Centromerization

K.H. Andy Choo

Centromere formation is a complex process that involves the packaging of DNA into a centromere-unique chromatin, chemical modification and the seeding of kinetochore and associated proteins. The early steps in this process, in which a chromosomal region is marked for centromerization (that is, to become resolutely committed to centromere formation), are unusual in that they can apparently occur in a DNA-sequence-independent manner. Current evidence indicates the involvement of epigenetic influences in these early steps. A number of epigenetic mechanisms that can affect centromere chromatin organization have been proposed. Here, the characteristics of these mechanisms and their relative roles as possible primary triggers for centromerization are discussed in the light of recent data.

The centromere is a well-known landmark on the chromosome. Ultrastructurally, it takes the form of a distinct primary constriction on the condensed metaphase chromosome of a higher eukaryote. The constricted region comprises a differentiated chromatin structure consisting of DNA and protein complexes (the kinetochores) onto which microtubules bind to effect proper chromosome movements. Recent literature has extensively reviewed the DNA and protein constituents of the centromere[1–8]. Molecularly, except in a small number of cases such as the well-defined 125-bp 'point' centromere of the budding yeast or 'regional' centromere of a non-essential 420-kb *Drosophila* minichromosome, the precise location of the centromere is often not well defined. This is because of current technical difficulty in delineating the functional limits of the *cis*-acting DNA sequences of a centromere. For the present discussion, the centromere is defined as the chromosomal region containing the DNA that provides the core centromere activity and might include some related DNA sequences whose centromere role might or might not be currently obvious (Box 1). The term 'pericentric' is used to describe a broader chromosomal region that includes both the centromere and flanking DNA, generally of unrelated sequences or functions. As we gain knowledge of the complex and fascinating properties of the centromeric DNA and its protein components, the rudimentary question of how centromerization of a chromosomal domain is achieved remains largely unanswered.

Andy Choo is at The Murdoch Institute, Royal Children's Hospital, Flemington Road, Parkville 3052, Melbourne, Australia.
E-mail: choo@cryptic.rch.unimelb.edu.au

DNA sequence requirement and epigenetic control

Four observations suggest that specific DNA sequences are not required for centromerization, strongly implicating the influence of epigenetic mechanisms[1,9–12]. First, the centromere of a higher eukaryote typically contains hundreds to thousands of kilobases of some tandemly repeated DNA sequences. Analysis of these repeats in organisms from insects, plants and fungi to mammals and other vertebrates reveals no obvious nucleotide sequence conservation, suggesting that a universal centromere sequence does not exist, or has at least defied current recognition (Fig. 1a). Second, in mitotically stable human isodicentric chromosomes containing two identical, well-separated regions of centromeric α-satellite DNA, only one active centromere is formed, suggesting that α-satellite alone is not sufficient for centromerization and that previously active centromeres can be silenced[13] (Fig. 1b). On such isodicentric chromosomes, reducing the distance separating the two centromeres (for example, to <10 Mb) results in both centromeres being active, suggesting that centromere activities on these chromosomes are controlled epigenetically rather than by mutational changes. Third, the discovery that functional neocentromeres can be formed from normally non-centromeric DNA on human and *Drosophila* chromosomes indicates that kinetochores can be assembled at a variety of genomic DNA sequences other than the usual centromeric satellite repeats (Fig. 1c). Direct sequencing of the core functional region of a neocentromere and comparison with the progenitor DNA sequence from which the neocentromere DNA has descended reveals 100% identity and provides direct support for a non-mutation-driven, epigenetic mechanism of neocentromerization (A.E. Barry and K.H.A. Choo, unpublished). Fourth, in the fission yeast *Schizosaccharomyces pombe*, truncated centromeric plasmids, which generally do not form functional centromeres, can infrequently assemble an active centromere upon transformation into cells, suggesting a sequence-independent centromerization mechanism (Fig. 1d).

Although there is now little doubt that many different DNA sequences are capable of nucleating centromere formation, it is less clear what bearing the primary nucleotide sequence actually has on this process. The fact that in all organisms studied to date one or a small number of specific primary centromere sequences predominate[1] suggests a biological preference for certain types of DNA. It is possible that only DNA with certain primary sequences that allow the formation of special secondary, tertiary or scaffold structures can become centromeres. A recent study comparing nucleotide sequences *in silico* has demonstrated common structural features between the apparently unrelated primary sequences of α-satellite, a human neocentromere and the centromere of the budding yeast *Saccharomyces cerevisiae*[14]. Such features include double dyad symmetries of a particular size and short conserved base motifs adjacent to these dyad symmetries. It would be of great interest to determine whether these sequence features provide a structural code for the normal centromere DNA of other species and for the DNA sequences of other neocentromeres.

Chromatin modification

If centromerization does not depend on a strict primary DNA sequence, it presumably occurs through

0962-8924/00/$ – see front matter
PII: S0962-8924(00)01739-6

modifications at the chromatin level. The existence of specialized centromere chromatin is clear from several lines of evidence. In *S. pombe*, the central centromeric domain forms an unusual chromatin structure essentially devoid of typical nucleosomal packaging. The Mis6 protein has been shown to play a crucial role in this chromatin structure since mutations in the *mis6* gene disrupt this structure, producing normal nucleosomal arrays and impairing equal chromosomal segregation[9]. Similarly, an especially compact, nuclease-resistant chromatin structure containing the histone H3-related protein Cse4p resides over a 220–250-bp region of the *S. cerevisiae* centromere[1]. The detection of a conserved homologue of Cse4p, CENP-A, in higher eukaryotic centromeres suggests the existence of a similar centromere-specific chromatin in diverse organisms.

A number of cellular mechanisms are known to influence chromatin dynamics and could therefore be directly involved in centromerization. These mechanisms are summarized in Fig. 2. The following discussion will focus on recent data from a number of centromere DNA-binding proteins.

BOX 1 – GLOSSARY OF TERMS

Acentric – lacking a centromere.

Centromere – chromosomal domain containing a functional centromere (defined by primary constriction, kinetochore–protein association, microtubule binding and/or segregation properties). The boundary of this domain is often not well-defined and might contain DNA of related sequences that do not directly participate in kinetochore formation (e.g. subsets of human α-satellite DNA).

Chromatin – higher-order complex of nucleic acids (DNA and RNA) and proteins (histones and non-histones) constituting the eukaryotic chromosomes.

Dicentric – containing two active centromeres.

Epigenetic mechanism – mechanism imprinting an alternative, heritable phenotypic expression of a DNA, occurring without detectable changes in primary nucleotide sequence.

Heterochromatin – chromosomal region showing maximal condensation in the nucleus during interphase, classically defined by tandemly repeated DNA, generally late replicating and transcriptionally suppressed.

Higher-order structure – secondary or tertiary structures of a DNA produced by bending, looping, folding, etc. and/or through association with proteins (chromatin structure).

Holocentric – consisting of kinetochores dispersed along a chromosome, attaching to spindles at many sites along its length rather than at a single differentiated point.

Isodicentric – consisting of two identical centromeres; found on isochromosomes with mirror-image arms.

Kinetochore – DNA–protein complex cuffing the two faces of an active centromere that point towards the spindle poles, providing opposite attachment sites for microtubules.

Major satellite – mouse pericentric DNA, probably not directly involved in centromere function.

Minor satellite – probably the functional DNA of the mouse centromere, unrelated in sequence to α-satellite except for the 17-bp CENP-B-binding motifs.

Neocentromere – ectopic centromere formed at a location other than normal centromeres. Human neocentromeres contain no α-satellite sequences, confer mitotic functions and show patterns of kinetochore-associated proteins identical to those of normal centromeres[10,54].

Nucleosome – core component of eukaryotic chromosomes; contains histones H2A, H2B, H3 and H4 wrapped by a DNA segment.

Pseudodicentric – containing two centromere DNA domains, of which only one is functionally active.

α-Satellite – human centromeric repeats, only a subset of which appears to be involved in centromere function.

Centromere DNA-binding proteins

Direct sequestration of a centromere-specific, DNA-binding protein could be a very effective way to mark a chromosomal domain for centromerization. Five proteins – CENP-A, CENP-B, CENP-C, CENP-G and CENP-H – are unique and intrinsic to the centromere[1,4,38]. Three of these proteins – CENP-A, CENP-B and CENP-C – have demonstrated DNA-binding activity and are potential candidates for centromerization marking. CENP-A is a histone H3-like protein that is conserved in mammals, *Caenorhabditis elegans* and *S. cerevisiae*[1,39]. In mammals and *S. cerevisiae*, CENP-A binding is confined to a defined centromeric region on the chromosome, whereas, in the nematode, in which the chromosomes are holocentric (Box 1), CENP-A binding occurs throughout the length of a chromosome[39]. Two observations indicate that uncoupling CENP-A expression from histone H3 expression in S phase is important in determining the function of CENP-A[33]. First, CENP-A synthesis is associated with centromere replication during mid-S to early G2 phase, whereas histone H3 expression peaks early in S phase. Second, expression of CENP-A under a histone H3 promoter fails to localize the protein at the normal centromere or seed additional illegitimate centromeres. The observed later-than-histone H3 expression timing of CENP-A has led to the hypothesis that a late-replicating region of a chromosome binds CENP-A preferentially and results in the assembly of a centromere (Fig. 2c).

Cenpa (symbol for mouse *CENP-A*) gene knockout[40] and *hcp-3* (*CENP-A* homologue in *C. elegans*) gene suppression[39] results in severe mitotic segregation problems and early embryonic death. Of particular interest, immunofluorescence analysis of *Cenpa* knockout mice reveals dispersion of unsequestered Cenpb and Cenpc (for mouse CENP-B and CENP-C, respectively) proteins throughout the interphase cell nucleus, formation of morphologically slightly fuzzy but otherwise discrete Cenpb foci on centromeres and the absence of centromeric Cenpc. These results suggest that (1) Cenpa is not essential for the localization of Cenpb at the centromere, although its presence appears necessary for the ultimate compactness

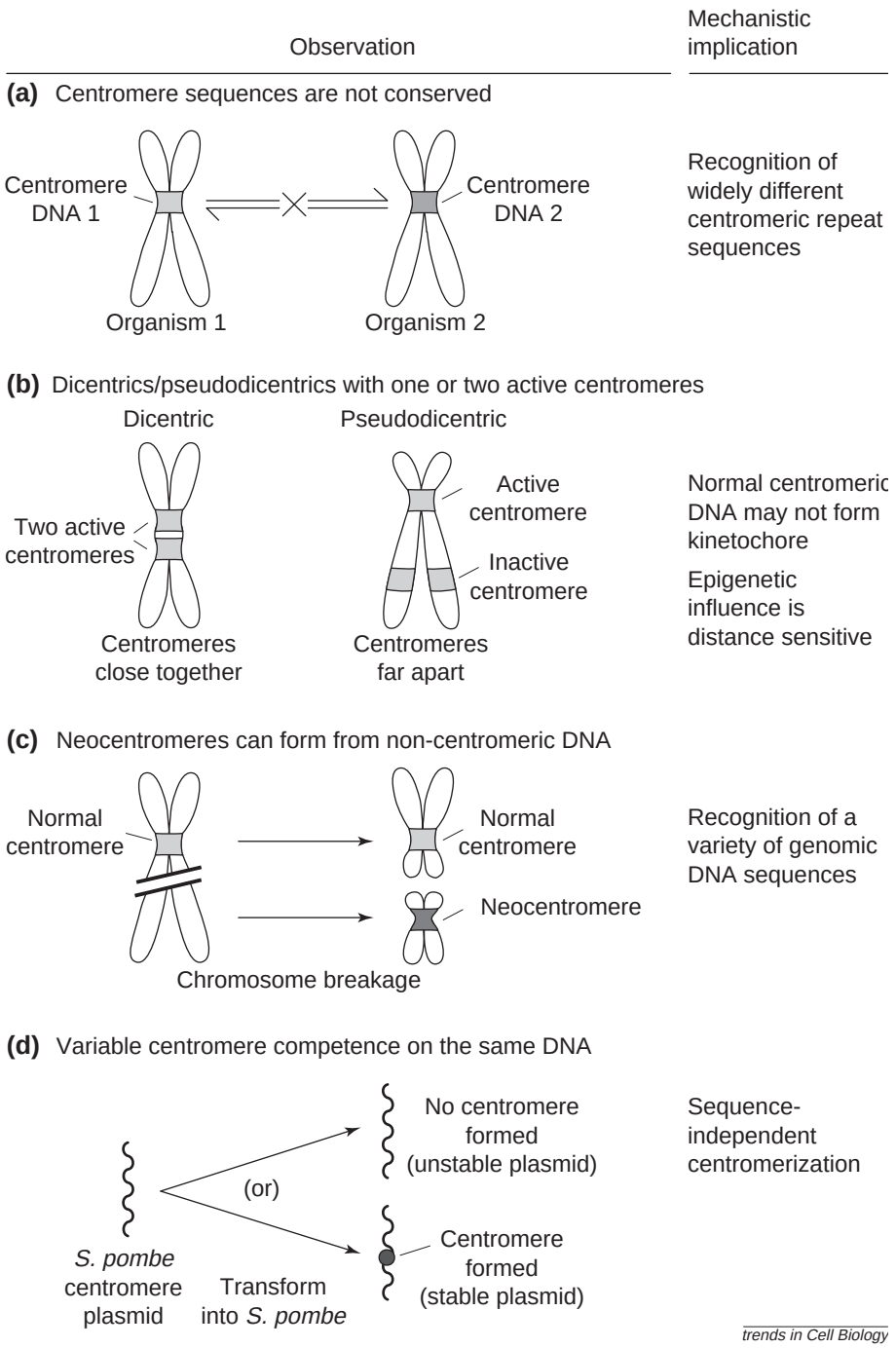

FIGURE 1

Observations implicating an epigenetic control model for centromerization. (a) Absence of a conserved centromere DNA sequence in different organisms [and within an organism; see (c)] suggests a mechanism for centromerization that can recognize many different sequences. (b) In isodicentric chromosomes, the presence of two functionally active centromeres when the centromeres are close together, but not when they are far apart, suggests that centromerization is epigenetically controlled and the possibility that the control mechanism can spread over a distance beyond the core centromere region (see Fig. 2d). (c) Neocentromere formation at many different non-centromeric sites suggests a versatility of centromerization that is independent of both the primary nucleotide sequence and the highly repetitive nature of the normal centromeric DNA. (d) Variable expression of centromere competence from a plasmid construct containing truncated *Schizosaccharomyces pombe* centromere DNA suggests an epigenetic control mechanism that is not dependent on the underlying DNA sequence.

of the assembled structure, and (2) Cenpa is essential for the kinetochore targeting of Cenpc, presumably by providing an emerging centromere-specific chromatin structure through its early nucleosomal organizing role at interphase[40]. In another study, Van Hooser *et al.*[41] demonstrated that, whereas the inner centromere domain of a hamster chromosome contains an interspersion of CENP-B with phosphorylated histone H3, normal histone H3 phosphorylation is absent from the CENP-A-containing kinetochore subdomain. These studies suggest a distinct organizational control for the kinetochore and inner centromere domain, and a key role of CENP-A in promoting kinetochore formation (Fig. 3).

CENP-B localizes to the human and mouse centromeres through direct binding to a 17-bp CENP-B box sequence found in the human α-satellite and mouse minor satellite DNA[1]. The localization of this protein on both active and inactive centromeres in pseudo-dicentric and -multicentric chromosomes suggests that deposition of this protein is insufficient for centromerization. The lack of an obvious mitotic and meiotic phenotype in *Cenpb* knockout mice[43], and the absence of this protein on active human neocentromeres and the normal centromere of the Y chromosome, further suggest that the protein is dispensable or functionally redundant. What then is the role of CENP-B? One possibility is to organize centromeric satellite DNA[1] (Fig. 3). A recent study has described a severe reproductive phenotype involving defective uterine epithelial cells in some *Cenpb* knockout mouse strains[43]. Further investigation of this unusual phenotype might help to shed some light on the role of this protein *in vivo*.

Gene disruption experiments demonstrate that CENP-C is essential for proper mitotic segregation and cell survival[44,45]. Several studies have attempted to investigate the role of CENP-C in centromere assembly. *In vivo* expression of CENP-C truncation mutants demonstrates that the protein has an autonomous centromere-targeting domain which, as shown by *in vitro* assays, also possesses DNA-binding property[42]. This DNA-binding property features a generalized affinity of CENP-C for DNA. However, it remains to be seen whether some specific centromere-associated sequences or sequence conformations are linked to the DNA-binding activity of this protein. When overexpressed in chicken DT40 cells using a heterologous promoter, the protein coats the chromosome arms but does not seed new centromeres or disrupt the native centromeres[46], indicating that overexpression of CENP-C alone is insufficient to precipitate centromerization. However, it is possible that, for centromere nucleation to occur,

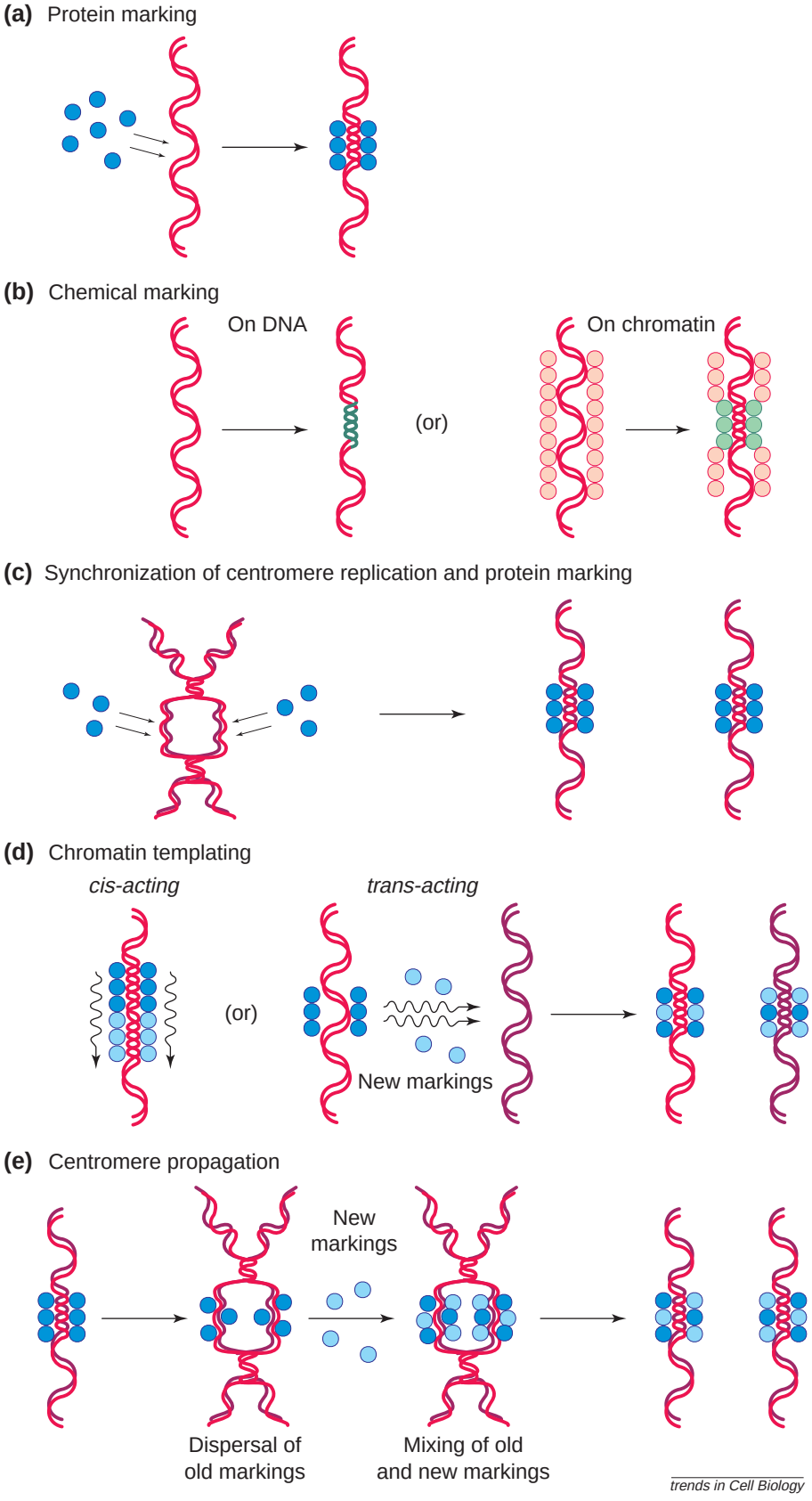

FIGURE 2

Epigenetic mechanisms that can potentially alter chromatin structure and facilitate centromerization. (a) Sequestration of a centromere DNA-binding protein (e.g. CENP-A, CENP-C, PARP; discussed in the text) might result in higher-order conformational changes that are necessary for centromerization. (b) Chromatin remodelling can be achieved through chemical modification of centromeric DNA or its associated histones and non-histone proteins[9,12,15]. (1) Methylation. Molecular analysis indicates that human and mouse pericentric heterochromatin is enriched in 5-methylcytosine or the methyl-CpG-binding protein (MeCP2), with variable methylation levels detected on the α-satellite DNA of different human chromosomes[16–18]. (2) Acetylation. Studies on the acetylation status of the centromeres and/or pericentric regions of *Schizosaccharomyces pombe*[19], humans[20–22] and *Drosophila*[23] have demonstrated underacetylation of histones H3 and/or H4, suggesting that the assembly of a functional centromere is imprinted in the hypoacetylated state of centromeric chromatin. Recent studies on the mechanism of gene silencing indicate that the methyl-CpG-binding proteins MeCP2, MBD2 and MBD3 exist as complexes with the MeCP2-Sin3, MeCP1 and Mi-2 histone deacetylase complexes, respectively, providing a causal link between DNA methylation, histone deacetylation and chromatin modification that might be relevant to the regulation of centromere activity[24–26]. (3) Phosphorylation, in particular that of histones H1 and H3, is known to be required for proper segregation and condensation of chromosomes during mitosis and meiosis[15,27–30]. Phosphorylation appears to begin in the pericentric heterochromatin during G2, progresses sequentially throughout the genome in association with chromosome condensation, persists in chromatin until anaphase and disappears following mitosis. It is not known whether the specialized centromeric H3 variant, CENP-A, is directly phosphorylated, although comparison of potential phosphorylation sites between the amino acid sequences of normal H3 and CENP-A raises the possibility that the latter also becomes phosphorylated during mitosis[15] (see text and Fig. 3). (4) Poly(ADP-ribosyl)ation. Although a recent study has localized the catalytic enzyme poly(ADP-ribose) polymerase (PARP) to the mammalian centromere[31], it remains to be determined whether this enzyme carries out the expected poly(ADP-ribosyl)ation reaction at the centromere (see text). (4) Ubiquitination. Not much is known about the role of this modification at the centromere, although a recent study[32] speculates that the function of CENP-C might be regulated by SUMO-1 or a similar ubiquitin-like protein (see text). (c) Synchronization of the timing of centromere replication and the expression/deposition of a centromere-specific protein onto nascent DNA duplexes has been proposed to provide a sequence-independent mechanism of centromerization[33,34]. The cell-cycle expression profile and histone H3-like property of CENP-A makes this protein a good contender for this role. (d) A pre-existing centromere chromatin might provide the template for the *cis*- or *trans*-spreading of a centromere-active factor to imprint adjacent or remote chromosomal regions with centromere properties[12,35–37]. (e) The chromatin templating model described in (d) can be adapted to explain the autonomous propagation of a centromere state during chromosomal replication. A key step in this scheme requires the initial dispersal of old centromere markings onto nascent daughter duplexes to provide bridging imprints of centromere properties for the recruitment of freshly synthesized markings[28].

CENP-C might need to be in stoichiometric equilibrium with some other proteins, or that the timing of expression under the heterologous promoter might have been wrong. It is known that the expression of CENP-C under its normal promoter varies during the cell cycle, rising progressively from S phase through G2 and mitosis, with the increase peaking at G1[47]. How this expression profile affects the role of CENP-C at the centromere is unclear, but the observation has led to the proposal that, in addition to a role in mitosis, CENP-C has a further role that might be related to cell-cycle control in the G1 phase.

In another study, specific degradation of CENP-C is shown to occur following binding of the herpes simplex virus regulatory protein Vmw110 to the centromere[32]. Since Vmw110 can induce the loss of a number of undefined proteins conjugated to ubiquitin-like proteins such as SUMO-1, it is speculated that, through modifications by SUMO-1 or a similar ubiquitin-like protein, CENP-C might in some way be involved in epigenetic marking of the active centromere[32]. Some support for this speculation comes from the suppression of temperature-sensitive mutations in MIF2, the *S. cerevisiae* homologue of CENP-C, by overexpression of the budding yeast homologue of SUMO-1, Smt-3.

In addition to the above proteins, a number of non-histone chromosomal proteins acting at the DNA/chromatin level have been localized to the pericentric regions of chromosomes[1,9,48]. These include TopoII (DNA topoisomerase II, which is involved in chromosome condensation and sister chromatid disjunction), HMG-I (high-mobility group protein-I, which is involved in nucleosomal positioning or phasing through binding of A:T nucleotide tracks), HP-1 (heterochromatin protein-1, which is involved in transcriptional repression), SU(VAR) proteins (suppressors of variegation, which are involved in transcriptional repression) and PARP [poly(ADP-ribose) polymerase, which is involved in ADP ribosylation]. These proteins are centrally involved in the regional organization of chromatin and the regulation of the activities of chromatin-dependent proteins or genes and might have important roles in the regulation of centromerization or centromere activity. Of particular interest is PARP, since this protein shows preferential affinity-binding to a 9-bp double-stranded sequence found in human α-satellite DNA and to an unrelated 28-bp tandem repeat found in a human neocentromere[31]. Immunofluorescence studies have demonstrated the localization of this protein to the normal centromere of different mammals, as well as to human neocentromeres and the active, but not the inactive, centromere of a dicentric chromosome. Absence of binding to the α-satellite DNA of the inactive centromere of a human dicentric chromosome, and to the q12 heterochromatic region of the Y chromosome, suggests that neither sequence nor heterochromatic state are sufficient on their own for PARP binding. In the mouse, PARP binding often extends into pericentromeric regions (Fig. 4), suggesting the possibility that sequestration of PARP to a preferred centromeric site (minor satellite DNA) might facilitate the 'spreading' of the protein to a surrounding less-preferred site (major satellite DNA) (see Fig. 2d; *cis*-spreading of centromere marking factor).

What is the role of PARP at the centromere? Previous studies have identified PARP as a multifunctional and highly conserved nuclear enzyme[49]. The enzyme catalyses the transfer of ADP-ribose moieties from its substrate NAD^+ to protein acceptors (heteromodification) or to the enzyme itself (automodification) in a poly(ADP-ribosyl)ation reaction that is both ubiquitous in higher eukaryotes and postulated to play a role in central nuclear processes, such as DNA repair, differentiation, transcriptional regulation, antirecombination and genome stability, modulation of chromosomal scaffold structure and function, and apoptosis. DNA binding appears to be required for maximal stimulation of basal PARP activity, with the enzyme showing preferential binding to single-stranded DNA breaks (in a sequence-independent manner), supercoiled DNA, DNA loops and cruciform structures. In addition to DNA binding, PARP interacts with and modifies the functions of many target proteins, including histones, topoisomerases, low- and high-mobility group (HMG) proteins, DNA polymerases and ligases[49]. The potential role of PARP in regulating chromosomal structure and function is highlighted by two studies, one identifying a PARP homologue (tankyrase) and ADP-ribosylation activity at human telomeres[50], the other demonstrating that PARP-deficient mouse cells display telomere shortening and severe chromosomal instability characterized by increased chromosome fusion and aneuploidy[51]. The observation that at least some of the PARP knockout mouse strains appear mitotically and meiotically normal further implies functional redundancy of the protein[52].

The strong binding of PARP to different centromere DNA sequences described by Earle *et al.*[31] suggests the possibility of a versatile mechanism of mammalian centromere recognition that is independent of primary DNA sequences. Can PARP bring about centromerization or other specific changes to a centromere? Does centromeric PARP binding merely reflect recruitment of the enzyme by DNA damage resulting from stress forces placed on DNA/chromatin at active centromeres, especially during mitosis? Answers to these questions might come in part from future studies to determine whether increased PARP binding is accompanied by an increase in poly(ADP-ribosyl)ation at the centromere or whether PARP has a completely different role at the centromere. Overexpression of a full-length PARP in cells might allow testing of whether ectopic centromere formation (that is, centromerization of normally non-centromeric DNA) occurs. A detailed assessment of the distribution pattern of PARP on the centromere during the cell cycle would also be helpful in establishing the requirement of this protein at different stages of the cell cycle.

Choosing between the different mechanisms

Centromerization normally concerns the propagation of an existing centromere, during which newly replicated daughter DNA duplexes are marked and assembled into functional centromeres. Only rarely,

such as during chromosomal rearrangements leading to neocentromere formation, does the need to centromerize a normally non-centromeric DNA arise. The centromerization requirements for normal centromere propagation and for neocentromere formation are likely to be different at the outset since the former can work from a pre-existing template, whereas the latter has to create one. It is uncertain how a chromosomal site is chosen for neocentromerization. Since repeated occurrences of neocentromeres have been described on some human chromosomes[10], in particular the long arms of chromosomes 13 and 15, careful mapping should indicate whether these neocentromeric sites occur at random, or whether some preferred sites or hotspots exist. The recent report that centromere competence is innate to satellite-containing blocks of heterochromatin, demonstrated by experimentally detaching such heterochromatin blocks from a natural centromere[53], raises the possibility that neocentromere competence might be similarly innate to chromosomal regions with heterochromatin-like properties. The observation that most neocentromeres are located at apparently euchromatic regions (except those at Yq12)[10] suggests at first that neocentromerization is not dependent on (or exclusive of) satellite-containing blocks of heterochromatin. However, the detection of the heterochromatin-associated protein M31 (murine homologue of HP-1) at non-heterochromatic neocentromeres[54] indicates that some aspects of heterochromatin are found at these neocentromeres. Future studies could determine what potential roles these heterochromatin-like properties might have at the neocentromere. For example, are they responsible for a possible delay in the replication timing of neocentromere domains, as predicted by the late-replication model[33,34]?

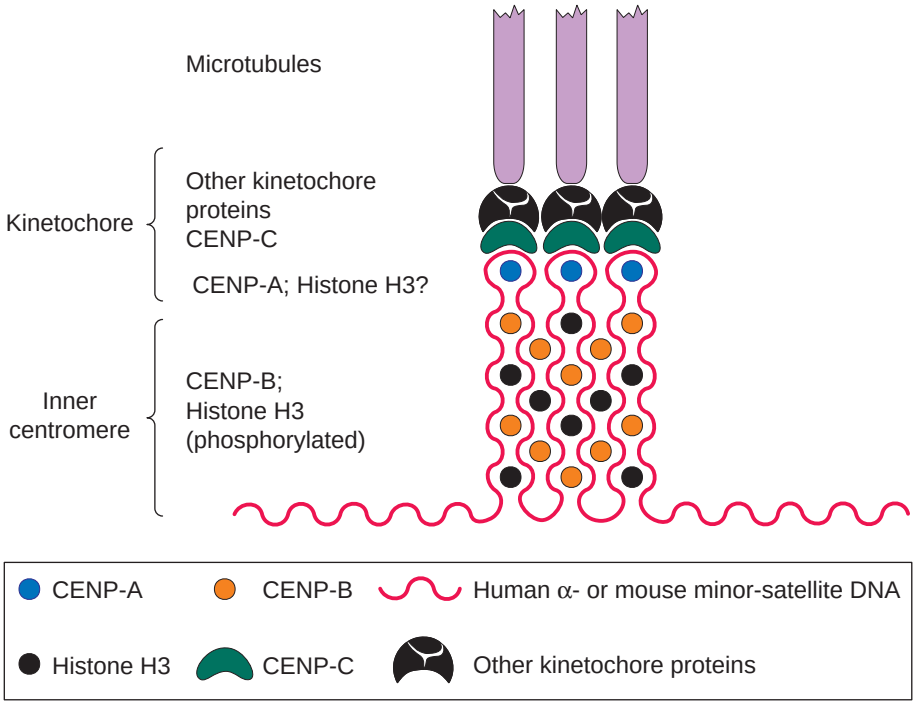

FIGURE 3

Molecular organization of a normal human or mouse centromere[40,41]. Within the inner centromere domain, the assembly of human centromeric α-satellite or mouse minor satellite DNA is facilitated by CENP-B and phosphorylated histone H3. This step is not dependent upon CENP-A, although CENP-A appears necessary for the establishment of a fully condensed centromere state. CENP-A binding to a subset of these satellite DNA sequences might be responsible for the formation of a pre-kinetochore nucleosomal structure, upon which CENP-C and other kinetochore proteins can associate[40]. CENP-C association might occur through direct contact with a specific DNA conformation generated by the CENP-A–DNA nucleosomal complex, which could explain the observed DNA-binding property of CENP-C[42]. Although use of immunofluorescence has demonstrated the exclusion of phosphorylated histone H3 moiety from the CENP-A-containing kinetochore domain[41], it is unclear whether histone H3 is actually found within this domain or whether CENP-A is itself phosphorylated (see Fig. 2b).

At present, there are insufficient data to pinpoint a definitive triggering mechanism for centromerization. Most of the aforementioned chromatin-modifying mechanisms are not unique to the centromere. For

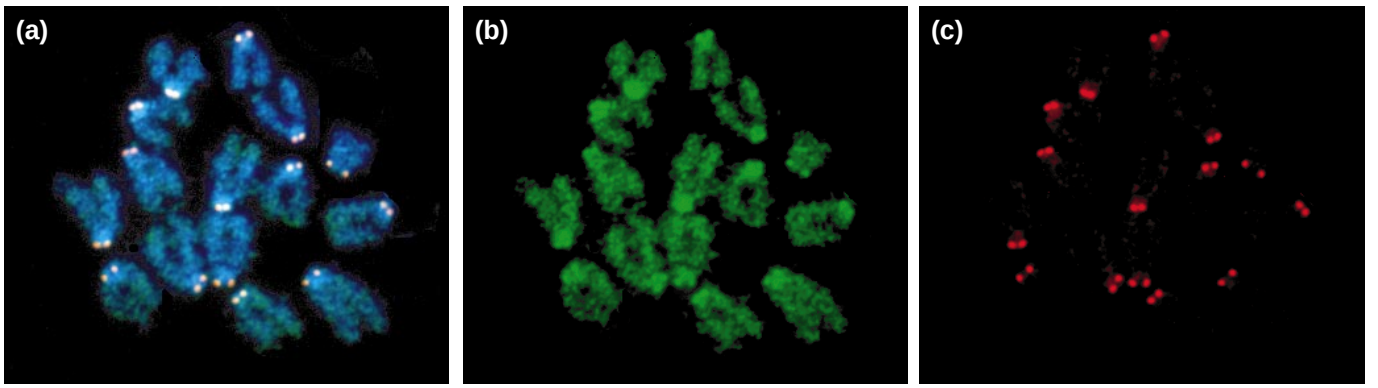

FIGURE 4

Immunofluorescence analysis of poly(ADP-ribose) polymerase (PARP) on mouse chromosomes. Strong co-localization of PARP (green) with CENP-A/CENP-B (red) is seen at the centromere, whereas a more generalized PARP signal is detected on the chromosome arms. On many chromosomes, the significantly larger area of enhanced PARP staining relative to that for the CENP-A/CENP-B signal suggests that centromeric PARP binding might spread to surrounding regions. (a) Merged image; (b, c) split images for green and red, respectively.

example, hypoacetylation, poly(ADP-ribosyl)ation, late replication and chromatin templating are features also shared to varying degrees by the telomere, non-centromeric heterochromatin and/or pockets of euchromatin (for example, different silencing states). Such overlapping roles at other chromosomal domains suggest that these mechanisms might not, on their own, be sufficient to provide the primary drive for centromerization. More likely, the primary driver would involve a centromere-specific factor, whose activity, possibly regulated by one or more of the other 'shared' mechanisms, would provide a chromosomal marking that would unequivocally endorse centromerization. CENP-A is an appealing candidate for this role, but a better understanding of the properties of this protein, and those of other intrinsic centromere constituents, will be necessary before a clearer picture emerges.

References

1 Choo, K.H.A. (1997) *The Centromere*, Oxford University Press
2 Allshire, R.C. (1997) Centromeres, checkpoints and chromatid cohesion. *Curr. Opin. Genet. Dev.* 7, 264–273
3 Clarke, L. (1998) Centromeres: proteins, protein complexes, and repeated domains at centromeres of simple eukaryotes. *Curr. Opin. Genet. Dev.* 8, 212–218
4 Craig, J.M. *et al.* (1998) Mammalian centromere: DNA sequence, protein composition, and role in cell cycle progression. *Exp. Cell Res.* 246, 249–262
5 Rieder, C.L. and Salmon, E.D. (1998) The vertebrate cell kinetochore and its role during mitosis. *Trends Cell Biol.* 8, 310–317
6 Skibbens, R.V. and Hieter, P. (1998) Kinetochores and the checkpoint mechanism that monitors for defects in the chromosome segregation machinery. *Annu. Rev. Genet.* 32, 307–337
7 Amon, A. (1999) The spindle checkpoint. *Curr. Opin. Genet. Dev.* 9, 69–75
8 Dobie, K.W. *et al.* (1999) Centromere proteins and chromosome inheritance: a complex affair. *Curr. Opin. Genet. Dev.* 9, 206–217
9 Karpen, G.H. and Allshire, R.C. (1997) The case for epigenetic effects on centromere identity and function. *Trends Genet.* 13, 489–496
10 Choo, K.H.A. (1997) Centromere DNA dynamics: latent centromeres and neocentromere formation. *Am. J. Hum. Genet.* 61, 1225–1233
11 Murphy, T.D. and Karpen, G.H. (1998) Centromeres take flight: alpha satellite and the quest for the human centromere. *Cell* 93, 317–320
12 Wiens, G.R. and Sorger, P.K. (1998) Centromeric chromatin and epigenetic effects in kinetochore assembly. *Cell* 93, 313–316
13 Sullivan, B.A. and Willard, H.F. (1998) Stable dicentric chromosomes with two functional centromeres. *Nat. Genet.* 20, 227–228
14 Koch, J. (2000) Neocentromeres and alpha satellite: a proposed structural code for functional human centromere DNA. *Hum. Mol. Genet.* 9, 149–154
15 Strahl, B.D. and Allis, C.D. (2000) The language of covalent histone modifications. *Nature* 403, 41–46
16 Lewis, J.D. *et al.* (1992) Purification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA. *Cell* 69, 905–914
17 Mitchell, A.R. *et al.* (1996) Epigenetic control of mammalian centromere protein binding: does DNA methylation have a role? *J. Cell Sci.* 109, 2199–2206
18 Miniou, P. *et al.* (1997) α-Satellite DNA methylation in normal individuals and in ICF patients: heterogeneous methylation of constitutive heterochromatin in adult and fetal tissues. *Hum. Genet.* 99, 738–745
19 Ekwall, K. *et al.* (1997) Transient inhibition of histone deacetylation alters the structural and functional imprint at fission yeast centromeres. *Cell* 91, 1021–1032
20 Jeppesen, P. *et al.* (1992) Antibodies to defined histone epitopes reveal variations in chromatin conformation and underacetylation of centric heterochromatin in human metaphase chromosomes. *Chromosoma* 101, 322–332
21 O'Neill, L.P. and Turner, B.M. (1995) Histone H4 acetylation distinguishes coding regions of the human genome from heterochromatin in a differentiation-dependent but transcription-independent manner. *EMBO J.* 14, 3946–3957
22 Johnson, C.A. *et al.* (1998) Distinctive patterns of histone H4 acetylation are associated with defined sequence elements within both heterochromatic and euchromatic regions of the human genome. *Nucleic Acids Res.* 26, 994–1001
23 Turner, B.M. *et al.* (1992) Histone H4 isoforms acetylated at specific lysine residues define individual chromosomes and chromatin domains in *Drosophila* polytene nuclei. *Cell* 69, 375–384
24 Ng, H-H. and Bird, A. (1999) DNA methylation and chromatin modification. *Curr. Opin. Genet. Dev.* 9, 158–163
25 Ng, H-H. *et al.* (1999) MBD2 is a transcriptional repressor belonging to the MeCP1 histone deacetylase complex. *Nat. Genet.* 23, 58–61
26 Wade, P.A. *et al.* (1999) Mi-2 complex couples DNA methylation to chromatin remodelling and histone deacetylation. *Nat. Genet.* 23, 62–66
27 Gurley, L.R. (1978) Histone phosphorylation and chromatin structure during mitosis in Chinese hamster cells. *Eur. J. Biochem.* 84, 1–15
28 Wolffe, A. (1995) *Chromatin: Structure and Function*, Academic Press
29 Hendzel, M.J. (1997) Mitosis-specific phosphorylation of histone H3 initiates primarily within pericentric heterochromatin during G2 and spreads in an ordered fashion coincident with mitotic chromosome condensation. *Chromosoma* 106, 348–360
30 Roth, S.Y. and Allis, C.D. (1992) Chromatin condensation: does histone H1 phosphorylation play a role? *Trends Biochem. Sci.* 17, 93–98
31 Earle, E. *et al.* (2000) Poly(ADP–ribose) polymerase at active centromeres and neocentromeres at metaphase. *Hum. Mol. Genet.* 9, 187–194
32 Everett, R.D. *et al.* (1999) Specific destruction of kinetochore protein CENP-C and disruption of cell division by herpes simplex virus immediate-early protein Vmw110. *EMBO J.* 18, 1526–1538
33 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
34 Csink, A.K. and Henikoff, S. (1998) Something from nothing: the evolution and utility of satellite repeats. *Trends Genet.* 14, 200–204
35 Williams, B.C. *et al.* (1998) Neocentromere activity of structurally acentric minichromosomes in *Drosophila*. *Nat. Genet.* 18, 30–37
36 Choo, K.H.A. (1998) Turning on the centromere. *Nat. Genet.* 18, 3–4
37 Grunstein, M. (1998) Yeast heterochromatin: regulation of its assembly and inheritance by histones. *Cell* 93, 325–328
38 Sugata, N. *et al.* (1999) Characterisation of a novel kinetochore protein, CENP-H. *J. Biol. Chem.* 274, 27343–27346
39 Buchwitz, B.J. *et al.* (1999) A histone-H3-like protein in *C. elegans*. *Nature* 401, 547–548
40 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
41 Van Hooser, A.A. *et al.* (1999) The mammalian centromere: structural domains and the attenuation of chromatin modeling. *FASEB J.* 13 (Suppl.), S216–S220
42 Yang, C. *et al.* (1996) Identification of overlapping DNA-binding and centromere-targeting domains in the human kinetochore protein CENP-C. *Mol. Cell. Biol.* 16, 3576–3586
43 Fowler, K.J. *et al.* (2000) Uterine dysfunction and genetic modifiers in centromere protein B-deficient mice. *Genome Res.* 10, 30–41
44 Fukagawa, T. and Brown, W.R.A. (1997) Efficient conditional mutation of the vertebrate CENP-C gene. *Hum. Mol. Genet.* 6, 2301–2308
45 Kalitsis, P. *et al.* (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
46 Fukagawa, T. *et al.* (1999) CENP-C is necessary but not sufficient to induce formation of a functional centromere. *EMBO J.* 18, 4196–4209
47 Knehr, M. *et al.* (1996) Cellular expression of human centromere protein C demonstrates a cyclic behavior with highest abundance in the G1 phase. *Proc. Natl. Acad. Sci. U. S. A.* 93, 10234–10239
48 Aagaard, L. *et al.* (1999) Functional mammalian homologues of the *Drosophila* PEV-modifier Su(var)3–9 encode centromere-associated proteins which complex with the heterochromatin component M31. *EMBO J.* 18, 1923–1938
49 Oei, S.L. *et al.* (1997) The role of poly(ADP-ribosyl)ation. *Rev. Physiol. Biochem. Pharmacol.* 131, 127–173
50 Smith, S. *et al.* (1999) Tankyrase, a poly(ADP-ribose) polymerase at human telomeres. *Science* 282, 1484–1487
51 d'Adda di Fagagna, F. *et al.* (1999) Functions of poly(ADP-ribose) polymerase in controlling telomere length and chromosomal stability. *Nat. Genet.* 23, 76–80
52 Oliver, F.J. *et al.* (1999) Poly(ADP-ribose) polymerase in the cellular response to DNA damage, apoptosis, and disease. *Am. J. Hum. Genet.* 64, 1282–1288
53 Platero, J.S. *et al.* (1999) A distal heterochromatic block displays centromeric activity when detached from a natural centromere. *Mol. Cell.* 4, 995–1004
54 Saffery, R. *et al.* (2000) Human centromeres and neocentromeres show identical distribution patterns of over 20 functionally important kinetochore associated proteins. *Hum. Mol. Genet.* 9, 175–185

Acknowledgements

I apologize to the colleagues whose primary references have been omitted because of space restrictions, but every effort has been made to ensure that their work is adequately referenced by the reviews or secondary papers cited. I am grateful to Jeff Craig, Richard Saffery and Anthony Lo for critical discussions, and NHMRC, AMRAD and AusIndustry for funding support.