Assembly of Variant Histones into Chromatin

Steven Henikoff1 and Kami Ahmad2

1Howard Hughes Medical Institute, Fred Hutchinson Cancer Research Center, Seattle, Washington 98109; email: steveh@fhcrc.org
2Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115; email: kami_ahmad@hms.harvard.edu

Key Words
nucleosome, chromatin remodeling, histone replacement, epigenetics, centromeric chromatin

Abstract
Chromatin can be differentiated by the deposition of variant histones at centromeres, active genes, and silent loci. Variant histones are assembled into nucleosomes in a replication-independent manner, in contrast to assembly of bulk chromatin that is coupled to replication. Recent in vitro studies have provided the first glimpses of protein machines dedicated to building and replacing alternative nucleosomes. They deposit variant H2A and H3 histones and are targeted to particular functional sites in the genome. Differences between variant and canonical histones can have profound consequences, either for delivery of the histones to sites of assembly or for their function after incorporation into chromatin. Recent studies have also revealed connections between assembly of variant nucleosomes, chromatin remodeling, and histone post-translational modification. Taken together, these findings indicate that chromosome architecture can be highly dynamic at the most fundamental level, with epigenetic consequences.
INTRODUCTION

Over the past decade we have witnessed a renaissance of interest in core histones with the general realization that these four simple components of nucleosomal octamers, histones H2A, H2B, H3 and H4, are also key players in basic nuclear processes. This is especially true for post-translational modifications of histones, which have been implicated both in modulating chromatin architecture, and in the regulation of transcription (Brownell et al. 1996, Jennewin & Allis 2001, Turner et al. 1992). The realization that addition or removal of these modifications can facilitate gene activation or silencing has fueled considerable excitement in the chromatin field.

Much less attention has been paid to the differentiation of chromatin by the incorporation of variant histones, which have been implicated both in modulating chromatin architecture, and in the regulation of transcription (Brownell et al. 1996, Jennewin & Allis 2001, Turner et al. 1992). The realization that addition or removal of these modifications can facilitate gene activation or silencing has fueled considerable excitement in the chromatin field.

Much less attention has been paid to the differentiation of chromatin by the incorporation of variant histones (Table 1). These are separately encoded forms of canonical histones that are distinguished by sequence differences (Malik & Henikoff 2003). Many variant histones are simply polymorphic versions of the major canonical forms that are assembled into bulk chromatin behind the replication fork. However, other variants are found to have profound differences that distinguish them from canonical forms, either in the way that they are deposited or the way that they function after deposition, or both. In the past few years, we have come to realize that this latter class of histone variants and the special machineries that deposit them play important roles in chromatin differentiation and epigenetic maintenance. The recent study of variants and their assembly has begun to reveal a highly dynamic picture of chromatin, in which processes of post-transcriptional modification appear to be coupled to processes of histone replacement.

In our review, we examine the basis for this more dynamic view of chromatin by considering recent studies on particular core histone variants and chromatin assembly complexes. We explore the possibility that the processes that replace histones at active genes and that propagate chromatin states when DNA replicates involve the concerted action of nucleosome remodeling and histone modification activities. Although the study of chromatin dynamics is technically demanding, we expect that the rapid improvements in molecular biology, cytogenetics, and genomics technologies mean that this area of research is still in its infancy.

Centromeric Chromatin is Identified by a Special Histone H3 Variant

Every eukaryotic chromosome requires a centromere for it to segregate at mitosis, and the uniqueness of this structure has facilitated the cytological identification of centromere-specific protein components. The first such components were identified as epitopes of autoimmune antibodies (Earnshaw & Rothfield 1985), and one of these, CENP-A, was found to be a histone H3 homolog that copurifies with nucleosomes (Palmer et al. 1991, Palmer et al. 1987). Genetic studies
Table 1  Histone variants and associated chromatin assembly complexes

<table>
<thead>
<tr>
<th>Histones</th>
<th>Features</th>
<th>Assembled by organism</th>
</tr>
</thead>
<tbody>
<tr>
<td>Archaeal histones</td>
<td>Ancestral histone fold proteins without tails found in singly wrapped tetrameric units that comprise nucleosome particles.</td>
<td>Unknown</td>
</tr>
<tr>
<td>H2A, H2B</td>
<td>Canonical core histones encoded by replication-coupled genes.</td>
<td>FACT (yeast, Drosophila)</td>
</tr>
<tr>
<td>H2AZ</td>
<td>H2A variant found in nearly all eukaryotes that has a diverged self-interaction domain.</td>
<td>SWR1 (yeast), Tip60 (Drosophila)</td>
</tr>
<tr>
<td>macroH2A</td>
<td>Vertebrate-specific H2A variant with a C-terminal globular domain.</td>
<td>Unknown</td>
</tr>
<tr>
<td>H2A-Bbd</td>
<td>Vertebrate-specific H2A variant that is widely distributed. Relatively deficient on the inactive X-chromosome.</td>
<td>Unknown</td>
</tr>
<tr>
<td>H2AX</td>
<td>H2A form with an SQ[E/D] Ø (Ø = hydrophobic) C-terminal motif that becomes serine phosphorylated at sites of double-stranded breaks.</td>
<td>INO80 (yeast)</td>
</tr>
<tr>
<td>H3, H4</td>
<td>Canonical core histones encoded by replication-coupled genes.</td>
<td>CAF-1 (plants, animals, fungi)</td>
</tr>
<tr>
<td>H3.3 (H3.2 in plants)</td>
<td>H3 variant that replaces H3 and differs at position 31 and at a few residues on helix 2 that allow deposition outside of replication.</td>
<td>HIRA (mammals)</td>
</tr>
<tr>
<td>Packaging histones</td>
<td>Core and linker histone variants adapted for tight packaging of DNA in sperm and pollen in some organisms.</td>
<td></td>
</tr>
</tbody>
</table>

reveal that mammalian CENP-A and its counterparts in other eukaryotes (generically referred to as CenH3s) are absolutely required for assembly of the proteinaceous kinetochore to which the spindle microtubules attach at mitosis and meiosis (Blower & Karpen 2001, Buchwitz et al. 1999, Howman et al. 2000, Stoler et al. 1995). Antibodies against CenH3s from both plants and animals have been used to map centromeres (Alonso et al. 2003, Lo et al. 2001, Nagaki et al. 2004), because in these organisms the centromere is not determined by DNA sequence but rather by the presence of centromeric chromatin. In fact, human neocentromeres that show no resemblance in DNA sequence to native alpha-satellite-containing centromeres are nevertheless packaged in CENP-A-containing nucleosomes (Alonso et al. 2003, Amor et al. 2004). These observations imply that the constant location of the centromere in all cells of an organism through millions of years of evolution is maintained by the faithful assembly of CenH3-containing chromatin. Apparently centromeres are maintained indefinitely by the action of a chromatin assembly process.

What features of CenH3s are recognized for assembly into chromatin? Swaps between CENP-A or Cse4p and H3 identified the core region as being crucial (Keith et al. 1999, Shelby et al. 1997), although further inferences were complicated by the possibility that some H3-specific residues might be incompatible with CenH3 function. A more refined approach is to use heterologous CenH3s; in the case of Drosophila CenH3 (Cid), this led to the identification of Loop I as being both necessary and sufficient for Cid localization to centromeres (Vermaak et al. 2002) (Figure 1). Thus Loop I of human CENP-A is included within a region inferred to be more compact than the corresponding regions of H3 (Black et al. 2004). The possibility that contacts between Loop I and centromeric DNA are important for correct assembly of centromeric nucleosomes is implied by the evidence for adaptive evolution of Loop I in Drosophila and Arabidopsis (Cooper & Henikoff 2004, Malik & Henikoff 2001). However, mammalian CENP-A is different in that no adaptive evolution is seen (Talbert et al. 2004), and heterologous CenH3s can localize to human centromeres (Henikoff et al. 2000, Wieland et al. 2004). In fact, yeast Cse4p can even functionally replace CENP-A (Wieland et al. 2004), unlike the
Regions of H3 and H2A variants responsible for chromatin differentiation. Surface-accessible residues (shaded blue) were annotated from the 1.9 Å X-ray crystal structure (Davey et al. 2002). N-terminal tails that are not included in nucleosome models but are presumed to be accessible are also indicated (shaded green). Unshaded residues are occluded by DNA or by other histones within the octamer. Schematics for the secondary structure of each histone is indicated, and regions of the histones and their variants with the functions described in the text are indicated.

situation for Drosophila bipeptinata Cid, which requires Drosophila melanogaster Loop I to localize to D. melanogaster centromeres (Vermaak et al. 2002). These differences between organisms might be attributable to the other centromere-specific DNA-binding protein, CENP-C, which is adaptively evolving in mammals and plants, but which has not been identified in Drosophila (Talbert et al. 2004). It has been proposed that CenH3s and CENP-Cs adapt the rapidly evolving centromeric satellites to the conserved kinetochore machinery (Malik & Henikoff 2001), in which case proteins that are not adaptively evolving, such as CENP-A and Cse4p, would not require species-specific interactions to package centromeric chromatin.

Assembly of CenH3-containing nucleosomes is independent of replication (Ahmad & Henikoff 2001a, Shelby et al. 2000). The process presumably initiates with interactions between DNA and CenH3 Loop I of CenH3-H4 units, and when the full core is assembled, the CenH3 N-terminal tail would interact with linker DNA (Vermaak et al. 2002) (Figure 2). In the canonical nucleosome, the H3 N-terminal tail exits between the DNA helices and contacts the DNA minor groove where the DNA leaves the nucleosome core (Luger et al. 1997). In contrast to the nearly invariant tail of canonical H3, which is constrained by the density of post-translational modification sites, the CenH3 N-terminal tails are extraordinarily diverse, differing in length and sequence to such an extent that they cannot be aligned between distant species (Malik & Henikoff 2003). Minor groove-binding motifs have been detected...
in CenH3 N-terminal tails (Malik et al. 2002), suggesting that contacts with the minor groove stabilize centromeric nucleosome assembly. In support of this possibility, the N-terminal tails of some CenH3s are adaptively evolving (Malik & Henikoff 2001, Talbert et al. 2002). In *Drosophila Cid*, regions of adaptive evolution are interspersed with patches of sequence conservation that might correspond to sites of interaction with the conserved kinetochoore machinery (Keith et al. 1999, Malik et al. 2002). Centromeric chromatin would then mature with the recruitment of CENP-C, a DNA-binding protein that nevertheless depends upon CenH3 for centromeric localization (Moore & Roth 2001, Sugimoto et al. 1994, Yang et al. 1996).

It is striking that centromeric chromatin is organized as interspersed stretches of CenH3- and H3-containing nucleosomes (Blower et al. 2002). It is difficult to envision processive assembly leading to an interspersed organization, which would require switching back and forth between substrates. Alternatively, CenH3- and H3-containing nucleosomes might be assembled at different times in the cell cycle (Ahmad & Henikoff 2002b). Canonical H3 incorporates behind the replication fork, whereas both human and fly CenH3s deposit in an RI manner. Just how CenH3 assembly is maintained at sites of pre-existing CenH3 in the absence of DNA sequence determinants is a major unanswered question. A speculative possibility is that anaphase tension on CenH3-containing chromatin causes adjacent nucleosomes to unravel and subsequent chromatin repair incorporates new CenH3 (Ahmad & Henikoff 2002b, Mellone & Allshire 2003). It is also possible that unique patterns of H3 modifications found in interspersed and flanking stretches of nucleosomes predispose centromeres for deposition of CenH3 (Sullivan & Karpen 2004). How such patterns arise might depend on modes of assembly of H3-containing nucleosomes, which we turn to next.

**Active Genes are Sites of H3.3 Replacement**

Another universal H3 variant, H3.3, is so similar to canonical H3 that its distinctive properties were not realized until recently. In
animals, H3.3 differs from H3 at only four amino acid positions (Figure 1), consistent with the view that it is essentially a constitutive version of canonical H3 (Yu & Gorovsky 1997). Indeed, H3.3 is the dominant H3-subtype in nondividing differentiated cells in vertebrates (Pina & Suau 1987, Urban & Zweidler 1983) and in endoreplicating cells during development of a chordate (Chioda et al. 2004). Nevertheless, we found that three of the four differences between H3 and H3.3 determine nucleosome assembly behavior: Changes from the H3 to the H3.3 form allowed RI assembly (Ahmad & Henikoff 2002c). This difference between H3 and H3.3 defines a pathway for RI assembly that is distinct from the RC pathway whereby canonical histones are assembled into bulk DNA. Another difference is that RC assembly of either H3 or H3.3 requires the N-terminal tail, which RI does not.

Confirmation that alternative RC and RI assembly pathways exist came from the purification of H3- and H3.3-containing complexes (Tagami et al. 2004). The well-studied RC assembly complex, CAF-1, copurifies with H3, whereas the replication-independent histone chaperone, HIRA, copurifies with H3.3. Both complexes include a common H4-binding component, RbAp48, a homolog of which is also essential for the assembly of Schizosaccharomyces pombe CenH3 (Hayashi et al. 2004). Interestingly, the assembly form of histones in both human CAF-1 and HIRA complexes was an H3•H4 (or H3.3•H4) dimer (Tagami et al. 2004), not a tetramer as might have been expected from the existence of (H3•H4)2 tetramers in solution (Wolffe 1992). This implies that histone heterodimers are substrates in all known assembly complexes (Figure 3), consistent with a common origin of eukaryotic nucleosome assembly mechanisms. Core eukaryotic histones evolved from structurally similar archaeal histones that assemble to form tetrameric nucleosomes closely resembling (H3•H4)2 tetramers produced in vitro (Pereira & Reeve 1998), and it will be interesting to determine whether there are biochemical similarities in assembly as well.

H3.3 is enriched in active chromatin (Hendzel & Davie 1990), and the basis for enrichment has been elucidated by cytological studies using epitope-tagged versions of H3.3. We showed that RI assembly of H3.3 localizes to active, but not inactive, rDNA arrays and to euchromatin, but not heterochromatin, in Drosophila (Ahmad & Henikoff 2002c). In human cells, H3.3 incorporates at a transgene array that has been induced to transcribe (Janicki et al. 2004). The concomitant loss of heterochromatic markers in both Drosophila and human cells demonstrates that the process of H3.3 deposition at active genes accompanies replacement of pre-existing histones. The replacement process can be rapid, occurring on the order of an hour at a transgene array observed in living cells (Janicki et al. 2004).

RI assembly of H3.3•H4 provides an attractive mechanism for the resetting and perpetuation of histone modifications (Ahmad & Henikoff 2002c). RC assembly leaves a mixture of old and new nucleosomes on daughter strands and, at active genes, this would be a mixture of H3.3- and H3-containing nucleosomes. If H3.3 is post-translationally modified in a way that suits transcriptional activity, then the mixture of histones after replication would continue to promote transcription (Henikoff et al. 2004). Continued transcriptional activity would then result in nucleosome replacement over the body of genes, leading to differentiation of chromatin whereby active regions are packaged in H3.3-containing nucleosomes. Indeed, there is enough H3.3 in Drosophila Kc cell chromatin to densely package transcribed regions (McKittrick et al. 2004). Furthermore, the enrichment in bulk H3.3 of lysine modifications that have been found to correlate with active transcription (McKittrick et al. 2004, Waterborg 1990) provides a connection between histone replacement via the RI pathway and changes in properties of chromatin by post-transcriptional modification.
It remains to be determined whether the form of H3.3•H4 that is assembled at active genes is post-translationally modified prior to or during assembly. Nevertheless, the strong correlations among lysine modifications found in active chromatin of flies and yeast are most easily understood if assembly and modification are concerted processes (Workman & Abmayr 2004). Associations between histone-modifying enzymes and elongating RNA polymerases are also consistent with concerted assembly and modification. In yeast, histone replacement can occur within a minute at sites of active transcription (Schwabish & Struhl 2004), which delimits the processes of histone modification and histone assembly to the same short-time interval.

**Nucleosomes are Disassembled at Promoters**

Nucleosome replacement during transcriptional elongation might also play a role in gene regulation during development. Transcription within the mouse beta-globin locus control region has been proposed to potentiate activation (Gribnau et al. 2000), and transcription through a *Drosophila* Polycomb response element derepresses expression of genes in cis (Bender & Fitzgerald 2002, Drewell et al. 2002, Hogg & Karch 2002, Rank et al. 2002). Inactive chromatin can obstruct the binding of transcription factors to their target sites, but once bound, the result is a heritable state of activity (Ahmad & Henikoff 2001b). Heritable activation of previously silent chromatin by transcription-coupled replacement of histones is one way that the RI assembly process might lead to inheritance of an epigenetic state (Ahmad & Henikoff 2002a).

Although histone modification is the best-established change in chromatin states upon activation of a gene, recent studies have shown that the distribution of nucleosomes can change as well. Activation of the yeast PHO5 gene leads to loss of the nucleosome at the PHO5 promoter (Boeger et al. 2003, Reinke & Horz 2003). The nucleosome is not simply moved aside, rather it unrolls (Figure 3), a process that is facilitated by the ASF1 histone chaperone (Adkins et al. 2004, Boeger et al. 2004). This process is evidently not limited to PHO5 because depletion of nucleosomes at promoters has been found to occur genome-wide in yeast (Bernstein et al. 2004, Lee et al. 2004). In addition, transiting RNA polymerases displace nucleosomes, leading to variation in nucleosome occupancy over the

**Figure 3**

A model for RI replacement with histone variants (Henikoff et al. 2004). Parallels between H2A•H2B and H3•H4 replacement processes suggest a common underlying mechanism, where a large molecular machine (either RNA polymerase or a SWI/SNF remodeler) partially or completely unrolls a nucleosome during transit. The result is either retention of heterodimeric subunits, such as the FACT-facilitated transfer of H2A•H2B from in front of RNA polymerase to behind (Belotserkovskaya et al. 2003, Formosa et al. 2002), or loss of a heterodimer. In the latter case, chromatin repair replaces the lost heterodimer with either H3.3•H4 (top) or H2AZ•H2B (bottom). Failure to repair will result in nucleosome eviction and reduced nucleosome densities, such as has been observed at promoters and in the body of highly transcribed genes in yeast.
body of genes (Kristjuhan & Svejstrup 2004, Schwabish & Struhl 2004). These findings raise questions about the implicit assumption that the distribution of histone modifications using chromatin immunoprecipitation (ChIP) accurately reflects their relative density along the DNA (Hanlon & Lieb 2004). The action of chromatin remodeling machines, DNA-binding proteins, and RNA polymerases can affect local nucleosome densities such that this assumption is valid, at best, as a first approximation. Thus transcription-coupled replacement or loss of histones could underlie some of the patterns of histone modifications that have been reported. This interpretation might extend to the mapping of histone-modifying enzymes, which are cross-linked to their target sites by formaldehyde, a reagent that primarily cross-links primary amines (Nagy et al. 2003, Solomon & Varshavsky 1985), which are especially abundant and accessible on the histones themselves.

Chromatin Remodeling Machines Replace H2A Variants

Structural alignments of nucleosomal subunits reveal ancestral homology between H3•H4, H2A•H2B, and archaean dimeric units, where H3 aligns with H2A and H4 with H2B (Pereira & Reeve 1998). This structural equivalence might underlie the fact that H3 and H2A have diverse variant forms, whereas H4 and H2B have (almost) none (Malik & Henikoff 2003). Of the H2A variants, H2AZ is conspicuous in having a single evolutionary origin very early in eukaryotic evolution. Although H2AZ is essential in animals (Faast et al. 2001, van Dal & Elgin 1992), it is nonessential in budding yeast (Dhillon & Kamakaka 2000, Jackson & Gorovsky 2000), and this has facilitated its in vivo study. Thus we know that H2AZ can act as a transcriptional activator and an antisilencer at different loci (Meneghini et al. 2003, Santisteban et al. 2000). Whether these functions of H2AZ generalize to plants and animals is not known. Surprisingly, mammalian H2AZ shows a heterochromatic distribution and interacts with HP1 (Fan et al. 2004, Rangasamy et al. 2003), making it difficult to draw firm conclusions about conserved roles for H2AZ in chromatin function.

Whereas the mechanism whereby H2AZ affects chromatin remains uncertain, much has been learned about how it is deposited into chromatin, thanks to a combination of in vivo and in vitro studies in yeast. H2AZ is assembled by the SWR1 complex, and mutations in the swr1 gene, which encodes a key component of this complex, show gene expression phenotypes similar to those of H2AZ (brt1) mutants (Kobor et al. 2004, Krogan et al. 2003, Mizuguchi et al. 2004). In vitro, SWR1 replaces H2A•H2B dimers in nucleosomes with H2AZ•H2B dimers (Mizuguchi et al. 2004). The fate of the leaving H2A•H2B dimer is unknown, so that it is premature to refer to this process as an exchange implying a reciprocal event, as opposed to replacement, which does not. The actual mechanism of H2A•H2B replacement might be very similar to that for H3•H4 replacement, in which chromatin is perhaps repaired after loss of a heterodimeric subunit during transit of a large complex such as RNA polymerase or an ATP-dependent remodeling complex (Figure 3).

Replacement of H2A•H2B with H2AZ•H2B requires ATP, as expected from the fact that the SWR1 subunit is a member of the SWI/SNF family of ATP-dependent chromatin remodelers. This finding has important implications both for understanding the assembly of histone variants, and for the understanding of chromatin remodeling. It remains uncertain as to just how actions of various SWI/SNF family members observed in vitro relate to their functions in vivo; but in the case of SWR1, the in vitro replacement of H2A with H2AZ and the supporting in vivo data provide unequivocal evidence for this specific function. We look forward to other examples of specific roles played by chromatin remodelers in nucleosome assembly or disassembly.
The *Drosophila* SWR1 ortholog performs an analogous H2AZ replacement reaction as part of the Tip60 complex (Kusch et al. 2004). An intriguing twist arises from the role of *Drosophila* H2AZ in repair of double-stranded breaks (Madigan et al. 2002). In most eukaryotes, this role is played by the H2AX variant, which is otherwise similar to canonical H2A except for the presence of a four-amino acid C-terminal motif (SQ[D/E]Ø, where Ø represents a hydrophobic amino acid). The *Drosophila* version of H2AZ (H2AvD) has evolved an H2AX-like motif and as a result functions similarly in double-strand break repair. In diverse eukaryotes, H2AX is phosphorylated on the serine of the H2AX C-terminal motif at sites of double-strand breaks, and phosphorylation spreads rapidly to other H2AXs along the chromosome, an event that is important for recruitment of break repair machinery (Fernandez-Capetillo et al. 2004, Rogakou et al. 1998). In vitro, Tip60 specifically binds phosphorylated H2AvD, acetylates it at Lys5 and replaces it with an unphosphorylated H2AvD (Kusch et al. 2004). This combination of activities suggests that the function of Tip60 is to remodel chromatin at sites of double-strand breaks, while restoring the ground state by effectively erasing the phosphorylation mark.

It is likely that a similar process of ATPase-catalyzed replacement occurs in yeast, where the INO80 chromatin remodeling complex is recruited to H2AX when it is phosphorylated following a double-strand break (Morrison et al. 2004, van Attikum et al. 2004). INO80 is an ATPase distinct from SWR1 and is consistent with yeast H2AX, which is actually the canonical version of H2A, distinct from H2AZ. Therefore, two different remodeling machines with distinct H2A substrates appear to have evolved in different organisms to assume similar roles in DNA repair.

Other H2A variants are lineage-specific. The macroH2A histone is a vertebrate-specific variant unique among histones in having an additional globular domain. This C-terminal 200-amino acid domain is homologous to a broad class of polynucleotide and peptide hydrolases, raising the possibility that macroH2A alters chromatin via the action of a tethered enzyme (Allen et al. 2003). macroH2A is enriched in regions of the mammalian inactive X chromosome that are associated with determinants of facultative silencing, including Xist RNA and H3 trimethyl lysine-27 (Chadwick & Willard 2004). In contrast, another vertebrate-specific variant, H2A IBd, shows a cytological distribution that indicates an association with active chromatin (Chadwick & Willard 2001). These patterns are suggestive of functional differentiation of variant-containing chromatin.

How Is Variant Structure Related to Function in Chromatin?

As we have seen, histone variants are distinguished from canonical histones both by their mode of assembly into nucleosomes and by their properties in chromatin. Ultimately, processes such as transcription and replication must alter the structure of nucleosomes to expose DNA, and this must involve regulating DNA-histone affinities. The structure of the nucleosome suggests ways that exposure can happen. The protein octamer can be divided into three surfaces, each with distinctive roles: (a) a perimeter ramp that underlies the superhelix of wrapped DNA, (b) the two exposed faces of the disk, and (c) flexible N- and C-terminal tails of the histones that extend out of the nucleosome. The roles of these surfaces are becoming apparent, as are the ways that histone modifications and variants alter accessibility of specific nucleosomes. For example, acetylation of N-terminal lysines may neutralize DNA-histone interactions (Turner et al. 1992). An excellent case has been made that modifications decorating the protein superhelical ramp sterically disrupt DNA-histone contacts (Cosgrove et al. 2004, Freitas et al. 2004). A second mode-of-action for modifications is as binding sites for proteins that remold or restrict nucleosomes. Modifications
on the histone tails and the exposed disk faces are clearly suitable for display to incoming factors, and both repressive partners (e.g., HP1) and activating ones (e.g., SWI/SNF nucleosome remodelers) are known (Jenuwein & Allis 2001).

Studies of the conserved histone variant H2AZ support these paradigms. The region of *Drosophila* H2AZ homolog (H2AvD) that is essential for development lies in the docking domain (Figure 4), where H2A interacts with the H3•H4 dimer within the nucleosome (Clarkson et al. 1999). This specialized docking domain of H2AZ presents a binding site for HP1 exposed on the face of the nucleosome (Fan et al. 2004). Thus the primary sequence differences in the variant create a new binding site in nucleosomes. Additionally, this same region shifts the underlying H3 αN helix (Suto et al. 2000). The H3 αN helix also contacts DNA, and the changes in the H2AZ docking domain appear to alter these DNA contacts, thus subtly destabilizing the nucleosome. Other differences in the Loop 1 region of H2AZ appear to configure the composition of nucleosomes. Comparison of crystal structures shows that the H2A Loop 1 will clash with H2AZ Loop I within the nucleosome (Figure 4). Thus nucleosomes homotypic for H2A or for H2AZ are structurally preferable. This incompatibility between H2A and H2AZ implies that the SWR1-catalyzed replacement of one H2A•H2B dimer by H2AZ•H2B will facilitate replacement of the other H2A•H2B dimer.

Despite these insights gleaned from structures, it is not yet clear how altered structural features of H2AZ-containing nucleosomes lead to their diverse roles inferred from in vivo studies. Comparison of variant and canonical nucleosomes based on their physical properties has led to different conclusions by different laboratories: Some observations are consistent with a destabilizing role for H2AZ (Abbott et al. 2001), whereas others are consistent with greater stabilization (Fan et al. 2002, Park et al. 2004). It is possible that higher order interactions are more important than

---

**Figure 4**
Location of differences between canonical and variant histones H3 (blue) and H2A (brown) shown on their three-dimensional structures. Segments that differ are highlighted in yellow.
particle stabilization for in vivo behavior of H2AZ-containing nucleosomes, as suggested by the interaction with HP1 (Fan et al. 2004). Differences in higher order interactions involving H2AZ could help rationalize the different in vivo behaviors implied from studies in yeast and animals.

There is more limited evidence for how other variants of H2A affect function. The H2AX variant is identical to H2A throughout most of the protein, with primary sequence differences limited to the defining C-terminal four-amino acid motif. Thus this variant functions by providing a new phosphorylation site. The macroH2A variant carries a large, globular C-terminal domain that impedes transcription factor binding in vitro, and its histone fold domain interferes with ATP-dependent remodeling (Angelov et al. 2003). Both attributes are consistent with macroH2A playing a role in facultative silencing of the inactivated X chromosome (Costanzi & Pehrson 1998). In contrast, nucleosomes containing the H2A\textsuperscript{Bbd} variant are more accessible than canonical nucleosomes (Angelov et al. 2004), consistent with the striking depletion of H2A\textsuperscript{Bbd} on the inactive X chromosome (Chadwick & Willard 2001).

Some variants of H3 discussed in this review also show evidence of structural differentiation. CenH3s are the most extreme, where positive selection is thought to act as an adaptor between the rapidly evolving centromeric satellite sequences and the conserved kinetochore apparatus (Malik & Henikoff 2001). Rapid substitutions are focused on sites of DNA-histone contacts and may alter the affinity of centromeric nucleosomes for underlying satellite sequences. The Cid Loop 1 that is required for targeting to the centromere is on the disk face, consistent with a role as an exposed binding surface (Vermaak et al. 2002), and the diverged N-terminal tails of all CenH3 histones are thought to be platforms for binding kinetochore components (Malik & Henikoff 2003). In contrast, H3.3 is an example of a variant that appears to be structurally almost identical to its canonical counterpart (Figure 4). The cluster of core residues that specify assembly pathway are located behind the sheath of water residues that are structured by the DNA double helices (Davey et al. 2002) and thus are not accessible in the complete nucleosome. The single remaining difference, Ala31 in H3 versus Ser/Thr31 in H3.3, is nearly universal, suggestive of phosphoregulation of H3.3 in as-yet unidentified processes. However, the interchangeability of RC and RI forms in \textit{Tetrahymena} (Yu & Gorovsky 1997), the presence of only H3.3 in yeasts and molds (Malik & Henikoff 2003), and the dominance of H3.3 in certain cell lineages (Chioda et al. 2004, Pina & Suau 1987, Urban & Zweidler 1983) suggest minimal structural differentiation between H3 and H3.3. Rather, a consistent requirement is seen for an RI H3-subtype in all eukaryotes, with the addition of a distinct RC-only form in multicellular eukaryotes. This phylectic pattern—and the evidence that H3.3 assembly is coupled to transcription—suggests that organisms with large, mostly silent genomes have evolved the RC-only form to package silent chromatin (Ahmad & Henikoff 2002c).

However, the functional distinctions between the two subtypes do not seem to be because of alterations of exposed surfaces in a static model of the nucleosome. We will need to consider the dynamics of nucleosome assembly to understand the role of the H3.3 variant.

Functions During Nucleosome Assembly

A number of protein-binding sites on histones are inaccessible in the complete nucleosome structure, implying that the sites are functional only in structural intermediates. One example of this is the RbAp48 subunit of CAF-1, which binds to the αN helix of histone H4 (Vermaak et al. 1999). This interaction probably occurs as histones are being delivered for deposition because the helix will be buried in...
the completely assembled nucleosome. Surprisingly, a number of complexes that regulate transcription and operate on nucleosomal templates also contain RbAp48. This suggests that nucleosomal intermediates with exposed internal portions may be regulatory targets (Vermaak et al. 2003).

Similarly, the cluster of core residues that are not accessible in the complete nucleosome have been proposed to form a binding surface in predeposition complexes that delivers H3.3 to DNA because they are essential for RI deposition of the histone (Ahmad & Henikoff 2002c). However, the inaccessibility of these residues in H3.3 in the nucleosome leaves little to distinguish it from the canonical H3 histone. The observation that H3.3 can deposit in actively transcribed chromatin at any time implies that these regions are structurally and compositionally dynamic, with nucleosomes continually disassembled and then reassembled (Figure 3). If this leaves little time to complete nucleosome assembly before a new round of replacement begins, transcribed chromatin would remain in a partially assembled state. Thus the repetition of RI assembly alone would be sufficient to confer high DNA accessibility. This function does not require that the H3.3 variant generates a specialized nucleosome structure. In this view, the lack of structural specialization in H3.3 is a key feature that results from the need for H3.3 to perform the same role as the major histone H3, i.e., to package DNA. The sequence identity between the two H3-subtypes would be especially important as genes become repressed, because transcription will have enriched the H3.3 variant in chromatin that must now be packaged in a silent configuration.

A number of observations suggest that nucleosomal intermediates may be critical for active chromatin. As mentioned above, the very process of targeted RI assembly suggests that intermediates will be common in transcribed chromatin. Structures consistent with split nucleosomes are indeed observed at highly transcribed genes (Lee & Garrard 1991), and in vivo cross-linking studies also imply that buried nucleosomal surfaces are exposed in active chromatin (Jackson 1978). Finally, a number of transcription-promoting factors have subunits with high histone-binding affinities, which are thought to assist chromatin binding. However, similar histone affinities are found in the protein chaperones that deliver histones for nucleosome assembly (Akey & Luger 2003). Indeed, some transcription factor complexes contain free histones (Keener 1997), and thus do not appear to be using their histone-binding subunits to bind chromatin. Instead, they might act by assisting nucleosome assembly. A nucleosome disassembly role has been reported for one of these factors, ASF1 (Adkins et al. 2004). These observations are consistent with the idea that nucleosomal intermediates are prevalent in active chromatin and distinguish it from inactive regions.

What Are the Functional Consequences of Multiple Nucleosome Assembly Pathways?

The bulk of nucleosome assembly occurs during DNA replication, and experiments with extracts defined a conserved set of CAFs that support histone deposition specifically on replicating DNA (Verreault et al. 1996). The expectation from these studies was that DNA replication without RC nucleosome assembly would be lethal. Thus it was surprising to find that null mutations for CAF components in budding yeast are viable (Enomoto & Berman 1998, Kaufman et al. 1997). CAF mutants show defects in telomere silencing and in DNA damage repair, but have normal packaging of chromatin. Thus other chromatin assembly pathways must compensate when RC assembly is defective. Indeed, other RC and RI nucleosome assembly activities have been identified. For example, the RCAF complex also supports RC assembly and enhances the activity of CAF. The lack of lethal phenotypes for chromatin assembly mutants suggests that any gaps in chromatin are filled by these other pathways. As most of the yeast genome is
transcriptionally active, gaps left after replication can be filled by transcription-coupled RI assembly. However, compensation also occurs in organisms with more complex genomes. EASCLATATA mutations eliminate CAF in Arabidopsis and cause some meristematic defects, but plants remain viable, indicating functional redundancy (Kaya et al. 2001). Therefore, even transcriptionally silent heterochromatin is being duplicated by alternative pathways in this mutant. A simple model is that alternative assembly pathways are capable of working on any gapped templates, but normally do not act on regions where CAF rapidly completes nucleosome assembly in S phase. Blocking CAF function in mammalian cells stimulates aDNA damage checkpoint (Hoek & Stillman 2003, Nabatiyan & Krude 2004), consistent with the idea that unpackaged DNA is not tolerated or is easily damaged.

Compensation between nucleosome assembly pathways can also explain why defects in the CAF and HIR nucleosome assembly activities result in mis-targeting of the centromeric Cse4p histone (Sharp et al. 2002). A compensation effect is also consistent with the effects of altering the expression of histone variants. Overexpression of centromeric histones in yeast, mammals, or Drosophila results in its deposition in euchromatin, as if extra CenH3 fills gaps in transcriptionally active chromatin (Ahmad & Henikoff 2002b, Collins et al. 2004, Shelby et al. 1997). This suggests that there is a balance between all nucleosome assembly pathways in these organisms, even though they normally act on distinct parts of the genome.

CAF mutants in both yeast and Arabidopsis package DNA into chromatin and are viable; however, they show defective telomeric silencing and unstable developmental fates. Reduction of an RbAp48 component of CAF in Arabidopsis also causes spectacular epigenetic defects (Hennig et al. 2003), although nulls for some of these are lethal (Kohler et al. 2003). These phenotypes point to critical links between the mode of nucleosome assembly and the inheritance of epigenetic states. For example, if each chromatin assembly complex recruits specific chromatin-modifying enzymes, epigenetic patterns would not be preserved when alternate assembly pathways duplicate a chromatin region.

Indeed, there is evidence that the propagation of heterochromatin and DNA replication are linked in this way. Proper heterochromatic localization of HP1 depends in part on binding to CAF-1 at heterochromatic replication forks, as if HP1 is recruited for loading onto H3K9-methylated nucleosomes (Quivy et al. 2004). Moreover, CAF-1 delivers histone H3 pre-methylated at lysine-9 to replication forks at sites of methylated DNA (Sarraf & Stancheva 2004). Both mechanisms for perpetuating heterochromatin require that CAF-1 be used to duplicate the chromatin during replication. These considerations predict that CAF mutations in plants are accumulating replacement H3.2 histones in heterochromatin because RC assembly fails and gap-filling, using RI pathways, compensates. Whereas the phenotypes for elimination of CAF-1 are more severe in mammalian cells, this may reflect its more critical role in epigenetic control of essential processes.

CONCLUSIONS

The differentiation of nucleosomes by incorporation of variant histones must be as ancient as the eukaryotes themselves, insofar as a CenH3-containing centromere is a defining feature of eukaryotic chromosomes. The fact that centromeric chromatin is maintained in the same chromosomal position for millions of years, yet can shift spontaneously to an unrelated DNA sequence, is the most extreme example imaginable of faithful epigenetic inheritance. Nevertheless, the rapid evolution of both the highly repetitive centromeric satellite DNA and the CenH3 variant that evidently adapts to it implies that centromeric stability is dynamically maintained. Yet little is known about how centromeric nucleosomes are assembled and propagated through the cell cycle.
Variant histones, such as CenH3s, are deposited independently of replication, in contrast to the bulk of chromatin, which is assembled by RC assembly of canonical histones. The use of distinct assembly pathways provides a simple means of differentiating chromatin, by targeting replacement of nucleosomes throughout the cell cycle. Transcription itself appears to catalyze the targeting of the H3.3 replacement variant, and the biochemistry of this process has begun to reveal insights into both variant deposition and the possible maintenance of associated active histone modifications. Disruption of nucleosomes is not limited to transcription-coupled replacement, because nucleosomes are “evicted” from promoters upon gene activation.

These dynamic processes of nucleosome replacement and eviction can help account for the abundance and diversity of ATP-dependent chromatin remodeling complexes. Indeed, replacement of the H2AZ variant is catalyzed by one such complex, and it seems likely that other H2A variants are associated with dedicated members of the SWI/SNF family of ATPases. In this way, nucleosome assembly and remodeling, formerly assumed from in vitro work to be distinct processes, are now seen to be activities of the same in vivo process.

Once incorporated into chromatin, most histone variants have distinct structural properties that are likely to profoundly alter chromatin. Differences between H2A and H2AZ in the docking domain can affect nucleosome integrity, and the large globular domain of macroH2A is suspected to have enzymatic function. H3.3 is the exception, because the only difference from H3 that is exposed in the nucleosome is a single tail residue of uncertain significance; rather, differences in posttranslational modifications that are found to distinguish H3 from H3.3 are more likely to affect nucleosome properties.

The study of histone variants and the multiple biochemical processes that deposit them into chromatin has led to new insights into chromatin dynamics. The effects of these processes on gene regulation and chromosome behavior have yet to be elucidated, but the availability of powerful new tools promises to change that. Most importantly, the excitement generated by these new insights has fueled a resurgence of interest in histone variants after decades of relative neglect. We look forward to the deeper insights into eukaryotic biology that now appear to be just around the corner.

**SUMMARY POINTS**

1. Variants of histones H3 and H2A differentiate chromatin at centromeres, active genes, and heterochromatin.
2. Nucleosomes characterized by a special H3 variant identify the centromeres of every eukaryotic chromosome.
3. The replacement histone, H3.3, marks actively transcribed loci by replication-independent nucleosome assembly.
4. Gene activation is accompanied by disassembly of a nucleosome at the promoter.
5. A chromatin remodeling machine replaces the conserved histone variant, H2AZ.
6. Epigenetically silenced chromatin is enriched or depleted in abundance of diverse H2A variants.
7. Variant structure can affect properties of chromatin.
8. The operation of multiple nucleosome assembly pathways has important implications for nucleosome dynamics.

LITERATURE CITED


Jackson JD, Gorovsky MA. 2000. Histone H2A.Z has a conserved function that is distinct from that of the major H2A sequence variants. Nucleic Acids Res. 28:3811–16

150 Henikoff • Ahmad


