Histone variants—the structure behind the function

Juan Ausío

Advance Access publication date 10 May 2006

Abstract

In recent years, the chromatin field has witnessed a renewed interest in histone variants as pertaining to their structural role, but mainly because of the functional specificity they impart to chromatin. In this review, I am going to discuss several of the most recent structural studies on core histone (H2A.Bbd, H2A.Z, H2A.X, macroH2A, H3.3, CENP-A) and linker histone variants (histone HI microheterogeneity) focusing on their role in nucleosome stability and chromatin fibre dynamics with special emphasis on their possible functional implications. The data accumulated to date indicates that histone variability plays an important role in the histone-mediated regulation of chromatin metabolism. Understanding and deciphering the underlying structural amino acid code behind such variability remains one of the most exciting future challenges in chromatin research.

Keywords: histones; chromatin; nucleosome; stability; dynamics

INTRODUCTION

In 1947, at the peak over the controversy of the histone/DNA nature of the genetic material Edgar and Ellen Stedman [1] wrote:

‘... the old view of the composition of the cell nucleus, that the basic protein of cell nuclei can provide the material of which genes are composed is one which, to say the least, has never seem satisfying. While histone is usually regarded as, and probably is, a fairly simple protein, its degree of complexity is not, in fact, known and it may therefore been argued that it is sufficiently complex to fulfill this function’.

This inconclusive, yet hesitant statement, appears in retrospective to be quite timely as there is now plenty of evidence in support of the important structural and epigenetic role of the ‘fairly simple’ histones in mediating gene expression. Furthermore, histones also play important additional roles in other metabolic chromatin functions, such as maintenance of the chromosome integrity (DNA repair), DNA recombination (meiosis) and the process of DNA replication itself. In fact, it has been the proposal of a histone language and deciphering of a potential histone code [2, 3] resulting from histone post-translational modifications (PTMs) [4, 5] and histone variants [6] that have been in part responsible for bringing much recent attention and renewed interest to the chromatin field.

At the time when the conventional nomenclature of histones was finally established at the CIBA Foundation Symposium on Structure and Function of Chromatin in 1974 [7], there was already evidence to suggest that in addition to each of the five major histone types (H1, H2A, H2B, H3 and H4) [8] each of them included some minor variant forms that could in turn be subject to PTMs (i.e. acetylation, methylation, phosphorylation, ubiquitination and poly-ADP-ribosylation) (see [9–11] for early reviews). From a structural point of view, histone variants (which are often replication-independent) [12] can be classified into homomorphous and heteromorphous families [13, 14] depending on the extent of their amino acid sequence departure from...
the main canonical isoforms (whose genes are usually expressed during the S phase of the cell cycle). Homomorphous variants involve only a few amino acid changes (i.e. H2A.1 and H2A.2; H3.1, H3.2 and H3.3) whilst heteromorphous variants involve changes that affect larger portions of the histone molecule [i.e. H2A.X, H2A.Z, macroH2A (mH2A), H2A Barr body-deficient (H2A.Bbd) and centromeric protein A (CENP-A)].

Histones and their variants are responsible for organizing the chromatin complex [15] and within this context they can be classified into core histones (H2A/H2B, H3/H4) and linker histones (histones of the H1 family). Core histones consist of a dimerizing central histone-fold domain [16] which is flanked by N- and C-terminal unstructured domains commonly referred to as ‘tails’, and as their name indicate they are responsible for creating a protein ‘core’ around which 146 bp of DNA are wrapped in approximately two left-handed superhelical turns giving rise to a particle that is known with the name of nucleosome core particle (NCP). Linker histones consist of a winged helix motif [17] flanked by N- and C-terminal tails, bind to the linker DNA connecting adjacent nucleosomes and are responsible for the higher order of folding into a 30 nm chromatin fibre.

In this review, I am going to focus on several recently unveiled dynamic structural aspects of histone variants, both at the nucleosome and at the chromatin fibre levels, with emphasis on their role and possible implications for their defined chromatin function.

**DYNAMIC CHROMATIN**

The two major functions of the chromatin nucleoprotein complex are: folding of DNA to fit the genome within the limited space available in the nucleus of the eukaryotic cell and modulation of the chromatin metabolism [15]. Implicit to these functions is the notion that this complex cannot be structurally rigid but rather a highly dynamic system [4]. Thus, although a lot of the highly valuable structural information about the nucleosome has been obtained from the crystallographic analysis of NCP in recent years [18, 19], this information is quite limited when trying to understand the dynamics of chromatin. Several highly complementary structural methods of analysis, have been, or are currently being used to help overcome the intrinsic limitations of the crystallographic data. In what follows next, I am going to briefly refer to some of the technology used for this purpose as it pertains to chromatin.

**Analytical ultracentrifuge**

Sedimentation velocity analysis in the analytical ultracentrifuge is extremely sensitive to changes in the tertiary and quaternary structure of macromolecular complexes and, hence is very suitable for the study of chromatin [20]. In addition to monitoring conformational changes of NCPs and nucleosome array (chromatin fibre) folding, the method is also useful for determining chromatin stability (DNA/histone dissociation fluctuations) [20].

**Fluorescence recovery after photobleaching (FRAP)**

This method utilizes cells that have been either transiently or stably transfected with green fluorescence protein (GFP)-tagged histones. They are then analysed using confocal microscopy, after bleaching certain areas of the nucleus, and the fluorescence recovery (molecular diffusion) from the unbleached to the bleached areas is monitored [21, 22]. The method allows for the determination of the rates at which histones associate and dissociate from the chromatin template in situ within the living cell and, hence provides a useful tool for the study of chromatin dynamics in vivo.

**Fluorescence resonance energy transfer (FRET)**

As with the analytical ultracentrifuge, this represents a rather classical method of analysis which still retains an enormous potential for structural analysis [24]. With this technique, two regions of the chromatin complex are labelled with different fluorescent labels. When one of these two fluorophores (donor) is excited by an incident light beam (at a suitable wavelength), and a second fluorophore (acceptor) is at a close distance (20–100 Å), an excited state energy from the former can be transferred to the acceptor. Monitoring this process allows one to determine the relative position (movement) of the two fluorophores. The technique has been recently applied to the study of NCP dynamics [25–27].

**Restriction enzyme accessibility assays**

This method monitors the accessibility of restriction enzyme cutting sites in chromatin complexes reconstituted with sequence-defined DNA templates. It has been applied to the characterization of DNA unwrapping from NCP [28, 29] and nucleosome arrays [30].
One potential problem with this technique is the lack of information on the role that the high concentration of divalent ions required for the restriction enzyme digestions has on the intrinsic dynamic properties of the chromatin complexes.

**Single molecule analysis**

In this type of analysis, individual molecules or complexes are analysed using a variety of techniques: atomic force microscope (AFM), optical tweezers (OT) and magnetic tweezers (MT) [31]. This is perhaps one of the most powerful, recently developed set of techniques for the study of chromatin stability and folding [32]. Recently, it has been used in combination with FRET to study the fluctuations of the DNA wrapping in the nucleosome [33]. However, its use in the analysis of chromatin complexes consisting of histone variants has not been forthcoming yet and it may turn out to be very informative.

HISTONE H2A VARIANTS SPECIALIZE THE NUCLEOSOME FOR DEDICATED FUNCTIONS

The heteromorphic members of this histone H2A type (Figure 1A) exhibit significantly major sequence changes at both the N- and C-terminal ends of their molecules with substantially larger predominance of those affecting the carboxy end [4, 14]. While the implications of the N-terminal heterogeneity still remain unclear, most of the recent work has centred its attention on the variability that affects the C-terminal domain. Indeed, early studies by Eickbush and coworkers [34] took advantage of an endogenous protease that selectively cleaves H2A.1 between V(114) and L (115) releasing the last 15 amino acids, to show that this region is critical for the stability of the histone core octamer [35]. An observation which might have consequences for nucleosome stability. Also, this is the region that exits the NCP close to its dyad axis [18, 36] at a position near the binding site of the winged helix domain of the linker histones in the nucleosome [37] and its disruption of this binding could affect the folding of the chromatin fibre.

Histone H2A.Bbd destabilizes the histone octamer

Histone H2A.Bbd appears to have occupied a primitive position in the lineage of the histone H2A family [38], but the occurrence of the gene and its expression did not occur until late in evolution during the appearance of vertebrates [39]. This histone variant which has only 48% identity to canonical H2A (Figure 1A) was first described in 2001 by Chadwick and Willard [40] and is shown to be excluded from the inactive X chromosome from where its name was derived.

The recombinant version of the protein does not form stable octamers when reconstituted with equinmolar amounts of H3-H4-H2B [41, 42], although the mixture is able to form NCPs in the presence of DNA [41, 43]. However, H2A.Bbd-containing NCPs exhibit a relaxed conformation compared with the native NCPs under a broad range of salt concentrations as determined by analytical ultracentrifuge and by FRET [41, 43]. This is most likely the result of the flanking DNA being released from the NCP constraints [41] with the whole particle adopting a more open structure which is highly reminiscent of what is observed in NCPs with the highly acetylated histones [4]. Interestingly, H2A.Bbd was found to be co-localized with acetylated histones during both metaphase and interphase. Thus, it is possible to envision the two structural effects acting synergistically to produce a chromatin domain that is more amenable to transcription [44]. Indeed, FRAP experiments using GFP–H2A.Bbd transfected cells showed that this variant exchanges more rapidly than canonical H2A in chromatin, a fact that would be consistent with the exchange of H2A–H2B dimers associated with transcription (see subsequently) and with the lower stability of the H2A.Bbd histone octamer. The high dynamic flexibility of NCPs consisting of this histone variant when compared with NCPs consisting of other H2A variants is clearly reflected on the ease with which the former can be assembled and disassembled by nucleosome assembly protein (NAP-I) [45].

Histone H2A.Z—a histone variant with controversial roles

If there are any histone variants that have been most extensively studied from a functional perspective, these are without a doubt H2A.Z and H2A.X. Because of the extensive literature available in this and the following section, I am going to refer mainly to the recently published reviews focusing on the more limited amount of structural information available for both of these variants.
Figure 1: Histone H2A variants and their involvement in nucleosome dynamics and conformation. (A) Amino acid sequence alignment of heteromorphous human H2A histone variants. The degree of similarity among the different sequences is proportional to the dark intensity of the shaded areas. Also shown are the α-helices and β-turns corresponding to the histone-fold domain and macro domain of mH2A. (B) Nucleosomes are highly dynamic nucleoprotein complexes [4, 25]. In solution, the DNA associated with the histone octamer rapidly fluctuates between a folded (closed) (1) and an open (2) conformation [25, 33, 138]. Histone variants can influence this process which likely plays a very important role in the process of histone H2A–H2B dimer exchange (3) involved in chromatin metabolism. The different intensities of the grey arrows depict the importance of the H2A variants in shifting the equilibrium towards an open (1) or a tightly bound histone octamer (2). The shorter arrows in the case of H2A.X–H2B dimers indicate the low level of H2A.X exchange in the cell [77]. The question mark indicates where the relationship is still in question. Histone H2A–H2B displacement actively participates in replication [140], transcription [141], and DNA repair [142]. Displacement of the histone H2A–H2B dimers leads to a characteristic enhanced nuclease sensitivity in the region close to the dyad axis of the nucleosomal DNA (black arrowhead) [183]. (1) and (2) have been modelled after the crystallographic structure of the NCP [18]. For clarity, the two H2A–H2B dimers have been removed in (3).
To date, histone H2A.Z is one of the few variants that has been shown to be indispensable for survival [46]. At the functional level, the role of this histone variant has been [47] and still remains extremely controversial, as claims for its involvement in gene inactivation [48] and activation (or both) [49, 50] are currently being made. In strong support of the silencing role comes the recent association of H2A.Z with heterochromatin binding protein (HP) (i.e. HP1α) [51]. Equally supportive of an active role in transcription is the finding of the association of this variant with chromatin remodelling complexes involved in transcription such as SWR1 [52–55].

The functional controversy has been mirrored by the structural characterization of NCP consisting of H2A.Z. The first crystallographic structure of H2A.Z-containing NCPs suggested the existence of a ‘subtle destabilization of the interaction between the (H2A.Z–H2B) dimer and the (H3–H4) tetramer’ and predicted that the coexistence of canonical H2A–H2B and H2A.Z–H2B dimers within the same nucleosome is unlikely due to changes in the interfaces between the two types of dimers [56]. In agreement with the NCP destabilizing role, it was shown that NCPs reconstituted with recombinant-H2A.Z–H2B dimers exhibited an ionic strength-dependent reduced stability as analysed in the analytical ultracentrifuge [57]. H2A.Z-containing NCPs appear to have an enhanced thermal-dependent mobility [58]. However, a more recent characterization of the salt-dependent stability using FRET by the same group that crystallized the NCP indicated that H2A.Z stabilizes the histone octamer within the NCP [59]. Suggesting a H2A.Z-chromatin stabilizing role are also the early results from Jim Davie’s lab using salt-elution of histones from hydroxyapatite (HAP)-adsorbed chromatin which showed that H2A.Z eluted at much higher salt concentrations than the canonical H2A–H2B dimer [60]. Nevertheless, all these data are seemingly hard to reconcile with the intrinsically reduced stability of the H2A.Z–H2B dimer itself [61].

Hence, like the functional studies, those dealing with the structural role of H2A.Z appear to be equally unsettled. A recent article from our group has shown the existence of two H2A.Z variants in chicken erythrocytes which only differ by three amino acids [62]. The presence of different subtypes of H2A.Z and/or the occurrence of differential PTMs (such as acetylation) [49, 63–65] may help to explain some of the structural and functional discrepancies described previously.

**Histone H2A.X—maintaining chromosome integrity**

Because of its connection with double-strand break (DSB) DNA repair and genome integrity, the functional role of this histone variant has been studied extensively [38, 66–68]. Histone H2A.X has also been involved in apoptosis [69], variable (diversity) joining [V(D)J] recombination [70], meiosis [71, 72] and replication [73]. Given all these important functional implications, it is not surprising that histone H2A.X has coevolved in parallel with the major canonical H2A variant [38]. Despite all this, the structural studies with this variant are lacking and only a limited amount of information is available. It is possible that like H2A.Z, an important structural role of this variant is mediated by its own PTMs and/or those of other histone nucleosomal partners [38] in addition to its interactions with non-nucleosomal partners such as the chromatin remodelling complexes Inositol 80-containing complex (INO80) and Swi2/Snf2-related adenosine triphosphate (SWR1) [74, 75].

In metazoans, H2A.X is evenly distributed throughout the genome with approximately one to two H2A.X molecules every ten nucleosomes [76]. FRAP results have shown that H2A.X has a low diffusional mobility [77]. Upon DSB damage, the H2A.X histones of extensive regions flanking the damaged site become reversibly phosphorylated [78]. Indeed, it was the discovery that the very C-terminal SQEY motif of mammalian H2A.X becomes phosphorylated by DNA-dependent protein kinase (DNA-PK) upon DSB damage [79] that brought much attention to this variant. Whether this phosphorylation has any effect on the structural characteristics of the NCP or the folding of the chromatin fibre remains yet to be elucidated. The only indirect information in this regard comes from a study carried out in yeast [80], where mutants were created in which the serine 129 of the native SQEL motif was replaced by glutamic acid (EQEL). The results [by supercoiling, micrococcal nuclease (MNase) digestions] obtained with these mutants suggested that chromatin adopts a more relaxed structure upon H2A.X phosphorylation [14, 80] which would be consistent with the homologous recombination (HR) repair mechanism in this...
organism. However, it remains to be determined whether this is also true in metazoans [which preferentially use non-homologous end-joining (NHEJ) for repair and HR in meiosis]. Whether or not this phosphorylation, which takes place in the vicinity of the histone H1 binding to the nucleosome, has any effect on such binding also needs to be established.

MacroH2A—a highly dynamic chromatin silencer

Like H2A.Bbd, mH2A is another recently discovered H2A variant, which, in contrast to the former, was first identified by its occurrence in the X-inactivated (Xi) chromosome of mammalian females [81–86]. Phylogenetically, mH2A has been a recent evolutionary acquisition which appeared during the evolution of vertebrates [38]. Approximately, one of every 30 nucleosomes in the cell contains one such variant [87]. The mH2A appears to be absent from the terminally differentiated cells such as the nucleated erythrocytes of fish, amphibian, reptiles and birds and from mature spermatozoa [88]. However, in addition to the inactive female X-chromosome, mH2A is also found in the tissues of vertebrate males that do not undergo this form of chromosome compensation [89]. Indeed, despite its absence from mature spermatozoa, mH2A has been found to be present during spermatogenesis (meiosis) [72, 90, 91] where it is associated with pericentric chromatin [92].

The mH2A is a heteromorphous H2A variant that has a tri-partite organization consisting of an N-terminal H2A-like histone-fold followed by a non-histone region (NHR) that contains a random coil and a C-terminal highly structured globular domain (Figure 1A). The latter has been called the ‘macro’ domain and shows high similarity to yeast macro domains [93]. Recent experimental evidence has shown that this domain can differentially bind ADP-ribosyl metabolites [88, 93, 94]. The NHR comprises two-thirds of the molecular mass of the whole protein.

This histone variant exists as two isoforms which are the product of alternative gene splicing: mH2A1.1, that accumulates throughout differentiation and development, and mH2A1.2, that has a constant level of expression [95]. The splicing site is located in a portion of the gene corresponding to the ligand binding pocket, suggesting a difference in the ability of the isoforms to bind nucleotide ligands [94]. Though the structural differences are minor, mH2A1.2 cannot bind nucleotides while mH2A1.1 can bind NAD metabolites [94]. Nevertheless, the functional significance of this is still unclear.

At the structural level, the crystallographic structure of the reconstituted NCP consisting of the H2A-like portion [96] as well as the macro domain have been recently published [94, 96]. Sedimentation velocity analysis demonstrated that in solution, the mH2A-NCPs had a very asymmetric conformation as a result of the NHR extending outwards. This is also supported by the ready accessibility of DNase I at the site where the random coil domain of the NHR exits the particle close to the dyad axis [88]. Sedimentation analysis using sucrose gradients revealed that mH2A stabilizes the NCP [89] in agreement with previous results using salt-dependent dissociation of histones from HAP-adsorbed chromatin [87, 88]. In this regard, it is important to notice that this variant has been shown to interfere with transcription factor binding and switch/sucrose non-fermenting (SWI/SNF) nucleosome remodelling [97]. Two-dimensional (2D)-polyacrylamide gel electrophoresis (PAGE), western blot analysis of the native mH2A-containing NCPs from chicken liver revealed the existence of two distinct populations, one of which consisted of an acid-labile PTM which was conditionally ascribed to poly-ADP ribosylation [89]. The low electrophoretic mobility fraction was found to be highly enriched in the MNase-resistant heterochromatin fraction [89].

HISTONE H3 VARIANTS

Histone H3.3—the functional implications of small compositional changes

Histone H3.3 provides a very good example of homomorphous variation. In humans, only four amino acids are different in the composition of this histone when compared with the major canonical H3.1. In H3.3, S31 replaces A31 within the N-terminal domain and A87, G90, S96 replace S87, M90, C96 within the second helix of the histone-fold domain. During the course of evolution, the split between the two types appears to have occurred several times, and hence they cannot be considered separate lineages [98]. Despite all this, the two variants exhibit important functional differences. In contrast to H3.1, H3.3 has been shown to
accumulate in some tissues during development in both humans and mice [99, 100], and it is enriched in actively transcribed regions of the genome [101] where it replaces H3.1 during transcriptional elongation [102]. The variant H3.3 also plays a very important role in spermatogenesis [72, 103].

Part of the involvement of this histone replacement variant in chromatin dynamics is mediated by the chaperone protein histone regulator A (HIRA). The HIRA complex participates in the assembly of H3.3–H4 dimers [104] into NCPs during the processes of transcription [104], replication [104] and in the assembly of the male pronucleus immediately after fertilization of the egg by sperm [105]. Remarkably however, recent evidence suggests that the amino acid differences between this histone variant and the canonical variant H3.1 may be enough to facilitate the NCP assembly/disassembly process alone [45]. The molecular details of how such small sequence compositional variability can account for this observation awaits further study.

CENP-A—defining the nucleosome structure of the centromeres

The hallmark of centromeric chromatin, CENP-A is a histone variant which replaces H3 in the NCP [106] (see [107–110] for recent reviews). As with H2A.Z, this variant is indispensable for survival [111]. This variant has also been referred to as chromatin-associated protein CSE4 (capping enzyme suppressor 4-p; Cse4p) in yeast or centromere identifier (Cid) in Drosophila. It is one of the most rapidly evolving members of the histone H3 family which cannot be considered an independent lineage, but rather, an ‘orphan’ that has arisen several times during evolution [98] as a result of an adaptive process that affects the whole molecule [112].

From a structural point of view, the protein consists of a highly variable and essential N-terminal sequence (which varies inter-species in both length and composition) [113] followed by a histone-fold domain, which in the case of human CENP-A, displays 62% identity to the canonical H3 counterpart [114]. Mutagenic analysis in yeast has shown that both domains perform different essential roles [115].

At the nucleosome level, the information we have about the conformation imparted by this variant has been sparse but forthcoming in recent years. In vitro reconstitution experiments with CENP-A purified from HeLa cells were able to produce NCPs with DNase I footprints and AFM images almost identical to those of native NCPs [116]. Using deuterium exchange/mass spectrometry and hydrodynamic analysis, it has been recently shown that CENP-A and histone H4 form tetramers that are more compact and rigid than those assembled from canonical counterparts [117]. These features have been attributed to the synergistic action of the first loop (L1) and second α-helix (α-2) of the histone-fold of CENP-A and they are suggested to be responsible for targeting the CENP-A–H4 complexes to centromeres [117] in a way that is independent of the underlying DNA sequence. It is important to notice here that the sequence variability of H3.3 also affected the α-2 region of the corresponding histone-fold.

At the level of the chromatin fibre, the information available is much more limited. It is worth mentioning in this regard, that the long N-terminal region of histone H3 interacts with the linker DNA and is involved, together with histone H1, in the maintenance of the higher order structure of the chromatin fibre [118, 119]. Hence, it is quite possible that this region of CENP-A, which has been postulated to bind to the narrow groove of DNA [120], plays a critical role in further defining the topological characteristics of centromeric chromatin which has been shown to be more compact than pericentric or bulk chromatin [121]. Whether this is further facilitated by the additional binding of other centromeric proteins (CENP-B, CENP-C, CENP-H, CENP-I) [122, 123] and/or requires histone H1 has not yet been established.

OTHER CORE HISTONE VARIANTS

The sperm-specific H2B variants

In great contrast to H2A and H3 histones, H2B and particularly H4 exhibit a low level of amino acid sequence variability, a fact that has been ascribed, at least in part, to the preferential role that these histones play in maintaining the interactions that hold together the histone octamer within the NCP [124]. Still, histone H2B has a relatively long N-terminal tail that protrudes from the NCP between the two DNA gyres [at superhelix location (SHL) 4.5 and −2.5/−3] [18], can form inter-nucleosomal histone–DNA interactions [125] and plays an important role in NCP mobility and dynamics [126, 127]. Therefore, histone variation within this domain may have important chromatin conformational implications.
Interestingly, most of the variability observed in H2B seems to occur in the sperm of vertebrate [90] and invertebrate [128] organisms and in male gametes in plants [129, 130]. In humans, three testis-specific variants have been described: TH2B [131], human testis-specific H2B (hTSH2B) [132] and H2B family member W testis-specific (H2BFWT) [133].

The long N-terminal domains of echinoderm sperm histones consist of highly characteristic repetitive motifs, SPR/KR/K [128], that are well-known to bind to the minor groove of DNA [134]. These tails may participate in maintaining a distinctive NCP organization [135] leading to a high extent of chromatin compaction observed in the sperm of these organisms. This compaction probably results from partial DNA charge neutralization and by establishing internucleosomal interactions.

As for vertebrate variants, hTSH2B is the only H2B human sperm histone variant whose role in nucleosome conformation has been recently characterized in detail [136]. As in the case of H2A.Bbd, it was found that hTSH2B lowered the stability of the histone octamer without affecting its ability to produce NCPs that were structurally and dynamically indistinguishable from particles consisting of canonical H2B [136]. The source of sequence variation of hTSH2B resides mainly in the occurrence of amino acid transitions to S/T residues. It is thus possible that the structural effects of this H2B variant on the NCP and chromatin fibre are phosphorylation mediated. However, despite this rather isolated work, structural studies on other vertebrate sperm H2B variants are for their most part still lacking, and hence their implications for chromatin are unknown.

In an interesting twist to the H2A.Z research, it has very recently been found that in Trypanosoma brucei, this histone variant pairs with a rather specialized H2BV variant [137]. The intriguing possibility of this being the case in higher eukaryotes deserves further exploration. As pointed out by the authors of this article, there are several potential H2B candidates in the human genome for this role.

**CORE HISTONE VARIANTS AND CHROMATIN DYNAMICS**

In solution and within the cell, nucleosomes are highly dynamic structures. In *vitro*, the conformation of the NCP fluctuates between folded and unfolded states [25, 33] (Figure 1B) with an affinity of the histone octamer for the template DNA that depends on both temperature and octamer concentration [138] as well as the ionic conditions [139]. *In vivo*, most of the metabolic functions of chromatin involve the dynamic exchange of histone H2A–H2B dimers. There is now evidence of different sorts to support this exchange during DNA replication [140], transcription [141] and repair [142]. Although in many instances the process is facilitated or assisted by ATP-dependent [143, 144] and independent chromatin assembly complexes (such as NAP-1, nucleoplasmin, chromatin assembly factor-1(CAF-1) or SWI/SNF, or facilitates transcription (FACT) just to mention a few) [145–148], the intrinsic NCP plasticity in solution and the incorporation of histone variants also play a critical role.

The sections described above provide a glimpse of how core histone variants themselves can actively participate in nucleosome dynamics. Core histones of the H2A family of variants provide several good examples. The FRET studies on H2A.Bbd [43] indicate that this variant can exchange very readily between different chromatin domains in a way that most likely drives the equilibrium between the closed (1) and open conformation (2) of the NCP shown in Figure 1B towards an open nucleosome conformation (2) as indicated by the thickness of arrow 2 in this figure.

In contrast to H2A.Bbd, the salt resilience toward NCP destabilization displayed by mH2A [88, 89] suggests that in this instance the equilibrium is shifted in the opposite direction (Figure 1B). It is important to note here that the shift in the equilibrium appears to be only temporary. The presence of mH2A during spermiogenesis and its absence from terminally differentiated cells suggests that the inactivation role of this variant is transient. It is likely that this variant dynamically exchanges to temporarily maintain certain repressed states of chromatin, but that other variants are responsible for a more permanent structures. In agreement with this, it has been shown that mH2A exhibits a mutually exclusive relationship with linker histones, suggesting some functional redundancy between these proteins [88].

Despite its even distribution throughout the genome and the high extent of similarity to the canonical form (Figure 1A), H2A.X appears to exhibit a highly restricted mobility [77] (Figure 1B). With H2A.Z the situation appears to be more complex. The most recent structural results imply an intra- and
inter-nucleosomal stabilizing role for this variant [27, 149] that in the latter instance has been ascribed to an enhanced interaction between the acidic C-terminal patch of this variant and histone H4 of adjacent nucleosomes [51]. Nevertheless, although supported by early crystallographic data [18], this interpretation contrasts with the polymorphic packing of nucleosomes within the recently published structure of a tetranucleosome complex [150], which is likely to be enhanced within the more structurally heterogeneous chromatin fibre.

A theoretical study using elastic network models [151] based on the plethora of NCP crystal structures currently available [19] has shown that in general, histone variants exhibit higher motilities and weaker correlations between internal motions in NCPs than those displayed by the canonical counterparts [151]. However, the extent with which the amino acid changes involved in histone variation affect the protein–protein interfaces and local alterations of the ionic environment [152] to contribute to this enhanced plasticity, still need to be experimentally determined. This knowledge is imperative for a complete understanding of the detailed molecular events involved in the dynamic equilibria shown in Figure 1B.

**HISTONE H1—MICROHETEROGENEITY MATTERS**

Eleven different linker histone variants have been identified to date in mammals: Seven somatic variants: (H1.1–H1.5, H1x and H1x) [4, 153–155], three spermatogenic variants: H1t [156], H1T2 [157] and HILS1 [158, 159] and an oocyte-specific H1foo [160]. Highly specialized histone H1 variants are also expressed in the sperm of many invertebrates including the long histone H1s from the sperm of echinoderms [128] and the highly specialized protamine–like protein I (PL-I) that are thought to be evolutionary linked to protamines (see Eirin-Lopez et al. [161] for a recent review). Except for H1.1–H1.5, most of the other histone H1 variants are expressed in a replication-independent cell cycle and development-dependent way [162]. In this section, I am going to focus mainly on recent functional and structural aspects of somatic histone H1 microheterogeneity. For a recent review on developmentally regulated histone H1 variants, the reader is referred to Khochbin [162].

Interestingly, the concept of histone variants within this family and of histone sequence variability itself had its origins in the early work on calf thymus H1 histones by Kinkade and Cole [163, 164] who noted that somatic histone H1 exhibited sequence microheterogeneity (Figure 2A). Although Cole [165, 166] himself wrote several reviews about the possible implications of this histone H1 microheterogeneity (see also for a more recent review [167]), evidence for a multifaceted functionality of this family both at the level of somatic microheterogeneity (H1.1–H1.5) (Figure 2A) and at the level of differentiation-specific variants (H10, H1t) has long remained elusive. Indeed, and quite unexpectedly, experimentally costly knock out-experiments carried out in different organisms in recent years seem to suggest that the function of this variability may have a large overlapping redundancy (reviewed in [4, 168]).

Despite this, the non-random distribution of linker histone variants in the genome has now been well-documented [169] and mammalian linker histone variants have been shown to differentially affect gene expression in vivo [170]. Not only that, but histone H1.2 is able to preferentially bind to a regulatory sequence of the gene for histone H3.2 [171] has been directly involved in DNA DSB-induced apoptosis [172], and interacts with Msx1 (a transcription factor involved in myogenic gene expression) [173]. Furthermore, there is recent evidence showing that the long-term evolution of the replication-dependent and replication-independent somatic linker histone variants (H1.1–H1.5) has taken place though a process of ‘birth-and-death’ evolution with a strong purifying selection [174, 175]. Accordingly, the degree of interspecific conservation of these variants is higher than that observed at the intraspecific level [174]; an observation which also comes in support of specialized structural and functional roles of these variants. Thus, although the functional specificity of linker histone variants does not appear to be as well-defined as with core histone variants, all of the above evidence suggests that their structural variability is functionally relevant.

The less defined functional character of linker histones can perhaps be better understood from the structural details of their association with chromatin. In comparison with core histones, linker histones have long been known to be less tightly associated with chromatin and much more mobile [176], a fact that has been recently corroborated by in situ FRAP experiments [21, 22]. The rapid and dynamic exchange of histone H1 between different segments
Figure 2: Histone H1 microheterogeneity and chromatin dynamics. (A) Amino acid sequence alignment of the human somatic H1 histones (H1.1–H1.5) to highlight the extent of microheterogeneity between these proteins. The degree of similarity among the different sequences is proportional to the dark intensity of the shaded areas. The nomenclature followed for the designation of these variants is that of Albig and his colleagues [154] with the nomenclature of Seyedin and Kistler [184] shown in parentheses. The α-helices and β-turns of the winged helix globular domain are indicated. (B) Schematic representation of the dynamic model of linker histone binding to chromatin [177]. Histone H1 microheterogeneous variants (black dots) are continuously exchanging at different rates [178] between folded and unfolded regions of chromatin (indicated by the reversible arrows). Local concentration may change in response to different stimuli (initiation of transcription, transcriptional elongation, DNA repair, DNA replication) that may involve selective PTMs of the linker histones being exchanged as well as core histone PTMs and the differential presence of core histone variants themselves. This dynamic process is in addition influenced by a broader network of chromatin interacting proteins [177, 180] such as heterochromatin binding proteins (HP) (i.e. HP1-α) and methyl binding domain proteins (MBD) such as MeCP2 and high mobility group proteins (HMG).
of chromatin (Figure 2B) is a very important determinant of the extent of fibre folding and interchromatin fibre association. A model for the dynamic association of linker histones with chromatin [177] is shown in Figure 2B. Accordingly, histone H1 is dynamically associated with the chromatin fibre in a transient mode. Recent FRAP experiments have shown the existence of ‘dramatic’ differences in the binding affinities of the different H1 subtypes [178]. Hence, their movement and local distribution is likely to vary from one type to another. It will additionally depend on other factors such as: core histone variant (and PTM) composition of individual nucleosomes [14] along the fibre, transcription activation pathways [177], and/or the extent of the histone H1 PTMs themselves (i.e. phosphorylation) [178, 179]. Furthermore, linker histones may operate in conjunction with a ‘network’ of other chromatin-binding proteins such as high mobility group (HMG) proteins, HP1 or methyl CpG-binding protein (MeCP2) [177,180] so as to define permissive (euchromatin) and repressive (heterochromatin) DNA domains. Therefore, as it was stated at the beginning of this section, if the main role of linker histones is to maintain and stabilize the folding of chromatin, it should not be entirely surprising that there is some degree of redundancy despite their well-supported structural and functional specialization.

CONCLUSIONS
Structural evidence gathered so far indicates that the amino acid sequence variability of histones constrained and maintained throughout evolution can, by itself or in conjunction with PTMs, regulate chromatin dynamics in a way that has important functional implications. The epigenetic language of the histone PTMs (histone code) is ultimately determined by this prevalent underlying (amino acid code) of the different histone variants. At the core histone level, a representative example of this can be found in the SQE[Y/L/F] motif at the C-terminal end of H2A.X that provides a sequence motif specificity for the DNA-PK involved in double-stranded DNA repair [38]. At the linker histone level, the yet to be clearly established sequence motif responsible for the apoptotic specificity of H1.2 [172] provides another example.

The histone-fold, the winged helix domain and the ‘disordered’ N- and C-terminal tails [181] are ‘simple’ structures that could be achieved through many different amino acid combinations. The fact that the amino acid locations in the molecule are so conserved throughout evolution within the different variants indicate that most of the amino acid variation observed is functionally relevant. Deciphering the details of the amino acid code responsible for the manifold structural and functional implications of all the histone variants such as, for instance, the amino acid motif involved in the recognition of H2A.Z by SWR1 and/or INO80 may still take some time. This effort, however, may prove very rewarding in understanding the overall epigenetic dimension of these ‘fairly simple proteins’ we call histones.

In the interim, it may also be worthwhile to pursue some additional questions raised by several sections of this review. For instance, do the different docking domains of H2A variants prevent them from forming heterologous NCPs consisting of H2A–H2B dimers with mixed variants? What are the histone H2B variant partners of the different H2A variants? In this regard, the recent finding that H2A.Z interacts with a highly specialized H2BV in T. brucei [137] is very insightful and hopefully similar analyses carried out in metazoans will soon follow. What is the structural and functional relevance of H2A.Z sequence microheterogeneity? Does the compositional microheterogeneity of H3 variants affect the structure of NCPs? Given the instability of the H2A.Z–H2B dimer, how does the replacement of the canonical dimer take place? In the case of H3.3, it appears that this variant interacts with chaperone proteins that may mediate its exchange with canonical H3 during elongation in actively transcribed genes. The recent observation that this variant may operate synergistically with other variants such as H2A.Bbd to facilitate the assembly and disassembly process [45] sheds important light on some of the possible mechanisms involved. However, considering that H3.3 is by far, much more abundant than H2A.Bbd, the full implications of H3.3 involvement in the loss of NCP stability and the molecular details of the exchange process need yet to be established.

Considering the intensity of the research, currently ongoing, on the topic of this review which can be easily attested by the large number of other recent reviews in the same area of chromatin research [38, 39, 124, 152, 182], the answer to several of these questions are likely to be forthcoming soon.
Histone variants

Key Points
- Chromatin is a dynamic nucleoprotein complex whose structural characterization requires the use of appropriate biochemical tools (such as, for instance, analytical ultracentrifugation, FRAP and FRET) in addition to the powerful static data provided by X-ray crystallography.
- The histone H2A family includes a broad spectrum of variants whose functional implications for chromatin have been extensively characterized. However, the structural aspects behind the function are in many instances still not clear and/or controversial (i.e. H2A.Z and phosphorylated H2A.X).
- Core histone variants and their PTMs play a critical role in the modulation of chromatin dynamics.
- Histone HI microheterogeneity exhibits a non-random distribution in the genome and is an important component of chromatin metabolism. The latter is mediated by the highly dynamic differential association of the distinct HI variants with chromatin.

Acknowledgements
I would like to thank Ron Finn and José María Eirín Lopez for their skilful computer assistance and help in the preparation of the figures and to Wade Abbott, Lindsay Frehlick and Anita Thambirajah for their critical and careful reading of the manuscript. This work was supported by Canadian Institutes of Health Research (CIHR) grant MOP-57718.

References


