\Delta 100$ or any of the longer mutants but caused a substantial increase of Δ N-CENP-A-V5-His total protein levels (Figure 1B-C, Supplementary Figure S1B-E). We concluded that the residues between 101aa and 125aa of the CENP-A N-tail contain information that influences protein stability. As previously reported, overexpressed CENP-A^{FL}-V5-His incor-

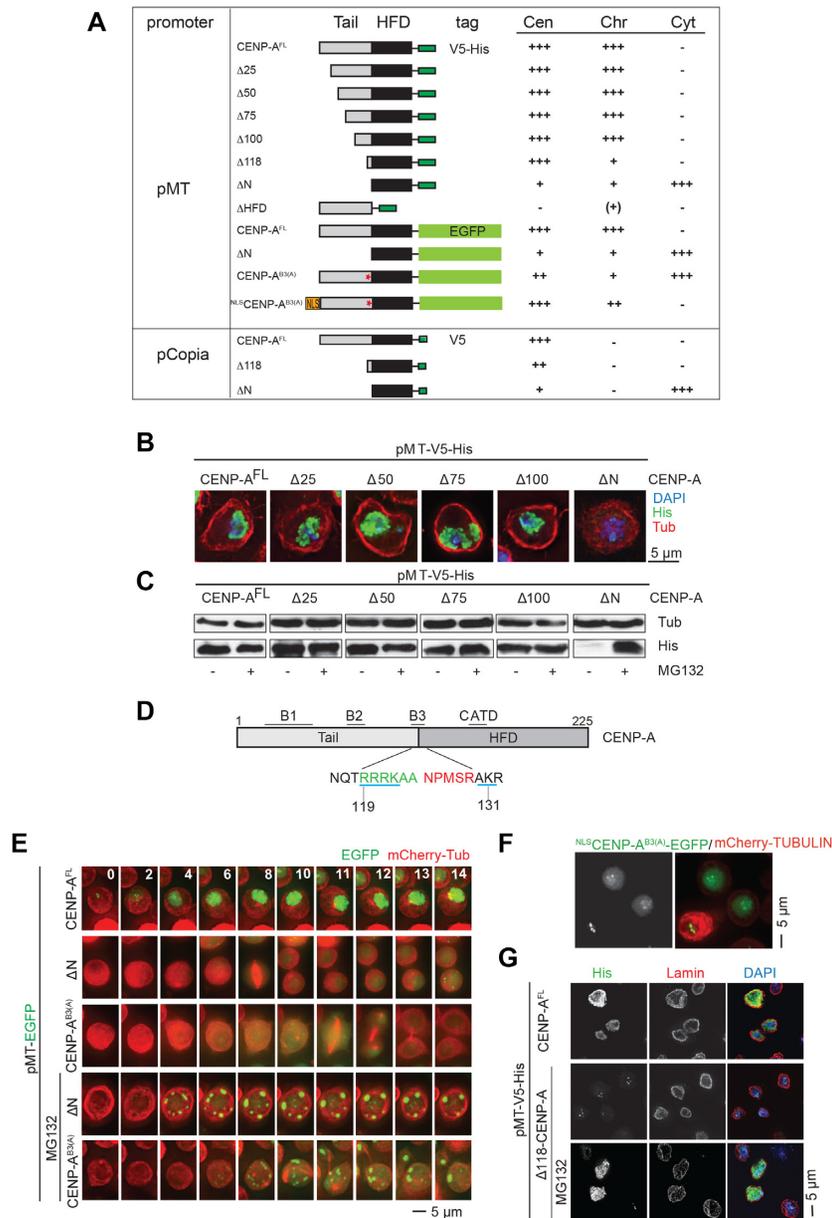


Figure 1. N-terminal tail and B3 domain of CENP-A are required for stability and nuclear localization. **(A)** Schematic representation of constructs used as stably transfected S2 cell lines in this study. The predominant localizations of the different constructs in the presence of a functional proteasome to centromeres (Cen), chromatin (Chr), or cytosol (Cyt) is schematically indicated by the following symbols: (+): hardly detectable, +: weak but detectable signal, ++: clearly detectable signal, +++: strong signal. **(B)** Representative images of cells expressing either full length CENP-A (CENP-A^{FL}) or different N-terminal truncated CENP-A-V5-His constructs (Δ25, Δ50, Δ75, Δ100, ΔN = Δ125). DAPI in blue, α-His in green and α-Tubulin (Tub) in red. Scale bar: 5 μm. **(C)** Corresponding western blot analysis of overexpressed CENP-A^{FL} or different N-terminal truncated CENP-A-V5-His constructs with or without MG132 treatment. Tubulin (Tub) served as loading control. **(D)** Model of CENP-A primary structure (tail: aa 1–125, HFD: aa 126–225). Positions of conserved domains are indicated (B1–3, CATD). Sequence of B3 is shown in green (tail) and red (HFD). Residues potentially encoding a bipartite nuclear import signal are underlined in blue. **(E)** Stills from time lapse microscopy of cells expressing mCherry tubulin and CENP-A^{FL}-EGFP, ΔN-CENP-A-EGFP or CENP-A^{B3(A)}-EGFP (induced with 1 mM CuSO₄) in the presence or absence of MG132 (see supplemental movies 1–5). CuSO₄ was added 1 h prior to time point t = 0. EGFP in green and mCherry-tubulin in red. Scale bar: 5 μm. **(F)** Depiction of CENP-A^{B3(A)}-EGFP with the conjugated NLS at the N-terminal site and representative images for its localization after induction. EGFP in green and mCherry-tubulin in red. Scale bar: 5 μm. **(G)** Representative images of induced Δ118–CENP-A-V5-His and CENP-A^{FL}-V5-His construct in the presence and absence of MG132. DAPI in blue, α-His in green and α-Lamin in red. Scale bar: 5 μm.

porated into non-centromeric regions along chromosome arms (11,43) and so did the tail-truncated CENP-A constructs $\Delta 25$, $\Delta 50$, $\Delta 75$ and $\Delta 100$ (Figure 1A, B and Supplementary Figure S1B). In contrast, ΔN -CENP-A-V5-His was virtually undetectable in transfected cells and metaphase chromosome spreads, and only a minority of ΔN -CENP-A-V5-His cells exhibited a low centromeric signal as confirmed by CENP-C co-immunostaining (Figure 1A, B, Supplementary Figure S1B–E). Additionally, about 12% of cells displayed weak cytoplasmic foci of ΔN -CENP-A-V5-His after prolonged high induction (e.g. overnight), which were never observed for CENP-A^{FL}-V5-His (Supplementary Figure S1B–C). Mitotic chromosome spreads showed that ΔN -CENP-A-V5-His localized very weakly to centromeric regions only, whereas CENP-A^{FL}-V5-His incorporated broadly into non-centromeric chromatin under identical expression conditions (Supplementary Figure S1B and C).

The protein region between aa 101 and 125 of CENP-A contains an arginine-rich motif within the conserved block 3 (B3) domain (44) (Figure 1D). Sequence analysis suggested that this motif contains a bipartite nuclear localization signal (NLS). To address the potential significance of this motif in nuclear localization, we made inducible stable cell lines that co-expressed a triple alanine mutant of the arginine-rich motif (CENP-A^{B3(A)}-EGFP) or CENP-A^{FL}-EGFP and ΔN -CENP-A-EGFP together with mCherry-tubulin and imaged these stable cell lines live for 16 h (Figure 1E and movies 1–5). CENP-A^{FL}-EGFP localized to centromeres already in non-induced, leaky cells as previously described (43). Induced overexpression led to fast nuclear accumulation of CENP-A^{FL}-EGFP (Figure 1E and movie 1). ΔN -CENP-A-EGFP, on the other hand, showed no GFP signal for several hours after induction and only a slowly increasing, faint EGFP signal throughout the cells after prolonged (e.g. overnight) imaging (Figure 1E and movie 2). Small cytoplasmic foci were again observed after prolonged induction with CuSO₄, probably protein that the proteasome has not been able to degrade yet. Indeed, proteasome inhibition caused an enlargement of these ΔN -CENP-A-EGFP foci in the cytoplasm (Figure 1E and movie 4). Similar to ΔN -CENP-A-EGFP, CENP-A^{B3(A)}-EGFP was only weakly visible in the cytoplasm and nucleus after induction (Figure 1E and movie 3). Upon proteasome inhibition, CENP-A^{B3(A)}-EGFP cells accumulated cytoplasmic foci perfectly recapitulating ΔN -CENP-A-EGFP appearance under these conditions (Figure 1E and movie 5). The accumulation of cytoplasmic protein after proteasome inhibition was also detectable in ΔN -CENP-A-V5-His but never in CENP-A^{FL}-V5-His cells (Supplementary Figure S1E). To confirm that the B3 motif functions as an NLS of CENP-A, we placed the NLS of SV40 (PPKKRKRKV) onto the N-terminus of CENP-A^{B3(A)}-EGFP to create NLS-CENP-A^{B3(A)}-EGFP and stably co-transfected this construct with mCherry-tubulin into S2 cells. Indeed, GFP was detectable in the nucleus and its preferred localization was virtually identical to ectopically expressed CENP-A^{FL} (Figure 1F). We concluded from this set of experiments that the B3 domain of CENP-A contains an NLS and that cytoplasmic CENP-A is subject to proteasomal degradation.

We noticed that even though CENP-A^{B3(A)}-EGFP cannot be imported into the nucleus and is subject to proteasomal degradation, protein levels were always slightly higher than the ΔN -CENP-A-EGFP construct. To distinguish nuclear import from protein stability, we created a ‘minimal tail’ construct containing only the B3 and the HFD of CENP-A ($\Delta 118$ -CENP-A) (Figure 1A and G). Indeed, $\Delta 118$ -CENP-A-V5-His efficiently localized to the nucleus in all transfected cells. However, $\Delta 118$ -CENP-A levels were strongly reduced compared to CENP-A^{FL} or the longer forms of CENP-A, both under overexpressed (pMT), as well as under low, constitutive expression conditions (pCopia) (Figure 1A, G and Supplementary Figure S1F). Importantly, MG132 treatment resulted in ectopic $\Delta 118$ -CENP-A-V5-His incorporation throughout chromatin indistinguishable from CENP-A^{FL}-V5-His (Figure 1G). This data confirmed that nuclear localization depends on the arginine-rich motif. Nevertheless, protein stability was still affected by the lack of the first 117 amino acids, suggesting either a protein stabilizing function of the CENP-A tail or simply a destabilization of the protein by an artificially exposed arginine stretch (N-end rule) (45). Interestingly, a clustal omega analysis showed that the triple arginines are evolutionarily conserved through five *Drosophila* species and also in humans, suggesting that this NLS region is highly conserved (Supplementary Figure S4A–C).

CENP-A loading factors influence the location of tail-depleted CENP-A

Canonical histones carry their NLS in their N-terminus and deletion of one NLS within a H3/H4 or H2A/H2B dimer does not prevent import due to a compensatory effect of the dimerization partner’s histone tail (46,47). We hypothesized that co-expression of pre-nucleosomal interaction partners may influence ΔN -CENP-A localization or degradation dynamics. Co-expression of histone H4-V5-His led to prominent nuclear localization of H4-V5-His but without major changes in ΔN -CENP-A-EGFP cytoplasmic localization or stability (Figure 2A), indicating that the NLS of H4-V5-His is insufficient to rescue nuclear import of a CENP-A/H4 complex (Figure 2A). Interestingly, co-expression of CENP-A^{FL}-V5-His and ΔN -CENP-A-EGFP completely restored ΔN -CENP-A-EGFP localization and stability within the nucleus similar to CENP-A^{FL}, suggesting that CENP-A may be present as a dimer in a pre-nucleosomal complex (Figure 2A).

The chaperones RbAp48 (also known as p55 in *Drosophila*) and CAL1 have been described to complex with pre-nucleosomal CENP-A (30,48). Co-expression of either RbAp48-V5-His or CAL1-V5-His with ΔN -CENP-A-EGFP led to a rescue of the phenotype of ΔN -CENP-A-EGFP with, however, remarkable differences: RbAp48-V5-His co-expression caused nuclear localization of ΔN -CENP-A-EGFP that resembled full-length CENP-A-EGFP overexpression whereas CAL1 co-expression resulted in nuclear ΔN -CENP-A-EGFP accumulations to several foci that resemble centromeric CENP-A signals of interphase cells (Figure 2A).

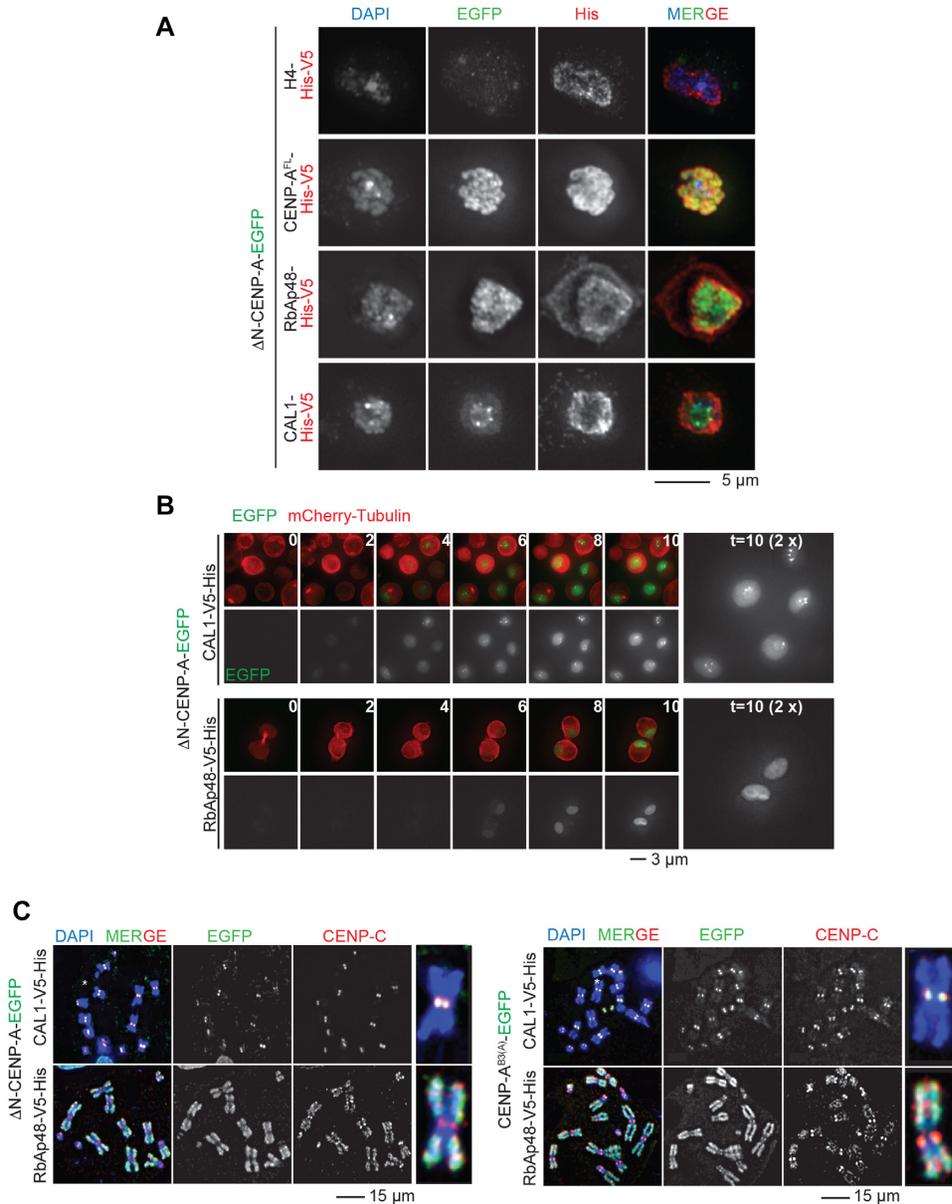


Figure 2. Full-length CENP-A and CENP-A loading factors stabilize Δ N-CENP-A or CENP-A^{B3(A)}, safeguard it to the nucleus and ectopic chromosome arms. (A) Representative images of cells co-overexpressing Δ N-CENP-A-EGFP and either H4-V5-His, CENP-A^{FL}-V5-His, RbAp48-V5-His or CAL1-V5-His. DAPI in blue, EGFP in green and α -His in red. Scale bar: 5 μ m. (B) Stills from time-lapse microscopy of cells co-expressing mCherry-tubulin and Δ N-CENP-A-EGFP with CAL1-V5-His (Supplementary movie 7) or RbAp48-V5-His (Suppl. movie 8) Time points in min are indicated. EGFP in green and mCherry-tubulin in red. Scale bar: 3 μ m. On the right, 2 \times zoom image for $t = 10$ min is shown. (C) Representative images of mitotic chromosomes co-overexpressing Δ N-CENP-A-EGFP or CENP-A^{B3(A)}-EGFP together with CAL1-V5-His or RbAp48-V5-His. DAPI in blue, EGFP in green and α -CENP-C in red. Scale bar: 15 μ m.

We repeated the co-expression of either RbAp48-V5-His or CAL1-V5-His with the N-terminal deleted CENP-A construct Δ N-CENP-A-EGFP and performed time lapse microscopy with cells co-expressing mCherry-Tubulin. CAL1-V5-His co-expression mediated centromeric incorporation of Δ N-CENP-A-EGFP, quickly accumulating after induction without the passage through mitosis (48); RbAp48-V5-His co-expression also mediated a nuclear localization of Δ N-CENP-A-EGFP throughout the nucleus with no centromeric specificity (Figure 2B and movies 6–7). We tested several other known histone chaperones for

their ability to load Δ N-CENP-A-EGFP, including NASP, DLP, and XNP by live cell imaging (Supplementary Figure S2A and movies 8–13) but none showed rescue of the Δ N-CENP-A or CENP-A^{B3(A)} phenotypes.

To confirm the different loading patterns of mutant CENP-A cells co-expressing RbAp48-V5-His or CAL1-V5-His, we analyzed the distribution of Δ N-CENP-A-EGFP or CENP-A^{B3(A)}-EGFP and of the inner kinetochore component CENP-C on mitotic chromosome spreads (Figure 2C). EGFP localization was strikingly different between RbAp48-V5-His and CAL1-V5-His co-expressing cells:

CAL1-V5-His loaded Δ N-CENP-A-EGFP and CENP-A^{B3(A)}-EGFP exclusively to centromeric sites. Immunofluorescence staining with CENP-C-specific antibodies showed that CENP-C was localized to centromeres only (Figure 2C). In contrast, RbAp48-V5-His expressing cells displayed a strong EGFP signal along the entire chromosome arms, with little signal at pericentric heterochromatin. CENP-C was recruited to many of the ectopic EGFP sites, suggesting the formation of ectopic kinetochores as previously shown for full-length CENP-A overexpression (43). Some of the RbAp48-V5-His co-expressing cells displayed low centromeric signal; however, additional depletion of endogenous CAL1 by RNAi led to the specific loss of centromeric Δ N-CENP-A-EGFP signal (Supplementary Figure S2B). Taken together, mislocalization of tail-deleted or NLS mutant CENP-A is rescued by co-expressing full-length CENP-A, the histone binding protein RbAp48, or the loading factor CAL1. Centromere specificity of CENP-A loading is mediated exclusively by CAL1 in *Drosophila* S2 cells; ectopic loading, however, requires RbAp48.

The NuRD complex associates with CENP-A

Even though RbAp48 has been identified as a bona fide interaction partner of CENP-A in *Drosophila* cultured cells (30), the RbAp48-containing protein complex that associates with CENP-A is unknown. To identify the complex responsible for the observed RbAp48-dependent non-centromeric loading of CENP-A, we expressed CENP-A^{B3(A)}-EGFP alone (low cytoplasmic GFP signal) or CENP-A^{B3(A)}-EGFP in combination with RbAp48-V5 (ectopic nuclear GFP signal), as well as CENP-A^{FL}-GFP (ectopic nuclear GFP signal) in S2 cells and determined their interaction partners by cross-linking protein complexes before immunoprecipitation and mass spectrometry (Xlink IP-MS) to identify all interactions, including transient interactions (49). Pulled-down complexes were detectable by western blot (WB) as higher molecular weight bands when the crosslinks were preserved at 65°C and largely disrupted when the crosslinks were reversed at 95°C (Figure 3A). The pairwise comparison of LFQ intensities for detected proteins indicated an enrichment of 566 proteins over control samples (LFQ ratio > 2; Supplementary Table S1). Among the detected proteins, 114 of those were either known centromeric proteins or previously identified enriched centromeric proteins (50). Gene Ontology (GO) terms analysis further implicated that CENP-A^{B3(A)}-EGFP interacts with chromatin-associated factors upon RbAp48 co-expression, resembling overexpressed CENP-A^{FL} whereas CENP-A^{B3(A)}-EGFP without RbAp48 co-expression resulted in pulled down proteins that did not enrich for nuclear and chromatin-associated GO terms (Figure 3B). Further filtering (LFQ CENP-A^{FL} vs LFQ CENP-A^{B3(A)} > 2) reduced the candidate interacting factors to 119 enriched proteins (Supplementary Table S2). Among those, we found factors already known to interact with non-centromeric CENP-A in other species, such as DEK and FACT complex subunits (19,51) (Supplementary Table S2). Importantly, we identified several components of the RbAp48-associated ‘Nucleosome Remodeling and Deacetylase’ complex NuRD

to be enriched with CENP-A^{B3(A)}-EGFP/RbAp48-V5 and CENP-A^{FL}-GFP when compared to CENP-A^{B3(A)}-EGFP (Figure 3C and Supplementary Tables S1 and S2). The *Drosophila* NuRD complex contains seven subunits: the histone deacetylase proteins Rpd3, the histone-binding proteins RbAp48, the metastasis-associated proteins MTA1-like, the methyl-CpG-binding domain protein MBD2/3-like and the chromodomain-helicase-DNA-binding protein Mi-2 (ortholog of human CHD3) (52). Except for MBD2/3-like, all subunits were enriched in our mass spectrometry data sets of nuclear CENP-A-GFP.

To confirm the interaction of the NuRD complex components with nuclear CENP-A, we performed co-immunoprecipitation (IP) using the same cell lines as for the mass spectrometry analysis. The interactions with the NuRD complex components Mi-2, MTA1-like, and RbAp48 were strongest with CENP-A^{FL}-EGFP but also detectable with CENP-A^{B3(A)}-EGFP with and without RbAp48 co-expression (Figure 3D). Even though we did not detect an interaction between RbAp48 and CENP-A by a yeast two hybrid approach, we confirmed a specific interaction between domains of MTA1 and CENP-A by this approach (Supplementary Figure S6E). These data further indicate that the NuRD components, indeed, interact with CENP-A-EGFP.

The NuRD complex is essential for ectopic CENP-A localization

To initially characterize a functional involvement of the NuRD complex in ectopic CENP-A loading, we depleted the two essential NuRD components Mi-2 and MTA1-like by RNA interference (RNAi) in cells expressing the inducible CENP-A-V5-His constructs. A significant reduction was apparent at the protein level by immunoblotting after 4 days of RNAi (Figure 4A, D and Supplementary Figure S5B). CENP-A-V5-His levels were strongly reduced, including the higher running CENP-A band that we previously identified as a mono-ubiquitylated form of CENP-A (53) (Figure 4A). In the absence of either Mi-2 or MTA1-like, we observed a significant reduction of ectopic CENP-A^{FL} by immunostaining on metaphase chromosome spreads (Figure 4B-C). In the presence of the NuRD complex components (Brown RNAi control), overexpressed CENP-A^{FL} widely incorporated into chromosome arms. In the absence of NuRD complex components (Mi-2 or MTA1-like RNAi), the levels of ectopic CENP-A on chromosome arms were strongly reduced but centromeric CENP-A levels remained readily visible (Figure 4B). Importantly, the levels of RbAp48 were stable after Mi-2 RNAi (Supplementary Figure S4E).

We next analyzed if the amount and cellular localization of nuclear CENP-A^{B3(A)}-EGFP levels co-expressed with RbAp48 were affected in cells depleted of NuRD complex components. We induced CENP-A^{B3(A)}-EGFP in Mi-2 or MTA1-like depleted cells and inhibited the proteasome before separating the cell into a soluble (containing cyto- and nucleoplasmic proteins) and a chromatin fraction (Figure 4D). Mi-2 and MTA1-like were predominantly found in the soluble fraction of control and depleted cell extracts. In control cells (Bw), the majority of overexpressed CENP-A^{B3(A)}-

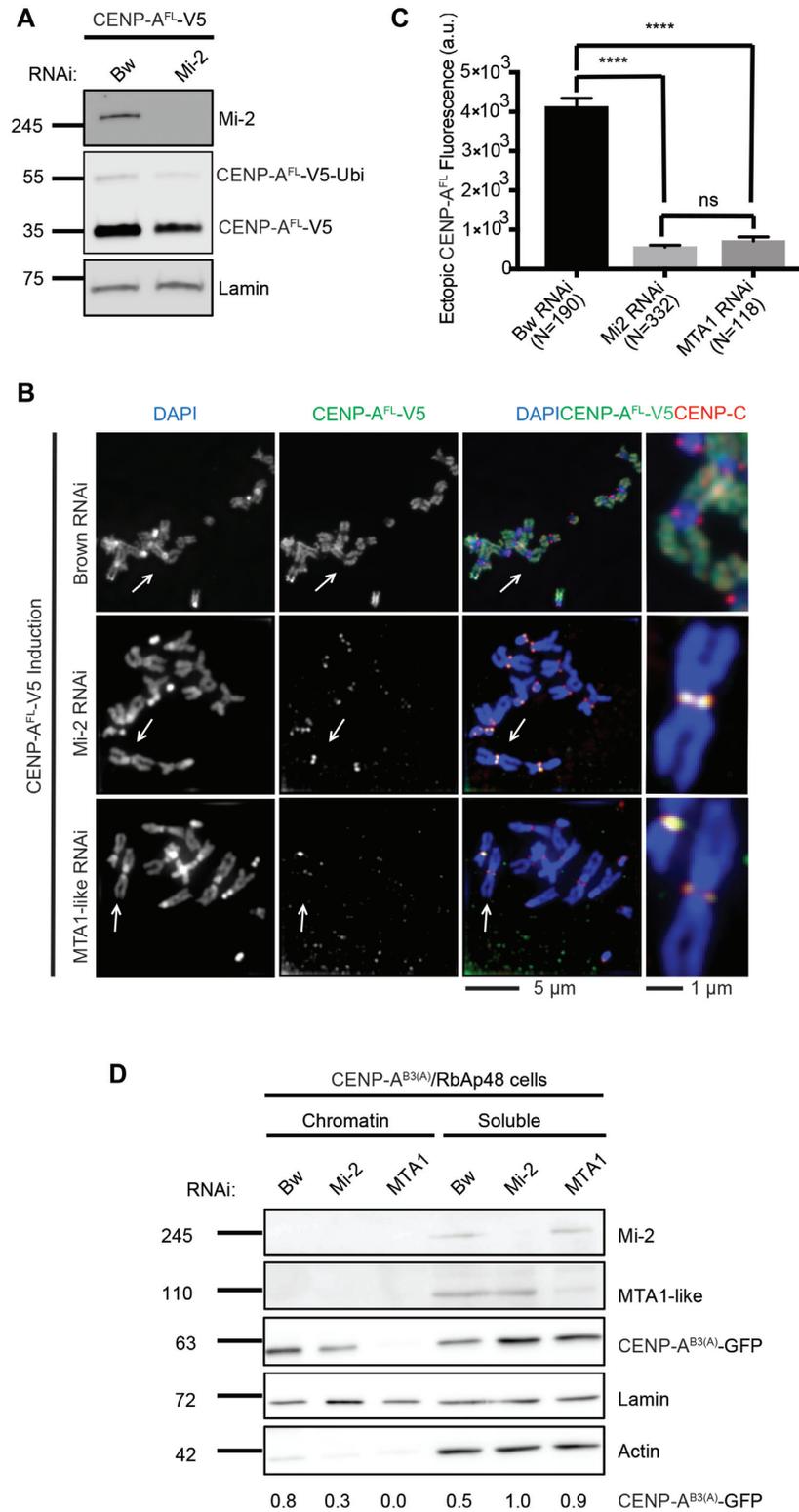


Figure 4. Ectopic CENP-A localization is reduced upon Mi-2 and MTA1-like depletion. **(A)** Western blot of Mi-2 knockdown by RNAi upon CENP-A^{FL-V5} induction (0.05 mM) probed with α -Mi-2, α -lamin, and α -V5 antibodies. Lamin served as loading control. **(B)** Representative images of mitotic chromosome spreads overexpressing CENP-A^{FL-V5} under Brown (Bw, control), Mi-2 and MTA1-like depletion. DAPI in blue, CENP-A^{FL-V5} in green and CENP-C in red. Scale bars: 5 μ m and 1 μ m. **(C)** Quantification of ectopic CENP-A^{FL} fluorescence per chromosome from chromosomes spreads as shown in B. Each box represents the mean of three independent experiments, and the error bars indicate the SEM. Fluorescence is measured in arbitrary units (a.u.). Student's t test: **** $P < 0.0001$ and ns: not significant. **(D)** Western blot of proteins fractionated into chromatin-associated and soluble proteins from cells co-overexpressing CENP-A^{B3(A)}-GFP and RbAp48-V5-His in Brown, Mi-2, MTA1-like depleted cells probed with α -Mi-2, α -MTA1-like, α -GFP, α -lamin and α -actin. The band intensity of CENP-A^{B3(A)}-GFP was measured by ImageJ and normalized to lamin for chromatin fraction and normalized to actin for soluble fraction.

EGFP was within the chromatin fraction even though a significant amount was also detectable in the soluble fraction under proteasome inhibition. Importantly, when we depleted Mi-2 or MTA1-like, a large proportion of CENP-A^{B3(A)}-EGFP protein was not present in the chromatin fraction anymore but within the soluble fraction (Figure 4D). This result is in line with our results of mitotic chromosome spreads (Figure 4B) that shows that only centromeric CENP-A^{FL}-V5 remains chromatin associated upon the depletion of NuRD complex component. These results further indicate that a functional NuRD complex is required for the localization of CENP-A to ectopic chromatin sites.

The NuRD complex was previously shown to bind to the H3 N-terminal tail (54). To test if the NuRD complex still binds H3 when CENP-A is overexpressed or whether H3 and CENP-A compete for NuRD we performed co-IP experiments (Supplementary Figure S3A and B). We immunoprecipitated Mi-2 upon CENP-A overexpression and found that the Mi-2-CENP-A interaction is very strong and that there is a decreased interaction with H3 when CENP-A is overexpressed. We further immunoprecipitated H3 or CENP-A-V5 upon CENP-A overexpression and found that Mi-2 was more efficiently co-IPed with V5 than with H3 further suggesting that CENP-A may have a stronger affinity to NURD than H3 (Supplementary Figure S3A and B).

NuRD complex components were indicated to be up-regulated in cancer (55). For instance, MTA1/2/3 are highly overexpressed in tumors and correlate with tumor metastasis (56). Given that CENP-A is overexpressed in many different human tumor entities (15,17), we tested whether NuRD subunits are also misexpressed in the same tumors. Using the UCSC XENA tool and ICGC database we compared the correlation in expression across all available cancer samples. We selected gene expression for CENP-A, SMARCA1, BPTF, MTA1, MTA2, MTA3, HDAC1, RBBP4 and RBBP7. Pearson's Rho comparison of CENP-A with each other component's gene expression was performed. We found that the expression of several NuRD complex components, including MTA1/2/3, RBBP4/7 and HDAC1 strongly correlates with CENP-A overexpression. On the contrary, NURF-specific remodeling subunits, which also contains RbAp48 as a subunit, negatively correlates with CENP-A, supporting a correlation between CENP-A and the NuRD remodeling complex, and that the interaction may be also important for misincorporation of CENP-A in human cancer cells (Supplementary Figure S4F). Overall, the ectopic localization of overexpressed CENP-A through NuRD complex interaction might have an important role in cancer.

RbAp48 mutant for MTA1-like binding prevents CENP-A ectopic loading

So far, we assumed that the interaction of NuRD with ectopic CENP-A is via its binding partner RbAp48. To test this, we co-transfected CENP-A^{B3(A)}-EGFP with a RbAp48 mutant construct that produces a protein with five point mutations, E361Q, D362N, E364Q, D365N and L35Y, which has previously been shown to block its interaction with MTA1-like and all the other NuRD subunits

(57). We created S2 cell lines expressing the RbAp48 mutant alone or together with CENP-A^{B3(A)}-EGFP. Indeed, mutant RbAp48 lost the physical interaction with its direct interacting partner MTA1-like as described previously (57) (Figure 5A). However, it retained most of its interaction with overexpressed CENP-A^{B3(A)}-EGFP and the CAF-1 subunit p180 as determined by Western blotting (Figure 5A). Importantly, even though mutant RbAp48 can still associate with CENP-A^{B3(A)}-EGFP by Western blot, we observed an about 10-fold decrease in CENP-A^{B3(A)}-EGFP localization to non-centromeric sites along mitotic chromosome arms when compared to wild type RbAp48 (Figure 5B and C). The only detectable CENP-A^{B3(A)}-EGFP signal visible was centromeric when co-expressed with the MTA1-like binding-deficient RbAp48 and not ectopically as when co-overexpressed with wild type RbAp48. This further indicates that MTA1-like is involved in CENP-A ectopic loading (Figure 5B). Overall, we conclude from these set of experiments that the NuRD complex is required for non-centromeric CENP-A localization to chromosome arms.

The nucleosome remodeling activity of Mi-2 is required for CENP-A misincorporation

Given that NuRD is an ATP-dependent nucleosome remodeler and our data suggests that its ATPase subunit Mi-2 is involved in incorporation of CENP-A at ectopic sites we tested whether Mi-2's remodeling activity plays a role. To address this, we utilized the previously characterized Mi-2 wild type and ATPase catalytic domain loss of function mutants (H1151R, R1161Q), an ATPase motor brace region gain of function mutant (H1196Y) and a CHD domain loss of function mutant (R552Q) (58) and made the CuSO₄-inducible pMT-constructs. Together with the different pMT-CENP-A constructs, we co-transfected those Mi-2 wild type and catalytic inactive/active pMT-constructs and created stable S2 cell lines. We then tested the localization behavior of overexpressed CENP-A when co-expressed with these different constructs. Even though the expression level of the Mi-2 constructs were highly variable (Supplementary Figure S3C), the effects on CENP-A localization on metaphase spreads were remarkable: CENP-A^{FL} was strongly misincorporated into chromatin arms upon co-expression of Mi-2 WT or the gain of function mutation H1196Y, similar to what has been observed for CENP-A misexpression without any co-expression. However, upon co-expression of ATPase inactive or the dominant-negative mutants, the levels of CENP-A misexpression on chromosome arms dropped significantly (Figure 6A and B). This suggests that the catalytic ATPase remodeling activity of Mi-2 plays an important role in CENP-A mislocalization to chromosome arms. Since the wild type Mi-2 was still present in these cells, we speculate that those mutants bear some dominant-negative effects that causes the observed phenotypes.

It was previously reported that NuRD complex components play a role in chromatin condensation (59). In order to rule this out, we compared the nuclear size in CENP-A overexpressing cells depleted for Mi-2 or MTA1-like with control (Bw)-depleted cells. We did not detect a significant

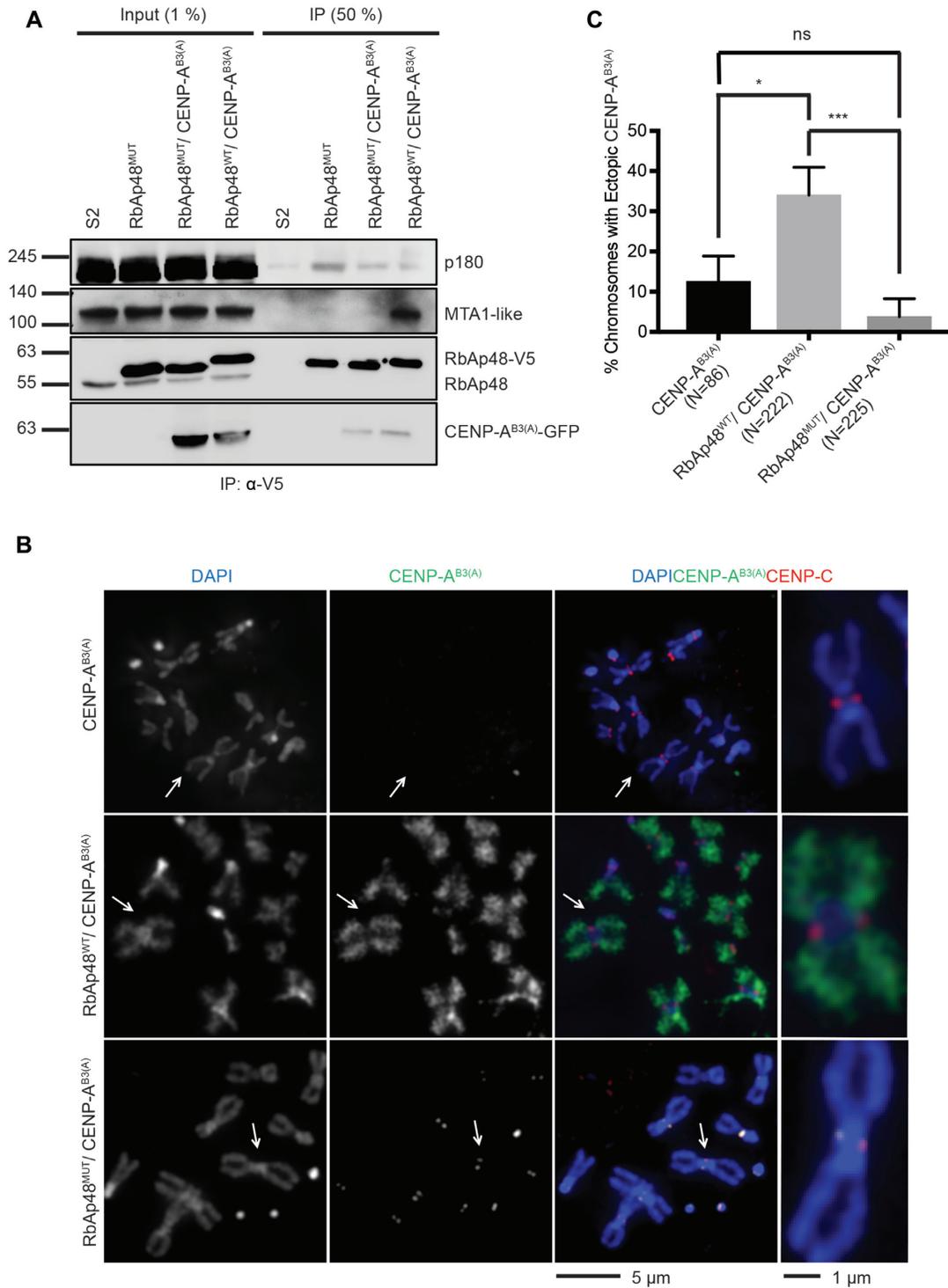


Figure 5. Ectopic CENP-A localization is abolished upon loss of NuRD interaction. (A) Western blot of cells expressing either RbAp48^{MUT} only, RbAp48^{MUT}/CENP-A^{B3(A)} or RbAp48^{WT}/CENP-A^{B3(A)} probed and immunoprecipitated with α-V5 and probed with NuRD complex components (α-Mi-2, α-MTA1-like, α-RbAp48 and α-CENP-A). Untransfected S2 cells served as control. (B) Representative images of mitotic chromosome spreads overexpressing CENP-A^{B3(A)}-GFP alone, CENP-A^{B3(A)}-GFP/RbAp48^{WT}-V5, or CENP-A^{B3(A)}-GFP/RbAp48^{MUT}-V5. DAPI in blue, α-V5 in green and α-CENP-C in red. Scale bars: 5 μm and 1 μm. (C) Quantification of percentage of chromosomes with ectopic CENP-A^{B3(A)} as shown in B. Each box represents the mean of three independent experiments, and the error bars indicate the SEM. Student's *t* test: * *P* < 0.05, *** *P* < 0.001 and ns: not significant.

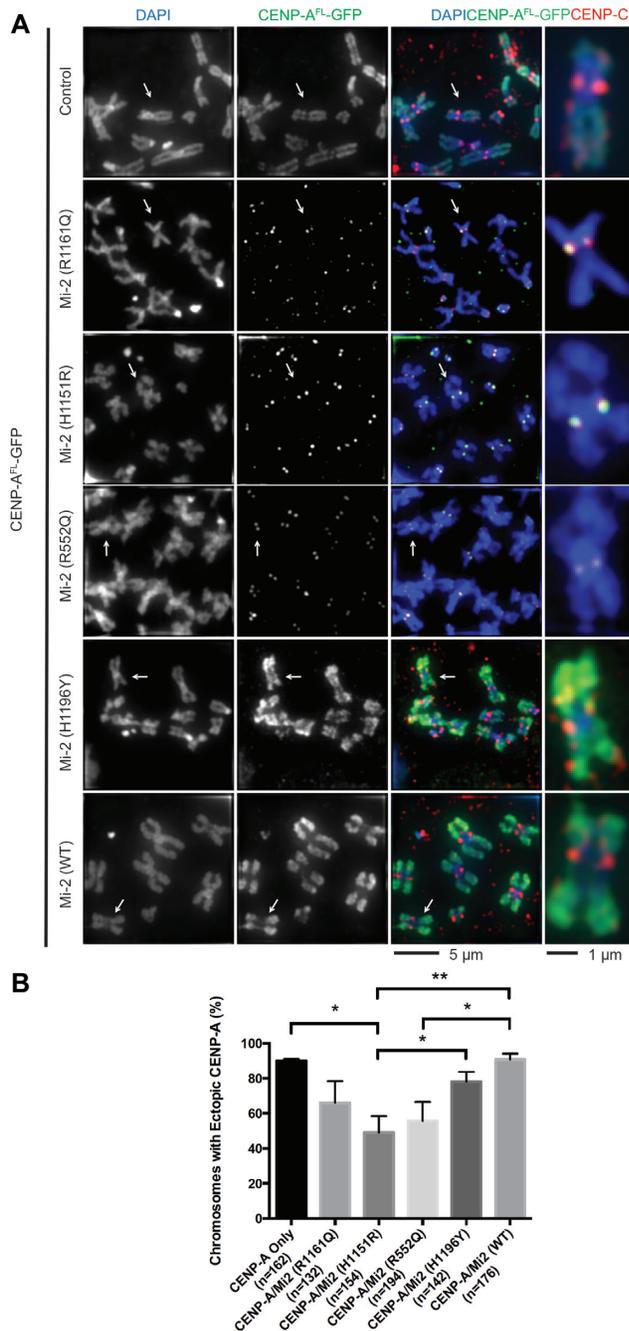


Figure 6. Mi-2 catalytic activity is required for CENP-A mislocalization. (A) Representative images of mitotic chromosome spreads co-expressing CENP-A^{FL}-GFP with catalytically inactive or active Mi-2-V5-His constructs as indicated. DAPI in blue, GFP in green and CENP-C in red. Scale bars: 5 μ m and 1 μ m. (B) Quantification of the percentage of chromosomes with ectopic CENP-A in cells co-expressing CENP-A^{FL}-GFP with catalytically inactive or active Mi-2-V5-His constructs. Each box represents the mean of three independent experiments. The error bars indicate the SEM. Student's *t* test: ** $P < 0.01$ and * $P < 0.05$.

difference in nuclear area upon knockdown of NuRD components (Supplementary Figure S5C) suggesting that the condensation of CENP-A misincorporated chromatin does not alter upon the depletion of NuRD components.

RbAp48 interaction with MTA1-like is required for ectopic CENP-A localization

We noticed that CENP-A^{B3(A)}-EGFP protein was comparably stable when co-expressed with mutant RbAp48 or wild type RbAp48 even though CENP-A^{B3(A)}-EGFP was virtually absent from ectopic sites on metaphase chromosomes (Figure 5A–C). We, therefore, analyzed the protein localization and stability of CENP-A^{B3(A)}-EGFP by immunofluorescence in cultured cells. CENP-A^{B3(A)}-EGFP was retained in the cytoplasm when overexpressed with mutant RbAp48 (Figure 7A and B) but did not get degraded by the proteasome as observed in CENP-A^{B3(A)}-EGFP without RbAp48 co-expression. The nuclear signal of CENP-A^{B3(A)}-EGFP was as low as in cells that do not co-express RbAp48 (Figure 7B). To analyze whether the lack of nuclear import depends on the association with the NuRD complex, we used live cell imaging of cells co-expressing CENP-A^{B3(A)}-EGFP and RbAp48-V5-His that were depleted of either Mi-2 or MTA1-like, and measured the rate of GFP nuclear localization over time. Interestingly, the depletion of MTA1-like caused a significant decrease in nuclear localization whereas the depletion of Mi-2 had no significant effect on CENP-A^{B3(A)}-EGFP localization even though the levels of Mi-2 and MTA1-like proteins were both significantly reduced (Figure 7C and D). Similarly, CENP-A^{FL}-V5-His nuclear localization was also reduced in Mi-2 or MTA1-like depleted cells (Supplementary Figure S5A, B and D). To test whether the difference is due to a defective import and proteasomal degradation, we performed live cell analysis in the same cells and added MG132 to inhibit the proteasome after RNA-depletion of NURD subunits. In the absence of MTA1-like, CENP-A^{B3(A)}-EGFP was, indeed, retained in the cytoplasm and degraded by the proteasome (Figure 7E and movies 14–16). We concluded from this set of experiments that the ectopic localization of CENP-A^{B3(A)}-EGFP requires the NuRD subunits RbAp48, Mi-2 and MTA1-like but that Mi-2 is dispensable for nuclear localization of CENP-A^{B3(A)}-EGFP.

The N-terminal Tail of CENP-A is required for NuRD-mediated ectopic localization

The conserved MTA1-like protein of the NuRD complex contains a SANT domain that is present in proteins implicated in binding of histone N-terminal tails (54). To investigate a potential role of the N-terminal tail of CENP-A in NuRD interaction, we examined the localization of the tail-deleted Δ N-CENP-A-EGFP or tail-containing CENP-A^{B3(A)}-EGFP upon co-expression of MTA1-like. Co-expression of MTA1-like perfectly rescued the nuclear localization of CENP-A^{B3(A)}-EGFP protein but not of the Δ N-CENP-A-EGFP (Figure 8A, Supplementary Figure S6A–C and movies 17–18). Importantly, mitotic chromosome spreads of Δ N-CENP-A-EGFP did not show any visible EGFP signal on chromosomes, whereas the CENP-A^{B3(A)}-EGFP readily incorporated into chromosome arms when MTA1-like was co-expressed (Figure 8B–C). To clarify whether a direct interaction of the N-terminal tail and MTA1-like is a likely reason for this difference in localization behavior, we co-immunoprecipitated MTA1-like with either

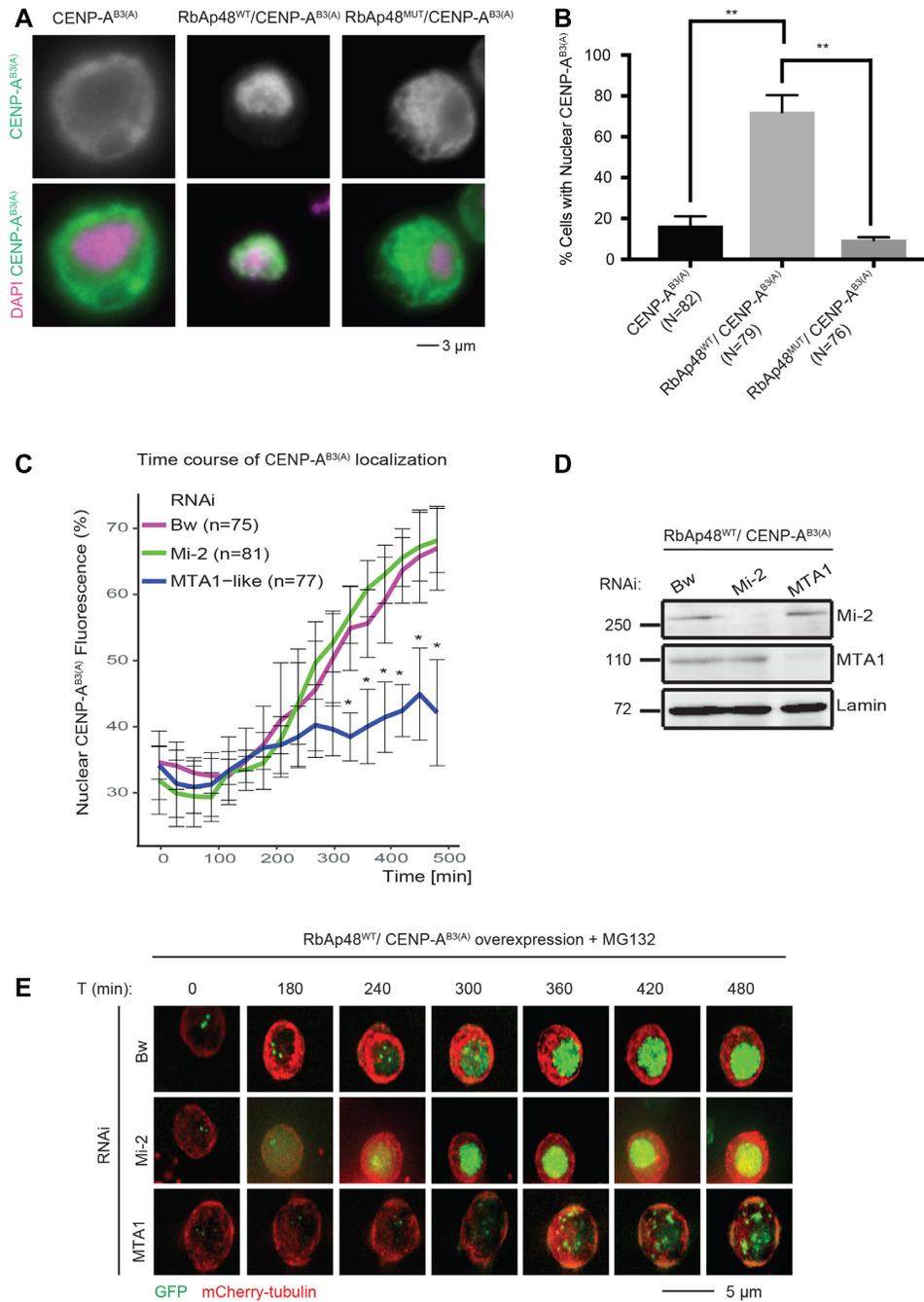


Figure 7. Nuclear localization of overexpressed CENP-A is abrogated upon loss of NuRD interaction or MTA1-like depletion. **(A)** Representative images (single image from a z-stack) of cells overexpressing CENP-A^{B3(A)}-GFP alone, CENP-A^{B3(A)}-GFP/RbAp48^{WT}-V5, or CENP-A^{B3(A)}-GFP/RbAp48^{MUT}-V5. DAPI in magenta, α -GFP in green. Scale bars: 3 μ m. **(B)** Quantification of percentage of cells with nuclear CENP-A^{B3(A)}. Each box represents the mean of three independent experiments, and the error bars indicate the SEM. Student's *t* test: ***P* < 0.01. **(C)** Quantification for percentage of nuclear CENP-A^{B3(A)} fluorescence compared to total cellular CENP-A^{B3(A)} fluorescence in time lapse microscopy of cells co-expressing CENP-A^{B3(A)}-GFP and RbAp48^{WT}-V5 under control (brown), Mi-2 or MTA1-like depletion by RNAi. Each single y-axis value represents the mean of three independent experiments, and the error bars indicate the SEM. Student's *t* test: **P* < 0.05. **(D)** Western blot of Mi-2 or MTA1-like knockdown upon CENP-A^{B3(A)}-GFP and RbAp48^{WT}-V5 co-overexpression (1 mM) probed with α -Mi-2, α -MTA1-like and α -lamin. Lamin served as loading control. **(E)** Stills from time lapse microscopy of S2 cells expressing mCherry-tubulin and CENP-A^{B3(A)}-GFP and RbAp48^{WT}-V5 under control (brown), Mi-2, or MTA1-like depletion by RNAi in the presence of MG132 (see supplemental movies 6–8). CuSO₄ was added 1 h prior to time point *t* = 0. GFP in green and mCherry-tubulin in red. Scale bar: 5 μ m.

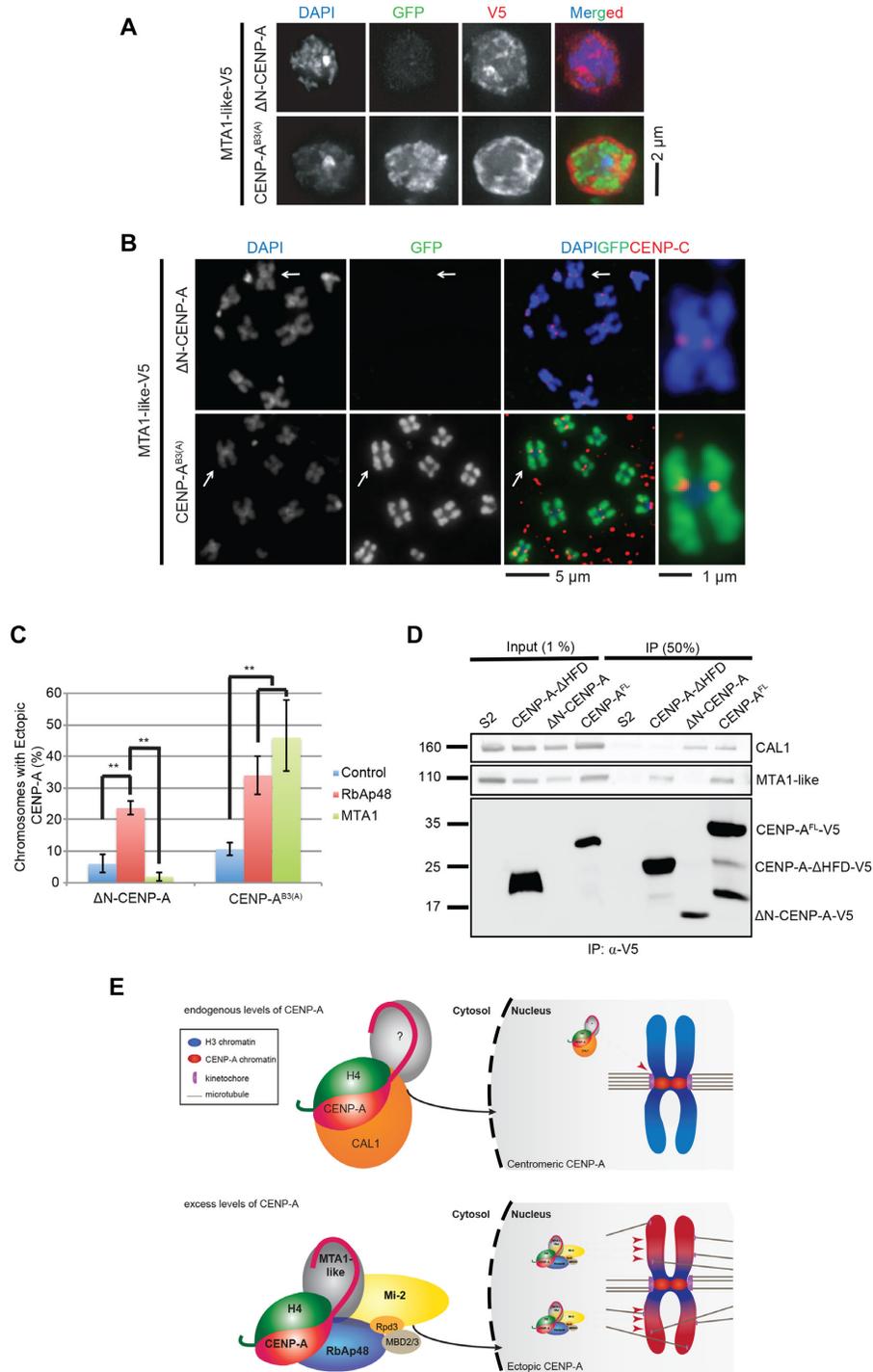


Figure 8. CENP-A interacts with MTA1-like through its N-tail and localizes to the nucleus and ectopic sites. **(A)** Representative images of cells co-expressing MTA1-like-V5-His either with Δ N-CENP-A-GFP or CENP-A^{B3(A)}-GFP. DAPI in blue, α -GFP in green and α -V5 in red. Scale bar: 2 μ m. **(B)** Representative images of mitotic chromosome spreads co-expressing either MTA1-like-V5/ Δ N-CENP-A-GFP or MTA1-like-V5/CENP-A^{B3(A)}-GFP. DAPI in blue, α -GFP in green and α -CENP-C in red. Scale bars: 5 μ m and 1 μ m. **(C)** Quantification of percentage of chromosomes with ectopic CENP-A in cells expressing either Δ N-CENP-A-GFP (control), CENP-A^{B3(A)}-GFP (control), MTA1-like-V5/ Δ N-CENP-A-GFP, MTA1-like-V5/CENP-A^{B3(A)}-GFP, RbAp48-V5/ Δ N-CENP-A-GFP, or RbAp48-V5/CENP-A^{B3(A)}-GFP. Each box represents the mean of three independent experiments, and the error bars indicate the SEM. Student's *t* test: ** *P* < 0.01. **(D)** Western blot of α -V5 co-immunoprecipitation from cells expressing either the N-terminus of CENP-A (CENP-A- Δ HFD-V5), Δ N-CENP-A-V5 or CENP-A^{FL}-V5 probed with α -CAL1 (positive control), α -MTA1-like, and α -V5. Untransfected S2 cells served as control. **(E)** Model for NuRD complex-mediated misincorporation of overexpressed CID. Pre-nucleosomal endogenous CENP-A associated with H4 is recognized by CAL1 through CENP-A HFD domain, escorted to the nucleus targeted exclusively to centromeres. A potential role of CENP-A N-terminal tail in this process is not characterized. Excess CENP-A interacts with the NuRD complex, imported into the nucleus and incorporated into chromatin with no affinity to centromeric regions. The interaction of the NuRD complex with CENP-A is mediated by RbAp48 and MTA1-like, which binds specifically to the N-terminal region of CENP-A.

CENP-A^{FL}, Δ N-CENP-A-V5 or only the tail of CENP-A with the histone-fold domain removed (CENP-A- Δ HFD-V5). The inducible expression of those constructs further showed that, unlike Δ N-CENP-A, CENP-A- Δ HFD is relatively stable and not degraded by the proteasome (Supplementary Figure S6D). Furthermore, CENP-A^{FL} co-immunoprecipitated MTA1-like efficiently but Δ N-CENP-A-V5 did not co-immunoprecipitate MTA1-like. Importantly, CENP-A- Δ HFD also co-immunoprecipitated MTA1-like even though weaker than CENP-A^{FL} (Figure 8D). Notably, we also detected a direct protein interaction between histone tail-binding MTA1-SANT domain and CENP-A N-tail by a yeast two hybrid assay (Supplementary Figure S6E). This set of experiments indicates that MTA1-like binds to the N-terminus of CENP-A and that the tail is at least to some extent important for efficient ectopic loading by an interaction with the histone tail binding protein MTA1-like.

DISCUSSION

We have presented evidence that the *Drosophila* N-terminus of CENP-A, and chromatin remodeling proteins which interact with it, are required for protein localization, stability and loading to non-centromeric chromatin. These findings indicate that the N-terminal tail of CENP-A is critical for non-centromeric incorporation but dispensable for centromeric localization as suggested previously (11,35). CENP-A lacking the N-terminus is not able to efficiently locate to the nuclear compartment most likely due to a missing nuclear localization signal. CENP-A protein that is retained in the cytoplasm is recognized by the ubiquitin proteasome machinery and efficiently degraded. Our co-expression analysis showed that the loss of the N-terminus can be tolerated in cells with an excess of CENP-A-interacting chaperones and reveals that CENP-A incorporation to different chromatin sites is controlled by the availability of those chaperones: RbAp48 facilitates ectopic CENP-A incorporation whereas CAL1 is the major driving force of focused centromere-specific CENP-A loading. We further showed that the RbAp48-containing NuRD complex is involved in non-centromeric loading that depends on an interaction with the long N-terminal tail of CENP-A.

The N-terminus of *Drosophila* CENP-A is required for protein localization.

Previous work in human, yeast and plant showed that the CENP-A N-terminus is implicated in the regulation of CENP-A (32,39,60); however, an impact on localization or stability was not described (11,35). Vermaak and co-workers analyzed a CENP-A truncation that still contained the B3 motif (deletion of aa 1–118) (35). The authors detected centromeric signals as we describe here for the Δ 118-CENP-A construct. Moreno-Moreno et al. deleted the entire N-terminus, however the analytical emphasis was on the quality of CENP-A localization (11). Since Δ 125-CENP-A exhibits weak centromere localization in a minority of cells, their transient transfection analysis could have led to them missing the observation that most transfected cells did not show any signal.

Tail deletion causes cytoplasmic retention of CENP-A most likely because of an incomplete nuclear localization signal (NLS). Several lines of evidence support this hypothesis: (a) sequence analysis suggests a bipartite NLS composed of the arginine-rich motif and residues KR 131/132, (b) mutation of the basic residues at position 119–121 mimics the distribution of Δ N-CENP-A, (c) B3-containing Δ 118-CENP-A restores centromeric signal and (d) fusion of an NLS-containing H3 tail to the CENP-A HFD restores nuclear signal. Histones encode their NLS in their N-terminal tails (46,47). The loss of a single tail or NLS in the H3–H4 histone dimer has been shown to be tolerated due to a compensatory function in nuclear import of the remaining tail/NLS (61). However, one tail can play a predominant role over the other as recently observed for *P. polycephalum* H4 (62). This is interesting since co-expressed H4 did not significantly rescue Δ N-CENP-A import, suggesting that the N-terminal tail of CENP-A might play the predominant role in CENP-A-H4 dimers. We, therefore, propose that the CENP-A N-terminal tail is the major regulator of nuclear localization, and that cytoplasmic CENP-A-H4 dimers may form oligomeric complexes, or, alternatively CENP-A-H4 may form tetramers that require at least one CENP-A tail for efficient nuclear import. The (CENP-A-H4)₂ tetrameric complex would explain our observation that full-length CENP-A can rescue Δ N-CENP-A import.

Tail-deleted CENP-A protein is trapped in the cytoplasm, degraded by the proteasome and accumulates in large aggregates when the proteasome is inhibited. We rule out misfolding as a potential cause of Δ N-CENP-A protein degradation because the tail-containing mutant CENP-A^{B3(A)}-EGFP behaves similarly to Δ N-CENP-A. Furthermore, the human CENP-A nucleosome high-resolution structure was obtained from crystals that contained an N-terminal tail-deleted version of CENP-A comparable to our Δ N-CENP-A construct with no consequences for crystal formation of the remaining globular part of CENP-A (63).

Histone chaperones regulate CENP-A destination and quantity

The two histone chaperones CAL1 and RbAp48 physically interact with CENP-A (24,30,48,64). Whether RbAp48 is also directly implicated in centromere deposition of CENP-A in *Drosophila* was unclear. Our experiments demonstrate that both chaperones can safeguard mutant CENP-A protein, facilitating its nuclear localization and preventing its premature degradation. However, we can exclude a major involvement of RbAp48 in centromeric CENP-A targeting: even though tail-mutated CENP-A protein amount and cytoplasmic mislocalization were efficiently rescued by RbAp48, centromere-specific targeting of CENP-A was not mediated by RbAp48. In contrast, the presence of elevated CAL1 focused CENP-A incorporation precisely to centromeric chromatin, suggesting that CAL1 is able to discriminate between bulk chromatin and centromeres and deposits CENP-A only at centromeric regions, which is consistent with its role as the CENP-A chaperone in *Drosophila* (24,25).

It was previously suggested that CAL1 not only acts as the rate-limiting CENP-A loading factor but may also con-

tain a ‘priming’ function for centromeric regions prior to CENP-A loading, similar to what has been suggested in human cells (48,65). Overexpressed CAL1 could then not only bring more CENP-A to centromeric regions but also provide larger regions competent for CENP-A incorporation when overexpressed. Overexpression of CENP-A without CAL1 co-expression leads to massive incorporation into ectopic sites along the chromosome arms. However, an increase of overexpressed CENP-A at centromeric regions is minimal when CAL1 is at endogenous levels (43). We conclude from our data that this ectopic incorporation is independent of CAL1 and the specificity of CENP-A for centromeres is mediated by its loading factors CAL1 without any intrinsic centromere specificity of CENP-A itself. Targeting more CENP-A to centromeres upon overexpression of CAL1 or localization of CENP-A primarily on chromosome arms upon overexpression of RbAp48 or MTA1-like suggests that CAL1 and NuRD compete for CENP-A loading, which is further supported by the observation that CENP-A centromeric localization is mostly bypassed in CAL1-depleted and RbAp48 overexpressing cells and predominantly loaded to ectopic sites.

NuRD as a regulator of CENP-A incorporation outside of centromeres

Our data indicate that RbAp48 interacts with ectopically localized CENP-A in association with the NuRD complex. The NuRD complex has been implicated in both, gene silencing but also gene activation. The complex has been identified as a regulator of developmental transcription as well as in cell growth regulation during DNA replication in S-phase (52). This is interesting because ectopic CENP-A in human cells enriches at sites of high histone exchange (19). It is, therefore, conceivable that insect and human cells display a similar preference for chromatin regions that require a high remodeling activity, and the NuRD complex has been implicated in regions that need to create a high density of nucleosomes (55).

We showed that the B3 motif of CENP-A, consisting of positively charged residues, functions as the nuclear localization sequence of CENP-A. Loss of nuclear localization leads to efficient proteasomal degradation of CENP-A. Consistently, arginine and lysine-rich motifs in the N-terminal tail of human CENP-A have been shown to be required for nuclear import, DNA contact and centromeric targeting in human cells (66). Interestingly, the N-terminus of CENP-A interacts with the NuRD complex subunit MTA1-like. MTA1-like contains BAH and SANT domains, which have been implicated in the recognition of nucleosomes or as readers of histone tails (67,68). Importantly, the minimal tail construct Δ 118-CENP-A that contains the NLS, but lacks the remaining N-terminus, entered the nucleus but did not localize ectopically, further suggesting that the N-terminal tail of CENP-A is required for NuRD-dependent ectopic incorporation.

RbAp48 has also been implicated in binding to H3 N-terminal tails (69). However, since we see an interaction with RbAp48 and a rescue of the Δ N-CENP-A phenotype, we suggest that RbAp48 has some preference for the histone fold domain of CENP-A. Mammalian CENP-A is

phosphorylated at S68 and ubiquitinated at K124, which regulate its binding to a prenucleosomal assembly complex (70,71). Similarly, MTA1 binding to H3 N-terminal tail is dependent on PTMs (72). The interaction of MTA1-like with the N-terminus of CENP-A may, therefore, depend on the presence or absence of PTMs on the N-terminal tail of CENP-A. This needs to be tested in the future.

Although we characterized the essential role of NuRD complex in CENP-A mislocalization, it is not clear whether a NuRD-mediated histone deacetylase function by Rpd3 is involved in this mechanism. Acetylation is known to be required for licensing of *de novo* CENP-A deposition, remodeling and transcription at centromeres (30,65,73–75). For instance, in the absence of HAT1-mediated H4 acetylation in human cells, the CENP-A loading factor HJURP fails to localize to centromeres, and CENP-A is misincorporated outside of centromeric chromatin (31). HAT1 has also been described to play a role in CENP-A loading in *Drosophila* (29) and HAT-dependent acetylation can be antagonized by Rpd3 deacetylation in JNK/AP-1-dependent transcription (76). Based on these studies, one may speculate that the level of histone acetylation of CENP-A nucleosomes affects the stability, localization specificity and perhaps their eviction rate from non-centromeric sites; however, further investigation will be required to elucidate the relation between deacetylation and CENP-A ectopic localization.

What will be interesting to determine is whether the observed NuRD-dependent CENP-A localization to non-centromeric chromatin reflects a physiological function of CENP-A. If there are specialized deposition pathways for non-centromeric CENP-A, we postulate that there are non-centromeric functions for chromatin-associated CENP-A. For instance, CENP-A has been shown to be essential in fully differentiated, mitotically inactive cells of the *Drosophila* midgut (77). In addition, physiological low levels of non-centromeric CENP-A have been described before and its presence may be dynamic during the cell cycle (8,41). This is further supported by our finding that CENP-A and NURD misexpression correlates in many different human cancer entities.

In human cells, the H3.3-specific histone chaperone DAXX has been implicated in non-centromeric CENP-A loading (19,21,78). Ectopic CENP-A nucleosomes have been suggested to be heterotypic containing one copy of histone H3.3 and one of CENP-A (19). Our live cell analysis showed that the fly ortholog DLP had no obvious effect on CENP-A loading, perhaps indicating a specific function of DAXX-mediated CENP-A loading and heterotypic CENP-A nucleosomes in humans. RbAp48 has been linked to CENP-A nucleosomes in other organisms and both, misexpression of RbAp48 and ectopic CENP-A incorporation were independently from each other linked to cancer formation in humans (15,74,79). We like to speculate that a deregulation of NuRD-mediated non-centromeric CENP-A may be detrimental for cells, and that it is worth elucidating the physiology of non-centromeric CENP-A loading in the future.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We are thankful to Ernest Laue and Wei Zhang for the RbAp48-mutant construct, Alexander Brehm for NuRD antibodies and Mi-2 constructs, ZMBH Mass Spec facility for help with Xlink-MS experiment and data analysis, Holger Lorenz for help with image analysis, Diana Doda for technical assistance, and the Erhardt lab members for suggestions and discussions. We thank Elmar Schiebel, Gary Karpen, Rebecca Wade, Mehmet Ali Ozturk and Alexander Brehm for discussions and Aubry Miller for critical comments on the manuscript. E.D., M.S.B. and D.B. were members of the HBIGS graduate school.

FUNDING

Deutsche Forschungsgemeinschaft [EXC81 (CellNetworks), ER576/2-2, SFB1036]; European Research Council [ERC-CoG- 682496] (cenRNA) (to S.E.). Funding for open access charge: European Research Council. *Conflict of interest statement.* None declared.

REFERENCES

- Allshire, R.C. and Karpen, G.H. (2008) Epigenetic regulation of centromeric chromatin: old dogs, new tricks? *Nat. Rev. Genet.*, **9**, 923–937.
- McKinley, K.L. and Cheeseman, I.M. (2016) The molecular basis for centromere identity and function. *Nat. Rev. Mol. Cell Biol.*, **17**, 16–29.
- Westhorpe, F.G. and Straight, A.F. (2013) Functions of the centromere and kinetochore in chromosome segregation. *Curr. Opin. Cell Biol.*, **25**, 334–340.
- 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.
- Talbert, P.B. and Henikoff, S. (2010) Histone variants—ancient wrap artists of the epigenome. *Nat. Rev. Mol. Cell Biol.*, **11**, 264–275.
- Mendiburo, M.J., Padeken, J., Fulop, S., Schepers, A. and Heun, P. (2011) Drosophila CENH3 is sufficient for centromere formation. *Science*, **334**, 686–690.
- Fukagawa, T. and Earnshaw, W.C. (2014) Neocentromeres. *Curr. Biol.*, **24**, R946–R947.
- Nechemia-Arbely, Y., Miga, K.H., Shoshani, O., Aslanian, A., McMahon, M.A., Lee, A.Y., Fachinetti, D., Yates, J.R. 3rd, Ren, B. and Cleveland, D.W. (2019) DNA replication acts as an error correction mechanism to maintain centromere identity by restricting CENP-A to centromeres. *Nat. Cell Biol.*, **21**, 743–754.
- Hewawasam, G., Shivaraju, M., Mattingly, M., Venkatesh, S., Martin-Brown, S., Florens, L., Workman, J.L. and Gerton, J.L. (2010) Psh1 is an E3 ubiquitin ligase that targets the centromeric histone variant Cse4. *Mol. Cell*, **40**, 444–454.
- Moreno-Moreno, O., Medina-Giro, S., Torras-Llort, M. and Azorin, F. (2011) The F box protein partner of paired regulates stability of Drosophila centromeric histone H3, CenH3(CID). *Curr. Biol.*, **21**, 1488–1493.
- Moreno-Moreno, O., Torras-Llort, M. and Azorin, F. (2006) Proteolysis restricts localization of CID, the centromere-specific histone H3 variant of Drosophila, to centromeres. *Nucleic Acids Res.*, **34**, 6247–6255.
- Collins, K.A., Furuyama, S. and Biggins, S. (2004) Proteolysis contributes to the exclusive centromere localization of the yeast Cse4/CENP-A histone H3 variant. *Curr. Biol.*, **14**, 1968–1972.
- Moreno-Moreno, O., Torras-Llort, M. and Azorin, F. (2019) The E3-ligases SCFPpa and APC/CCdh1 co-operate to regulate CENP-ACID expression across the cell cycle. *Nucleic Acids Res.*, **47**, 3395–3406.
- Zhang, W., Mao, J.H., Zhu, W., Jain, A.K., Liu, K., Brown, J.B. and Karpen, G.H. (2016) Centromere and kinetochore gene misexpression predicts cancer patient survival and response to radiotherapy and chemotherapy. *Nat. Commun.*, **7**, 12619.
- Tomonaga, T., Matsushita, K., Yamaguchi, S., Oohashi, T., Shimada, H., Ochiai, T., Yoda, K. and Nomura, F. (2003) Overexpression and mistargeting of centromere protein-A in human primary colorectal cancer. *Cancer Res.*, **63**, 3511–3516.
- Filipescu, D., Naughtin, M., Podsypanina, K., Lejour, V., Wilson, L., Gurard-Levin, Z.A., Orsi, G.A., Simeonova, I., Toufektchan, E., Attardi, L.D. *et al.* (2017) Essential role for centromeric factors following p53 loss and oncogenic transformation. *Genes Dev.*, **31**, 463–480.
- Shrestha, R.L., Ahn, G.S., Staples, M.I., Sathyan, K.M., Karpova, T.S., Foltz, D.R. and Basrai, M.A. (2017) Mislocalization of centromeric histone H3 variant CENP-A contributes to chromosomal instability (CIN) in human cells. *Oncotarget*, **8**, 46781–46800.
- 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.
- Lacoste, N., Woolfe, A., Tachiwana, H., Gareau, 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.
- Barnhart, M.C., Kuich, P.H., Stellfox, M.E., Ward, J.A., Bassett, E.A., Black, B.E. and Foltz, D.R. (2011) HJURP is a CENP-A chromatin assembly factor sufficient to form a functional de novo kinetochore. *J. Cell Biol.*, **194**, 229–243.
- Nye, J., Sturgill, D., Athwal, R. and Dalal, Y. (2018) HJURP antagonizes CENP-A mislocalization driven by the H3.3 chaperones HIRA and DAXX. *PLoS One*, **13**, e0205948.
- Hewawasam, G.S., Dhatchinamoorthy, K., Mattingly, M., Seidel, C. and Gerton, J.L. (2018) Chromatin assembly factor-1 (CAF-1) chaperone regulates Cse4 deposition into chromatin in budding yeast. *Nucleic Acids Res.*, **46**, 4440–4455.
- Nye, J., Melters, D.P. and Dalal, Y. (2018) The Art of War: harnessing the epigenome against cancer [version 1; peer review: 4 approved]. *FI000Res*, **7**, 141.
- 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.
- Chen, C.-C., Dechassa, M.L., Bettini, E., Ledoux, M.B., Belisario, C., Heun, P., Luger, K. and Mellone, B.G. (2014) CAL1 is the Drosophila CENP-A assembly factor. *J. Cell Biol.*, **204**, 313–329.
- Clapier, C.R., Iwasa, J., Cairns, B.R. and Peterson, C.L. (2017) Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes. *Nat. Rev. Mol. Cell Biol.*, **18**, 407–422.
- Clapier, C.R. and Cairns, B.R. (2009) The biology of chromatin remodeling complexes. *Annu. Rev. Biochem.*, **78**, 273–304.
- Hoek, M. and Stillman, B. (2003) Chromatin assembly factor 1 is essential and couples chromatin assembly to DNA replication in vivo. *Proc. Natl. Acad. Sci. U.S.A.*, **100**, 12183–12188.
- Boltengagen, M., Huang, A., Boltengagen, A., Trixl, L., Lindner, H., Kremser, L., Offterdinger, M. and Lusser, A. (2016) A novel role for the histone acetyltransferase Hat1 in the CENP-A/CID assembly pathway in Drosophila melanogaster. *Nucleic Acids Res.*, **44**, 2145–2159.
- Furuyama, T., Dalal, Y. and Henikoff, S. (2006) Chaperone-mediated assembly of centromeric chromatin in vitro. *Proc. Natl. Acad. Sci. U.S.A.*, **103**, 6172–6177.
- Shang, W.H., Hori, T., Westhorpe, F.G., Godek, K.M., Toyoda, A., Misu, S., Monma, N., Ikeo, K., Carroll, C.W., Takami, Y. *et al.* (2016) Acetylation of histone H4 lysine 5 and 12 is required for CENP-A deposition into centromeres. *Nat. Commun.*, **7**, 13465.
- Black, B.E., Jansen, L.E., Maddox, P.S., Foltz, D.R., Desai, A.B., Shah, J.V. and Cleveland, D.W. (2007) Centromere identity maintained by nucleosomes assembled with histone H3 containing the CENP-A targeting domain. *Mol. Cell*, **25**, 309–322.
- Maddox, P.S., Corbett, K.D. and Desai, A. (2011) Structure, assembly and reading of centromeric chromatin. *Curr. Opin. Genet. Dev.*, **22**, 139–147.
- Malik, H.S. and Henikoff, S. (2003) Phylogenomics of the nucleosome. *Nat. Struct. Biol.*, **10**, 882–891.

35. Vermaak,D., Hayden,H.S. and Henikoff,S. (2002) Centromere targeting element within the histone fold domain of Cid. *Mol. Cell Biol.*, **22**, 7553–7561.
36. Sekulic,N. and Black,B.E. (2012) Molecular underpinnings of centromere identity and maintenance. *Trends Biochem. Sci.*, **37**, 220–229.
37. Chen,Y., Baker,R.E., Keith,K.C., Harris,K., Stoler,S. and Fitzgerald-Hayes,M. (2000) The N terminus of the centromere H3-like protein Cse4p performs an essential function distinct from that of the histone H3 fold domain. *Mol. Cell Biol.*, **20**, 7037–7048.
38. Torras-Llort,M., Medina-Giro,S., Moreno-Moreno,O. and Azorin,F. (2010) A conserved arginine-rich motif within the hypervariable N-domain of Drosophila centromeric histone H3 (CenH3) mediates BubR1 recruitment. *PLoS One*, **5**, e13747.
39. 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.
40. Ravi,M., Shibata,F., Ramahi,J.S., Nagaki,K., Chen,C., Murata,M. and Chan,S.W. (2011) Meiosis-specific loading of the centromere-specific histone CENH3 in Arabidopsis thaliana. *PLoS Genet.*, **7**, e1002121.
41. 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.
42. Shukla,M., Tong,P., White,S.A., Singh,P.P., Reid,A.M., Catania,S., Pidoux,A.L. and Allshire,R.C. (2018) Centromere DNA Destabilizes H3 nucleosomes to promote CENP-A deposition during the cell cycle. *Curr. Biol.*, **28**, 3924–3936.
43. 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.
44. Malik,H.S., Vermaak,D. and Henikoff,S. (2002) Recurrent evolution of DNA-binding motifs in the Drosophila centromeric histone. *Proc. Natl. Acad. Sci. U.S.A.*, **99**, 1449–1454.
45. Varshavsky,A. (2019) N-degron and C-degron pathways of protein degradation. *Proc. Natl. Acad. Sci. U.S.A.*, **116**, 358–366.
46. Morgan,B.A., Mittman,B.A. and Smith,M.M. (1991) The highly conserved N-terminal domains of histones H3 and H4 are required for normal cell cycle progression. *Mol. Cell Biol.*, **11**, 4111–4120.
47. Schuster,T., Han,M. and Grunstein,M. (1986) Yeast histone H2A and H2B amino termini have interchangeable functions. *Cell*, **45**, 445–451.
48. Mellone,B.G., Grive,K.J., Shteyn,V., Bowers,S.R., Oderberg,I. and Karpen,G.H. (2011) Assembly of Drosophila centromeric chromatin proteins during mitosis. *PLoS Genet.*, **7**, e1002068.
49. Klockenbusch,C. and Kast,J. (2010) Optimization of formaldehyde cross-linking for protein interaction analysis of non-tagged integral beta1. *J. Biomed. Biotechnol.*, **2010**, 927585.
50. Barth,T.K., Schade,G.O., Schmidt,A., Vetter,I., Wirth,M., Heun,P., Thomae,A.W. and Imhof,A. (2014) Identification of novel Drosophila centromere-associated proteins. *Proteomics*, **14**, 2167–2178.
51. Deyter,G.M. and Biggins,S. (2014) The FACT complex interacts with the E3 ubiquitin ligase Psh1 to prevent ectopic localization of CENP-A. *Genes Dev.*, **28**, 1815–1826.
52. Bouazoune,K. and Brehm,A. (2006) ATP-dependent chromatin remodeling complexes in Drosophila. *Chromosome Res.*, **14**, 433–449.
53. Bade,D., Pauleau,A.L., Wendler,A. and Erhardt,S. (2014) The E3 ligase CUL3/RDX controls centromere maintenance by ubiquitinating and stabilizing CENP-A in a CAL1-dependent manner. *Dev. Cell*, **28**, 508–519.
54. Boyer,L.A., Langer,M.R., Crowley,K.A., Tan,S., Denu,J.M. and Peterson,C.L. (2002) Essential role for the SANT domain in the functioning of multiple chromatin remodeling enzymes. *Mol. Cell*, **10**, 935–942.
55. Lai,A.Y. and Wade,P.A. (2011) Cancer biology and NuRD: a multifaceted chromatin remodelling complex. *Nat. Rev. Cancer*, **11**, 588–596.
56. Toh,Y. and Nicolson,G.L. (2009) The role of the MTA family and their encoded proteins in human cancers: molecular functions and clinical implications. *Clin. Exp. Metast.*, **26**, 215–227.
57. Alqarni,S.S., Murthy,A., Zhang,W., Przewloka,M.R., Silva,A.P., Watson,A.A., Lejon,S., Pei,X.Y., Smits,A.H., Kloet,S.L. *et al.* (2014) Insight into the architecture of the NuRD complex: structure of the RbAp48-MTA1 subcomplex. *J. Biol. Chem.*, **289**, 21844–21855.
58. Kovac,K., Sauer,A., Macinkovic,I., Awe,S., Finkernagel,F., Hoffmeister,H., Fuchs,A., Muller,R., Rathke,C., Langst,G. *et al.* (2018) Tumour-associated missense mutations in the dMi-2 ATPase alters nucleosome remodelling properties in a mutation-specific manner. *Nat. Commun.*, **9**, 2112.
59. Fasulo,B., Deuring,R., Murawska,M., Gause,M., Dorighi,K.M., Schaaf,C.A., Dorsett,D., Brehm,A. and Tamkun,J.W. (2012) The Drosophila Mi-2 chromatin-remodeling factor regulates higher-order chromatin structure and cohesin dynamics In Vivo. *PLoS Genet.*, **8**, e1002878.
60. Ravi,M., Kwong,P.N., Menorca,R.M., Valencia,J.T., Ramahi,J.S., Stewart,J.L., Tran,R.K., Sundaresan,V., Comai,L. and Chan,S.W. (2010) The rapidly evolving centromere-specific histone has stringent functional requirements in Arabidopsis thaliana. *Genetics*, **186**, 461–471.
61. Blackwell,J.S. Jr., Wilkinson,S.T., Mosammaparast,N. and Pemberton,L.F. (2007) Mutational analysis of H3 and H4 N termini reveals distinct roles in nuclear import. *J. Biol. Chem.*, **282**, 20142–20150.
62. Ejllassi-Lassalette,A., Mocquard,E., Arnaud,M.C. and Thiriet,C. (2011) H4 replication-dependent diacetylation and Hat1 promote S-phase chromatin assembly in vivo. *Mol. Biol. Cell*, **22**, 245–255.
63. 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.
64. Schittenhelm,R.B., Althoff,F., Heidmann,S. and Lehner,C.F. (2010) Detrimental incorporation of excess Cenp-A/Cid and Cenp-C into Drosophila centromeres is prevented by limiting amounts of the bridging factor Cal1. *J. Cell Sci.*, **123**, 3768–3779.
65. 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.
66. Jing,R., Xi,J., Leng,Y., Chen,W., Wang,G., Jia,W., Kang,J. and Zhu,S. (2017) Motifs in the amino-terminus of CENP-A are required for its accumulation within the nucleus and at the centromere. *Oncotarget*, **8**, 40654–40667.
67. Schmidberger,J.W., Sharifi Tabar,M., Torrado,M., Silva,A.P., Landsberg,M.J., Brillault,L., AlQarni,S., Zeng,Y.C., Parker,B.L., Low,J.K. *et al.* (2016) The MTA1 subunit of the nucleosome remodeling and deacetylase complex can recruit two copies of RBBP4/7. *Protein Sci.*, **25**, 1472–1482.
68. Torchy,M.P., Hamiche,A. and Klaholz,B.P. (2015) Structure and function insights into the NuRD chromatin remodeling complex. *Cell Mol. Life Sci.*, **72**, 2491–2507.
69. Millard,C.J., Varma,N., Saleh,A., Morris,K., Watson,P.J., Bottrill,A.R., Fairall,L., Smith,C.J. and Schwabe,J.W. (2016) The structure of the core NuRD repression complex provides insights into its interaction with chromatin. *Elife*, **5**, e13941.
70. Yu,Z., Zhou,X., Wang,W., Deng,W., Fang,J., Hu,H., Wang,Z., Li,S., Cui,L., Shen,J. *et al.* (2015) Dynamic phosphorylation of CENP-A at Ser68 orchestrates its cell-cycle-dependent deposition at centromeres. *Dev. Cell*, **32**, 68–81.
71. Niikura,Y., Kitagawa,R., Ogi,H., Abdulle,R., Pagala,V. and Kitagawa,K. (2015) CENP-A K124 Ubiquitylation Is Required for CENP-A Deposition at the Centromere. *Dev. Cell*, **32**, 589–603.
72. Nair,S.S., Li,D.Q. and Kumar,R. (2013) A core chromatin remodeling factor instructs global chromatin signaling through multivalent reading of nucleosome codes. *Mol. Cell*, **49**, 704–718.
73. Bergmann,J.H., Jakubsche,J.N., Martins,N.M., Kagansky,A., Nakano,M., Kimura,H., Kelly,D.A., Turner,B.M., Masumoto,H., Larionov,V. *et al.* (2012) Epigenetic engineering: histone H3K9 acetylation is compatible with kinetochore structure and function. *J. Cell Sci.*, **125**, 411–421.
74. Hayashi,T., Fujita,Y., Iwasaki,O., Adachi,Y., Takahashi,K. and Yanagida,M. (2004) Mis16 and Mis18 are required for CENP-A loading and histone deacetylation at centromeres. *Cell*, **118**, 715–729.
75. Chen,C.C., Bowers,S., Lipinszki,Z., Palladino,J., Trusiak,S., Bettini,E., Rosin,L., Przewloka,M.R., Glover,D.M., O'Neill,R.J. *et al.* (2015) Establishment of centromeric chromatin by the CENP-A assembly factor CAL1 requires FACT-mediated transcription. *Dev. Cell*, **34**, 73–84.

76. Miotto, B., Sagnier, T., Berenger, H., Bohmann, D., Pradel, J. and Graba, Y. (2006) Chameau HAT and DRpd3 HDAC function as antagonistic cofactors of JNK/AP-1-dependent transcription during *Drosophila* metamorphosis. *Genes Dev.*, **20**, 101–112.
77. Garcia Del Arco, A., Edgar, B.A. and Erhardt, S. (2018) In vivo analysis of centromeric proteins reveals a stem Cell-Specific asymmetry and an essential role in differentiated, Non-proliferating cells. *Cell Rep.*, **22**, 1982–1993.
78. Athwal, R.K., Walkiewicz, M.P., Baek, S., Fu, S., Bui, M., Camps, J., Ried, T., Sung, M.-H. and Dalal, Y. (2015) CENP-A nucleosomes localize to transcription factor hotspots and subtelomeric sites in human cancer cells. *Epigenet. Chromatin*, **8**, 2.
79. Kong, L., Yu, X.P., Bai, X.H., Zhang, W.F., Zhang, Y., Zhao, W.M., Jia, J.H., Tang, W., Zhou, Y.B. and Liu, C.J. (2007) RbAp48 is a critical mediator controlling the transforming activity of human papillomavirus type 16 in cervical cancer. *J. Biol. Chem.*, **282**, 26381–26391.
80. Fields, S. and Song, O.-K. (1989) A novel genetic system to detect protein–protein interactions. *Nature*, **340**, 245–246.
81. Knop, M., Siegers, K., Pereira, G., Zachariae, W., Winsor, B., Nasmyth, K. and Schiebel, E. (1999) Epitope tagging of yeast genes using a PCR-based strategy: more tags and improved practical routines. *Yeast*, **15**, 963–972.
82. Rothbauer, U., Zolghadr, K., Muyldermans, S., Schepers, A., Cardoso, M.C. and Leonhardt, H. (2008) A versatile nanotrap for biochemical and functional studies with fluorescent fusion proteins. *Mol. Cell Proteomics*, **7**, 282–289.
83. Cox, J. and Mann, M. (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.*, **26**, 1367–1372.
84. Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M.Y., Geiger, T., Mann, M. and Cox, J. (2016) The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat. Methods*, **13**, 731–740.