Review TRENDS in Cell Biology Vol.16 No.2 February 2006

Chromosome Segregation and Aneuploidy series

Centromere formation: from epigenetics to self-assembly

Christopher W. Carroll and Aaron F. Straight

Department of Biochemistry, Stanford University, Beckman Building, Rm. 409, 279 Campus Drive, Stanford, CA 94305-5307, USA

This review is part of the *Chromosome segregation and Aneuploidy* series that focuses on the importance of chromosome segregation mechanisms in maintaining genome stability. Centromeres are specialized chromosomal domains that serve as the foundation for the mitotic kinetochore, the interaction site between the chromosome and the mitotic spindle. The chromatin of centromeres is distinguished from other chromosomal loci by the unique incorporation of the centromeric histone H3 variant, centromere protein A. Here, we review the genetic and epigenetic factors that control the formation and maintenance of centromeric chromatin and propose a chromatin self-assembly model for organizing the higher-order structure of the centromere.

Introduction

Cell division is essential for the reproduction and development of all organisms. In unicellular species, each cell division produces a new organism, whereas in multicellular species, cell division generates new cells that differentiate into distinct tissues and cell types. In both cases, it is crucial that each new cell is genetically identical to the cell from which it was derived, thus ensuring that information crucial to the reproduction and development of the organism is not lost.

Eukaryotic cells have evolved complex structural and regulatory mechanisms that maintain the genetic integrity of the organism. During each cell division cycle, chromosomes are precisely duplicated during DNA replication and the resulting sister chromatids are equally segregated during mitosis. Central to this accurate segregation is the assembly of a single site for microtubule attachment, called the kinetochore, on each sister chromatid. Kinetochores form the primary interface between chromosomes and the mitotic spindle and mediate microtubule-dependent chromosome movements during mitosis. Kinetochores also sense errors in chromosome attachment to the mitotic spindle and respond by activating the spindle assembly checkpoint (SAC), which delays the metaphase-to-anaphase transition until all chromosomes achieve bipolar spindle attachment (reviewed in Refs [1–3]).

Here, we focus on the centromere, the specialized region of the chromosome that directs kinetochore assembly. The centromere was first described cytologically as the primary constriction on vertebrate chromosomes and was later characterized as a chromosomal region that had a reduced recombination frequency [4]. Recently, many of the molecular features that distinguish centromeres from other chromatin have been defined. Centromeric DNA is extremely diverse among eukaryotes, ranging from the simple ~125-bp centromere of *Saccharomyces cerevisiae* to the highly repetitive α-satellite sequences of vertebrates. Despite these differences in nucleotide sequence, all centromeres share a unique chromatin composition that is characterized by the incorporation of the histone 3 (H3) variant centromere protein A (CENP-A) within nucleosomes of centromeric chromatin. This centromeric chromatin is essential for the formation of a functional kinetochore in all eukaryotes. Current experimental evidence indicates that DNA sequence and epigenetic factors that modify centromeric chromatin contribute to centromere assembly and inheritance, but the mechanisms by which they do so are not well understood. Here, we discuss current literature and models regarding the genetic and epigenetic factors that control centromere assembly and inheritance, with an emphasis on recent advances in yeast and vertebrate systems.

Centromeric DNA

Centromeric DNA is extremely variable among species. At the extremes are the simple 125-bp 'point' centromeres of *S. cerevisiae* and holocentric centromeres of *Caenorhabditis elegans*, which span the entire length of the chromosome. Most eukaryotes, however, have complex centromeres that consist of long stretches of repetitive DNA that are highly divergent between different species (Figure 1). The best-characterized complex centromeres are those of the fission yeast *Schizosaccharomyces pombe*, for which the complete nucleotide sequence of the centromeric DNA of all three chromosomes is known [5]. Each centromere in *S. pombe* contains a non-repetitive central sequence (*cnt*) ~4–7 kb in length that is flanked by centromere-specific innermost repeats (*imr*), which together form the site of kinetochore assembly. This central domain is surrounded by long tandem arrays of outer repeats (*otr*) that are common to all three centromeres in the organism, resulting in total centromere lengths of ~40–100 kb [6]. Plasmids that contain the unique central core and a portion of the *otr* domain are stably inherited and assemble functional kinetochores

Corresponding author: Straight, A.F. (astraigh@stanford.edu).

 0962-8924/$ - see front matter doi:10.1016/j.tcb.2005.12.008

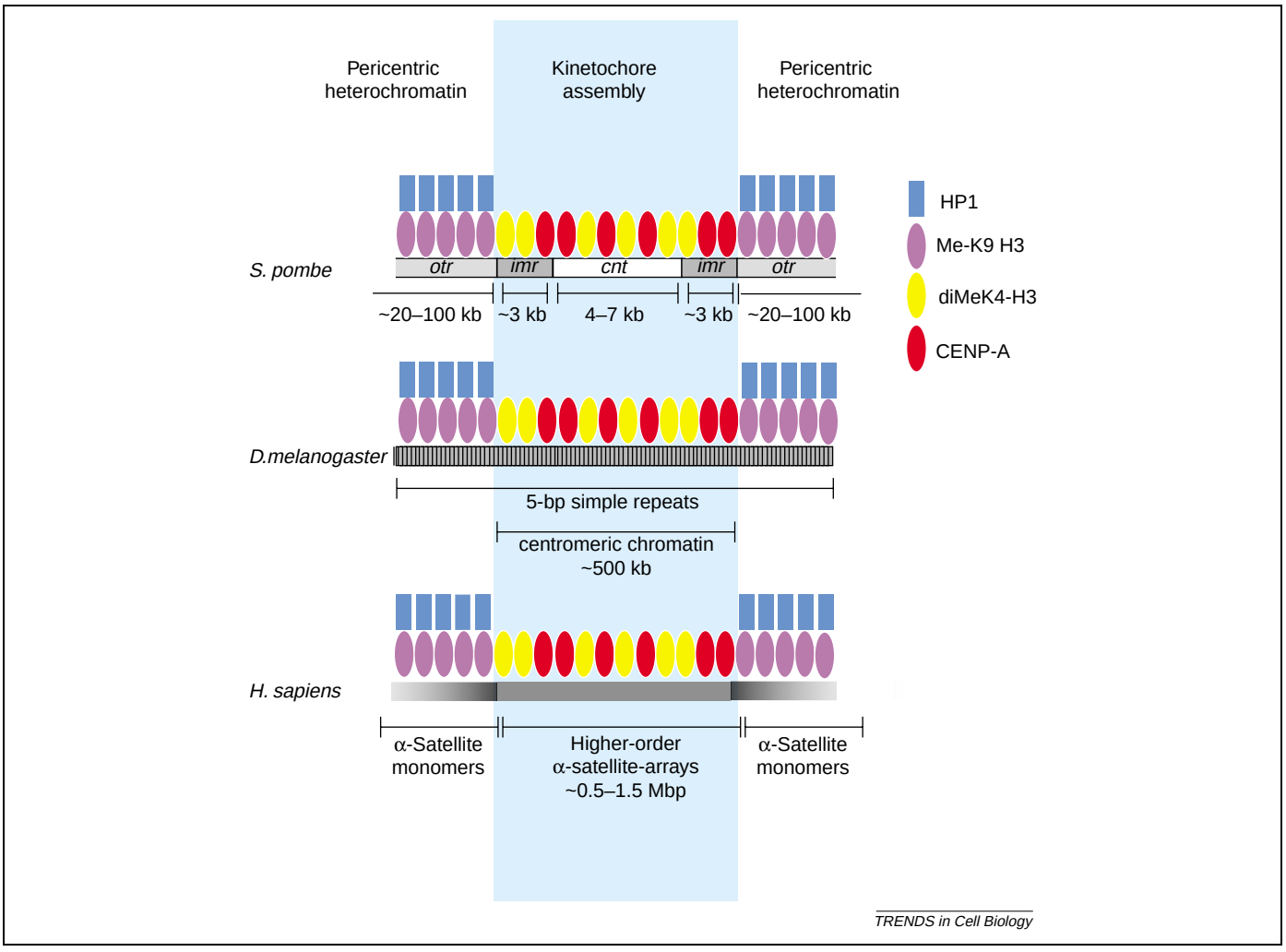

Figure 1. The nucleic acid and protein features of centromeric chromatin from *S. pombe, D. melanogaster* and human. The centromeric DNA from each organism differs in size, and each contains distinct repetitive DNA. Despite significant differences in the DNA composition, the overall chromatin structure of the centromere is conserved among these species.

when introduced into cells. However, no single DNA sequence is essential for kinetochore assembly, indicating that DNA sequence and epigenetic mechanisms mediate centromere assembly in *S. pombe* [7–9].

Centromeres in metazoan species are more complex than those in yeasts. The mapping and large-scale sequencing of a substantial portion of a functional *Drosophila melanogaster* centromere have identified several islands of complex sequence within a long, otherwise contiguous region of simple repetitive DNA [10,11]. Centromeres of human chromosomes contain large arrays of tandemly repeated 171-bp α-satellite DNA that can span several megabases. However, the complete nucleotide sequence of a metazoan centromere has not yet been determined. Sequencing of a portion of the human X-chromosome centromere indicated that functional centromere sequences in humans are homogeneous repeats of α-satellite DNA [12]. However, without the complete sequence of a centromere, the presence of unique, non-alphoid DNA in vertebrate centromeres cannot be ruled out.

α-Satellite DNA cloned from human chromosomes and synthetic arrays of a human α-satellite repeat supports centromere formation when introduced into cultured cells [13,14]. Many human α-satellite repeats contain a conserved binding site for the sequence-specific DNA-binding protein known as centromere protein B (CENP-B) [15], indicating that α-satellite DNA promotes centromere formation by recruiting sequence-specific DNA-binding proteins. Consistent with this, mutation of the CENP-B binding site reduces the efficiency of mammalian artificial chromosome formation on synthetic alphoid arrays [16]. However, the role of DNA sequence in vertebrate centromere function is currently a matter of debate. Mice that lack CENP-B are viable and do not display significant defects in chromosome segregation [17,18]. Moreover, rare human neocentromeres have been observed, such as the centromere of the mardel(10) chromosome, on which a new centromere was activated in a previously non-centromeric region [19]. These neocentromeres form without α-satellite DNA and do not bind CENP-B but do assemble fully functional kinetochores that are stably inherited through many generations of cell division [20,21]. Thus, specific DNA sequences and, in particular, repetitive DNA do not seem to be essential for centromere formation but could provide a favorable environment for establishing centromeric chromatin, possibly by recruiting sequence-specific

binding proteins. Once established, however, it is likely that epigenetic mechanisms ensure the stable propagation of centromeres.

Features of centromeric chromatin

CENP-A and centromeric chromatin

Despite the diversity in size and sequence of centromeric DNA, the overall architecture and composition of centromeric chromatin is similar between different species (Figure 1). As previously stated, one hallmark of functional centromeres is the replacement of H3 in centromeric nucleosomes by the H3 variant CENP-A. CENP-A is found exclusively in centromeric chromatin and mutation of CENP-A in all eukaryotes results in a complete failure in chromosome segregation [22–27]. Cells that lack CENP-A fail to recruit most known kinetochore components to the centromere in interphase and mitosis [28].

The requirement for CENP-A in kinetochore formation, the dependence of most other kinetochore proteins on CENP-A for their localization and the incorporation of CENP-A only within active centromeres make CENP-A a good candidate for providing an epigenetic mark that specifies centromere position. However, several other kinetochore proteins, including centromere protein C (CENP-C), centromere protein I (CENP-I) and centromere protein H (CENP-H), are also exclusively targeted to centromeres and are required for kinetochore function [29–31]. In addition, recent work has shown that H3-containing nucleosomes that are di-methylated on lysine 4 of H3 (diMe-K4 H3) are also present within centromeric chromatin [32–34]. diMe-K4 H3 is typically associated with euchromatin but its function in the normally transcriptionally silent chromatin of the centromere is unknown.

Pericentric heterochromatin

CENP-A-containing chromatin is usually embedded within a large domain of heterochromatin, called pericentric heterochromatin, which is also required for the accurate segregation of chromosomes during mitosis (Figure 1). The assembly and inheritance of pericentric heterochromatin is thought to be partly dependent on the methylation of H3 on lysine 9 (Me-K9 H3) by the histone methyltransferase Su(var)3–9, the binding of Me-K9 H3 by the chromodomain-containing protein HP1, and the RNA interference (RNAi) machinery (reviewed in Ref. [35]). In *S. pombe*, pericentric chromatin is assembled on the repetitive sequences of the *otr* region of centromeric DNA and is crucial for sister chromatid cohesion. *S. pombe* cells that lack the HP1 homolog *Swi6*, and are therefore deficient in pericentric heterochromatin, are unable to recruit the chromosome cohesin subunit Rad21 to centromeres and fail to maintain centromere cohesion [36,37]. However, cohesion of the chromosome arms is unaffected. Cohesins associate with pericentric heterochromatin in mammalian cells as they do in yeast, and several mammalian neocentromeres and artificial chromosomes contain pericentric heterochromatin, highlighting the close functional relationships between these chromatin domains [38,39]. In cell lines that lack Su(var)3–9, HP1 no longer associates with pericentric heterochromatin and cohesion is lost between pericentric regions of sister chromatids [40].

Centromeric and pericentric chromatin are functionally distinct

Centromeric chromatin and the surrounding pericentric heterochromatin seem to function independently. Genes that are integrated within the *cnt*, *imr*, or surrounding *otr* domains in *S. pombe* centromeres are transcriptionally silenced owing to the assembly of kinetochores or pericentric heterochromatin, respectively [41]. Mutations that disrupt pericentromeric heterochromatin derepress the expression of a reporter gene integrated within the *otr* domain, but not an identical reporter gene located within the unique central core. By contrast, the expression of a reporter gene that has been integrated into the *cnt* or the *imr* region of the centromere is derepressed by mutations that disrupt kinetochore function; however, an identical reporter present in pericentromeric chromatin is not expressed under the same conditions [6]. Deletion of the Su(var)3–9 methyltransferase in flies and mammals, or the deletion of the Dicer ribonuclease in mammals, disrupts pericentric heterochromatin and causes defects in chromosome cohesion but does not affect the localization of centromere antigens such as CENP-A [22,40,42]. In *D. melanogaster* cells that are depleted of CENP-A, centromeric chromatin and kinetochore function is lost but HP1 can still localize to pericentric heterochromatin [22].

Chromatin boundaries within centromeres

How are the distinct domains of centromeric chromatin and pericentric heterochromatin insulated from one another? At the mating type locus in *S. pombe*, the spread of heterochromatin is limited by defined DNA sequences that demarcate the boundaries between adjacent chromatin domains, and similar mechanisms might function at centromeres [43]. *S. pombe* centromeres have an unusual nucleosome structure within the unique *imr–cnt* central domain. Digestion of *S. pombe* chromatin with micrococcal nuclease, which non-specifically cleaves DNA between nucleosomes, results in a smeared appearance of *imr–cnt* DNA, rather than the regular repeating ladder that is seen after micrococcal nuclease digestion of euchromatin; this could reflect an atypical spacing between adjacent nucleosomes [7,44]. CENP-A is a possible candidate for maintaining these barriers; in CENP-A mutants, cells are unable to assemble the unique chromatin structure characteristic of centromeres and ordered nucleosomes invade the central core chromatin [27]. However, the absence of CENP-A does not lead to the ectopic spread of heterochromatin into the central core, shown by the derepression of normally silenced reporter genes within the central core of the centromere in the absence of CENP-A. Interestingly, many tRNA genes are found at the boundary between centromeric chromatin and pericentric heterochromatin in *S. pombe* [45]. The binding of regulatory factors to the promoters of tRNA genes or the active transcription of these tRNAs might help define

the boundary between centromeric and pericentromeric chromatin domains.

What determines the boundary between centromeric chromatin and the adjacent pericentric heterochromatin in vertebrates is less clear. α-Satellite arrays associated with centromere function are highly homogeneous, whereas satellite DNA that flanks the centromere is more divergent at the sequence level, has fewer binding sites for CENP-B and is interspersed with non-alphoid DNA [12]. However, only a subset of the homogenous arrays of α-satellite sequence is typically associated with CENP-A, with the remainder being packaged in pericentric heterochromatin, indicating that DNA sequence is unlikely to establish the barrier between centromeric chromatin and the surrounding pericentric heterochromatin. One possibility is that CENP-A itself inhibits the ectopic spread of pericentric heterochromatin within vertebrate centromeres. Histone variants can restrict the spread of heterochromatin in other contexts. In *S. cerevisiae*, for example, the histone H2A variant H2A.Z inhibits the spread of heterochromatin from telomeres and the mating type locus [46]. The mechanisms that vertebrate cells use to maintain the distinction between centromeric chromatin and pericentric heterochromatin have yet to be determined.

How are centromeres stably propagated?

One of the most interesting questions in understanding centromere function is how centromeres are stably propagated from one generation to the next. Propagation requires two steps: first, the centromere must in some way be marked as the proper site for centromeric chromatin formation; and second, that mark must be recognized by the factors that assemble the centromeric chromatin (Figure 2). The best candidate for an epigenetic mark that specifies the centromere is CENP-A [47,48]. During each round of DNA replication, CENP-A is distributed to each of the replicated sister chromatids [49], thereby ensuring that each sister chromatid can direct the assembly of a kinetochore in the ensuing mitosis. Mechanisms must therefore exist that specifically target newly synthesized CENP-A to centromeric chromatin during each cell cycle.

Analyses of the determinants required for the loading of CENP-A into centromeric chromatin in human cells, *D. melanogaster* and yeast have demonstrated that the histone-fold domain is required for its centromere-specific deposition [49–52]. Recently, the region of CENP-A necessary for centromere targeting, known as the CENP-A targeting domain (CATD), was further defined to include the L1 linker and α2 helix of CENP-A [53]. When substituted into the analogous region of H3, the CATD was sufficient to confer replication-independent centromere-specific targeting upon the chimeric molecule. It is still unclear, however, whether the incorporation of the CATD–H3 chimeric molecules into centromeres was because of dimerization with endogenous CENP-A–histone 4 (H4) or if homogeneous tetramers of these chimeras could incorporate into centromeres. Thus, it is still possible that other sequences outside the CATD are also required for centromere targeting of CENP-A. Reconstitution of the CENP-A–histone 4 (H4) complex demonstrated that the CENP-A–H4 tetramer was structurally distinct from a tetramer of H3 and H4 [53]. The CENP-A–H4

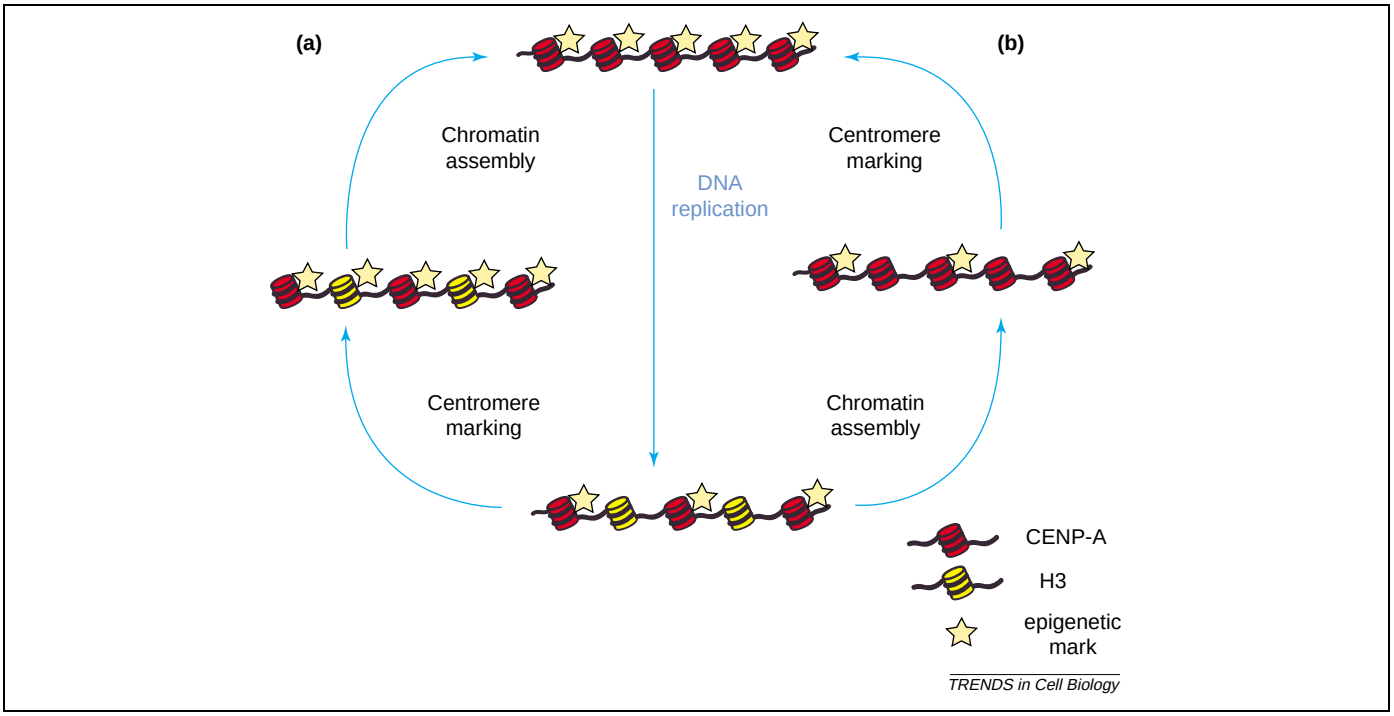

Figure 2. Models for the specification and assembly of centromeric chromatin. Centromeric chromatin must be uniquely marked to distinguish it from other chromatin. Mechanisms must also exist that recognize this mark and specifically deposit new components into the existing centromeric chromatin. The semi-conservative nature of DNA replication dilutes the components of the centromere and the mark that specifies the position during each cell cycle. Both must be restored to maintain the integrity of centromeric chromatin. In principle, the restoration of centromeric chromatin after DNA replication could occur in two ways. First, replicated centromeric chromatin could be marked, which would subsequently promote the recruitment of new CENP-A (**a**). Second, new CENP-A could be deposited within the centromeric chromatin, which would promote the new marking of centromeric chromatin (**b**).

tetramer showed slower backbone proton exchange, indicating a conformational rigidity in the CENP-A–H4 tetramer compared with the H3–H4 tetramer. Interestingly, the L1 linker and α2-helix region required for the centromere targeting of CENP-A was also the region of CENP-A that conferred the conformational rigidity on CENP-A–H4 tetramers, indicating that the unique structural properties of the CENP-A–H4 tetramer might be linked to its deposition within centromeric chromatin. The structural difference conferred by the CATD might be sufficient to mark the centromere as the site for new CENP-A deposition.

Whereas the kinetochore localization of many proteins is dependent upon CENP-A, at least one conserved protein, Mis12, is specifically localized to kinetochores independently of CENP-A [54,55]. Therefore, additional mechanisms must exist that also mark the site for centromeric chromatin assembly. Chromatin immunoprecipitation (ChIP) analysis in *S. pombe* showed that the central core of the centromere contained an increase in diMe-K4 H3, a modification that is generally associated with euchromatin, and a decrease in Me-K9 H3, a modification that is present in the surrounding pericentric heterochromatin [32]. A similar modification pattern has been observed in *D. melanogaster* and human cells by immunolocalization of modified histones on extended chromatin fibers [34]. Although this modification state is similar to that found in euchromatin, centromeres from *S. pombe*, *D. melanogaster* and human cells all have a reduced level of H3 and H4 acetylation at centromeres, resembling heterochromatic regions [34,56]. Although any single modification of conventional H3 is unlikely to mark the position of the centromere, a unique combination of modifications could serve as a landmark that specifies centromere position.

Many genetic screens in *S. pombe* have identified several genes, including *Mis6, Mis15–18*, *Ams2, Sim4* and *Mal2*, that, when mutated, cause a failure in the centromeric localization of CENP-A [27,57–60]. The loss of CENP-A indicates that these genes are important for maintaining the integrity of centromeric chromatin. Consistent with this possibility, *Mis6, Mis15–17, Sim4* and *Mal2* associate with the central core of centromeres throughout the cell cycle [27,58–60]. Immunoprecipitation experiments indicated that the encoded proteins function in the context of larger proteins complexes; *Mis6* has been shown to associate with *Sim4* and *Mis15* and *Mis17*; and *Mis16* and *Mis18* interact with one another [58,60]. Epistasis analysis indicated that *Mis16* and *Mis18* are required for the centromere recruitment of *Mis6, Mis15* and *Mis17* (and therefore presumably *Sim4*), which in turn are required for the localization of CENP-A [58]. Interestingly, in *Mis16* or *Mis18* mutants, acetylation of H3 and H4 increases within the central domain of the centromere [58].

Several of the genes that control CENP-A localization to centromeres in *S. pombe* are conserved in vertebrates (Table 1). *Schizosaccharomyces pombe Mis16* is homologous to the p46 and p48 subunits (also known as RbAp46/RBBP7 and RbAp48/RBBP4) of chromatin assembly factor 1 (CAF-1), which is required for replication-coupled nucleosome assembly. Depletion of these proteins by RNAi in human cells results in a loss of CENP-A from centromeres, indicating a role for RbAp46 and RbAp48 in targeting CENP-A to centromeres in vertebrates [58]. The RbAp46 and RbAp48 proteins are also subunits of several other chromatin modifying and remodeling complexes that have histone acetylation, histone deacetylation and nucleosome-positioning activities [61]. It is unclear which of these RbAp46/48-associated activities is relevant to the centromeric localization of CENP-A.

Table 1. Human homologs of *S. pombe* proteins required for CENP-A localization[a]

S. pombe	Human
Cnp1	CENP-A
Mis6	CENP-I
Sim4	CENP-H
Mis12	hMis12
Mis16	RbAp46 and RbAp48

[a]For further information, see main text.

Some of the genes that are required for CENP-A incorporation at centromeres in *S. pombe* do not appear to have the same role in other eukaryotes. The vertebrate homologs of the *S. pombe Mis6* and *Sim4* proteins, CENP-I and CENP-H respectively, are not required for the centromere-specific recruitment of CENP-A in chicken DT40 cells [29,30]. Instead, the recruitment of CENP-I and CENP-H to centromeric chromatin is dependent upon CENP-A in both human and chicken DT40 cells [26,54]. This reciprocal relationship also exists between *S. cerevisiae* and *S. pombe* in that the *S. cerevisiae Mis6* homolog, Ctf3, is found at centromeres but is not required to target CENP-A to *S. cerevisiae* centromeres. Instead, the centromeric targeting of Ctf3 is dependent upon CENP-A [62].

The epigenetic mark that identifies the centromere is unknown, but Allshire and colleagues have recently proposed a model that couples the functional activity of kinetochores during mitosis with reinforcing the location of the centromere [56]. The authors propose that the successful marking and propagation of centromere identity is a direct result of the kinetochore successfully engaging microtubules during mitosis. In this model, some direct output of kinetochore function, for example, tension across a pair of sister chromosomes or microtubule binding, is converted into an epigenetic change that signals the site of CENP-A deposition in the subsequent cell cycle. This model is attractive as it could help to explain the discrepancies between proteins that are involved in CENP-A loading in *S. pombe* compared with those of other species. In principle, any protein that impairs the function of the kinetochore could in turn result in a loss of proper centromere maintenance in the next cell cycle.

Recognizing the site of centromere formation

The site of centromere formation could be marked by structural differences in centromeric histones, histone modifications or some other mechanism. Regardless of the nature of the mark, the existing centromere must be recognized by the machinery that deposits new CENP-A specifically at centromeres. Recent work from several

experimental systems indicates that the histone variants H3.3 and H2A.Z are loaded onto DNA by chromatin assembly factors that are distinct from conventional histone-loading factors. The purification of proteins that interact with H3 and H3.3 showed that these histone proteins are associated with different histone chaperone complexes [63]. Canonical H3 was found in a complex with all three subunits of CAF-1, a histone chaperone previously shown to be required for the DNA-replication-coupled assembly of histones. By contrast, the H3.3 complex did not contain either the p150 or p60 subunits of CAF-1, but it did contain histone regulator A (HirA), which has been implicated in replication-independent chromatin assembly. Using a histone-loading assay in *Xenopus laevis* egg extracts, it was shown that the p150 subunit of CAF-1 was required for the loading of H3 on replicating DNA only, whereas HirA was required for the replication-independent loading of H3.3 [63]. Similarly, H2A.Z from *S. cerevisiae* interacts with a large chromatin-remodeling complex that includes the Swi/Snf family ATPase Swr1, the subunits of which are necessary for the incorporation of H2A.Z into chromatin [64–66].

It is not known to what extent a similar chromatin-remodeling complex mediates CENP-A deposition. The demonstration that RbAp46 and RbAp48 are required for CENP-A targeting in yeast and humans indicates that such activities are likely to exist. Recently, mutation of Hrp1, which is a member of the chromo-helicase/ATPase (CHD) subfamily of Swi/Snf-related proteins, was shown to derepress silencing in the central core domain and *otr* domains of *S. pombe* centromeres [67]. Loss of Hrp1 reduces the amount of CENP-A at centromeres but the unique nucleosome-laddering pattern of the *cnt* domain is not disrupted in Hrp1 mutants, and *Hrp1* is not an essential gene. Therefore, other mechanisms must also contribute to the loading of CENP-A into centromeric chromatin.

Higher-order centromeric chromatin assembly

As cells enter mitosis, chromosomes condense and the centromeric chromatin undergoes a dramatic reorganization so that the CENP-A-containing chromatin is organized along the exterior face of each sister chromatid. This organization is crucial for the proper assembly of kinetochores so that they are optimally positioned for microtubule capture and bipolar alignment on the mitotic spindle. Experiments in which mitotic chromosomes were hypotonically stretched and stained with antibodies directed against centromere antigens showed that the centromere protein-containing chromatin is not continuous [68]. Rather, centromeric chromatin is interrupted by chromatin that lacks centromere antigens. This led to the proposal of a repeat subunit model for kinetochore formation in mitosis that required the association of the centromere-protein-containing chromatin into a higher-order structure [68]. Careful analysis of stretched chromatin has revealed that blocks of nucleosomes ~10–50 kb in length that contain either CENP-A or diMe-K4 H3 are interspersed within the centromeres of human and *D. melanogaster* chromosomes [34,69]. A similar pattern was also demonstrated for nucleosomes that form the chromatin of a well-characterized human neocentromere [33]. However, examination of mitotically condensed chromosomes showed that CENP-A and diMe-K4 H3 reside in spatially distinct locations, with the diMe-K4 H3 occupying the chromosome interior and the CENP-A-containing chromatin facing the chromosome exterior [34,69]. These observations support the repeat subunit model for kinetochore formation and are consistent with the assembly of a higher-order centromeric chromatin structure. The mechanisms through which linearly interspersed blocks of CENP-A nucleosomes are resolved into cohesive units during mitosis are not yet understood.

What mechanisms might assemble the higher-order structure of the centromeric chromatin? Chromosome condensation through the activity of the condensin complex is known to be important for the mitotic chromosome structure and two condensin complexes that differ in their non-SMC subunits have been identified so far [70]. Condensin II localizes to the inner kinetochore in vertebrate mitotic chromosomes and to the kinetochore plate in *C. elegans* [40,71]. In the absence of condensin, chromosome structure is disrupted in *C. elegans*, *X. laevis* egg extract chromosomes and in chicken and human cells [17,40,71–73]. However, in all of these cases (and in *D. melanogaster* cells that lack condensin), centromere formation, CENP-A localization and microtubule attachment continue [74,75], indicating that condensin activity is important for the global structure of the chromosome but does not directly organize CENP-A-containing chromatin.

Centromere self-assembly

Here, we propose a model for centromere organization driven by the self-assembly of CENP-A-containing nucleosomes (Figure 3). Chromatin fibers that contain phased nucleosomes self-assemble through two distinct mechanisms [76]: first, in a single nucleosome array, interactions between two nucleosomes, called intranucleosomal interactions, drive the assembly of the higher-order structure of the 30-nM fiber, which has been observed *in vivo* [77,78]. Intranucleosomal assembly requires divalent cations and the N-terminal tail of H4, and is stabilized by linker histones, including histone 1 (H1) [76,79]; second, internucleosomal assembly can occur through the oligomerization of two (or more) distinct chromatin fibers, resulting in larger arrays of chromatin fibers [76]. The oligomerization of chromatin fibers requires slightly higher concentrations of divalent cations than do intranucleosomal interactions and this difference is thought to reflect long-range fiber–fiber interactions that could be important for the assembly of higher-order chromosome structures.

Histone variants can affect the propensity of reconstituted chromatin fibers to adopt either of these higher-order structures. For example, H2A.Z promotes intranucleosomal assembly but inhibits internucleosomal assembly relative to conventional H2A [76]. We suggest that internucleosomal assembly, in which linearly separated blocks of CENP-A-containing chromatin laterally

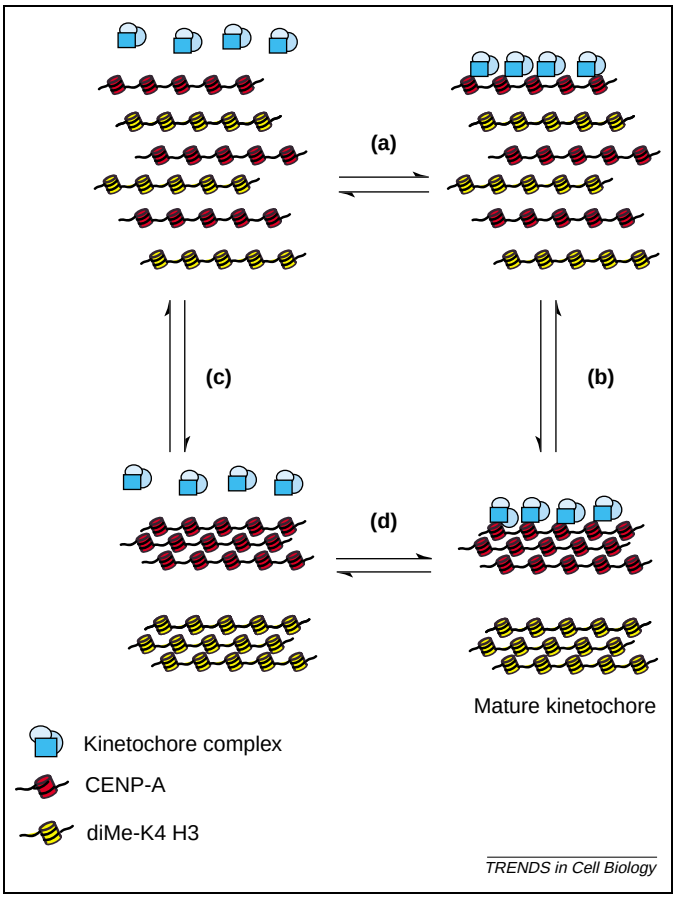

Figure 3. Model for the assembly of a mature kinetochore. Linear separated blocks of CENP-A-or H3-containing chromatin must be resolved in mitotically condensed chromosomes. Each block of CENP-A-containing chromatin could independently recruit kinetochore proteins (**a**) that would promote the formation of the higher-order chromatin structure observed in mitotic chromosomes (**b**). Alternatively the self-assembly of chromatin fibers could drive the resolution of CENP-A-containing chromatin into a higher-order structure (**c**). This higher-order structure would provide a platform for the assembly of kinetochore proteins (**d**), which in turn would stabilize the higher-order chromatin structure.

self-assemble and exclude conventional H3-containing chromatin to form a uniform chromatin lattice, could drive the observed higher-order structure of centromeric chromatin that is present in mitotically condensed chromosomes. Such interactions could operate in interphase and mitosis but might be reinforced by the assembly of the mitotic kinetochore to stabilize the centromeric chromatin against the forces applied during mitotic chromosome segregation.

Concluding remarks

The analysis of the genomic sequences of various eukaryotes has revolutionized our understanding of DNA-sequence-dependent control of chromosome architecture and function; however, many aspects of chromosome biology are controlled by sequence-independent epigenetic phenomena. In particular, the organization of functional chromosomal domains required for biological processes such as dosage compensation, gene silencing and chromosome segregation is governed by epigenetic factors. The centromere provides an excellent experimental example for exploring the epigenetic control of chromatin specialization. Important aspects of centromere formation, such as how the histone variant CENP-A is uniquely targeted to the centromere and how centromeric chromatin directs kinetochore assembly, have yet to be elucidated. Genetic analysis has identified many proteins that influence centromere formation, and biochemical analysis has uncovered fundamental principles of chromatin assembly. Our current challenge is to understand the molecular mechanisms and epigenetic regulation of centromere formation and how these general principles of chromosome organization relate to chromosome missegregation and aneuploidy.

References

1 Chan, G.K., *et al.* (2005) Kinetochore structure and function. *Trends Cell Biol.*
2 Cleveland, D.W. *et al.* (2003) Centromeres and kinetochores: from epigenetics to mitotic checkpoint signaling. *Cell* 112, 407–421
3 Maiato, H. *et al.* (2004) The dynamic kinetochore–microtubule interface. *J. Cell Sci.* 117, 5461–5477
4 Beadle, G.W. (1932) A possible influence of the spindle fibre on crossing-over in *Drosophila*. *Proc. Natl. Acad. Sci. U. S. A.* 18, 160–165
5 Wood, V. *et al.* (2002) The genome sequence of *Schizosaccharomyces pombe*. *Nature* 415, 871–880
6 Blackwell, C. *et al.* (2004) The *Schizosaccharomyces pombe* HIRA-like protein Hip1 is required for the periodic expression of histone genes and contributes to the function of complex centromeres. *Mol. Cell Biol.* 24, 4309–4320
7 Clarke, L. and Carbon, J. (1980) Isolation of a yeast centromere and construction of functional small circular chromosomes. *Nature* 287, 504–509
8 Hahnenberger, K.M. *et al.* (1989) Construction of functional artificial minichromosomes in the fission yeast *Schizosaccharomyces pombe*. *Proc. Natl. Acad. Sci. U. S. A.* 86, 577–581
9 Struhl, K. *et al.* (1979) High-frequency transformation of yeast: autonomous replication of hybrid DNA molecules. *Proc. Natl. Acad. Sci. U. S. A.* 76, 1035–1039
10 Sun, X. *et al.* (2003) Sequence analysis of a functional *Drosophila* centromere. *Genome Res.* 13, 182–194
11 Sun, X. *et al.* (1997) Molecular structure of a functional *Drosophila* centromere. *Cell* 91, 1007–1019
12 Schueler, M.G. *et al.* (2001) Genomic and genetic definition of a functional human centromere. *Science* 294, 109–115
13 Haaf, T. *et al.* (1992) Integration of human α-satellite DNA into simian chromosomes: centromere protein binding and disruption of normal chromosome segregation. *Cell* 70, 681–696
14 Harrington, J.J. *et al.* (1997) Formation of de novo centromeres and construction of first-generation human artificial microchromosomes. *Nat. Genet.* 15, 345–355
15 Masumoto, H. *et al.* (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
16 Ohzeki, J. *et al.* (2002) CENP-B box is required for de novo centromere chromatin assembly on human alphoid DNA. *J Cell Biol.* 159, 765–775
17 Hudson, D.F. *et al.* (2003) Condensin is required for nonhistone protein assembly and structural integrity of vertebrate mitotic chromosomes. *Dev. Cell* 5, 323–336
18 Kapoor, M. *et al.* (1998) The cenpB gene is not essential in mice. *Chromosoma* 107, 570–576
19 Voullaire, L.E. *et al.* (1993) A functional marker centromere with no detectable α-satellite, satellite III, or CENP-B protein: activation of a latent centromere? *Am. J. Hum. Genet.* 52, 1153–1163
20 Barry, A.E. *et al.* (1999) Sequence analysis of an 80 kb human neocentromere. *Hum. Mol. Genet.* 8, 217–227
21 Slater, H.R. *et al.* (1993) Confirmation of trisomy 22 in two cases using chromosome painting: comparison with t(11;22). *Am. J. Med. Genet.* 46, 434–437

22 Blower, M.D. and Karpen, G.H. (2001) The role of *Drosophila* CID in kinetochore formation, cell-cycle progression and heterochromatin interactions. *Nat. Cell Biol* 3, 730–739

23 Buchwitz, B.J. *et al*. (1999) A histone-H3-like protein in *C. elegans*. *Nature* 401, 547–548

24 Howman, E.V. *et al*. (2000) Early disruption of centromeric chromatin organization in centromere protein A (Cenpa) null mice. *Proc Natl Acad Sci. U. S. A.* 97, 1148–1153

25 Meluh, P.B. *et al*. (1998) Cse4p is a component of the core centromere of *Saccharomyces cerevisiae*. *Cell* 94, 607–613

26 Regnier, V. *et al*. (2005) CENP-A is required for accurate chromosome segregation and sustained kinetochore association of BubR1. *Mol. Cell Biol.* 25, 3967–3981

27 Takahashi, K. *et al*. (2000) Requirement of Mis6 centromere connector for localizing a CENP-A-like protein in fission yeast. *Science* 288, 2215–2219

28 Amor, D.J. *et al*. (2004) Building the centromere: from foundation proteins to 3D organization. *Trends Cell Biol.* 14, 359–368

29 Fukagawa, T. *et al*. (2001) CENP-H, a constitutive centromere component, is required for centromere targeting of CENP-C in vertebrate cells. *EMBO J.* 20, 4603–4617

30 Nishihashi, A. *et al*. (2002) CENP-I is essential for centromere function in vertebrate cells. *Dev. Cell* 2, 463–476

31 Tomkiel, J. *et al*. (1994) CENP-C is required for maintaining proper kinetochore size and for a timely transition to anaphase. *J. Cell Biol.* 125, 531–545

32 Cam, H.P. *et al*. (2005) Comprehensive analysis of heterochromatin- and RNAi-mediated epigenetic control of the fission yeast genome. *Nat. Genet.* 37, 809–819

33 Chueh, A.C. *et al*. (2005) Variable and hierarchical size distribution of L1-retroelement-enriched CENP-A clusters within a functional human neocentromere. *Hum. Mol. Genet.* 14, 85–93

34 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

35 Grewal, S.I. and Moazed, D. (2003) Heterochromatin and epigenetic control of gene expression. *Science* 301, 798–802

36 Bernard, P. *et al*. (2001) Requirement of heterochromatin for cohesion at centromeres. *Science* 294, 2539–2542

37 Nonaka, N. *et al*. (2002) Recruitment of cohesin to heterochromatic regions by Swi6/HP1 in fission yeast. *Nat. Cell Biol.* 4, 89–93

38 Grimes, B.R. *et al*. (2004) Assembly and characterization of heterochromatin and euchromatin on human artificial chromosomes. *Genome Biol.* 5, R89

39 Saffery, R. *et al*. (2000) Human centromeres and neocentromeres show identical distribution patterns of >20 functionally important kinetochore-associated proteins. *Hum. Mol. Genet.* 9, 175–185

40 Guenatri, M. *et al*. (2004) Mouse centric and pericentric satellite repeats form distinct functional heterochromatin. *J. Cell Biol.* 166, 493–505

41 Allshire, R.C. (1995) Elements of chromosome structure and function in fission yeast. *Semin. Cell Biol.* 6, 55–64

42 Fukagawa, T. *et al*. (2004) Dicer is essential for formation of the heterochromatin structure in vertebrate cells. *Nat. Cell Biol.* 6, 784–791

43 Noma, K. *et al*. (2001) Transitions in distinct histone H3 methylation patterns at the heterochromatin domain boundaries. *Science* 293, 1150–1155

44 Takahashi, K. *et al*. (1992) A low copy number central sequence with strict symmetry and unusual chromatin structure in fission yeast centromere. *Mol. Biol. Cell* 3, 819–835

45 Takahashi, K. *et al*. (1991) A large number of tRNA genes are symmetrically located in fission yeast centromeres. *J. Mol. Biol.* 218, 13–17

46 Meneghini, M.D. *et al*. (2003) Conserved histone variant H2A.Z protects euchromatin from the ectopic spread of silent heterochromatin. *Cell* 112, 725–736

47 Vafa, O. and Sullivan, K.F. (1997) Chromatin containing CENP-A and α-satellite DNA is a major component of the inner kinetochore plate. *Curr. Biol.* 7, 897–900

48 Warburton, P.E. *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

49 Shelby, R.D. *et al*. (1997) Assembly of CENP-A into centromeric chromatin requires a cooperative array of nucleosomal DNA contact sites. *J. Cell. Biol.* 136, 501–513

50 Keith, K.C. *et al*. (1999) Analysis of primary structural determinants that distinguish the centromere-specific function of histone variant Cse4p from histone H3. *Mol. Cell Biol.* 19, 6130–6139

51 Malik, H.S. *et al*. (2002) Recurrent evolution of DNA-binding motifs in the *Drosophila* centromeric histone. *Proc. Natl. Acad. Sci. U. S. A.* 99, 1449–1454

52 Sullivan, K.F. *et al*. (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

53 Black, B.E. *et al*. (2004) Structural determinants for generating centromeric chromatin. *Nature* 430, 578–582

54 Goshima, G. *et al*. (2003) Human centromere chromatin protein hMis12, essential for equal segregation, is independent of CENP-A loading pathway. *J. Cell Biol.* 160, 25–39

55 Goshima, G. *et al*. (1999) Proper metaphase spindle length is determined by centromere proteins Mis12 and Mis6 required for faithful chromosome segregation. *Genes Dev.* 13, 1664–1677

56 Mellone, B.G. and Allshire, R.C. (2003) Stretching it: putting the CEN(P-A) in centromere. *Curr Opin. Genet. Dev.* 13, 191–198

57 Chen, E.S. *et al*. (2003) A cell cycle-regulated GATA factor promotes centromeric localization of CENP-A in fission yeast. *Mol. Cell* 11, 175–187

58 Hayashi, T. *et al*. (2004) Mis16 and Mis18 are required for CENP-A loading and histone deacetylation at centromeres. *Cell* 118, 715–729

59 Jin, Q.W. *et al*. (2002) The mal2p protein is an essential component of the fission yeast centromere. *Mol. Cell Biol.* 22, 7168–7183

60 Pidoux, A.L. *et al*. (2003) Sim4: a novel fission yeast kinetochore protein required for centromeric silencing and chromosome segregation. *J. Cell Biol.* 161, 295–307

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

62 Measday, V. *et al*. (2002) Ctf3p, the Mis6 budding yeast homolog, interacts with Mcm22p and Mcm16p at the yeast outer kinetochore. *Genes Dev.* 16, 101–113

63 Tagami, H. *et al*. (2004) Histone H3.1 and H3.3 complexes mediate nucleosome assembly pathways dependent or independent of DNA synthesis. *Cell* 116, 51–61

64 Kobor, M.S. *et al*. (2004) A protein complex containing the conserved Swi2/Snf2-related ATPase Swr1p deposits histone variant H2A.Z into euchromatin. *PLoS Biol* 2, E131

65 Krogan, N.J. *et al*. (2003) A Snf2 family ATPase complex required for recruitment of the histone H2A variant Htz1. *Mol. Cell* 12, 1565–1576

66 Mizuguchi, G. *et al*. (2004) ATP-driven exchange of histone H2AZ variant catalyzed by SWR1 chromatin remodeling complex. *Science* 303, 343–348

67 Walfridsson, J. *et al*. (2005) The CHD remodeling factor Hrp1 stimulates CENP-A loading to centromeres. *Nucleic Acids Res.* 33, 2868–2879

68 Zinkowski, R.P. *et al*. (1991) The centromere–kinetochore complex: a repeat subunit model. *J. Cell Biol.* 113, 1091–1110

69 Blower, M.D. *et al*. (2002) Conserved organization of centromeric chromatin in flies and humans. *Dev. Cell* 2, 319–330

70 Ono, T. *et al*. (2003) Differential contributions of condensin I and condensin II to mitotic chromosome architecture in vertebrate cells. *Cell* 115, 109–121

71 Hagstrom, K.A. *et al*. (2002) C. elegans condensin promotes mitotic chromosome architecture, centromere organization, and sister chromatid segregation during mitosis and meiosis. *Genes Dev.* 16, 729–742

72 Stear, J.H. and Roth, M.B. (2002) Characterization of HCP-6, a C. elegans protein required to prevent chromosome twisting and merotelic attachment. *Genes Dev.* 16, 1498–1508

73 Wignall, S.M. *et al*. (2003) The condensin complex is required for proper spindle assembly and chromosome segregation in Xenopus egg extracts. *J. Cell Biol.* 161, 1041–1051
74 Dej, K.J. *et al*. (2004) Mutations in the *Drosophila* condensin subunit dCAP–G: defining the role of condensin for chromosome condensation in mitosis and gene expression in interphase. *Genetics* 168, 895–906
75 Savvidou, E. *et al*. (2005) Drosophila CAP-D2 is required for condensin complex stability and resolution of sister chromatids. *J. Cell Sci.* 118, 2529–2543
76 Fan, J.Y. *et al*. (2002) The essential histone variant H2A.Z regulates the equilibrium between different chromatin conformational states. *Nat. Struct. Biol.* 9, 172–176
77 Dorigo, B. *et al*. (2004) Nucleosome arrays reveal the two-start organization of the chromatin fiber. *Science* 306, 1571–1573
78 Huynh, V.A. *et al*. (2005) A method for the in vitro reconstitution of a defined "30 nm" chromatin fibre containing stoichiometric amounts of the linker histone. *J. Mol. Biol.* 345, 957–968
79 Dorigo, B. *et al*. (2003) Chromatin fiber folding: requirement for the histone H4 N-terminal tail. *J. Mol. Biol.* 327, 85–96