Computational, biochemical, and NMR-driven structural studies on histone variant H2A.B

Zhang, H.

Chapter 1. General introduction

The nucleosome, a DNA-protein complex, is well known today as the basic unit of chromatin. However, the discovery and characterization of it took more than a hundred years of unveiling each building block of the nucleosome step by step. The existence of histones, the protein components of the nucleosome, was first recognized in 1884 by Kossel [1]. It was not until the 1960s that the five histone types were observed [2], now known as H1, H2A, H2B, H3 and H4. DNA, on the other hand, was first isolated by Miescher in 1869, and its double-helix structure was revealed by Watson and Cricks more than eighty years later in 1953 [3]. With the identification and characterization of both protein and DNA components available, studies emerged in the 1970s to understand the fundamental basis of chromatin structure [4-7]. In 1974, a complex of eight histone molecules and about 200 DNA base pairs was proposed by Kornberg as the repeating unit of the chromatin structure [8]. Finally, twenty-three years later in 1997, the first high resolution crystal structure of the isolated nucleosome core particle was solved by Luger *et al*. [9] Two copies of each of the four core histones H2A, H2B, H3, and H4 form a histone octamer, around which wraps 147 bp of DNA to form the basic unit of chromatin: the nucleosome (Figure 1.1a). These nucleosomes can be further coiled into thicker chromatin fibers, which are further compacted and folded into the typical X-shaped chromosomes during metaphase. This packaging of DNA is dynamic and modulates nuclear processes. At lowest level, the presence of nucleosomes forms a roadblock, limiting DNA accessibility for polymerases, transcription factors, and other DNA binding proteins. Modulation of nucleosome structure and position allows the manipulation of this roadblock and thus ultimately the regulation of DNA transcription or repair. A vast range of proteins cooperate to achieve this regulation. Even if the resulting chromatin biology is highly complex, just like mosaic art, no matter how splendid it is, the art piece is made of simple building blocks. And just as the color and shapes of these building blocks determine the range of mosaics that can be made, the fundamental molecular properties of the nucleosome and chromatin factors determine the dynamic network of interactions that underlie chromatin biology. In this regard, the determination of the crystal structure of nucleosome in 1997 has truly been a scientific breakthrough discovery (Figure 1.1a) [9]. Since then it has been shown that the architecture and structure of the nucleosome

is nearly independent of DNA sequence and constant among different species. Yet the efficiency of this roadblock can be altered through a number of mechanisms, most notable through incorporation of post translational modifications (PTMs) and the replacement of the canonical histones by histone variant proteins.

PTMs alter nucleosome structure and function

PTMs are covalent modifications that can be dynamically added on and removed from residue side chains as a means to regulate protein function in response to the cell environment. While histone PTMs are mostly known for serving as binding anchors for various chromatin factors, a number of histone PTMs, in particular those that occur in the histone core rather than the tail, act by influencing nucleosome structure directly. Core modifications often alter histone-DNA and/or histone-histone interactions thus changing nucleosome dynamics and stability. For example, acetylation of H3 lysine 56, which is located close to the nucleosomal DNA, interrupts the local histone-DNA interaction, increases DNA accessibility, and ultimately affects gene expression and DNA repair [10-11]. Similarly, acetylation of H4 lysine 91 removes an essential salt bridge to H2B E68 in the tetramer-dimer interface, resulting in higher sensitivity toward DNA digestion by micrococcal nuclease [12]. This modification occurs before the H3-H4 tetramer is assembled on DNA and is thus suggested to regulate nucleosome assembly[13]. These examples illustrate how small structural changes, i. e. addition of a simple acetyl group, can have pronounced effects on overall nucleosome stability and function.

Histone variants alter nucleosome structure and function

The largest structural changes are brought about by substitution of the canonical histones by variant histones with different amino acid sequence. For example, macroH2A, an H2A variant, contains an additional ~250 amino acids at its C-terminus that includes a macro domain and thus offers new functionalities to the nucleosome [14-15].

Unlike canonical histones, whose synthesis and assembly into nucleosomes is highly coupled with DNA replication, histone variants are expressed and incorporated into chromatin in a replication-independent manner. In addition, histone variants are usually encoded by single copy genes whereas the canonical forms are encoded by multiple copy genes. Variants are less abundant but play a vital role in regulating chromatin function by replacing their canonical forms from nucleosome at defined genomic loci. For example, centromeres contain specialized nucleosomes that include histone H3 variant CENP-A. Due to the presence of this variant these nucleosomes have an altered structure and can form interactions with other centromeric proteins. The crystal structure of the centromeric nucleosomes (pdb: 3AN2) showed that it only organizes 121 bp DNA due to the shorter αN helix of the variant [16] and the resulting higher flexibility of entry/exit nucleosomal DNA may facilitate the binding of centromeric DNA-binding proteins.

The first histone variant, sperm H2B, was identified in the 70s of 20th century [17]. There are now in total 18 groups of histones variants recorded in Histone Database for H2A, H2B, H3 and H4 [18]. Overall, H2A and H3 have the most variants, while H2B and H4 have only a few variants. Sequence differences to the canonical form can be large, such as in the case of macroH2A, but are typically limited to few amino acid substitutions. For example, histone variant H2A.X contains an additional C-terminal motif that contains a serine that is phosphorylated in the response to DNA damage, thereby signaling and recruiting DNA repair proteins to the damage site [19].

The incorporation of histone variants is mediated by nucleosome remodelers and dedicated histone chaperone proteins. For H2A.Z, a variant involved in transcription regulation and mostly positioned at promoters, it has been shown that Swr1, part of the SWR1 remodeling complex, can load the H2A.Z-H2B dimer in two steps on a tetrasome, the complex of DNA and the H3-H4 tetramer [20]. SWR1 is thought to first unpeel the DNA from the nucleosome, using energy from ATP, releasing the canonical H2A-H2B dimers in the process, after which the variant dimer is incorporated.

The structures of several variant nucleosomes have been determined and some have shown surprisingly little structural changes to the canonical form. While a wide array of biochemical and biophysical

data have indicated that H2A.Z forms mobile and dynamic nucleosomes, the crystal structure of the H2A.Z nucleosome (pdb: 1F66) is virtually identical to that of the canonical nucleosome [21], highlighting the need for complementary studies that are sensitive to dynamics.
This thesis is focused on H2A.B, which is one of the most divergent histone variants [22]. Incorporation of H2A.B into nucleosomes has been shown to induce a more open conformation of the nucleosome in which the DNA is more accessible [23]. Functionally, it has been linked to various processes, including RNA splicing and active transcription [24]. To provide a solid fundamental basis to understand its biological function, a study on the structure and dynamics of the H2A.B-H2B dimer and the H2A.B nucleosome is described in this thesis. The remainder of this introduction describes the current knowledge on this enigmatic histone variant.

Discovery of histone variant H2A.B

As human females have two X chromosomes, while males have one X and one Y chromosome, one of the two X chromosomes from female somatic cells is randomly subjected to permanent inactivation to ensure comparable gene expression from the X chromosome [25]. This sex-determining system is consistent in most of mammals. Inactivation of one of the X-chromosomes is achieved by heterochromatin formation and further compaction and results in the formation of so-called Barr bodies in the nucleus. While the exact mechanism of inactivation is not fully resolved yet, a crucial role is played by the long non-coding RNA Xist [26]. In addition, the inactive X (Xi) chromatin contains specialized nucleosomes with H2A variant macroH2A[27]. This variant is thought to promote the repression of transcription by amino acid substitutions within the histone fold that stabilize the nucleosome, as well as by promoting recruitment of histone deacetylases (HDACs) via the macrodomain [28]. In the hope to find a counterpart to macroH2A that is repelled from Xi, Chadwick and Willard searched the genome database and found a group of distant H2A homologs [22, 29]. Expression of the identified gene was confirmed by Northern blot and reverse transcription PCR in a variety of cell lines and tissues. Transfection of

a GFP fusion protein into primary human fibroblasts showed that it was excluded from the Xi chromosome in female cells. Further analysis proved that the protein is a nucleosomal core histone [22]. The identified protein was originally named H2A.Bbd, for Barr-body deficient H2A variant. The simplified name H2A.B has been slowly adopted during the past two decades. H2A.B contains 114 amino acids and shares only 48% identity to canonical H2A, making it one of the most divergent histone variants up to date.

The sequence of H2A.B differs strongly from canonical H2A

Up to date, fifteen H2A.B sequences have been reported over twelve mammalian species, and their sequence alignment against human canonical H2A is shown in Figure 1.1c [18]. Three key differences between the H2A.B family and canonical H2A emerge from sequence comparisons. First, most sequence differences occur in the N-terminal tail of H2A.B, which is also a highly variable region within the H2A.B family. A common overall feature of the H2A.B N-terminal tail is the lack of lysine residues as opposed to canonical H2A. Histone lysines are common targets for PTMs, such as acetylation, methylation, and ubiquitination, which are essential in defining chromatin function. In the case of H2A, N-terminal K13/15 ubiquitination plays a crucial role in the DNA damage response [30-31], while N-terminal K5 acetylation is essential in transcription regulation and chromatin remodeling [32]. Notably, also the histone fold domain of H2A.B is scarce in lysines, with only one lysine present instead of four in the canonical H2A. Together, this absence of modifiable residues suggests that the variant is regulated in a manner different from other members of H2A family and that it exerts its specific impact on chromatin function independent of PTMs.

The second key difference in H2A.B is its missing C-terminal tail and truncated C- terminal docking domain. Absence of the C-terminal tail again removes a number of well-known modification sites in H2A, in particular K118/K119 sites that are involved in transcriptional repression [33]. In the canonical nucleosome, the H2A docking domain mediates the interaction between H2A-H2B dimer and H3-H4 tetramer

[9]. In particular, the segment absent in H2A.B binds to the H3 αN helix in the canonical nucleosome. Furthermore, the truncated docking domain in H2A.B contains several sequence differences compared to H2A, raising the question whether this will also translate into structural differences in the H2A.B nucleosome.

A third common feature of H2A.B compared to other H2A histones is the reduced negative charge of the acidic patch, a negatively charged area on nucleosome surface formed by six H2A residues (E56, E61, E64, D90, E91, and E92) and two H2B residues (E102 and E110) [9]. The acidic patch functions as a binding surface for the H4 tail from a neighboring nucleosome, thereby mediating higher order chromatin folding [34]. Moreover, the acidic patch is the key binding site for a wide variety of nucleosome binding proteins, such as LANA, RCC1, SIR3, HMGN2, etc. [35-38]. Just as the H4 tail, these proteins all use a strategically located arginine residue to form hydrogen bonds with acidic patch residues[39]. In human H2A.B three of the six residues that form acidic patch are either neutralized or charge-swapped: E61 to K, E91 to R, and E92 to L. Within the H2A.B family, several species have also a D90 to N substitution, while mouse H2A.B retains E92 in its sequence shifted by one residue position. The reduced negative charge on H2A.B-nucleosome surface will disfavor or even inhibit binding of the many acidic-patch binding chromatin factors including the H4 tail, thus affecting chromatin function and higher-order structure (see also below).

Last but not least, it has been noticed that H2A.B is less basic compared to H2A, due to the loss of lysines and a series of substitution of charged residues [23, 40] (see Figure 1.1c). The reduced electropositive potential may affect DNA binding directly. Moreover, the extensively altered pattern of charged residues in H2A.B may influence the stability of its dimer with H2B and/or its octamer with H2B, H3, and H4 [41]. A more detailed examination of this hypothesis is discussed in **Chapter 3**. In addition, we note that one of minor groove anchoring arginines in H2A is shifted one residue position in most H2A.B sequences, which may alter the interaction with nucleosomal DNA. This is further investigated in **Chapter 4**.

H2A.B induces opening structure at DNA entry/exit ends of nucleosome

To investigate how the sequence differences in H2A.B impact nucleosome structure, Bao et al. reconstituted H2A.B nucleosome *in vitro* and studied the organization of the nucleosomal DNA using fluorescence resonance energy transfer (FRET) experiments and micrococcal nuclease (MNase) digestion assays [23]. Both methods indicated that H2A.B nucleosome have a relatively open structure in which DNA entry/exit ends are far away from each other, even at low ionic strength. In particular, the MNase digestion assays showed that only ~118 bp were left intact after digestion of the free DNA ends while 146 bp are protected in the conventional case. Similar results were obtained later using the high-affinity, artificial '601' DNA sequence rather than the 5S DNA sequence [42]. Subsequent studies using atomic force microscopy (AFM) confirmed that less DNA is tightly bound to the histone octamer core in H2A.B nucleosomes, but arrived at ~130 bp DNA being protected [42-43], possibly due to different experimental conditions. Electron cryo-microscopy, atomic force microscopy, SAXS, and small-angle neutron scattering experiments have further confirmed the less organized DNA entry/exit ends in the variant nucleosome (see Figure 1.1b) [42, 44-46]. In an attempt to dissect the structural basis for the opened structure, various mutation studies have been conducted. A chimeric protein made by replacing H2A N-terminal tail with the H2A.B N-terminal tail showed no changes of the DNase digestion profile, proving that the elongated N-terminal tail of H2A.B has no impact on the opening structure of the variant nucleosome [42]. Similarly, a chimeric protein containing H2A histone fold and the H2A.B docking domain as well as a series of H2A mutants in which the docking domain was truncated at different positions (marked as triangles in Figure 1.1c) were constructed [23, 42, 47]. These studies indicated that the truncation itself is not responsible for the open structure of H2A.B nucleosome. They also suggested that this region does not significantly contribute to nucleosome stability, but rather that the sequence difference within the truncated domain are responsible for this. In addition, replacement of the H2A.B docking domain with the H2A docking domain could rescue the orientation of the nucleosomal DNA entry/exit ends, further supporting the crucial

role of H2A docking domain in stabilizing the nucleosome structure [42]. It should be noted that it has not been tested whether extension of the H2A.B docking domain with the missing H2A C-terminal region can similarly rescue the formation of a closed nucleosome.

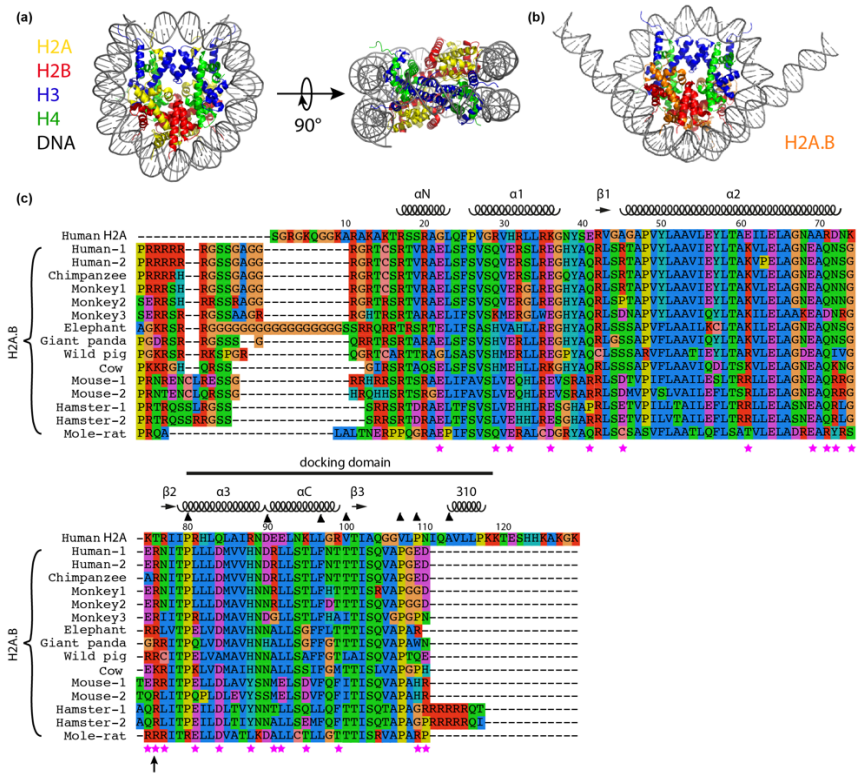

Figure 1.1. Nucleosome structure and sequence alignment of histone variant H2A.B with canonical H2A from human. (a) Top (left) and side (right) views on the crystal structure of canonical nucleosome (pdb: 2PYO). Color coding of the histones indicated in the figure. **(b)** Model of H2A.B nucleosome, made by opening the 15 bp of entry/exit DNA ends of the canonical crystal structure. **(c)** Sequence alignment of fifteen H2A.B sequences with human canonical H2A. Secondary structure elements based on the nucleosome crystal structure are indicated above the sequence. The docking domain of H2A is marked as black bar. Substitutions that alter the charge within the histone fold domain, based on the difference of human H2A.B and H2A, are marked as pink star below the sequence. The shifted minor groove arginine is marked as black arrow below the sequence. C-terminal docking domain truncation sites from literature are marked as black triangles above the sequence.

H2A.B prevents chromatin fiber formation

To examine the impact of H2A.B on higher-order chromatin structure, a few studies investigated the folding of H2A.B nucleosomal arrays. As discussed above, H2A.B nucleosomes will have a reduced acidic patch on their surface. Since the acidic patch is a crucial binding platform for the N-terminal H4 tail from an adjacent nucleosome to mediate the higher-order compaction of chromatin fibers [48], the loss of the three glutamic acids in the acidic patch of H2A.B nucleosomes can be expected to interrupt the nucleosome-nucleosome interactions. Indeed, H2A.B nucleosomes arrays were shown to fold as "beads on a string", rather than a compacted fibre [49], in sharp contrast to arrays of histone variant H2A.Z nucleosomes, which fold more readily into fibres than canonical ones due to the more extended acidic patch of H2A.Z [50]. Another study demonstrated by ultracentrifugation sedimentation experiments that an H2A.B mutant in which the three glutamic acids are restored, H2A.B-EE/E, fold as a fibre and that this folding of a H2A.B-EE/E nucleosome array is due to the restored acidic patch and not due to stably wrapping of 146 bp DNA within the mutant nucleosome [34]. *In vivo* data for chromatin conformation containing H2A.B nucleosomes is lacking. However, H2A.B was found to be enriched at the actively transcribed genes in cells, which suggests H2A.B incorporation is related to a less compacted chromatin structure *in vivo* [51-52].

Assembly, disassembly and remodelling of H2A.B nucleosomes *in vivo*

In the studies described above H2A.B nucleosomes were reconstituted *in vitro*. While canonical nucleosomes can be reconstituted from refolded histone octamers, the H2A.B containing octamer is not stable, so that reconstitution is carried out from refolded H2A.B-H2B dimers and H3-H4 tetramers that are deposited on DNA using salt gradient dialysis. How H2A.B nucleosomes are formed *in vivo* is yet unknown. While histone variants are typically incorporated by dedicated chaperones and/or remodelers, no such protein has been identified for H2A.B. However, histone chaperone NAP-I can mediate reversible

dimer exchange of H2A.B-H2B from the nucleosome at physiological ionic strength *in vitro* [53].
Interestingly, H2A.B-H2B dimers may be much more dynamically incorporated than for canonical dimers. Photobleaching experiments on cells expressing GFP-H2A and GFP-H2A.B showed faster fluorescence recovery for H2A.B at the photobleached area, indicating higher mobility of the variant compared to the canonical form [45, 54]. Such high dynamic transfer of H2A.B is perhaps directly related to the instability of H2A.B nucleosome as measured by sedimentation and force-extension experiments [42, 54].
Viewing the H2A.B nucleosome as a more labile roadblock, one could hypothesize that there is also a reduced need for remodeling of the nucleosome, at least for remodelers that typically open chromatin such as SWI/SNF. Indeed, while SWI/SNF binds to the H2A.B nucleosome with the same efficiency as to the canonical nucleosome, it is not capable of remodeling the variant nucleosome [44]. Experiments using domain swapped mutants showed that neither the elongated N-terminal tail of H2A.B nor docking domain are fully responsible for preventing SWI/SNF remodeling [42]. For the related RSC remodeler, the truncated docking domain of H2A.B turned out to play a larger role in the resistance to remodeling [47]. Still, these results suggest that the histone fold part of H2A.B plays a role in preventing the remodeling ability of at least SWI/SNF. Interestingly, since the acidic patch has been implicated in remodeling by the ISWI remodeler (unrelated to SWI/SNF), the reduced acidic patch in H2A.B may be a molecular reason for the failure in remodeling.

H2A.B associates with active transcription and RNA splicing

The open DNA entry/exit end conformation of H2A.B nucleosome, together with the less compacted chromatin structure of H2A.B nucleosome arrays suggests the DNA in the variant chromatin is more accessible and thus more amenable to transcription, DNA replication, and repair. Indeed, increased transcription rate has been demonstrated *in vitro* using the H2A.B nucleosome arrays compared to the canonical arrays [34]. An extensive study by Tolstorukov *et al*. using *in vivo* ChIP-

seq data from Hela cells showed that H2A.B was enriched in the bodies of actively expressed genes and its enrichment correlated with the gene expression level [51]. Since depletion of H2A.B caused both up- and down-regulated gene expression and even a net decrease in transcription, the mild effects of H2A.B depletion suggested H2A.B is not directly regulating transcription. The authors used mass spectrometry to identify proteins that are specific to H2A.B enriched chromatin, revealing a number of RNA processing factors and spliceosome components. By depleting H2A.B, the frequency of exons included in mature transcript was elevated and the intronic read density was increased, which are consistent with a decrease in splicing efficiency [51]. Similarly, mouse H2A.B was found to be enriched at actively transcribed genes and associated with mRNA splicing [24, 55-56]. In particular, one study demonstrated by using RNA-pull down assays that unlike canonical H2A or H2A.Z, the H2A.B N-terminal tail possesses RNA binding affinity and confirmed a direct interaction *in vivo* through cross-linking and immune-precipitation assays [56]. Further support for involvement in RNA splicing comes from H2A.B knock-out mouse study, showing loss of proper RNA Pol II localization and changes in pre-mRNA splicing [57]. Altogether there is substantial evidence for a role of H2A.B in RNA splicing, although the precise molecular details are still to be elucidated. In addition, H2A.B is also suggested to function in cell memory[58] and DNA repair[59].

H2A.B-like H2A variants: H2A.L, H2A.P and H2A.Q

In recent years, several other H2A variants have been identified within the mammalian lineage that like H2A.B have a shortened C-terminus with a truncated docking domain. These variants H2A.L, H2A.Q, H2A.P are closely related to H2A.B and supposed to share a common ancestor, H2A.R [60]. These short histone variants reside on the X-chromosome in most species [60-61] . As H2A.B, these short H2A variants lack the acidic patch that is characteristic for canonical H2A. Other features of H2A.B are mirrored also in the other members of the short H2A histone variant family. Indeed, H2A.L incorporated nucleosomes have been reported to associate with 130 bp DNA and its nucleosome array shows a beads on the string conformation [62]. Even more dramatic

than in H2A.B, extensive charge-altering substitutions in H2A.P reduce the theoretical isoelectric point below 5, where it is 11 for canonical H2A [60]. Two conserved arginine residues in H2A that contact the DNA minor-groove are lost in H2A.P and H2A.Q proteins, similar to the altered position of one of these in H2A.B. Finally, these short H2A variants do not appear in all mammals, suggesting they are functionally non-essential or redundant with each other and they may have different functions between different mammals [60].

Perspective and outline of this thesis

To summarize, H2A.B is one of the most divergent histone variants, which induces an open structure in its nucleosome and is involved in active gene transcription and mRNA splicing. While the functional role of H2A.B is slowly emerging, high-resolution structural characterization of the H2A.B nucleosome is still lacking. Yet, detailed knowledge on the structure and dynamics of the variant dimer and nucleosomes is essential for understanding its functional properties. While the crystal structure of the H2A.B-H2B heterodimer was solved very recently [63], a structure for the H2A.B nucleosome is still unavailable and unlikely to be solved at high resolution by crystallography or cryo-electron microscopy, due to the reduced DNA organization, higher dynamics and lower stability. In this work, we aimed to investigate the structure and dynamics of the H2A.B-H2B dimer and nucleosome at atomic resolution using NMR spectroscopy. Over the course of this decade, and in particular thanks to the introduction of NMR and isotope labeling techniques suitable for high molecular weight systems, nucleosomes have become amenable to NMR studies (for a recent review see reference [64]). **Chapter 2** reviews the state-of-the-art isotope labeling methods that have extended the capability of solution state NMR to study macromolecular complexes. **Chapter 3** describes the solution structure of the H2A.B-H2B as well as a mutational study correlating surface charge of the heterodimer to its thermostability. The structural and dynamics properties of the H2A.B nucleosome are reported in **Chapter 4**, highlighting the impact of a position change at a minor groove anchoring arginine in the variant nucleosome. Next, the acidity of the canonical acidic patch is

investigated in **Chapter 5** to provide experimental mapping of the electrostatics. The thesis is concluded with a general discussion of the findings in **Chapter 6** followed by a summary.

References

1. Kossel, A., Über eine peptoartigen Bestandteil des zellkerns. *Z. Phyciol. Chem.* **1884,** *8*, 511-515.

2. Phillips, D. M.; Johns, E. W., A Fractionation of the Histones of Group F2a from Calf Thymus. *Biochem J* **1965,** *94*, 127-30.

3. Watson, J. D.; Crick, F. H., Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature* **1953,** *171* (4356), 737-8.

4. Bram, S.; Ris, H., On the structure of nucleohistone. *J Mol Biol* **1971,** *55* (3), 325-36.

5. Sahasrabuddhe, C. G.; Van Holde, K. E., The effect of trypsin on nuclease-resistant chromatin fragments. *J Biol Chem* **1974,** *249* (1), 152-6.

6. Williamson, R., Properties of rapidly labelled deoxyribonucleic acid fragments isolated from the cytoplasm of primary cultures of embryonic mouse liver cells. *J Mol Biol* **1970,** *51* (1), 157-68.

7. Hewish, D. R.; Burgoyne, L. A., Chromatin sub-structure. The digestion of chromatin DNA at regularly spaced sites by a nuclear deoxyribonuclease. *Biochem Biophys Res Commun* **1973,** *52* (2), 504-10.

8. Kornberg, R. D., Chromatin structure: a repeating unit of histones and DNA. *Science* **1974,** *184* (4139), 868-71.

9. Luger, K.; Mader, A. W.; Richmond, R. K.; Sargent, D. F.; Richmond, T. J., Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature* **1997,** *389* (6648), 251-60.

10. Xu, F.; Zhang, K.; Grunstein, M., Acetylation in histone H3 globular domain regulates gene expression in yeast. *Cell* **2005,** *121* (3), 375-85.

11. Zhang, L.; Serra-Cardona, A.; Zhou, H.; Wang, M.; Yang, N.; Zhang, Z.; Xu, R. M., Multisite Substrate Recognition in Asf1-Dependent Acetylation of Histone H3 K56 by Rtt109. *Cell* **2018,** *174* (4), 818-830 e11.

12. Ye, J.; Ai, X.; Eugeni, E. E.; Zhang, L.; Carpenter, L. R.; Jelinek, M. A.; Freitas, M. A.; Parthun, M. R., Histone H4 lysine 91 acetylation a core domain modification associated with chromatin assembly. *Mol Cell* **2005,** *18* (1), 123-30.

13. Yang, X.; Yu, W.; Shi, L.; Sun, L.; Liang, J.; Yi, X.; Li, Q.; Zhang, Y.; Yang, F.; Han, X.; Zhang, D.; Yang, J.; Yao, Z.; Shang, Y., HAT4, a Golgi apparatus-anchored B-type histone acetyltransferase, acetylates free histone H4 and facilitates chromatin assembly. *Mol Cell* **2011,** *44* (1), 39-50.

14. Pehrson, J. R.; Fried, V. A., MacroH2A, a core histone containing a large nonhistone region. *Science* **1992,** *257* (5075), 1398-400.

15. Kozlowski, M.; Corujo, D.; Hothorn, M.; Guberovic, I.; Mandemaker, I. K.; Blessing, C.; Sporn, J.; Gutierrez-Triana, A.; Smith, R.; Portmann, T.; Treier, M.; Scheffzek, K.; Huet, S.; Timinszky, G.; Buschbeck, M.; Ladurner, A. G., MacroH2A histone variants limit chromatin plasticity through two distinct mechanisms. *EMBO Rep* **2018,** *19* (10).

16. Tachiwana, H.; Kagawa, W.; Shiga, T.; Osakabe, A.; Miya, Y.; Saito, K.; Hayashi-Takanaka, Y.; Oda, T.; Sato, M.; Park, S. Y.; Kimura, H.; Kurumizaka, H., Crystal structure of the human centromeric nucleosome containing CENP-A. *Nature* **2011,** *476* (7359), 232-5.

17. Strickland, M.; Strickland, W. N.; Brandt, W. F.; Von Holt, C.; Wittmann-Liebold, B., The complete amino-acid sequence of histone H2B(3) from sperm of the sea urchin Parechinus angulosus. *Eur J Biochem* **1978,** *89* (2), 443-52.

18. Draizen, E. J.; Shaytan, A. K.; Marino-Ramirez, L.; Talbert, P. B.; Landsman, D.; Panchenko, A. R., HistoneDB 2.0: a histone database with variants--an integrated resource to explore histones and their variants. *Database (Oxford)* **2016,** *2016*.

19. Ismail, I. H.; Hendzel, M. J., The gamma-H2A.X: is it just a surrogate marker of double-strand breaks or much more? *Environ Mol Mutagen* **2008,** *49* (1), 73-82.

20. Giaimo, B. D.; Ferrante, F.; Herchenrother, A.; Hake, S. B.; Borggrefe, T., The histone variant H2A.Z in gene regulation. *Epigenetics Chromatin* **2019,** *12* (1), 37.

21. Suto, R. K.; Clarkson, M. J.; Tremethick, D. J.; Luger, K., Crystal structure of a nucleosome core particle containing the variant histone H2A.Z. *Nat Struct Biol* **2000,** *7* (12), 1121-4.

22. Chadwick, B. P.; Willard, H. F., A novel chromatin protein, distantly related to histone H2A, is largely excluded from the inactive X chromosome. *J Cell Biol* **2001,** *152* (2), 375-84.

23. Bao, Y.; Konesky, K.; Park, Y. J.; Rosu, S.; Dyer, P. N.; Rangasamy, D.; Tremethick, D. J.; Laybourn, P. J.; Luger, K., Nucleosomes containing the histone variant H2A.Bbd organize only 118 base pairs of DNA. *EMBO J* **2004,** *23* (16), 3314-24.

24. Soboleva, T. A.; Nekrasov, M.; Pahwa, A.; Williams, R.; Huttley, G. A.; Tremethick, D. J., A unique H2A histone variant occupies the transcriptional start site of active genes. *Nat Struct Mol Biol* **2011,** *19* (1), 25-30.

25. Sierra, I.; Anguera, M. C., Enjoy the silence: X-chromosome inactivation diversity in somatic cells. *Curr Opin Genet Dev* **2019,** *55*, 26-31.

26. Loda, A.; Heard, E., Xist RNA in action: Past, present, and future. *PLoS Genet* **2019,** *15* (9), e1008333.

27. Costanzi, C.; Pehrson, J. R., Histone macroH2A1 is concentrated in the inactive X chromosome of female mammals. *Nature* **1998,** *393* (6685), 599-601.

28. Chakravarthy, S.; Gundimella, S. K.; Caron, C.; Perche, P. Y.; Pehrson, J. R.; Khochbin, S.; Luger, K., Structural characterization of the histone variant macroH2A. *Mol Cell Biol* **2005,** *25* (17), 7616-24.

29. Naylor, J. A.; Buck, D.; Green, P.; Williamson, H.; Bentley, D.; Giannelli, F., Investigation of the factor VIII intron 22 repeated region (int22h) and the associated inversion junctions. *Hum Mol Genet* **1995,** *4* (7), 1217-24.

30. Horn, V.; Uckelmann, M.; Zhang, H.; Eerland, J.; Aarsman, I.; le Paige, U. B.; Davidovich, C.; Sixma, T. K.; van Ingen, H., Structural basis of specific H2A K13/K15 ubiquitination by RNF168. *Nat Commun* **2019,** *10* (1), 1751.

31. Mattiroli, F.; Vissers, J. H.; van Dijk, W. J.; Ikpa, P.; Citterio, E.; Vermeulen, W.; Marteijn, J. A.; Sixma, T. K., RNF168 ubiquitinates K13-15 on H2A/H2AX to drive DNA damage signaling. *Cell* **2012,** *150* (6), 1182-95.

32. Poplawski, A.; Hu, K.; Lee, W.; Natesan, S.; Peng, D.; Carlson, S.; Shi, X.; Balaz, S.; Markley, J. L.; Glass, K. C., Molecular insights into the recognition of N-terminal histone modifications by the BRPF1 bromodomain. *J Mol Biol* **2014,** *426* (8), 1661-76.

33. Zhou, W.; Zhu, P.; Wang, J.; Pascual, G.; Ohgi, K. A.; Lozach, J.; Glass, C. K.; Rosenfeld, M. G., Histone H2A monoubiquitination represses transcription by inhibiting RNA polymerase II transcriptional elongation. *Mol Cell* **2008,** *29* (1), 69-80.

34. Zhou, J.; Fan, J. Y.; Rangasamy, D.; Tremethick, D. J., The nucleosome surface regulates chromatin compaction and couples it with transcriptional repression. *Nat Struct Mol Biol* **2007,** *14* (11), 1070-6.

35. Mucke, K.; Paulus, C.; Bernhardt, K.; Gerrer, K.; Schon, K.; Fink, A.; Sauer, E. M.; Asbach-Nitzsche, A.; Harwardt, T.; Kieninger, B.; Kremer, W.; Kalbitzer, H. R.; Nevels, M., Human cytomegalovirus major immediate early 1 protein targets host chromosomes by docking to the acidic pocket on the nucleosome surface. *J Virol* **2014,** *88* (2), 1228-48.

36. Makde, R. D.; England, J. R.; Yennawar, H. P.; Tan, S., Structure of RCC1 chromatin factor bound to the nucleosome core particle. *Nature* **2010,** *467* (7315), 562-6.

37. Armache, K. J.; Garlick, J. D.; Canzio, D.; Narlikar, G. J.; Kingston, R. E., Structural basis of silencing: Sir3 BAH domain in complex with a nucleosome at 3.0 A resolution. *Science* **2011,** *334* (6058), 977-82.

38. Kato, H.; van Ingen, H.; Zhou, B. R.; Feng, H.; Bustin, M.; Kay, L. E.; Bai, Y., Architecture of the high mobility group nucleosomal protein 2-nucleosome complex as revealed by methyl-based NMR. *Proc Natl Acad Sci U S A* **2011,** *108* (30), 12283-8.

39. Kalashnikova, A. A.; Porter-Goff, M. E.; Muthurajan, U. M.; Luger, K.; Hansen, J. C., The role of the nucleosome acidic patch in modulating higher order chromatin structure. *J R Soc Interface* **2013,** *10* (82), 20121022.

40. Eirin-Lopez, J. M.; Ishibashi, T.; Ausio, J., H2A.Bbd: a quickly evolving hypervariable mammalian histone that destabilizes nucleosomes in an acetylation-independent way. *FASEB J* **2008,** *22* (1), 316-26.

41. Strickler, S. S.; Gribenko, A. V.; Gribenko, A. V.; Keiffer, T. R.; Tomlinson, J.; Reihle, T.; Loladze, V. V.; Makhatadze, G. I., Protein stability and surface electrostatics: a charged relationship. *Biochemistry* **2006,** *45* (9), 2761-6.

42. Doyen, C. M.; Montel, F.; Gautier, T.; Menoni, H.; Claudet, C.; Delacour-Larose, M.; Angelov, D.; Hamiche, A.; Bednar, J.; Faivre-Moskalenko, C.; Bouvet, P.; Dimitrov, S., Dissection of the unusual structural and functional properties of the variant H2A.Bbd nucleosome. *EMBO J* **2006,** *25* (18), 4234-44.

43. Montel, F.; Fontaine, E.; St-Jean, P.; Castelnovo, M.; Faivre-Moskalenko, C., Atomic force microscopy imaging of SWI/SNF action: mapping the nucleosome remodeling and sliding. *Biophys J* **2007,** *93* (2), 566-78.

44. Angelov, D.; Verdel, A.; An, W.; Bondarenko, V.; Hans, F.; Doyen, C. M.; Studitsky, V. M.; Hamiche, A.; Roeder, R. G.; Bouvet, P.; Dimitrov, S., SWI/SNF remodeling and p300-dependent transcription of histone variant H2ABbd nucleosomal arrays. *EMBO J* **2004,** *23* (19), 3815-24.

45. Arimura, Y.; Kimura, H.; Oda, T.; Sato, K.; Osakabe, A.; Tachiwana, H.; Sato, Y.; Kinugasa, Y.; Ikura, T.; Sugiyama, M.; Sato, M.; Kurumizaka, H., Structural basis of a nucleosome containing histone H2A.B/H2A.Bbd that transiently associates with reorganized chromatin. *Sci Rep* **2013,** *3*, 3510.

46. Sugiyama, M.; Arimura, Y.; Shirayama, K.; Fujita, R.; Oba, Y.; Sato, N.; Inoue, R.; Oda, T.; Sato, M.; Heenan, R. K.; Kurumizaka, H., Distinct features of the histone core structure in nucleosomes containing the histone H2A.B variant. *Biophys J* **2014,** *106* (10), 2206-13.

47. Shukla, M. S.; Syed, S. H.; Goutte-Gattat, D.; Richard, J. L.; Montel, F.; Hamiche, A.; Travers, A.; Faivre-Moskalenko, C.; Bednar, J.; Hayes, J. J.; Angelov, D.; Dimitrov, S., The docking domain of histone H2A is required for H1 binding and RSC-mediated nucleosome remodeling. *Nucleic Acids Res* **2011,** *39* (7), 2559-70.

48. Dorigo, B.; Schalch, T.; Kulangara, A.; Duda, S.; Schroeder, R. R.; Richmond, T. J., Nucleosome arrays reveal the two-start organization of the chromatin fiber. *Science* **2004,** *306* (5701), 1571-3.

49. Montel, F.; Menoni, H.; Castelnovo, M.; Bednar, J.; Dimitrov, S.; Angelov, D.; Faivre-Moskalenko, C., The dynamics of individual nucleosomes controls the chromatin condensation pathway: direct atomic force microscopy visualization of variant chromatin. *Biophys J* **2009,** *97* (2), 544-53.

50. Fan, J. Y.; Rangasamy, D.; Luger, K.; Tremethick, D. J., H2A.Z alters the nucleosome surface to promote HP1alpha-mediated chromatin fiber folding. *Mol Cell* **2004,** *16* (4), 655-61.

51. Tolstorukov, M. Y.; Goldman, J. A.; Gilbert, C.; Ogryzko, V.; Kingston, R. E.; Park, P. J., Histone variant H2A.Bbd is associated with active transcription and mRNA processing in human cells. *Mol Cell* **2012,** *47* (4), 596-607.

52. Ioudinkova, E. S.; Barat, A.; Pichugin, A.; Markova, E.; Sklyar, I.; Pirozhkova, I.; Robin, C.; Lipinski, M.; Ogryzko, V.; Vassetzky, Y. S.; Razin, S. V., Distinct distribution of ectopically expressed histone variants H2A.Bbd and MacroH2A in open and closed chromatin domains. *PLoS One* **2012,** *7* (10), e47157.

53. Okuwaki, M.; Kato, K.; Shimahara, H.; Tate, S.; Nagata, K., Assembly and disassembly of nucleosome core particles containing histone variants by human nucleosome assembly protein I. *Mol Cell Biol* **2005,** *25* (23), 10639-51.

54. Gautier, T.; Abbott, D. W.; Molla, A.; Verdel, A.; Ausio, J.; Dimitrov, S., Histone variant H2ABbd confers lower stability to the nucleosome. *EMBO Rep* **2004,** *5* (7), 715-20.

55. Chen, Y.; Chen, Q.; McEachin, R. C.; Cavalcoli, J. D.; Yu, X., H2A.B facilitates transcription elongation at methylated CpG loci. *Genome Res* **2014,** *24* (4), 570-9.

56. Soboleva, T. A.; Parker, B. J.; Nekrasov, M.; Hart-Smith, G.; Tay, Y. J.; Tng, W. Q.; Wilkins, M.; Ryan, D.; Tremethick, D. J., A new link between transcriptional initiation and pre-mRNA splicing: The RNA binding histone variant H2A.B. *PLoS Genet* **2017,** *13* (2), e1006633.

57. Anuar, N. D.; Kurscheid, S.; Field, M.; Zhang, L.; Rebar, E.; Gregory, P.; Buchou, T.; Bowles, J.; Koopman, P.; Tremethick, D. J.; Soboleva, T. A., Gene editing of the multi-copy H2A.B gene and its importance for fertility. *Genome Biol* **2019,** *20* (1), 23.

58. Wu, B. J.; Dong, F. L.; Ma, X. S.; Wang, X. G.; Lin, F.; Liu, H. L., Localization and expression of histone H2A variants during mouse oogenesis and preimplantation embryo development. *Genet Mol Res* **2014,** *13* (3), 5929-39.

59. Sansoni, V.; Casas-Delucchi, C. S.; Rajan, M.; Schmidt, A.; Bonisch, C.; Thomae, A. W.; Staege, M. S.; Hake, S. B.; Cardoso, M. C.; Imhof, A., The histone variant H2A.Bbd is enriched at sites of DNA synthesis. *Nucleic Acids Res* **2014,** *42* (10), 6405-20.

60. Molaro, A.; Young, J. M.; Malik, H. S., Evolutionary origins and diversification of testis-specific short histone H2A variants in mammals. *Genome Res* **2018,** *28* (4), 460-473.

61. Ferguson, L.; Ellis, P. J.; Affara, N. A., Two novel mouse genes mapped to chromosome Yp are expressed specifically in spermatids. *Mamm Genome* **2009,** *20* (4), 193-206.

62. Syed, S. H.; Boulard, M.; Shukla, M. S.; Gautier, T.; Travers, A.; Bednar, J.; Faivre-Moskalenko, C.; Dimitrov, S.; Angelov, D., The incorporation of the novel histone variant H2AL2 confers unusual structural and functional properties of the nucleosome. *Nucleic Acids Res* **2009,** *37* (14), 4684-95.

63. Dai, L.; Xie, X.; Zhou, Z., Crystal structure of the histone heterodimer containing histone variant H2A.Bbd. *Biochem Biophys Res Commun* **2018,** *503* (3), 1786-1791.

64. Kerfah, R.; Plevin, M. J.; Sounier, R.; Gans, P.; Boisbouvier, J., Methyl-specific isotopic labeling: a molecular tool box for solution NMR studies of large proteins. *Curr Opin Struct Biol* **2015,** *32*, 113-22.