HACKING THE CENTROMERE CHROMATIN CODE:

dissecting the epigenetic regulation of centromere identity

Jan H. Bergmann

Thesis Submitted for the Degree of Doctor of Philosophy

The University of Edinburgh

July 2010

I declare that this thesis is composed entirely by myself, and that the work presented herein is my own, with contributions made by others being clearly indicated.

Jan Bergmann

Edinburgh

July 2010

I thank my father, whose pursuit of knowledge and insight, determined spirit and imperturbable approach to life has been an inspiration that forms the fundamental basis for the person I am today.

Acknowledgements

I would like to sincerely thank Bill for giving me the opportunity and the freedom to pursue my research interests in a highly stimulating and exciting environment. Most importantly, I am grateful for his continuous and devoted support and valuable feedback both on my scientific and personal progress. I am grateful for his guidance through these years, and could not have wished for a better supervisor and Doktorvater.

I am grateful to Robin and Irina, who as part of my thesis panel have offered stimulating discussions and valuable suggestions.

All the members of the Earnshaw and Schirmer labs, past and present, for their help, support, suggestions and friendship, for making life in and out of work ever so much more enjoyable, and whom I will miss greatly. In particular, I would like to thank Melpi, for our serene and fun conversations, and for her kindly refilling my coffee mug throughout each day.

I thank Laura, for the fantastic character she is, and for her friendship that I value greatly and would not want to miss, as it has in so many facets enriched my life within this past year.

My mother, for never losing her faith in me, and for supporting me in every decision I made. Who has bestowed on me her values of life, and who has fundamentally provided me with the will and strength to pursue this career.

Words cannot describe the happiness that fills my heart for having Kate in my life. I thank her for being with me in all those moments of joy, and standing by my side in times of despair. I could not wish for anything more than sharing my future with her.

Abstract

The centromere is a specialized chromatin domain that serves as the assembly site for the mitotic kinetochore structure, thereby playing a fundamental role in facilitating the maintenance of the genetic information. A histone H3 variant termed CENP-A is specifically found at all active centromeres. Beyond this, however, little is known about how and to which extent the chromatin environment of centromeres modulates and contributes towards centromere identity and function.

Here, I have employed a novel Human Artificial Chromosome (HAC), the centromere of which can be targeted by fusions to the tet repressor, to characterize the chromatin environment underlying active kinetochores, as well as to specifically probe the role of this environment in the maintenance of kinetochore structure and function. My data demonstrate that centromeric chromatin resembles the downstream regions of actively transcribed genes. This includes the previously unrecognized presence of histone H3 nucleosomes methylated at lysine 36 within the chromatin underlying functional kinetochores. Targeted manipulation of this chromatin through tethering of a heterochromatin-seeding transcriptional repressor results in the inactivation of HAC kinetochore function concomitant with a hierarchical disassembly of the structure. Through an even more selective engineering of the HAC centromere chromatin, I have provided evidence supporting a critical role for nucleosomes dimethylated at lysine 4 on histone H3 in facilitating local transcription of the underlying DNA.

Tethering of different chromatin-modifying activities into the HAC kinetochore collectively reveals a critical role for both, histone H3 dimethylated on lysine 4 and low-level, non-coding transcription in the maintenance of the CENP-A chromatin domain. On one hand, repression of centromeric transcription negatively correlates with the maintenance of

CENP-A and ultimately results in the loss of kinetochore function. On the other hand, increasing kinetochore-associated RNA polymerase activity to within physiological levels for euchromatin is associated with rapid loss of CENP-A from the HAC centromere. Together, my data point towards the requirement for a delicate balance of transcriptional activity that is required to shape and maintain the chromatin environment of active centromeres.

Abbreviations

BAC	bacterial artificial chromosome
bp	base pair(s)
BSA	bovine serum albumine
BSr	blasticidin S resistance
CATD	centromere targeting domain
CCAN	constitutive centromere-associated network
cDNA	complementary DNA
CENP	centromere protein
ChIP	chromatin immunoprecipitation
CIP	calf intestinal phosphatase, alkaline
cRPMI	complete RPMI medium 1640
DAPI	4’.6’-diamidino-2-phenylindole
DNA	deoxyribonucleic acid
dNTP	deoxynucleotidetriphosphate
Dox	doxycycline
dsDNA	double-stranded deoxyribonucleic acid
EDTA	ethylenediaminetetraacetic acid
EYFP	enhanced yellow fluorescent protein
FISH	fluorescence *in situ* hybridization
FITC	fluorescein isothiocyanate
HAC	human artificial chromosome
HP1	heterochromatin hrotein 1
KAP1	krueppel-associated protein 1
kb	kilobase(s)
KRAB	krueppel-associated box
LB	lysogeny broth (alt.: Luria Bertani broth)
LSD1	lysine-specific demethylase 1
M	molar
mg	milligram

ml	millilitre
mM	milimolar
NEB	New England Biolabs
nM	nanomolar
OD	optical density
ORF	open reading frame
PBS	phosphate-buffered saline
PBS-T	PBS/Tween
PCR	polymerase chain reaction (alt.: pure chance reaction)
PMSF	phenylmethylsulfonyl fluoride
PHD	plant homeodomain
RBCC	ring-B-box-coiled-coil
RNA	ribonucleic acid
PFA	para-formaldehyde
RT-PCR	reverse transcription polymerase chain reaction
SDS	sodium-dodecyl sulfate
sec	second(s)
TE	Tris-EDTA buffer
U	(enzyme) units
UV	ultra violet
µg	microgram
µl	microlitre
µM	micromolar

Contents

Introduction

A CHROMOSOMES, MITOSIS AND KINETOCHORES: AN OVERVIEW

A.1 The Packaging of DNA in Chromosomes

DNA is the primary carrier of the genetic information of an organism. In humans, the combined length of a cell's DNA amounts to about two meters. Incorporation of this material into nuclei with diameters of about 5-10 µm requires a tremendous extent of packaging in the form of chromosomes. Chromosomes are highly organized structures made up of chromatin, comprising DNA, RNA and a large amount of structural and non-structural proteins. The basic packaging unit of DNA within chromosomes is formed by the nucleosome, comprising of a compact octamer with two copies each of the core histones H3, H4, H2A and H2B, wrapped by 146 bp of DNA (Luger et al., 1997). As discussed below, the nucleosome is the primary target for a large variety of post-translational modifications involved in the regulation of chromatin structure and function.

A series of additional, poorly defined higher-order packaging arrangements occurs to achieve the highest compaction of DNA in the form of the mitotic chromosome. A first step was considered to be the assembly of nucleosomes into a 30 nm fibre, facilitated by the linker histone H1. However, the concept of a structurally ordered 30 nm fibre remains a matter of debate (Maeshima and Eltsov, 2008). Additional compaction is achieved through the looping of chromatin fibres along a central lattice termed the chromosome scaffold, formed largely by non-histone proteins such as topoisomerase II and condensin (Hudson et al., 2009; Swedlow and Hirano, 2003). It is notable to point out that mitotic chromosomes themselves display variable compaction rates in interphase and across different stages of mitosis (Lleres

et al., 2009), emphasizing the obvious requirement for a dynamic regulation of the chromatin state to accommodate various chromatin-associated processes and structural needs associated with different stages of the cell cycle.

A.2 The Stages of Mitosis

Actively dividing cells face a considerable challenge, in that they need not only to faithfully replicate the genetic information associated with their chromosomes, but also to ensure that, in the process of cell division, each daughter cell inherits exactly the same complement of this genetic information. Errors in the regulation or mechanics of the process of sister chromatid segregation cause aneuploidy, which frequently results in apoptosis, and is further implicated in genetic abnormalities associated with birth defects and cancer (Kops et al., 2005; Storchova and Pellman, 2004).

Based on cytological analysis of cellular and chromatin structure, mitosis can be readily divided into five distinct phases. These mitotic phases are termed prophase, prometaphase, metaphase, anaphase and telophase, with the latter being concomitant with cytokinesis. Most aspects of mitosis are regulated through a highly complex interplay of various kinases and phosphatases that frequently trigger cataclysmic all-or-nothing cascades mediating the progression from one mitotic state to the next (reviewed in (Nigg, 2001; Sullivan and Morgan, 2007)).

Prophase

At the onset of prophase, the largely amorphous chromatin characteristic of the interphase cell begins to condense into the highly-compacted mitotic chromosomes. These are formed by two sister chromatids, the result of previous DNA replication during S phase. The sister

chromatids remain attached largely along the chromosome arms and their centromeres, the latter being mediated by pericentromeric heterochromatin containing heterochromatin protein (HP)1 and associated cohesin complexes (Grewal and Jia, 2007; Pidoux and Allshire, 2005). The centrosome, which is the organizing centre for the cell's microtubule cytoskeleton, is replicated early in interphase. In parallel to chromosome condensation, the centrosomes begin to move to opposite sites of the cell in preparation for the assembly of a bi-polar mitotic microtubule spindle.

Prometaphase

The initiating event of prometaphase is the disassembly of the nuclear envelope, which until this point served the compartmentalization of the chromatin and the separation of chromatin-associated processes from the cytoplasm and *vice versa*. The loss of this barrier allows microtubules emanating from the centrosomes to extend into the chromosomal mass. These microtubules are highly dynamic, rapidly alternating between polymerization and de-polymerization in the absence of stabilizing end-on contacts of the microtubule plus ends (Desai and Mitchison, 1997). It is this dynamic mode of the microtubules that eventually allows them to establish and maintain contact with kinetochores of individual sister chromatids. A combination of microtubule pulling forces and lateral sliding of sister chromatids mediated by kinetochore-associated motor proteins results in a net movement and accumulation of chromosomes at the centre of the cell between the two centrosomes. Concomitantly, non-kinetochore microtubules of opposing centrosomes engage in anti-parallel interactions, resulting in the establishment of an anti-parallel microtubule array forming the backbone of the mitotic spindle.

During late prometaphase, it becomes increasingly critical for both sister kinetochores to be attached to microtubule bundles emanating from

opposite centrosomes in order to achieve bi-orientation, while eliminating both syntelic and merotelic attachments in which both sister kinetochores are attached to the same pole or in which a single kinetochore is bound by microtubules originating from opposing poles, respectively. A central factor in this process is the chromosomal passenger complex containing Aurora B kinase, which at this mitotic stage localizes to the centromeric chromatin between sister kinetochores. It is believed that a combination of Aurora B kinase-mediated control of outer kinetochore components required for regulating microtubule interaction, as well as spatial separation of Aurora B kinase from its substrates upon establishment of microtubule pulling forces, are key to achieving bi-orientation (Liu et al., 2009; Ruchaud et al., 2007).

Metaphase

As more and more individual microtubules establish contact with kinetochores, the pulling forces on sister kinetochores resulting from oppositely anchored microtubule bundles, now termed k-fibres, result in the alignment of all chromosomes at an equidistant position between the two centrosomes. This structure is termed the metaphase plate. It is also in metaphase that an essential, kinetochore-associated process ensures all kinetochores are attached to microtubules prior to anaphase onset. In a molecular manner that is not yet entirely resolved, the “spindle assembly checkpoint” is active at unattached kinetochores, and prevents global activation of the anaphase promoting complex until all chromosomes have achieved bi-orientation (Musacchio and Salmon, 2007; Nasmyth, 2005).

Anaphase

Upon satisfaction of the spindle assembly checkpoint, the anaphase promoting complex is activated and targets key substrates, including securin for proteasomal degradation. Until this point, securin maintained separase in an inactive form, and degradation of securin consequently results in the

accumulation of active separase. Separase in turn cleaves cohesin, thereby allowing segregation of sister chromatids subject to k-fibre pulling forces resulting from microtubule de-polymerization. INCENP, another component of the chromosomal passenger complex, translocates from kinetochores to the spindle midzone where it is believed to facilitate stability of the non-kinetochore microtubule bundle (McCollum, 2004; Ruchaud et al., 2007). Towards the later stage of anaphase, together with other force-generating events, these non-kinetochore microtubules are thought to be involved in mediating the pushing-apart of centrosomes into opposite directions, thereby further increasing the distance between the two chromatid masses.

Telophase and Cytokinesis

Key events in telophase are the reversal of chromosome compaction and the re-establishment of the nuclear envelope around the chromosomal mass. Neither process is particularly well understood. Concomitant with telophase occurs the establishment and subsequent ingression of the cleavage furrow mediated by a contractile ring formed from actin filaments and myosin. Centrally positioned between the two daughter chromatin masses, the cleavage furrow ingresses to form the midbody structure resulting in the subsequent abscission of the cell membrane and the completion of cytokinesis with the establishment of two separate cells.

A.3 The Kinetochore

The kinetochore is a complex proteinacious structure that assembles on the centromeric chromatin of each sister chromatid and acts as hub for a variety of critical mitotic events, most importantly the capture of and the attachment to spindle microtubules. The kinetochore contains microtubule-binding and motor proteins, coordinates and facilitates movement along spindle microtubules, and serves as a signalling platform for the spindle

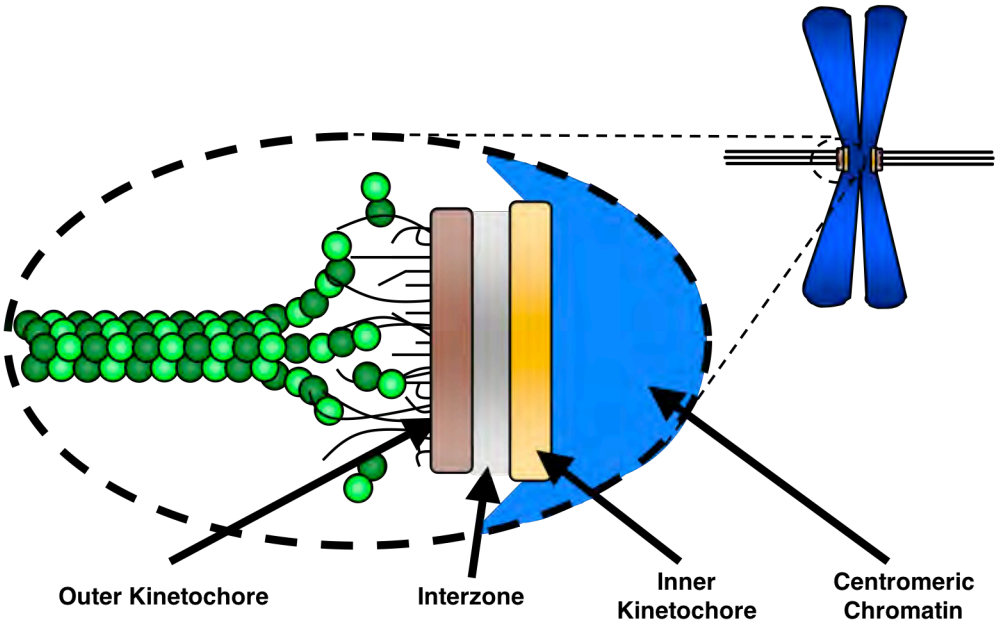
assembly checkpoint to regulate the metaphase-to-anaphase transition (Fig. A). To accommodate these functions, the kinetochore is consequently composed of a vast number of proteins that assemble into spatio-temporally tightly controlled protein networks (Chan et al., 2005; Cheeseman and Desai, 2008; Maiato et al., 2004)

Based on transmission electron microscopy analysis of mitotic chromosomes, the kinetochore is composed of a tri-laminar structure composed of an electron-dense inner and an outer plate separated by an interzone that is less electron dense (Brinkley and Stubblefield, 1966). Localization studies have facilitated the grouping of kinetochore protein complexes into inner and outer kinetochore components. The core of the inner kinetochore largely comprises a set of proteins that are closely associated with the underlying centromeric chromatin throughout the cell cycle. These include CENP-A and the components of the constitutive centromere-associated network, which are discussed in detail below. Together, these form the pre-kinetochore structure which in turn serves as the basis for the kinetochore assembly process commencing in late G2 and early mitosis (Cheeseman and Desai, 2008; Maiato et al., 2004).

The outer kinetochore is fully assembled only during mitosis and its structural composition is considerably more dynamic than that of the inner kinetochore (Cheeseman and Desai, 2008). While the architectural relationship between outer kinetochore components is starting to emerge, regulation of the temporal component in this assembly process is currently largely unknown. The outer kinetochore provides the primary end-on interaction with the plus ends of kinetochore microtubules. The number of microtubules attaching to a kinetochore is variable across species, ranging from as little as one microtuble in budding yeast to around 25 in humans. Various microtubule-dependent motor proteins including the kinesin family

member CENP-E, but also factors regulating microtubule dynamics including CLASP1, are located at the outer kinetochore (Chan et al., 2005).

An important component in organizing the structural make-up of the outer kinetochore is the KNL1-Mis12-Ndc80 (KMN) network. The tetrameric Mis12 complex is directly recruited by components of the constitutive centromere-associated network and in turn facilitates recruitment of the KNL1 and Ndc80 sub-complexes (Kwon et al., 2007; Liu et al., 2006). Both KNL1 and Ndc80 mediate interaction with spindle microtubules in an Aurora B kinase-regulated manner that allows fine-tuning of the mitotic kinetochore / microtubule attachments (Cheeseman et al., 2006; Welburn et al., 2010). The KMN network is also a key factor in assembly of additional outer kinetochore components, including those comprising the spindle assembly checkpoint (Chan et al., 2005), and assumes a central role in both, mechanical and regulatory control of chromosome segregation.



Outer Kinetochore		Interzone	Inner Kinetochore	Centromeric Chromatin
Ndc80	Plk1	3F3/2 antigen	CENP-A	INCENP
Bub1	PP1		CCAN	Survivin
BubR1	RZZ complex		Mis12	Aurora B
Bub3	Dynein		KNL1	Borealin
Mad1	CLASPs			HP1
Mad2	CLIP-170			PICH
Cdc20	EB1			MCAK
CENP-E	Nup107-160			
RanGAP1	Mel28			
RanBP2	Rama complex			
Crm1	Bod1			
CENP-F	HURP			
APC	Spindly			
MT attachment; k-fibre stability; regulation of MT dynamics SAC		*tension sensor*	*kinetochore assembly; kinetochore size*	*chromatid cohesion; structural support; MT attachment correction; bi-orientation*

Figure A| Molecular organization of the mammalian kinetochore. Known centromere and kinetochore proteins are listed. The functional properties of individual compartments are indicated in italics. MT: microtubule; SAC: spindle assembly checkpoint. Adapted and modified from Maiato et al. 2004.

B EPIGENETICS AND REGULATION OF CHROMATIN ARCHITECTURE

B.1 The Basic Concept of "Epigenetic" Regulation

Traditionally, the term "epigenetics" as coined by Conrad Waddington was closely associated with the process of development, or more specifically the question of how the same genotype can give rise to different phenotypes. Multi-cellular organisms are composed of a large variety of distinct cell types displaying, in part, tremendously different gene expression profiles and phenotypic characteristics, despite comprising an identical complement of genetic information and originating from the very same zygote. It was therefore recognized early on that cellular fate must be determined by mechanisms independent of the primary DNA sequence alone. Later, modern epigenetics was redefined as "the study of mitotically and / or meiotically heritable changes in gene function that cannot be explained by changes in DNA sequence" (Riggs et al., 1996).

Within recent years, the term "epigenetic regulation" has been used more frequently in a context somewhat diverged from this definition, to include structural and functional differences of chromosomal regions other than genes. Furthermore, it is now established that many of the marks that were originally regarded as "stable" are in fact subject to targeted modification or removal within the cell. I will therefore use the term epigenetics to collectively refer to "the structural adaptation of chromosomal regions so as to register, signal or perpetuate altered activity states" (Bird, 2007). This may include the regulation of telomeres (Blasco, 2007), DNA replication (Mechali, 2001), DNA repair (Bassal and El-Osta, 2005) and, as argued below, centromeres.

It is now well established that two major chromatin-associated pathways, involving the addition of chemical groups either to bases of the DNA directly (Bird, 2002) or to distinct amino acids within nucleosomes (Kouzarides, 2007), respectively, contribute to epigenetic processes. These modifications do not occur uniformly across the genome, resulting in the functional and structural compartmentalization of the chromatin in form of an “epigenome” that may be highly variable across different cell types. Recent work has identified other layers of epigenetic mechanisms, which are mediated in first instance by transcriptional activity (see below).

With respect to structural compartmentalization, two main states of interphase chromatin are classically distinguished, termed “euchromatin” and “heterochromatin”. Through microscopic examination, these appear comparatively decondensed and compacted, respectively. Functionally, euchromatin is generally associated with a transcriptionally active or permissive state, whereas heterochromatin is considered to be antagonistic to transcription. Various studies have uncovered a correlation between certain histone modifications and transcriptional states. Generally, histone hyper-acetylation, as well as methylation of histone H3 on lysine (K) 4, 36 and 79 is considered to be characteristic for transcriptionally active euchromatin, whereas histone hypo-acetylation and methylation of H3K9, K27, K64 as well as H4K20 are associated with a transcriptionally repressive, heterochromatic environment (Berger, 2007; Daujat et al., 2009; Sims et al., 2003).

B.2 Histone Lysine Methylation

Far beyond a simple entity of DNA packaging, the nucleosome is subject to a staggering variety of post-translational modifications. The best-studied covalent histone modifications include acetylation (Shahbazian and Grunstein, 2007), methylation (Shi and Whetstine, 2007), phosphorylation

(Nowak and Corces, 2004) and ubiquitination (Weake and Workman, 2008). Several additional modifications have been discovered on histones, including sumoylation, ADP-ribosylation, and proline isomerization. Combined with the afore-mentioned modifications, these are targeted to more than 70 different amino acid residues within the nucleosome, largely functioning to modulate chromatin architecture (Taverna et al., 2007). Not all modifications are present within the same nucleosome at a given time, and complexity is further enhanced through in part distinct agonistic and antagonistic relationships between individual modifications (for reference, see for example (Kim et al., 2009; Kirmizis et al., 2007; Nishioka et al., 2002; Rudolph et al., 2007; Zippo et al., 2009).

Histone modifications, most notably acetylation, methylation and phosphorylation, serve as platforms for the recruitment of a large variety of adaptor and effector proteins containing binding motifs recognizing distinct histone modification states (Taverna et al., 2007). Combined, these post-translational modifications place the nucleosome at the centre of a plethora of signalling pathways that regulate diverse structural and functional aspects of the chromatin (Jenuwein and Allis, 2001; Kouzarides, 2007).

Among the known enzymes catalyzing the above histone modifications, methyltransferases acting on either lysine or arginine residues are the most specific and diverse with respect to the amino acid residues targeted (Kouzarides, 2007). Furthermore, lysine residues can exist in monomethylated (me1), dimethylated (me2) or trimethylated (me3) states, providing the most versatile modification with respect to interpretation through the nuclear machinery. In contrast to previous assumptions that histone lysine methylation represents a permanent modification, it is now known that most if not all lysine methylation states are reversible through highly specific demethylases (Shi and Whetstine, 2007).

Distinct methylation states have been associated with critical chromatin functions, and are perhaps best characterized with respect to local transcriptional activity. For example, hyper-methylation of H3K4 is strongly enriched around the transcription start sites of actively transcribed genes and non-coding RNAs (Guenther et al., 2007; Guttman et al., 2009). The transcription elongation factor CHD1 binds to methylated H3K4 both *in vivo* and *in vitro* (Pray-Grant et al., 2005; Sims et al., 2005). H3K4me3 is also implicated in directly recruiting a histone acetyl-transferase activity in yeast (Martin et al., 2006), thereby potentially creating an environment favourable for transcription. In contrast, methylation of H3K27 is linked to facultative heterochromatin formation and gene repression (Lee et al., 2006). The silencing effect through methylation of H3K27 is well established and dynamically mediated through binding of downstream repressive complexes of the Polycomb group (Agger et al., 2007; Lee et al., 2006; Schwartz and Pirrotta, 2008).

Particularly well studied is a role of H3K9me3 in constitutive heterochromatin. In mammals, pericentromeric H3K9 trimethylation is catalyzed by the Suv39h histone methyl-transferases (Peters et al., 2001). H3K9me3 recruits heterochromatin protein (HP)1 to modulate higher-order chromatin architecture (Lachner et al., 2001). Amongst other factors, HP1 in turn facilitates recruitment of the H4K20-specific histone methyl-transferase Suv4-20h (Schotta et al., 2004), thereby re-enforcing local heterochromatic character. However, methylation of H3K9 also serves to illustrate an evident degree of functional plasticity between histone modifications: mice lacking Suv39h display a loss of pericentromeric H3K9me3 levels, which appears to be compensated in part by a conversion of local H3K27me1 to H3K27me3 (Peters et al., 2003). In addition, the recent detection of H3K9me3 as well as HP1γ within the coding region of transcribed genes (Vakoc et al., 2005; Vakoc et al., 2006) strongly emphasizes that the earlier, strict equation of H3K9me3 with transcriptionally inactive heterochromatin is too simplistic.

B.3 Transcription in the Modulation of Chromatin Architecture and Function

It has been known for a long time that transcription contributes to the metabolism of the cell through the generation of different RNA species, most prominently protein-coding messenger RNAs as well as structural and functional RNAs in form of ribosomal and transfer RNAs, respectively. While this certainly still holds true, recent technical advances, most notably the advent of next-generation (deep) sequencing technologies, have revealed that these transcripts merely represent the tip of the transcriptional iceberg. The human genome comprises only about 1.5 % protein-coding genetic information. Strikingly however, the human ENCODE project suggested that across different cell types, in excess of 90 % of the genomic DNA is actively transcribed (Birney et al., 2007). In fact, it is now established that a complex variety of long and short non-coding RNAs directly regulate cellular processes including chromatin structure, nuclear architecture as well as gene expression, the latter at both the pre- and post-transcriptional level (for an excellent review, see Amaral et al (Amaral and Mattick, 2008)). Most of these non-coding RNAs are developmentally regulated and differentially expressed throughout differentiation (for reference, see for example (Cawley et al., 2004; Dinger et al., 2008).

However, it is now furthermore established that the activity of polymerases themselves regulates important aspects of chromatin structure and function independent of the primary transcripts, through the transient recruitment of a multitude of chromatin-remodelling activities (see below). Taken together, non-coding transcription forms a vital and highly diverse entity in the over-all function of the nucleus

Non-coding Transcripts in the Regulation of Nuclear Function

Non-coding transcripts have been arbitrarily divided into short (<200 nucleotides) and long (>200 nucleotides) non-coding RNAs. To date, short non-coding RNAs have been most intensively studied. There are several types of short non-coding RNAs, which generally are the product of longer precursor RNA molecules processed in a pathway involving a distinct family of RNA binding and processing factors collectively referred to as the RNA interference machinery (Farazi et al., 2008). The majority of these, including short interfering (si), micro (mi) and piwi (pi) RNAs, in turn act post-transcriptionally to promote the cleavage of complementary messenger or non-coding RNAs through directing RNases of the Argonaute and Piwi families (Aravin et al., 2007; Carthew and Sontheimer, 2009; Zhao et al., 2008). In addition to this post-transcriptional gene silencing mechanism, siRNAs and piRNAs are also implicated in regulating transcription on the chromatin level, which may depend on the targeting of chromatin-modifying activities through interaction of associated small RNAs with nascent transcripts (Moazed, 2009; Zaratiegui et al., 2007). For example, as outlined below, siRNAs derived from anti-parallel transcripts of outer repeat sequences at fission yeast centromeres contribute towards mediating local chromatin structure.

There is also increasing evidence for a direct role of long non-coding RNAs in the regulation of nuclear function. An example is the requirement of the *MEN*ε/β RNA for formation of paraspeckles (Sunwoo et al., 2009). Long non-coding RNAs can also function on the chromatin level through guiding histone-modifying activities: For example, the *Air* non-coding RNA is implicated in transcriptional silencing in *cis* through interactions with the H3K9-specific histone methyl-transferase G9a (Nagano et al., 2008). *HOTAIR* was recently shown to interact with the Polycomb-repressive complex 2, which contains a H3K27-specific methyl-transferase, to mediate a *trans*-repressive effect across 40kb of the *HOXD* locus (Rinn et al., 2007).

Finally, the antagonistic action of two long non-coding RNAs is implicated in X chromosome silencing during dosage compensation (Leeb et al., 2009).

RNA Polymerase-mediated Regulation of Chromatin Architecture

It is now clear that RNA polymerase II itself can mediate distinct structural aspects of the local chromatin state through transient recruitment of various chromatin-modifying activities. Central to this function is the C-terminal domain of RNA polymerase II, which is comprised of multiple repeats of a conserved hepta-peptide sequence. Serine residues within this motif are subject to dynamic phosphorylation and de-phosphorylation events that are closely associated with the transcriptional status of the polymerase (Buratowski, 2009): Newly-loaded, promoter-associated RNA polymerase is largely unphosphorylated. As the polymerase engages into elongation mode, serine 5 is hyper-phosphorylated. In the ongoing course of elongation, serine 5 phosphorylation levels decrease gradually, while serine 2 phosphorylation becomes more prominent. The differentially phosphorylated C-terminal domain of RNA polymerase serves the recruitment of various factors required for processing of the primary RNA transcript, including 5' capping, splicing and 3' poly-adenylation, as well as factors modulating nucleosome modifications (Egloff and Murphy, 2008).

In yeast, the interplay of RNA polymerase-associated factors and the modification of the underlying chromatin is well documented, and the relationship of elongating RNA polymerase activity with chromatin architecture and function is starting to emerge. At the promoter and 5' region of transcribed genes, serine 5-phosphorylated RNA polymerase recruits the Paf1 complex which in turn serves as platform for the recruitment of a variety of chromatin remodelling complexes. These include Rad6, which mediates mono-ubiquitination of histone H2B (Wood et al., 2003), the COMPASS complex containing the H3K4-specific methyl-transferase Set1, as well as

Dot1, which is required for methylation of H3K79 (Krogan et al., 2003a; Ng et al., 2003). Rad6-mediated H2B ubiquitination is required to facilitate the activity of Set1 and Dot1 (Dover et al., 2002). As illustrated above, hyper-methylated H3K4 is in turn implicated in further recruiting factors modulating local chromatin architecture favouring transcription.

Further downstream in actively transcribed genes, serine 2 phosphorylation of the C-terminal domain is associated with the recruitment of the H3K36-specific methyl-transferase Set2 (Krogan et al., 2003b). Methylation of H3K36 in turn is implicated in the recruitment of an Rpd3-containing histone deacetylase complex (Carrozza et al., 2005; Joshi and Struhl, 2005; Keogh et al., 2005). The Set2 / Rpd3 pathway is thought to restore the local chromatin architecture after passage of RNA polymerase, which is facilitated by transient acetylation of the preceding nucleosome (Workman, 2006). Functionally, loss of this pathway can result in the occurrence of aberrant transcripts initiating from within the coding region (Carrozza et al., 2005; Joshi and Struhl, 2005), emphasizing an important role in maintenance of the local chromatin structure.

C THE CENTROMERE

The centromere is a specialized chromatin domain that both, serves as a site for and regulator of the assembly of the mitotic kinetochore structure, thereby playing a pivotal role in the maintenance of the genetic information. At mitotic chromosomes, the centromere is readily detectable as the primary constriction, further highlighting an additional role in the cohesion of sister chromatids.

The primary make-up of centromeric DNA is highly diverged across different species, owing to rapid and in part differential evolution (Henikoff et al., 2001). The simplest centromere, found in the budding yeast *Saccharomyces cerevisiae*, comprises 125 bp of DNA that are sufficient to facilitate kinetochore assembly in a sequence-dependent manner (McAinsh et al., 2003). In contrast to the “point” centromeres of *S. cerevisiae*, however, most other eukaryotes, including the fission yeast *Schizosaccharomyces pombe*, possess “regional” centromeres of varying length (Pluta et al., 1995). These are typically composed of repetitive DNA sequence elements, with centromere size increasing across evolutionary more complex species, ranging from 35-110kb in *S. pombe* (Pidoux and Allshire, 2004) to 0.3-5 Mbp in humans (Cleveland et al., 2003; Schueler and Sullivan, 2006). A notable exception to point and regional centromeres is the formation of “holocentric” kinetochore structures along the full length of chromatids in *Caenorhabditis elegans* (Albertson and Thomson, 1982).

Despite the heterogeneous composition of their underlying sequences, metazoan centromeres display conservation of some general features, including a bias towards A / T-rich base composition of the underlying DNA, flanking heterochromatic domains and a similar protein composition of their associated kinetochore structures (Morris and Moazed, 2007; Vos et al., 2006). Importantly, centromeres of all species examined contain a specific

histone H3 variant termed CENP-A (Carroll and Straight, 2006), which, as discussed in detail below, is thought to be a key determinant of the centromere locus.

C.1 DNA Sequence and Centromere Identity

In humans, centromeres are typically formed on chromosome-specific higher-order “alphoid” DNA arrays composed of 171 bp monomer units termed alpha-satellite (Willard, 1985) that are tandemly arranged in a directional, head-to-tail fashion (for review, see (Willard, 1990)). Alphoid sub-families can be distinguished based on their monomer sequence and composition, with many being present at the centromeres of more than one chromosome (Choo et al., 1991). These highly-ordered type I satellites can span several hundreds of kilobases and are flanked by more diverged, monomeric type II satellites that are frequently interspersed with other repetitive elements, such as long and short interspersed elements (LINEs and SINEs, respectively) (Ikeno et al., 1994; Rudd et al., 2006; Schueler et al., 2001) (Fig. B). CENP-A associates selectively with type I satellite DNA (Ando et al., 2002; Lam et al., 2006). Furthermore, several type I satellite monomers contain a 17 bp DNA sequence termed the CENP-B box that facilitates binding of the conserved centromere protein, CENP-B (Earnshaw et al., 1987; Ikeno et al., 1994; Masumoto et al., 1989). A comparable organization is apparent at mouse centromeres, where a core of tandemly-arranged, CENP-B box-containing minor satellites is associated with centromere function and is flanked by larger major satellite arrays.

Despite functional endogenous centromeres apparently favouring alpha-satellite DNA, several findings demonstrate that this sequence context, including the localization of CENP-B, merely represents a preference rather than a determinant for the centromere locus. At mitotically stable dicentric chromosomes that contain two spatially distinct alpha-satellite regions as a

consequence of chromosome fusion or translocation events, CENP-B is present at both alphoid sites, whereas detectable levels of CENP-A and other essential kinetochore components localize to the “active” site only (Sugata et al., 2000; Sullivan and Schwartz, 1995; Warburton et al., 1997). Furthermore, neocentromeres can form and assemble fully functional kinetochores in the absence of alphoid DNA, recruiting all kinetochore components except for CENP-B (Alonso et al., 2007; Alonso et al., 2003; Sugata et al., 2000; Warburton et al., 1997). Based on these findings, maintenance of centromere identity and function is considered to be regulated by epigenetic mechanisms (see for example (Cleveland et al., 2003)).

From the findings summarized above, it is evident that CENP-B is neither a necessary requirement for the maintenance of centromere identity, nor for assembly of a functional kinetochore structure. Besides the stable maintenance of neocentromeres in the absence of CENP-B boxes, this is further supported by the lack of CENP-B from the centromere of the endogenous Y chromosome both in human and mouse (Earnshaw et al., 1989; Masumoto et al., 1989), and perhaps most strikingly illustrated by the viability of CENP-B knock-out mice in the absence of any detectable mitotic defects (Hudson et al., 1998). Instead, as further discussed below, binding of CENP-B, as well as an A / T-rich alphoid input DNA, is required for the *de novo* establishment of centromeres in the context of formation of human artificial chromosomes. How and to which extent this requirement for CENP-B translates into a functional role *in vivo*, where *de-novo* centromere formation is rare and typically does not involve naked alphoid DNA, remains elusive and a subject of debate. Despite the apparent absence of mitotic and meiotic phenotypes in male CENP-B null mice, sperm count in these animals was found to be significantly reduced (Hudson et al., 1989). While of no immediate consequence in the short term, mild phenotypes in the germ line may explain the observed evolutionary conservation of the CENP-B gene.

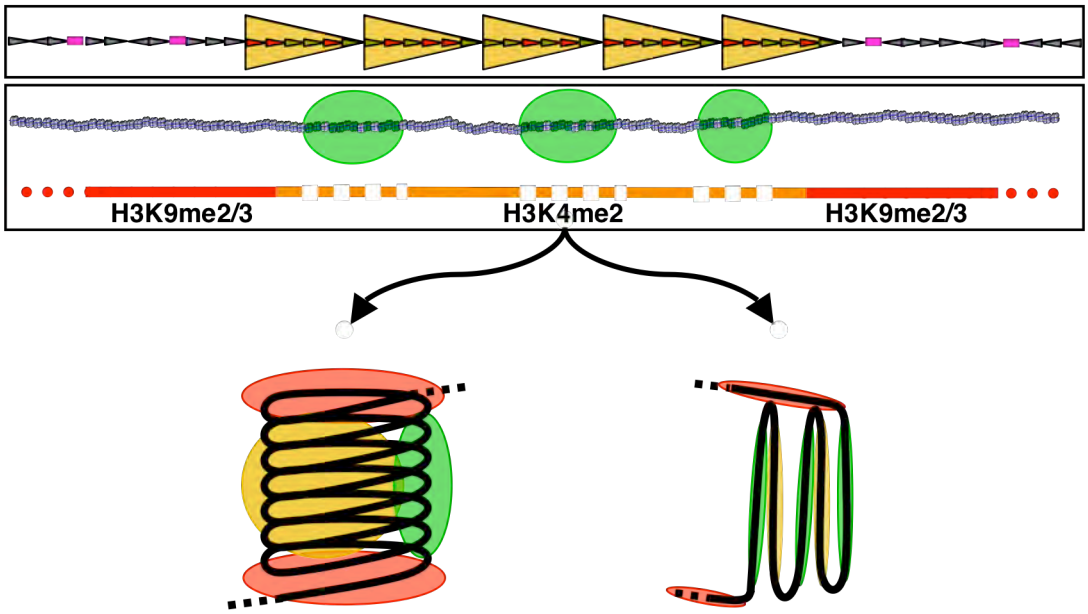


Figure B| Organization of mammalian centromeric DNA and chromatin. **Top:** Higher-order type I alpha-satellite arrays (large yellow triangles) containing alpha-satellite monomers carrying CENP-B boxes (small red triangles) are flanked by more diverged type II satellite arrays lacking higher-order periodicity and frequently containing other types of repetitive DNA such as LINE elements (purple boxes). **Centre:** CENP-A nucleosomes (green) assemble on a portion of the type I alpha-satellite array and are organized into distinct domains (green ovals) interspersed with H3 nucleosomes. Based on fibre FISH analysis (Sullivan and Karpen, 2004), H3K4me2 is present within and between the CENP-A domains, whereas more distal chromatin is characterized by hypermethylated H3K9 nucleosomes. **Bottom:** Two models for a higher-order folding of centromeric chromatin in mitotic chromatids. The solenoid model (adapted from Sullivan and Karpen, 2004) and an anti-parallel array based on the proposed boustrophedon organization of chicken centromere chromatin (Ribeiro et al., 2010) is schematically drawn on the left and right, respectively. Green: relative CENP-A nucleosome positioning; yellow: relative H3K4me2 nucleosome positioning; red: H3K9 hyper-methylated nucleosomes.

The Centromere-specific Histone H3 Variant CENP-A

CENP-A was among the first centromere proteins to be identified using antibodies from autoimmune patient sera (Earnshaw and Rothfield, 1985), and subsequently shown to be a variant form of histone H3 (Palmer et al., 1991; Sullivan et al., 1994). Homologues of CENP-A were identified in other eukaryotes, including budding yeast Cse4 (Meluh et al., 1998), fission yeast Cnp1 (Takahashi et al., 2000), *C. elegans* HCP3 (Buchwitz et al., 1999) and *Drosophila* CID (Henikoff et al., 2000). In all species examined, CENP-A localizes exclusively to active centromeres, including those of human artificial chromosomes (see below) (Ohzeki et al., 2002; Okamoto et al., 2007), stable dicentric chromosomes and neocentromeres (Alonso et al., 2007; Warburton et al., 1997).

CENP-A nucleosomes are associated with a series of centromere proteins termed the CENP-A nucleosome-associated complex (NAC), comprising CENP-B, -C, -H, -M, -N, -T and -U (Foltz et al., 2006). The NAC forms part of a larger group of proteins that are constitutively present at the centromere (Foltz et al., 2006; Izuta et al., 2006; Okada et al., 2006). CENP-A itself is required for localization of almost all other known constitutive centromere proteins (Carroll et al., 2009; Goshima et al., 2003; Hori et al., 2008a), placing it at the base of a hierarchical pathway towards the assembly of the mitotic kinetochore structure. Over-expression of CENP-A results in its ectopic localization on chromosome arms, and may be accompanied by recruitment of at least a subset of kinetochore components, including CENP-C (Heun et al., 2006; Van Hooser et al., 2001). As expected for a nucleosomal histone, CENP-A forms a stable component of centromeric chromatin, as determined by salt extraction (Ando et al., 2002), fluorescence recovery after photo-bleaching (Hemmerich et al., 2008) and fluorescence pulse-chase experiments experiments (Jansen et al., 2007). Together, these findings have prompted the proposal that CENP-A may be (the) epigenetic

determinant maintaining centromere identity and function (Allshire and Karpen, 2008; Black and Bassett, 2008; Cleveland et al., 2003; Warburton et al., 1997).

While CENP-A and canonical histone H3 share considerable homology across their C-terminal histone domains, their N-termini are remarkably divergent. As discussed above, the N-terminal tail of histone H3 is a primary target for a large variety of post-translational modifications mediating various chromatin-based signalling pathways. Remarkably, besides Aurora B kinase-mediated phosphorylation during mitosis (Zeitlin et al., 2001), little is known about other post-translational modifications occurring at the N-terminus of CENP-A. In fact, the N-terminus of CENP-A is highly diverged across different species both in terms of its amino acid sequence and length, indicating little evolutionary conservation, and further suggesting its redundancy in the role of CENP-A at centromeres.

Existing evidence suggests that human CENP-A assembles into homotypic nucleosomes *in vivo* (Shelby et al., 1997), although a minor fraction of CENP-A may be present within heterotypic CENP-A / H3 nucleosomes (Foltz et al., 2006). On the level of tetrameric sub- and mature octameric nucleosomes, CENP-A confers a higher compaction and conformational rigidity compared to canonical histone H3 (Black et al., 2007a; Black et al., 2004). This rigidity is mediated by the CENP-A-specific histone fold domain that forms the interface with histone H4. As further illustrated below, this domain also mediates the localization of CENP-A to centromeres, and was therefore termed the CENP-A targeting domain (CATD) (Black et al., 2004). Notably, as discussed later, the CATD also serves as substrate recognition site for a specific histone chaperone (Foltz et al., 2009) and is sufficient to facilitate targeting of chimeric histone $H3^{CATD}$ to centromeres (Black et al., 2004). Strikingly, expression of the latter histone $H3^{CATD}$ chimera

in HeLa cells rescues the severe depletion phenotype of CENP-A RNAi (Black et al., 2007a).

Together, these findings strongly suggest that the primary function of CENP-A is mediated by its unique histone fold domain. Consistently, a recent study demonstrated direct recognition and binding of this domain by CENP-N, a member of the constitutive centromere-associated network (see below) acting upstream in facilitating the recruitment of most other components of this network (Carroll et al., 2009). Furthermore, it is tempting to speculate that the structural properties of CENP-A nucleosomes may enable the local chromatin to accommodate the requirements for both flexibility and stability in light of the forces exerted by kinetochore-attached microtubules during mitosis.

C.2 A Network of Constitutive Centromere Proteins forms the Basis for Kinetochore Assembly

Data acquired in particular within the past five years have established that the basis for assembly of the mitotic kinetochore superstructure is formed by a complex, in part interdependent network of 16 proteins, comprising CENP-C, -H, -I, -K to -U, -W, and -X (Amano et al., 2009; Foltz et al., 2006; Hori et al., 2008a; Hori et al., 2008b; McClelland et al., 2007; Okada et al., 2006) (Fig. C). In contrast to the temporally controlled recruitment of other kinetochore components, these proteins, together with CENP-A, localize to centromeres throughout the cell cycle and have therefore been termed the constitutive centromere-associated network (CCAN). Within the CCAN, the majority of the proteins are organized into distinct sub-complexes. Interestingly, some of these components, such as the CENP-O sub-complex (see below), are non-essential for cell survival in chicken DT40 cells (Okada et al., 2006). These findings suggest some

degree of redundancy within the CCAN with respect to formation of a functional kinetochore architecture in otherwise unperturbed cells.

Combined, the CCAN is largely responsible for recruitment and assembly of the KNL1-Mis12-Ndc80 (KMN) network in mitosis, which in turn promotes regulated recruitment of additional outer kinetochore components, and facilitates interaction with spindle microtubules (Cheeseman et al., 2008; Cheeseman et al., 2004; Goshima et al., 2003; Kline et al., 2006) reviewed in (Cheeseman and Desai, 2008). Consistently, depletion of individual CCAN members in vertebrates, with exception of the afore-mentioned CENP-O sub-complex, causes severe mitotic defects (Amano et al., 2009; Foltz et al., 2006; Hori et al., 2008a; Hori et al., 2008b; McClelland et al., 2007; Okada et al., 2006). Recent studies have started to investigate individual functions of CCAN members based on more subtle phenotypic analysis (see for example Hori et al., 2008 and Amano et al., 2010) and provided evidence that the CCAN assemles in a non-linear manner. However, the pronounced interdependency of CCAN members and sub-complexes with respect to their centromeric localization renders phenotypic analysis of individual components a challenge.

CENP-C

The perhaps most-studied member of the CCAN is CENP-C (Saitoh et al., 1992), although many gaps in understanding its contribution towards the architecture of centromeric chromatin and the kinetochore still remain to be filled. CENP-C is an important factor in maintaining proper mitotic progression and is essential for viability of cultured cells as well as mouse embryos (Fukagawa and Brown, 1997; Fukagawa et al., 1999; Kalitsis et al., 1998; Kwon et al., 2007; Tomkiel et al., 1994). The central and C-terminal portions of CENP-C independently confer its targeting to centromeres (Politi et al., 2002; Song et al., 2002; Trazzi et al., 2002; Yang et al., 1996). The

central domain of CENP-C also mediates binding to DNA *in vitro*, although no preferential target DNA sequence motifs could be identified (Sugimoto et al., 1994; Yang et al., 1996). Interestingly, although CENP-A is required for localization of CENP-C to centromeres, CENP-C immunoprecipitation experiments following extensive micrococcal nuclease digestion to the mono-nucleosome level recovered canonical histone H3, but not CENP-A. CENP-A was, however, recovered after less stringent digestion conditions, suggesting that CENP-C binds to centromeric alpha-satellite DNA in the vicinity of CENP-A nucleosomes, but does not directly associated with them (Hori et al., 2008a).

Injection of anti-CENP-C antibodies or conditional knock-out of CENP-C resulted in a reduction of the kinetochore size at mitotic chromosomes (Hori et al., 2008a; Tomkiel et al., 1994). CENP-C directly interacts with Mis12 (Cheeseman et al., 2004; Kwon et al., 2007) and is required for Mis12 as well as Ndc80 targeting into the kinetochore (Kwon et al., 2007; Liu et al., 2006), providing a direct link to the assembly of the outer kinetochore. CENP-C is furthermore required for the localization of various checkpoint proteins as well as the kinesin CENP-E (Liu et al., 2006). CENP-C was shown to be reciprocally required for the localization of CENP-A in *Drosophila* (Erhardt et al., 2008; Goshima et al., 2007), whereas depletion of CENP-C in chicken cells apparently does not affect centromeric CENP-A levels (Hori et al., 2008a; Okada et al., 2006), indicating some divergence in the assembly pathways of the inner kinetochore between different species.

The CENP-H/I Complex

CENP-H was discovered in 1999 as a novel, constitutive kinetochore protein in mice (Sugata et al., 1999). Based on sequence homology, human CENP-H was identified shortly after, demonstrating similar characteristics with respect to constitutive, centromeric localization (Sugata et al., 2000).

CENP-H was subsequently shown to be required for localization of CENP-C in interphase, but not mitotic chicken cells (Fukagawa et al., 2001). Together with CENP-I (Nishihashi et al., 2002), CENP-H interacts with a series of other CCAN members to form the conserved CENP-H/I complex (Okada et al., 2006). CENP-H/I complex components were reported to be required for Ndc80 localization into the kinetochore (Cheeseman et al., 2008; Okada et al., 2006). Based on reciprocal interaction studies and analysis of the phenotypic severity, the CENP-H/I complex was further sub-divided into the CENP-H class (CENP-H, -I, -K, -L), CENP-M class (CENP-M) and CENP-O class (CENP-O, -P, -Q, -R -U) (Okada et al., 2006).

As pointed out above, CENP-O class DT40 knock-out cells are viable. However, CENP-O-depleted cells demonstrate delays in the progression through mitosis, and members of the CENP-O class, with the exception of CENP-R, are required for efficient recovery from nocodazole-induced spindle damage (Hori et al., 2008b). In contrast to CENP-O, knock-out of CENP-H, CENP-I, CENP-K or CENP-M is lethal, resulting in severe chromosome alignment defects and a mitotic arrest both in DT40 and HeLa cells (Fukagawa et al., 2001; Nishihashi et al., 2002; Okada et al., 2006). Notably, CENP-H, -M and -U were previously co-purified with TAP-tagged CENP-A in HeLa cells (Foltz et al., 2006), and conditional knock-out of CENP-H or CENP-M in DT40 cells resulted in reduced centromeric incorporation of transiently expressed CENP-A-GFP (Okada et al., 2006). Thus, although established centromeric CENP-A nucleosomes are refractory to depletion of most CCAN members, components of the CCAN are clearly required for maintenance of CENP-A levels and centromere identity.

The CENP-S/X Complex

CENP-S and CENP-X were found to interact in DT40 and HeLa cells, and their centromeric localization showed a mutual dependency (Amano et

al., 2009). CENP-S knock-out DT40 cells are viable but display prolonged mitoses reminiscent of the CENP-O class phenotype (Amano et al., 2009). However, recovery from spindle damage in CENP-S knock-out cells was more complete than in the CENP-O knock-out background, suggesting distinct roles within the CCAN. Furthermore, while localization of CENP-S was abrogated in cells depleted of CENP-T or CENP-K, CENP-S displayed normal centromeric staining in CENP-U knock-out cells, placing the CENP-S/X sub-complex into a distinct CCAN compartment than the CENP-O class (Amano et al., 2009).

At CENP-S and CENP-X deficient DT40 kinetochores, levels of KNL1 and Ndc80, but not Mis12, were mildly reduced. HeLa cells depleted of CENP-X by RNAi showed a more exacerbated phenotype, with many cells displaying severe chromosome alignment defects. This was paralleled by a more prominent reduction of Ndc80 and KNL1 levels (Amano et al., 2009).

The CENP-T/W Complex

Recently, an intriguing CCAN sub-complex was identified, comprising CENP-T and CENP-W (Hori et al., 2008a). Centromeric localization of CENP-T and CENP-W is mutually dependent, and is upstream of most other CCAN members, with the notable exception of CENP-C. DT40 knock-outs of either CENP-T or CENP-W are not viable, and conditional depletion of these proteins results in a drastic accumulation of cells in prometaphase, consistent with loss of a detectable outer kinetochore plate on the electron microscopy level at mitotic chromosomes (Hori et al., 2008a). Both, CENP-T and CENP-W contain a histone-like fold domain. These domains are required for binding of CENP-T and CENP-W to centromeric DNA *in vivo* and *in vitro* (Hori et al., 2008a). Intriguingly, similar to CENP-C, immunoprecipitation after micrococcal nuclease digestion of chromatin to the mono-nucleosome level demonstrated that the CENP-T/W complex associates with nucleosomes

containing canonical histone H3, but not CENP-A (Hori et al., 2008a), together highlighting the existence of distinct, non-redundant pathways tethering the CCAN to centromeric chromatin.

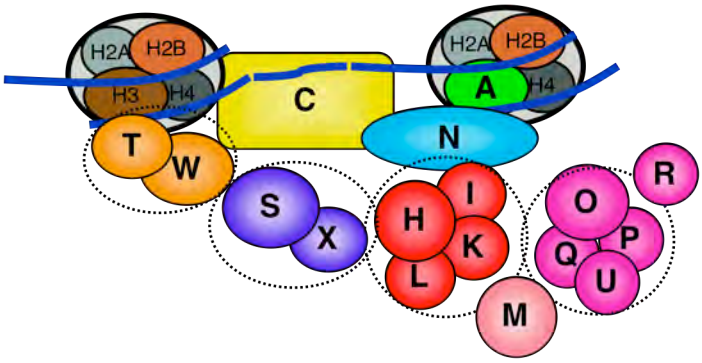


Figure C| Molecular composition of the CCAN. Schematic histone H3 and CENP-A nucleosomes are shown together with the 16 components of the CCAN known to date. Dotted circles comprise members of a distinct CCAN sub-complex as noted in the main text.

CENP-N

CENP-N was originally identified as a component of the CENP-A NAC in pull-downs of TAP-tagged CENP-A nucleosomes obtained after extensive micrococcal nuclease digestion (Foltz et al., 2006). In contrast to most other NAC components tested, CENP-N was subsequently shown to directly bind to CENP-A nucleosomes assembled on alpha-satellite or non-alphoid DNA *in vitro* (Carroll et al., 2009). Importantly, in the same study, nucleosomes assembled with histone $H3^{CATD}$ chimeras containing the CENP-A histone fold domain were bound by CENP-N with indistinguishable affinity from wild-type CENP-A. *In vivo*, CENP-N point mutations reducing interaction with CENP-A target less efficiently to centromeres. These point mutants further cause a reduction of centromeric levels of CENP-H, -I and -K, a dependency also observed after RNAi-mediated depletion of CENP-N (Carroll et al., 2009; McClelland et al., 2007). The CENP-N-dependent localization of the CENP-H class sub-complex *in vivo* is likely mediated by direct interaction of the CENP-N C-terminus with CENP-L (Carroll et al., 2009). Together, these results indicate that CENP-N provides an intimate, physical link between the "epigenetic" mark formed by CENP-A, and the assembly of an important part of the CCAN connecting centromeric chromatin with the outer kinetochore structure.

C.3 Maintenance and Deposition of CENP-A

Specific Timing of CENP-A Deposition

In light of the stable association of CENP-A with centromeres and its vital role upstream of most other constitutive centromere components, maintenance of centromeric character is considered to be tightly linked to the maintenance of CENP-A. The deposition of CENP-A at centromeres in the course of the cell cycle has therefore been a major focus of research. In actively dividing cells, the bulk of canonical histone H3 is assembled into

chromatin during S phase (Ahmad and Henikoff, 2002; Tagami et al., 2004) in a process mediated by chromatin-assembly factor (CAF-)1 (Tagami et al., 2004). In contrast, early studies established that CENP-A deposition is uncoupled from DNA replication (Shelby et al., 2000; Sullivan and Karpen, 2001). CENP-A expression peaks in G2, and CENP-A expressed ectopically during S phase is not incorporated into centromeres (Shelby et al., 1997). Instead, ectopically expressed CENP-A displays global chromatin association, suggesting that a more specific, temporarily controlled mechanism must ensure proper deposition of newly-synthesized CENP-A at centromeres.

Recently, elegant quench-pulse-chase experiments in HeLa cells surprisingly established that deposition of newly-synthesized CENP-A molecules occurs late in mitosis and subsequent G1 (Jansen et al., 2007). Quantitative fluorescence microscopy in *Drosophila* embryos revealed incorporation of CENP-A to occur during anaphase (Schuh et al., 2007). Importantly, this timing means that centromeres contain only half the amount of their original CENP-A molecules when entering mitosis, as existing data suggests that CENP-A nucleosomes are distributed onto replicated sister centromeres in a semi-conservative manner (Jansen et al., 2007; Shelby et al., 2000). The fate of the vacant CENP-A nucleosome space produced during S phase is currently unclear. As also discussed by others (Allshire and Karpen, 2008), this gap may likely be filled with canonical H3-containing nucleosomes, or be maintained as a nucleosome-free region. Alternatively, replicated centromeres may carry atypical histone tetramers containing CENP-A, H4, H2B and H2A. The existence of interphase CENP-A half-nucleosomes has been proposed by a recent study in *Drosophila* (Dalal et al., 2007), although the interpretability of the results is disputed (Black and Bassett, 2008). The unique features of the CENP-A / H4 interface (Black et al., 2004) discussed above may convey structural stability to these tetramers.

The CENP-A Deposition Machinery

A series of deletion and substitution studies have identified the minimal region within the CENP-A histone fold domain, CATD, that confers targeting to centromeres (Black et al., 2004; Black et al., 2007b; Shelby et al., 1997; Vermaak et al., 2002). Interestingly, this domain corresponds to the domain conferring structural rigidity to CENP-A nucleosomes (Black et al., 2004). Substitution of the analogous domain in canonical histone H3 with that of CENP-A is sufficient to facilitate localization of the chimeric H3 to centromeres (Black et al., 2004; Black et al., 2007b), suggesting that putative CENP-A chromatin assembly factors specifically recognize this domain as their substrate. Consistently, two concomitant studies reported the identification of HJURP (Holiday-junction recognition protein) as a chaperone that facilitates deposition of pre-nucleosomal CENP-A at centromeres (Dunleavy et al., 2009; Foltz et al., 2009). HJURP directly binds to CENP-A, and the CENP-A centromere targeting domain is both required and sufficient for this interaction (Foltz et al., 2009).

Further studies have substantially increased the list of factors that are required for CENP-A localization to and maintenance at centromeres. Of particular note is the trimeric hMis18 complex, which interacts with the histone chaperone RbAp46 and RbAp48 (Fujita et al., 2007). This complex transiently localizes to HeLa centromeres from anaphase into G1 (Fujita et al., 2007; Maddox et al., 2007), thereby slightly preceding the reported time window of CENP-A deposition (Jansen et al., 2007). Knock-down of either hMis18 component, or combined knock-down of RbAp46 and RbAp48 causes a marked reduction in centromeric CENP-A (Fujita et al., 2007; Hayashi et al., 2004; Maddox et al., 2007). In *S. pombe*, spMis16 and spMis18, the homologue of mammalian RbAp46/48, are also required for CENP-A loading (Hayashi et al., 2004). Notably, deletion of either spMis16 or spMis18 results in a pronounced increase in acetylation levels of centromeric nucleosomes (Hayashi et al., 2004). Together, Mis18 and RbAp46/48 have

been suggested to contribute towards “licensing” of centromeric chromatin for subsequent deposition of CENP-A (Fujita et al., 2007). Notably, combined downregulation of RbAp46/48 negatively affects HJURP stability (Dunleavy et al., 2009), highlighting a tight connection between these pathways in CENP-A depostion.

Additional factors required for CENP-A localization that have been identified include *Drosophila* CAL1, a putative ubiquitin binding protein (Erhardt et al., 2008; Goshima et al., 2007). However, homologues in other eukaryotes have not been identified. The histone chaperone Rsf1 interacts with CENP-A in interphase, and was subsequently shown to be able to mediate incorporation of CENP-A into nucleosomes *in vitro* (Perpelescu et al., 2009). Interestingly, loss of centromeric CENP-A signal during Rsf1 knock-down is dependent on a high-salt wash step, suggesting that Rsf1 might act downstream of an initial recruitment of CENP-A to the centromere, where it subsequently facilitates the stable incorporation of CENP-A into centromeric chromatin (Perpelescu et al., 2009).

The CCAN can Mediate Deposition of CENP-A

While established centromeric CENP-A clearly acts upstream of the CCAN, recent work highlights a role for certain CCAN members in mediating the deposition of newly-synthesized CENP-A. As mentioned above, a part of the CENP-H/I complex is required for the incorporation of ectopically expressed CENP-A-GFP at chicken centromeres (Okada et al., 2006). The underlying mechanism remains to be determined.

More recently, fluorescent pulse-chase experiments after knock-down of CENP-N indicate its requirement for CENP-A deposition (Carroll et al., 2009). Given the direct interaction of CENP-N with both, CENP-A and the CENP-H class sub-complex, it will be interesting to determine if CENP-A

deposition is mediated by the kinetochore-associated or the nucleoplasmic pool of CENP-N. With respect to the latter, it is notable that knock-down of CENP-N results in reduced global levels of CENP-A (Carroll et al., 2009). Endogenous CENP-A levels are also reduced upon stable, ectopic expression of CENP-A (Jansen et al., 2007), suggesting that levels of CENP-A are tightly controlled presumably through a pathway involving protein-degradation. Interaction of pre-nucleosomal CENP-A with CENP-N may stabilize CENP-A, although a direct interaction of CENP-N with DNA-free CENP-A / H4 tetramers has not been observed *in vitro* (Carroll et al., 2009).

Transcription Elongation-associated Factors in the Maintenance of CENP-A

Several recent studies identified the Facilitates Chromatin Transcription (FACT) chromatin remodelling complex within pull-downs of CENP-A (Foltz et al., 2006; Izuta et al., 2006; Okada et al., 2009). In eukaryotes, FACT is associated with DNA polymerase α as well as elongating RNA polymerase II, and is thought to facilitate transient nucleosome destabilization by removing a nucleosomal H2A / H2B dimer (Sims et al., 2004). FACT is further suggested to mediate the re-integration of H2A / H2B dimers into the nucleosome, thereby playing a pivotal role not only in facilitating polymerase activity, but also in maintaining accurate chromatin architecture in the wake of the elongating polymerase (Sims et al., 2004). In chicken cells, retention of FACT staining was reported at interphase and mitotic centromeres following pre-extraction with a moderately high triton concentration, (Okada et al., 2009). Interestingly, conditional depletion of individual components of the FACT complex abrogated centromeric targeting of CENP-A-GFP. The same study also demonstrated a critical role for the ATP-dependent, transcription elongation-associated chromatin remodelling factor CHD1 for CENP-A deposition at chicken and HeLa centromers (Okada et al., 2009). Together, these data may point towards a requirement for non-

coding transcription and associated chromatin-remodelling activities in the incorporation of newly-synthesized CENP-A.

A Proof-reading Mechanism for CENP-A Deposition?

CAF-1 or HIRA histone chaperones that mediate interphase assembly of H3.1 and H3.3, respectively, (Tagami et al., 2004) are absent from CENP-A pull-downs (Foltz et al., 2006). Instead, the presence of a CENP-A-specific histone chaperone, in combination with the limited time window of CENP-A incorporation primed during mitosis, would be predicted to facilitate precise spatio-temporal control of the deposition of newly-synthesized CENP-A. However, while the specific delivery of CENP-A to and its incorporation at centromeres is clearly a fundamental requirement for maintaining centromere identity, it is noteworthy to mention that an inverse process is likely to exist that corrects for mis-targeted CENP-A ectopically incorporated into chromosomal regions other than centromeres. This is emphasized by the finding that simple over-expression of CENP-A results in mis-targeting of CENP-A (Heun et al., 2006; Van Hooser et al., 2001), suggesting the saturation of a putative proof-reading pathway. At present, the existence of such pathway that is capable of removing CENP-A from arm chromatin remains elusive.

C.4 The Centromere Chromatin Domain

Clearly, the mere presence of nucleosomes containing a specific histone variant renders centromeres a unique compartment distinct from bulk chromatin. At budding yeast point centromeres, a single CENP-A nucleosome directs the assembly of the mitotic kinetochore structure (Furuyama and Biggins, 2007). In higher eukaryotes, centromeric chromatin is a more complex domain displaying peculiar features. Initial studies of *Drosophila* and human kinetochore chromatin fibres revealed a surprising

pattern displaying blocks of CENP-A nucleosomes interspersed with nucleosomes containing canonical histone H3 (Blower et al., 2002). Discrete domains of CENP-A nucleosomes were also detected at functional neocentromers (Alonso et al., 2007; Chueh et al., 2005), suggesting that regional blocks of either nucleosome compartment may be required for kinetochore function.

CENP-A is Embedded in a Chromatin Domain with a Higher-order Conformation

As determined by deconvolution immunofluorescence microscopy, the chromatin underlying functional mitotic kinetochores shows an asymmetric arrangement, with CENP-A nucleosomes aligning along the pole-ward-facing surfaces of sister chromatids. The chromatin between sister kinetochores and underlying the polarized CENP-A aggregates is made up exclusively from H3 nucleosomes (Blower et al., 2002; Sullivan and Karpen, 2004). Together with the peculiar arrangement of CENP-A nucleosomes on extended fibres, these observations led to suggestion of a model in which kinetochore chromatin is organized into a higher-order structure, such as a cylindrical solenoid (Blower et al., 2002) (Fig. B). Recent quantitative electron microscopy analysis in human HT1080 cells concluded that at least one additional instance of higher-order folding is required to achieve the necessary chromatin compaction fitting the measured dimensions of the three-dimensional centromere constriction and CENP-A domain (Marshall et al., 2008). A distinct model based on a multi-layer boustrophedon-like arrangement, that is a series of anti-parallel folds of the chromatin fibre, has been put forward after quantitative super-resolution microscopy analysis of chicken DT40 kinetochore fibres (Ribeiro et al., 2010). It has to be noted that a critical factor in any empirical modelling of the kinetochore chromatin domain is the number of CENP-A molecules present at each kinetochore, which was calculated as 24 (12 homotypic nucleosomes) per chicken DT40

centromere (Ohta et al., 2010), ranging up to semi-quantitative estimates of 30,000 (15,000 homotypic nucleosomes) in HeLa cells (Black et al., 2007b). The latter number may be an over-estimate, as data from recent fluorescence quantification predicts on average 100 CENP-A molecules (50 homotypic nucleosomes) at human kinetochores (Dani Bodor, Lars Jansen, unpublished data).

The Chromatin Environment of Centromeres

As illustrated above, pericentromeric heterochromatin in various species is highly enriched for distinct degrees of methylation at H3K9, H3K27 and H4K20, a feature that is tightly linked to both pericentromeric structure and function. These chromatin domains do not contain detectable levels of CENP-A, and while over-expression of CENP-A results in its localization to chromosome arms, ectopic CENP-A is not observed within pericentromeric heterochromatin (Heun et al., 2006; Van Hooser et al., 1999), suggesting that local constraints prevent CENP-A deposition or maintenance. Indeed, a study assessing kinetochore fibres in both *Drosophila* and HeLa cells, established that histone H3 nucleosomes within the CENP-A chromatin domain display a post-translational modification pattern that is distinct from that of flanking heterochromatic domains (Sullivan and Karpen, 2004). Based on these studies, centromeric chromatin is characterized by distinctive histone hypo-acetylation and undetectable levels of H3K4me3, but also lacks appreciable hyper-methylation of H3K9 (Sullivan and Karpen, 2004). CENP-A chromatin may be directly flanked by H3K9me2, while H3K9me3 domains appear to be restricted to more distal positions on extended fibres (Sullivan and Karpen, 2004).

An intriguing finding was the presence of H3K4me2 nucleosomes within *Drosophila* and human CENP-A chromatin domains, and their localization to the chromatin directly underlying CENP-A aggregates at mitotic

chromosomes (Sullivan and Karpen, 2004) (Fig. B). H3K4me2 nucleosomes have also been found at fission yeast centromeres (Cam et al., 2005) and, albeit at low levels, at human neocentromeres (Alonso et al., 2010). Based on chromatin immunoprecipitation and immuno-fluorescence / *in situ* hybridization experiments, prominent levels of CENP-A and H3K4me2 are associated with type I satellite DNA of various human centromeres (Lam et al., 2006), as well as alphoid arrays of human artificial chromosomes (Lam et al., 2006; Nakano et al., 2008). H3K4me2 nucleosomes within the CENP-A chromatin domain have subsequently received considerable attention, and currently several models and hypotheses attribute this mark a critical role in kinetochore chromatin higher-order structure or the maintenance of CENP-A nucleosomes (see for example a recent review by Allshire and Karpen (Allshire and Karpen, 2008)). It has to be noted, however, that levels of the H3K4me2 mark at different centromeres and neocentromeres or across different cell lines appear to be highly variable and may even be undetectable by different analytical means (Alonso et al., 2010; Sullivan and Karpen, 2004; Vagnarelli et al., 2008). Potential functional roles of centromeric H3K4me2, if existing, may therefore be redundant with other local histone modifications.

Transcription at Centromeres

As pointed out above, the process of transcription as well as non-coding transcripts themselves can influence chromosome structure and function in several distinct ways. At fission yeast centromeres, a role for transcription is well established. In this system, transcription of outer repeat sequences yields double-stranded RNA, feeding into an RNAi pathway mediated by Argonaute that is required for the establishment of pericentromeric heterochromatin through the H3K9 methyl-transferase Clr4 (Volpe et al., 2003; Volpe et al., 2002). Recent data indicates that this pathway is further required for *de novo* establishment, but not subsequent maintenance, of a flanking CENP-A domain on naked centromeric input DNA (Folco et al.,

2008). Notably however, in the context of *de novo* CENP-A establishment, this pathway can be cut short by direct tethering of Clr4 to the chromatin (Kagansky et al., 2009), suggesting that the sole purpose of outer repeat transcription is likely to facilitate local heterochromatin formation. However, whether centromeric transcription or transcripts also play a functional role in higher eukaryotes is presently unknown.

Two recent proteomics screens identified components of the transcription-associated FACT complex in CENP-A pull-downs from HeLa cells (Foltz et al., 2006; Obuse et al., 2004). Furthermore, low levels of transcripts derived from mouse major and minor satellites, as well as from human alpha-satellite repeats, have been detected in several studies (Bouzinba-Segard et al., 2006; Efroni et al., 2008; Frescas et al., 2008; Kanellopoulou et al., 2005; Martens et al., 2005). Neocentromeres, including their CENP-A chromatin domain, tolerate transcription of underlying genes, although levels of RNA polymerase activity at these genes have not been determined (Saffery et al., 2003). In case of murine major and minor satellite transcripts, RNase1 digestion indicates the presence of a double-stranded RNA fraction (Kanellopoulou et al., 2005; Martens et al., 2005) that may serve as a substrate for processing by the RNAi machinery. Indeed, a short (~20-40 nucleotides) species of minor satellite RNA indicative of Dicer processing is detected in mouse embryonic stem cells (Kanellopoulou et al., 2005). Deletion of Dicer results in a strong reduction in the levels of these short RNAs, concomitant with an increase in long (>200 nucleotides) minor satellite transcripts and a decrease in global H3K9me3 levels. However, no defects in chromosome number or structure were observed in Dicer null cells (Kanellopoulou et al., 2005).

Similarly, conditional depletion of Dicer in a DT40 cell hybrid stably carrying a copy of human chromosome 21 resulted in a marked increase in the abundance of chromosome 21-derived type I alpha-satellite as well as

pericentromeric satellite transcripts (Fukagawa et al., 2004). In these cells, Dicer depletion resulted in mis-localization of HP1, premature sister chromatid separation and an increase in cells with aberrant chromosome 21 copy numbers. The structure of the inner kinetochore did not appear to be affected. A concern in interpretation of these data is that the major phenotypic effects described occurred only about four days after loss of detectable Dicer protein levels, a time span sufficient for up to ten additional cell divisions in the background of the fast-growing DT40 cells.

Ectopic expression of a mouse minor satellite repeat unit was reported to result in reduction of sister chromatid cohesion, an undefined increase of prometa- and metaphase cells, and mislocalization of Aurora B kinase and HP1γ (Bouzinba-Segard et al., 2006). However, conclusions as to whether this represents a transcript-specific defect, or is merely the consequence of an unspecific mechanism such as saturation of the RNA degradation machinery, cannot be drawn, as no controls involving expression of scrambled or non-centromeric RNA were included.

Together, these data support the idea of active transcription of vertebrate centromeric satellite DNA, with at least a fraction of these transcripts serving as substrates for Dicer. However, a relevant function of processed centromeric transcripts remains largely elusive. Notably, all studies to date affect post-transcriptional processes, thereby leaving centromeric transcription *per se* intact.

D HUMAN ARTIFICIAL CHROMOSOMES

D.1 Generation and Biology of Human Artificial Chromosomes

A human artificial chromosome (HAC) is an ectopic mini-chromosome generated *de novo* from naked input DNA containing synthetic or natural higher-order alpha-satellite arrays (Harrington et al., 1997; Ikeno et al., 1998; Nakano et al., 2008). HACs are therefore different from endogenous mini-chromosomes derived through natural truncation events or targeted engineering (Grimes and Cooke, 1998), in that they typically do not contain host-derived sequence elements. HACs assemble apparently normal kinetochore structures, as reflected by recruitment of inner and outer kinetochore components paralleled by high mitotic stability (Ikeno et al., 1998; Nakashima et al., 2005). Dynamics of HAC sister chromatid movements during mitosis are furthermore largely comparable to endogenous chromosomes (Tsuduki et al., 2006). In contrast to endogenous chromosomes, HACs are small and not essential for cell viability, forming favourable model systems for the study of various aspects of chromosomal structure and function. Moreover, their high mitotic stability and capacity to stably carry hundreds of kilobases of exogenous DNA have gained them considerable attention with respect to their potential function as non-integrating vectors for human gene therapy (Basu and Willard, 2005).

First-generation human artificial chromosomes were generated by co-transfection of cloned chromosome 17 higher-order alpha-satellite arrays together with telomeric repeat sequences and random genomic fragments into human HT1080 cells (Harrington et al., 1997). Subsequently, a team led by Hiroshi Masumoto at Nagoya University in Japan succeeded to from *de novo* human artificial chromosomes from a single, linear input DNA construct, comprising about 100 kb of chromosome 21-derived type I alpha-satellite

(α21-I) DNA cloned in a modified yeast artificial chromosome vector fitted with human telomeric sequences (Ikeno et al., 1998), thereby overcoming the inherent uncertainty of HAC composition when co-transfecting random genomic fragments. HAC formation was efficient and observed in about 30 % of stable transfectants. Cytological as well as molecular analysis suggested extensive multimerization of the original input DNA, which is now recognized as a common feature in the formation of HACs and may represent a functional requirement for HAC formation or maintenance.

Soon after this pioneering work, Ebersole and co-workers reported the generation of HACs from circular input DNA lacking telomeric repeats, containing 70 kb α21-I DNA in a minimal phage P1 artificial chromosome vector (Ebersole et al., 2000). Surprisingly, HAC formation efficiency using this circular input DNA yielded HAC-carrying transformants at comparable frequency to using linear input DNA containing telomeric sequences. Immunofluorescence *in situ* hybridization suggested multimerization of the input DNA, but also demonstrated that these HACs had not acquired host telomeric DNA, suggesting their ectopic maintenance as a circular chromosome. Nevertheless, these HACs displayed high mitotic stability even in the absence of selection (Ebersole et al., 2000).

HACs derived from linear or circular input DNA have since served to study requirements for *de novo* centromere formation. A fundamental factor for HAC generation is the type of alpha-satellite used. HACs formed efficiently using different vectors containing higher-order type I alphoid DNA including that of chromosome 17 and 21 (Grimes et al., 2002; Harrington et al., 1997; Ikeno et al., 1998; Masumoto et al., 1998). However, for as yet unknown reasons, not all type I arrays support high-frequency HAC formation, with the X chromosome-derived type I array being less efficient than those of chromosomes 17 or 21 (Schueler et al., 2001). The Y chromosome type I alpha-satellite, or chromosome 21 type II satellite DNA

did not result in detectable formation of HACs at all (Harrington et al., 1997; Ikeno et al., 1998; Masumoto et al., 1998). As both of these arrays do not contain CENP-B boxes, binding of CENP-B therefore appears to play a pivotal role in establishing *de novo* centromeres on naked input DNA. This was formally demonstrated by the finding that input α21-I satellite arrays containing mutations in their CENP-B boxes abolishing association with CENP-B did not support HAC formation (Ohzeki et al., 2002). Furthermore, reducing the density of functional CENP-B boxes on the input alphoid array negatively correlated with HAC formation (Okamoto et al., 2007). Precisely how CENP-B regulates establishment of centromere structure is currently unknown, but may involve nucleosome phasing (Ando et al., 2002) or modulation of the local chromatin character (Okada et al., 2007).

D.2 The Alphoid^{tetO} Human Artificial Chromosome

HACs have been successfully used both for the study of chromatin modifications associated with the input DNA (Nakashima et al., 2005), as well as the dynamics associated with seeding and maintenance of CENP-A (Okamoto et al., 2007) during *de novo* HAC formation. However, a key question in the field of centromere biology is if and to what extent the chromatin state underlying established kinetochores governs aspects of centromere identity and function.

To provide a tool for the specific analysis as well as manipulation of a single kinetochore and its associated chromatin domain in living cells, our lab, in collaboration with Hiroshi Masumoto at Nagoya University and Vladimir Larionov at the National Institutes of Health, have designed and generated a novel HAC, the kinetochore of which is assembled on a synthetic higher-order alphoid^{tetO} array containing the sequence of the *Escherichia coli* tetracycline operator (tetO) in every other alphoid monomer (Nakano et al., 2008). The tetO sequence is not present within the host cell

genome, thereby distinguishing the alphoid^{tetO} HAC centromere from any other genomic region, including endogenous centromeres containing related alpha-satellite arrays. Furthermore, the tetO serves as highly specific binding site for the *E. coli* tetracycline repressor (tetR), and has been used for tetracycline / doxycycline-reversible tethering of fusion constructs to the tetR within mammalian cells (Freundlieb et al., 1999; Gossen and Bujard, 1992).

Generation of the Alphoid^{tetO} Human Artificial Chromosome

The seed of the alphoid^{tetO} higher order array was formed by a 342 bp alphoid dimer comprising a CENP-B box-containing chromosome 17-derived alpha-satellite monomer extended by an entirely synthetic alphoid DNA unit based on an alpha-satellite consensus sequence. In place of a CENP-B box, the synthetic unit carried the 42 bp tetO. By means of conventional cloning, this dimer was extended into a head-to-tail-cloned array spanning about 3.5 kb. Subsequent expansion into a higher-order array captured in an artificial chromosome vector was performed using an elaborate method described by Ebersole et al. (Ebersole et al., 2005); Initial extension of the 3.5 kb alphoid^{tetO} array was performed using phage Φ29 DNA polymerase-mediated rolling circle amplification. Products derived from the rolling circle reaction were subsequently transformed into a recombination-proficient yeast strain, along with a linearized targeting vector carrying a yeast artificial chromosome cassette (*HIS3*, *CEN6*, *ARSH4*), a bacterial artificial chromosome cassette (*Cm*, *ori F*) as well as a blasticidin S resistance cassette for selection in yeast, *E. coli* and mammalian cells, respectively. Ultimately, a vector carrying an alphoid^{tetO} array spanning about 50 kb was isolated, and the circular construct was transfected into HT1080 cells (Fig. D).

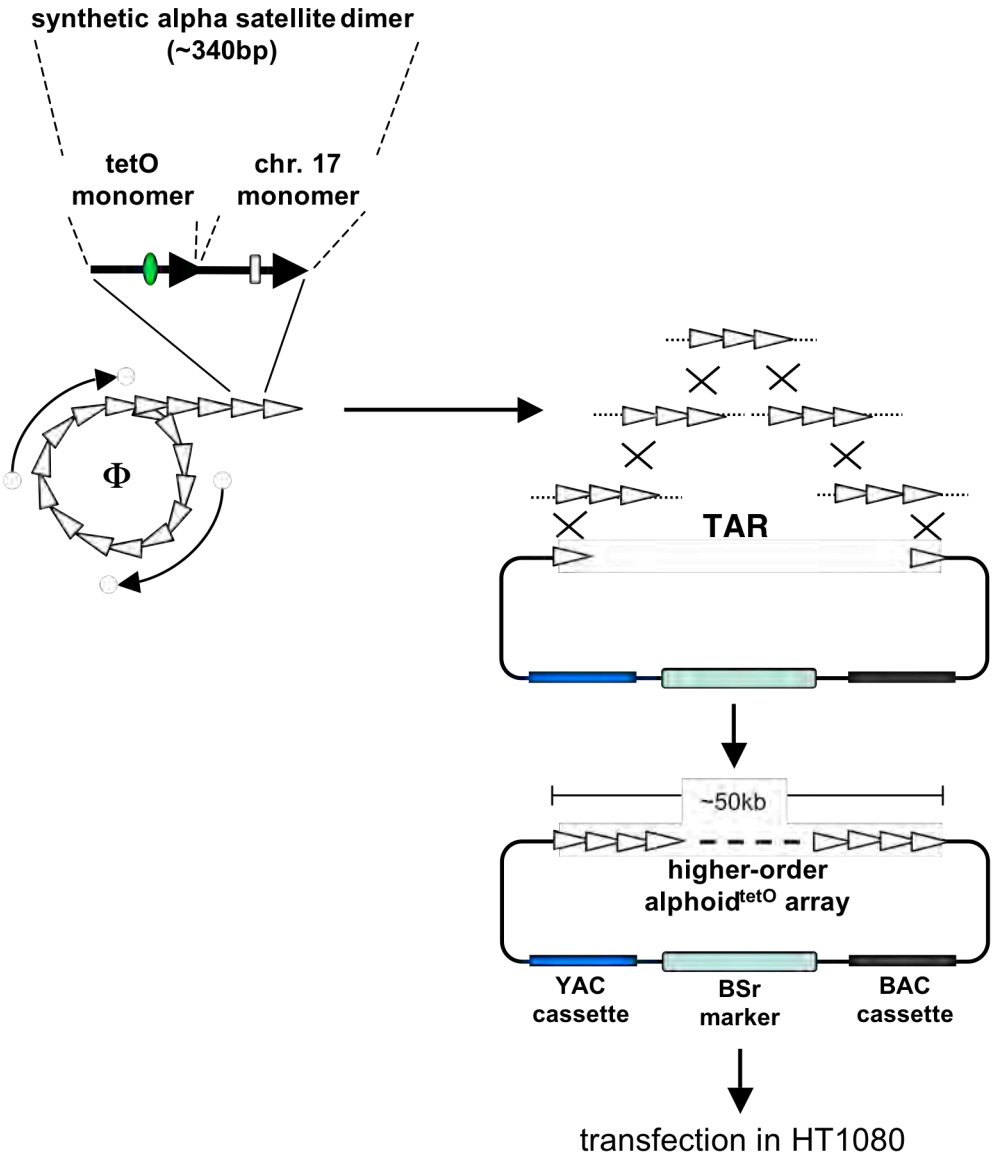


Figure D| Generation and structure of the alphoid^{tetO} HAC vector. Schematic workflow illustrating the generation of the 50 kb higher-order alphoid^{tetO} repeat in which every other alphoid monomer contains the tet operator (tetO) sequence in place of a CENP-B box. The process is illustrated in the main text. Φ: rolling circle amplification; TAR: transformation-associated recombination.

Despite the fully-synthetic nature of every second alphoid monomer, HAC formation was observed, albeit at low frequency. Molecular analysis suggested that the alphoid^{tetO} HAC comprised in excess of 45 copies of the input DNA, with a roughly conserved alphoid-to-backbone ratio. The HAC displayed mitotic stability comparable to that of endogenous chromosomes in the absence of selection. Importantly, expression of tetR fused to EYFP in cells containing a single HAC copy revealed co-localization of a discrete EYFP spot signal with staining for various kinetochore components both in interphase and throughout mitosis (Fig. E), demonstrating the assembly of a fully-functional kinetochore structure on the synthetic alphoid^{tetO} array.

Manipulation of the Chromatin of a Single Kinetochore **in vivo**

A remarkable feature of the alphoid^{tetO} HAC kinetochore was its apparent structural integrity despite the direct tethering of tetR-EYFP into the underlying chromatin. In fact, HAC stability was not compromised even after continuous targeting with tetR-EYFP for 30 days (Nakano et al., 2008). Furthermore, tethering of tetR-EYFP for up to 14 days did not result in detectable alterations of the alphoid^{tetO}-associated chromatin, which was characterized by high levels of CENP-A and H3K4me2, modest levels of H3K9me3 and extremely low levels of H3K4me3.

Subsequent experiments established that the alphoid^{tetO} HAC system can indeed be used for the specific and conditional manipulation of the associated centromeric chromatin environment (Nakano et al., 2008). Within seven days, targeting of a fusion construct of tetR and the effective domain of a transcriptional repressor (tTS) resulted in doxycycline-sensitive increase of H3K9me3 levels associated with the alphoid^{tetO} array, concomitant with a prominent reduction of H3K4me2 levels. Tethering of tTS further increased local binding of HP1, demonstrating a prominent heterochromatinization of the HAC chromatin. Importantly, these effects were paralleled by a reduction

of CENP-A levels and those of other kinetochore components, misalignment of mitotic HACs and ultimately the loss of the HAC from the population within 14 days, indicative of inactivation of its kinetochore. Consistent with an adverse effect of local heterochromatin formation on centromere function, direct targeting of HP1 fused to tetR was sufficient also to inactive the alphoid^{tetO} HAC kinetochore.

Notably, tethering of a transcriptional activator (tTA) into the HAC kinetochore also induced mosaic disruption of HAC kinetochore structure (Nakano et al., 2008). Overall however, the impact of the tTA on HAC kinetochore chromatin and stability were mild compared to those of tTS. Nevertheless, together these data provide strong support for the hypothesis that the underlying chromatin state is critical for kinetchore structure and function.

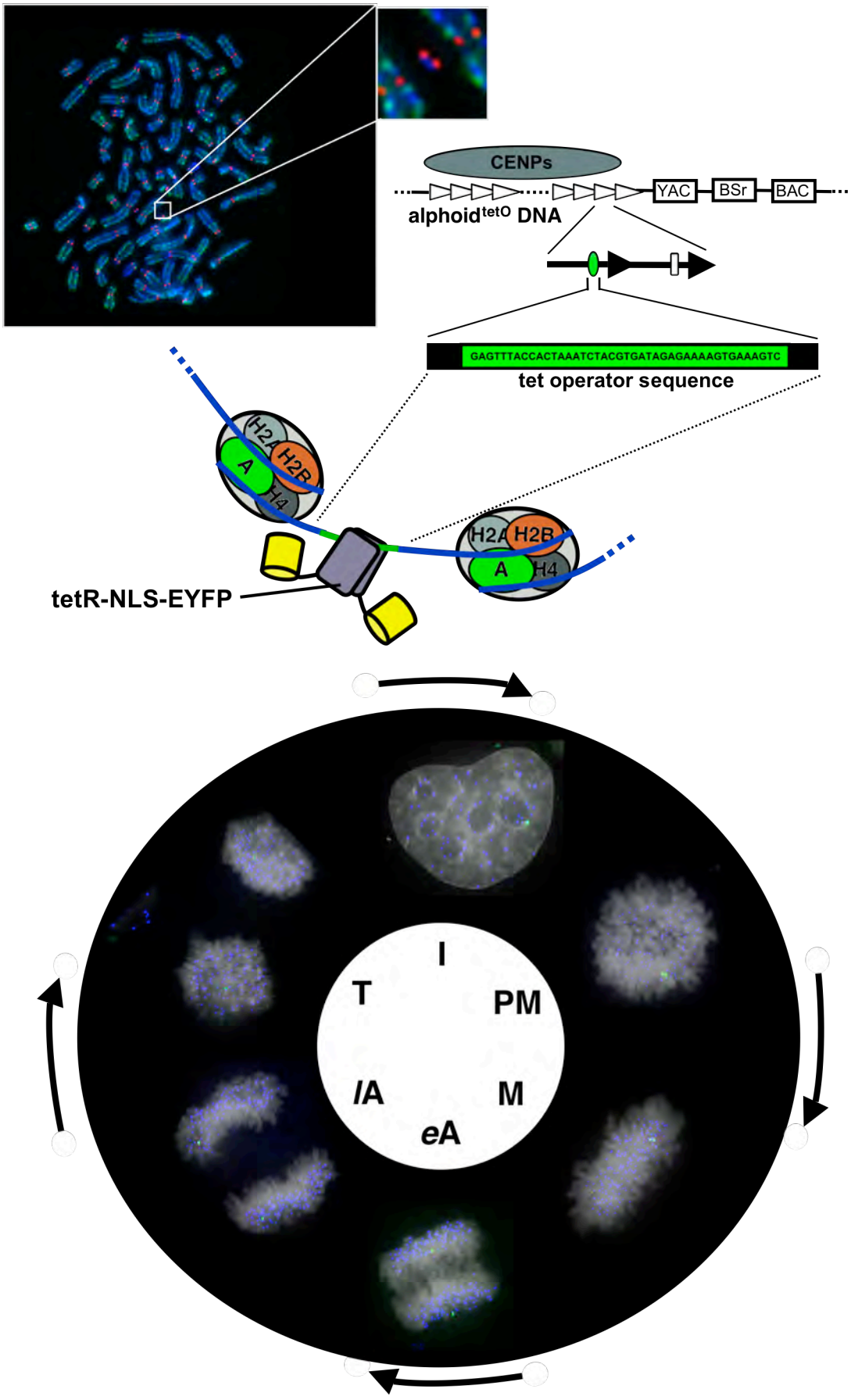

CENPs
YAC
BSr
BAC
alphoid^tetO DNA
GAGTTTACCACTAAATCTACGTGATAGAGAAAAGTGAAAGTC
tet operator sequence
H2A
H2B
H4
A
tetR-NLS-EYFP
I
T
PM
/A
M
eA

Figure E| Targeting of the alphoid^{tetO} HAC in living cells. The top inset shows the alphoid^{tetO} HAC generated in HT1080 cells (red: CENP-A). The schematic drawing shows one unit of the input plasmid (see Fig. D). The alphoid^{tetO} array assembles a functional kinetochore structure. The tetO sequence in every second alphoid monomer can be bound by tet repressor (tetR) fusion constructs. A dimer or tetR-EYFP is shown here to bind the tetO in the vicinity of alphoid^{tetO} HAC CENP-A nucleosomes. In cells, a discrete EYFP spot signal (green) can be seen co-localizing with CENP-A (blue) in every stage of the cell cycle. I: interphase; PM: prometaphase; M: metaphase; *e*A: early anaphase; *l*A: late anaphase; T: telophase.

E AIMS OF THIS WORK

As outlined above, post-translational histone modifications play a critical role in shaping and regulating the local chromatin environment, as well as modulating important aspects of nuclear structure and function. Centromeres are unique chromatin compartments, in that they contain a specific histone H3 variant together with blocks of H3-containing nucleosomes that display a modification pattern distinct from those found within the chromatin compartment distal to the CENP-A domain. It is therefore feasible to postulate that the local chromatin state contributes, directly or indirectly, to the over-all structure and function of centromeres. Until the establishment of the alphoid^{tetO} HAC, this hypothesis was difficult if not impossible to test in a biologically relevant context.

My motivation, and the over-all aim of my research, was to acquire novel insight into the nature of the centromere chromatin compartment with particular emphasis on the structure-function relationship of local chromatin state, centromere identity and kinetochore function. Employing the unique alphoid^{tetO} HAC system, I wanted to address the questions outlined below:

What is the Chromatin Environment of an Active Centromere?

Recent genome-wide chromatin immunoprecipitation studies have significantly enhanced the understanding of chromatin compartments formed by the interplay of different synergistic as well as antagonistic histone modifications. I therefore initially wanted to comprehensively analyse the chromatin environment of the well-defined alphoid^{tetO} HAC centromere. The relative profile of histone modifications at this genomic locus may facilitate the prediction of chromatin-remodelling activities shaping and maintaining this domain. Furthermore, these data form an essential basis for the selection of specific catalytic activities to be used for targeted engineering of the

underlying chromatin, as most known histone-modifying enzymes prefer specific modified histone residues as substrate.

What is the Mechanism of tTS-mediated Kinetochore Disruption?

As illustrated above, the tTS is an efficient antagonist of alphoid^{tetO} HAC kinetochore structure. The existing evidence suggests that excessive heterochromatin character itself may be incompatible with maintenance of kinetochore function. However, modifications upstream of HP1 recruitment may be sufficient to destabilize the HAC kinetochore structure. tTS is thought to recruit the molecular scaffolding protein KAP1 which in turn comprises distinct functional domains recruiting a variety of chromatin-modifying activities. By direct tethering of KAP1 and KAP1 domains, I wished to determine if and to what extent tTS-mediated kinetochore disruption can be dissected. Furthermore, I was interested in gaining additional insight into the process of tTS / KAP1-mediated deactivation of the HAC kinetochore.

What is the Consequence of Depletion of Centromeric H3K4me2?

Centromeric H3K4me2 has been put forward as a key modification in the maintenance or function of centromeric structure. Prominent levels of H3K4me2 are also found at the synthetic alphoid^{tetO} HAC centromere. By engineered reduction of local H3K4me2 levels, I wanted to assess the associated consequences on maintenance of centromere character and kinetochore structure with the aim to provide experimental insight into a possible role of this modification at centromeres.

What Other Chromatin Aspects Determine Alphoid Fate?

In the majority of cases, transfection of HAC vectors carrying alphoid arrays results in their integration into a host chromosome. Despite the underlying primary sequence being identical, these ectopic integrations do

not assemble or retain pronounced levels of CENP-A, and do not facilitate kinetochore formation. Comparing the histone modification profile at an alphoid^{tetO} integration site with that observed at the functional alphoid^{tetO} HAC centromere, I expected to identify aspects different between these two sites that may act as key elements in determining the respective fate of the associated arrays. Identification of histone modifications displaying relative differences between the integration and the HAC alphoid^{tetO} arrays would then serve as a basis for future reciprocal engineering of both, the integration site and the established HAC, to determine their impact on the functional status of the alphoid^{tetO} array.

Materials and Methods

F MATERIALS

F.1 Bacterial and Eukaryotic Cell Culture

Cell Lines

HT1080-derived AB2.2.18.21 (Nakano et al., 2008), 1C7 (Cardinale et al., 2009) and 1C7-derived cells were maintained in complete RPMI Medium 1640 (+L-Glutamine) (Invitrogen), supplemented with 10 % FBS (Invitrogen) (cRPMI). cRPMI was further supplemented with 100 U/ml penicillin G and 100 µg/ml streptomycin sulfate (Invitrogen). All HAC-containing cell lines where grown in the presence of 4 µg/ml blasticidin S (Invitrogen). Cells were grown at 37°C in humidified atmosphere containing 5% CO_2.

Bacterial* Escherichia coli *Strains

Chemically competent *E. coli* TOP10 cells (F^- *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) ϕ80*lac*ZΔM15 Δ *lac*X74 *deo*R *rec*A1 araD139 Δ(ara-leu)7697 galU galK rpsL endA1) (in-house) or *E. coli* DAM^- (ara-14 leuB6 fhuA31 lacY1 tsx78 glnV44 galK2 galT22 mcrA dcm-6 hisG4 rfbD1 R(zgb210::Tn10) TetS endA1 rspL136 (StrR) dam13::Tn9 (CamR) xylA-5 mtl-1 thi-1 mcrB1 hsdR2) purchased from NEB were routinely grown at 37°C on LB-Agar plates or in liquid LB shaking at 180rpm. Resistance to kanamycin sulfate or ampicillin (Sigma-Aldrich) was selected for in the presence of 50 µg/ml or 100 µg/ml antibiotic, respectively.

F.2 Buffers, Solutions and Reagents

All buffers and solutions were prepared using double-distilled water. Chemicals and reagents used were purchased from Sigma-Aldrich or Merck,

with the exception of formamide solution (Fisher), SDS solution (Severn Biotech Ltd) and Tween-20 (BioRad).

Name	Composition
10x DNA loading buffer	40% Sucrose; 2mg/ml Orange G
Denhardt's solution, 100x	20g/l Ficoll 400; 20g/l BSA (fraction V); 20g/l Polyvinylpyrrolidone; pH7.0
Dilution Buffer 1, for ChIP	50mM Tris-HCl pH8.0; 2mM EDTA; 0.2% SDS; 134mM NaCl; 0.88% Triton X-100; 0.088% Sodium-Deoxycholate
Dilution Buffer 2, for ChIP	50mM Tris-HCl pH8.0; 167mM NaCl; 1.1% Triton X-100; 0.11% Sodium-Deoxycholate
Elution Buffer, for ChIP	10mM Tris-HCl pH8.0; 1mM EDTA; 1% SDS
Hybridization Buffer, for FISH	50% Formamide; 2x SSC; 1x Denhardt's solution; 0.1% SDS; 10% Dextran Sulfate
IP Buffer (1.7x), for ChIP	20mM Tris-HCl pH8.0; 1M NaCl; 1mM EDTA; 0.08% SDS; 1.7% Triton X-100; 33% glycerol; 1mM DTT
KCM buffer	10mM Tris-HCl pH8.0; 120mM KCl; 20mM NaCl; 0.5mM EDTA; 0.1% Triton X-100
LB	1% tryptone; 0.5% yeast extract; 10mM NaCl; pH7.4
Lysis Buffer, for ChIP	10mM Tris-HCl pH8.0; 10mM NaCl; 0.5% NP-40
PBS	65mM Na_2PO_4; 8.8mM KH_2PO_4; 137mM NaCl; 2.7mM KCl; pH7.4
PBS/Tween (PBS-T)	PBS; 0.2% Tween-20
Post-Elution Buffer, for ChIP	10mM Tris-HCl pH8.0; 9mM EDTA; 600mM NaCl

Protease Inhibitor Cocktail (CLAP)	1µg/ml each chymostatin, leupeptin, antipain, pepstatin
ProtK Buffer, for ChIP	10mM Tris-HCl pH8.0; 1mM EDTA; 0.67% SDS; 450µg/ml Proteinase K
RIPA-150 /-500	50mM Tris-HCl pH8.0; 150 / 500mM NaCl; 1mM EDTA; 0.1% SDS; 1% Triton X-100; 0.1% Sodium-Deoxycholate
S.O.C.	2% tryptone; 0.5% yeast extract; 10mM NaCl; 2.5mM KCl; 10mM $MgCl_2$; 10mM $MgSO_4$; 20mM glucose; pH7.0
Sonication Buffer, for ChIP	20mM Tris-HCl pH8.0; 1mM EDTA; 1mM DTT
SSC, 20x	3M NaCl; 300mM Sodium Citrate; pH7.0
TAE	40mM Tris-acetate; 1mM EDTA; pH8.0
TBS	25mM Tris; 137mM NaCl; 3mM KCl; pH7.4
TE	10mM Tris; 1mM EDTA; pH8.0
Wash Buffer, for ChIP	20mM Tris-HCl pH8.0; 600mM NaCl; 1mM EDTA; 0.05% SDS; 1% Triton X-100; 20% Glycerol; 1mM DTT

F.3 Oligonucleotide primers

All oligonucleotide oligonucleotides were synthesized by and purchased from Sigma-Aldrich. Desalted and dried oligonucleotides were resuspended in double-distilled water and stored as 100µM stock solution at -20℃.

Cloning

Description	Sequence (5' - 3')	Notes
HPBD_1_Fwd	ATGGCCCCTCCAAGAGCCCC	

HPBD_1_Rev	CCCTCCGCAAGAGCCATAAGC	
KRBD_1_Fwd	TTCGGCCGCAGCCTCGGCCT	
KRBD_1_Rev	GAGGGGCCATGGGTGCAGGG	
LSD1_F	AAAAAAGATCTGAGATGTTATCTGGGAAGAAGGC	5’ BglII
LSD1_R	AAAAAAGATCTATGCATCTGTCTCACATGCTTGG	5’ BglII
p65_F	TCCGGAGCGTCGACCCCGGGG	5’ BspEI
p65_R	TCCGGATTAGGAGCTGATCTGACTCAGC	5’ BspEI
PHD_1_Fwd	GGTGGCCCGGGAACCCTGGA	
PHD_1_Rev	GGGGCCATCACCAGGGCCAC	
VP16-RE_F	AATAGGATCCTCCGCGTACAGCCG	5’ BamHI
VP16-RE_R	ATAAGGATCCCTACCCACCGTACTCG	5’ BamHI

Real-Time PCR

Description	**Sequence (5’ - 3’)**
5SrDNA_F	CCGGACCCCAAAGGCGCACGCTGG
5SrDNA_R	TGGCTGGCGTCTGTGGCACCCGCT
actin_F	GCCGGGACCTGACTGACTAC
actin_R	AGGCTGGAAGAGTGCCTCAG
bsr_F	CAGGAGAAATCATTTCGGCAGTAC
bsr_R	TCCATTCGAAACTGCACTACCA
chr21_F	GTCTACCTTTTATTTGAATTCCCG
chr21_R	AGGGAATGTCTTCCCATAAAAACT
PABPC-10_F	CTAGCATCCGTGGGCCAAGAG

PABPC-10_R	CTCTTCCCCAACCCCAGCAAAAT
PP1A-3_F	AGTGTGGGGAAGAAAGGAGCCG
PP1A-3_R	CTAAAAGAGACAGTGTTGGTCAGGC
sat2_F	TCGCATAGAATCGAATGGAA
sat2_R	GCATTCGAGTCCGTGGA
tetO_F	CCACTCCCTATCAGTGATAGAGAA
tetO_R	TCGACTTCTGTTTAGTTCTGTGCG
tetO_R2	GTTAAACTCAGTCGTCACCAAGAG

Site-Directed Mutagenesis

Description	**Sequence (5’ - 3’)**
LSD_mut661F	GATTTGGCAACCTTAACGCGGTGGTGTTGTGTTTTG
LSD_mut661R	CAAAACACAACACCACCGCGTTAAGGTTGCCAAATC

Sequencing

Description	**Sequence (5’ - 3’)**	**Target / Construct**
CenA-SEQ1_F	CCGAGTTACTCTCTTCCCAAAG	CENP-A ORF
CenA-SEQ2_R	GAGCTAGACACCACCCGGT	CENP-A ORF
GFP 5’	CATGGTCCTGCTGGAGTTCGT	tetR-EYFP / tYIP backbone
KAPSEQ_A	CAGAGCGTCCTGGCACTAACTC	KAP1 ORF
KAPSEQ_Arev	GAGTTAGTGCCAGGACGCTCTG	KAP1 ORF
KAPSEQ_B	TCCCCGCCTGGCCTCACC	KAP1 ORF
KAPSEQ_Brev	GGTGAGGCCAGGCGGGGA	KAP1 ORF
KAPSEQ_C	AGCCTCGGCAGCAGCGGC	KAP1 ORF
LSD-SEQ1_F	GAAACTGGAATAGCAGAGAC	LSD1/(K661A) ORF
LSD-SEQ1_R	GTCTCTGCTATTCCAGTTTC	LSD1/(K661A) ORF
LSD-SEQ2_F	CTTTTGGAAGCCAGGGATC	LSD1/(K661A) ORF

LSD-SEQ2_R	GATCCCTGGCTTCCAAAAG	LSD1/(K661A) ORF
LSD-SEQ3_F	CAAATACTTGATTGGCATTTTG	LSD1/(K661A) ORF
LSD-SEQ3_R	CAAAATGCCAATCAAGTATTTG	LSD1/(K661A) ORF
LSD-SEQ4_F	GTACCTCAGCCCAAAGAAAC	LSD1/(K661A) ORF
LSD-SEQ4_R	GTTTCTTTGGGCTGAGGTAC	LSD1/(K661A) ORF
M13F(-21)	CGTTGTAAAACGACGGCCAGT	pGEM-T easy backbone
M13R	TTTCACACAGGAAACAGCTATGA	pGEM-T easy backbone
pIRES_5prime	GGGAGGTCTATATAAGCAGAGC	pIRES-Puro2 / pCenA-SNAP-IP backbone
tYIP_3prime	GGTGAATATCAAATCCTCCTCG	tYIP backbone
tYIP_SEQ1	CCGACAATTGCATGAAGAATCTGC	tYIP backbone
tYIP_SEQ2	CTCTGGCTAACTAGAGAACCC	tYIP backbone
tYIP_SEQ3	ATGTGATAGACGCTGCACCACC	tYIP backbone
tYIP_SEQ4	CTGGAGTACAACTACAACAGCC	tYIP backbone

F.4 Antibodies

Primary Antibodies

Antibody (host)	**Source**	**Working Dilution (application)**
CENP-A (rabbit)	Manuel Valdivia	1:250 (IF)
CENP-A, AN1 (mouse)	Hiroshi Masumoto	1:250 (IF) 1:100 (ChIP)
CENP-C, R554 (rabbit)	in house	1:500 (IF)
CENP-H, R1276 (rabbit)	in house	1:250 (IF)
Dsn1 (rabbit)	Iain Cheeseman	1:1,000 (IF)
H2B, ab52484 (mouse)	Abcam	5µg (ChIP)

H3, total, ab10799 (mouse)	Abcam	5µg (ChIP)
H3K27ac, CMA009 (mouse)	Hiroshi Kimura	1:10 (ChIP)
H3K27me1, 1B3 (mouse)	Hiroshi Kimura	1:10 (ChIP)
H3K27me2, 5D1 (mouse)	Hiroshi Kimura	1:50 (ChIP)
H3K27me3, 1E7 (mouse)	Hiroshi Kimura	1:50 (ChIP)
H3K36me1, 1H1 (mouse)	Hiroshi Kimura	1:10 (ChIP)
H3K36me2, 2C3 (mouse)	Hiroshi Kimura	1:50 (ChIP) 1:10 (IF)
H3K4me0, CMA001 (mouse)	Hiroshi Kimura	1:10 (ChIP)
H3K4me1, CMA002 (mouse)	Hiroshi Kimura	1:10 (ChIP)
H3K4me2, 07-030 (rabbit)	Upstate	1µg (ChIP)
H3K4me2, CMA003 (mouse)	Hiroshi Kimura	1:50 (ChIP)
H3K4me3, 05-745 (rabbit)	Upstate	1µg (ChIP)
H3K4me3, CMA004 (mouse)	Hiroshi Kimura	1:50 (ChIP)
H3K9ac, 07-352 (rabbit)	Upstate	1:100 (IF)
H3K9ac, CMA005 (mouse)	Hiroshi Kimura	1:10 (ChIP)
H3K9ac, R607 (rabbit)	Bryan Turner	1:200 (IF)
H3K9me1, CMA006	Hiroshi Kimura	1:10 (ChIP)
H3K9me2, CMA007 (mouse)	Hiroshi Kimura	1:10 (ChIP)
H3K9me3, 07-523 (rabbit)	Upstate	1µg (ChIP)
H3K9me3, 2F3 (mouse)	Hiroshi Kimura	1:50 (ChIP)
H4K16ac, R251 (rabbit)	Bryan Turner	1:200 (IF)
H4K8ac, R403 (rabbit)	Bryan Turner	1:200 (IF)
KNL1 (rabbit)	Iain Cheeseman	1:1,000 (IF)
Mis12 (rabbit)	Mitsuhiro Yanagida	1:500 (IF)
normal mouse IgG	Calbiochem	10µg (ChIP)
RNA Pol II CTD, ab817 (rabbit)	Abcam	5µg (ChIP)
RNA PolII Ser2p, PC26B5 (mouse)	Hiroshi Kimura	1:10 (ChIP)

Secondary Antibodies

Secondary antibodies used for immunofluorescence analysis were conjugated to either FITC, TexasRed or Cy5. All secondary antibodies for immunofluorescence analysis were purchased from Jackson ImmunoResearch Laboratories. Biotinylated anti-avidin antibody for FISH analysis was purchased from Vector Labs and used at a 1:100 dilution.

F.5 Commercial Kits

Description (catalogue number)	**Manufacturer**
DNeasy Blood & Tissue Kit (69504)	Qiagen
QIAfilter Plasmid Midi Kit (12243)	Qiagen
BioNick DNA Labeling System (18247015)	Invitrogen
QIAprep Spin Miniprep Kit (27106)	Qiagen
QIAquick Gel Extraction Kit (28704)	Qiagen
QIAquick PCR Purification Kit (28104)	Qiagen
QuikChange Site-Directed Mutagensis Kit (200518)	Stratagene
Transcriptor High Fidelity cDNA Synthesis Kit (05091284001)	Roche
pGEM-T Easy Vector System I (A1360)	Promega

G MOLECULAR BIOLOGY TECHNIQUES

G.1 Preparation of Bacterial Plasmid DNA

Transformation of Chemically Competent Escherichia coli

Chemically competent *E. coli* were thawed on ice. Typically, an appropriate amount of freshly ligated or super-coiled plasmid DNA was added to 50-100 µl of competent cells and incubated for 20 minutes on ice prior to a 45 sec heat-shock on a 42°C thermoblock. Cells were subsequently allowed to recover for 2 minutes on ice, and 500-900 µl S.O.C. medium (Invitrogen) prewarmed to 37°C were added. Cells were incubated with gentle agitation at 37°C for 45 minutes prior to plating out on pre-warmed LB-Agar plates containing either 50 µg/ml kanamycin sulfate or 100 µg/ml ampicillin, depending on the plasmid-encoded antibiotic resistance. Plates were incubated at 37°C over night.

Isolation of Plasmid DNA from Escherichia coli

Bacterial over night cultures were grown at 37°C by inoculating LB medium supplemented with the relevant antibiotic at the concentrations indicated above with cells derived from a single colony picked from LB-Agar plates. Small-scale (mini-prep) or large-scale (midi-prep) isolation of super-coiled plasmid DNA from *E. coli* cells was subsequently performed using the respective plasmid preparation kit supplied by Qiagen. 1.5 ml or 50 ml over night culture were centrifuged at 4,000g, and the bacterial pellet was processed for mini- or midi-preps, respectively, essentially according to the manufacturer's instructions. Plasmid DNA destined for transfection into eukaryotic cells was typically subjected to a 30 minute treatment with endo-toxin removal buffer (Qiagen) during the process of the midi-prep, as recommended by the manufacturer.

Isolated plasmid DNA was resuspended in double-distilled H_2O, and concentration and quality of midi-prep DNA was determined by spectrophotometric measurement of the optical density (OD) at 260 nm and 280 nm using a Beckman DU530 UV/Vis spectrophotometer.

G.2 Molecular Cloning of Plasmid Constructs

Proof-reading Polymerase Chain Reactions

Where appropriate, the DNA sequence of interest was amplified using proof-reading PCR. To this end, the Expand High Fildelity PCR System (Roche) was employed using sequence-specific oligonucleotide primers. A suitable amount of random hexamer-primed cDNA or, where available, plasmid DNA was used as reaction template. PCR reactions in a final volume of 50 µl were prepared with double-distilled water and contained 2.6U of the supplied enzyme mix, 200 µM of each dNTP, 200 nM of each primer and 1.5 mM $MgCl_2$ in a 1x dilution of the reaction buffer provided.

Reactions were run on a Biometra T3000 thermocycler. Typically, following a two minute initial denaturation / enzyme heat-activation step at 94°C, denaturing of dsDNA was performed for 10 sec at 98°C followed by 30 sec annealing at a primer-specific temperature and subsequent elongation at 72°C for one minute per target kb. This cycle was repeated 20-25 times und completed by a final extension at 72°C for 10 minutes. Specificity and yield of the PCR reaction was subsequently analyzed by agarose gel electrophoresis. Sequence validity of the PCR product was verified by sequencing after subcloning of the product, as described below.

Agarose Gel Electrophoresis and Gel Purification of DNA Fragments

Agarose gels between 0.8 and 3 % agarose (Sigma-Aldrich) were prepared with TAE buffer and were supplemented with 0.3 µg/ml ethidium

bromide. Gel electrophoresis of DNA in a final 1x dilution of Orange G DNA loading buffer was performed in TAE buffer at a constant voltage of 100-120 V for the required time. Desired DNA fragments were excised in agarose gel blocks under visualization with low-intensity UV light. Recovery and clean-up of DNA from excised gel blocks was performed using a gel extraction kit provided by Qiagen following the manufacturer's instructions.

Restriction Digestion of Plasmid or PCR-generated DNA

An appropriate amount of super-coiled plasmid DNA or gel-purified PCR product generated with oligonucleotide primers containing 5' restriction enzyme target sites was digested with the required restriction endonuclease purchased from NEB. Typically, 50 µl reactions were prepared with double-distilled water and contained 5-10 U of the relevant endonuclease and a final 1x dilution of the recommended restriction enzyme buffer provided by NEB. Where recommended, BSA was added to a final concentration of 100 µg/ml. Reactions were incubated at the required temperature for a minimum of two hours prior to analysis on and purification of the digested product from an agarose gel after electrophoresis.

De-Phosphorylation of Restriction Enzyme-digested DNA

Where required, removal of the 5' phosphate group of restriction enzyme-digested plasmid DNA was performed using alkaline Calf Intestinal Phosphatase (CIP) (NEB). The volume of restriction enzyme digestion reactions was adjusted to 100 µl using a 1x dilution of the relevant restriction enzyme buffer to yield a DNA concentration equal to or less than 0.5 µg / 10 µl. 5 U CIP were added, and the reaction was incubated for one hour at 37°C followed by gel purification of the product.

Blunting of Single-stranded DNA Ends

Blunting of protruding single-stranded DNA ends obtained after restriction enzyme digestion was performed by filling in of 5' overhangs using the DNA polymerase I Klenow fragment (NEB). Gel-purified DNA was used in a 50 µl reaction containing 5 U Klenow fragment, 200 µM of each dNTP and 100 µg/ml BSA in a final 1x dilution of NEB restriction enzyme buffer 2. The reaction was incubated for 10 minutes at room temperature prior to heat-inactivation for a further 10 minutes at 75°C after the addition of EDTA to a final concentration of 10 µM. DNA was subsequently purified using a PCR clean-up kit provided by Qiagen, essentially according to the manufacturer's instructions.

DNA Ligation

Ligation of DNA fragments with complementary, single-stranded DNA overhangs was typically performed using Quick T4 DNA ligase (NEB) in a final reaction volume of 10 µl containing 0.5 µl of the Quick T4 DNA ligase and a final 1x dilution of the provided Quick Ligation Buffer. Concentration of vector and insert DNA was estimated from the fluorescence intensity on an agarose gel. A molar ratio of vector-to-insert DNA between 1:3 and 1:6 relative to 50 µg vector DNA was used. Reactions were incubated at room temperature for 20 minutes prior to transformation of the total reaction volume into *E. coli* cells.

Ligation of blunt-ended DNA fragments was performed analogously, but using 400 U T4 DNA ligase (NEB) in an over night reaction at 16°C.

Subcloning of PCR Products

Gel-purified PCR products eluted in double-distilled water were ligated into pGEM-T Easy (Promega) after 3' A-tailing as described below, according

to the manufacturer's instructions. Following ligation reactions, the DNA was transformed into *E. coli*, plated on ampicillin-containing LB agar plates, and plasmids were extracted from individual antibiotic-resistant colonies by mini-prep for subsequent sequence verification.

Standard 3' A-tailing reactions of proof-reading enzyme-generated PCR products prior to subcloning into pGEM-T Easy were performed essentially as recommended by Promega. In brief, 1/20th of the gel-purified PCR products was used in a 10 µl reaction containing 5U Taq polymerase (Roche), 2.5 mM $MgCl_2$ and 200 µM dATP in a final 1x dilution of the Taq reaction buffer supplied by Roche. The reaction was incubated at 72°C for 30 minutes on a thermocycler, and 2 µl tailed product were used for immediate ligation into pGEM-T Easy.

Sequencing of Plasmid Constructs

Sequencing reactions were performed based on the di-deoxynucleotide method, using the BigDye v3.1 Cycle Sequencing Kit provided by Applied Biosystems. 100-250 ng super-coiled plasmid DNA, 1µM of the sequencing primer and 4 µl BigDye mix were used in a final reaction volume of 10 µl. Following an initial denaturation step for two minutes at 96°C, cycle parameters were 30 sec denaturation at 96°C, 15 sec annealing at 50°C and four minutes extension at 60°C, for a total of 25 cycles.

Sequencing reactions were further processed and loaded for analysis on an ABI 3730 DNA Analyzer (Applied Biosystems) by The GenePool Sequencing Facility (The University of Edinburgh). Sequence analysis of associated ABI sequence files was performed using Sequencher software (Gene Codes Corporation). Seriously guys - are you actually reading this?!

G.3 Expression Constructs

tetR-EYFP Expression Constructs

tetR-EYFP: expressing a fusion of the tet repressor containing the SV40 nuclear localization signal (NLS) to EYFP under control of a CMV promoter, corresponds to TetR:EYFP described previously in Nakano et al. (Nakano et al., 2008).

tYIP: to create a vector expressing the above fusion construct and conferring resistance to puromycin through an internal ribosomal entry site, tetR-EYFP was digested with BsaHI and BspEI. The resulting 1.5 kb fragment containing the coding sequence for tetR-NLS-EYFP was subsequently ligated into pIRESpuro2 (Clontech) digested with ClaI and BspEI.

tYIP-tol2: the 1.2 kb EcoRI fragment of pGEM-T-tol2 (Kumiko Samejima) containing the tol2 cassette was blunted and cloned into the de-phosphorylated NruI site of tYIP.

Generation of KAP1-derived tetR-EYFP Fusion Constructs

tetR-EYFP-KAP1: pC3-FLAG-KAP1 (a kind gift of David Schultz) was digested with EagI and EcoRI. The 2.7 kb fragment containing the coding sequence corresponding to wild-type KAP1 amino acids 38-835 was blunted and ligated into the blunted and de-phosphorylated BglII site of tetR-EYFP (Note: the 2.7 kb fragment contains an additional ~150 bp sequence downstream of the KAP1 stop codon that were not indicated in the pC3-FLAG-KAP1 sequence provided by D. Schulz).

tetR-EYFP-RBCC: the sequence coding for the N-terminal KAP1 Krueppel-associated box binding domain containing the Ring-B-box-coiled-

coil (RBCC) motif was amplified from pC3-FLAG-KAP1 using KRBD_1_Fwd/Rev. The 1.2 kb PCR product was subcloned into pGEM-T Easy, excised with SmaI and EcoRI and cloned blunt into the BglII site of tetR-EYFP. *[The final cloning step of this construct was performed by Stefano Cardinale.]*

tetR-EYFP-HP1BD: the sequence coding for the central Heterochromatin Protein (HP)1 binding domain of KAP1 was amplified from pC3-FLAG-KAP1 using HP1BD_1_Fwd/Rev. The 0.5 kb PCR product was subcloned into pGEM-T Easy, excised with NcoI and EcoRI and cloned blunt into the BglII site of tetR-EYFP.

tetR-EYFP-PHD/Bromo: the sequence coding for the C-terminal PHD/bromo-like domain motif of KAP1 was amplified from pC3-FLAG-KAP1 using PHD_1_Fwd/Rev. The 0.7 kb PCR product was subcloned into pGEM-T Easy, excised with SmaI and EcoRI and cloned blunt into the BglII site of tetR-EYFP.

tetR-EYFP-KAP1[38-559]: this construct was generated by converting the codon of amino acid residue 560 of KAP1 into a stop codon, employing site-directed mutagenesis using tetR-EYFP-KAP1 as template. *[This construct was generated by Stefano Cardinale.]*

tYIP-KAP1: the 2.5 kb BspEI/BclI fragment of tetR-EYFP-KAP1 was ligated into tYIP digested with BspEI and BamHI.

Generation of LSD1-derived tetR-EYFP Fusion Constructs

tYIP-LSD1: human LSD1 cDNA (NCBI accession BC048134) was purchased in pBluescript from SourceBioScience / GeneService. The full-length coding sequence of LSD1 was amplified with LSD1_F/R and the PCR

product digested with BglII. The 2.6 kb fragment was ligated into the BamHI restriction site of tYIP. For stable transfection, an analogous construct was generated using tYIP-tol2 as target vector.

tYIP-LSD1(K661A): alteration of the coding sequence of LSD1 to yield the indicated amino acid substitution was introduced into the LSD1 cDNA in pBluescript by site-directed mutagenesis using LSD_mut661F/R. tYIP-LSD1(K661A) was then generated analogous to tYIP-LSD1 above. For stable transfection, an analogous construct was generated using tYIP-tol2 as target vector.

Generation of tetR-EYFP Fusions with Acidic Activator Domains

tYIP-VP16: the coding sequence for the VP16 acidic activation domain was amplified from pTet-On (Clontech) using VP16-RE_F/R. The PCR product was subcloned into pGEM-T Easy. The plasmid was digested with BamHI and the 0.4 kb fragment was ligated into the BamHI site of tYIP.

tYIP-p65: the p65 acidic activation domain was amplified using p65_F/R from NYE108 (a kind gift of Andrew Belmont) and sub-cloned into pGEM-T Easy. The 0.1 kb BspEI fragment was ligated into the BspEI site of tYIP. *[This construct was generated by Julia Jakubsche under my direct supervision.]*

Generation of a 3xHA-SNAP-tagged CENP-A Expression Construct

pCenA-SNAP-IP: pLJ184 (a kind gift of Lars Jansen), containing the coding sequence for a CENP-A-SNAP-3xHA fusion construct was digested with BglII an NotI. The 1.1 kb fragment was ligated into pIRESpuro2 (Clontech) digested with BamHI and NotI.

G.4 Reverse Transcriptase Polymerase Chain Reaction

Reverse Transcriptase (RT)-PCR was typically performed on total RNA extracted with TRIzol reagent (Invitrogen), using the Transcriptor High Fidelity cDNA Synthesis Kit supplied by Roche, essentially according to the manufacturer's instructions. All reaction steps were carried out on a Biometra T3000 thermocycler. For real-time RT-PCR analysis of centromeric transcripts, synthesis of cDNA from 2 µg RNA was primed using 60 µM random hexamer primers, and reverse transcription was performed for 30 minutes at 50 ℃.

G.5 Real-time Polymerase Chain Reaction Analysis

Real-time PCR was performed using a SYBR Green master mix containing the hot start enzyme, dNTPs and SYBR Green dye (JumpStart, Sigma), in a final reaction volume of 20 µl. All oligonucleotide primers were used at a final concentration of 400 nM. Reactions were run on a LightCycler 480 system (Roche) in 96 well plates supplied by Roche. Cycle parameters are provided in Appendix I. Some early ChIP analysis was performed on an iCycler system (BioRad) in 96 well plates supplied by Eurogentec. For every plate and primer pair used, a log-scale standard curve was prepared from a serial dilution of relevant template DNA to determine and account for differential reaction efficiencies. Specificity of reactions was validated by product melting curve analysis. Reaction crossing points (LightCycler 480) or cycle thresholds (iCycler) were determined using the corresponding software, and values were exported to, further processed and analyzed in Excel (Microsoft). Crossing points (LightCycler 480) were determined using the 2^{nd} derivative maximum algorithm. Experiments run on either system yielded comparable over-all results.

Real-time PCR Analysis of Chromatin Immunoprecipitation Experiments

Typically, 2 µl of Input or ChIP'ed DNA resuspended in 45 µl TE were used in each reaction. To account for differential primer efficiencies and quality of sonicated chromatin, standard curves were prepared from the corresponding Input material. The following primer pairs were used for analysis (refer to section D.3): tetO_F/R for the alphoid^{tetO} array; chr21_F/R for the α21-I satellite locus (alphoid$^{chr.21}$); bsr_F/R for the blasticidin S resistance marker; 5SrDNA_F/R for endogenous 5S ribosomal DNA locus; sat2_F/R for the pericentromeric satellite 2 repeats; PABPC-10_F/R for the +10 kb region of the *PABPC1* gene; PP1A-3_F/R for the +3 kb region of the *PP1A* gene.

For each antibody and locus analyzed by ChIP, the amount of ChIP'ed DNA was calculated and presented as percentage of input material added to the beads. For histone ChIP analysis of the same loci within the same cell line over time or under different experimental conditions, these values were further normalized to the % of input value of the α21-I satellite locus following subtraction of background signal (IgG). For RNA polymerase II Ser2p ChIP, background-subtracted % of input values were normalized to those of *PP1A*.

Real-time RT-PCR Analysis

A cDNA amount equivalent to ~70 ng (alphoid^{tetO} and alphoid$^{chr.21}$) or ~7 ng (BSr and β-actin) input RNA was used per reaction. Standard curves were prepared from genomic DNA extracted from the corresponding cell line grown under standard culture conditions, thereby normalizing relative transcript copy numbers to the copy number of the corresponding genomic loci. The following primer pairs were used for analysis (refer to section D.3): tetO_F/R2 for the alphoid^{tetO} array; chr21_F/R for the α21-I satellite locus

(alphoid$^{chr.21}$); bsr_F/R for the blasticidin S resistance marker and actin_F/R for the β-actin gene.

After normalizing to the genomic locus copy number as illustrated above, transcript levels were further normalized to those of β-actin to account for variations in the amount of cDNA used. Data is presented either as relative transcript level or as copy number relative to that of alphoid^{tetO} at the initial time point.

H CELL BIOLOGY TECHNIQUES

H.1 Transfection with Plasmid DNA

Transient Transfections

Transient transfections were performed using either Fugene HD (Roche) or Fugene 6 (Roche). Cells were grown in cRPMI and typically seeded in 6 well plates so as to achieve 80-90 % or 50-60 % confluency, respectively, on the day of transfection. The transfection complex was prepared essentially according to the manufacturer's instructions in 100 µl OptiMEM (Invitrogen), containing 3 µl of the transfection reagent and 1 µg of high-quality supercoiled plasmid DNA. Up- or downscaling of the transfection complex mixture was done proportionally to the surface area of the culture vessel used. Were desired, puromycin was added one day after transfection for 24 hours or longer to enrich for transfected cells. Transfected cells used for RT-PCR or ChIP analysis were typically selected for in 2 µg/ml or greater than 3µg/ml puromycin, respectively.

Stable Transfections

Stable transfections were carried out in 6cm cell culture dishes using Fugene HD in a transfection mix prepared in 250 µl OptiMEM (Invitrogen), containing 15 µl transfection reagent and 5 µg high-quality supercoiled plasmid DNA, essentially as recommended by the manufacturer. 1C7-KAP1 cells were generated using tYIP-KAP1. 1C7-LSD1^{WT} and 1C7-LSD1^{K661A} cells were generated using tYIP-tol2-LSD1 or tYIP-tol2-LSD1(K661A), respectively.

Where the tol2 retrotransposon system was used to facilitate stable integration of the plasmid DNA into the host genome, the transfection mix contained 1µg vector DNA expressing the tol2 enzyme (Kumiko Samejima).

One day after transfection, the medium was replaced with fresh cRPMI containing the required drugs for selection, and stable clones were allowed to grow in the presence of selection until reaching confluency. Isolation of stable clonal cell lines was subsequently performed by standard limiting dilution in 96 well plates. 1C7-derived stable cell lines were routinely maintained in cRPMI containing 4 µg/ml blasticidin S, 1-2 µg/ml puromycin and 1 µg/ml doxycycline.

H.2 Flowcytometric Analysis

Standard Flowcytometric Analysis

For standard flowcytometric analysis of adherent cell lines expressing fluorescent fusion constructs, cells were harvested using TrypLE Express (Invitrogen), washed twice in pre-warmed D-PBS (Invitrogen) and fixed in 2 % PFA/PBS for five minutes at room temperature. Cells were washed again in PBS, and flowcytometric analysis was carried out using a FACSCalibur flow cytometer (BD BioSciences) using CellQuest Software (BD BioSciences).

Cell Cycle Analysis Using Propidium Iodide Staining

Cells were harvested using TrypLE Express (Invitrogen) and washed once in D-PBS (Invitrogen). Cells were then resuspended in ice-cold D-PBS to a final concentration of $1x10^6$ / ml and 1 ml of the cell suspension was added drop-wise to 9 ml ice-cold 70 % ethanol under constant agitation. Cells were incubated in ethanol for 16 h at 4°C. Ethanol-fixed cells were washed twice in ice-cold PBS and subsequently stained in freshly prepared propodium iodide staining solution (0.2 % Triton X-100 / PBS; 50 µg/ml RNAseA; 20 µg/ml propidium iodide) for 15 minutes at 37°C prior to flowcytometric analysis on a FACSCalibur flow cytometer (BD BioSciences) using CellQuest Software (BD BioSciences).

H.3 Preparation of Chromosome Spreads

Spreads of mitotic HT1080 chromosomes for FISH analysis were directly prepared from cells grown on poly-lysine slides. Cells were seeded on poly-lysine slides and incubated for six hours in 0.1 µg/ml colcemid (KaryoMax, Invitrogen) at 37°C. Slides were rinsed with pre-warmed D-PBS, cells were treated in hypotonic buffer (75mM KCl) for 10 minutes at 37°C, and subsequently fixed in ice-cold 3:1 methanol:acetic acid for 10 minutes at -20°C. Slides were stored in fresh fixative at -20°C until processed for FISH staining.

H.4 Immunofluorescence Analysis

Direct and Indirect Immunofluorescence Staining

Cells expressing fluorescently labelled fusion constructs and/or to be processed for indirect immunofluorescence staining were grown on sterile no. 1.5 glass cover slips. Cover slips were washed twice in pre-warmed D-PBS and cells were subsequently fixed in freshly prepared 4% PFA (Electron Sciences) in PBS for 10 minutes at room temperature. Cells were permeabilized for five minutes in 0.2 % Triton X-100 / PBS and either counterstained in 1 µg/ml DAPI (Sigma) in PBS and mounted on Superfrost glass slides in VectaShield (Vector Labs), or further labelled with antibodies as follows.

Fixed and permeabilized cells were pre-blocked in 3 % BSA/PBS-T for 15 minutes at 37°C prior to incubation with appropriately diluted primary antibody in 1 % BSA / PBS-T for one hour at 37°C. Cells were washed in PBS-T and subsequently incubated for 30 minutes at 37°C with fluorophore-conjugated secondary antibody (Jackson Labs) diluted 1:200 in 1 % BSA / PBS-T. After excess antibody was washed off in PBS, cells were counterstained and mounted as above.

Staining of Unfixed Mitotic Chromosomes

Staining of unfixed mitotic chromosomes was essentially performed according to a protocol described in Keohane et al (Keohane et al., 1996). In brief, cells were incubated in 0.1 µg/ml colcemid (KaryoMax, Invitrogen) for two hours at 37℃. Mitotic cells were collected by shake-off and washed once in pre-warmed D-PBS. Cells were resuspened in 75mM KCl to a concentration of $2x10^5$ / ml and incubated for 10 minutes at room temperature. 100-200µl cell suspension were then cytospun on Superfrost glass slides at 800g for five minutes at 4℃. Slides were immersed in 4℃ KCM buffer for 10 minutes prior to incubation with appropriately diluted primary antibody in 1 % BSA / KCM for one hour at 4℃. Cells were washed in KCM buffer and incubated for another hour with fluorophore-conjugated secondary antibody diluted 1:200 in 1 % BSA/KCM at 4℃. Excess antibody was washed off with KCM buffer and cells were fixed for 10 minutes in 4% PFA (Electron Sciences) in KCM buffer at room temperature prior to counterstaining in 1 µg/ml DAPI in PBS and mounting in VectaShield.

Fluorescence In Situ *Hybridization*

Biotinylated probe DNA to specifically label the alphoid^{tetO} HAC backbone was generated by nick translation using the BioNick Labeling System supplied by Invitrogen, essentially according to the manufacturer's instructions, using 1 µg of BAC32-2mer DNA as template in each reaction. Reactions were allowed to proceed for two hours at 16℃, and probe yield and fragment size was assessed by agarose gel electrophoresis. Probe DNA was purified using a G25 spin column (GE Healthcare) essentially as recommended by the manufacturer. Purified probe DNA was supplemented with 50 µg salmon sperm DNA (Invitrogen) per reaction, ethanol-precipitated and subsequently resuspended in hybridization buffer.

Methanol/acetic acid-fixed chromosome preparations where aged by dying the slides at room temperature over night. 150-300 ng of biotinylated probe DNA initially denatured at 95℃ were added to each slide and covered with a cover glass. Samples and probe where subsequently denatured at 75℃ for two minutes. Following annealing at 39℃ over night, slides were extensively washed in 0.1x SSC buffer at 65℃ followed by one wash in room temperature 4x SCC buffer containing 0.1 % Tween-20. Primary labelling of annealed probe DNA was done using FITC-avidin (Promega) at a 1:500 dilution for 30 minutes at 37℃. Slides were washed in 4x SSC buffer and incubated with biotinylated anti-avidin antibody at 37℃ for 30 minutes prior to washes in 4x SSC and secondary labelling with FITC-avidin as above. Slides were counterstaind with DAPI as described above and subsequently mounted in VectaShield.

Immunofluorescence Microscopy

Brightfield Deconvolution Microscopy was typically performed using a DeltaVision Core system (Applied Precision), based on an inverted Olympus IX-71 microscope stand equipped with an Olympus UPlanSApo 100x oil immersion objective (NA 1.4) and a 250 W Xenon light source. Camera (CoolSnap HQ, Photometrix), shutter and stage were controlled through SoftWorx (Applied Precision). Image z-series were collected with a spacing of 0.2 or 0.3 µm and subsequently subject to deconvolution using a conservative algorithm in SoftWorx. For display purposes, images are presented as either maximum intensity projections or single focal planes.

Fluorescence Signal Quantification

All fluorescence signal quantification was performed on deconvolved images acquired at a 1x1 binning, using identical exposure conditions for each experimental subset. To avoid variation in signal intensities due to

chromatic and spherical lens aberration, HACs were centered and a 512 x 512 pixel field of view was imaged. Only cells displaying a single HAC were used for quantification of the associated signal. Furthermore, for CENP signal quantification, only those cells that showed a clear overlap of CENP staining and EYFP (HAC) signal were considered. Quantification was performed in ImagePro 6.3 or ImagePro Plus (Media Cybernetics) employing different custom-written macros *[kindly composed by David A. Kelly]*, details of which are provided in Appendix II.

H.5 Doxycycline Wash-Out Experiments

To prevent binding of tetR fusion constructs to the alphoid^{tetO} array, all stably expressing cell lines were grown in the presence of 1 µg/ml doxycyline in the culture medium, which was changed every two days. In the interest of data reproducibility, a defined protocol to initiate targeting of tetR fusion constructs was established. Cells were seeded in cell culture dishes, and wash-out was initiated the subsequent day on cultures of ~50 % confluency. Culture medium was removed and the cell layer was rinsed twice with excess D-PBS pre-warmed to 37°C. Cells were then incubated in cRPMI without doxycycline for 30 minutes at 37°C. These steps were repeated once followed by a final two washes in D-PBS. Cells were then incubated in cRPMI supplemented with the required concentration of puromycin, but lacking both doxycyline and blasticidin S, for 16 hours at 37°C. Following this, cells were rinsed twice with pre-warmed D-PBS, harvested with TrypLE Express (Invitrogen) and re-seeded in cell culture dishes as required.

H.6 SNAP-based Quench-Pulse-Chase Experiments

1C7 cells were seeded on coverslips and transiently co-transfected with 0.8 µg of either tetR-EYFP fusion construct and 0.2 µg pCenA-SNAP-IP using 3 µl Fugene 6 (Roche) in a transfection mix of 100 µl OptiMEM,

essentially according to the manufacturer's instructions. The next day, thymidine (Sigma) was added to a final concentration of 4 mM, and transfected cells were enriched for by selection in the presence of 1 µg/ml puromycin. 20 hours after addition of thymidine, existing SNAP-tag was quenched for 20 minutes in medium containing 10 µM non-fluorescent bromothenylpteridine (SNAP-Cell Block, NEB), and cells were subsequently released from thymidine arrest. 20 hours after release, newly-synthesized SNAP-tagged CENP-A was fluorescently labelled in medium containing 3 µM TMR-Star (NEB) for 30 minutes. Cells were fixed and processed for fluorescence microscopy as described above two hours after labelling.

I BIOCHEMICAL TECHNIQUES

I.1 Extraction of Total Cellular RNA

Total RNA was extracted from sub-confluent cultures using TRIzol reagent (Invitrogen), essentially according to the manufacturer's instructions, except that RNA was precipitated using Isopropanol instead of ethanol. RNA precipitates were then resuspended in double-distilled water, and RNA concentration and qualitiy was determined by measuring the absorbance spectrum on a Nanodrop2000 system (Thermo Scientific). OD260/280 ratios of greater than 1.95 were routinely achieved.

I.2 Extraction of Genomic DNA

Sub-confluent cell cultures were harvested, cells were washed once in D-PBS and then resuspended in TE buffer containing 0.5 % SDS. RNAse A was added to a finale concentration of 20 µg/ml and lysates were incubated for one hour at 37°C. Subsequently, proteinase K was added to a final concentration of 100 µg/ml, and lysates incubated over night at 50°C. Genomic DNA was isolated by standard Phenol/Chloroform extraction, ethanol-precipitated and resuspended in TE buffer. Concentration and quality of the DNA was assessed by spectrophotometric determination of the OD at 260nm and 280nm using a Beckman DU530 UV/Vis spectrophotometer.

Genomic DNA used for quantitative determination of HAC copy number in transiently transfected 1C7 cells was isolated using the DNeasy Blood and Tissue Kit provided by Qiagen.

I.3 Chromatin Immunoprecipitation Analysis

Chromatin Immunoprecipitation using Anti-Mouse IgG Dynabeads

Exponentially growing cells where washed in D-PBS (Invitrogen), harvested with TrypLE Express (Invitrogen), resuspended in D-PBS to a final concentration of $1x10^6$ / ml and crosslinked in 1 % formaldehyde (Fischer Scientific) for 5 minutes at room temperature, followed by quenching with 0.5M glycine for an additional 5 minutes. Cells where washed in TBS and $5x10^6$ cells where lysed in Lysis Buffer containing protease inhibitors (1 µg/ml CLAP; 0.5 µg/ml aprotinin; 1 mM PMSF) for 10 minutes on ice. Nuclei where briefly washed in Lysis Buffer containing protease inhibitors and resuspended in 300 µl Dilution Buffer 1 containing protease inhibitors. Chromatin was sheared by sonication in a Bioruptor sonicator (Diagenode) for 14 cycles of 30 sec on / 30sec off at a high setting at 4°C. Supernatants where diluted with 300 µl Dilution Buffer 1, 500 µl Dilution Buffer 2 and 500 µl RIPA buffer containing 150 mM NaCl (RIPA-150) containing protease inhibitors. Anti-mouse IgG Dynabeads M-280 (Invitrogen) where pre-blocked in PBS / 0.5 % BSA for 30 minutes at 4°C and subsequently coupled with the relevant antibodies for 4-6 hours in RIPA-150 / 0.5 % BSA at 4°C, washed twice in RIPA-150 / 0.5 % BSA, and 500 µl of the sheared chromatin was incubated with the beads over night at 4°C.

Beads where washed twice with RIPA-150 containing protease inhibitors, followed by two washes in RIPA-500 and a final wash in TE pH8.0. Antibody/chromatin complexes where eluted at 65°C in TE / 1 % SDS. An equal volume of Post-Elution Buffer was added and crosslinks reversed at 65°C over night. Samples where treated with RNAse A and Proteinase K followed by phenol/chloroform extraction. 20 µg glycogen (Roche) were added, and DNA was precipitated using isopropanol. DNA precipitates were finally resuspended in TE.

Chromatin Immunoprecipitation using Protein G Sepharose Beads

Crosslinking of cells was performed essentially as described above, except that the final concentration of formaldehyde was adjusted to 0.5 %. $5x10^6$ crosslinked cells were resuspended in 400 µl Sonication Buffer containing protease inhibitors (1.5 µM aprotinin; 10 µM leupeptin) and 40µM MG132. Cells were sonicated using a Bioruptor sonicator (Diagenode) for 8 cycles of 30 sec on / 30 sec off at a high setting at 4°C. Sonicated chromatin was cleared by centrifugation. 80 µl chromatin was supplemented with 120 µl 1.7x IP Buffer and incubated with the relevant antibodies over night at 4°C.

40 µl Protein G Sepharose beads (Invitrogen) were pre-blocked in Wash Buffer containing 5 % BSA for 30 minutes at 4°C and subsequently resuspended with an equal volume Wash Buffer. Pre-blocked beads were added to the antibody/chromatin suspension and incubated for 45 minutes at 4°C. Beads were spun down at 2,000rpm for one minute and washed three times in Wash Buffer containing protease inhibitors and MG132. Beads were then incubated in 100 µl TE / 1 % SDS for 15 minutes at 65°C. Supernatants containing eluted antibody/chromatin complexes were mixed with 150 µl ProtK Buffer and incubated for three hours at 37°C followed by an over night incubation at 65°C to reverse crosslinks. Samples were treated with 100 µg/ml RNAse A for one hour at 37°C, and DNA was isolated by phenol/chloroform extraction, precipitated with isopropanol in the presence of 20 µg glycogen (Roche) and resuspended in TE buffer as above.

Results

J THE CHROMATIN LANDSCAPE OF ACTIVE CENTROMERES

J.1 Background

The centromere is commonly defined as a specialized chromatin domain. This view is reinforced by the specific localization of nucleosomes containing the histone H3 variant CENP-A to active centromeres, including those of human artificial chromosomes as well as neocentromeres (Alonso et al., 2003; Masumoto et al., 1998; Warburton et al., 1997). Recent work has provided evidence that CENP-A-containing nucleosomes are organized into discontinuous domains and are interspersed with nucleosomes containing canonical histone H3 (Blower et al., 2002). This “centrochromatin” is characterized by the absence of detectable lysine acetylation, H3K4me3 or H3K9me3, but displays the presence of nucleosomes dimethylated on H3K4 (Lam et al., 2006; Sullivan and Karpen, 2004). As pointed out previously, post-translational histone modifications act at least in part to define a chromatin environment that may serve nuclear structure and function through signalling in a large variety of pathways. While it is therefore plausible to assume that histone modifications present within the centromeric chromatin domain constitute a comparable functional entity that may be required for the maintenance of centromere identity, little or nothing is known about if and to which extent centrochromatin affects centromere structure or kinetochore function. Initial studies involving the targeted manipulation of the alphoid^{tetO} HAC centromere as summarized above strongly support that the underlying chromatin state is indeed important for kinetochore function (Nakano et al., 2008). I will provide and discuss additional evidence for this view in the subsequent sections of this thesis.

Earlier studies assessing histone modifications associated with mammalian satellite DNA were largely biased towards the investigation of stereotypical marks associated with classical euchromatic or heterochromatic character (Lam et al., 2006; Peters et al., 2003; Sullivan and Karpen, 2004). This provides a highly simplistic picture of the local chromatin state that is prone to misinterpretation. The aim of this section is therefore to provide a more comprehensive view of the state of centromeric chromatin through the use of a wide panel of specific and well-characterized antibodies.

As result of the comparatively small size of alpha-satellite input DNA, functional centromeres of HACs may essentially be regarded as minimal or concentrated centromeres. Importantly, sequence-specific features of HAC alpha-satellite DNA allow us to assess a single-copy centromere *in vivo*, thereby by-passing the inherent multi-ploidy of genomic regions. This could in principle further enhance the sensitivity of sequence-based analysis by a factor of two or greater relative to the diploid or multiploid nature of endogenous chromosomes. In addition, endogenous centromeres frequently share the same or similar alpha-satellite monomers, rendering sequence specificity even more difficult. I therefore reasoned that analysis of the alphoid^{tetO} HAC centromere will provide a distilled view of the epigenetic centromere state that - if such a common state exists - would allow to identify common and conserved features of the underlying chromatin and to postulate a functional regulatory interplay of the local chromatin based on established findings from non-centromeric genomic regions.

J.2 Results

Active Centromeres Share a Distinct Histone Modification Profile and Display the Signature of RNA Polymerase II Activity

To establish the histone modification profile associated with active centromeres, ChIP analysis followed by quantification using real-time PCR,

employing a large collection of well-characterized monoclonal antibodies ((Kimura et al., 2008) and unpublished data; a kind gift of Hiroshi Kimura) was performed on asynchronously growing HT1080-derived AB2.2.18.21 cells (Nakano et al., 2008). FISH analysis using a DNA probe specific for the BAC vector sequence confirmed that these cells carried the alphoid^{tetO} HAC as a single copy (Fig. 1A). Co-staining of CENP-A and CENP-C overlapped with the distinct HAC-associated EYFP signal on the mitotic HAC in AB2.2.18.21 cells transiently expressing tetR-EYFP, further confirming that a normal kinetochore structure is assembled on the alphoid^{tetO} array (Fig. 1B). As expected, ChIP analysis showed strong enrichment for CENP-A compared to the negligible enrichment when using unspecific IgG on both the synthetic (alphoid^{tetO}) HAC centromere and the endogenous α21-I satellite of chromosome 21 (alphoid$^{chr.21}$) (Fig. 2). In contrast, the endogenous 5S rDNA locus failed to show any significant association with CENP-A.

Strikingly, both synthetic and endogenous centromeric loci showed a comparable tendency for relative enrichment or depletion across the majority of histone modifications analyzed. H3K4me1, H3K9me3, H3K27me1 and H3K27me3 were clearly detectable at both centromeres. Histone modifications associated with the promoter region of actively transcribed genes, including H3K4me3, acetylation of H3K9 and acetylation of H3K27, were present only at very low levels at both the synthetic and the chromosome 21 centromeres. Unmethylated H3K4, H3K4me2, H3K36me1 and H3K27me2 were less clear-cut, with H3K4me2 displaying high levels at the alphoid^{tetO} HAC centromere and low, but detectable levels associated with the considerably larger chromosome 21 centromere locus.

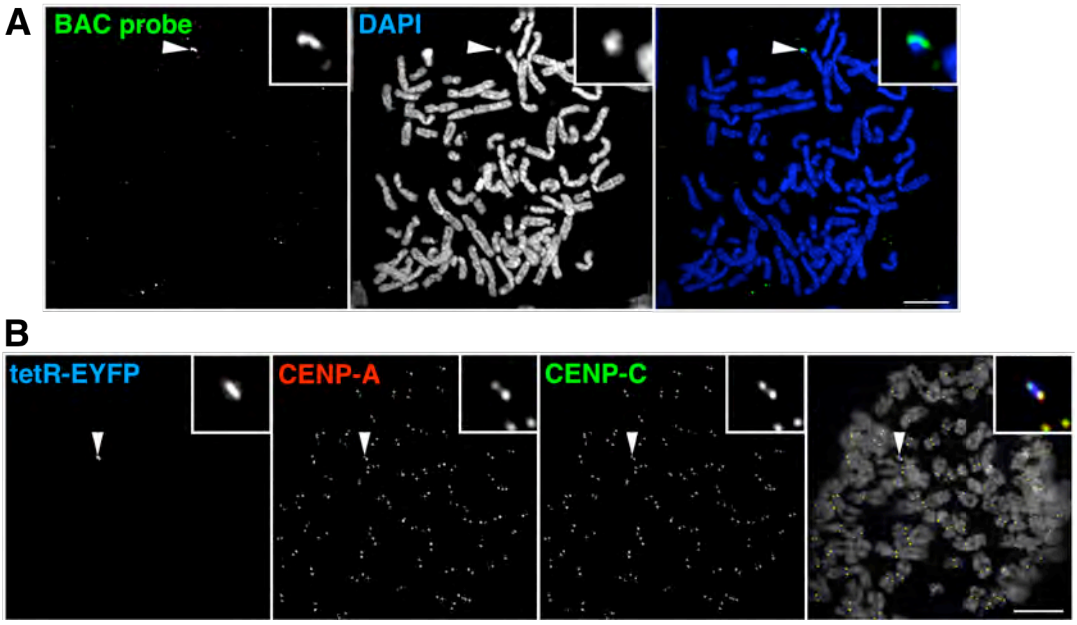


Figure 1| The alphoid^{tetO} HAC in AB2.2.18.21 cells. A) FISH analysis of mitotic AB2.2.18.21 cells using a DNA probe (green) specific for the BAC vector backbone identifies the single-copy alphoid^{tetO} HAC. **B)** AB2.2.18.21 cells were transiently transfected with a construct expressing tetR-EYFP. The EYFP signal marking the alphoid^{tetO} array of the HAC co-localizes with distinct staining for CENP-A and CENP-C. Scale bar: 5 µm.

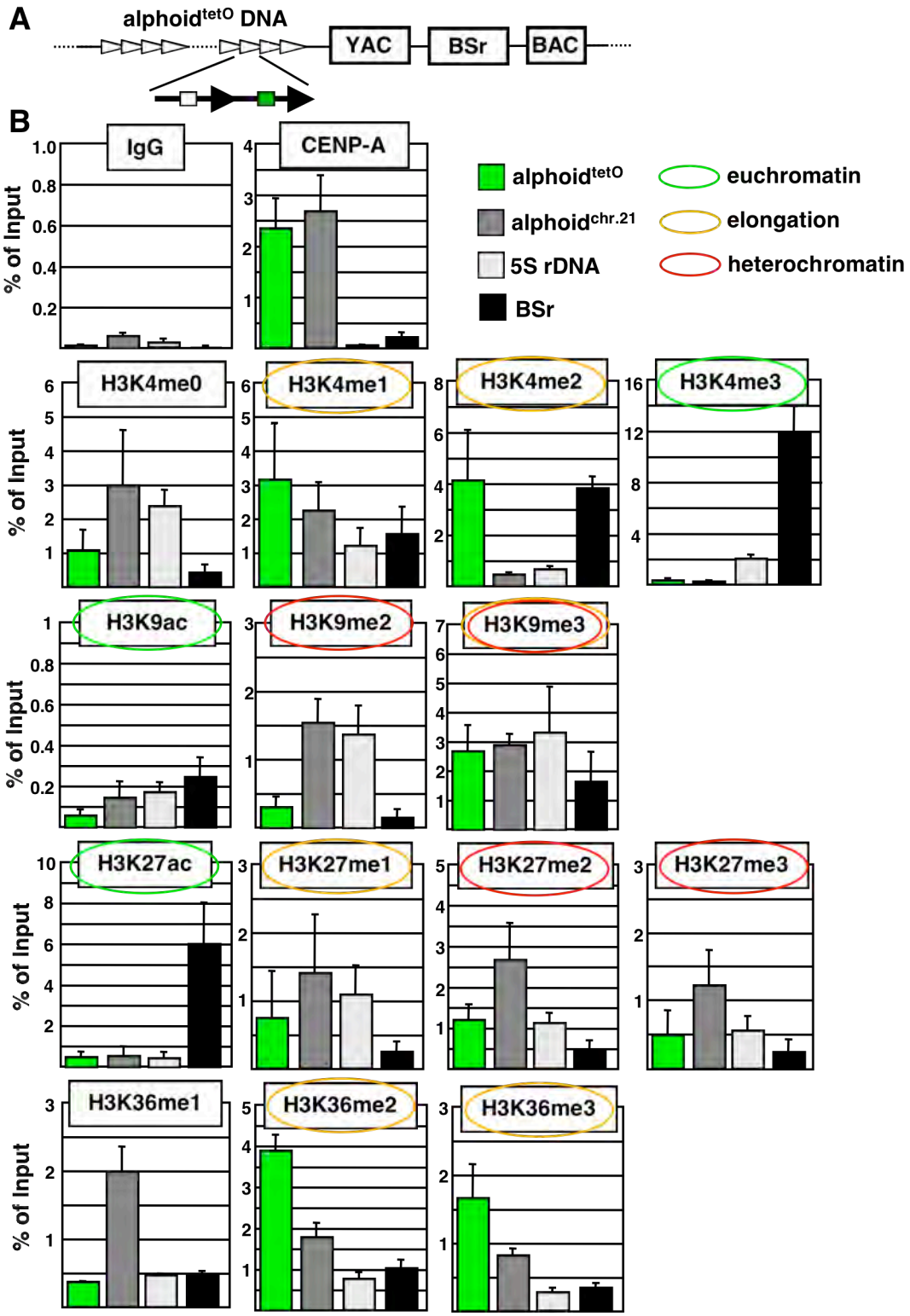


Figure 2| Chromatin of active centromeres displays the signature of RNA polymerase II activity. *Figure legend continued overleaf.*

Figure 2 continued

A) Schematic drawing of the alphoid^{tetO} HAC input vector indicating the higher-order array of synthetic alphoid monomers, containing the tet operator sequence (green boxes) in every second monomer. **B)** ChIP analysis of HT1080 cells using monoclonal antibodies with the indicated reactivity or normal mouse IgG as control. Oligonucleotide primers specific for the HAC (alpoid^{tetO}) or endogenous chromosome 21 (alphoid$^{chr.21}$) centromeres, the actively transcribed blasticidin S resistance marker (BSr) on the HAC vector backbone and the endogenous 5S rDNA locus are used for analysis. ChIP'ed and input material was analyzed by quantitative real-time PCR and is expressed as percentage of the amount of input DNA (% of Input). Histone modifications associated with "strong" euchromatic character or heterochromatin are highlighted in green and red, respectively. Histone modifications associated with elongating RNA polymerase and found within the body of transcribed genes (see main text) are highlighted in yellow. Note the differential scaling of individual panels. Data represents mean and standard deviation of three independent ChIP experiments.

Surprisingly, both centromeres showed a distinct enrichment for H3K36me2 and H3K36me3, histone modifications that together with hypoacetylation, H3K4me1 and H3K4me2 are signature marks found within the body of actively transcribed vertebrate genes (Barski et al., 2007; Schneider et al., 2004; Vakoc et al., 2006). Real-time RT-PCR analysis of total AB2.2.18.21 RNA was therefore performed to assess if alphoid^{tetO} transcripts could be detected. Strikingly, not only where low levels of alphoid^{tetO} and alhoid$^{chr.21}$ transcripts detectable, but normalization of the relative transcript levels to the copy number of the corresponding genomic loci revealed virtually identical transcript-to-template ratios for both centromeres, implying a comparable level of transcriptional activity (Fig. 3A). In a control experiment, detection of these transcripts was sensitive to low doses of actinomycin D (Fig. 3B).

Methylated H3K36 Forms Part of the Mammalian CENP-A Chromatin Domain

Detection of H3K36 methylation within active centromeres was surprising. To independently assess the localization of methylated H3K36 at mammalian centromeres, immunofluorescence labelling of unfixed HT1080 (AB2.2.18.21) chromosomes was performed. Distinct staining for H3K36me2 clearly overlapped with staining for CENP-A on mitotic chromosomes (Fig. 4A). Importantly, investigation of extended kinetochore chromatin fibres revealed that nucleosomes methylated at H3K36 were interspersed with CENP-A-containing nucleosomes, and also extended beyond the CENP-A nucleosome domain (Fig. 4B). In contrast, and consistent with above ChIP data as well as previously published results (Sullivan and Karpen, 2004), centromeric acetylation of H3K9 was virtually undetectable by immunofluoresence (Fig. 4A). Together, I conclude that methylated H3K36 forms a genuine, previously unrecognized component of the chromatin underlying functional kinetochores.

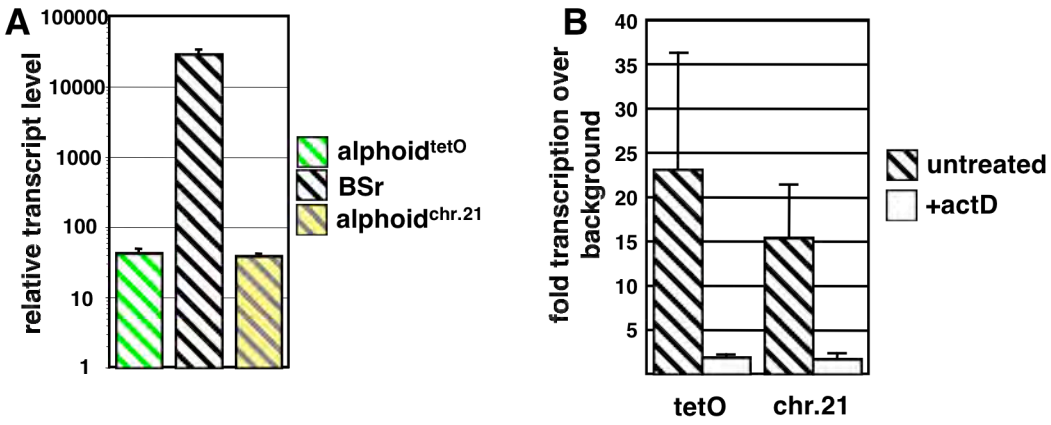


Figure 3| Low levels of polymerase activity generate centromeric non-coding transcripts. A) Real-time RT-PCR analysis of total RNA extracted from AB2.2.18.21 cells identifies low-levels of transcripts derived from the synthetic HAC and the endogenous chromosome 21 centromeres. Transcript levels are normalized to those of β-actin, expressed relative to the copy number of the corresponding genomic loci and assigned an arbitrary value. Data shown represents the mean and standard deviation of three independent experiments. **B)** Analysis of alphoid^{tetO} (tetO) and chromosome 21 (chr.21) centromeric transcripts in AB2.2.18.21 cells as in (A) grown in the presence or absence of actinomycin D (actD) for 16 hours. Specific transcript levels are divided by the background (no RT) to yield fold expression over background.

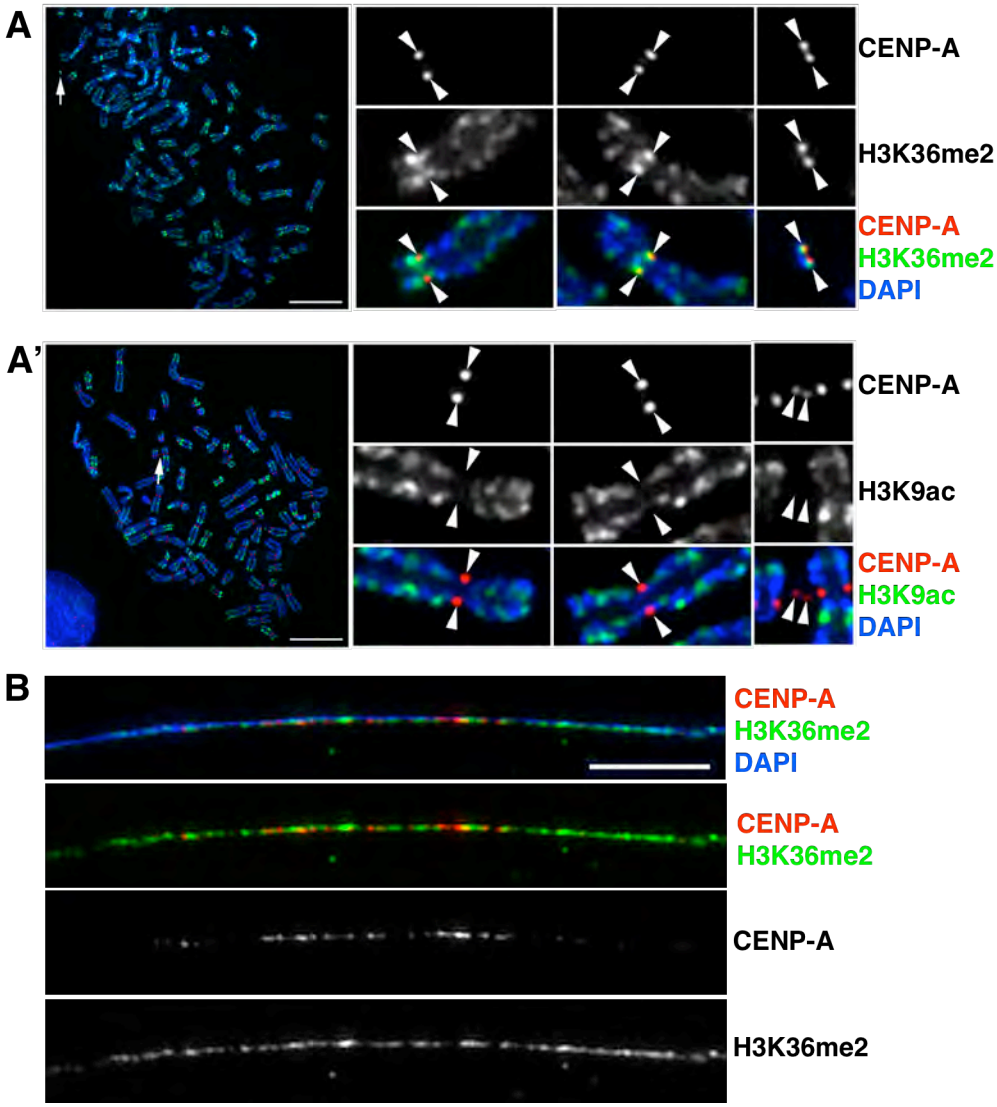


Figure 4| H3K36me2 forms part of centromere chromatin. A) Immunofluorescence co-staining of unfixed HT1080 (AB2.2.18.21) mitotic chromosomes using antibodies against CENP-A (red) and either H3K36me2 (A) or H3K9ac (A') (green). Magnified panels display individual endogenous chromosomes (left and centre panels) and the alphoid^{tetO} HAC (right panel), as determined by its size (white arrow in the spread). **B)** Stretched chromatin fibre stained for CENP-A (red) and H3K36me2 (green). Interspersed pattern of H3K36me2 and CENP-A was reproducible across several independent chromosome preparations and a representative image is depicted. Scale bars: 5 µm.

The Chromatin Environment of Active Centromeres is Maintained after Transfer into a Different Nuclear Background

1C7 cells are the product of polyethylene glycol-mediated fusion of AB2.2.18.21 cells with HeLa cells, and stably carry the alphoid^{tetO} HAC as a single copy (Stefano Cardinale). The HAC in 1C7 cells recruits all kinetochore components tested for and consistently shows proper metaphase alignment as well as correct segregation in anaphase (data not shown), demonstrating that the kinetochore assembled at the synthetic centromere remains functional. To assess if structural and functional integrity of the HAC kinetochore is paralleled by conservation of the underlying centromeric chromatin, parallel ChIP analysis was performed in 1C7 and AB2.2.18.21 cells to assess the levels of CENP-A and various histone modifications associated with the alphoid^{tetO} array. Strikingly, the alphoid^{tetO} HAC centromere displayed virtually identical CENP-A occupancy after transfer (Fig. 5). Furthermore, the HAC centromere retained high levels of H3K4me2, while detection of H3K4me3 remained close to background. The chromatin formed on the alphoid^{tetO} array also retained methylation of H3K9, although absolute levels of associated H3K9me3 were somewhat lower in 1C7 compared to AB2.2.18.21 cells. These data demonstrate that functional centromeres retain their basic chromatin landscape even in cells with remarkably different growth profiles and nuclear organisation.

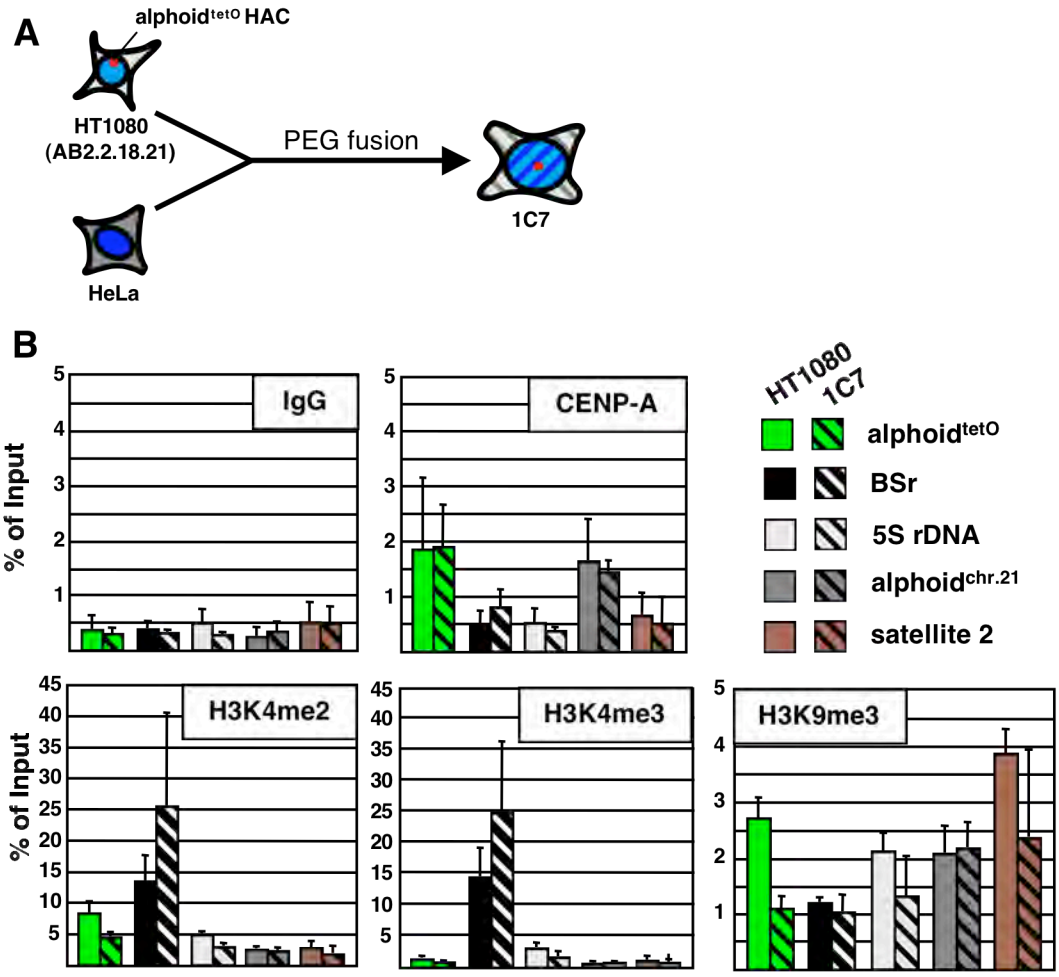


Figure 5| The HAC centromere chromatin is conserved after transfer into a different nuclear context. A) Polyethylene glycol (PEG)-mediated fusion of AB2.2.18.21 and HeLa cells yields the clonal cell line 1C7, stably carrying the alphoid^{tetO} HAC as a single copy. *[Cell fusion, isolation and initial characterization of 1C7 cells was performed by Stefano Cardinale].* **B)** ChIP analysis of AB2.2.18.21 and 1C7 cells assessing association of CENP-A and the indicated histone modifications at the genomic loci noted in figure 2 as well as at the heterochromatic satellite 2 repeats. Note that in contrast to the ChIP in figure 2, this assay was performed using polyclonal anti-histone antibodies, and chromatin was ChIP'ed using protein G agarose beads rather than anti-mouse dynabeads. Data is presented as in figure 2 and derived from two or more independent ChIP experiments.

J.3 Discussion

The present study provides the most comprehensive analysis of centromeric chromatin with respect to post-translational histone modifications analyzed to date. The results presented in this section demonstrate a remarkable degree of similarity in the histone modification profile defining active centromeres beyond the specific presence of CENP-A nucleosomes. In fact, the majority of modifications analyzed display a comparable tendency for relative enrichment at or depletion from both the synthetic alphoid^{tetO} centromere and that of endogenous chromosome 21. Importantly, transfer of the alphoid^{tetO} HAC into a different nuclear context essentially demonstrates conservation of the associated centromeric chromatin environment, both on a structural and functional level, and highlights centromere chromatin as an epigenetic entity in the tightest definition. In light of the highly dynamic regulation of histone methylation states (Shi and Whetstine, 2007), these findings therefore further point towards the requirement for a common mechanism that shapes and, importantly, maintains this chromatin environment.

Of particular note is the lack of “strong” euchromatic character at active centromeres, reflected by low or undetectable levels of H3K4me3, acetylation of H3K9 or H3K27. Whether this represents merely a passive by-product of low-level non-coding transcription through the centromere (see below) or is of direct functional consequence for the maintenance of centromeric structure and function remains to be determined. Interestingly in this respect is the absence of marked levels of CENP-A associated with the actively transcribed BSr marker on the HAC vector backbone that shows high levels of H3 acetylation and hypermethylation of H3K4. Although the relative absence of CENP-A from the BSr locus may result from a preference of CENP-A deposition at alphoid DNA, this observation may indeed support the notion that euchromatic character needs to be limited for the maintenance of

CENP-A. I will be directly addressing this question in the subsequent chapter M of the present work.

The above ChIP and subsequent immunofluorescence studies surprisingly identified hypermethylation of H3K36 as genuine component of centromere chromatin. In budding yeast, methylation of H3K36 is directly associated with the process of transcriptional elongation (Keogh et al., 2005; Vakoc et al., 2006). This modification is also found throughout the body of actively transcribed mammalian genes (Barski et al., 2007; Vakoc et al., 2006). These chromatin domains are further characterized by the presence of H3K4me1/me2 and H3K27me1. Co-transcriptional methylation of both H3K4 and H3K36 is directly implicated in the recruitment of HDAC activities, thereby resulting in local hypoacetylation and the formation of a characteristic chromatin signature found at transcribed genes and non-coding RNAs (Barski et al., 2007; Guttman et al., 2009; Mikkelsen et al., 2007; Schneider et al., 2004; Vakoc et al., 2006). In fact, corresponding genomic chromatin state maps of methylation of H3K36, have recently facilitated the prediction and discovery of a multitude of long non-coding transcripts (Guttman et al., 2009). Remarkably, the above ChIP data demonstrate that, in large part, centromeres essentially resemble this specific chromatin domain. Consistently, both synthetic alphoid^{tetO} and endogenous chromosome 21 centromeres are actively transcribed. It is therefore feasible to suggest that the local transcriptional activity, might be directly implicated in the maintenance of the centromere chromatin domain, perhaps by indirect modification of the chromatin environment through mechanisms analogous to those described at transcribed genes. I will provide evidence supporting this hypothesis in section L. It is noteworthy to point out that H3K36 methylation does not indicate immediate transcriptional activity, but rather represents a consequence of RNA polymerase activity that may persist as a memory state throughout the remained of the cell cycle.

The hypothesis that centromeric transcription maintains the local chromatin domain would furthermore explain the intermediate nature of centromeric H3K4me2 levels. Lam and co-workers detect moderately high levels of this modification at the centromeres of chromosomes 7 and 17 in human HT1080 cells (Lam et al., 2006), and consistently, the alphoid^{tetO} HAC centromere displays strong enrichment for this mark. Levels of H3K4me2 at the chromosome 21 centromere however are low, although still detectable. In contrast to methylation of H3K36, which is essentially present throughout the entire body of active open reading frames (Vakoc et al., 2006), dimethylation of H3K4 is strongly biased towards their 5' regions. This inevitably means that relative enrichment or "density" of the latter mark associated with a given transcribed genomic locus as detected by ChIP will inversely correlate with the size of this region. More generally, perceived abundance of certain histone modifications at centromeric loci detected by ChIP would also be subject to other forms of non-uniform distribution of these marks. For example, "centrochromatin" occupies only a portion of the total type I alpha satellite region of a given centromere (Lam et al., 2006). Real-time PCR analysis suggests that the relative genomic copy number of the alphoid$^{chr.21}$ is about 10 times higher than that of the synthetic alphoid^{tetO} (data not shown), and could therefore explain the corresponding differences in the perceived H3K4me2 density, as well as differing over-all levels of certain other histone modifications if the transcript length is correspondingly increased.

The level of transcription through both, HAC and chromosome 21 centromeres, is extremely low. Furthermore, the rate of centromere transcription based on the near-identical levels of the transcripts corresponding to the HAC and endogenous arrays appears to be similar, suggesting that transcriptional activity at the centromere is tightly controlled. An important subject that remains to be addressed in the future is the origin of transcription through endogenous centromeres. Study of these significantly hampered by the extensive size and nature of the underlying higher-order

repeat sequences. A possible source of elongating RNA polymerase II is read-through transcription from upstream transcribed elements, for example LINE-1 elements present within the pericentromeric heterochromatin domain. In this light, it is notable that hypermethylation of H3K9 does not necessarily impede with transcription elongation, as H3K9me3 has recently been described within the open reading frame of *PABPC1*, which is transcribed at high levels (Vakoc et al., 2006). Alternatively, transcription might originate from an as-yet unidentified cryptic promoter element present within the alpha-satellite repeat arrays. 5' rapid amplification of cDNA ends combined with new DNA sequencing technologies placing an emphasis on either increased sensitivity towards low-abundance targets or generation of extensive, individual sequence reads will eventually be able to provide an answer to this question.

K HIERARCHICAL DISASSEMBLY OF KINETOCHORE STRUCTURE BY CHROMATIN MODIFIERS

K.1 Background

In earlier studies employing the alphoid^{tetO} HAC in HT1080 cells (Nakano et al., 2008), tethering of a fusion of tetR to the Krueppel-associated box (KRAB) domain of the transcription factor Kid-1 (tTS; (Freundlieb et al., 1999)) resulted in dramatic inactivation of HAC kinetochore function and the near-complete loss of the HAC from the cell population within 14 days. On the level of the underlying chromatin, the tTS induced substantial changes at the HAC, including an increase in H3K9me3 levels associated with the alphoid^{tetO} array, a decrease in local H3K4me2 and the loss of CENP-A nucleosomes. Tethering of the tTS further resulted in substantial recruitment of heterochromatin protein (HP)1α, demonstrating the establishment of a strong heterochromatic character.

The conserved KRAB domain is found in a large subset of zinc-finger transcription factors, and has been shown to recruit the transcriptional co-repressor KRAB-Associated Protein 1 (KAP1; also known as KRIP-1 and TIF1β) (Friedman et al., 1996; Kim et al., 1996; Moosmann et al., 1996). KAP1 in turn acts as a molecular scaffold that directly interacts with a variety of chromatin modifying activities (Le Douarin et al., 1996; Lechner et al., 2000; Ryan et al., 1999; Schultz et al., 2002; Schultz et al., 2001). A short amino acid sequence within the central part of KAP1 mediates binding to HP1 (Le Douarin et al., 1996; Lechner et al., 2000; Ryan et al., 1999). At its C-terminus, KAP1 contains a tandem PHD finger and a bromo-like domain that cooperatively facilitate interaction with the NuRD histone deacetylase complex subunit Mi-2α (Schultz et al., 2001). The PHD finger appears sufficient for the recruitment of the H3K9-specific methyltransferase SETDB1 (Schultz et al., 2002). An N-terminal RING-finger / B-box / coiled-coil (RBCC)

motif is required for interaction with the KRAB domain, but appears dispensable for transcriptional repression of a reporter construct (Friedman et al., 1996; Kim et al., 1996; Moosmann et al., 1996). Direct tethering of KAP1 to a stably integrated reporter construct triggered changes in local chromatin architecture reflecting those observed at the HAC after tTS tethering, including loss of H3K4 methylation, an increase in H3K9 methylation and recruitment of HP1 (Sripathy et al., 2006). These findings are therefore consistent with the model of KAP1 mediating changes in chromatin architecture downstream of the tTS.

My initial aim was to test the hypothesis that direct tethering of KAP1 indeed interferes with HAC kinetochore function, as well as to gain a better understanding of the associated process resulting in disruption of the kinetochore structure. The modular domain architecture of KAP1 is in principle ideally suited to generate a series of truncations fused to tetR in order to assess the contribution of individual functional domains towards the deleterious over-all effect of tTS or KAP1. The data presented in this section identifies KAP1 as a potent antagonist of the structural maintenance of functional kinetochores. I furthermore show that KAP1 mediates kinetochore disassembly, at least in part, through a CENP-A independent pathway.

K.2 Results

[Some data presented in this section was obtained and/or analyzed by Stefano Cardinale. S.C.'s contribution is acknowledged in the relevant paragraphs and figure legends.]

KAP1 Tethering Rapidly Disrupts HAC Kinetochore Structure and Function

To determine if KAP1, the likely downstream effector of tTS, is sufficient for the disruption of the alphoid^{tetO} HAC kinetochore, I generated a plasmid expressing EYFP-labelled tetR fused to amino acids 38 through 835 of human KAP1 (tetR-EYFP-KAP1), including the KAP1 RBCC motif, HP1 binding domain and C-terminal PHD / bromo-like domains (Fig. 6A). For simplicity, I will refer to this fusion construct as “full-length” KAP1. I went on to transiently transfect 1C7 cells with plasmids expressing either tetR-EYFP or tetR-EYFP-KAP1, and assessed staining for the inner kinetochore component CENP-C at the HAC in cells with comparable expression levels of the fusion constructs. As expected, all interphase HACs targeted by tetR-EYFP displayed solid staining for associated CENP-C (Fig. 6B). Strikingly however, tethering of KAP1 for as little as two days resulted in loss of detectable CENP-C staining at interphase HAC kinetochores in a fraction of cells analyzed (Fig. 6B’). A comprehensive analysis *[performed by S.C.]* of KAP1-tethered HACs four days after transfection revealed that 75 % of HACs had lost detectable levels of CENP-C or CENP-H (Fig. 6C). 14 % of HACs also failed to display CENP-A staining, demonstrating a complete disruption of inner kinetochore structure. Interestingly, these data imply that loss of CENP-C precedes loss of CENP-A. Indeed, basic line profile fluorescence quantification showed that CENP-C staining was frequently reduced or completely lost at HACs that retained CENP-A levels comparable to those of tetR-EYFP-targeted control HACs (*[S.C.]*, data not shown).

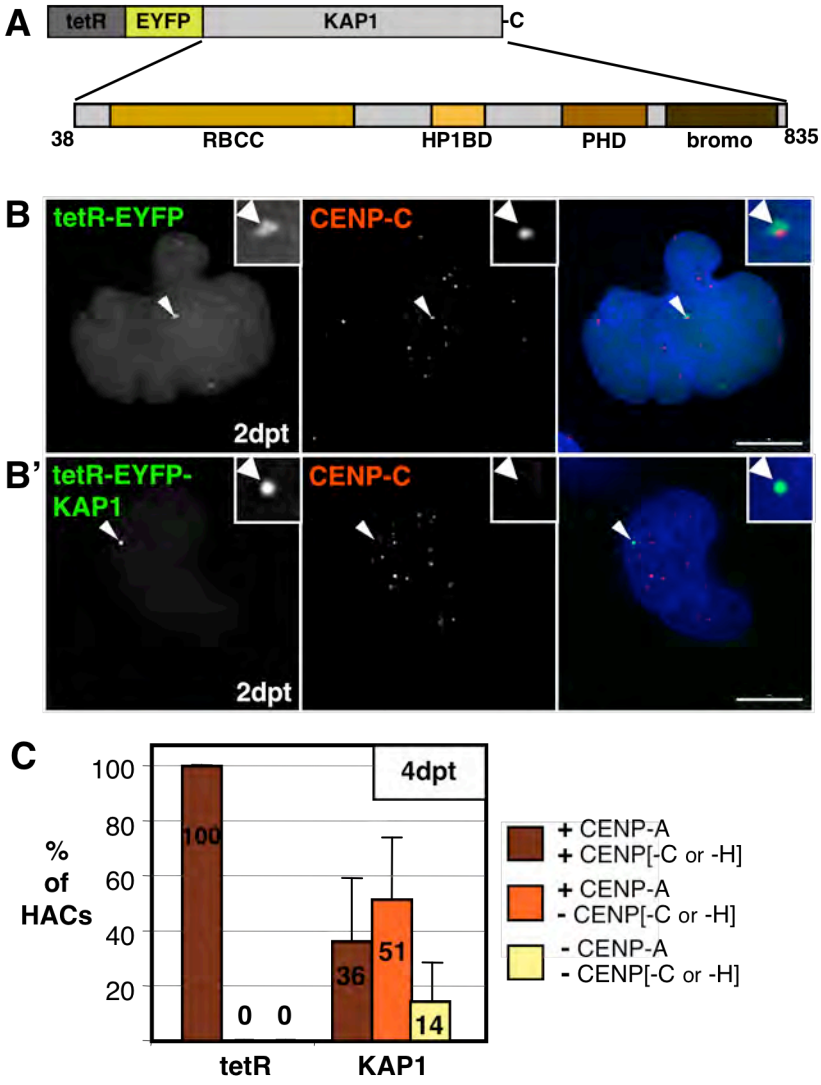


Figure 6| KAP1 tethering induces hierarchical loss of HAC inner kinetochore structure. A) Schematic drawing of tetR-EYFP-KAP1. **B)** IF analysis of 1C7 cells two days after transfection with constructs expressing tetR-EYFP (B) or tetR-EYFP-KAP1 (B') and stained for CENP-C (red). Scale bar: 10 µm. **C)** Four days after transfection with constructs expressing tetR-EYFP (tetR) or tetR-EYFP-KAP1 (KAP1), 1C7 cells were co-stained for CENP-A and either CENP-C or CENP-H. The presence (+) or absence (-) of either kinetochore component at the interphase HAC was quantified *[this analysis was performed by S.C.]*.

To investigate HAC kinetochore disruption mediated by KAP1 tethering in greater detail, I generated a 1C7-derived cell line stably expressing tetR-EYFP-KAP1 (1C7-KAP1). Cells were grown in the presence of doxycyline to prevent binding of the fusion construct to the HAC. Targeting of the construct into the HAC kinetochore was initiated by washing out the drug using the protocol defined in figure 7A. This protocol reproducibly resulted in maximal binding of the fusion construct to the HAC within 24 hours, as determined by direct quantification of the HAC-associated EYFP signal (data not shown). Importantly, expression levels of the construct remain identical before and after doxycycline wash-out, so that any effects observed at the HAC kinetochore after drug wash-out are the consequence of a direct interaction of the construct with the HAC.

Using this experimental system, I performed an immunofluorescence time course experiment to monitor the levels of CENP-A and CENP-C staining associated with the HAC kinetochore after inducing its targeting by tetR-EYFP-KAP1. In order to report this data in an unbiased, quantitative manner, I measured the average HAC-associated CENP fluorescence in each cell. Staining of endogenous centromeres did not vary significantly across the different time points analyzed (data not shown). In contrast, levels of CENP-A at the HAC kinetochore decreased gradually over time (Fig. 7B, D and D'). This resulted in a loss of detectable CENP-A from 24 % of HACs analyzed by four days after doxycycline wash-out. Strikingly, levels of CENP-C decreased much faster than those of CENP-A (Fig. 7C, E and E'). Within 48 hours, more than 30 % of HACs lacked measurable CENP-C association. By 72 hours, almost 80 % of HACs were devoid of CENP-C staining. None of the HACs analyzed at four days retained CENP-C, a time point at which the majority of HACs still showed detectable CENP-A levels. These data confirm that as a consequence of KAP1 tethering, CENP-C is lost more readily from the HAC kinetochore than CENP-A.

Analysis of mitotic 1C7-KAP1 cells revealed HAC-specific defects closely correlating with the CENP-C levels measured above. 48 hours after KAP1 tethering was induced, a time point at which CENP-C levels were low, but reproducibly detectable at most HACs, almost all (94 %) targeted HACs displayed normal alignment on the metaphase plate (Fig. 8A and C). Individual HAC sister chromatids were clearly resolved and showed tension (Fig. 8A). In contrast, by 72 hours of targeting by KAP1, when CENP-C signals at the HAC were largely lost, the HAC failed to achieve metaphase alignment in almost 50 % of cells analyzed (Fig. 8B, B' and C). Interestingly, HAC sister kinetochores were frequently unresolved (Fig. 8B'), suggesting abnormal compaction of the underlying centromeric chromatin. Notably, interphase CENP-A levels at this time point remained readily detectable at the majority of HACs (see above), further suggesting that KAP1-induced disruption of kinetochore function occurs primarily through a CENP-A independent mechanism.

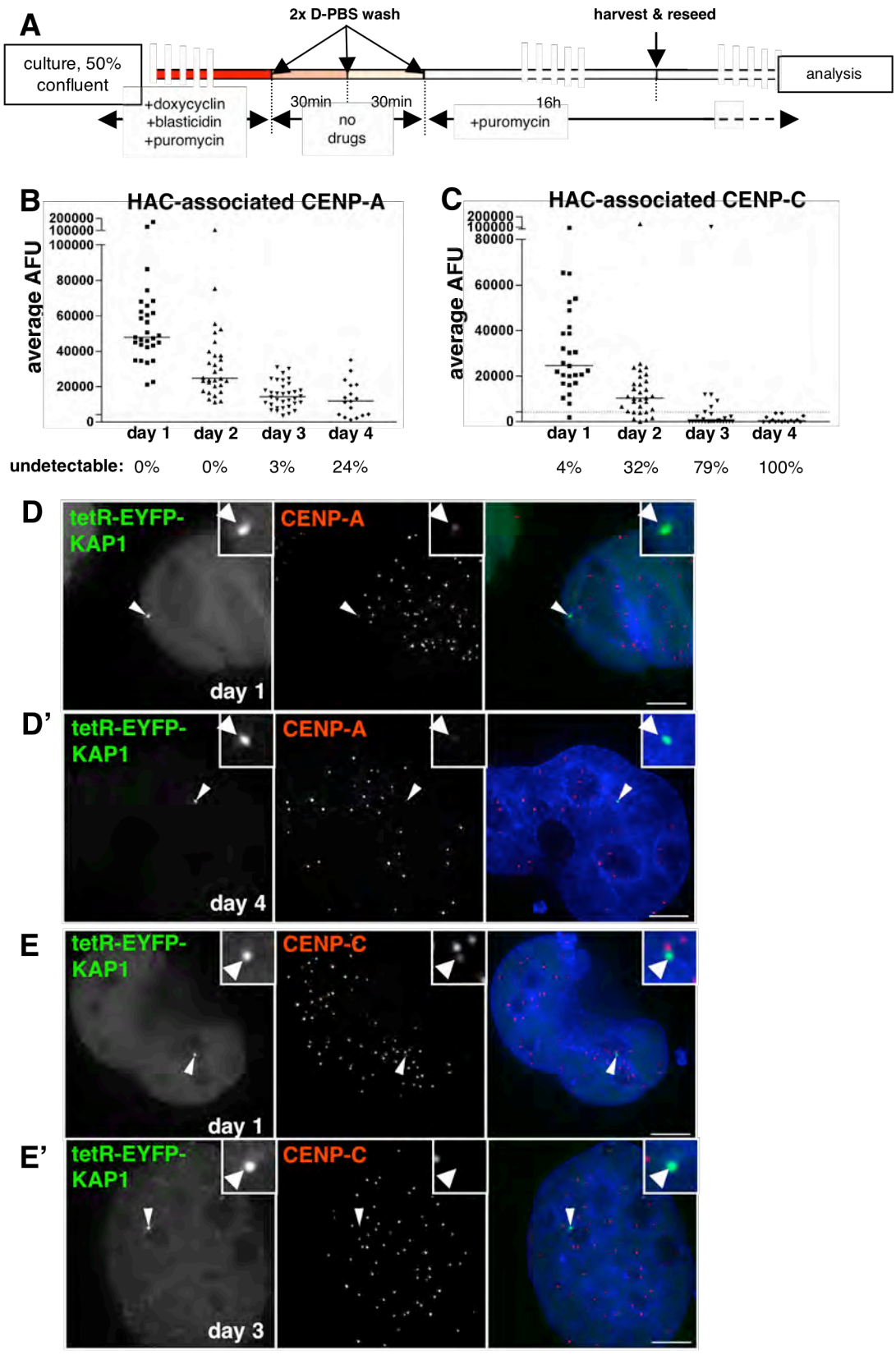


Figure 7| CENP-C is lost more readily from the HAC kinetochore during KAP1 tethering than CENP-A. *Figure legend continued overleaf.*

Figure 7 continued.

A) Work flow of doxycycline wash-out experiments to induce targeting of tetR-EYFP-KAP1 into the HAC kinetochore in stable 1C7-KAP1 cells. **B)** Quantification of CENP-A fluorescence signals associated with the HAC at the indicated time points after doxycycline wash-out. Cells were stained with an antibody against CENP-A, and the background-subtracted, average arbitrary fluorescence units (AFU) are plotted. Solid lines indicate the median value. The dotted line indicates the limit of detection of specific fluorescence signal over nuclear background, determined as the 95% confidence interval of the spread of background-subtracted background values. For each time point, percentages of cells lacking detectable HAC-associated CENP-A staining are indicated. **C)** Quantification of HAC-associated CENP-C staining as in (B). **D)** IF analysis of 1C7-KAP1 cells stained for CENP-A one (D) and four (D') days after doxycyline wash-out. Cells represent the median values of the corresponding time points in the quantification in (B). Scale bar: 5 µm. **E)** IF analysis as in (D), using an antibody against CENP-C at one (E) and three (E') days after wash-out.

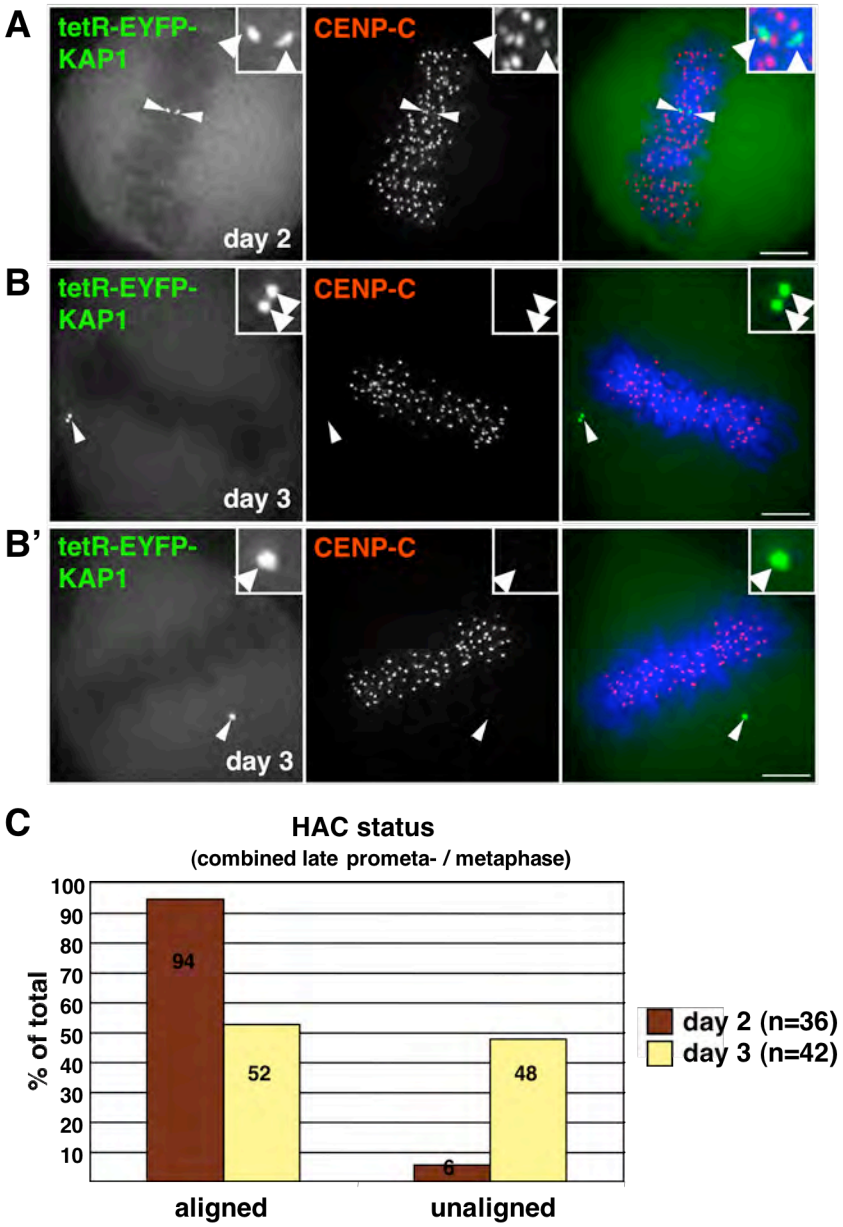


Figure 8| KAP1 tethering interferes with HAC structure and kinetochore function. A) IF analysis of a metaphase 1C7-KAP1 cell stained for CENP-C two days after doxycyline wash-out. Arrowheads depict HAC sister kinetochores under tension. Scale bar: 5 µm. *Figure legend continued overleaf.*

Figure 8 continued.

B) IF analysis as in (A) at three days after wash-out. Arrowheads depict resolved HAC sister chromatids of the unaligned HAC. Frequently at this time point, HAC sister chromatids appeared hyper-compacted (arrowhead in B'). C) Quantification of HAC fate in individual 1C7-KAP1 cells by IF analysis. HACs were scored as aligned (as in A) or unaligned (as in B) in late prometaphase and metaphase, at the indicated time points after doxycycline wash-out.

KAP1 Tethering Causes Unequal HAC Sister Kinetochore Structure

During analysis of KAP1-targeted mitotic HACs, I often noticed a clear bias in the CENP-C staining associated with the two HAC sister kinetochores. More frequently than not, one EYFP-labelled sister kinetochore showed markedly lower CENP-C levels than the other. This trend was readily observed at HACs in metaphase (Fig. 9A) and anaphase (Fig. 9A'). This was highly unusual, as mitotic HACs targeted with tetR-EYFP alone or other fusion constructs generally did not show this phenotype (for example, see Fig. 17A-A'", section L). An imbalance was also observed for outer kinetochore components KNL-1 and hMis12 (data not shown). Quantification of CENP-C fluorescence signals at both sister kinetochores strikingly demonstrated that most HAC kinetochores had differences in CENP-C levels exceeding a factor of 1.5, with some showing differences of a factor greater than two (Fig. 9B). Importantly, associated EYFP signals were typically distributed evenly, demonstrating that unequal kinetochore structures were not the result of different amounts of KAP1 tethered into the sister kinetochores.

A Concerted Action of KAP1 Domains is Required for Maximal HAC Kinetochore Disruption

KAP1 tethering results in the rapid destabilization of the HAC inner kinetochore structure. As outlined above, KAP1 combines a series of well-defined functional domains. In order to further characterize the basis of KAP1-mediated kinetochore disruption, I was therefore interested in determining if a given domain of KAP1 exerts a dominant antagonistic effect on HAC kinetochore maintenance. To this end, I generated a series of constructs expressing tetR-EYFP fused to different segments of KAP1 (Fig. 10A). TetR-EYFP-RBCC comprises the KAP1 N-terminal RBCC motif. Further constructs included fusions with truncations of KAP1 containing its HP1 binding domain (tetR-EYFP-HP1BD) or the C-terminal PHD and bromo-

like domains (tetR-EYFP-PHD/Bromo). An additional construct, tetR-EYFP-KAP1[38-559] including the RBCC motif and HP1 binding domain of KAP1 was generated by Stefano Cardinale.

To assess effects of different KAP1 domain fusions on the structural integrity of the HAC inner kinetochore, 1C7 cells were transfected with these constructs, tetR-EYFP or tetR-EYFP-KAP1, followed by immunofluoresence analysis of CENP-A, CENP-C or CENP-H four days after transfection (Fig. 10B) *[this work and analysis was carried out by S.C.].* In contrast to tetR-EYFP, all constructs were able to disrupt HAC kinetochore structure to some extent. Individual domains displayed a hierarchical disruption in a manner similar to full-length KAP1, with CENP-C or CENP-H being lost more readily than CENP-A. However, the severity of structural defects caused was highly variable depending on the KAP1 domain tethered. Surprisingly, the KAP1 HP1 binding domain had only a very weak effect at the HAC, with 94 % of HAC kinetochores appearing structurally normal. At the remaining HACs, CENP-C or CENP-H, but not CENP-A, were lost. The tetR-EYFP-PHD/Bromo fusion was somewhat more potent, inducing loss of CENP-C or CENP-H from almost 20 % of HACs analyzed. However, as with the HP1 binding domain, this construct was unable to induce loss of CENP-A. Unexpectedly, the RBCC motif of KAP1 was sufficient to disrupt HAC kinetochore structure in 45 % of cells analyzed, including the loss of detectable CENP-A in 12 % of cases. A likely explanation may be recruitment of endogenous KAP1 through oligomerization mediated by the RBCC domain (Peng et al., 2002). Interestingly, kinetochore disruption using this construct was less efficient compared to full-length KAP1, suggesting an additive adverse effect of KAP1 oligomerization. Extending the RBCC motif to include the downstream HP1 binding domain (tetR-EYFP-KAP1[38-559]) fully restored the disruptive effect of full-length KAP1 tethering.

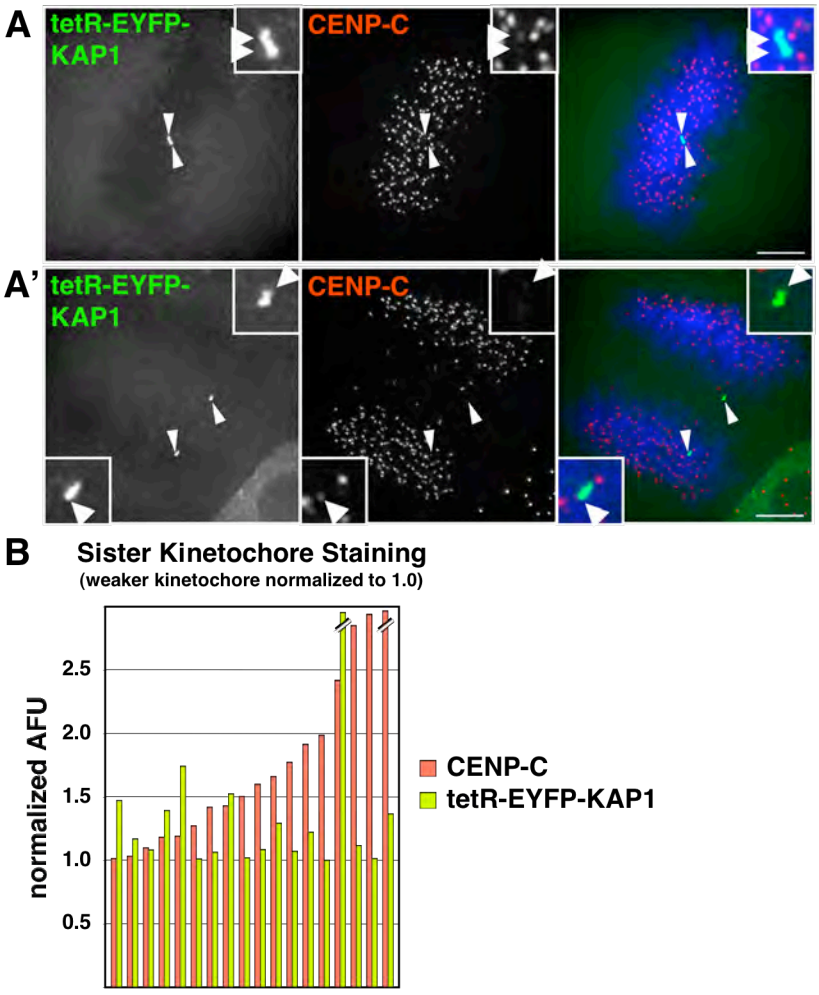


Figure 9| KAP1 induces unequal kinetochore architecture at tethered HACs. A) IF analysis of mitotic 1C7-KAP1 cells two days after doxycycline wash-out, stained for CENP-C. Unequal levels of CENP-C are evident at the individual HAC sister kinetochores (arrowheads) in metaphase (A) and anaphase (A') cells. Note that the lagging HAC sister chromatid in (A') also shows the weaker CENP-C staining. Scale bars: 5 µm. **B)** EFYP (tetR-EYFP-KAP1) fluorescence signals and staining of CENP-C associated with mitotic HAC kinetochores as in (A) were quantified two days after wash-out in individual cells, and plotted as the ratio of the brighter sister over the other.

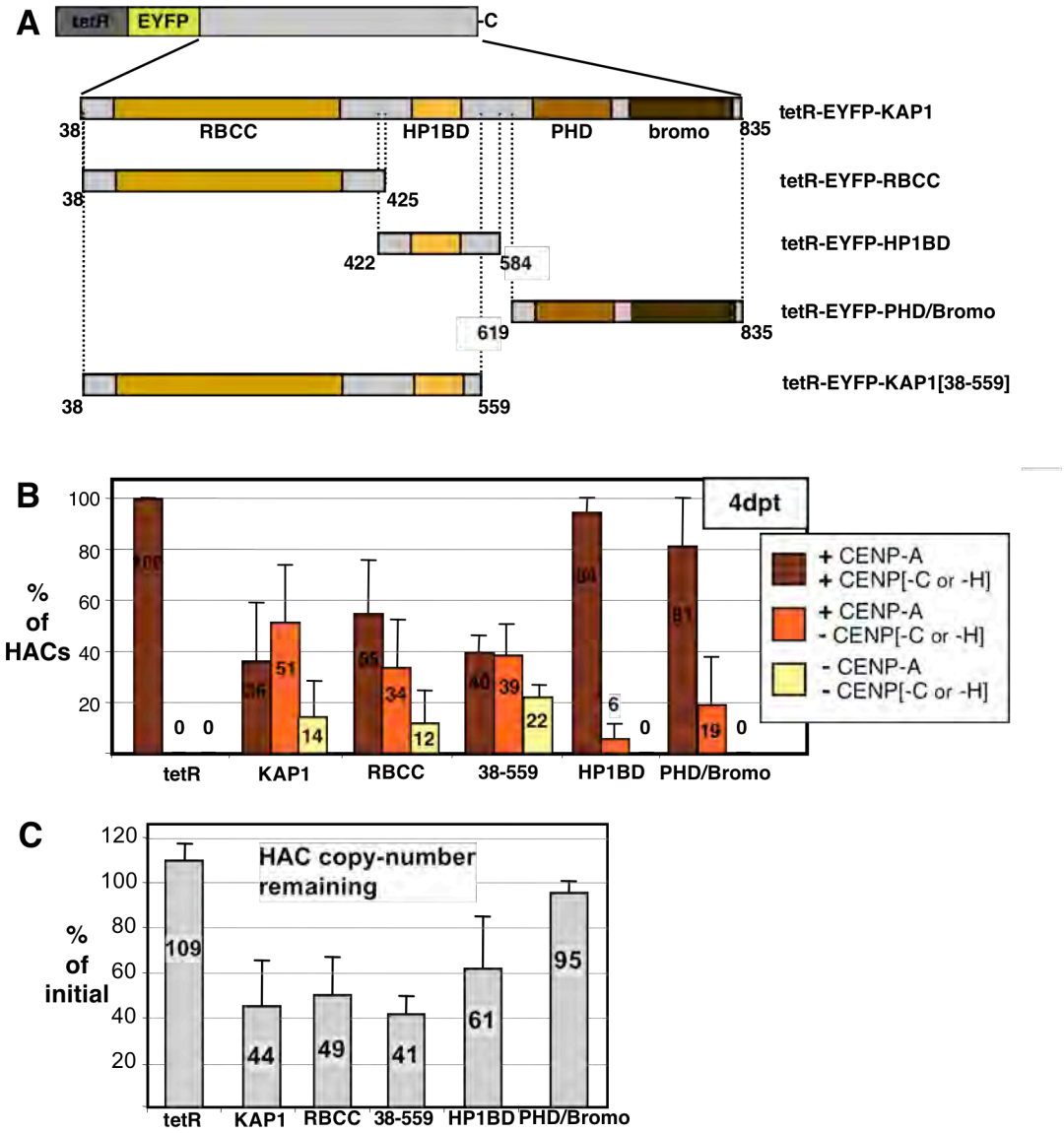


Figure 10| A combinatorial effect of KAP1 domains is required for maximal HAC kinetochore disruption. A) Schematic drawing of different KAP1 domain fusions to tetR-EYFP. **B)** Four days after transfecting expression constructs of tetR-EYFP or either of the above KAP1 domain fusions, 1C7 cells were co-stained for CENP-A and either CENP-C or CENP-H. The presence (+) or absence (-) of either kinetochore component at the interphase HAC was quantified. *[This analysis was performed by S.C.]* **C)** HAC copy numbers as determined by real-time PCR 12 days after expressing tetR-EYFP or either of the above fusion constructs in 1C7 cells. HAC copies are plotted as percentage relative to untransfected 1C7 cells.

To determine the longer-term impact of tethering individual KAP1 domains into the HAC kinetochore, 1C7 cells were transfected with the different constructs and maintained in culture under continuous selection for the puromycin marker present in the vector backbone. After twelve days, I extracted genomic DNA from the puromycin-resistant cells and performed real-time PCR analysis to determine the HAC copy number in each population relative to untransfected control cells. Consistent with previous data (Nakano et al., 2008) and the lack of detectable adverse effects on kinetochore structure when tethering tetR-EYFP, the cell population expressing this construct did not show loss of the HAC (Fig. 10C). In contrast, tethering of either full-length KAP1, the KAP1 RBCC domain or KAP1[38-559] resulted in prominent reduction of HAC copy numbers (loss of 56 %, 51 % and 59 %, respectively) from the population, closely reflecting the subtle differences in their ability to disrupt HAC kinetochore structure as illustrated above. The combined PHD and bromo-like domains showed only a mild decrease (5 %) in HAC copies. Surprisingly, long-term tethering of the HP1 binding domain of KAP1, which displayed only a very mild effect on HAC kinetochore structure after four days, resulted in a considerable loss of HAC copies (39 %), suggesting that the associated process of HAC kinetochore disruption extends beyond a four day time course. Together, these data suggest that KAP1-mediated interference with kinetochore structure and function is a result of the combination of its individual subdomains. A concerted action of individual KAP1 domains also appears to be essential for the over-all chromatin remodelling activity of KAP1 at a chromatin template (Sripathy et al., 2006).

K.3 Discussion

Tethering of the tTS into the HAC kinetochore is incompatible with maintenance of HAC kinetochore structure, and results in loss of CENP-A nucleosomes from the HAC centromeric chromatin (Nakano et al., 2008).

The data presented in this section establish that direct tethering of KAP1, the putative downstream effector of the tTS, is sufficient to rapidly destabilize HAC kinetochore architecture and function. Interestingly, in the course of KAP1 tethering, CENP-C is displaced from the HAC kinetochore more readily than CENP-A. Furthermore, HACs start to show kinetochore defects at a time point at which CENP-A is present at readily detectable levels. The initial mechanism of kinetochore disruption triggered by KAP1 as such appears to be independent of CENP-A-dependent assembly of the HAC inner kinetochore structure. Loss of HAC kinetochore function instead appears to correlate with loss of detectable CENP-C. An essential role for CENP-C in maintaining proper kinetochore structure and function is well established (Hori et al., 2008a; Kalitsis et al., 1998; Tomkiel et al., 1994).

On the molecular level, the tTS ectopically induces a character of strong, constitutive heterochromatin at the HAC centromere (Nakano et al., 2008). Direct tethering of KAP1 to chromatin results in local hypoacetylation, induces hypermethylation of H3K9, initiates recruitment of all three HP1 isoforms and increases nucleosome occupancy, indicative of local chromatin condensation (Sripathy et al., 2006). KAP1 furthermore acts as transcriptional repressor *in vivo*, and its direct tethering is sufficient to silence transcription of ectopic as well as chromatinized reporter constructs (Friedman et al., 1996; Kim et al., 1996; Lechner et al., 2000; Moosmann et al., 1996; Sripathy et al., 2006).

Transcription of alpha-satellite DNA may be necessary for the maintenance of centromere identity, as hypothesized by others (see for example (Allshire and Karpen, 2008)). Disruption of HAC kinetochore structure during tethering of KAP1 could therefore be a consequence of repression of the low-level transcription at the alphoid^{tetO} array (see section J). As illustrated below, targeting of the alphoid^{tetO} HAC by a lysine-specific demethylase also represses HAC centromeric transcription (section L).

Importantly, this process is independent of ectopic heterochromatin formation. In this context, destabilization of the HAC kinetochore occurs both over a longer period of time, and without the apparent hierarchical mechanism observed during KAP1 tethering. I therefore hypothesize that rapid, KAP1-mediated disruption of the HAC kinetochore occurs through a mechanism other than transcriptional repression. Instead, KAP1-mediated chromatin remodelling may impose changes in the chromatin structure that are incompatible with assembly or function of the complex kinetochore architecture, or interfere with any suggested higher-order arrangement of centromeric chromatin (Blower et al., 2002; Marshall et al., 2008; Ribeiro et al., 2010).

Recent data reveal that CENP-C associates with domains of centromere chromatin containing canonical histone H3 (Hori et al., 2008a). In contrast to CENP-A, which forms a stable component of kinetochores throughout the cell cycle (Hemmerich et al., 2008; Jansen et al., 2007), CENP-C displays a dynamic exchange at the interphase pre-kinetochore (Hemmerich et al., 2008). Changes in post-translational modifications of centromeric H3-containing nucleosomes may disrupt association of CENP-C with the local chromatin, unfavourably shifting the equilibrium of CENP-C binding and resulting in net loss of this component from the HAC pre-kinetochore.

An intriguing observation made during the course of these studies is the asymmetric kinetochore structure present at mitotic HAC sister chromatids during tethering of KAP1. This may indicate an antagonising effect of KAP1-mediated chromatin changes on pre-kinetochore replication during or after S-phase, and mechanistically provide the basis of KAP1-mediated disruption of the HAC kinetochore. Duplication of the interphase kinetochore architecture established in G1 in the course of chromosome duplication is poorly investigated. Notably, in immuno-FISH studies, alpha-

satellite duplexes indicative of replicated sister centromeres often showed ACA staining associated with only one of the two spot signals (Haaf and Ward, 1994). Analogously, others and I sometimes observed interphase tetR-EYFP doublets with unilateral staining for a given pre-kinetochore marker. Together, these findings might indicate that at least some components of the pre-kinetochore are seeded subsequent to sister centromere replication, rather than concomitantly. Alterations of the local chromatin state introduced by KAP1 could thereby prevent efficient reassembly of pre-kinetochore structure following exit from S-phase.

L TRANSCRIPTIONAL ACTIVITY CORRELATES WITH THE MAINTENANCE OF THE CENP-A CHROMATIN DOMAIN

L.1 Background

In 2004, a single study reporting the presence of H3K4me2 nucleosomes within the CENP-A chromatin domain of both, human and *Drosophila* chromosomes (Sullivan and Karpen, 2004) sparked intense speculations about a possible functional role of this modification within centromere chromatin that persist to the present day in the absence of any supporting experimental data (Allshire and Karpen, 2008; Dunleavy et al., 2005; Schueler and Sullivan, 2006). As H3K4me2 was found to cluster in between sister kinetochores, underlying the poleward-facing CENP-A nucleosomes on mitotic chromosomes (Sullivan and Karpen, 2004), this modification might directly or indirectly mediate folding of the kinetochore domain into a suggested solenoidal three-dimensional super-structure (Blower et al., 2002; Sullivan and Karpen, 2004). Alternatively, H3K4me2 was suggested to act as a "placeholder" mark to facilitate the deposition of CENP-A nucleosomes subsequent to centromere DNA replication in S phase (Schueler and Sullivan, 2006). Another role may include barrier formation towards the flanking pericentromeric region, acting to prevent the spreading of excessive heterochromatin character into the CENP-A-occupied domains. A role in antagonizing H3K9 methylation was suggested for H3K4me3 (Huang et al., 2006; Nishioka et al., 2002), and a recent study in *Drosophila* indicates a requirement for removal of H3K4me2 prior to establishment of chromosomal heterochromatin domains (Rudolph et al., 2007). Direct tethering of HP1 (Nakano et al., 2008) or the heterochromatin-inducing KAP1 (see section K above) into the alphoid^{tetO} HAC kinetochore highlighted the negative correlation between strong heterochromatic character and kinetochore function and emphasize the requirement to spatially and functionally separate centrochromatin from pericentromeric heterochromatin.

In the experiments described in this section, I initially set out to assess if centromeric H3K4me2 is a functional requirement for the structural integrity of the HAC kinetochore *in vivo*. Studies within the past few years have established that most if not all histone methylation marks in chromatin can be actively removed through a large variety of specific histone demethylase enzymes (Shi and Whetstine, 2007). Amongst these, Lysine-Specific Demethylase (LSD) 1 was shown to selectively target H3K4me1 and H3K4me2 (Shi et al., 2004). By generating a fusion construct of LSD1 to tetR-EYFP, I have created a molecular tool to specifically deplete H3K4me2 from the alphoid^{tetO} HAC centromere. The data presented in this section implicates centromeric H3K4me2 in facilitating local RNA polymerase activity and uncovers a surprising relationship between transcription of an active centromere and the maintenance of the local CENP-A chromatin domain.

L.2 Results

Centromeric H3K4me2 is not Required for Immediate Kinetochore Function

To address the possible role of H3K4me2 within centrochromatin, I generated an expression construct encoding tetR labelled by EYFP and fused to full-length human LSD1 (tetR-EYFP-LSD1; Fig. 11A). When expressed in 1C7 cells and assessed by immunofluorescence microscopy, tetR-EYFP marking the alphoid^{tetO} centromere overlaps with CENP-A as well as distinct H3K4me2 staining at both interphase and mitotic HACs (Fig. 11B and 11C). In contrast, expression of tetR-EYFP-LSD1 resulted in a reduction of H3K4me2 staining at the HAC centromere to nuclear background in interphase cells (Fig. 11B') at two days after transfection, and H3K4me2 staining was virtually undetectable on tetR-EYFP-LSD1-targeted mitotic HACs (Fig. 11C'). A control construct containing an enzymatically inactive point mutation of LSD1, LSD1 (K661A) (Stavropoulos et al., 2006), failed to

display this dramatic decrease of H3K4me2 at the HAC centromere (Fig. 11B" and 11C). Together, these results confirm that tetR-EYFP-LSD1 is catalytically active, and more importantly, that this construct can be used to specifically deplete H3K4me2 from the single HAC centromere *in vivo*.

To allow a more defined and quantitative analysis of LSD1-mediated effects at the HAC centromere, I generated 1C7 cell lines stably expressing tetR-EYFP-LSD1 or the K661A mutant fusion construct (1C7-LSD1^{WT} and 1C7-LSD1^{K661A}, respectively). Cells were grown in the presence of doxycycline to prevent binding of the fusion construct to the HAC, and experimental targeting through washing out of the drug was induced as described for 1C7-KAP1 cells in section K (Fig. 7). Strikingly, highly sensitive ChIP analysis of 1C7-LSD1^{WT} cells showed that within 24 hours of washing out doxycycline, H3K4me2 was virtually completely depleted from the HAC centromere, and remained absent in the presence of the fusion construct by three days after drug wash-out (Fig 12B). In stark contrast, cells expressing the K661A mutant construct displayed only an insignificant reduction of alphoid^{tetO}-associated H3K4me2 within the initial 24 hours (Fig. 12B'), despite expressing the fusion construct at two-fold higher levels compared to 1C7-LSD1^{WT} cells (Fig. 12A). Notably, LSD1(K661A) tethering eventually resulted in reduced centromeric H3K4me2 after three days, consistent with a loss of co-transcriptional methylation of this residue as described below.

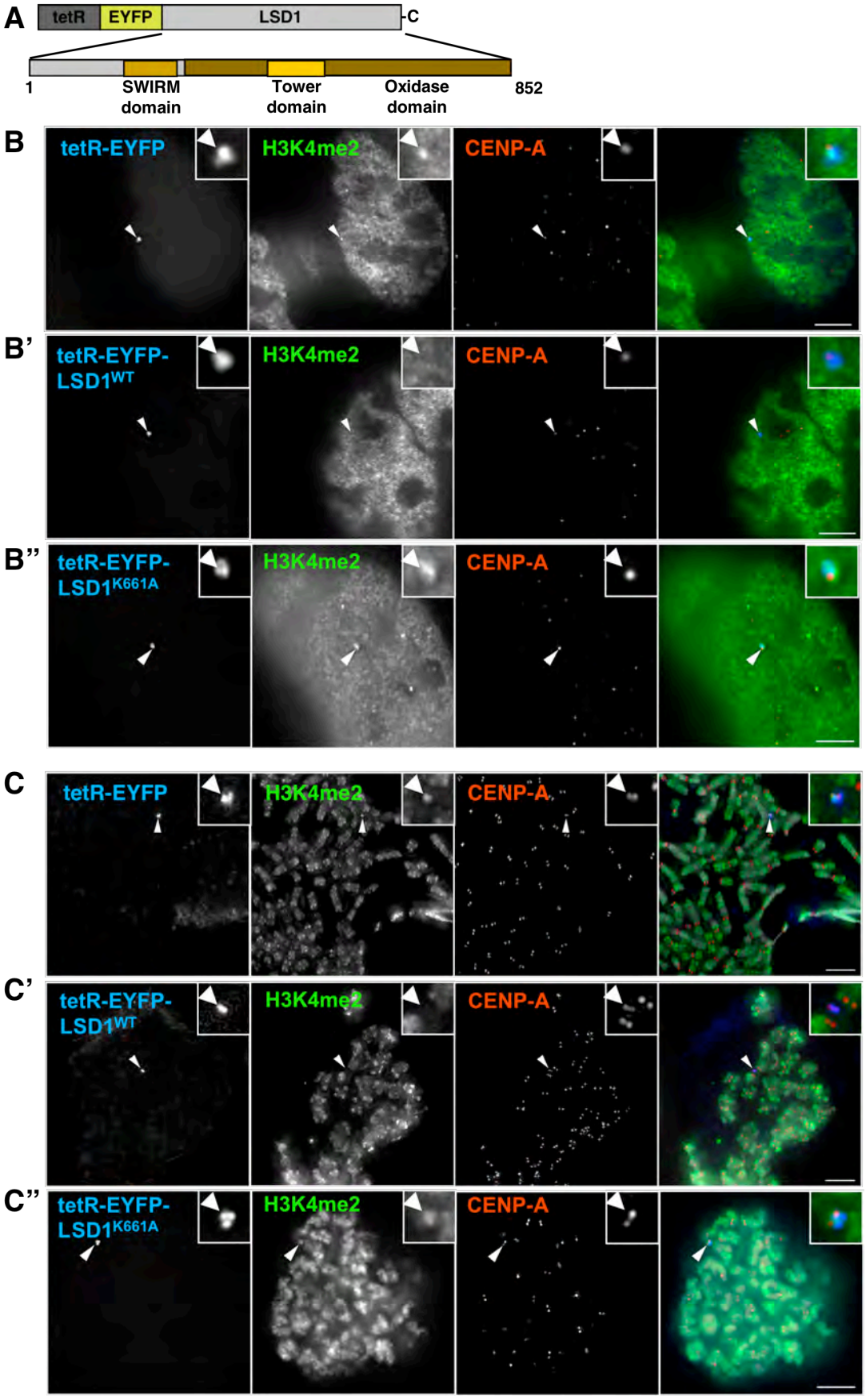

A
tetR
EYFP
LSD1
C
1
SWIRM domain
Tower domain
Oxidase domain
852
B
tetR-EYFP
H3K4me2
CENP-A
B'
tetR-EYFP-LSD1WT
H3K4me2
CENP-A
B''
tetR-EYFP-LSD1K661A
H3K4me2
CENP-A
C
tetR-EYFP
H3K4me2
CENP-A
C'
tetR-EYFP-LSD1WT
H3K4me2
CENP-A
C''
tetR-EYFP-LSD1K661A
H3K4me2
CENP-A

Figure 11| TetR-EYFP-LSD1 as tool to deplete H3K4me2 at the HAC centromere. A) Schematic drawing of tetR-EYFP-LSD1. **B)** Immunofluorescence analysis of interphase 1C7 cells two days after transfection with constructs expressing tetR-EYFP (B), tetR-EYFP-LSD1^{WT} (B') or the catalytically inactive tetR-EYFP-LSD1^{K661A} fusion construct (B")(blue pseudocolor). Cells were stained with antibodies against H3K4me2 (green) and CENP-A (red). Arrowheads depict the HAC. **C)** Immunofluorescence analysis of mitotic 1C7 chromosomes as in B. Scale bar: 5 µm.

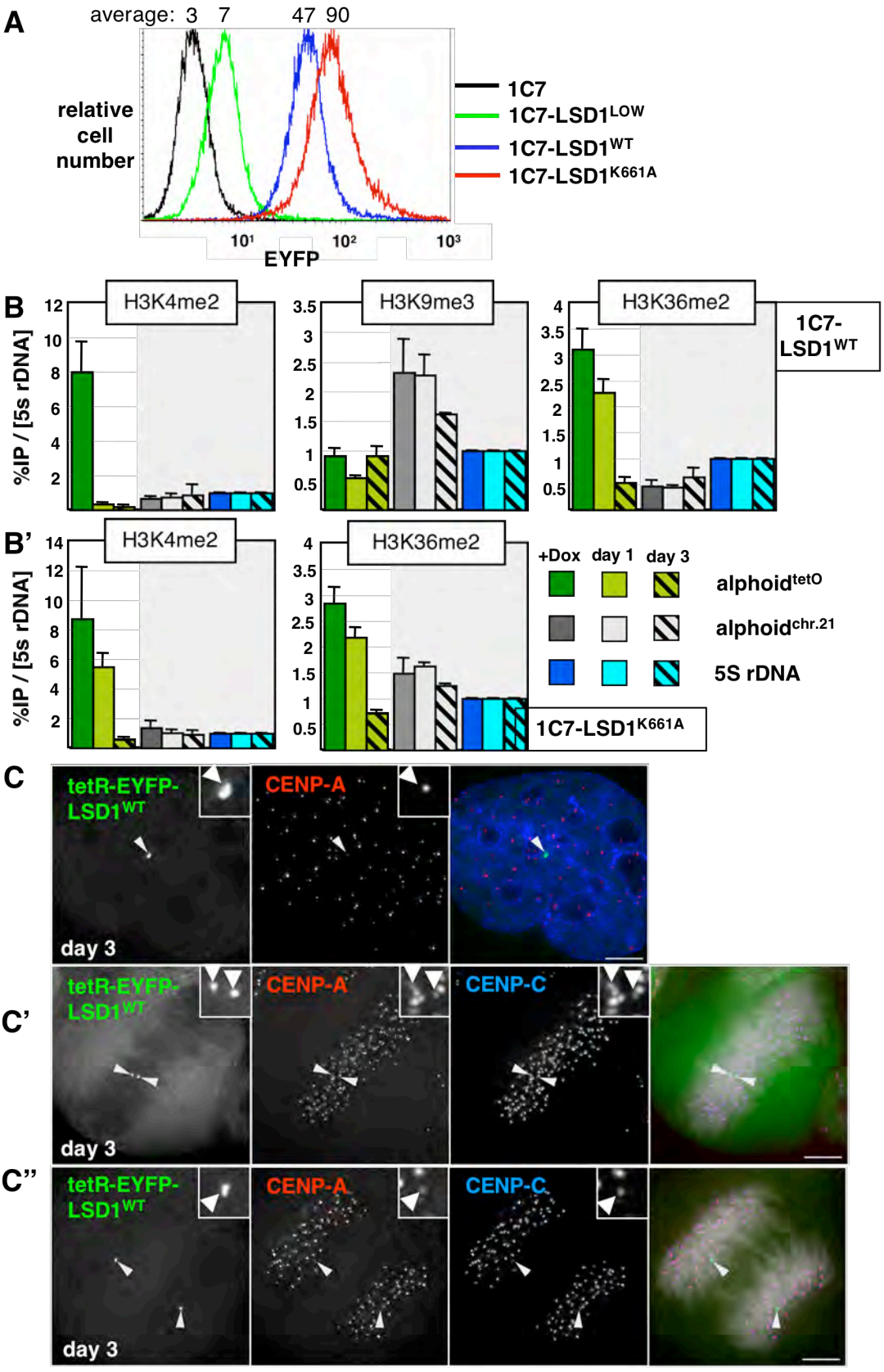

A
average: 3 7 47 90
relative cell number
1C7
1C7-LSD1LOW
1C7-LSD1WT
1C7-LSD1K661A
10¹ 10² 10³
EYFP
B
%IP / [5s rDNA]
H3K4me2
H3K9me3
H3K36me2
1C7-LSD1WT
B'
%IP / [5s rDNA]
H3K4me2
H3K36me2
+Dox day 1 day 3
alphoidtetO
alphoidchr.21
5S rDNA
1C7-LSD1K661A
C
tetR-EYFP-LSD1WT
CENP-A
day 3
C'
tetR-EYFP-LSD1WT
CENP-A
CENP-C
day 3
C''
tetR-EYFP-LSD1WT
CENP-A
CENP-C
day 3

Figure 12| Centromeric H3K4me2 is not required for a functional kinetochore structure. A) Flowcytometric analysis of expression levels of 1C7-LSD1 cell lines expressing tetR-EYFP-LSD1^{WT} at low (1C7-LSD1^{LOW}; green) or high (1C7-LSD1^{WT}; blue) levels, and of 1C7 cells stably expressing the fusion of tetR-EYFP with the catalytic mutant LSD1 (1C7-LSD1^{K661A}; red). Parental 1C7 cells expressing no fusion construct are indicated (black). Average fluorescence values are indicated above the histograms. **B)** ChIP analysis with antibodies of the indicated specificities in 1C7-LSD1^{WT} (B) and 1C7-LSD1^{K661A} (B') cells prior to (+Dox) and at the indicated time points after wash-out of doxycyline. Oligonucleotide primers specific for the synthetic HAC centromere (alphoid^{tetO}), the chromosome 21 centromere (alphoid$^{chr.21}$) or the endogenous 5S rDNA locus were used for real-time PCR quantification of precipitated material. Percentages of precipitated material relative to the input were normalized to those of the 5S rDNA locus, which forms a genomic region with an invariable histone modification profile. Data normalization was performed to account for higher variability in the efficiency between ChIP experiments in the context of 1C7 cells, which is likely owing to the comparatively low HAC to chromatin ratio. Data represents the mean and standard deviation of two or more independent ChIP experiments. **C)** Immunofluorescence analysis of 1C7-LSD1^{WT} cells in interphase (C), metaphase (C') and anaphase (C") three days after washing out doxycyline. Cells were stained with antibodies against CENP-A (red; C, C', C") and CENP-C (blue; C', C"). Arrowheads depict the HAC (C) or individual HAC sister chromatids (C', C"). Scale bar: 5 µm.

Interestingly, levels of H3K9me3 at the HAC centromere dropped slightly within the first 24 hours of targeting by the LSD1 construct in 1C7-LSD1^{WT} cells (Fig. 12B). The levels of this mark recovered again by day 3, however. Importantly, the lack of centromeric H3K4me2 did not result in a marked increase of centromeric H3K9me3 levels, arguing against a role of this modification in active barrier formation towards pericentromeric heterochromatin.

Staining for CENP-A in 1C7-LSD1^{WT} cells on day 3 after doxycycline wash-out revealed the expected, compacted and spherical shape at both, interphase and mitotic HACs (Fig. 12C, 12C' and 12C"). Similarly, mitotic CENP-C structure at the HAC did not appear to be perturbed detectably (Fig. 12C' and 12C"). Importantly, LSD1-targeted HAC kinetochores remained functional at this time point, with HAC sister chromatids aligning under tension on the metaphase plate (Fig. 12C') and segregating normally in subsequent anaphase (Fig. 12C"). Together, these data demonstrate that centromeric H3K4me2 is not directly required to maintain folding of the CENP-A chromatin domain, nor for immediate kinetochore function.

Centromere Tethering of LSD1 Interferes with the Long-term Maintenance of Centromeric Structure

As pointed out above, tethering of tetR-EYFP does not interfere with the HAC centromere structure or kinetochore function either after transient expression or during stable targeting into the HAC kinetochore for up to 30 days ((Cardinale et al., 2009; Nakano et al., 2008) and data not shown). However, upon careful inspection of my immunofluorescence data, I noticed a decrease of HAC-associated CENP-C staining in 1C7 cells expressing tetR-EYFP-LSD1, compared to cells expressing tetR-EYFP only (Fig. 13A and 13A'). Assessing cells displaying comparable expression levels of the fusion constructs, I quantified the CENP-C signal associated with the

alphoid^{tetO} centromere four days after transfection with the relevant constructs. Reproducibly, CENP-C staining intensities at HACs targeted with tetR-EYFP-LSD1 were significantly reduced over those targeted with tetR-EYFP only (Fig. 13B). Importantly, tethering of the K661A mutant construct failed to cause any significant difference in the HAC-associated CENP-C staining compared to tetR-EYFP (Fig. 13A" and 13B).

To unambiguously assess the loss of centromeric structure mediated through LSD1 catalytic activity in more detail, I performed an immunofluorescence time-course experiment in a heterogeneously expressing population of stable 1C7-LSD1 cells. Staining for either CENP-A or CENP-C and subsequent quantification of the fluorescence signal at the HAC centromere at one, three and five days after doxycyline wash-out revealed a gradual loss over time of both CENP-A and CENP-C from the HAC centromere (Fig. 14A-D). Importantly, CENP-A and CENP-C levels appeared to decrease concomitantly. This was different from the hierarchical disassembly resulting from tetR-mediated tethering of KAP1, in which loss of CENP-C preceded loss of CENP-A at the HAC kinetochore (see section K).

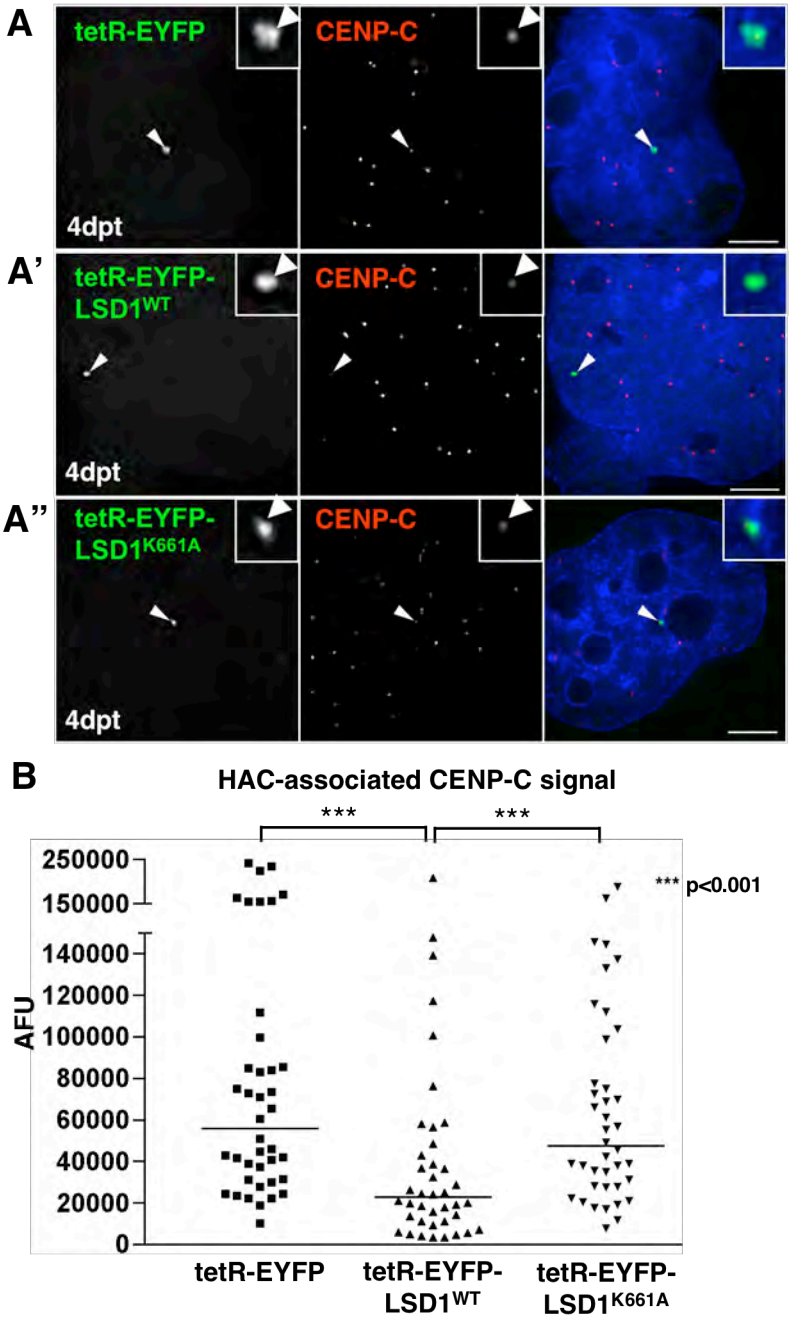


Figure 13| Tethered LSD1 enzymatic activity interferes with long-term maintenance of HAC kinetochore structure. A) Immunofluorescence analysis of 1C7 cells four days after transfection with constructs expressing tetR-EYFP (A), tetR-EYFP-LSD1^{WT} (A') or tetR-EYFP-LSD1^{K661A} (A''). Cells were stained for CENP-C (red) and reflect the median staining levels determined in the quantification in (B). Arrowheads depict the HAC. Scale bar: 5 µm. *Figure legend continued overleaf.*

Figure 13 continued.

(B) Quantification of the HAC-associated CENP-C staining in cells transfected as in (A). Arbitrary fluorescence units (AFU) are plotted. Differences in CENP-C staining between HACs targeted by tetR-EYFP and tetR-EYFP-LSD1^{WT} or tetR-EYFP-LSD1^{WT} and tetR-EYFP-LSD1^{K661A} are statistically significant ($p<0.001$; Mann-Whitney test).

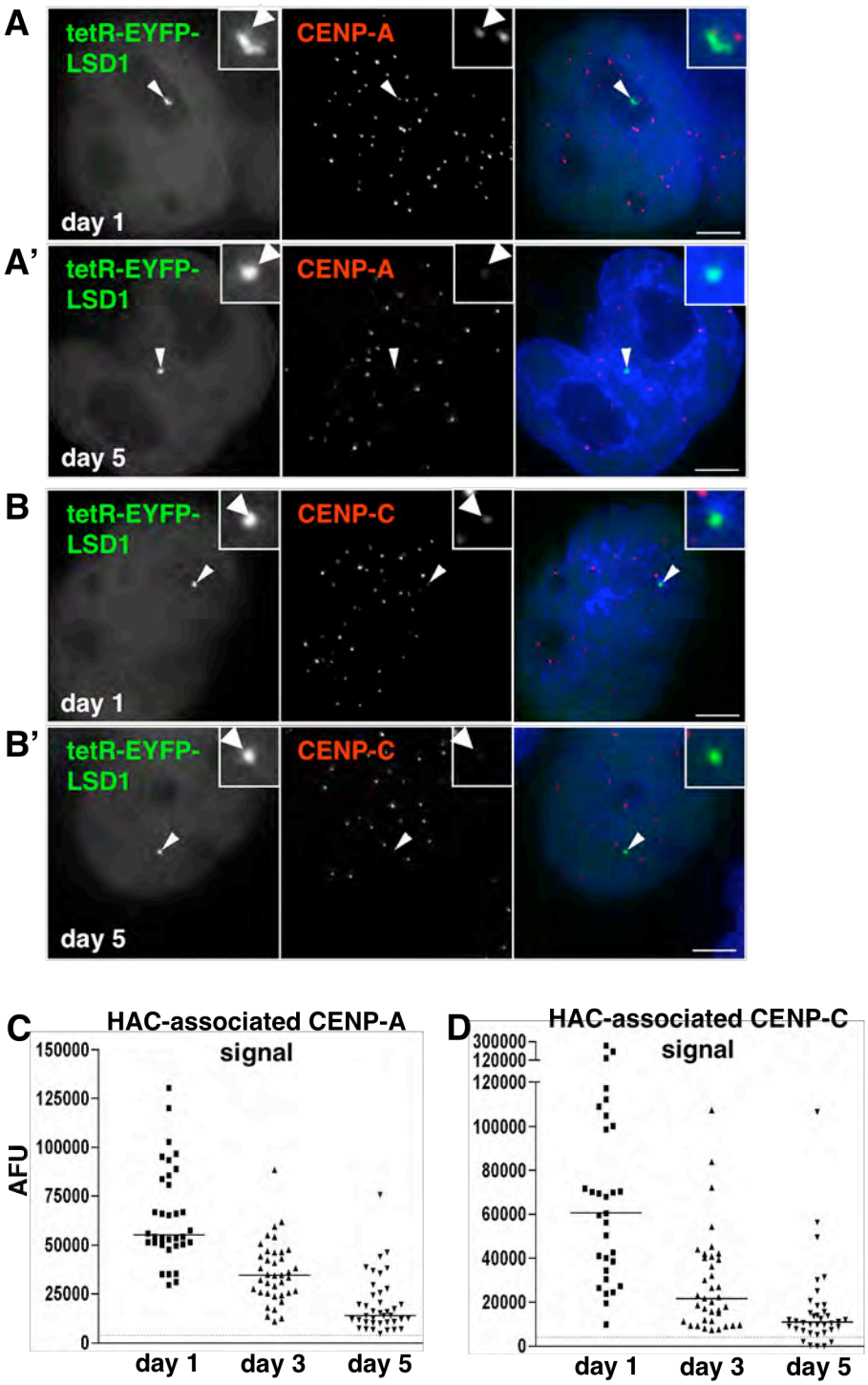


Figure 14| Tethered LSD1 induces a gradual, concommitant loss of CENP-A and CENP-C from the HAC kinetochore. *Figure legend continued overleaf.*

Figure 14 continued.

A, B) Immunofluorescence analysis of 1C7 cells stably expressing tetR-EYFP-LSD1^{WT} at one (A, B) and five (A', B') days after washing out doxycycline. Cells were stained with antibodies against CENP-A (A, A') or CENP-C (B, B') and examples chosen to represent the median values of the corresponding quantifications in C and D, respectively. Arrowheads depict the HAC. Scale bar: 5 µm. **C, D)** Quantification of HAC-associated CENP-A (C) and CENP-C (D) fluorescence staining at the indicated time points after doxycycline wash-out in the time course experiment in (A) and (B).

Notably, the dynamics observed in these time-course experiments suggested a passive loss of CENP-A from the HAC centromere, presumably as a result of inefficient incorporation of newly-synthesized CENP-A molecules. This incorporation normally occurs in a defined time window at the transition from mitosis into G1 phase (Jansen et al., 2007). To directly confirm this hypothesis, I performed SNAP-tag based quench-pulse-chase experiments (Jansen et al., 2007) by co-transfecting 1C7 cells with a plasmid expressing CENP-A-SNAP-3xHA and the constructs expressing either tetR-EYFP or tetR-EYFP-LSD1. Transfected cells were subsequently arrested in S phase using thymidine. During the block, existing SNAP-tagged CENP-A molecules were irreversibly quenched and, after thymidine release, newly-synthesized CENP-A molecules were labelled using the fluorescent TMR-Star substrate (Fig. 15A and section G.6 of the Materials and Methods). I analyzed incorporation of TMR-Star-labelled CENP-A in subsequent G1 by fluorescence microscopy (Fig. 15B and 15B'). To account for differential expression levels of the SNAP-tagged CENP-A construct, I quantified the HAC-associated TMR-Star signal relative to the average TMR-Star signal measured at endogenous centromeres. Consistent with the hypothesis that LSD1 activity interferes with efficient incorporation of CENP-A into the HAC centromere, tethering of LSD1 to the HAC centromere resulted in a statistically significant reduction of HAC-associated TMR-Star signal compared to tethering of tetR-EYFP alone (Fig. 15C).

Repression of Centromere Transcription Alters the CENP-A Chromatin Domain

Native LSD1 interacts directly with CoREST, thereby forming part of the histone deacetylase (HDAC)-1 or HDAC-2 containing BHC complex that is involved in repression of neuronal gene promoters (Hakimi et al., 2002; Humphrey et al., 2001; Lee et al., 2005; Shi et al., 2003). As LSD1(K661A) retains its ability to form part of the BHC complex (Lee et al., 2005), this

construct can be used to dissociate the contribution of BHC-mediated repression from the H3K4-demethylating, intrinsic activity of LSD1 towards the loss of centromeric CENP-A. To assess if LSD1 represses transcription through the HAC centromere, I performed real-time RT-PCR analysis in 1C7-LSD1^{WT} as well as 1C7-LSD1^{K661A} stable cell lines. The wild-type LSD1 fusion construct proved to be an extremely efficient repressor, reducing alphoid^{tetO} transcript copy numbers by more than 70 % within the first 24 hours after washing out of doxycycline (Fig. 16A). By day 3, transcripts derived from the HAC centromere were barely detectable. Consistently, levels of RNA polymerase II at the alphoid^{tetO} array, despite being extremely low, were reproducibly detectable prior to doxycycline wash-out, but essentially undetectable 24 hours thereafter (Fig. 16B). The LSD1 K661A mutant construct also induced repression of the HAC centromere (Fig. 16A). Importantly however, this construct was significantly less efficient at repressing HAC centromere transcription, particularly within the initial 24 hours, during which transcript levels decreased only by about 30 %. Furthermore, in contrast to the wild-type LSD1 fusion, 1C7-LSD1^{K661A} cells retained moderate (>30 %) levels of transcriptional activity at the alphoid^{tetO} centromere even after 3 days of targeting. These data show that in the context of the fusion constructs used, LSD1 H3K4-demethylating activity strongly augments the repressive effect at the HAC centromere. Importantly, repression of HAC centromere transcription in 1C7-LSD1^{K661A} cells was paralleled by a gradual decrease in the levels of H3K4me2 and H3K36me2 (Fig. 12B'). As methylation of H3K36 and H3K4 occurs co-transcriptionally (Keogh et al., 2005; Krogan et al., 2003a; Ng et al., 2003), these data therefore implicate centromeric non-coding transcription in the maintenance of the local chromatin state.

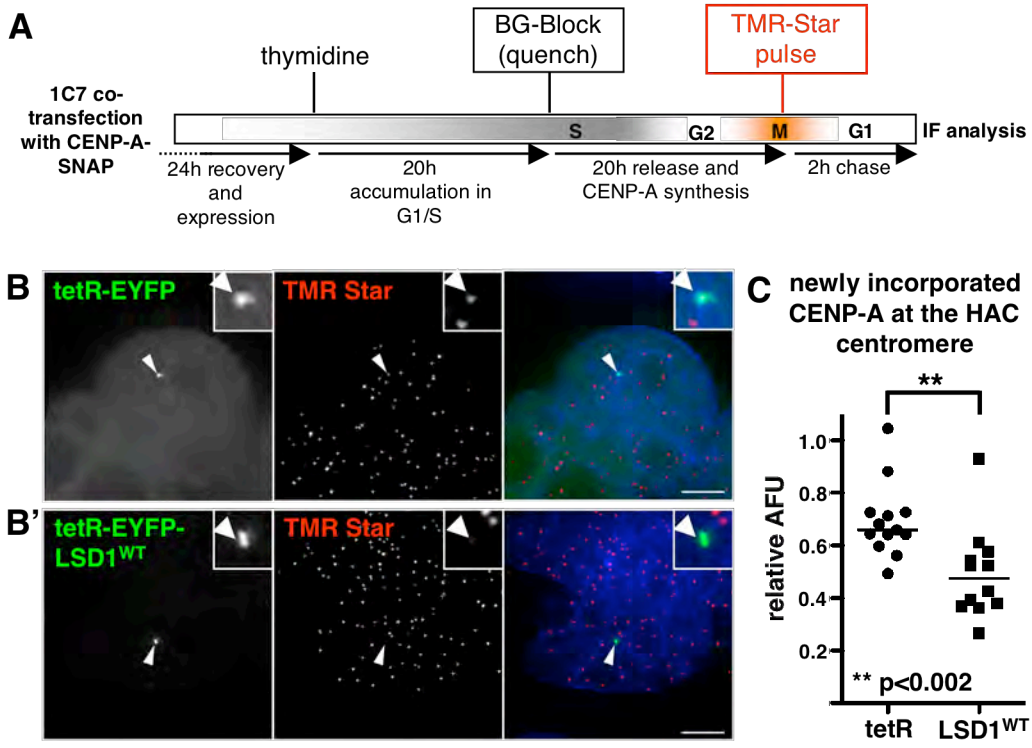


Figure 15| LSD1 tethering reduces the efficiency of CENP-A loading into the HAC kinetochore. A) Schematic of the work-flow to determine incorporation of newly synthesized CENP-A molecules into the HAC kinetochore. **B)** Immunofluorescence analysis of SNAP-tagged CENP-A in 1C7 cells labelled with the TMR-Star substrate (red) in cells expressing tetR-EYFP (B) or tetR-EYFP-LSD1^{WT} (B'). Arrowheads indicate the HAC. Scale bar: 5 µm. **C)** Quantification of HAC-associated TMR-Star fluorescence signal relative to the average signal measured at all endogenous centromeres in cells expressing tetR-EYFP (tetR) or tetR-EYFP-LSD1^{WT} (LSD1^{WT}). Differences in TMR-Star signal between these two data sets are significant ($p<0.002$; t-test).

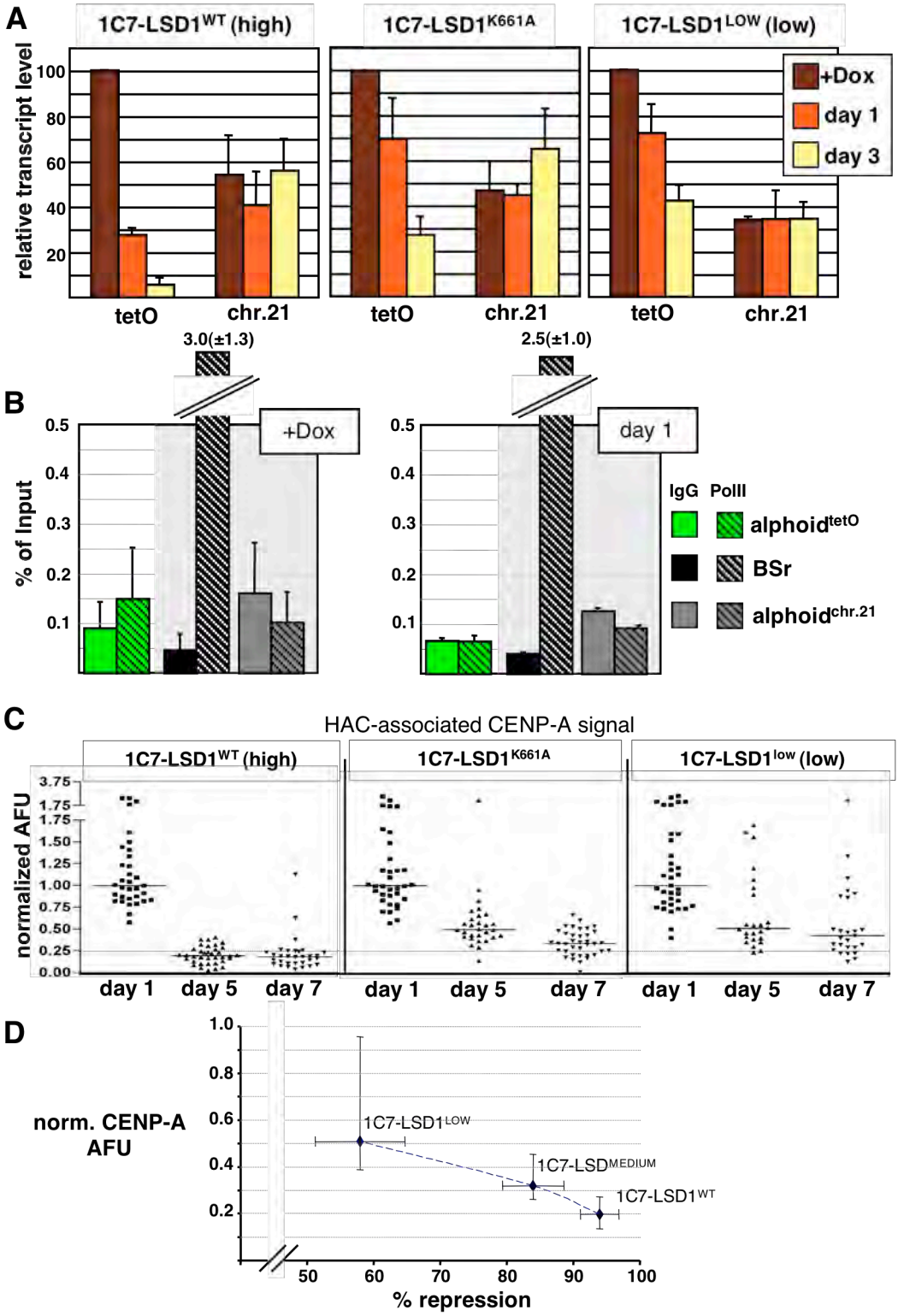


Figure 16| Repression of HAC centromere transcription negatively correlates with maintenance of CENP-A. *Figure legend continued overleaf.*

Figure 16 continued.

A) Real-time RT-PCR analysis of HAC (tetO) and chromosome 21 (chr. 21) centromeric transcripts in the indicated cell lines prior to (+Dox), and one or three days after doxycycline wash-out. Expression levels are normalized to the genomic copy number and levels of β-actin, and plotted relative to the +Dox level of the alphoid^{tetO} transcripts. Data represent mean and S.E.M. of three or more independent experiments. Differences in the repressive activity of wild-type (high) and mutant LSD1 fusions is statistically significant at both time points ($p=0.015$ and $p<0.001$ for day1 and day3, respectively; t test) **B)** ChIP analysis of 1C7-LSD1^{WT} cells prior to (+Dox) and one day after doxycycline wash-out using unspecific IgG or an antibody against RNA polymerase II (PolII). Data represents the mean and S.E.M. or two or more independent experiments. **C)** Immunofluorescence analysis of the indicated cell lines using an antibody against CENP-A and subsequent quantification of the HAC-associated CENP-A signal at the indicated time points after doxycycline wash-out. AFU values were normalized to the median AFU of the corresponding day 1 time point. For orientation, the dotted line indicates the 25 % mark. **D)** Immunofluorescence analysis and CENP-A signal quantification in three 1C7-derived cell lines stably expressing tetR-EYFP-LSD1^{WT} at different levels (HIGH > MEDIUM > LOW). The median AFU values and inter-quartile ranges measured five days after doxycycline wash-out are plotted on the y-axis against the mean repression of the alphoid^{tetO} and corresponding standard deviation on the x-axis, as determined by real-time RT-PCR on day three. Day 5 and day 3 endpoints were chosen and correspond to the time points displaying the strongest effect determined in "HIGH" cells.

To assess the long-term effects of LSD1- and LSD1(K661A) tethering on centromere structure, I quantified levels of HAC-associated CENP-A staining in a time-course experiment in 1C7-LSD1^{WT} and 1C7-LSD1^{K661A} cells. Strikingly, the mutant LSD1 fusion induced much slower loss of CENP-A from the HAC centromere compared to LSD1^{WT} (Fig. 16C). While by day 5 after doxycyline wash-out the majority of HACs in 1C7-LSD1^{WT} cells displayed a reduction of CENP-A staining of more than 80 % relative to the intensities at day 1, most 1C7-LSD1^{K661A} cells retained CENP-A staining of 50 % or higher. Indeedy, even by seven days after induced targeting, the majority of HACs retained CENP-A intensity values considerably greater than 25 % in cells expressing tetR-EYFP-LSD1^{K661A} (Fig. 16C).

Repression of Centromere Transcription and CENP-A Loss Show a Dose-dependent Correlation

I was intrigued by the strong correlation of the dynamics of transcriptional repression and the loss of CENP-A from the HAC centromere in 1C7-LSD1^{WT} and 1C7-LSD1^{K661A} cells, respectively. This could indicate a tight association between transcription through the centromere and the maintenance of CENP-A levels. LSD1 exerts a non-linear, dose-dependent repressive effect in reporter assays (Shi et al., 2004). I therefore isolated an additional 1C7-LSD1^{WT} clone (1C7-LSD1^{LOW}) that showed about 7-fold less tetR-EYFP-LSD1 expression based on flow cytometric analysis (Fig. 12A). Expression levels determined by flow cytometry correlated well with the steady-state amount of fusion construct present at the interphase HAC, as determined by direct fluorescence signal quantification of microscopic images (data not shown). Consistent with a dose-dependent adverse effect on transcription, 1C7-LSD1^{LOW} cells exerted a moderate repression at the alphoid^{tetO} centromere that was significantly milder compared to 1C7-LSD1^{WT} cells (Fig. 16A). Importantly, quantification of the HAC-associated CENP-A signal in 1C7-LSD1^{LOW} cells also revealed a decreased loss of CENP-A from

the HAC centromere over a 7 day time-course (Fig. 16C), with the majority of HACs retaining CENP-A signals well above 25 % of the values measured one day after doxycyline wash out. Isolation of a further 1C7-LSD1 clone (1C7-LSD1^{MEDIUM}) with a repressive effect intermediate to 1C7-LSD1^{LOW} and 1C7-lSD1^{WT} cells allowed me to plot alphoid^{tetO} transcriptional repression against the CENP-A levels measured at the HAC centromere in these three cell lines. Figure 16D strikingly illustrates the negative correlation between the extent of early centromeric repression and the maintenance of CENP-A levels.

To assess the impact of LSD1-mediated loss of centromere structure on kinetochore function, I analysed mitotic HACs in 1C7-LSD1^{WT} and 1C7-LSD1^{LOW} cells at different time points after doxycycline wash-out. As pointed out earlier, HAC kinetochores showed no obvious mitotic defects after only 3 days of tethering the fusion (Fig. 12C and 17A). However, by 7 days after doxycycline wash-out, HAC sister chromatids in 1C7-LSD1^{WT} cells frequently failed to align on the metaphase plate (Fig. 17A'). HACs were furthermore prone to missegregation in anaphase, reflected by 1C7-LSD1^{WT} cells frequently showing aberrant HAC copy numbers (Fig. 17A" and Fig. 17C). These observations demonstrate that gradual loss of CENP-A from the HAC centromere resulted in inactivation of its kinetochore. In contrast, HACs in 1C7-LSD1^{LOW} cells suffered mitotic abnormalities at the later time point only infrequently, with the majority of HACs aligning on the metaphase plate (Fig. 17A"'). In the latter case, only a few cells showed HAC-specific mitotic defects (Fig. 17B), and these generally did not display gain or loss of HAC copies (Fig. 17C), consistent with the retention of higher levels of CENP-A at the HAC centromere.

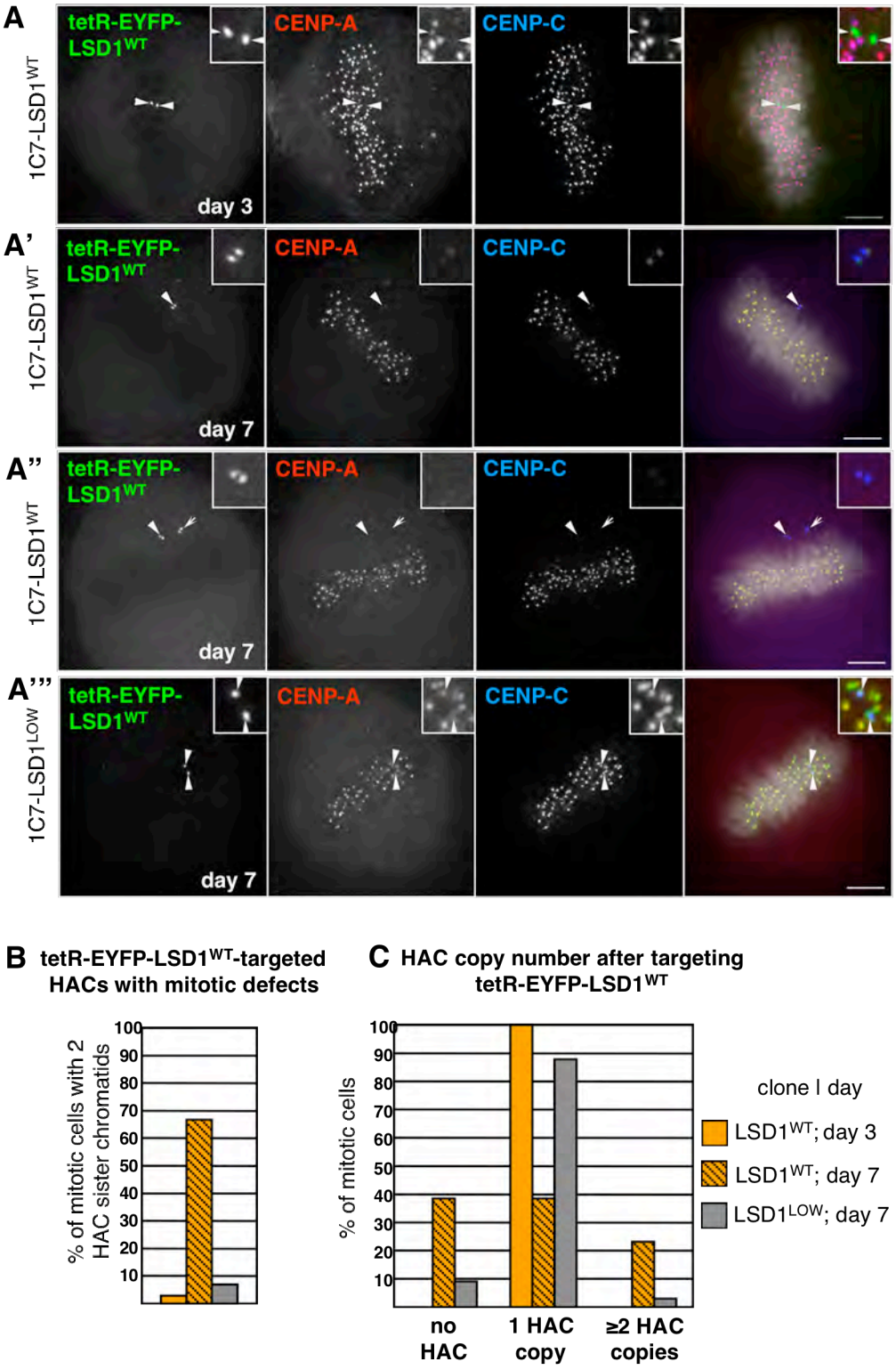


Figure 17| Targeting of LSD1 catalytic activity interferes with long-term HAC kinetochore function. *Figure legend continued overleaf.*

Figure 17 continued.

A) Immunofluorescence analysis of mitotic 1C7-LSD1^{WT} cells (A-A”) expressing tetR-EYFP-LSD1 at high levels and of 1C7-LSD1^{LOW} cells (A’”) expressing the fusion construct at low levels at the indicated time points after washing out doxycyline. Cells were co-stained for CENP-A (red) and CENP-C (blue). Arrowheads in (A) and (A’”) depict resolved HAC sister chromatids. Arrowheads in (A’) and (A”) depict unaligned HACs. The arrow in (A”) points to a second copy of the HAC. Scale bar: 5 µm. **B)** Quantification of HAC-specific mitotic abnormalities (unaligned metaphases as in A’ plus unequal segregation of HAC sister chromatids in anaphases, not pictured) in the indicated cell lines and time points. Only mitotic cells displaying a single HAC copy were scored. (n=36, n=15 and n=29 for clones LSD1^{WT}/day3, LSD1^{WT}/day7 and LSD1^{low}/day7, respectively) **C)** Quantification of HAC copy numbers as determined by the associated EYFP signals in the indicated cell lines and time points. (n=36, n=39 and n=33 for clones LSD1^{WT}/day 3, LSD1^{WT}/day 7 and LSD1^{low}/day7, respectively).

L.3 Discussion

In this section, I have employed a tetR fusion construct of the lysine-specific histone demethylase LSD1 as a molecular tool to probe, at the single alphoid^{tetO} HAC centromere, the relevance and role of H3K4me2. H3K4me2 was identified as a distinctive feature of centromere chromatin (Sullivan and Karpen, 2004) and as pointed out above, its presence within the chromatin underlying kinetochores has been subject to intense speculations. However, until now, the functional significance of centromeric H3K4me2 has not been accessible for direct interrogation. The present work represents the first targeted engineering of the chromatin underlying an operative kinetochore *in vivo*, and forms the foundation for future analysis of distinctive aspects of centromere chromatin using comparable enzymatic fusion constructs. Although particular conditions will likely depend on the nature of each tetR-associated enzymatic activity itself, tetR-EYFP-LSD1 highlights the remarkable efficiency of this approach, which reduced levels of alphoid^{tetO}-associated H3K4me2 to background as determined by ChIP analysis. Importantly, depletion of H3K4me2 depended on LSD1 catalytic activity, was specifically constrained to the targeted HAC centromere and did not detectably affect levels of this mark at other genomic loci.

Employing this system, my initial aim was to determine if centromeric H3K4me2 is directly required for the formation or maintenance of a functional higher-order arrangement of centrochromatin, such as the solenoid structure postulated by Gary Karpen and co-workers (Blower et al., 2002; Sullivan and Karpen, 2004). My finding that HAC kinetochores depleted of H3K4me2 in their underlying chromatin remain functional over short term strongly argues against this hypothesis. Within the resolution limit of the light microscope, HAC-associated CENP-A structure itself appeared unperturbed, while biochemical unravelling of CENP-A nucleosome arrays is readily detectable with comparable magnification power (Blower et al., 2002; Ribeiro et al.,

2010). These data further support the hypothesis that H3K4me2-independent factors are sufficient to maintain the highly compacted state of the CENP-A chromatin domain. Arguably, LSD1 activity tethered to the alphoid^{tetO} array might not result in absolute loss of local H3K4me2, with HAC centromeres retaining very low levels of this mark that are not detected in the ChIP assay. However, real-time PCR analysis following ChIP is sensitive enough to detect single copies of target DNA molecules, and the absence of considerable enrichment of alphoid^{tetO} DNA after pull-down with antibodies against H3K4me2 over the IgG background is strongly indicative of near-complete depletion of this modification. In light of the comparatively vast size of endogenous CENP-A chromatin domains (~70kb or longer (Alonso et al., 2007; Alonso et al., 2010; Lo et al., 2001; Ribeiro et al., 2010), it appears unlikely that such low levels of H3K4me2 would be sufficient to maintain the required degree of compaction observed at kinetochores.

The above data uncover an intriguing relationship between centromeric H3K4me2, local transcriptional activity and long-term maintenance of the CENP-A chromatin domain. In the context of HAC-tethered LSD1 fusion constructs, abrogation of local H3K4me2 significantly enhanced the repression of alphoid^{tetO} centromere transcription. This suggests that centromeric H3K4me2 could positively mediate local transcription of the alphoid^{tetO} array. Although the molecular mechanisms are presently unclear, methylation of H3K4 is thought to form a local chromatin mark that serves as “memory” for recent transcriptional activity (Li et al., 2007). This notion is emphasized by hypermethylation of H3K4 found at and downstream of promoters of actively transcribed genes (Barski et al., 2007; Mikkelsen et al., 2007; Schneider et al., 2004). Consistently, interphase patterns of H3K4 methylation appear largely conserved in mitotic chromatin, while RNA polymerase and other chromatin-remodelling activities are fundamentally evicted from the chromosomes (Blobel et al., 2009; Kouskouti and Talianidis, 2005). Compelling evidence for a direct role of H3K4

methylation in regulating transcriptional activity of individual genes after mitotic exit comes from a recent live cell study in *Dictyostelium*, where abrogation of H3K4 methylation results in the loss of similarity of transcriptional bursts from mother to daughter cells (Muramoto et al., 2010). Finally, the mere presence within a transcriptional repressor complex of an enzymatic activity, in form of LSD1, that specifically demethylates H3K4me2 (Shi et al., 2004) emphasizes a biologically relevant link between this modification and transcriptional maintenance. Interestingly, a chromatin-modifying pathway based on H3K4me2 was recently implicated in facilitating transcriptional elongation in yeast (Kim and Buratowski, 2009).

Importantly, the degree of repression of transcription through the HAC centromere strongly correlates with the loss of HAC-associated CENP-A and the ultimate disruption of kinetochore function. Interestingly, knock-down of the transcription elongation-associated chromatin remodelling factor CHD1 also results in a reduction of centromeric CENP-A levels (Okada et al., 2009). At the HAC centromere, transcriptional repression is further paralleled by a gradual loss of H3K4 and H3K36 methylation, consistent with the co-transcriptional establishment of these marks illustrated earlier. Together with the data presented and discussed in section J, I would like to propose that non-coding transcription through the centromere shapes and importantly maintains a chromatin environment that is “permissive” for the deposition of newly-synthesized CENP-A molecules. In my model (Fig. 18), low levels of RNA polymerase leave in their wake a characteristic chromatin environment through recruitment of the relevant nucleosome-modifying activities (see section J and references therein). This chromatin state directly or indirectly facilitates the incorporation of CENP-A into centromeric nucleosomes, thereby ensuring that centromere identity is faithfully maintained over generations. In addition to direct modification of the local chromatin environment by RNA polymerase-associated factors, non-coding transcripts might additionally shape aspects of the pericentromeric heterochromatin via

an RNAi pathway analogous to yeast (Volpe et al., 2003). Using locked nucleic acid antisense oligonucleotides specifically designed to sequester nuclear alphoid^{tetO} transcripts, I did not measure any detectable changes in the levels of CENP-A or CENP-C at the HAC centromere (data not shown). However, my experiments could not control for the efficiency of this approach *in vivo*.

On the functional level, the chromatin context might mediate the process termed “centromere licencing” during mitosis which depends on the Mis18 complex (Fujita et al., 2007; Hayashi et al., 2004; Maddox et al., 2007), or play a role during the HJURP-dependent loading of CENP-A itself at the transition from mitosis into G1 (Dunleavy et al., 2009; Foltz et al., 2009). Consequently, repression of centromeric transcription results in the loss of its special chromatin character and lead to subsequent failure to efficiently deposit CENP-A. This would cause a progressive loss of centromere identity and function.

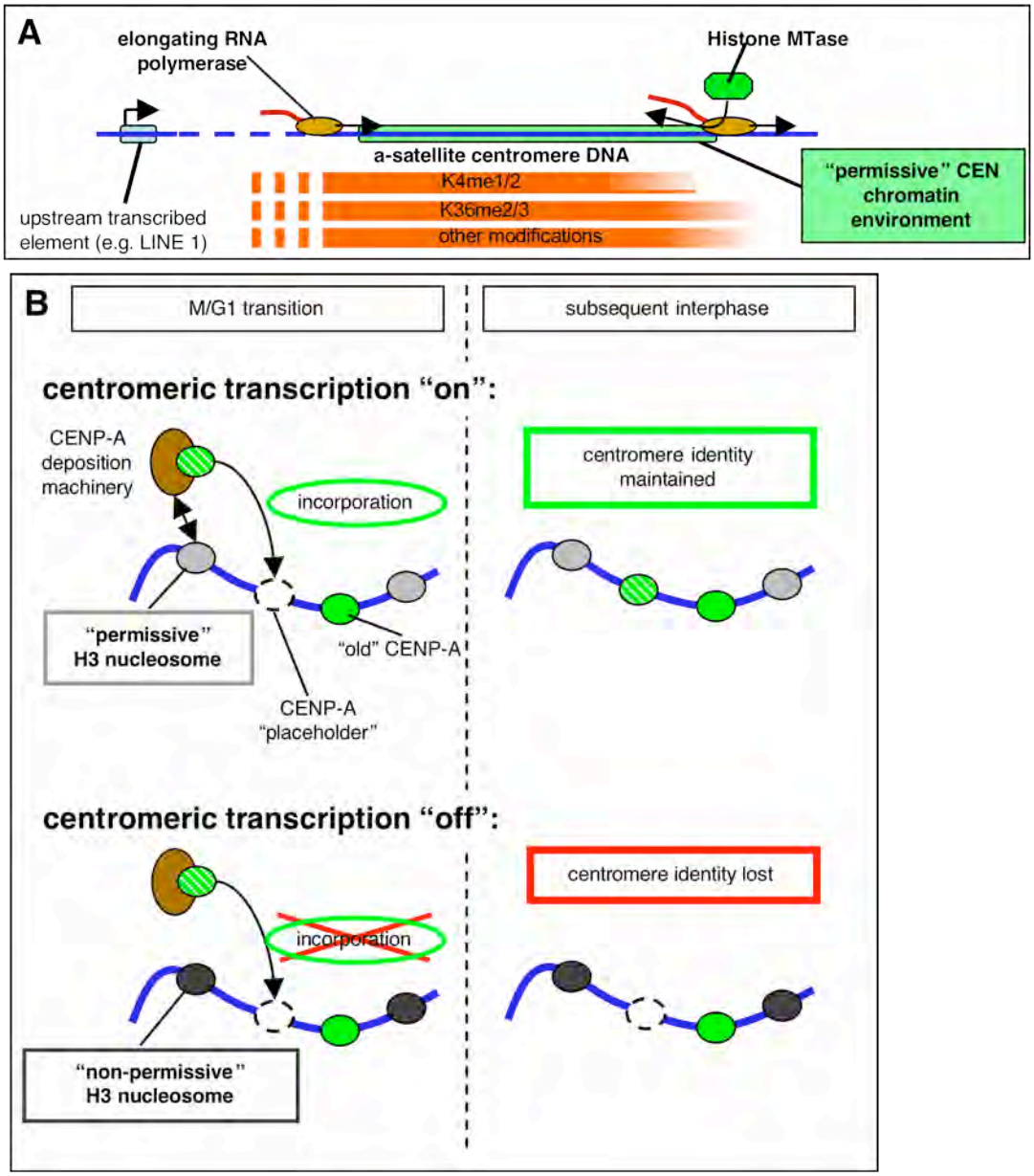


Figure 18| Model for a role of transcription in maintaining centromere chromatin character and identity. A) Low levels of elongating RNA polymerase leave in their wake the characteristic histone modification profile described in chapter 1. The resulting chromatin environment is "permissive" for centromere maintenance. **B)** This permissive chromatin directly or indirectly facilitates deposition of newly synthesized CENP-A molecules and thereby ensures maintenance of the CENP-A chromatin domain. Loss of local transcription results in the abrogation of the required chromatin character, resulting in loss of CENP-A deposition and centromere identity.

M RAPID AND SPECIFIC LOSS OF CENTROMERIC CENP-A NUCLEOSOMES FOLLOWING TETHERING OF A POTENT TRANSCRIPTIONAL ACTIVATOR

M.1 Background

Upon transfection of alphoid DNA arrays, the majority of stable transfectants exhibit ectopic integration of the input DNA into host chromosomes (Ikeno et al., 1998; Okamoto et al., 2007). However, in contrast to HACs that support formation of a functional kinetochore assembled on exactly the same primary input sequence, most of the integrated alphoid arrays do not recruit detectable levels of CENP-A or other kinetochore components (Nakano et al., 2003; Okamoto et al., 2007), suggesting that either *de novo* deposition of CENP-A or its maintenance at these loci is inefficient. Interestingly, CENP-A levels at integrated, heterochromatic α21-I-derived satellite arrays increase following transient exposure to the histone deacetylase inhibitor trichostatin A (Nakano et al., 2003; Okamoto et al., 2007), indicating that the local chromatin environment may be a critical factor with respect to regulating local CENP-A concentration.

As with other alphoid input DNA, transfection of HT1080 cells with the synthetic alphoid^{tetO} plasmid yielded a high proportion of stable transfectants carrying the construct integrated into host chromosomes and not supporting centromere function (Nakano et al., 2008). I reasoned that a comparative analysis of the alphoid^{tetO}-associated chromatin environment in HAC and integration cell lines would identify differences in individual histone modifications or the over-all chromatin state when the same array is active as a centromere or merely a passive insert in a chromosome arm. On the

chromatin level, these differences may be critical in mediating the fate of the alphoid^{tetO} array with respect to centromeric function.

In this section, I demonstrate that alphoid^{tetO} integration sites appear to be highly euchromatic. Employing a fusion of tetR-EYFP to the potent transcriptional activator VP16, I subsequently demonstrate rapid loss of CENP-A nucleosomes from the alphoid^{tetO} HAC centromere as a consequence of local VP16 activity.

M.2 Results

[Some data presented in this section was obtained and/or analyzed by Julia Jakubsche under my direct supervision. J.J.'s contribution is acknowledged in the relevant paragraphs and figure legends.]

The Alphoid^{tetO} Integration Site Displays a Euchromatic Character

In order to characterize the chromatin environment present at an alphoid^{tetO} integration site, I performed chromatin immunoprecipitation analysis in HT1080-derived AB2.5.30 cells analogous to the experiments described in section J. AB2.5.30 cells carry the alphoid^{tetO} vector stably integrated into a chromosome arm. The array recruits only low levels of CENP-A (Megumi Nakano, unpublished data), indicative of a defective CENP-A maintenance and propagation mechanism. Surprisingly, the nature of the chromatin associated with the integrated array was markedly euchromatic (Fig. 19A), characterized by pronounced acetylation of H3K27, high levels of H3K4me2 and significant levels of H3K4me3. In contrast, levels of H3K9me3 were low compared to the levels present at the alphoid^{tetO} HAC array. Strong histone acetylation and H3K4 hypermethylation was also observed at alphoid^{tetO} arrays in four other integration cell lines, all of which lacked appreciable levels of CENP-A (Megumi Nakano, unpublished data).

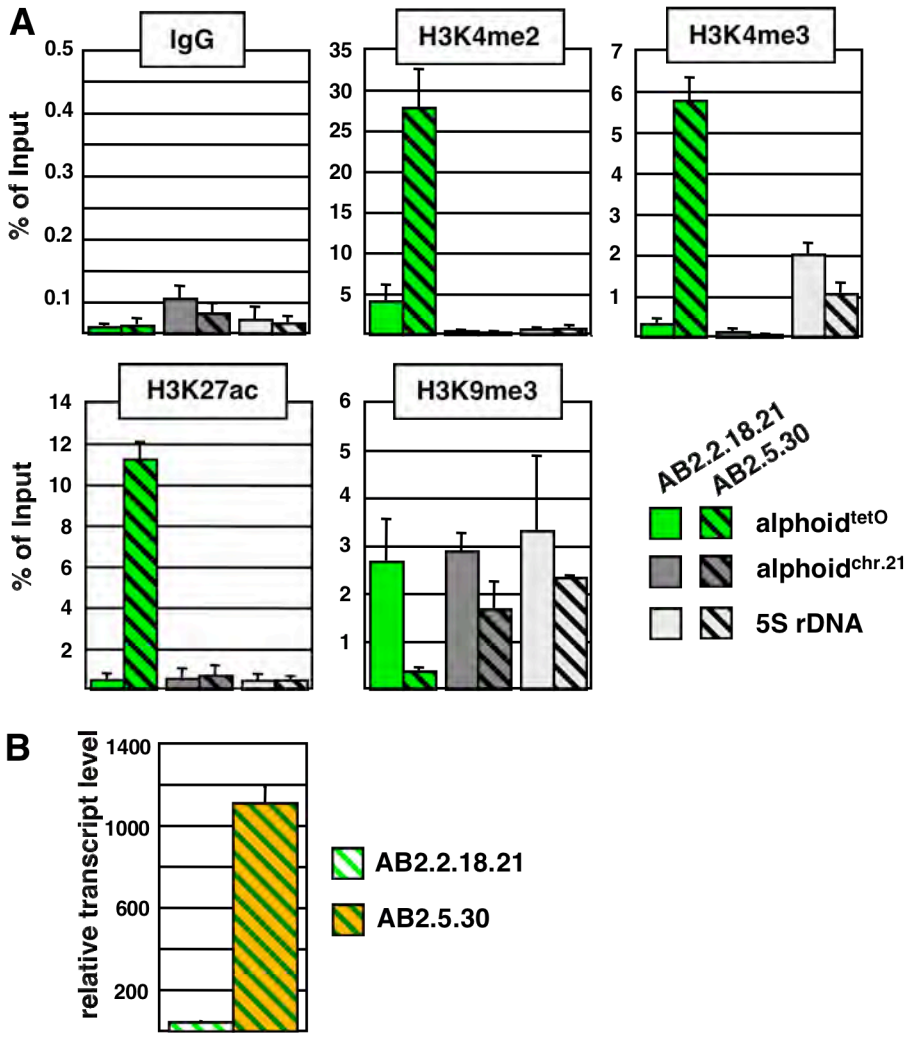


Figure 19| The alphoid^{tetO} integration site is euchromatic. A) ChIP analysis of AB2.2.18.21 (HAC) and AB2.5.30 (integration) cells using antibodies with the indicated specificities. The alphoid^{tetO}, chromosome 21 alphoid (alphoid$^{chr.21}$) and the endogenous 5S rDNA locus were assessed. The percentage of precipitated material is plotted. Data represents the mean and S.E.M. of at least two independent experiments. **B)** Real-time RT-PCR analysis of alphoid^{tetO} transcript copy numbers in AB2.2.18.21 and AB2.5.30 cells. Transcript levels are normalized to those of β-actin, expressed relative to the copy number of the corresponding genomic loci and assigned an arbitrary value. Data shown represents the mean and S.E.M. of three independent experiments.

Consistent with the more euchromatic histone modification profile, real-time RT-PCR analysis in AB2.5.30 cells revealed a steady-state alphoid^{tetO} transcript level exceeding that of HAC-bearing AB2.2.18.21 cells several-fold following normalization to genomic copy number (Fig. 19B).

Tethering of VP16 Induces Rapid Loss of CENP-A from the HAC Centromere

Above findings may point towards an adverse relationship between a euchromatic chromatin environment and the establishment of a functional CENP-A chromatin domain. In earlier studies, targeting a transcriptional activator, tTA-EYFP (tetR-VP16-EYFP), into the alphoid^{tetO} HAC kinetochore resulted in loss of the HAC from some of the cells analyzed (Nakano et al., 2008), which may be the result of inactivation of the associated kinetochore structure. In those experiments, however, effects of the tTA on the underlying chromatin, with respect to transcriptional activation and histone modifications, were mild or undetectable, respectively.

I reasoned that positioning of EYFP at the C-terminus of the tTA may interfere with its activity at the chromatin level. I therefore generated an alternative expression construct in which the viral VP16 trans-activation domain present in the tTA is fused to the C-terminus of tetR-EYFP (tetR-EYFP-VP16). To investigate effects of the VP16 activation domain on alphoid^{tetO} HAC centromere structure, 1C7 cells were transiently transfected with plasmids expressing either tetR-EYFP or tetR-EYFP-VP16. Strikingly, within only 24 hours following transfection, tethering of the VP16 construct into the HAC kinetochore resulted in a strong reduction of the HAC-associated interphase CENP-A staining compared to targeting of tetR-EYFP alone (Fig. 20A'). CENP-A levels were even further reduced over the course of the following 24 hours, with many of the targeted HAC kinetochores failing to display appreciable staining over the nuclear background (Fig.20A").

Consistent with the severe depletion of CENP-A from the underlying chromatin, 65 % of HACs targeted by tetR-EYFP-VP16 displayed mitotic defects at this latter time point, represented as a failure to achieve proper alignment on the metaphase plate or apparent mis-segregation during subsequent anaphase (Fig. 21). In contrast, none of the HACs targeted by tetR-EYFP suffered mitotic defects.

Judged by the associated EYFP signal, interphase HACs targeted by tetR-EYFP-VP16 generally appeared less compacted than those targeted by tetR-EYFP. A fraction of these HACs displayed a pronounced unfolding within the interphase nucleus (Fig. 20B), a feature commonly observed in response to tethering of various activation domains to chromatinized reporter arrays (Carpenter et al., 2005). Most notably, staining for CENP-A indicated that the CENP-A chromatin domain remained compacted, despite the pronounced unfolding of the remainder of the HAC (Fig. 20B). Despite the continued association of tetR-EYFP-VP16 with the chromatin, HACs always achieved compaction in mitosis (see Fig. 21B and data not shown), suggesting that mitotic chromatin condensation must be dominant over the chromatin-unfolding activities recruited by artificially tethered VP16.

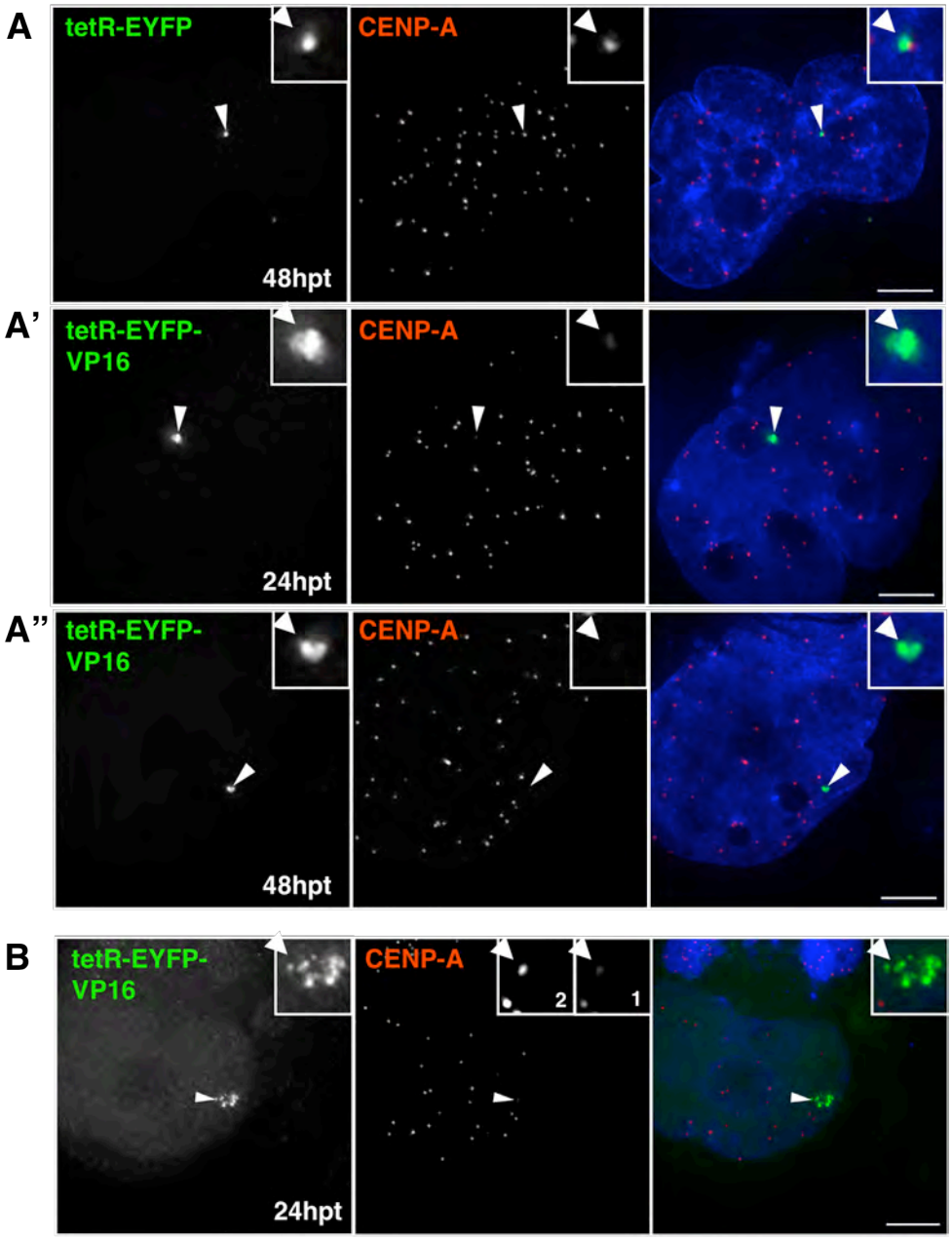


Figure 20| Tethering of VP16 causes rapid loss of CENP-A from the HAC centromere. A, B) Immunofluorescence analysis of 1C7 cells following transient expression of tetR-EYFP (A) or tetR-EYFP-VP16 (A', A", B) using an antibody against CENP-A at the indicated time points after transfection. Arrowheads depict the HAC. The HAC in (B) shows an example of the large-scale unfolding sometimes observed after VP16 tethering. Insets show the HAC-associated CENP-A signal at normal exposure (1) and after digital saturation (2) of the signal. Scale bar: 5 µm.

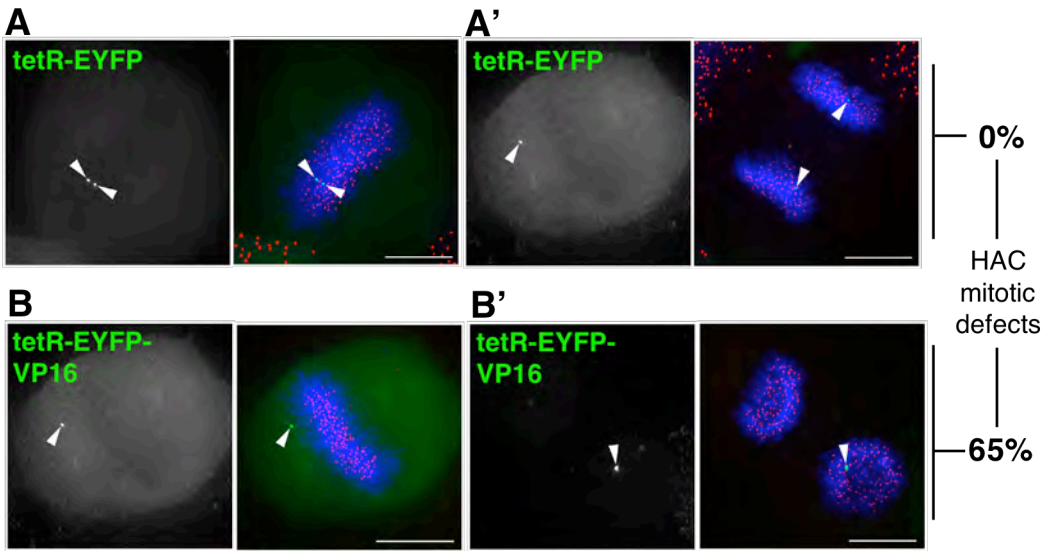


Figure 21| VP16 tethering causes HAC-specific mitotic defects. A, B) Immunofluorescence analysis of mitotic 1C7 cells transiently expressing tetR-EYFP (A, A') or tetR-EYFP-VP16 (B, B') 48h after transfection. (A) and (B) show metaphases, (A') and (B') show late anaphase and telophase cells, respectively. Arrowheads depict individual HAC kinetochores (A, A') or point to misaligned (B) or mis-segregated (B') HACs. HAC-specific defects were scored as unaligned (as in B) or mis-segregated (as in B') HACs and values represent the preliminary analysis of 17 cells each. Scale bar: 15 µm.

Chromatin Hyper-Acetylation is Compatible with CENP-A Maintenance

Acidic activation domains present in VP16 and other trans-activators interact with a variety of transcription factors and histone acetyl-transferases (Blair et al., 1994; Gerritsen et al., 1997; Hall and Struhl, 2002; Schmitz et al., 1995). Consistently, when compared to HACs targeted by tetR-EYFP, tethering of VP16 resulted in substantial acetylation of the underlying chromatin, as determined by immunofluorescent staining for H3K9ac (Fig. 22A). To determine the impact of VP16-tethering on HAC centromeric transcription, I performed real-time RT-PCR analysis of alphoid^{tetO} transcript levels two days following transient transfection of constructs expressing either tetR-EYFP or tetR-EYFP-VP16. While expression of tetR-EYFP resulted in a mild (2.5-fold) increase of the low levels of alphoid^{tetO} transcripts detected in untransfected cells, expression of the VP16 construct dramatically raised (150-fold) the corresponding transcript copy numbers (Fig. 22C). In contrast, transcription at the endogenous chromosome 21 centromere was essentially unaffected in the presence of either construct.

In light of the pronounced effect of the VP16 trans-activation domain on both HAC centromere chromatin acetylation and local transcription, I wanted to distinguish the relative contribution of either process towards displacement of CENP-A. The activation domain of the NF-κB subunit p65 (RelA) is similar to that of VP16 (Uesugi et al., 1997), and like VP16, p65 interacts with general transcription factors and a histone acetyl-transferase activity (Perkins et al., 1997). It was recently shown by Andrew Belmont and co-workers that direct tethering of the p65 C-terminal activation domain (aa520-550) or that of VP16 to an integrated lac operator array resulted in a comparable extent of unfolding of the underlying chromatin (Carpenter et al., 2005). In contrast to the strong transcriptional activation mediated by VP16 in reporter assays, however, p65(520-550) caused only a mild increase in transcriptional activity (Carpenter et al., 2005). I therefore went on to perform

transfection experiments using a construct containing the C-terminal trans-activation domain of p65 fused to the C-terminus of tetR-EYFP (tetR-EYFP-p65) *[This construct was generated by J.J.].*

Tethering of tetR-EYFP-p65 to the HAC centromere resulted in a moderate (10-fold) up-regulation of alphoid^{tetO} transcript copy numbers that was significantly lower than the induction observed when targeting tetR-EYFP-VP16 (Fig. 22C). In contrast, immunofluorescence analysis of HACs targeted by tetR-EYFP-p65 revealed a substantial increase of H3K9 acetylation levels at the HAC centromere. This increase was comparable to the effect observed with the VP16 construct (Fig. 22A' and A"). Staining for H4K8ac and H4K16ac at the HAC was comparable in cells expressing the VP16 and p65 constructs and did not display an appreciable increase over tetR-EYFP (data not shown). The large-scale unfolding of the HAC seen in some of the tetR-EYFP-VP16 expressing cells was observed only infrequently at HACs targeted by tetR-EYFP-p65. This contrasts with the observations reported at non-centromeric lac operator arrays (Carpenter et al., 2005). Importantly however, despite fully overlapping the staining for CENP-A at the HAC kinetochore, the strong acetylation facilitated by tetR-EYFP-p65 did not appear to perturb local CENP-A levels (Fig. 22A). In fact, immunofluorescence quantification of the HAC-associated CENP-A signals two days after transfection demonstrated that, in striking contrast to tetR-EYFP-VP16, levels of CENP-A were indistinguishable from HACs targeted by tetR-EYFP (Fig. 22B) *[Quantification of CENP-A levels following transfection of the tetR fusion constructs was performed by JJ]*.

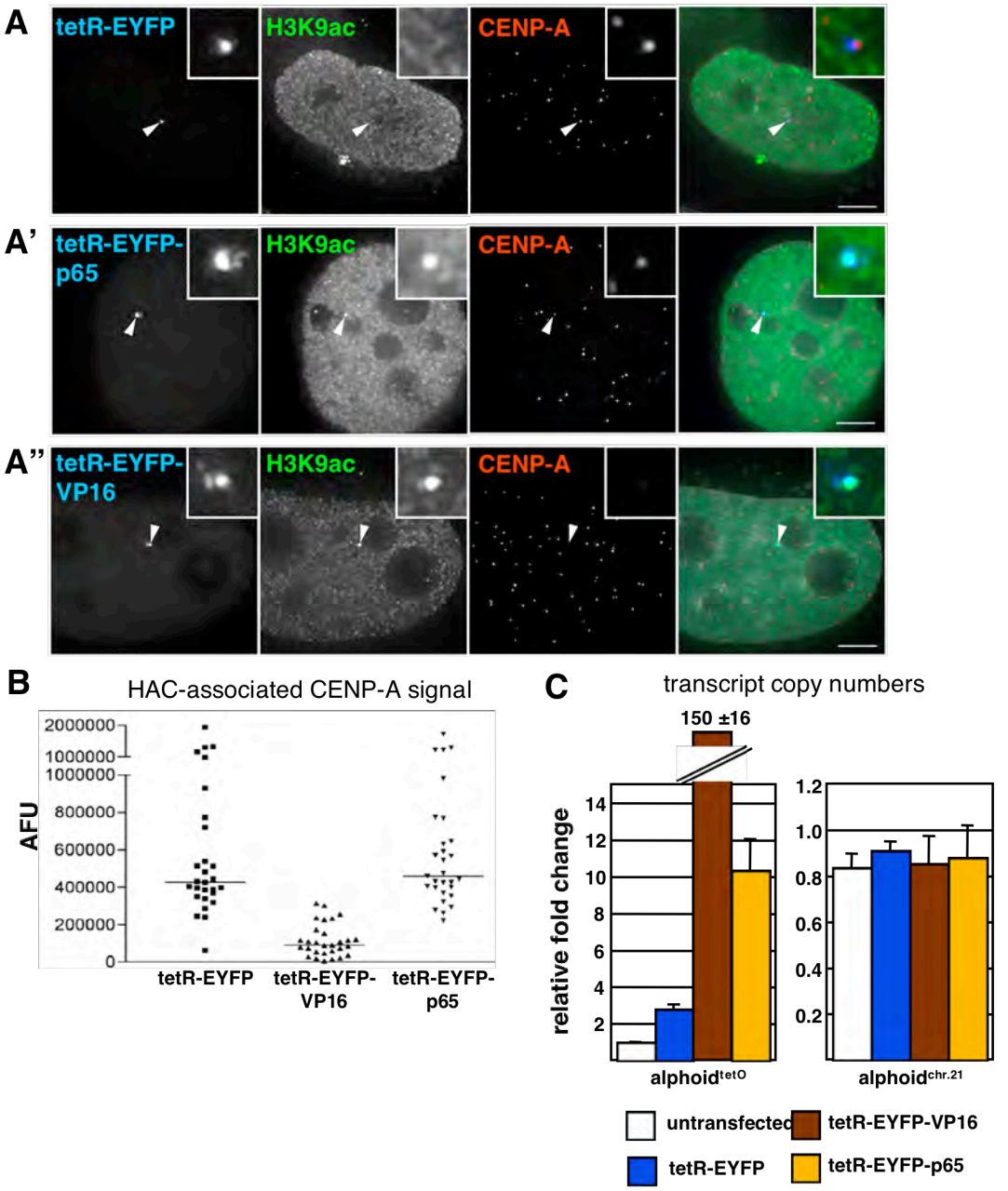


Figure 22| Tethering of VP16 and p65 affect the HAC chromatin in different ways. A) Immunofluorescence analysis of 1C7 cells transiently expressing tetR-EYFP (A), tetR-EYFP-p65 (A') or tetR-EYFP-VP16 (A") 48h after transfection using antibodies specific for H3K9ac (green) and CENP-A (red). Arrowheads depict the HAC. Scale bar: 5 µm. *Figure legend continued overleaf.*

Figure 22 continued.

B) Fluorescence signal quantification of HAC-associated CENP-A staining 48h after transfection as in (A). Arbitrary fluorescence units (AFU) are plotted. *[Quantification of CENP-A levels following transfection was performed by JJ].* **C)** Real-time RT-PCR analysis of alphoid^{tetO} and chromosome 21 (alphoid$^{chr.21}$) transcript copy numbers in untransfected 1C7 cells and 48h after transfection with the indicated constructs. Transcript copy numbers were normalized to β-actin levels and are plotted relative to alphoid^{tetO} levels in untransfected cells, the value of which was set to 1. Data represents the mean and S.E.M. of three independent experiments.

Together, these data demonstrate that hyper-acetylation of centromeric chromatin is not sufficient to cause displacement of CENP-A from the HAC kinetochore. Loss of CENP-A in the course of targeting tetR-EYFP-VP16 is therefore likely facilitated by a mechanism independent of VP16-mediated histone acetylation.

Tethering of VP16 Mediates Specific Displacement of CENP-A from the HAC Centromere

[Some data presented in this section is of a preliminary nature]

To characterize in more detail the impact of tethering VP16 into the HAC chromatin, I decided to employ ChIP analysis of 1C7 cells expressing this fusion construct. As several attempts to generate 1C7 populations stably expressing tetR-EYFP-VP16 were unsuccessful, I performed transient transfections followed by stringent selection for puromycin resistance conferred through an internal ribosome entry site. ChIP analysis 48 hours after transfection revealed a prominent reduction of centromere-associated CENP-A levels at HACs targeted by the VP16 fusion construct compared to untransfected 1C7 cells (Fig. 23A). Loss of CENP-A depended on binding of the fusion construct to the alphoid^{tetO} array. In contrast, levels of CENP-A associated with the endogenous chromosome 21 centromere were not affected by tetR-EYFP-VP16 expression. Importantly, tethering of VP16 specifically displaced CENP-A-containing nucleosomes, as alphoid^{tetO}-associated levels of canonical histone H3 were unaltered (Fig. 23A).

I performed parallel ChIP experiments using an antibody specific for RNA polymerase II phosphorylated on serine 2 of its C-terminal domain, which constitutes the cellular population of polymerases engaged in active elongation. In untransfected cells, levels of this specific polymerase state were below the detection limit of the ChIP assay at either the synthetic

alphoid^{tetO} or the endogenous chromosome 21 centromere (Fig. 23B). Elongating RNA polymerase was readily detected at the actively transcribed BSr marker on the HAC vector backbone, the endogenous poly-A binding protein gene (*PABPC1*) +10 kb locus (Vakoc et al., 2006) as well as the protein phosphatase A (*PP1A*) +3 kb locus. Expression of tetR-EYFP-VP16, however, drastically increased the levels of elongating polymerase at the HAC centromere (Fig. 23B). Polymerase occupancy also further increased at the BSr locus on the HAC vector backbone. As with loss of CENP-A, the increase in numbers of active polymerase molecules at the alphoid^{tetO} array strictly depended on binding of the VP16 fusion construct (Fig. 23B). Importantly, polymerase elongation activity at the HAC centromere following tethering of tetR-EYFP-VP16 was comparable to that present at the *PABPC1* locus and may therefore be considered to remain within physiological levels.

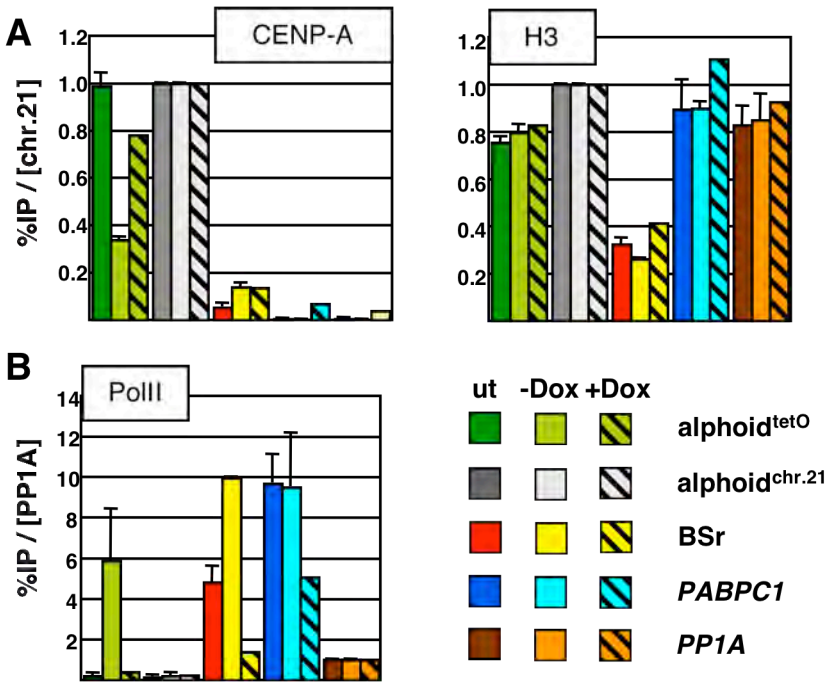


Figure 23| TetR-EYFP-VP16 specifically displaces CENP-A from the HAC centromere. A, B) ChIP analysis of untransfected (ut) 1C7 cells or 48h following transient expression of tetR-EYFP-VP16 in the presence (+Dox) or absence (-Dox) of doxycycline. Antibodies used were specific for CENP-A and total histone H3 (A), or RNA polymerase II phosphorylated on CTD serine 2 (B). Centromeres of the HAC (alphoid^{tetO}) and chromosome 21 (alphoid$^{chr.21}$) were assessed, together with the BSr maker, endogenous *PABPC1* or *PP1A* genes 10 kb or 3 kb downstream of the transcription start site, respectively. Background (IgG)-subtracted percentages of input values were normalized to alphoid$^{chr.21}$ (A) or *PP1A* (B) values, which showed the least variation between experimental conditions. Mean and S.E.M. of two or more independent experiments are plotted for untransfected and -Dox conditions. +Dox represents preliminary data of a single ChIP experiment.

M.3 Discussion

In this section, I initially demonstrated that, following integration into a host chromosome, a non-centromeric alphoid^{tetO} array may be highly euchromatic with respect to its associated chromatin state. This characteristic was confirmed at several independent alphoid^{tetO} integration sites that do not recruit CENP-A (Megumi Nakano, unpublished data), and is noteworthy as ectopically integrated α21-I-derived arrays had previously been reported to assemble pronounced levels of H3K9me3 (Okada et al., 2007; Okamoto et al., 2007). The fate of alphoid input DNA was suggested to be regulated at least in part by CENP-B, which exerts a dual function, capable of promoting *de novo* CENP-A assembly into HAC centromeres, but also prevention of CENP-A establishment or maintenance at integrated arrays paralleled by local H3K9- and DNA methylation (Okada et al., 2007). However, as both, the alphoid^{tetO} and the α21-I-derived alphoid arrays contain an approximately equal density of functional CENP-B boxes, their distinct propensity with respect to chromatin character at the corresponding integration sites is unlikely to be mediated by CENP-B. Integration of the alphoid^{tetO} vector may be favourable at pre-established euchromatic loci. Alternatively, the alphoid^{tetO} array may exhibit an intrinsic predisposition towards promoting a more “open” chromatin character.

Several studies, including the data presented in section E, have established a negative correlation between excessive heterochromatin character and the formation or maintenance of CENP-A domains (Nakano et al., 2008; Okada et al., 2007). The characterization of the chromatin environment of the alphoid^{tetO} integration site described in this section further indicates that one or several aspects of classical euchromatic character may also be incompatible with local CENP-A upkeep. Indeed, tetR-EYFP-VP16 forms the most potent antagonist of CENP-A maintenance at the HAC centromere tested to date. In an attempt to uncover the nature of this putative

antagonistic relationship, I demonstrate that, somewhat surprisingly, hyper-acetylation of centromeric chromatin is compatible with maintenance of CENP-A levels. It should be noted, however, that longer-term effects of nucleosome acetylation on kinetochore structure and function have not been investigated here.

During revision of this manuscript, quench-pulse-chase labelling experiments of SNAP-tagged CENP-A have confirmed that loading of CENP-A is virtually completely abolished within two days of targeting by tetR-EYFP-VP16 (manuscript in preparation). However, the rapid displacement of CENP-A from the HAC centromere in the presence of tethered VP16 further suggests an active eviction of CENP-A-containing nucleosomes. For example, occasional VP16-tethered HACs displayed levels of CENP-A staining close to the nuclear background within 24 hours (Julia Jakubsche, Diploma Thesis, 2009), a time period that is too short for passive dilution of existing CENP-A as consequence of inhibited loading. I was furthermore able to provide direct evidence supporting the hypothesis that VP16 additionally mediates active CENP-A eviction (manuscript in preparation). Acidic activation domains, including that of VP16, can interact with the yeast SWI/SNF chromatin remodelling complex (Boyer et al., 2000; Neely et al., 2002), although a direct interaction of human SWI/SNF with VP16 was not detected (Boyer et al., 2000). However, tethering of VP16 to an integrated lac operator array resulted in the localized enrichment of components of the SWI/SNF complex in mammalian cells (Memedula and Belmont, 2003). *In vitro* data using reconstituted oligonucleosomal reporter arrays recently suggested that Gal4-VP16-mediated recruitment of SWI/SNF may induce the destabilization of an adjacent nucleosome (Dechassa et al., 2010), although to my knowledge a displacement of nucleosomes *in vivo*, outside of a promoter context, has not been reported. I did not observe any loss of nucleosomes containing canonical histone H3 from the VP16-tethered HAC

centromere, demonstrating that in this experimental system, VP16 is unable to directly modulate nucleosome density.

An obvious speculation arising from the data presented above is that specific displacement of CENP-A following tethering of tetR-EYFP-VP16 may be mediated by the considerably increased local activity of RNA polymerase II. I am currently performing experiments investigating this hypothesis, by expressing tetR-EYFP-VP16 followed by treatment of cells with actinomycin D. Studies in yeast suggest that two distinct mechanisms can facilitate RNA polymerase elongation through chromatin *in vivo.* These involve either transient histone acetylation without loss of nucleosomes, or destabilization of nucleosome / DNA contacts (Kristjuhan and Svejstrup, 2004). The transcription elongation factor FACT has been implicated in destabilizing and chaperoning H2A / H2B dimers in the course of elongation (Belotserkovskaya et al., 2003; Orphanides et al., 1998; Orphanides et al., 1999), while the histone chaperone Spt6 associates with H3 / H4 dimers (Bortvin and Winston, 1996). In current models, FACT and Spt6 act in concert to promote the disassembly of nucleosomes in front of elongating RNA polymerases followed by restoration of nucleosomes in their wake to restore proper chromatin architecture (Selth et al., 2010; Workman, 2006) (Fig. 24). During this process, “old” nucleosome components can be replaced with “new” ones from the nucleoplasmic pool, a mechanism thought to underlie the deposition of the replication-independent histone variant H3.3 (Henikoff and Ahmad, 2005). Elevating transcriptional activity at centromeres may therefore result in active eviction specifically of CENP-A if either the pool of available CENP-A is limited relative to that of H3.3, or if the association of CENP-A with histone chaperones such as HJURP does not facilitate transcription-coupled nucleosome assembly. In this case, the expected outcome would be the replacement of CENP-A- with H3.3-nucleosomes, or a “nucleosome gap” due to instability of the resulting H2A / H2B hetero-tetramer, respectively (Fig. 24).

In light of the near-physiological level of RNA polymerase activity observe following tethering of VP16 to the HAC centromere, transcription-coupled eviction of CENP-A may be biologically relevant. With a vast amount of the genome being actively transcribed, this transcription may act in part as a “proof-reading” mechanism to displace CENP-A nucleosomes erroneously deposited outside a transcriptionally more tightly controlled centromeric region. In keeping with this, it is interesting to point out that ectopic expression of CENP-A predominantly causes its targeting to euchromatic regions, possibly reflecting saturation of a putative transcription-coupled eviction process.

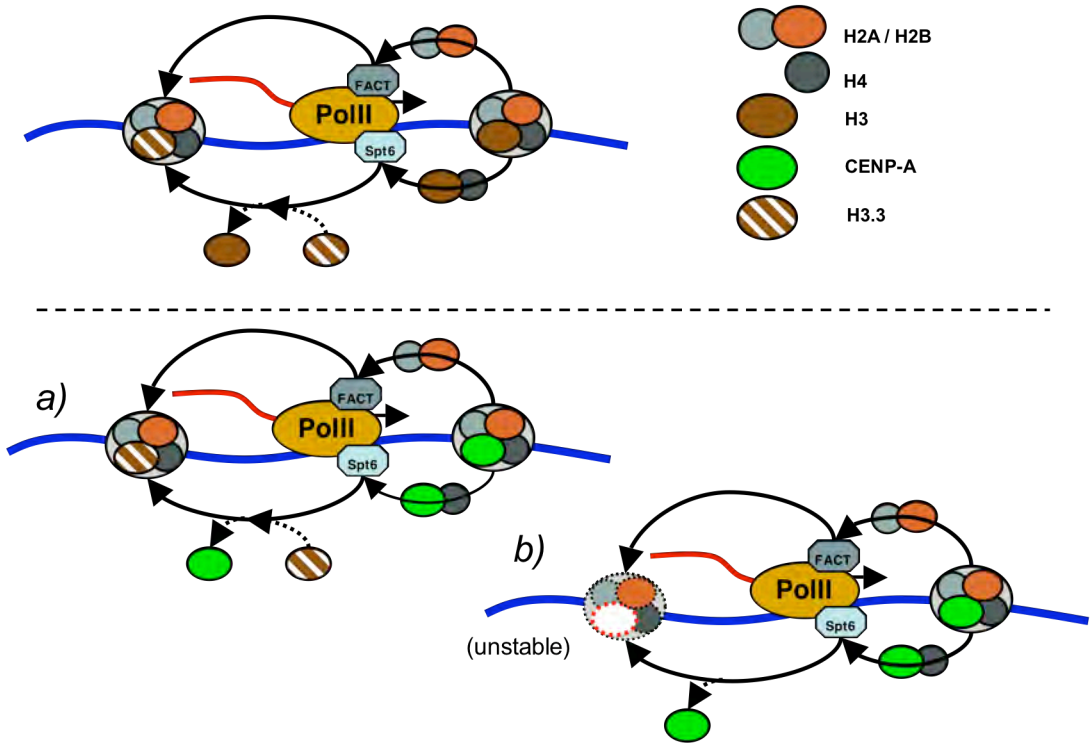


Figure 24| Active nucleosome eviction model for transcribing RNA polymerase. Top: At actively transcribed genes, FACT and Spt6 are implicated in the destabilization and re-assembly of nucleosomes during transcriptional elongation. During this process, H3.1 may be replaced by nucleoplasmic H3.3. **Bottom:** Excessive RNA polymerase elongation through CENP-A chromatin would be predicted to result in the replacement of CENP-A with H3.3 (a) or the loss of CENP-A and subsequent loss of a nucleosome (b) as described in the main text.

Conclusion and Outlook

The early recognition of CENP-A as a centromere-specific histone variant was a priming moment for the realization that the local chromatin may play a pivotal role in defining the centromere locus and contributing to kinetochore architecture and function. Subsequent characterization of centromere-associated histone modifications led to the speculation that a distinct chromatin state exists that, together with the presence of CENP-A, distinguishes centromeres from any other genomic locus. However, the nature of endogenous centromeres, notably their inaccessibility to direct manipulation as well as their intimate connection with cellular viability, essentially hampered any further research into the potential functional relationship between local chromatin state, CENP-A maintenance and kinetochore structure. These caveats have in large part prohibited the establishment of evidence-supported models with respect to this possible inter-relationship.

Earlier scientific break-through in form of the generation of the alphoid^{tetO} human artificial chromosome has enabled me to employ this system in order to characterize the relationship of chromatin state and centromere function on several levels, ranging from the descriptive investigation of “active” and “inactive” alpha satellite arrays to the targeted manipulation of the chromatin underlying a single functional kinetochore in living cells. The data presented above emphasize an obvious inter-dependency of local chromatin state and upkeep of kinetochore structure and function. Surprisingly, several independent experimental approaches combined place a spotlight on a possible requirement for local non-coding transcription and the preservation of the CENP-A chromatin domain. It would appear that a tightly regulated balance of transcriptional activity at centromeres may be central to the maintenance of CENP-A and other aspects of local chromatin character. Tipping this balance too far in either

direction, that is reducing transcriptional fidelity or augmenting occupancy of actively transcribing RNA polymerase molecules, appears detrimental for ongoing centromere identity.

Future challenges obviously include to prove, on the molecular level, this hypothesis inferred from correlative observations, as well as to establish and to dissect the relevant contribution of individual aspects of non-coding transcription towards centromere maintenance. The latter may depend on different, not necessarily mutually exclusive, processes. At present, a role for alphoid transcripts *per se* cannot be distinguished from RNA-independent contribution of transcription complexes, which in turn may involve the recruitment of histone-modifying activities to shape and maintain a certain chromatin environment, or that of chromatin remodelling complexes to actively promote the positioning and composition of local nucleosomes.

Several related questions will be more feasible to address in the immediate future and include a thorough characterization of endogenous alpha satellite transcripts with respect to their length and origin as well as the temporal regulation of their expression. Do these transcripts indeed span the majority of CENP-A-associated sequence, or do they map to only a fraction of alpha satellite in a centromere? Do centromeric transcripts associate with nuclear proteins and if so, which functional role can be inferred? Novel technological advances including super-resolution and single molecule microscopy may provide answers to these and other questions. New generation human artificial chromosomes containing for example alphoid-associated MS2 sequences that may facilitate the purification of corresponding transcripts would in addition be of extremely high value for the biochemical characterization of centromeric non-coding transcripts.

The present-generation alphoid^{tetO} HAC has proven to be a formidable system for the study of centromeric chromatin. It is tempting to further dissect

the local chromatin state, to investigate for example the role of hyper-methylated H3K36 in the maintenance of centromeric structure. Can the well-established molecular pathway linking this mark to maintaining a critical chromatin architecture throughout the body of transcribed genes be transferred to centromeres? What, if any, are the consequences of abolishing the putative H3K36 / HDAC loop at centromeres? On the other hand, converse experimental approaches are now becoming more feasible: can the chromatin environment at the “inactive” alphoid^{tetO} integration be manipulated towards increasing the levels of local CENP-A, or even towards the *de novo* formation of an ectopic kinetochore structure? Based on above findings, an obvious experiment would be to reduce, but not abolish, the transcriptional activity at the integration site and to monitor recruitment of CENP-A.

In conclusion, many aspects of chromatin and its functional relationship with centromeres remain to be investigated. In this respect, the understanding of both, *de novo* establishment as well as maintenance of centromere character have been significantly promoted by the use of human artificial chromosomes and derivatives. Human artificial chromosomes as models for centromeric chromatin character will likely continue to form an integral aspect of varied studies in the future.

References

Agger, K., Cloos, P.A., Christensen, J., Pasini, D., Rose, S., Rappsilber, J., Issaeva, I., Canaani, E., Salcini, A.E., and Helin, K. (2007). UTX and JMJD3 are histone H3K27 demethylases involved in HOX gene regulation and development. Nature *449*, 731-734.

Ahmad, K., and Henikoff, S. (2002). The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly. Mol Cell *9*, 1191-1200.

Albertson, D.G., and Thomson, J.N. (1982). The kinetochores of Caenorhabditis elegans. Chromosoma *86*, 409-428.

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

Alonso, A., Fritz, B., Hasson, D., Abrusan, G., Cheung, F., Yoda, K., Radlwimmer, B., Ladurner, A.G., and Warburton, P.E. (2007). Co-localization of CENP-C and CENP-H to discontinuous domains of CENP-A chromatin at human neocentromeres. Genome Biol *8*, R148.

Alonso, A., Hasson, D., Cheung, F., and Warburton, P.E. (2010). A paucity of heterochromatin at functional human neocentromeres. Epigenetics Chromatin *3*, 6.

Alonso, A., Mahmood, R., Li, S., Cheung, F., Yoda, K., and Warburton, P.E. (2003). Genomic microarray analysis reveals distinct locations for the CENP-A binding domains in three human chromosome 13q32 neocentromeres. Hum Mol Genet *12*, 2711-2721.

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

Amaral, P.P., and Mattick, J.S. (2008). Noncoding RNA in development. Mamm Genome *19*, 454-492.

Ando, S., Yang, H., Nozaki, N., Okazaki, T., and Yoda, K. (2002). CENP-A, -B, and -C chromatin complex that contains the I-type alpha-satellite array constitutes the prekinetochore in HeLa cells. Mol Cell Biol *22*, 2229-2241.

Aravin, A.A., Hannon, G.J., and Brennecke, J. (2007). The Piwi-piRNA pathway provides an adaptive defense in the transposon arms race. Science *318*, 761-764.

Barski, A., Cuddapah, S., Cui, K., Roh, T.Y., Schones, D.E., Wang, Z., Wei, G., Chepelev, I., and Zhao, K. (2007). High-resolution profiling of histone methylations in the human genome. Cell *129*, 823-837.

Bassal, S., and El-Osta, A. (2005). DNA damage detection and repair, and the involvement of epigenetic states. Hum Mutat *25*, 101-109.

Basu, J., and Willard, H.F. (2005). Artificial and engineered chromosomes: non-integrating vectors for gene therapy. Trends Mol Med *11*, 251-258.

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

Berger, S.L. (2007). The complex language of chromatin regulation during transcription. Nature *447*, 407-412.

Bird, A. (2002). DNA methylation patterns and epigenetic memory. Genes Dev *16*, 6-21.

Bird, A. (2007). Perceptions of epigenetics. Nature *447*, 396-398.

Birney, E., Stamatoyannopoulos, J.A., Dutta, A., Guigo, R., Gingeras, T.R., Margulies, E.H., Weng, Z., Snyder, M., Dermitzakis, E.T., Thurman, R.E.*, et al.* (2007). Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature *447*, 799-816.

Black, B.E., and Bassett, E.A. (2008). The histone variant CENP-A and centromere specification. Curr Opin Cell Biol *20*, 91-100.

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

Black, B.E., Jansen, L.E., 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.

Blair, W.S., Bogerd, H., and Cullen, B.R. (1994). Genetic analysis indicates that the human foamy virus Bel-1 protein contains a transcription activation domain of the acidic class. J Virol *68*, 3803-3808.

Blasco, M.A. (2007). The epigenetic regulation of mammalian telomeres. Nat Rev Genet *8*, 299-309.

Blobel, G.A., Kadauke, S., Wang, E., Lau, A.W., Zuber, J., Chou, M.M., and Vakoc, C.R. (2009). A reconfigured pattern of MLL occupancy within mitotic chromatin promotes rapid transcriptional reactivation following mitotic exit. Mol Cell *36*, 970-983.

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

Bortvin, A., and Winston, F. (1996). Evidence that Spt6p controls chromatin structure by a direct interaction with histones. Science *272*, 1473-1476.

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

Boyer, L.A., Logie, C., Bonte, E., Becker, P.B., Wade, P.A., Wolffe, A.P., Wu, C., Imbalzano, A.N., and Peterson, C.L. (2000). Functional delineation of three groups of the ATP-dependent family of chromatin remodeling enzymes. J Biol Chem *275*, 18864-18870.

Brinkley, B.R., and Stubblefield, E. (1966). The fine structure of the kinetochore of a mammalian cell in vitro. Chromosoma *19*, 28-43.

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.

Buratowski, S. (2009). Progression through the RNA polymerase II CTD cycle. Mol Cell *36*, 541-546.

Cam, H.P., Sugiyama, T., Chen, E.S., Chen, X., FitzGerald, P.C., and Grewal, S.I. (2005). Comprehensive analysis of heterochromatin- and RNAi-mediated epigenetic control of the fission yeast genome. Nat Genet *37*, 809-819.

Cardinale, S., Bergmann, J.H., Kelly, D., Nakano, M., Valdivia, M.M., Kimura, H., Masumoto, H., Larionov, V., and Earnshaw, W.C. (2009). Hierarchical inactivation of a synthetic human kinetochore by a chromatin modifier. Mol Biol Cell *20*, 4194-4204.

Carpenter, A.E., Memedula, S., Plutz, M.J., and Belmont, A.S. (2005). Common effects of acidic activators on large-scale chromatin structure and transcription. Mol Cell Biol *25*, 958-968.

Carroll, C.W., Silva, M.C., Godek, K.M., Jansen, L.E., 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., and Straight, A.F. (2006). Centromere formation: from epigenetics to self-assembly. Trends Cell Biol *16*, 70-78.

Carrozza, M.J., Li, B., Florens, L., Suganuma, T., Swanson, S.K., Lee, K.K., Shia, W.J., Anderson, S., Yates, J., Washburn, M.P.*, et al.* (2005). Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. Cell *123*, 581-592.

Carthew, R.W., and Sontheimer, E.J. (2009). Origins and Mechanisms of miRNAs and siRNAs. Cell *136*, 642-655.

Cawley, S., Bekiranov, S., Ng, H.H., Kapranov, P., Sekinger, E.A., Kampa, D., Piccolboni, A., Sementchenko, V., Cheng, J., Williams, A.J.*, et al.* (2004). Unbiased mapping of transcription factor binding sites along human chromosomes 21 and 22 points to widespread regulation of noncoding RNAs. Cell *116*, 499-509.

Chan, G.K., Liu, S.T., and Yen, T.J. (2005). Kinetochore structure and function. Trends Cell Biol *15*, 589-598.

Cheeseman, I.M., Chappie, J.S., Wilson-Kubalek, E.M., and Desai, A. (2006). The conserved KMN network constitutes the core microtubule-binding site of the kinetochore. Cell *127*, 983-997.

Cheeseman, I.M., and Desai, A. (2008). Molecular architecture of the kinetochore-microtubule interface. Nat Rev Mol Cell Biol *9*, 33-46.

Cheeseman, I.M., Hori, T., Fukagawa, T., and Desai, A. (2008). KNL1 and the CENP-H/I/K complex coordinately direct kinetochore assembly in vertebrates. Mol Biol Cell *19*, 587-594.

Cheeseman, I.M., Niessen, S., Anderson, S., Hyndman, F., Yates, J.R., 3rd, Oegema, K., and Desai, A. (2004). A conserved protein network controls assembly of the outer kinetochore and its ability to sustain tension. Genes Dev *18*, 2255-2268.

Choo, K.H., Vissel, B., Nagy, A., Earle, E., and Kalitsis, P. (1991). A survey of the genomic distribution of alpha satellite DNA on all the human chromosomes, and derivation of a new consensus sequence. Nucleic Acids Res *19*, 1179-1182.

Chueh, A.C., Wong, L.H., Wong, N., and Choo, K.H. (2005). Variable and hierarchical size distribution of L1-retroelement-enriched CENP-A clusters within a functional human neocentromere. Hum Mol Genet *14*, 85-93.

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

Dalal, Y., Wang, H., Lindsay, S., and Henikoff, S. (2007). Tetrameric structure of centromeric nucleosomes in interphase Drosophila cells. PLoS Biol *5*, e218.

Daujat, S., Weiss, T., Mohn, F., Lange, U.C., Ziegler-Birling, C., Zeissler, U., Lappe, M., Schubeler, D., Torres-Padilla, M.E., and Schneider, R. (2009). H3K64 trimethylation marks heterochromatin and is dynamically remodeled during developmental reprogramming. Nat Struct Mol Biol *16*, 777-781.

Dechassa, M.L., Sabri, A., Pondugula, S., Kassabov, S.R., Chatterjee, N., Kladde, M.P., and Bartholomew, B. (2010). SWI/SNF has intrinsic nucleosome disassembly activity that is dependent on adjacent nucleosomes. Mol Cell *38*, 590-602.

Desai, A., and Mitchison, T.J. (1997). Microtubule polymerization dynamics. Annu Rev Cell Dev Biol *13*, 83-117.

Dinger, M.E., Amaral, P.P., Mercer, T.R., Pang, K.C., Bruce, S.J., Gardiner, B.B., Askarian-Amiri, M.E., Ru, K., Solda, G., Simons, C.*, et al.* (2008). Long noncoding RNAs in mouse embryonic stem cell pluripotency and differentiation. Genome Res *18*, 1433-1445.

Dover, J., Schneider, J., Tawiah-Boateng, M.A., Wood, A., Dean, K., Johnston, M., and Shilatifard, A. (2002). Methylation of histone H3 by COMPASS requires ubiquitination of histone H2B by Rad6. J Biol Chem *277*, 28368-28371.

Dunleavy, E., Pidoux, A., and Allshire, R. (2005). Centromeric chromatin makes its mark. Trends Biochem Sci *30*, 172-175.

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.

Earnshaw, W.C., Ratrie, H., 3rd, and Stetten, G. (1989). Visualization of centromere proteins CENP-B and CENP-C on a stable dicentric chromosome in cytological spreads. Chromosoma *98*, 1-12.

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.

Earnshaw, W.C., Sullivan, K.F., Machlin, P.S., Cooke, C.A., Kaiser, D.A., Pollard, T.D., Rothfield, N.F., and Cleveland, D.W. (1987). Molecular cloning of cDNA for CENP-B, the major human centromere autoantigen. J Cell Biol *104*, 817-829.

Ebersole, T., Okamoto, Y., Noskov, V.N., Kouprina, N., Kim, J.H., Leem, S.H., Barrett, J.C., Masumoto, H., and Larionov, V. (2005). Rapid generation of long synthetic tandem repeats and its application for analysis in human artificial chromosome formation. Nucleic Acids Res *33*, e130.

Ebersole, T.A., Ross, A., Clark, E., McGill, N., Schindelhauer, D., Cooke, H., and Grimes, B. (2000). Mammalian artificial chromosome formation from circular alphoid input DNA does not require telomere repeats. Hum Mol Genet *9*, 1623-1631.

Efroni, S., Duttagupta, R., Cheng, J., Dehghani, H., Hoeppner, D.J., Dash, C., Bazett-Jones, D.P., Le Grice, S., McKay, R.D., Buetow, K.H.*, et al.* (2008). Global transcription in pluripotent embryonic stem cells. Cell Stem Cell *2*, 437-447.

Egloff, S., and Murphy, S. (2008). Cracking the RNA polymerase II CTD code. Trends Genet *24*, 280-288.

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.

Farazi, T.A., Juranek, S.A., and Tuschl, T. (2008). The growing catalog of small RNAs and their association with distinct Argonaute/Piwi family members. Development *135*, 1201-1214.

Folco, H.D., Pidoux, A.L., Urano, T., and Allshire, R.C. (2008). Heterochromatin and RNAi are required to establish CENP-A chromatin at centromeres. Science *319*, 94-97.

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

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

Frescas, D., Guardavaccaro, D., Kuchay, S.M., Kato, H., Poleshko, A., Basrur, V., Elenitoba-Johnson, K.S., Katz, R.A., and Pagano, M. (2008). KDM2A represses transcription of centromeric satellite repeats and maintains the heterochromatic state. Cell Cycle *7*, 3539-3547.

Freundlieb, S., Schirra-Muller, C., and Bujard, H. (1999). A tetracycline controlled activation/repression system with increased potential for gene transfer into mammalian cells. J Gene Med *1*, 4-12.

Friedman, J.R., Fredericks, W.J., Jensen, D.E., Speicher, D.W., Huang, X.P., Neilson, E.G., and Rauscher, F.J., 3rd (1996). KAP-1, a novel corepressor for the highly conserved KRAB repression domain. Genes Dev *10*, 2067-2078.

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.

Fukagawa, T., Mikami, Y., Nishihashi, A., Regnier, V., Haraguchi, T., Hiraoka, Y., Sugata, N., Todokoro, K., Brown, W., and Ikemura, T. (2001). CENP-H, a constitutive centromere component, is required for centromere targeting of CENP-C in vertebrate cells. Embo J *20*, 4603-4617.

Fukagawa, T., Nogami, M., Yoshikawa, M., Ikeno, M., Okazaki, T., Takami, Y., Nakayama, T., and Oshimura, M. (2004). Dicer is essential for formation of the heterochromatin structure in vertebrate cells. Nat Cell Biol *6*, 784-791.

Fukagawa, T., Pendon, C., Morris, J., and Brown, W. (1999). CENP-C is necessary but not sufficient to induce formation of a functional centromere. Embo J *18*, 4196-4209.

Furuyama, S., and Biggins, S. (2007). Centromere identity is specified by a single centromeric nucleosome in budding yeast. Proc Natl Acad Sci U S A *104*, 14706-14711.

Gerritsen, M.E., Williams, A.J., Neish, A.S., Moore, S., Shi, Y., and Collins, T. (1997). CREB-binding protein/p300 are transcriptional coactivators of p65. Proc Natl Acad Sci U S A *94*, 2927-2932.

Goshima, G., Kiyomitsu, T., Yoda, K., and Yanagida, M. (2003). Human centromere chromatin protein hMis12, essential for equal segregation, is independent of CENP-A loading pathway. J Cell Biol *160*, 25-39.

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

Gossen, M., and Bujard, H. (1992). Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. Proc Natl Acad Sci U S A *89*, 5547-5551.

Grewal, S.I., and Jia, S. (2007). Heterochromatin revisited. Nat Rev Genet *8*, 35-46.

Grimes, B., and Cooke, H. (1998). Engineering mammalian chromosomes. Hum Mol Genet *7*, 1635-1640.

Grimes, B.R., Rhoades, A.A., and Willard, H.F. (2002). Alpha-satellite DNA and vector composition influence rates of human artificial chromosome formation. Mol Ther *5*, 798-805.

Guenther, M.G., Levine, S.S., Boyer, L.A., Jaenisch, R., and Young, R.A. (2007). A chromatin landmark and transcription initiation at most promoters in human cells. Cell *130*, 77-88.

Guttman, M., Amit, I., Garber, M., French, C., Lin, M.F., Feldser, D., Huarte, M., Zuk, O., Carey, B.W., Cassady, J.P.*, et al.* (2009). Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. Nature *458*, 223-227.

Haaf, T., and Ward, D.C. (1994). Structural analysis of alpha-satellite DNA and centromere proteins using extended chromatin and chromosomes. Hum Mol Genet *3*, 697-709.

Hakimi, M.A., Bochar, D.A., Chenoweth, J., Lane, W.S., Mandel, G., and Shiekhattar, R. (2002). A core-BRAF35 complex containing histone

deacetylase mediates repression of neuronal-specific genes. Proc Natl Acad Sci U S A *99*, 7420-7425.

Hall, D.B., and Struhl, K. (2002). The VP16 activation domain interacts with multiple transcriptional components as determined by protein-protein cross-linking in vivo. J Biol Chem *277*, 46043-46050.

Harrington, J.J., Van Bokkelen, G., Mays, R.W., Gustashaw, K., and Willard, H.F. (1997). Formation of de novo centromeres and construction of first-generation human artificial microchromosomes. Nat Genet *15*, 345-355.

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.

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., and Ahmad, K. (2005). Assembly of variant histones into chromatin. Annu Rev Cell Dev Biol *21*, 133-153.

Henikoff, S., Ahmad, K., and Malik, H.S. (2001). The centromere paradox: stable inheritance with rapidly evolving DNA. Science *293*, 1098-1102.

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.

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.

Hori, T., Amano, M., Suzuki, A., Backer, C.B., Welburn, J.P., Dong, Y., McEwen, B.F., Shang, W.H., Suzuki, E., Okawa, K.*, et al.* (2008a). CCAN makes multiple contacts with centromeric DNA to provide distinct pathways to the outer kinetochore. Cell *135*, 1039-1052.

Hori, T., Okada, M., Maenaka, K., and Fukagawa, T. (2008b). CENP-O class proteins form a stable complex and are required for proper kinetochore function. Mol Biol Cell *19*, 843-854.

Huang, Y., Fang, J., Bedford, M.T., Zhang, Y., and Xu, R.M. (2006). Recognition of histone H3 lysine-4 methylation by the double tudor domain of JMJD2A. Science *312*, 748-751.

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.

Hudson, D.F., Marshall, K.M., and Earnshaw, W.C. (2009). Condensin: Architect of mitotic chromosomes. Chromosome Res *17*, 131-144.

Humphrey, G.W., Wang, Y., Russanova, V.R., Hirai, T., Qin, J., Nakatani, Y., and Howard, B.H. (2001). Stable histone deacetylase complexes distinguished by the presence of SANT domain proteins CoREST/kiaa0071 and Mta-L1. J Biol Chem *276*, 6817-6824.

Ikeno, M., Grimes, B., Okazaki, T., Nakano, M., Saitoh, K., Hoshino, H., McGill, N.I., Cooke, H., and Masumoto, H. (1998). Construction of YAC-based mammalian artificial chromosomes. Nat Biotechnol *16*, 431-439.

Ikeno, M., Masumoto, H., and Okazaki, T. (1994). Distribution of CENP-B boxes reflected in CREST centromere antigenic sites on long-range alpha-satellite DNA arrays of human chromosome 21. Hum Mol Genet *3*, 1245-1257.

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

Jansen, L.E., 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.

Jenuwein, T., and Allis, C.D. (2001). Translating the histone code. Science *293*, 1074-1080.

Joshi, A.A., and Struhl, K. (2005). Eaf3 chromodomain interaction with methylated H3-K36 links histone deacetylation to Pol II elongation. Mol Cell *20*, 971-978.

Kagansky, A., Folco, H.D., Almeida, R., Pidoux, A.L., Boukaba, A., Simmer, F., Urano, T., Hamilton, G.L., and Allshire, R.C. (2009). Synthetic heterochromatin bypasses RNAi and centromeric repeats to establish functional centromeres. Science *324*, 1716-1719.

Kalitsis, P., Fowler, K.J., Earle, E., Hill, J., and Choo, K.H. (1998). Targeted disruption of mouse centromere protein C gene leads to mitotic disarray and early embryo death. Proc Natl Acad Sci U S A *95*, 1136-1141.

Kanellopoulou, C., Muljo, S.A., Kung, A.L., Ganesan, S., Drapkin, R., Jenuwein, T., Livingston, D.M., and Rajewsky, K. (2005). Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing. Genes Dev *19*, 489-501.

Keogh, M.C., Kurdistani, S.K., Morris, S.A., Ahn, S.H., Podolny, V., Collins, S.R., Schuldiner, M., Chin, K., Punna, T., Thompson, N.J.*, et al.* (2005). Cotranscriptional set2 methylation of histone H3 lysine 36 recruits a repressive Rpd3 complex. Cell *123*, 593-605.

Keohane, A.M., O'Neill L, P., Belyaev, N.D., Lavender, J.S., and Turner, B.M. (1996). X-Inactivation and histone H4 acetylation in embryonic stem cells. Dev Biol *180*, 618-630.

Kim, J., Guermah, M., McGinty, R.K., Lee, J.S., Tang, Z., Milne, T.A., Shilatifard, A., Muir, T.W., and Roeder, R.G. (2009). RAD6-Mediated transcription-coupled H2B ubiquitylation directly stimulates H3K4 methylation in human cells. Cell *137*, 459-471.

Kim, S.S., Chen, Y.M., O'Leary, E., Witzgall, R., Vidal, M., and Bonventre, J.V. (1996). A novel member of the RING finger family, KRIP-1, associates with the KRAB-A transcriptional repressor domain of zinc finger proteins. Proc Natl Acad Sci U S A *93*, 15299-15304.

Kim, T., and Buratowski, S. (2009). Dimethylation of H3K4 by Set1 recruits the Set3 histone deacetylase complex to 5' transcribed regions. Cell *137*, 259-272.

Kimura, H., Hayashi-Takanaka, Y., Goto, Y., Takizawa, N., and Nozaki, N. (2008). The organization of histone H3 modifications as revealed by a panel of specific monoclonal antibodies. Cell Struct Funct *33*, 61-73.

Kirmizis, A., Santos-Rosa, H., Penkett, C.J., Singer, M.A., Vermeulen, M., Mann, M., Bahler, J., Green, R.D., and Kouzarides, T. (2007). Arginine methylation at histone H3R2 controls deposition of H3K4 trimethylation. Nature *449*, 928-932.

Kline, S.L., Cheeseman, I.M., Hori, T., Fukagawa, T., and Desai, A. (2006). The human Mis12 complex is required for kinetochore assembly and proper chromosome segregation. J Cell Biol *173*, 9-17.

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.

Kouskouti, A., and Talianidis, I. (2005). Histone modifications defining active genes persist after transcriptional and mitotic inactivation. Embo J *24*, 347-357.

Kouzarides, T. (2007). Chromatin modifications and their function. Cell *128*, 693-705.

Kristjuhan, A., and Svejstrup, J.Q. (2004). Evidence for distinct mechanisms facilitating transcript elongation through chromatin in vivo. Embo J *23*, 4243-4252.

Krogan, N.J., Dover, J., Wood, A., Schneider, J., Heidt, J., Boateng, M.A., Dean, K., Ryan, O.W., Golshani, A., Johnston, M.*, et al.* (2003a). The Paf1 complex is required for histone H3 methylation by COMPASS and Dot1p: linking transcriptional elongation to histone methylation. Mol Cell *11*, 721-729.

Krogan, N.J., Kim, M., Tong, A., Golshani, A., Cagney, G., Canadien, V., Richards, D.P., Beattie, B.K., Emili, A., Boone, C.*, et al.* (2003b). Methylation of histone H3 by Set2 in Saccharomyces cerevisiae is linked to transcriptional elongation by RNA polymerase II. Mol Cell Biol *23*, 4207-4218.

Kwon, M.S., Hori, T., Okada, M., and Fukagawa, T. (2007). CENP-C is involved in chromosome segregation, mitotic checkpoint function, and kinetochore assembly. Mol Biol Cell *18*, 2155-2168.

Lachner, M., O'Carroll, D., Rea, S., Mechtler, K., and Jenuwein, T. (2001). Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. Nature *410*, 116-120.

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

Le Douarin, B., Nielsen, A.L., Garnier, J.M., Ichinose, H., Jeanmougin, F., Losson, R., and Chambon, P. (1996). A possible involvement of TIF1 alpha and TIF1 beta in the epigenetic control of transcription by nuclear receptors. Embo J *15*, 6701-6715.

Lechner, M.S., Begg, G.E., Speicher, D.W., and Rauscher, F.J., 3rd (2000). Molecular determinants for targeting heterochromatin protein 1-

mediated gene silencing: direct chromoshadow domain-KAP-1 corepressor interaction is essential. Mol Cell Biol *20*, 6449-6465.

Lee, M.G., Wynder, C., Cooch, N., and Shiekhattar, R. (2005). An essential role for CoREST in nucleosomal histone 3 lysine 4 demethylation. Nature *437*, 432-435.

Lee, T.I., Jenner, R.G., Boyer, L.A., Guenther, M.G., Levine, S.S., Kumar, R.M., Chevalier, B., Johnstone, S.E., Cole, M.F., Isono, K.*, et al.* (2006). Control of developmental regulators by Polycomb in human embryonic stem cells. Cell *125*, 301-313.

Leeb, M., Steffen, P.A., and Wutz, A. (2009). X chromosome inactivation sparked by non-coding RNAs. RNA Biol *6*, 94-99.

Li, B., Carey, M., and Workman, J.L. (2007). The role of chromatin during transcription. Cell *128*, 707-719.

Liu, D., Vader, G., Vromans, M.J., Lampson, M.A., and Lens, S.M. (2009). Sensing chromosome bi-orientation by spatial separation of aurora B kinase from kinetochore substrates. Science *323*, 1350-1353.

Liu, S.T., Rattner, J.B., Jablonski, S.A., and Yen, T.J. (2006). Mapping the assembly pathways that specify formation of the trilaminar kinetochore plates in human cells. J Cell Biol *175*, 41-53.

Lleres, D., James, J., Swift, S., Norman, D.G., and Lamond, A.I. (2009). Quantitative analysis of chromatin compaction in living cells using FLIM-FRET. J Cell Biol *187*, 481-496.

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

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 A resolution. Nature *389*, 251-260.

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.

Maeshima, K., and Eltsov, M. (2008). Packaging the genome: the structure of mitotic chromosomes. J Biochem *143*, 145-153.

Maiato, H., DeLuca, J., Salmon, E.D., and Earnshaw, W.C. (2004). The dynamic kinetochore-microtubule interface. J Cell Sci *117*, 5461-5477.

Marshall, O.J., Marshall, A.T., and Choo, K.H. (2008). Three-dimensional localization of CENP-A suggests a complex higher order structure of centromeric chromatin. J Cell Biol *183*, 1193-1202.

Martens, J.H., O'Sullivan, R.J., Braunschweig, U., Opravil, S., Radolf, M., Steinlein, P., and Jenuwein, T. (2005). The profile of repeat-associated histone lysine methylation states in the mouse epigenome. Embo J *24*, 800-812.

Martin, D.G., Baetz, K., Shi, X., Walter, K.L., MacDonald, V.E., Wlodarski, M.J., Gozani, O., Hieter, P., and Howe, L. (2006). The Yng1p plant homeodomain finger is a methyl-histone binding module that recognizes lysine 4-methylated histone H3. Mol Cell Biol *26*, 7871-7879.

Masumoto, H., Ikeno, M., Nakano, M., Okazaki, T., Grimes, B., Cooke, H., and Suzuki, N. (1998). Assay of centromere function using a human artificial chromosome. Chromosoma *107*, 406-416.

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.

McAinsh, A.D., Tytell, J.D., and Sorger, P.K. (2003). Structure, function, and regulation of budding yeast kinetochores. Annu Rev Cell Dev Biol *19*, 519-539.

McClelland, S.E., Borusu, S., Amaro, A.C., Winter, J.R., Belwal, M., McAinsh, A.D., and Meraldi, P. (2007). The CENP-A NAC/CAD kinetochore complex controls chromosome congression and spindle bipolarity. Embo J *26*, 5033-5047.

McCollum, D. (2004). Cytokinesis: the central spindle takes center stage. Curr Biol *14*, R953-955.

Mechali, M. (2001). DNA replication origins: from sequence specificity to epigenetics. Nat Rev Genet *2*, 640-645.

Meluh, P.B., Yang, P., Glowczewski, L., Koshland, D., and Smith, M.M. (1998). Cse4p is a component of the core centromere of Saccharomyces cerevisiae. Cell *94*, 607-613.

Memedula, S., and Belmont, A.S. (2003). Sequential recruitment of HAT and SWI/SNF components to condensed chromatin by VP16. Curr Biol *13*, 241-246.

Mikkelsen, T.S., Ku, M., Jaffe, D.B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T.K., Koche, R.P.*, et al.* (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature *448*, 553-560.

Moazed, D. (2009). Small RNAs in transcriptional gene silencing and genome defence. Nature *457*, 413-420.

Moosmann, P., Georgiev, O., Le Douarin, B., Bourquin, J.P., and Schaffner, W. (1996). Transcriptional repression by RING finger protein TIF1 beta that interacts with the KRAB repressor domain of KOX1. Nucleic Acids Res *24*, 4859-4867.

Morris, C.A., and Moazed, D. (2007). Centromere assembly and propagation. Cell *128*, 647-650.

Muramoto, T., Muller, I., Thomas, G., Melvin, A., and Chubb, J.R. (2010). Methylation of H3K4 Is Required for Inheritance of Active Transcriptional States. Curr Biol.

Musacchio, A., and Salmon, E.D. (2007). The spindle-assembly checkpoint in space and time. Nat Rev Mol Cell Biol *8*, 379-393.

Nagano, T., Mitchell, J.A., Sanz, L.A., Pauler, F.M., Ferguson-Smith, A.C., Feil, R., and Fraser, P. (2008). The Air noncoding RNA epigenetically silences transcription by targeting G9a to chromatin. Science *322*, 1717-1720.

Nakano, M., Cardinale, S., Noskov, V.N., Gassmann, R., Vagnarelli, P., Kandels-Lewis, S., Larionov, V., Earnshaw, W.C., and Masumoto, H. (2008). Inactivation of a human kinetochore by specific targeting of chromatin modifiers. Dev Cell *14*, 507-522.

Nakano, M., Okamoto, Y., Ohzeki, J., and Masumoto, H. (2003). Epigenetic assembly of centromeric chromatin at ectopic alpha-satellite sites on human chromosomes. J Cell Sci *116*, 4021-4034.

Nakashima, H., Nakano, M., Ohnishi, R., Hiraoka, Y., Kaneda, Y., Sugino, A., and Masumoto, H. (2005). Assembly of additional heterochromatin distinct from centromere-kinetochore chromatin is required for de novo formation of human artificial chromosome. J Cell Sci *118*, 5885-5898.

Nasmyth, K. (2005). How do so few control so many? Cell *120*, 739-746.

Neely, K.E., Hassan, A.H., Brown, C.E., Howe, L., and Workman, J.L. (2002). Transcription activator interactions with multiple SWI/SNF subunits. Mol Cell Biol *22*, 1615-1625.

Ng, H.H., Robert, F., Young, R.A., and Struhl, K. (2003). Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. Mol Cell *11*, 709-719.

Nigg, E.A. (2001). Mitotic kinases as regulators of cell division and its checkpoints. Nat Rev Mol Cell Biol *2*, 21-32.

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

Nishioka, K., Chuikov, S., Sarma, K., Erdjument-Bromage, H., Allis, C.D., Tempst, P., and Reinberg, D. (2002). Set9, a novel histone H3 methyltransferase that facilitates transcription by precluding histone tail modifications required for heterochromatin formation. Genes Dev *16*, 479-489.

Nowak, S.J., and Corces, V.G. (2004). Phosphorylation of histone H3: a balancing act between chromosome condensation and transcriptional activation. Trends Genet *20*, 214-220.

Obuse, C., Yang, H., Nozaki, N., Goto, S., Okazaki, T., and Yoda, K. (2004). Proteomics analysis of the centromere complex from HeLa interphase cells: UV-damaged DNA binding protein 1 (DDB-1) is a component of the CEN-complex, while BMI-1 is transiently co-localized with the centromeric region in interphase. Genes Cells *9*, 105-120.

Ohta, S., Bukowski-Wills, J., Sanchez-Pulido, L., de Lima Alves, F., Wood, L.C., Chen, Z., Platani, M., Fischer, L., Hudson, D.F., Ponting, C.P.*, et al.* (2010). The protein composition of mitotic chromosomes determined using multi-classifier combinatorial proteomics. Cell *in press*.

Ohzeki, J., Nakano, M., Okada, T., and Masumoto, H. (2002). CENP-B box is required for de novo centromere chromatin assembly on human alphoid DNA. J Cell Biol *159*, 765-775.

Okada, M., Cheeseman, I.M., Hori, T., Okawa, K., McLeod, I.X., Yates, J.R., 3rd, Desai, A., and Fukagawa, T. (2006). The CENP-H-I complex is

required for the efficient incorporation of newly synthesized CENP-A into centromeres. Nat Cell Biol *8*, 446-457.

Okada, M., Okawa, K., Isobe, T., and Fukagawa, T. (2009). CENP-H-containing complex facilitates centromere deposition of CENP-A in cooperation with FACT and CHD1. Mol Biol Cell *20*, 3986-3995.

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.

Okamoto, Y., Nakano, M., Ohzeki, J., Larionov, V., and Masumoto, H. (2007). A minimal CENP-A core is required for nucleation and maintenance of a functional human centromere. Embo J *26*, 1279-1291.

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

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

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.

Peng, H., Feldman, I., and Rauscher, F.J., 3rd (2002). Hetero-oligomerization among the TIF family of RBCC/TRIM domain-containing nuclear cofactors: a potential mechanism for regulating the switch between coactivation and corepression. J Mol Biol *320*, 629-644.

Perkins, N.D., Felzien, L.K., Betts, J.C., Leung, K., Beach, D.H., and Nabel, G.J. (1997). Regulation of NF-kappaB by cyclin-dependent kinases associated with the p300 coactivator. Science *275*, 523-527.

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

Peters, A.H., Kubicek, S., Mechtler, K., O'Sullivan, R.J., Derijck, A.A., Perez-Burgos, L., Kohlmaier, A., Opravil, S., Tachibana, M., Shinkai, Y.*, et al.* (2003). Partitioning and plasticity of repressive histone methylation states in mammalian chromatin. Mol Cell *12*, 1577-1589.

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

Pidoux, A.L., and Allshire, R.C. (2004). Kinetochore and heterochromatin domains of the fission yeast centromere. Chromosome Res *12*, 521-534.

Pidoux, A.L., and Allshire, R.C. (2005). The role of heterochromatin in centromere function. Philos Trans R Soc Lond B Biol Sci *360*, 569-579.

Pluta, A.F., Mackay, A.M., Ainsztein, A.M., Goldberg, I.G., and Earnshaw, W.C. (1995). The centromere: hub of chromosomal activities. Science *270*, 1591-1594.

Politi, V., Perini, G., Trazzi, S., Pliss, A., Raska, I., Earnshaw, W.C., and Della Valle, G. (2002). CENP-C binds the alpha-satellite DNA in vivo at specific centromere domains. J Cell Sci *115*, 2317-2327.

Pray-Grant, M.G., Daniel, J.A., Schieltz, D., Yates, J.R., 3rd, and Grant, P.A. (2005). Chd1 chromodomain links histone H3 methylation with SAGA- and SLIK-dependent acetylation. Nature *433*, 434-438.

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

Riggs, A.D., Martienssen, R.A., and Russo, V.E.A. (1996). Introduction. Epigenetic mechanisms of gene regulation *Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York*, 1-4.

Rinn, J.L., Kertesz, M., Wang, J.K., Squazzo, S.L., Xu, X., Brugmann, S.A., Goodnough, L.H., Helms, J.A., Farnham, P.J., Segal, E.*, et al.* (2007). Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. Cell *129*, 1311-1323.

Ruchaud, S., Carmena, M., and Earnshaw, W.C. (2007). Chromosomal passengers: conducting cell division. Nat Rev Mol Cell Biol *8*, 798-812.

Rudd, M.K., Wray, G.A., and Willard, H.F. (2006). The evolutionary dynamics of alpha-satellite. Genome Res *16*, 88-96.

Rudolph, T., Yonezawa, M., Lein, S., Heidrich, K., Kubicek, S., Schafer, C., Phalke, S., Walther, M., Schmidt, A., Jenuwein, T.*, et al.* (2007). Heterochromatin formation in Drosophila is initiated through active removal of H3K4 methylation by the LSD1 homolog SU(VAR)3-3. Mol Cell *26*, 103-115.

Ryan, R.F., Schultz, D.C., Ayyanathan, K., Singh, P.B., Friedman, J.R., Fredericks, W.J., and Rauscher, F.J., 3rd (1999). KAP-1 corepressor protein interacts and colocalizes with heterochromatic and euchromatic HP1 proteins: a potential role for Kruppel-associated box-zinc finger proteins in heterochromatin-mediated gene silencing. Mol Cell Biol *19*, 4366-4378.

Saffery, R., Sumer, H., Hassan, S., Wong, L.H., Craig, J.M., Todokoro, K., Anderson, M., Stafford, A., and Choo, K.H. (2003). Transcription within a functional human centromere. Mol Cell *12*, 509-516.

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

Schmitz, M.L., Stelzer, G., Altmann, H., Meisterernst, M., and Baeuerle, P.A. (1995). Interaction of the COOH-terminal transactivation domain of p65 NF-kappa B with TATA-binding protein, transcription factor IIB, and coactivators. J Biol Chem *270*, 7219-7226.

Schneider, R., Bannister, A.J., Myers, F.A., Thorne, A.W., Crane-Robinson, C., and Kouzarides, T. (2004). Histone H3 lysine 4 methylation patterns in higher eukaryotic genes. Nat Cell Biol *6*, 73-77.

Schotta, G., Lachner, M., Sarma, K., Ebert, A., Sengupta, R., Reuter, G., Reinberg, D., and Jenuwein, T. (2004). A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. Genes Dev *18*, 1251-1262.

Schueler, M.G., Higgins, A.W., Rudd, M.K., Gustashaw, K., and Willard, H.F. (2001). Genomic and genetic definition of a functional human centromere. Science *294*, 109-115.

Schueler, M.G., and Sullivan, B.A. (2006). Structural and functional dynamics of human centromeric chromatin. Annu Rev Genomics Hum Genet *7*, 301-313.

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.

Schultz, D.C., Ayyanathan, K., Negorev, D., Maul, G.G., and Rauscher, F.J., 3rd (2002). SETDB1: a novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins. Genes Dev *16*, 919-932.

Schultz, D.C., Friedman, J.R., and Rauscher, F.J., 3rd (2001). Targeting histone deacetylase complexes via KRAB-zinc finger proteins: the PHD and bromodomains of KAP-1 form a cooperative unit that recruits a novel isoform of the Mi-2alpha subunit of NuRD. Genes Dev *15*, 428-443.

Schwartz, Y.B., and Pirrotta, V. (2008). Polycomb complexes and epigenetic states. Curr Opin Cell Biol *20*, 266-273.

Selth, L.A., Sigurdsson, S., and Svejstrup, J.Q. (2010). Transcript Elongation by RNA Polymerase II. Annu Rev Biochem *79*, 271-293.

Shahbazian, M.D., and Grunstein, M. (2007). Functions of site-specific histone acetylation and deacetylation. Annu Rev Biochem *76*, 75-100.

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.

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

Shi, Y., Lan, F., Matson, C., Mulligan, P., Whetstine, J.R., Cole, P.A., Casero, R.A., and Shi, Y. (2004). Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. Cell *119*, 941-953.

Shi, Y., Sawada, J., Sui, G., Affar el, B., Whetstine, J.R., Lan, F., Ogawa, H., Luke, M.P., Nakatani, Y., and Shi, Y. (2003). Coordinated histone modifications mediated by a CtBP co-repressor complex. Nature *422*, 735-738.

Shi, Y., and Whetstine, J.R. (2007). Dynamic regulation of histone lysine methylation by demethylases. Mol Cell *25*, 1-14.

Sims, R.J., 3rd, Belotserkovskaya, R., and Reinberg, D. (2004). Elongation by RNA polymerase II: the short and long of it. Genes Dev *18*, 2437-2468.

Sims, R.J., 3rd, Chen, C.F., Santos-Rosa, H., Kouzarides, T., Patel, S.S., and Reinberg, D. (2005). Human but not yeast CHD1 binds directly and selectively to histone H3 methylated at lysine 4 via its tandem chromodomains. J Biol Chem *280*, 41789-41792.

Sims, R.J., 3rd, Nishioka, K., and Reinberg, D. (2003). Histone lysine methylation: a signature for chromatin function. Trends Genet *19*, 629-639.

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.

Sripathy, S.P., Stevens, J., and Schultz, D.C. (2006). The KAP1 corepressor functions to coordinate the assembly of de novo HP1-demarcated microenvironments of heterochromatin required for KRAB zinc finger protein-mediated transcriptional repression. Mol Cell Biol *26*, 8623-8638.

Stavropoulos, P., Blobel, G., and Hoelz, A. (2006). Crystal structure and mechanism of human lysine-specific demethylase-1. Nat Struct Mol Biol *13*, 626-632.

Storchova, Z., and Pellman, D. (2004). From polyploidy to aneuploidy, genome instability and cancer. Nat Rev Mol Cell Biol *5*, 45-54.

Sugata, N., Li, S., Earnshaw, W.C., Yen, T.J., Yoda, K., Masumoto, H., Munekata, E., Warburton, P.E., and Todokoro, K. (2000). Human CENP-H multimers colocalize with CENP-A and CENP-C at active centromere--kinetochore complexes. Hum Mol Genet *9*, 2919-2926.

Sugata, N., Munekata, E., and Todokoro, K. (1999). Characterization of a novel kinetochore protein, CENP-H. J Biol Chem *274*, 27343-27346.

Sugimoto, K., Yata, H., Muro, Y., and Himeno, M. (1994). Human centromere protein C (CENP-C) is a DNA-binding protein which possesses a novel DNA-binding motif. J Biochem (Tokyo) *116*, 877-881.

Sullivan, B., and Karpen, G. (2001). Centromere identity in Drosophila is not determined in vivo by replication timing. J Cell Biol *154*, 683-690.

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

Sullivan, B.A., and Schwartz, S. (1995). Identification of centromeric antigens in dicentric Robertsonian translocations: CENP-C and CENP-E are necessary components of functional centromeres. Hum Mol Genet *4*, 2189-2197.

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.

Sullivan, M., and Morgan, D.O. (2007). Finishing mitosis, one step at a time. Nat Rev Mol Cell Biol *8*, 894-903.

Sunwoo, H., Dinger, M.E., Wilusz, J.E., Amaral, P.P., Mattick, J.S., and Spector, D.L. (2009). MEN epsilon/beta nuclear-retained non-coding RNAs are up-regulated upon muscle differentiation and are essential components of paraspeckles. Genome Res *19*, 347-359.

Swedlow, J.R., and Hirano, T. (2003). The making of the mitotic chromosome: modern insights into classical questions. Mol Cell *11*, 557-569.

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

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

Taverna, S.D., Li, H., Ruthenburg, A.J., Allis, C.D., and Patel, D.J. (2007). How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. Nat Struct Mol Biol *14*, 1025-1040.

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.

Tsuduki, T., Nakano, M., Yasuoka, N., Yamazaki, S., Okada, T., Okamoto, Y., and Masumoto, H. (2006). An artificially constructed de novo human chromosome behaves almost identically to its natural counterpart during metaphase and anaphase in living cells. Mol Cell Biol *26*, 7682-7695.

Uesugi, M., Nyanguile, O., Lu, H., Levine, A.J., and Verdine, G.L. (1997). Induced alpha helix in the VP16 activation domain upon binding to a human TAF. Science *277*, 1310-1313.

Vagnarelli, P., Ribeiro, S.A., and Earnshaw, W.C. (2008). Centromeres: old tales and new tools. FEBS Lett *582*, 1950-1959.

Vakoc, C.R., Mandat, S.A., Olenchock, B.A., and Blobel, G.A. (2005). Histone H3 lysine 9 methylation and HP1gamma are associated with transcription elongation through mammalian chromatin. Mol Cell *19*, 381-391.

Vakoc, C.R., Sachdeva, M.M., Wang, H., and Blobel, G.A. (2006). Profile of histone lysine methylation across transcribed mammalian chromatin. Mol Cell Biol *26*, 9185-9195.

Van Hooser, A.A., Mancini, M.A., Allis, C.D., Sullivan, K.F., and Brinkley, B.R. (1999). The mammalian centromere: structural domains and the attenuation of chromatin modeling. Faseb J *13 Suppl 2*, S216-220.

Van Hooser, A.A., Ouspenski, II, 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.

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.

Volpe, T., Schramke, V., Hamilton, G.L., White, S.A., Teng, G., Martienssen, R.A., and Allshire, R.C. (2003). RNA interference is required for normal centromere function in fission yeast. Chromosome Res *11*, 137-146.

Volpe, T.A., Kidner, C., Hall, I.M., Teng, G., Grewal, S.I., and Martienssen, R.A. (2002). Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. Science *297*, 1833-1837.

Vos, L.J., Famulski, J.K., and Chan, G.K. (2006). How to build a centromere: from centromeric and pericentromeric chromatin to kinetochore assembly. Biochem Cell Biol *84*, 619-639.

Warburton, P.E., Cooke, C.A., Bourassa, S., Vafa, O., Sullivan, B.A., Stetten, G., Gimelli, G., Warburton, D., Tyler-Smith, C., Sullivan, K.F.*, et al.* (1997). Immunolocalization of CENP-A suggests a distinct nucleosome structure at the inner kinetochore plate of active centromeres. Curr Biol *7*, 901-904.

Weake, V.M., and Workman, J.L. (2008). Histone ubiquitination: triggering gene activity. Mol Cell *29*, 653-663.

Welburn, J.P., Vleugel, M., Liu, D., Yates, J.R., 3rd, Lampson, M.A., Fukagawa, T., and Cheeseman, I.M. (2010). Aurora B phosphorylates spatially distinct targets to differentially regulate the kinetochore-microtubule interface. Mol Cell *38*, 383-392.

Willard, H.F. (1985). Chromosome-specific organization of human alpha satellite DNA. Am J Hum Genet *37*, 524-532.

Willard, H.F. (1990). Centromeres of mammalian chromosomes. Trends Genet *6*, 410-416.

Wood, A., Schneider, J., Dover, J., Johnston, M., and Shilatifard, A. (2003). The Paf1 complex is essential for histone monoubiquitination by the Rad6-Bre1 complex, which signals for histone methylation by COMPASS and Dot1p. J Biol Chem *278*, 34739-34742.

Workman, J.L. (2006). Nucleosome displacement in transcription. Genes Dev *20*, 2009-2017.

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.

Zaratiegui, M., Irvine, D.V., and Martienssen, R.A. (2007). Noncoding RNAs and gene silencing. Cell *128*, 763-776.

Zeitlin, S.G., Shelby, R.D., and Sullivan, K.F. (2001). CENP-A is phosphorylated by Aurora B kinase and plays an unexpected role in completion of cytokinesis. J Cell Biol *155*, 1147-1157.

Zhao, Y., He, S., Liu, C., Ru, S., Zhao, H., Yang, Z., Yang, P., Yuan, X., Sun, S., Bu, D.*, et al.* (2008). MicroRNA regulation of messenger-like noncoding RNAs: a network of mutual microRNA control. Trends Genet *24*, 323-327.

Zippo, A., Serafini, R., Rocchigiani, M., Pennacchini, S., Krepelova, A., and Oliviero, S. (2009). Histone crosstalk between H3S10ph and H4K16ac generates a histone code that mediates transcription elongation. Cell *138*, 1122-1136.

Appendix I

Real-time PCR Programme

The following parameters were used for all real-time PCR analysis on the LightCycler480 system (Roche) but is readily applicable to any other real time PCR machine.

Temperature	Time	Ramp Rate	Acquisition
		Hot Start	
95℃	5 min	4.4℃ / sec	none
		Amplification (x40)	
95℃	15 sec	4.4℃ / sec	none
61℃	15 sec	2.2℃ / sec	none
72℃	45 sec	4.4℃ / sec	single
		Melt Curve	
95℃	30 sec	4.4℃ / sec	none
55℃	30 sec	2.2℃ / sec	none
99℃		0.3℃ / sec	continuous (2/℃)

Appendix II

Macro for Quantifications of HAC-associated EYFP and TexasRed Fluorescence Signals

The following Macro was kindly composed by David A. Kelly and is intended for use in Image Pro 6.3 (Media Cybernetics). The output is the background-subtracted fluorescence signal within the custom-placed region of interests, calculated as the sum of all sections specified.

```
Option Explicit
Public StartFrame As Integer, EndFrame As Integer, NumFrames As Integer
Public TheGreen1(100) As Single, TheGreen2(100) As Single, TheBkGrd(100) As Single, BkAv As Single
Public TheRed1(100) As Single, TheRed2(100) As Single,TheBkGrdRed(100) As Single, BkAvRed As Single, RandomRed(500) As Single
Public mypt1 As POINTAPI, mypt1a As POINTAPI, mypt2 As POINTAPI
Public ValNum As Long
Sub Main
	Prep_Images
	PointsClick
	RedStuff
	RandomRedValues
	OutputData
End Sub
Sub Prep_Images
	ret = MsgBox("Open an Image",vbOkCancel)
	ret = IpSMOpen(SM_FILE, "\\129.215.237.150\s0679824\")

	'*Close unwanted images/dialogues*
	ret = IpCmpShow(0)		'Close colour composite dialogue
	ret = IpAppSelectDoc(3)	'Select colour composite image
	ret = IpDocClose()		'Close colour composite Image
	ret = IpAppSelectDoc(0)	'Select Blue Image
	ret = IpDocClose()		'Close Blue Image
	ret = IpSMShowNav(SM_SHOW)
	ret = IpAppSelectDoc(1) 'Select Green Image
	ret = MsgBox("Play through series and note which frames have the spots in")
End Sub
Sub PointsClick()
```

```
        Dim Howmany As Integer, x As Integer, y As Integer, z As Integer

        ret = IpAppSelectDoc(1) 'Select Green Image
        ret = IpSMPlay(SM_FIRST)        'Set file to first frame

        StartFrame = InputBox("Which Frame would you like to start from")
        StartFrame = StartFrame - 1
        EndFrame = InputBox("Which Frame is the end frame")
        NumFrames = EndFrame - StartFrame
        'Howmany = InputBox("How many points are there")

        ret = IpSeqPlay(StartFrame)
        ret = IpBlbLoadSetting("\\129.215.237.150\s0679824\Mitotics
Macro\JanPoint.env")
        ret = IpDcSet(DC_AUTO, 1)       'Turn on Auto-Collection in Data
Collector
        ret = IpDcMeasList(DC_LOAD,"\\129.215.237.150\s0679824\Mitotics
Macro\JanPointer.dcl")
        ret = IpDcUpdate(DC_RESET)    'Remove any values in data collector

        ret = MsgBox("Zoom in on spots you want to measure")
        'For x = 1 To Howmany

                '** First Spot to Measure    **
                ret = IpSeqPlay(StartFrame)
                ret = IpSMPlay(StartFrame)
                ret = IpDocClick("Select the point", mypt1)
                ipRect.Left = (mypt1.x - 4)
                ipRect.top = (mypt1.y - 4)
                ipRect.Right = (mypt1.x + 4)
                ipRect.bottom = (mypt1.y + 4)
                ret = IpAoiCreateEllipse(ipRect)

                For y = 1 To NumFrames
                        ret = IpBlbCount()
                        ret = IpDcGet(DC_DATA, 1, TheGreen1(y))
                        ret = IpSMPlay(SM_NEXT)
                        ret = IpDcUpdate(DC_RESET)    'Remove any values in
data collector
                Next y


                '**      Second Spot to Measure    **
                ret = IpSeqPlay(StartFrame)
                ret = IpSMPlay(StartFrame)
                ret = IpDocClick("Select the point", mypt1a)
```

```
                ipRect.Left = (mypt1a.x - 4)
                ipRect.top = (mypt1a.y - 4)
                ipRect.Right = (mypt1a.x + 4)
                ipRect.bottom = (mypt1a.y + 4)
                ret = IpAoiCreateEllipse(ipRect)

                For y = 1 To NumFrames
                        ret = IpBlbCount()
                        ret = IpDcGet(DC_DATA, 1, TheGreen2(y))
                        ret = IpSMPlay(SM_NEXT)
                        ret = IpDcUpdate(DC_RESET)     'Remove any values in
data collector
                Next y

'       Next x

        ret = IpWsZoom(100)
        ret = IpSeqPlay(StartFrame)
        ret = IpSMPlay(StartFrame)
        ret = IpDocClick("Select an area for Backround", mypt2)
        ipRect.Left = (mypt2.x - 4)
        ipRect.top = (mypt2.y - 4)
        ipRect.Right = (mypt2.x + 4)
        ipRect.bottom = (mypt2.y + 4)
        ret = IpAoiCreateEllipse(ipRect)

        For x = 1 To NumFrames
                ret = IpBlbCount()
                ret = IpDcGet(DC_DATA, 1, TheBkGrd(x))
                ret = IpSMPlay(SM_NEXT)
                ret = IpDcUpdate(DC_RESET)     'Remove any values in data
collector
        Next x

        '** Find Average sum background **
        For x = 1 To NumFrames
                BkAv = BkAv + TheBkGrd(x)
        Next x
        BkAv = BkAv/NumFrames
        '**Corrects the first spot
        For z = 1 To NumFrames
                        TheGreen1(z) = TheGreen1(z) - BkAv
        Next z

        For z = 1 To NumFrames
                        TheGreen2(z) = TheGreen2(z) - BkAv
```

```
        Next z
End Sub

Sub RedStuff
        Dim x As Integer, y As Integer, z As Integer
                ret = IpAppSelectDoc(2) 'Select Red Image
                '** First Spot to Measure    **
                ret = IpSeqPlay(StartFrame)
                ret = IpSMPlay(StartFrame)
ret = IpDocClick("Select the point", mypt1)
                ipRect.Left = (mypt1.x - 4)
                ipRect.top = (mypt1.y - 4)
                ipRect.Right = (mypt1.x + 4)
                ipRect.bottom = (mypt1.y + 4)
                ret = IpAoiCreateEllipse(ipRect)

                For y = 1 To NumFrames
                        ret = IpBlbCount()
                        ret = IpDcGet(DC_DATA, 1, TheRed1(y))
                        ret = IpSMPlay(SM_NEXT)
                        ret = IpDcUpdate(DC_RESET)     'Remove any values in
data collector
                Next y


                '**      Second Spot to Measure    **
                ret = IpSeqPlay(StartFrame)
                ret = IpSMPlay(StartFrame)
ret = IpDocClick("Select the point", mypt1a)
                ipRect.Left = (mypt1a.x - 4)
                ipRect.top = (mypt1a.y - 4)
                ipRect.Right = (mypt1a.x + 4)
                ipRect.bottom = (mypt1a.y + 4)
                ret = IpAoiCreateEllipse(ipRect)

                For y = 1 To NumFrames
                        ret = IpBlbCount()
                        ret = IpDcGet(DC_DATA, 1, TheRed2(y))
                        ret = IpSMPlay(SM_NEXT)
                        ret = IpDcUpdate(DC_RESET)     'Remove any values in
data collector
                Next y
        ret = IpSeqPlay(StartFrame)
        ret = IpSMPlay(StartFrame)
        For x = 1 To NumFrames
```

```
ret = IpDocClick("Select an area for Backround", mypt2)
                ipRect.Left = (mypt2.x - 4)
                ipRect.top = (mypt2.y - 4)
                ipRect.Right = (mypt2.x + 4)
                ipRect.bottom = (mypt2.y + 4)
                ret = IpAoiCreateEllipse(ipRect)

                ret = IpBlbCount()
                ret = IpDcGet(DC_DATA, 1, TheBkGrdRed(x))
                ret = IpSMPlay(SM_NEXT)
                ret = IpDcUpdate(DC_RESET)    'Remove any values in data
collector
        Next x

        '** Find Average sum background **
        For x = 1 To NumFrames
                BkAvRed = BkAvRed + TheBkGrdRed(x)
        Next x
        BkAvRed = BkAvRed/NumFrames
        '**Corrects the first spot
        For z = 1 To NumFrames
                        TheRed1(z) = TheRed1(z) - BkAvRed
        Next z
        For z = 1 To NumFrames
                        TheRed2(z) = TheRed2(z) - BkAvRed
        Next z

End Sub

Sub RandomRedValues
        Dim x As Integer

        ret = IpAppSelectDoc(2) 'Select Red Image
        ret = IpAoiShow(FRAME_NONE)
        ret       =       IpBlbLoadSetting("\\129.215.237.150\s0679824\Mitotics
Macro\JanB_RandomRed_mitotic.env")
        ret = IpDcUpdate(DC_RESET)    'Remove any values in data collector
        ret = IpBlbCount()

        ret = IpDcGet(DC_NUMROW, -1, ValNum)

        For x = 0 To (ValNum-1)
                ret = IpDcSet(DC_ROW, x)
                ret = IpDcGet(DC_DATA, 1, RandomRed(x))
        Next x
```

```
End Sub

Sub OutputData()
        Dim CellName As String, ValueName As String
        Dim x As Integer

        ret = MsgBox("OPen Excel")
        ret = IpDde(DDE_OPEN, "excel", "sheet1") 'Open DDE Link to Excel

        '***             First Green Spot      ***
        For x = 1 To NumFrames
                CellName = "R" + CStr(x) + "C1"
                ValueName = CStr(TheGreen1(x))
                ret = IpDde(DDE_PUT, CellName, ValueName)
        Next x

        '***             Second Green Spot ***
        For x = 1 To NumFrames
                CellName = "R" + CStr(x) + "C2"
                ValueName = CStr(TheGreen2(x))
                ret = IpDde(DDE_PUT, CellName, ValueName)
        Next x

        '***             First Red Spot        ***
        For x = 1 To NumFrames
                CellName = "R" + CStr(x) + "C3"
                ValueName = CStr(TheRed1(x))
                ret = IpDde(DDE_PUT, CellName, ValueName)
        Next x

        '***             Second Red Spot    ***
        For x = 1 To NumFrames
                CellName = "R" + CStr(x) + "C4"
                ValueName = CStr(TheRed2(x))
                ret = IpDde(DDE_PUT, CellName, ValueName)
        Next x

        '***             RandomRed Spots  ***
        For x = 1 To ValNum
                CellName = "R" + CStr(x) + "C5"
                ValueName = CStr(RandomRed(x))
                ret = IpDde(DDE_PUT, CellName, ValueName)
        Next x

End Sub
```

Macro for Quantifications of Endogenous Centromere TexaRed Fluorescence Signals

The following Macro was kindly composed by David A. Kelly and is intended for use in Image Pro Plus 7.0 (Media Cybernetics). User-defined intensity thresholds are used to determine the centre x-y of any spot-shaped structure followed by plotting a 9 pixel diameter region of interest around each centre. The output is the average fluorescence intensity of all regions of interest, calculated on a maximum intensity projection.

```
 ' Default Script
Public RedOne As Integer, RedProject As Integer
Public numobj As Long
Public centreX(1000) As Single, centreY(1000) As Single
Option Explicit
Sub StartPoint
        OpenFile
        ProcessFile
        MeasureSpots
End Sub
Sub OpenFile
        ret = MsgBox("Start Excel")
        ret = IpMacroStop("Please open image", MS_MODAL)          'Prompts
for the file
        ret = IpTemplateMode(1)
                'Opens a open file dialogue box
        ret = IpSMOpen(SM_FILE, "\\129.215.237.150\s0679824\")  'Sets    a
starting point for the open file dialogue
        ret = IpTemplateMode(0)
                'Closes the open file dialogue box
        ret = IpAppSelectDoc(0)
        ret = IpDocClose
        ret = IpAppSelectDoc(1)
        ret = IpDocClose
        ret = IpDocGet(GETACTDOC, 0, RedOne)
End Sub

Sub ProcessFile()

        Dim Y As Integer

        ret = IpEDFAdd(2)
```

```
        ret = IpEDFNew(2)
        ret = IpEDFSet(EDF_NORMALIZE, 0, 0)
        ret = IpEDFSet(EDF_CRITERIA, EDF_MAX_INTENSITY, 0)
        ret = IpEDFSet(EDF_ORDER, EDF_BOTTOMUP, 0)
        ret = IpEDFCreate(EDF_COMPOSITE)
        ret = IpEDFShow(0)
        ret = IpDocGet(GETACTDOC, 0, RedProject)
        ' Original function IpSCalSetUnit(1.000000, 0.064680)
        ret = IpSCalSetSng(SCAL_CURRENT_CAL, SCAL_SCALE_X,
1.000000)
        ' Original function IpSCalSetUnit(1.000000, 1.000000)
        ret = IpSCalSetSng(SCAL_CURRENT_CAL, SCAL_SCALE_Y,
1.000000)
        ret = IpSCalSetLong(SCAL_CURRENT_CAL, SCAL_APPLY, 0)
        ret = IpBlbLoadSetting("\\129.215.237.150\s0679824\red
macro\countset.env")
        ret = IpDcMeasList(DC_LOAD,"\\129.215.237.150\s0679824\red
macro\collectorset.dcl")
        ret = IpBlbShow(1)
        ret = IpSegShow(1)
        ret = MsgBox("Threshold the Image")

        '*************************************************************
        '*          Section to get x/y coordinates of all selected spots      *
        '*************************************************************
        ret = IpDcUpdate(DC_RESET)

        ret = IpBlbCount()
        ret = IpBlbUpdate(0)
        ret = IpBlbSplitObjects(1)
        ret = IpDcUpdate(DC_RESET)
        ret = IpBlbGet(GETNUMOBJEX, 0, 0, numobj) 'Gets the number of
spots in the image
        ret = IpDcUpdate(DC_FETCH)
        For Y = 0 To numobj -1
                '*      Get Centre x spot value for array  *
                ret = IpDcSet(DC_COL, 0)
                ret = IpDcSet(DC_ROW, Y)
                ReDim fData(10) As Single
                ret = IpDcGet(DC_DATA, 1, fData(0))
                centreX(Y) = fData(0)

        '*          Get Centre y spot value for array  *
                ret = IpDcSet(DC_COL, 1)
                ret = IpDcSet(DC_ROW, Y)
                ReDim fData(10) As Single
```

```
            ret = IpDcGet(DC_DATA, 1, fData(0))
            centreY(Y) = fData(0)
        Next Y

End Sub
Sub MeasureSpots
        Dim Y As Integer

        ret = IpDcMeasList(DC_LOAD,"\\129.215.237.150\s0679824\red
macro\collectorset.dcl")
        ret = IpBlbLoadSetting("\\129.215.237.150\s0679824\red
macro\intensityset.env")
        ret = IpDcUpdate(DC_RESET)
        ret = IpDcSet(DC_AUTO, 1)

        For Y = 0 To numobj - 1

            ipRect.Left = centreX(Y) - 4.5
            ipRect.top = centreY(Y) - 4.5
            ipRect.Right = centreX(Y) + 4.5
            ipRect.bottom = centreY(Y) + 4.5
            ret = IpAoiCreateEllipse(ipRect)
            ret = IpBlbCount()
            ret = IpBlbUpdate(0)
            ret = IpDcUpdate(DC_FETCH)

        Next Y
        ret = IpDcSaveData("",  S_Y_AXIS + S_X_AXIS + S_DDE)

End Sub
```