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

Zhang, H.

Chapter 4. Structure and dynamics of the H2A.B variant nucleosome

This chapter is based on:
Heyi Zhang, Yiran Lin, Jan Huertas, Vlad Cojocaru, Vincenzo Lobbia, Ulric B. le Paige, Hugo van Ingen. Structure and dynamics of the H2A.B variant nucleosome. *In preparation.*

Contributions of authors:
MNase digestion assays were performed with Yiran Lin; NMR study of H2A.B N-terminal tail in the nucleosome was performed with Vincenzo Lobbia; Energy minimization of H2A.B nucleosome models was performed by Jan Huertas and Vlad Cojocaru; 601 DNA used in nucleosome reconstitutions was provided by Ulric B. le Paige.

Abstract

Incorporation of histone variants into nucleosomes confers distinct functional properties to nucleosomes by their unique amino acid sequences, contributing to the regulation of chromatin functions such as transcription and DNA repair. Histone variant H2A.B is a highly divergent H2A variant and is strongly linked to active transcription and RNA splicing. H2A.B nucleosomes are known to have an 'open' conformation in which the nucleosomal DNA is partially released from the histone octamer core at the entry/ exit ends. Yet, the exact molecular reason for such opened structure is unknown, nor is there high-resolution structural and dynamical data of H2A.B nucleosomes available. Here, we report an NMR-driven investigation of the structure and dynamics of H2A.B nucleosomes, showing that the H2A.B truncated docking domain is folded within the variant nucleosome as in the canonical conformation and forms a stably folded interface with the H3-H4 subunit. Our NMR data further indicate that the H3 N-terminal tail in H2A.B nucleosomes has increased flexibility and is overall less bound to DNA compared to canonical nucleosomes. Finally, we identify a register shift of a DNA minor groove anchoring arginine in H2A.B and show using micrococcal nuclease digestion assays that this shift contributes to destabilization and DNA opening of the nucleosome. Together, these results provide new insights into the molecular properties of the H2A.B nucleosome.

Introduction

Histone proteins are among the most highly conserved proteins in nature, reflecting their cardinal importance in packaging and protecting the genetic material of eukaryotes in the form of chromatin. The bulk of chromatin is made of four canonical histones: H2A, H2B, H3 and H4, and together they form a histone octamer around which ~150 base pairs of DNA is wrapped, creating the nucleosome. To regulate chromatin biology, nucleosomes are the binding target of many nuclear proteins and are actively altered in various ways, one of which is by a special class of histone proteins, histone variants, to confer specific structural and functional properties to chromatin at particular locations in the genome.

Histone variants have been discovered for all four core histone types [1-2] and can substitute their canonical forms in the nucleosome, typically through the action of specific histone chaperones and remodeller proteins [3]. Through their distinct amino acid sequences, incorporation of histone variants bestows special functional and conformational properties to nucleosome. For example, histone H3 variant CENP-A is crucial in assembling the centrosome as it facilitates binding of centromeric DNA-binding proteins [4] and histone H2A variant H2A.X plays a major role in signalling DNA damage through phosphorylation of a unique serine in its C-terminus [5].

One of the most recently discovered and the most divergent histone variants is H2A variant H2A.B or H2A.Bbd (Barr body deficient) (Figure 4.1a). Originally identified as a variant that is excluded from the inactive X chromosome [6], it is now tightly associated to active transcription and mRNA splicing [7-12]. These functional roles may ultimately result from the pronounced promotion of open chromatin structure by H2A.B. When incorporated into nucleosomal arrays, H2A.B was shown to prevent formation of a compact chromatin fibre [13]. At the level of isolated nucleosomes, a series of biochemical and low-resolution structural studies have shown that, compared to canonical nucleosomes, nucleosomes containing H2A.B have a significantly shorter stretch of DNA tightly bound to the histone octamer (with the last turn of DNA dissociated from the octamer) and that the DNA ends are at a much wider angle relative to each other (Figure 4.1b,c) [14-16].

The sequence differences between canonical H2A and the variant H2A.B must underlie the molecular basis of its open nucleosomal conformation, yet up to date an exact description of this relationship is lacking. Since the last turn of nucleosomal DNA is bound to H3 in the canonical nucleosome, there is not an *a priori* clear-cut answer to this question. At only 48% sequence identity to canonical H2A, H2A.B's distinct features are the elongated and arginine rich N-terminal tail, a truncated C-terminal docking domain, the absence of acidic patch residues, and an overall less basic nature. Previous studies have pointed to the truncated C-terminal docking domain as main cause of the open nucleosome conformation [14-15, 17]. This region supports multiple interactions with H3 and H4 that stabilizes the histone octamer core and is truncated by ~ 19 residues in H2A.B resulting in a loss of several of these stabilizing interactions [18]. A series of experiments including MNase digestion and DNase footprinting assays, as well as FRET measurement of reconstituted nucleosomes using either H2A docking domain mutants or H2A/H2A.B chimera proteins resulted in the conclusion that H2A.B specific sequence differences from the α3 helix to the C-terminus are causing the open structure of the nucleosome either directly or indirectly [14-15, 17]. In addition, recent structures of the H2A.B-H2B heterodimers have pointed to the reduced positive electrostatic potential on the DNA binding surface which may promote DNA unwrapping [19] (see Chapter 3).

Here, we aimed to characterize the structural and dynamical properties of H2A.B nucleosomes at atomic resolution, combining solution-state NMR, biochemical assays and molecular dynamics (MD) simulations. We show that the elongated H2A.B N-terminal tail is highly dynamic whereas on the C-terminal side only the ultimate residue is highly mobile. Our NMR data indicate that the H2A.B truncated docking domain is folded and interfaced to H3-H4 as in canonical nucleosomes. NMR relaxation data support an opened nucleosome structure in which the DNA is at least transiently released from the octamer surface. As a result, the H3 N-terminal tail is more dynamic and overall less DNA-bound than in canonical nucleosomes, which is supported by a preliminary analysis of ongoing MD simulations. Finally, we identified a register shift of a key DNA binding arginine residue in H2A.B that contributes directly to unwrapping of the nucleosomal DNA. Together,

this study offers new insights into the unusual properties of the H2A.B nucleosome.

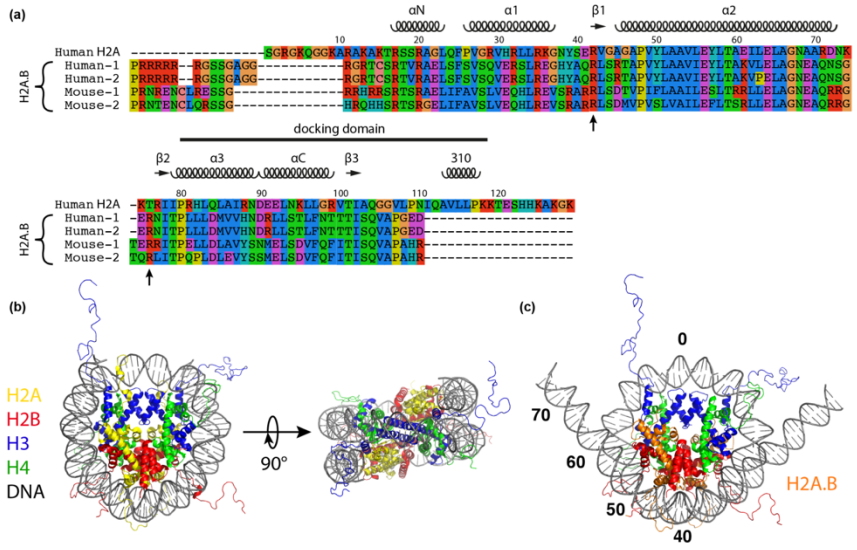

Figure 4.1. Histone variant H2A.B. (a) Sequence alignment of the most studied human and mouse H2A.B sequences with the human canonical H2A. Secondary structural elements of canonical H2A in the context of nucleosome are indicated above the sequences. The definition of docking domain is marked as a black bar. Positions of anchoring arginine's in canonical nucleosome are shown with arrows below the sequences. **(b)** Overall structure of the canonical nucleosome (PDB id: 1KX5). Colour coding is indicated. **(c)** A simplistic model for the H2A.B nucleosome with an arbitrary conformation for the terminal 15 bp of nucleosomal DNA. The dyad DNA base pair is marked as position 0 and the terminal three turns of the DNA are marked at positions 40-70.

Results

H2A.B N-terminal tail is highly flexible and transiently DNA-bound

To investigate the structural and dynamical properties of the elongated and arginine rich H2A.B N-terminal tail, we reconstituted H2A.B nucleosome using uniformly ^{15}N labelled H2A.B and unlabelled H2B, H3, H4, and DNA. This isotope-labelling pattern in combination with the large mass of nucleosome (> 200 kDa) permits the selective observation of the highly flexible parts of H2A.B. Thirteen peaks were observed in the 2D NH TROSY fingerprint spectrum (Figure 4.2a).

Comparison to the H2A.B assignments in the H2A.B-H2B heterodimer (see Chapter 3) allowed transfer of seven assignments for G11-G14, R17, T18 and D114, together with the blob assignment of the six consecutive arginine's (R2-R7). The two remaining peaks did not overlap with any peak in the dimer spectrum and could only be assigned tentatively to S9 and/or S10, and C19. Apart from the C-terminal residue D114, the observed resonances together constitute the majority of the twenty N-terminal residues, corresponding to the expected size of the N-terminal H2A.B tail and demonstrating the high flexibility of this region. No resonances for residues 21-25 could be observed suggesting the αN-helix is formed as in the canonical nucleosome and the H2A.B-H2B crystal structure [18-19], while it is highly flexible and disordered within the variant dimer free in solution (see Chapter 3).

Notably, resonances for R15 and G16 are absent and resonances for G14 and R17 have very low signal intensity, indicating the limited mobility for the GRGR box in the nucleosome which is possibly caused by its attachment to the nucleosomal DNA. Further evidence for transient DNA binding comes from the small chemical shift changes of the tail resonances in the nucleosome when compared to the dimer. Interestingly, these changes are in the downfield direction whereas upfield changes were observed for canonical H2A in a mononucleosome sediment [20], signifying a change in DNA interaction mode. In addition, in both sediment and solution, peak doubling was observed for several resonances in the canonical H2A tail, reflecting the interaction to the different sites of the non-palindromic 601-DNA [20-21]. For H2A.B such peak doubling is not apparent suggesting that at least for the observed resonances DNA binding is less intimate.

Assignment of H2A.B docking domain methyl TROSY signals

We next examined the conformation and dynamics of the H2A.B C-terminal docking domain within the nucleosome. Starting from the end of H2A.B αC helix, this region comprises residues F101 to D114. This region is unfolded and highly dynamic within the context of the dimer (see Chapter 3), yet the corresponding part in canonical H2A forms a defined interface to H3-H4 tetramer in nucleosome, including an intermolecular β-sheet (see Figure 4.1). We thus wanted to probe the conformation of this region in the variant nucleosome. The tail-focused

NMR experiment described above showed only one observable resonance for this region, the C-terminus residue D114 (Figure 4.2a). This suggests that residues F101-E113 have strongly reduced mobility within the context of the nucleosome and are thus most likely folded in a defined conformation.

To further examine the conformation of the docking domain, we used the methyl-TROSY approach [22] to allow site-specific investigation of the histone octamer core [23]. H2A.B nucleosomes were prepared with Ile-δ1-[$^{13}CH_3$], Leu, Val-[$^{13}CH_3$, $^{12}CD_3$]-H2A.B and perdeuterated H2B, H3, and H4 with unlabelled 601-DNA. The scarcity of protons in such sample dramatically slows the transverse relaxation of the methyl ^{1}H spins, thereby allowing to probe the Ile, Val, Leu methyl groups in such large system as the nucleosome. The methyl-TROSY spectrum of ILV-labelled H2A.B within the nucleosome, displayed in Figure 4.2b, shows overall decent chemical shift dispersion. Out of the 47 expected resonances for the 3 Ile, 13 Leu and 9 Val residues in H2A.B, 45 peaks could be identified, including three signals in the Ile-region. The signal intensities appear rather heterogenous, in particular for the Leu/Val region where several resonances have very low intensities.

Assignment of methyl-TROSY spectra is typically based on a combination of mutagenesis and NOESY spectra, by comparing the experimental connectivities in the proton network to that expected from the structure. In the case of H2A.B nucleosomes, NOESY experiments were of low quality, showing only few NOEs, thereby preventing complete assignment of the methyl signals. Instead, we focussed our analysis on the three isoleucine's in H2A.B, which give rise to well-dispersed and intense cross peaks in the nucleosome spectra. Since two of the isoleucine's are located at approximately opposite ends of the histone fold core (I59 and I82) and one (I106) is part of the docking domain, assignment of these three signals offers the opportunity to probe the structural and dynamical properties of H2A.B in the nucleosome at key locations.

In a straightforward approach, we compared the H2A.B nucleosome spectrum to that of the canonical nucleosome with known H2A assignments [23] (see Figure S4.3). Two isoleucines in H2A.B are conserved in canonical H2A; H2A.B I82 corresponds to H2A I77 and H2A.B I106 to H2A I101. The inset in Figure 4.2b shows that each of

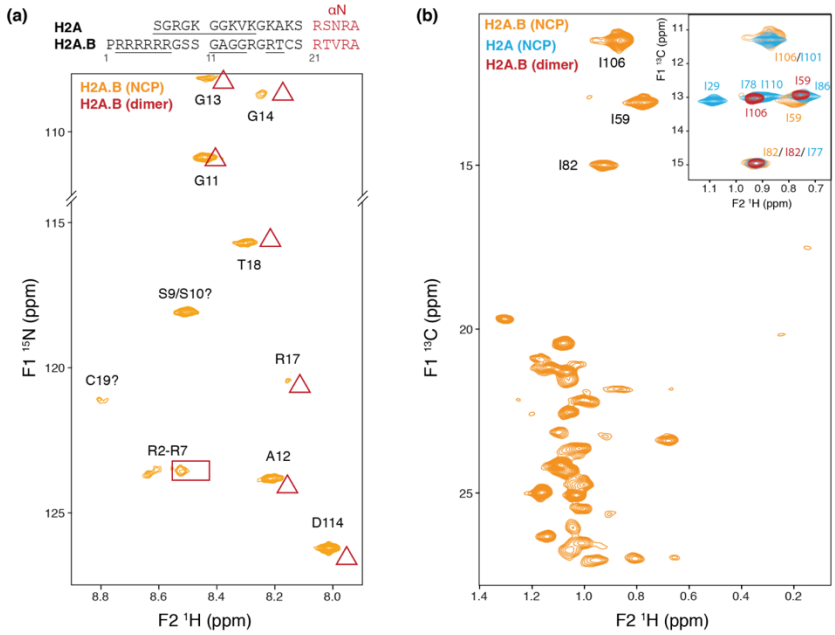

Figure 4.2. Amide and methyl-TROSY NMR spectra of H2A.B in the nucleosome. (a) Amide HN-TROSY spectrum of uniformly ^{15}N-labeled H2A.B in otherwise protonated nucleosomes. Assignments for tail residues in the H2A.B-H2B heterodimer are indicated with triangles for well resolved individual assignments and as a rectangle for overlapped group assignments. Tentative assignments indicated with ? were based on chemical shift correspondence to the dimer (S10) or based on chemical shift statistics of tail residues (S9 and C19, unassigned in the dimer). Amino acid sequences of H2A.B and H2A N-terminal tails are indicated on top, aligned with respect to the αN helix. Residues visible in canonical H2A nucleosomes are underlined. **(b)** Methyl-TROSY spectrum of ILV-labelled H2A.B with otherwise perdeuterated core histones and protonated DNA in the nucleosome. The inset shows the overlay of methyl-TROSY spectra Ile region of H2A.B in the nucleosome (orange) and dimer (red) and canonical H2A in nucleosome (cyan). Spectra were processed with the same parameters and overlaid by matching I82 of H2A.B to the corresponding I77 of H2A.

the H2A I77 and I101 overlaps with a H2A.B resonance, thus allowing to transfer their assignments to the corresponding residues in H2A.B. By exclusion, the remaining H2A.B isoleucine can be assigned to I59. This assignment is further supported by comparison to methyl-TROSY spectra of ILV-labelled H2A.B in H2A.B-H2B heterodimers (see Figure S4.3). There are two (nearly) overlapping signals, which match the histone fold core resonances for which little changes in chemical

environment are expected upon formation of the nucleosome. One resonance has very different chemical shifts in the two systems, corresponding to the docking domain isoleucine I106 (see inset Figure 4.2b).

Conserved docking domain structure of H2A.B in the variant nucleosome

The chemical shift difference of I106 between H2A.B in dimers and in nucleosomes suggests that the docking domain of H2A.B becomes folded upon incorporation of the variant dimer into the nucleosome. The close chemical shift correspondence to H2A I101 in canonical nucleosomes suggests the variant docking domain obtains the same fold as the canonical domain. To further investigate the conformation of the docking domain, we analysed the NOESY spectrum of H2A.B in the nucleosome. Among the few NOE cross peaks, a clear NOE was observed between I106 and a leucine or valine methyl group (Figure 4.3a). While the leucine and valine methyl signals are as of yet unassigned, the ^{13}C chemical shift (14.7 ppm) indicates the resonance belongs most likely to a valine. In addition, the distinct up-field ^{1}H chemical shift (0.138 ppm) indicates proximity of the corresponding methyl group to an aromatic ring. We next checked whether this NOE is consistent with folding of the H2A.B docking domain as in the canonical nucleosome structure. Using canonical H2A docking domain as a template for homology modelling, we find four methyl groups, from residues L87, V91, V109, that are within 6 Å ^{13}C-^{13}C distance from the I106 δ1-methyl group. Among these, the closest methyl group is from V91 at 4.1 Å, which is the only methyl within 5 Å, and the only one that is also close to an aromatic ring (see Figure 4.3b). We thus assigned the unknown resonances to V91, which is located on α3 helix of H2A.B. This I106-V91 NOE cross peak indicates that the docking domain of H2A.B is folded in a stable conformation and suggests that the H2A.B interface to H3-H4 is formed in the canonical manner, including the β-sheet with H4.

H2A.B docking domain dynamics

Having established evidence for stable folding of the docking domain within the context of the H2A.B variant nucleosome, we next turned to characterize the H2A.B dynamics within the nucleosome. Given the

partially unwrapped nature of the H2A.B nucleosome, one could expect that the DNA unwrapping also results in a destabilization of the histone octamer and thus an increase in histone dynamics. Indeed, the prevalence of low intensity peaks in the methyl-TROSY spectrum could result from line broadening due to conformational exchange on a micro- to milli-second time scale. Alternatively, the line broadening could reflect the density of the local proton network which will have a larger impact on the line width upon increasing overall rotational tumbling time.

Comparison to the peak intensity distributions of H2A.B methyl resonances in either nucleosomes or H2A.B-H2B dimers and of H2A resonances in canonical nucleosomes indicate that the preponderance of broadened signals is specific to H2A.B in the nucleosomal context (Figure 4.4a). To quantify this more rigorously, we measured both ^{1}H and ^{13}C-^{1}H multiple quantum transverse relaxation rates (R_2) of H2A.B isoleucine methyl groups in the variant nucleosome and compared these to relaxation rates of H2B core isoleucine methyls in the canonical nucleosome (unpublished data from Hugo van Ingen, Hidenori Kato, Yawen Bai, Lewis Kay). Histone H2B contains 9 isoleucines, of which 8 are part of the H2A-H2B histone fold core. As such, we expect the H2B isoleucine methyls to serve as equally valid reporters of the dimer dynamics as H2A. As shown in Figure 4.4b, while methyls in the variant nucleosome have similar multiple quantum R_2 values as in the canonical nucleosome, they have much higher ^{1}H R_2 values than in the canonical nucleosome, on average 82 s^{-1} vs. 47 s^{-1}. The elevated ^{1}H R_2 values in the variant nucleosome explain the relatively poor sensitivity of the NOESY experiment. Strikingly, the average ^{1}H R_2 of H2A.B methyl groups in buffer without any added salt matches well with rates obtained for the canonical nucleosome upon addition of 75 mM NaCl, conditions that are known to promote inter-nucleosome interactions in particular at NMR concentrations [24-25]. The higher R_2 for the variant could thus reflect a larger effective particle size, for example due to DNA unwrapping or inter-nucleosome interactions. However, due to the lack of acidic patch of H2A.B, it is unlikely that inter-nucleosome interactions are happening in the sample. Nevertheless, since a similar trend for the multiple quantum R_2 values is not observed, line

broadening due to conformational dynamics affecting predominantly the ^{1}H chemical shift cannot be ruled out.

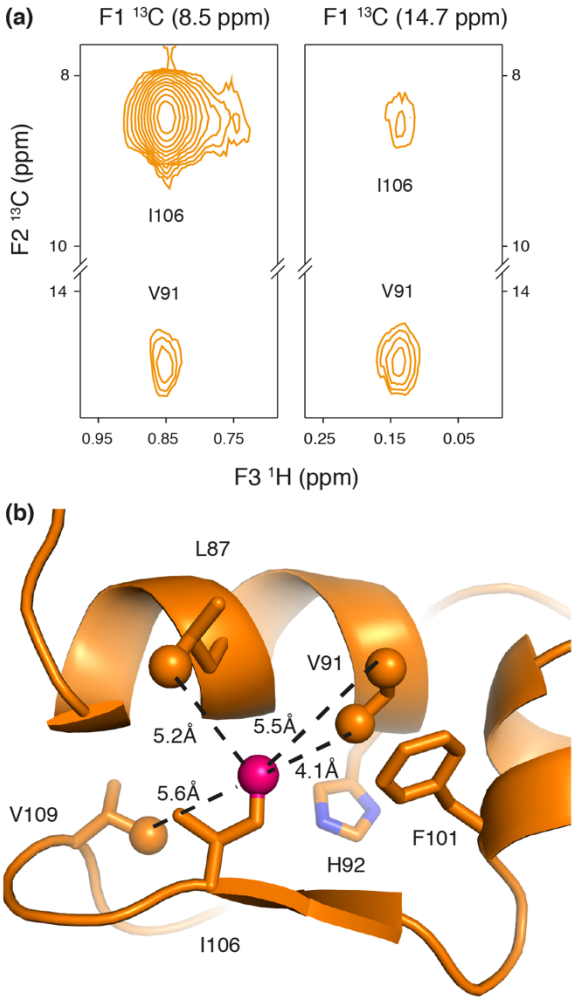

Figure 4.3. H2A.B docking domain structure. (a) Strips from the 3D CCH Me-NOESY showing NOE cross peaks between I106 to V91. **(b)** H2A.B structural model with docking domain modelled according to H2A structure in the canonical nucleosome. Methyl carbon of I106 is shown as pink sphere and H2A.B methyl carbons within 6 Å are shown as orange spheres. Residues and inter-methyl distances are marked. Side chains from aromatic residues (H92 and F101) are shown as sticks.

To investigate whether this specific broadening is due to slow local conformational changes, we recorded both ^{1}H single quantum and ^{13}C-^{1}H multiple quantum (MQ) CPMG relaxation dispersion experiments. No dispersion in ^{1}H R_2 rates could be observed up to a CPMG pulsing rate of 1000 Hz (Figure S4.4). While small dispersions due to conformational changes ($\Delta R_2 < 10\ s^{-1}$) cannot be excluded due to the limited sensitivity, these results speak against large conformational motions in the millisecond time scale regime. In the ^{13}C-^{1}H MQ experiment, similar flat dispersion profiles were observed for I59 and I82, indicating the lack of conformational changes for the two residues, whereas a slight dispersion of $R_{2,eff}$ value was observed for I106 (Figure 4.4c). Comparison against data recorded on H2A.B-H2B dimers shows that a similar slight dispersion is already present for I106 in the heterodimer (see inset Figure 4.4c). Importantly, comparison of the dimer and nucleosome data clearly show that the average $R_{2,eff}$ for I106 is similar to that for core isoleucines in the context of the nucleosome, while it was much lower within the dimer. This further supports stable folding of the H2A.B docking domain upon incorporation into the nucleosome.

Together, the NMR data on H2A.B nucleosomes point to a canonical-like histone octamer conformation. The H2A.B N-terminal tail is highly flexible as the H2A N-terminal tail, the truncated docking domain of H2A.B is folded as in H2A and there are no large-scale conformation motions present within the variant nucleosome. Yet, the elevated relaxation rates point to increased effective size of the variant nucleosome, likely as a result of partial DNA unwrapping.

Increased H3 tail dynamics in the H2A.B nucleosome

To further map the impact of H2A.B incorporation on histone octamer, we examined the H3 N-terminal tail since unwrapping of the nucleosomal DNA is likely to affect this tail foremost (see Figure 4.1). Using the same approach as for the H2A.B tail, we reconstituted the variant and canonical nucleosome using uniformly ^{15}N and ^{13}C labelled H3. Both amide NH-TROSY spectra showed well resolved peaks that were assigned to residues T3 to K36 based on previously published assignments[20] and standard triple resonance experiments (Figure 4.5a). Crucially, no additional peaks were observed for H3 in the H2A.B nucleosome, indicating that unwrapping of the DNA does not result in

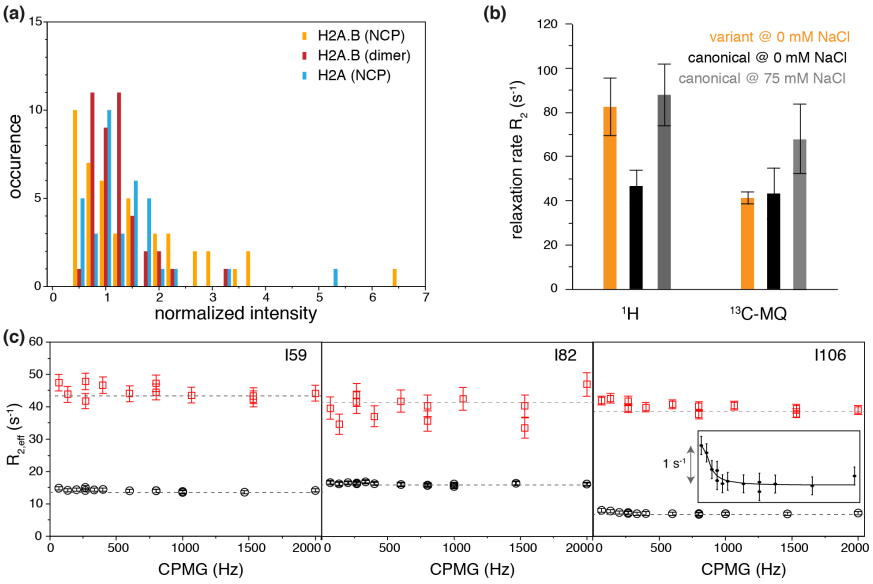

Figure 4.4. NMR relaxation data of H2A.B in NCP support stable folding of docking domain in an overall opened nucleosome conformation. (a) Histogram of normalized methyl-TROSY peak intensities for H2A.B in dimers and nucleosomes and H2A in nucleosomes. Peak intensities are normalized to the median intensity of each spectrum. **(b)** Average relaxation rates R_2 from ^{1}H (^{1}H) or ^{13}C-^{1}H multiple quantum (^{13}C-MQ) relaxation experiments for core isoleucine of H2A.B in variant nucleosome at no salt condition (orange bar), H2B in canonical nucleosome at no salt (black bar) or with salt (grey bar) conditions. **(c)** ^{13}C-^{1}H multiple quantum CPMG relaxation dispersion curves for H2A.B Ile residues in nucleosomes (red squares) and dimers (black circles). The inset highlights the small dispersion of $R_{2,eff}$ values obtained for I106 in the variant dimer.

extension of the tail. In particular, resonances for K38-P44, which pass in between the two DNA gyres in the canonical nucleosomes, are not observed for either of the nucleosomes. This indicates that this region remains immobile, likely attached to only a single DNA helix in the variant nucleosome.

More detailed inspection of the two H3 spectra reveals that the majority of resonances are shifted upfield in the canonical compared to the variant nucleosome and this effect is present throughout the tail (Figure 4.5a,b). An elegant study from the Selenko lab has shown that the H3 tail resonances generally shift upfield upon incorporation into the

nucleosome as a result of transient DNA binding [26]. The general trend of the peak displacements here indicates that on average the H3 tail is less bound to the nucleosomal DNA in the variant nucleosome. More detailed comparison of the H2A.B-nucleosome spectra to those of the free H3 peptide recorded by the Selenko lab [26] shows that several residues, including A29, A31 and V35 have very similar chemical shifts as in the free peptide, indicating there are likely unattached to DNA. This is in agreement with the simplistic model shown in Figure 4.1c, where the stretch of A29-V35 is far from the DNA ends and thus may not bind to it and represent more like the free peptide chemical environment. Other resonances, such as Q19-A21 and A25 have chemical shifts somewhere in between that of canonical nucleosomes and the free H3 peptide, suggesting a weakened DNA interaction. Yet a third group of residues, including K14 and A15, have chemical shifts distinct from either canonical nucleosome and free peptide. Together, the chemical shifts pattern observed for the H3 tail in the H2A.B nucleosome implies a distinctly altered DNA interaction that is overall less intimate. Notably, the glycine resonances, which were shown to be least involved in DNA binding, also have the smallest chemical shift differences between the two systems.

The reduced DNA binding of the H3 tail is further supported by a general decrease in local effective rotational correlation times (τ_c), derived from measured ^{15}N R_1 and R_2 relaxation rates [26], for the H3 tail in the variant nucleosome (Figure 4.5c). Overall, the local τ_c is reduced by ~1.5 ns for most of the H3 tail residues in the H2A.B nucleosome. Strikingly, the τ_c profile is highly similar between the two nucleosome types. This argues against full release of the H3 tail from the DNA, in which case a gradual decrease in τ_c would be expected from the nucleosome core towards the N-terminus. Rather these data speak to a maintained, but reduced DNA binding in the variant.

Register shift of a minor groove arginine anchor in the H2A.B nucleosome

With a canonical-like H2A.B conformation and no evidence for large conformational changes, the underlying molecular basis for the opened nucleosome structure remains unclear. In an attempt to identify a direct (contributing) factor, we carefully inspected the molecular model for the H2A.B variant nucleosome as shown in Figure 4.1. This revealed a

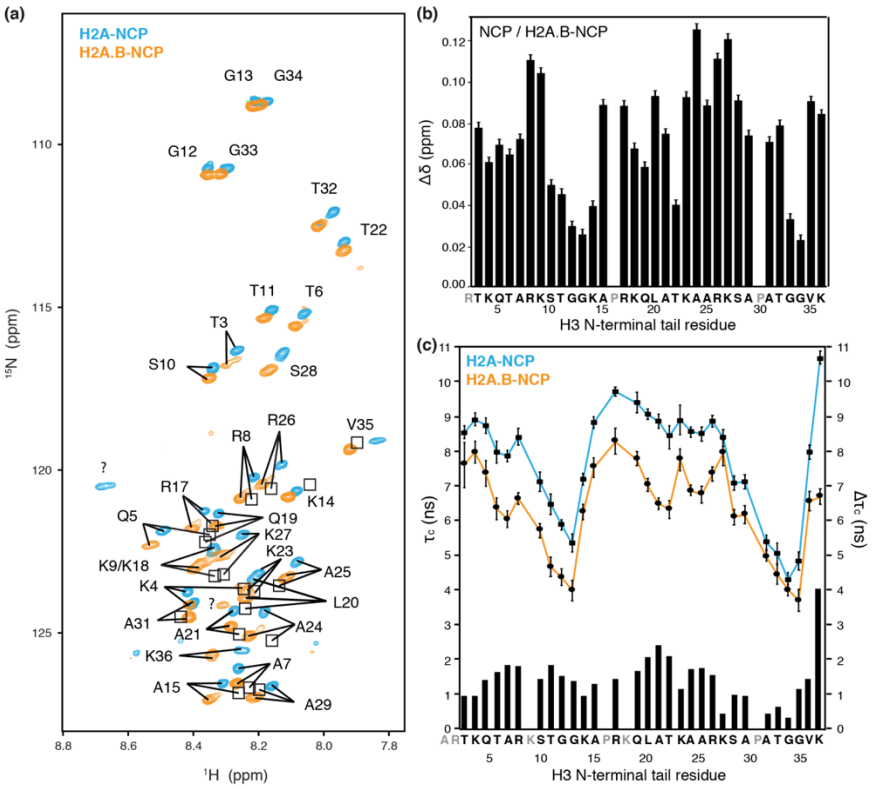

Figure 4.5. The H3 N-terminal tail has increased flexibility and reduced DNA-binding in Bbd-nucleosomes. (a) Overlay of TROSY spectra of nucleosomes containing ^{15}N-labelled H3 for both canonical (cyan) and H2A.B (orange) nucleosomes. Only highly flexible residues can be observed corresponding to the H3-tail. Peak positions for H3 tail peptide recorded by the Selenko lab [26] are marked as black boxes. **(b)** Chemical shift perturbation (Δδ) between canonical and variant nucleosomes for the H3 N-terminal residues. **(c)** Effective local rotational correlation times (τ_c) of H3 tail residues in canonical (squares/cyan) and variant (circles/orange) nucleosomes, together with differences in τ_c for each residue shown as bars ($\Delta\tau_c$).

register shift of an arginine that anchors to the DNA minor groove in the canonical nucleosome. This residue is in the L2 loop that connects the α2 and α3 helices. While in the canonical H2A sequence this is part of a sequence K74-T75-R76, the arginine is shifted by one position in the corresponding H2A.B sequence: E79-R80-N81. This register shift of the minor groove anchoring arginine causes a major clash between H2A.B R80 and the nucleosomal DNA backbone when in the

canonical, wrapped conformation (Figure 4.6a). Given the position of R80 in the L2 loop which is close to the DNA at super helix location (SHL) ±5.5, this clash could promote unwrapping of the nucleosomal DNA down to the next anchoring point at SHL ±4.5 [27], thus resulting in effective wrapping of 100-120 bp. Of note, the register shift of this arginine is conserved in H2A.B across most species (see Figure 4.1a and S1). To verify whether the arginine side chain conformation could adapt to the register shift and form a favourable interaction with the DNA nonetheless, two models for the H2A.B variant nucleosome were subjected to energy-minimization. While the H2A.B conformation is modelled based on the canonical H2A structure in the first model (Figure 4.6b), the other is based on the solution structure of the H2A.B-H2B histone fold core (Figure 4.6c), extended with its N- and C- tail region (including the docking domain) modelled based on canonical H2A. For both models and both copies of H2A.B in each model, R80 is not involved favourable interactions to the DNA after energy minimization (Figure 4.6b,c).

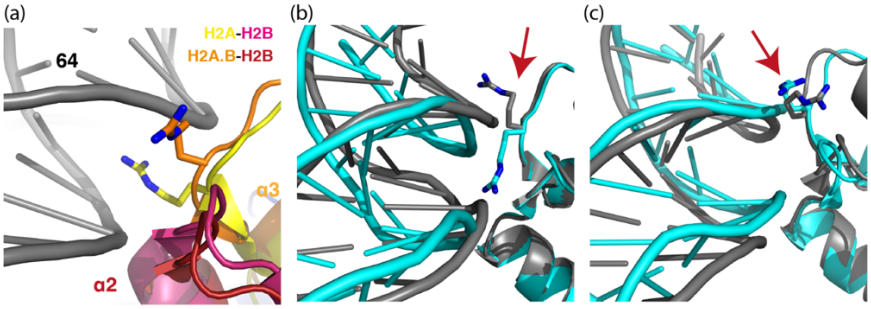

Figure 4.6. Register shift of the minor groove anchoring arginine in H2A.B. **(a)** Zoom on the α2-α3 loop of H2A.B superimposed with the canonical nucleosome structure highlight the clash of Arg80 with DNA. **(b, c)** Energy-minimized structural model zoomed on the α2-α3 loop showing an 'escaped' conformation of the Arg80. The two H2A.B α2-α3 loop regions in each model are superimposed, coloured cyan and grey respectively. The energy minimized model starting with canonical homology model is shown in (b), and starting with H2A.B-H2B dimer solution structure shown in (c).

H2A.B L2 loop destabilizes nucleosomal DNA wrapping

To test our hypothesis that the arginine register shift in the H2A.B L2 loop disrupts the local interaction between DNA and histone octamer, we compared electrophoretic mobilities and micrococcal nuclease

(MNase) DNA digestion patterns of nucleosomes reconstituted from an H2A mutant containing the arginine register shift to that of H2A.B nucleosomes. The H2A mutant (H2A_ERN) was constructed by replacing the L2 sequence KTR in H2A to the corresponding sequence ERN from H2A.B.

Nucleosomes containing H2A_ERN could successfully be reconstituted using both purified histone octamers and mixtures of H2A_ERN-H2B dimers and H3-H4 tetramers with DNA (Figure 4.7a). This indicates that the mutant can form stable octamers with other core histones as expected from the presence of the complete docking domain. Thanks to its opened conformation, reconstituted H2A.B nucleosomes show a reduced electrophoretic mobility and results in a wider band appearance than the canonical nucleosome, as was reported before [14]. The H2A_ERN nucleosome showed a markedly reduced mobility, similar to the variant, indicating the mutation resulted in a similar open conformation as the variant nucleosome. Nevertheless, the bands for the mutant nucleosomes are sharper than for the H2A.B nucleosome, indicating a more homogenous conformation for the mutant (Figure 4.7a).

We next subjected the three different types of nucleosomes to MNase treatment to assess their DNA accessibility. Samples taken at increasing lengths of MNase treatment show a progressively increased mobility and decreased intensity of the nucleosome band (Figure 4.7b). The increase in mobility reflects the trimming of the exposed DNA in the nucleosome particles, while the gradual disappearance of the band suggests disassembly of the nucleosome after excessive trimming of the nucleosomal DNA, possibly in combination with destabilization of the nucleosome during the electrophoresis and high salt exposure in the MNase quenching buffer. The lower part of most lanes shows the resulting bands from digested or undigested nucleosomal DNA fragments that are striped off from the octamer. The size of the digested DNA fragments indicates the number of base pairs that are protected in the nucleosome particle.

The canonical nucleosomes are the most resistant to MNase digestion, as shown by the modest loss of nucleosome band intensity (Figure 4.7b, lanes 1-6). As expected, these nucleosomes also protect the most of the nucleosomal DNA, as evidenced by the only gradual and modest increase in nucleosome mobility upon MNase treatment and the

gradual shift in size of the free DNA fragments from 167 to ~140 bp after 30 min treatment. Analysis of the DNA extracted after the protein digestion treatment shows fragments in the range of 130-150 bp (Figure 4.7c), which is typical for a properly folded, canonical nucleosome.

For H2A.B, MNase treatment resulted in rapid disappearance of the nucleosome band under our conditions (Figure 4.7b, lanes 13-18). Almost immediately after start of the DNA digestion, a free DNA band appeared at ~130 bp (estimated by comparison to free DNA band of canonical nucleosomes after 5 min MNase treatment), in line with the exposed nature of the DNA ends. Continued MNase digestion resulted to smaller fragments to around ~120 bp after 5 min of MNase digestion, as has been observed also in a previous study [15]. The DNA extracted at the end of the MNase treatment shows a broad size range, with most fragments around 80-120 bp.

For the H2A_ERN mutant nucleosomes, like for the variant, a free DNA band at ~130 bp appeared quickly after addition of the enzyme, indicating the DNA ends are readily accessible also in the mutant. After 10 min of MNase treatment the digested DNA band shifted towards ~110-120 bp, which was also the size of majority of DNA fragments extracted at the end of the treatment. Thus, these data demonstrate that the ERN substitution in the L2 loop increases DNA accessibility and decreases the protection of the nucleosomal DNA. While the digestion pattern of the mutant does not recapitulate the high sensitivity of the variant towards MNase, we conclude that the arginine register shift contributes to the opening of the DNA in H2A.B nucleosomes.

Discussion

Histone variant proteins can alter nucleosome functionality by their different amino acid sequences compared to canonical histones, ranging from substitution of few amino acids as in H3.3 to addition of a folded protein domain as in macroH2A. Resolving the impact of these sequence changes on the nucleosome is essential to understand the function of these variants in chromatin biology. The histone variant H2A.B present a particularly interesting case as one of the most divergent histone variants, and as a variant that affects nucleosome conformation in a profound way, causing partial unwrapping of the

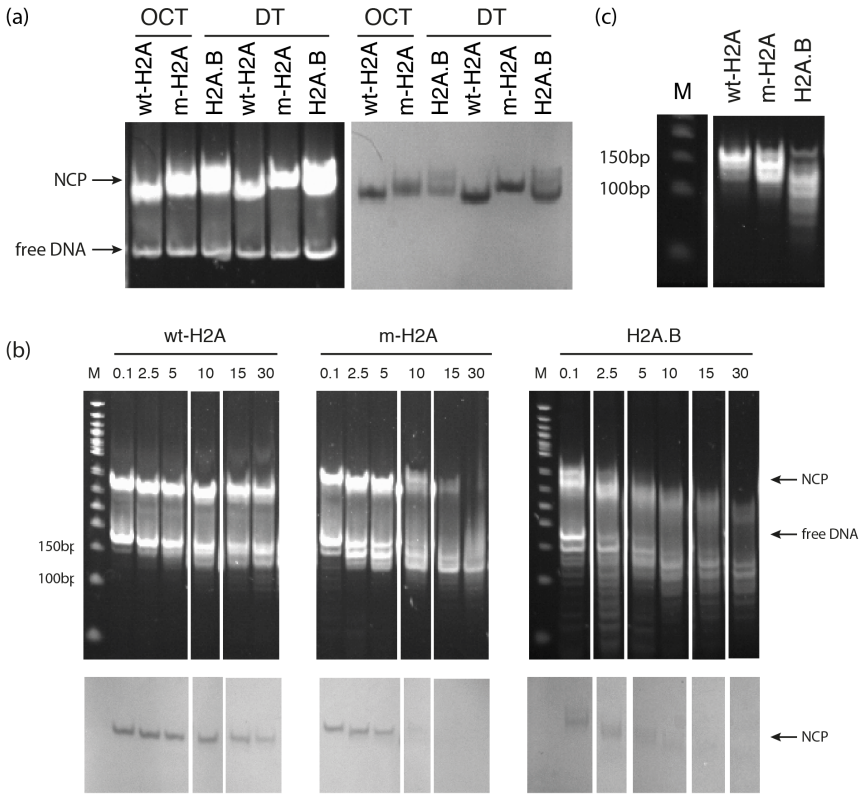

Figure 4.7. A register shift of a minor groove anchoring Arg causes destabilization of canonical nucleosomes. (a) Nucleosomes reconstituted using octamer and DNA (OCT) or dimer, tetramer and DNA (DT) were compared for canonical H2A (wt-H2A) and H2A_ERN (m-H2A) nucleosomes on the 5% native gel, stained for DNA using GelRed (left) and for proteins using Coomassie (right). (b) 5 µg of each H2A (wt-H2A), H2A_ERN (m-H2A) and H2A.B nucleosomes reconstituted by DT were digested with 0.6U MNase for the indicated amount of time (0.1, 2.5, 5, 10, 15, 30 min). MNase were deactivated before loading the sample on the 5% native gel, stained for DNA (GelRed, top) and protein (Coomassie, bottom). Lanes are reordered according to time of treatment for clarity. (c) Samples were deproteinized after 30 min and the extracted DNA was loaded on the 5% native gel.

nucleosomal DNA. Yet, how the sequence changes in H2A.B affect the histone octamer conformation is largely unaddressed. Undoubtedly, this is in part related to the dynamic nature of the particle, complicating both crystallography and cryo-EM efforts. Here, we used state-of-the-art solution NMR to map the structure and

dynamics of the key sites in the H2A.B nucleosome at atomic resolution. In addition, we showed that a shift of critical arginine residue by one position can directly affect DNA wrapping.

The first part of our study focussed on assessing the nature of H2A.B in the variant nucleosome, in a way to verify the simplistic model of the H2A.B nucleosome in Figure 4.1. Our previous study of the H2A.B-H2B heterodimer showed that the histone core maintains the canonical fold in the variant dimer, and demonstrated that the docking domain in H2A.B is unstructured and highly flexible after the αC helix, which is incompatible with nucleosome formation. Using methyl TROSY, we how here that I106 within the H2A.B docking domain experiences a large change in chemical environment upon incorporation into nucleosome. Relaxation and NOE data further showed that the I106 δ1 methyl group has relaxation rates comparable to histone core residues and show an NOE that is fully consistent with folding of the docking domain in the canonical fold. Only the very C-terminal residue of the H2A.B docking domain, which is predicted to protrude from the H3-H4 surface, remains highly flexible in the variant nucleosome. Taken together our NMR data indicate that, very much like canonical nucleosomes, the H2A.B docking domain is folded in the variant nucleosome in the same manner as H2A in canonical nucleosomes.

Moreover, while the H2A.B nucleosome particle is seen as a dynamic system due to the (transient) unwrapping of DNA giving rise to flexible DNA protrusions, this does not seem to be mirrored in increased dynamics for the histone octamer core. The variant dimer has been shown to have high mobility in cells suggesting a low barrier to exchange out of the nucleosome [11-12]. This may be caused by the reduced interactions between H2A.B and H3 because of the truncation of the docking domain and reduced net negative charge for the remaining part (see Chapter 3). Beautiful work from the Luger and Kay labs showed that destabilization of the dimer-tetramer interface can result in conformational changes that generate very large effects in CPMG relaxation dispersion studies [28]. None such effects were detected for the H2A.B nucleosome, suggesting no large conformational changes either in the H2A.B core region or the docking domain. The increased ^{1}H transverse relaxation rates are likely rather

due to increase in effective size of the H2A.B nucleosome due to the partial DNA unwrapping.

We found significant changes in conformational dynamics for the H3 N-terminal tail, most likely directly related to the unwrapping of the DNA. The overall downfield shift, indicative of changes in the local electrostatic environment [29], for the H3 tail resonances suggests that the DNA binding equilibrium is shifted towards the unbound state in the variant nucleosome. In particular, some resonances in the tail are have chemical shifts very similar to that free peptide, suggesting that some stretches of the tail may be free from the DNA in the H2A.B-nucleosome. However, since the profile of the residue-specific local effective correlation times was almost unchanged from the canonical nucleosome, we conclude that the tail remains associated to the DNA, rapidly sampling many different conformations. Given the length of the tail and the local nature of the DNA-tail interactions, this likely means that the tail as a whole remains DNA bound at all times, but the precise location of the contacts points switches rapidly in time and each contact spends significant time in an unbound state. The region where the chemcial shifts most strongly resemble that of the free peptide, A29-V35, may be free at all times, but nevertheless indirectly linked to the DNA, thus limiting the flexibility. We thus amend the cartoon of the H2A.B nucleosome of Figure 4.1 by noting that the N-terminal part of the H3 tail is in fact bound to the DNA, likely limiting the opening angle between the DNA ends.

The second part of this study addressed the molecular basis for the partial DNA unwrapping in the H2A.B nucleosome. Previous work showed that truncation of the docking domain itself is not a main causative factor for DNA unwrapping [14], but rather that sequence differences in the α3-helix to C-terminal are responsible [17]. Our work establishes that this region of H2A.B is stably incorporated into nucleosomes, most likely in the same conformation as canonical H2A. This points to a rather subtle and indirect effect as this region is not involved in any DNA contacts directly. We found that the shifted arginine position in the L2 loop of H2A.B contributes directly to the open nucleosomal conformation as it is incompatible with anchoring to the DNA minor groove in the wrapped state. A H2A mutant containing the ERN sequence from the H2A.B L2 loop forms a nucleosome that is more vulnerable to MNase digestion and protects

118 bp DNA just like H2A.B nucleosomes. Nevertheless, the H2A_ERN nucleosome does not recapitulate the reduced stability of the H2A.B nucleosomes, in accordance for an additional and major role for sequence differences within the docking domain. We previously pointed to a reduced electrostatic match between the H2A.B docking domain and H3-H4 that may subtly destabilize the dimer/tetramer interface (see Chapter 3). We hypothesize that the reduced grip of H2A.B on both DNA and H3 may permit the H3 αN helix to transiently detach from the octamer core and thereby promote opening of the DNA. Yet, this process cannot involve a highly dynamic state of the αN helix is as that should have resulted in observable resonances in the amide backbone based NMR experiments. Interestingly, the arginine shift identified here is also present in other truncated histone variants, suggesting that the underlying DNA unwrapping mechanism is conserved.

In summary, we found that the histone octamer in the H2A.B nucleosome is remarkably like the canonical nucleosome. We provide evidence that that the docking domain is folded in the canonical configuration and as rigid as the histone core, pointing to defined interaction to H3. We find that the opened state of the variant nucleosome correlates with increased flexibility and reduced DNA binding for H3 N-terminal. Finally, we here show that a register shift of a minor groove anchoring arginine in the H2A.B L2 loop is an important and direct contributing cause for DNA unwrapping in H2A.B nucleosomes. With these results we have provided new insights in the structure of the H2A.B nucleosome. Furthermore, we look forward to combining these data with the outcome of ongoing MD simulations to further dissect the molecular nature of this peculiar nucleosome.

Material and methods

Site-directed mutagenesis. The plasmid containing H2A_ERN was produced by site-direct mutagenesis on plasmid pET-21b containing the *Drosophila* H2A_K74E made in Chapter 3. A pair of complementary mutagenic primers was designed to introduce 2 mutations T75R and R76N at once. The reaction was performed using

Site-Direct Mutagenesis Kit (Novagen) according to the following thermocycling: 12 cycles of 95 °C for 30 s, 52 °C for 30 s, 68 °C for 8 min. After the reaction, the 1 μL of DpnI was added and incubated at 37 °C for 1 hr to remove the original template. The generated nicked circular DNA was used to transform *E.coli* DH5α cells. Plasmids were extracted from successful transformations and the mutations were verified by sequencing.

Histone protein production. *Drosophila melanogaster* (Dm.) canonical histones H2A (Uniprot-id: P84051), H2B (Uniprot-id: P02283), H3 (Uniprot-id: P02299), H4 (Uniprot-id: P84040), *Homo sapiens* (Hs.) H2A.B (Uniprot: P0C5Z0), and mutant H2A_ERN were expressed in *E. coli* BL21 Rosetta 2 (DE3) cells (Novagen) and purified under denaturing conditions from inclusion bodies by extraction in 6 M guanidine chloride, followed by size-exclusion chromatography in buffer A (7 M urea, 50 mM NaPi pH 7.5, 1 mM EDTA, 150 mM NaCl, 5mM BME) using a Superdex 75 column (GE) and ion exchange with a salt gradient from buffer A to buffer A supplemented with 1 M NaCl. Histones used for NMR studies were produced in D_2O or H_2O-based M9 minimal medium containing desired isotopes and amino acid precursors in case of ILV methyl-labelling [30]. Histones used for MNase assays were produced in LB medium.

DNA production. The same protocol as described in a previous publication [20] was followed to produce the 167 bp '601'-Widom DNA used in this study. In short, a pUC19 plasmid containing 12 repeats of a 167 bp '601'-Widom sequence was amplified in *E. coli* DH5α and extracted by alkaline lysis followed by isopropanol and ethanol precipitations prior to dissolve in TE buffer (20 mM Tris-HCl, 5 mM EDTA, 100 mM NaCl, pH 7.5) and purify by anion exchange chromatography. The purified plasmid was restricted using ScaI (ThermoFisher) and the restricted fragment was purified by anion exchange chromatography.

Nucleosome reconstitution. Histone dimers, tetramers or octamers were refolded from equimolar mixes of denatured purified histones by dialysis to 2 M NaCl at 4°C [31]. Histone dimers, tetramers or octamers

were subsequently purified by size-exclusion chromatography over a Superdex 200 column (GE) in 2 M NaCl buffer with an addition of 5 mM BME (see also Figure S4.2). H2A.B nucleosomes were reconstituted from a mix of H2A.B-H2B dimer, H3-H4 tetramer, and 167 bp '601' DNA at ratio 2:1:1. H2A nucleosomes and H2A_ERN nucleosomes were either made the same way as H2A.B nucleosomes or directly from a mix of octamer and 167 bp '601' DNA at ratio 1:1. The mixes were dialysed at 4 °C from high salt buffer (10 mM KPi pH 6.5, 2 M KCl) to low salt buffer (10 mM KPi pH 6.5, 250 mM KCl) over 18 hours using a pump. Efficiency of the reconstitutions was analysed with 15% SDS-PAGE and 5% native-PAGE (see also Figure 4.2 and Figure S4.2). Nucleosomes were buffer exchanged to lower salt concentration for NMR or MNase studies (see below).

MNase study. Nucleosomes containing H2A, H2A_ERN, or H2A.B reconstituted from dimer, tetramer and DNA were first dialyzed to buffer with 10mM NaPi pH 6.5, 50 mM KCl. 4.5µg of each nucleosome was mixed with MNase reaction buffer (10 mM Tris pH 7.5, 5 mM Ca^{2+}) to a final volume of 30 µL per reaction. Reactions were incubated at 37 °C with 0.6 U MNase. At 0.1, 2.5, 5, 10, 15, and 30 min, 4 µL of each reaction mix was sampled, and the reaction stopped by adding 1 µL MNase stop buffer (100 mM EDTA pH 8.0). Reaction mixes after 30 min were deproteinized by 0.6 U proteinase K at 55°C for 90 min. DNA was extracted by precipitation with Na-acetate and 96% ethanol and re-suspended with 0.2×TBE buffer. All samples were analysed with 5% native-PAGE.

Molecular modelling. The simplistic model of H2A.B nucleosome as shown in Figure 4.1c was created based on the canonical nucleosome structure (PDB id: 1KX5) as a template. The path of nucleosomal DNA entry/exit ends was constructed by replacing the terminal 13 bp with straight B-form DNA in PyMOL software [32]. This model is to match the reported 118 bp DNA protected by the H2A.B containing octamer [14].

The H2A.B nucleosome structure used to initiate MD simulations shown in Figure 4.6b is a homology model based on the canonical nucleosome structure (PDB id: 1KX5), or, for Figure 4.6c, was built using the H2A.B solution structure for the core region with C-terminal

and N-terminal structure modelled according to canonical conformation (PDB id: 1KX5) using MODELLER [33].

NMR spectroscopy. For H3 tail dynamics study, $^{15}N/^{13}C$ labelled H3 was reconstituted into nucleosomes with unlabelled H2A or H2A.B, H2B, H4 and DNA. Samples of H2A.B nucleosomes (32 μM) or H2A nucleosomes (93 μM) in NMR buffer 1 (10 mM KPi pH 6.5, 10 mM KCl, 7.5% D_2O) were measured at 303 K on a Bruker Avance III spectrometer equipped with a cryo-probe and operating at 22.3 T corresponding to 950 MHz ^1H Larmor frequency. Assignments for H3 tail residues in H2A.B nucleosomes were transferred from canonical nucleosomes published before [20]. T_1, T_2 relaxation experiments were analysed using PINT software [34].

For comparison of H2A.B N-terminal tail resonances between H2A.B-H2B heterodimer and H2A.B nucleosomes, ^{15}N labelled H2A.B was reconstituted into either dimers or nucleosomes with unlabelled H2B, H3, H4 and DNA. Samples of H2A.B-H2B dimer (20 μM) or H2A.B nucleosomes (20 μM) were carried out at 303 K in buffer 2 (20 mM NaPi pH 6.5, 5 % D_2O, 0.01 % NaN_3) on a Bruker Avance III spectrometer operating at 20.0 T corresponding to 850 MHz 1H Larmor frequency. Assignments of H2A.B were transferred from Chapter 3.

Methyl-labelled nucleosomes and methyl-labelled H2A.B-H2B dimers were prepared with Ile-δ1-[$^{13}CH_3$], Leu/ Val-[$^{13}CH_3$, $^{12}CD_3$]-H2A.B or H2A and perdeuterated H2B, H3, and H4 with unlabelled DNA. Data for methyl-labelled nucleosomes were collected in NMR buffer 3 (20 mM NaPi pH 6.2, 0.01% NaN_3, 0.5 mM PMSF, 100% D_2O), while methyl labelled H2A.B-H2B dimers were measured in NMR buffer 3 with 200mM NaCl. Methyl-TROSY spectra of H2A.B-H2B dimer (160 μM), H2A.B nucleosome (85 μM), canonical nucleosome (100 μM) were measured on a Bruker Avance III spectrometer operating at 14.1 T corresponding to 600 MHz 1H Larmor frequency. Nucleosomes spectra were recorded at temperature of 318 K, while dimer spectrum was recorded at 303 K. Spectra were processed to equal acquisition times with identical processing parameters. For each sample, the peak intensities were normalized based on the median value of each spectrum to allow comparison of the intensity distribution between the

different systems. N-terminal tail residues were excluded in this analysis.

CPMG relaxation dispersion experiments of H2A.B nucleosomes were performed at 318 K on a Bruker Avance III spectrometer operating at 14.1 T corresponding to 600 MHz 1H Larmor frequency. For the measurement of ^{1}H single quantum relaxation rates, a relaxation delay of 10 ms was used to collect 11 CPMG pulse frequencies ranging from 100-1000 Hz with three duplicate points for experimental error assessment. For the multiple-quantum CPMG relaxation experiments, a relaxation delay of 15 ms was used to collect 13 CPMG pulse frequencies ranging from 67- 2000 Hz with three duplicate points for experimental error assessment. All dispersion relaxation data were analysed using PINT [34].

NOESY experiments of H2A.B nucleosomes were measured with mixing time 200 ms at 318 K on a Bruker Avance III spectrometer operating at 22.3 T corresponding to 950 MHz 1H Larmor frequency, equipped with a cryo-probe.

Acknowledgements

We thank Lewis Kay (University of Toronto) for providing the CPMG relaxation dispersion pulse sequences.

References

1. 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*.

2. Long, M.; Sun, X.; Shi, W.; Yanru, A.; Leung, S. T. C.; Ding, D.; Cheema, M. S.; MacPherson, N.; Nelson, C. J.; Ausio, J.; Yan, Y.; Ishibashi, T., A novel histone H4 variant H4G regulates rDNA transcription in breast cancer. *Nucleic Acids Res* **2019**.

3. Talbert, P. B.; Henikoff, S., Histone variants on the move: substrates for chromatin dynamics. *Nat Rev Mol Cell Biol* **2017,** *18* (2), 115-126.

4. 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.

5. 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.

6. 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.

7. 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.

8. 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.

9. 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.

10. 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.

11. 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.

12. 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.

13. 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.

14. 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.

15. 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.

16. 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.

17. 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.

18. 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.

19. 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.

20. Xiang, S.; le Paige, U. B.; Horn, V.; Houben, K.; Baldus, M.; van Ingen, H., Site-Specific Studies of Nucleosome Interactions by Solid-State NMR Spectroscopy. *Angew Chem Int Ed Engl* **2018,** *57* (17), 4571-4575.

21. Zhou, B. R.; Feng, H.; Ghirlando, R.; Kato, H.; Gruschus, J.; Bai, Y., Histone H4 K16Q mutation, an acetylation mimic, causes structural disorder of its N-terminal basic patch in the nucleosome. *J Mol Biol* **2012,** *421* (1), 30-7.

22. Tugarinov, V.; Hwang, P. M.; Ollerenshaw, J. E.; Kay, L. E., Cross-correlated relaxation enhanced 1H[bond]13C NMR spectroscopy of methyl groups in very high molecular weight proteins and protein complexes. *J Am Chem Soc* **2003,** *125* (34), 10420-8.

23. 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.

24. Zheng, C.; Lu, X.; Hansen, J. C.; Hayes, J. J., Salt-dependent intra- and internucleosomal interactions of the H3 tail domain in a model oligonucleosomal array. *J Biol Chem* **2005,** *280* (39), 33552-7.

25. Korolev, N.; Allahverdi, A.; Yang, Y.; Fan, Y.; Lyubartsev, A. P.; Nordenskiold, L., Electrostatic origin of salt-induced nucleosome array compaction. *Biophys J* **2010,** *99* (6), 1896-905.

26. Stutzer, A.; Liokatis, S.; Kiesel, A.; Schwarzer, D.; Sprangers, R.; Soding, J.; Selenko, P.; Fischle, W., Modulations of DNA Contacts by Linker Histones and Post-translational Modifications Determine the Mobility and Modifiability of Nucleosomal H3 Tails. *Mol Cell* **2016,** *61* (2), 247-59.

27. Muthurajan, U. M.; Park, Y. J.; Edayathumangalam, R. S.; Suto, R. K.; Chakravarthy, S.; Dyer, P. N.; Luger, K., Structure and dynamics of nucleosomal DNA. *Biopolymers* **2003,** *68* (4), 547-556.

28. Kitevski-LeBlanc, J. L.; Yuwen, T.; Dyer, P. N.; Rudolph, J.; Luger, K.; Kay, L. E., Investigating the Dynamics of Destabilized Nucleosomes Using Methyl-TROSY NMR. *J Am Chem Soc* **2018,** *140* (14), 4774-4777.

29. Borgia, A.; Borgia, M. B.; Bugge, K.; Kissling, V. M.; Heidarsson, P. O.; Fernandes, C. B.; Sottini, A.; Soranno, A.; Buholzer, K. J.; Nettels, D.; Kragelund, B. B.; Best, R. B.; Schuler, B., Extreme disorder in an ultrahigh-affinity protein complex. *Nature* **2018,** *555* (7694), 61-66.

30. Tugarinov, V.; Kanelis, V.; Kay, L. E., Isotope labeling strategies for the study of high-molecular-weight proteins by solution NMR spectroscopy. *Nat Protoc* **2006,** *1* (2), 749-54.

31. Dyer, P. N.; Edayathumangalam, R. S.; White, C. L.; Bao, Y.; Chakravarthy, S.; Muthurajan, U. M.; Luger, K., Reconstitution of nucleosome core particles from recombinant histones and DNA. *Methods Enzymol* **2004,** *375*, 23-44.

32. The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.

33. Webb, B.; Sali, A., Comparative Protein Structure Modeling Using MODELLER. *Curr Protoc Protein Sci* **2016,** *86*, 2 9 1-2 9 37.

34. Ahlner, A.; Carlsson, M.; Jonsson, B. H.; Lundstrom, P., PINT: a software for integration of peak volumes and extraction of relaxation rates. *J Biomol NMR* **2013,** *56* (3), 191-202.

35. Taguchi, H.; Horikoshi, N.; Arimura, Y.; Kurumizaka, H., A method for evaluating nucleosome stability with a protein-binding fluorescent dye. *Methods* **2014,** *70* (2-3), 119-26.

Supplements

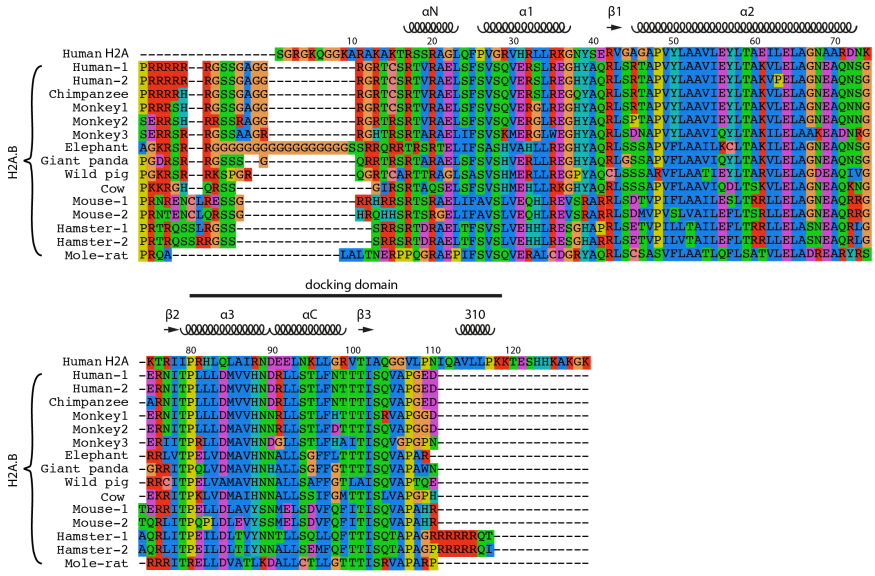

Figure S4.1. Sequence alignment of H2A.B with canonical H2A. Fifteen H2A.B sequences from different species that are available on *Histone Database 2.0* [1] are aligned and compared to human canonical H2A. Secondary structural elements of H2A in the canonical nucleosome structure are indicated above the sequences.

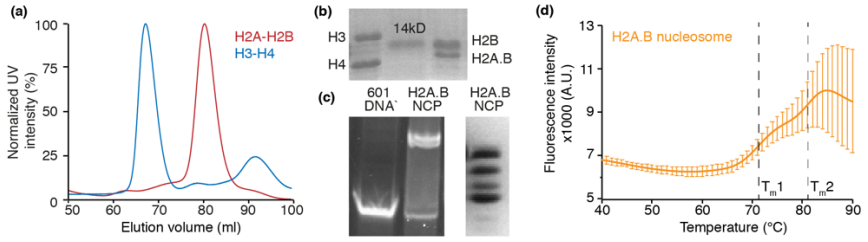

Figure S4.2. Reconstitution of H2A.B-nucleosomes. (a,b) Chromatograph and SDS-PAGE of purification of refolded H2A.B-H2B dimers and H3-H4 tetramers used in the nucleosome reconstitution. **(c)** Native PAGE (left) and SDS-PAGE (right) analysis of reconstituted H2A.B nucleosomes used for subsequent NMR studies. **(d)** Thermal melting curves of H2A.B-nucleosomes as measured in a SYPRO-orange based thermostability assay. Thermal shift assay showed the presence of two-phase melting transition, with a first transition occurring at 71 (T_m1) and a second at 81 °C (T_m2), correspondingly well to previously published

values [35]. These transitions correspond to stepwise release of the H2A.B-H2B dimer and H3 and H4 histones respectively.

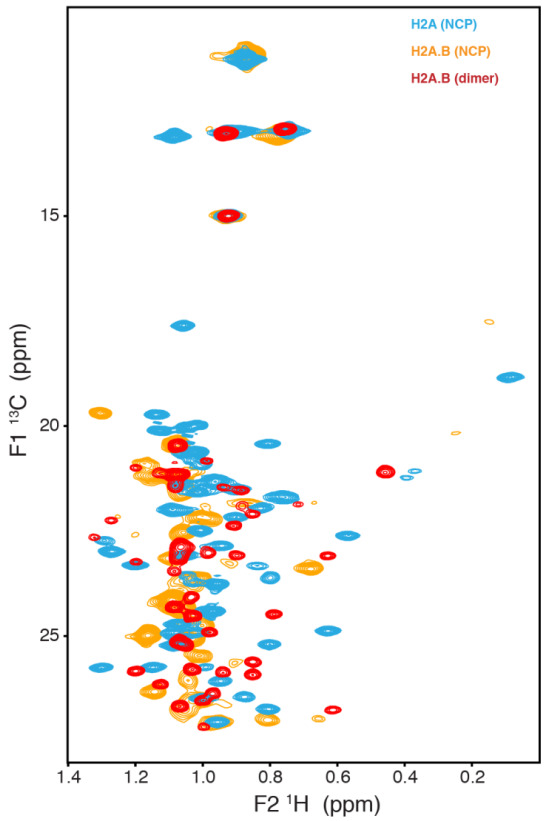

Figure S4.3. Complete Me-TROSY spectra overlay for H2A.B in either nucleosomes (orange) or H2A.B-H2B heterodimer (red), and H2A in canonical nucleosomes (cyan). The H2A.B spectra of nucleosome and heterodimer show significant changes in overall peak pattern. Similar differences between dimer and nucleosome chemical shifts where observed for canonical H2A and attributed to the formation of the large interfaces to the DNA, H3-H4 and the second copy of the H2A-H2B dimer in the nucleosome [20]. The changes observed here are thus consistent with formation of a well-folded nucleosome.

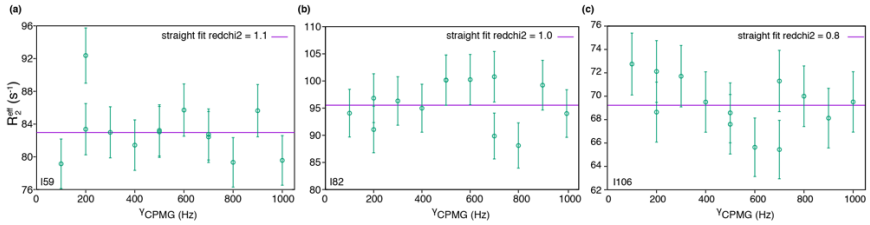

Figure S4.4 1H SQ CPMG relaxation dispersion profiles for H2A.B isoleucine's I59 (a), I82 (b), and I106 (c) in nucleosomes.