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

Zhang, H.

Chapter 3. Variant H2A.B-H2B histone dimer is more stable than the canonical dimer

This chapter is based on:
Heyi Zhang, Sofie Dubbeldam, Hugo van Ingen. Variant H2A.B-H2B histone dimer is more stable than the canonical dimer. *Submitted.*

Contributions of authors:
Buffer screening assays for thermostability comparison between histone dimers were carried out by Sofie Dubbelman.

Abstract

Histone variants incorporated nucleosomes with specific structural and functional properties offer the opportunity to regulate chromatin biology. Histone variant H2A.B is one of the most divergent histone variants known to date and has been shown to be associated with regulation of gene transcription, DNA replication, and RNA splicing. Compared to those with the canonical H2A, nucleosomes containing H2A.B display a 'loosened' conformation in which the entry and exit DNA are partly unwrapped from the histone octamer core. In this study, we studied the structure, dynamics and stability of the H2A.B-H2B heterodimer with solution NMR. Using backbone chemical shift and sparse NOE data in combination with CS-Rosetta modelling, we determined the solution structure of H2A.B-H2B. The dimer shows the canonical histone fold with highly flexible and disordered tails. The solution structure of the histone core is well-defined and matches the recently determined H2A.B-H2B crystal structure. The H2A.B-H2B core has an overall reduced electrostatic surface potential, compared to the canonical H2A-H2B dimer. Based on comparative thermal stability assays on a series of mutants, we show that H2A.B-H2B dimer has increased thermostability compared to the canonical dimer, and that stability correlates well with overall net charge of the dimer core. Structural analysis highlights that the distinct electrostatic properties of H2A.B-H2B dimer not only destabilize DNA binding but may also weaken the interaction with the H3-H4 tetramer. The increased dimer stability may facilitate the dynamic and chaperone-free exchange of the H2A.B-H2B dimer from the nucleosome.

Introduction

Eukaryotic cells have evolved a number of mechanisms to regulate the basic processes that need access to the DNA, such as transcription, replication and repair. Next to chromatin remodeling and the installation of post-translation modifications, the incorporation of variant histone proteins is a key mechanism to confer specific functional properties to chromatin at defined loci. These histone variants replace the abundant, canonical form of H2A, H2B, H3 or H4 histones that make up the histone octamer core in the nucleosome, the basic unit of chromatin. Histone variants, which unlike canonical histones can be expressed throughout the whole cell cycle, have been identified for all four core histones.[1,2,3] Due to their different amino acid sequence, histone variants can alter nucleosome structure and dynamics, promote or inhibit nucleosome-protein interactions, or provide different post-translational modification sites to regulate chromatin function.[4,5,6] For example, H3 variant CENP-A is essential for centromere identity,[7] while histone H2A variant H2A.X signals DNA damage by phosphorylation on its C-terminal serine which as a result recruits and accumulates DNA repair proteins to damage sites.[8] H2A.B is one of the most divergent H2A histone variants,[9-10] sharing only 48% sequence identity to the canonical H2A.[11] While most sequence differences with the canonical H2A occur in the N- and C-terminal tails, there are also a considerable number of substitutions in the core of the protein: 36 for the human variant and 45 for the mouse variant (Figure 3.1a). The majority of these substitutions involve reversal, loss or addition of charge. Notably, sequence conservation for H2A.B orthologues is significantly lower than for canonical H2A or H2B (Figure 3.1a, 1b).

As discussed in Chapter 1, a wide array of biochemical and low-resolution structural methods have been used to show that the DNA ends in H2A.B-incorporated nucleosomes are transiently unwrapped from the histone octamer, resulting in the stable association of 110-130 bp instead of 147bp DNA as in the conventional nucleosomes [12-14]. In addition, H2A.B nucleosome arrays are less compact and may thus improve transcription efficiency and facilitate DNA replication and repair [14-20,21]. Despite the importance of H2A.B in these essential cell functions, the molecular basis of how this variant confers its special

properties to the nucleosome is not fully resolved. The partly neutralized "acidic patch", a key interaction site for proteins that bind the canonical nucleosome[22] and site of nucleosome-nucleosome interactions, as well as the near-absence of lysine residues, which are main sites for post-translational modifications (see also Figure 3.1a), impact the functional properties of the H2A.B nucleosome. Nevertheless, these changes are unlikely to affect nucleosome structure directly. In this regard, previous studies have highlighted the impact of the altered and truncated C-terminal docking domain (see Figure 3.1a) of H2A.B on its less compact structure of nucleosomes[12-13, 23-24]. The histone fold core of H2A.B-H2B may thus additionally contribute to peculiar properties of the H2A.B nucleosome.

Recently, the crystal structure of the H2A.B-H2B dimer core was solved, which showed that dimer has the canonical histone fold and significantly reduced positive charge on its DNA binding surface compared to the canonical dimer.[25] For canonical H2A-H2B, the solution structure showed overall the same fold as in the nucleosome but with considerable structural heterogeneity, suggesting that intrinsic structural flexibility of the H2A-H2B dimer facilitates its association with DNA and other histone partners into nucleosomes.[26] The extent of intrinsical flexibility in H2A.B-H2B is unknown but may give clues to understand the destabilizing influence on the nucleosome and the increased exchange dynamics of this dimer.

Here, we studied the structure, dynamics and stability of the H2A.B-H2B dimer in solution. Using chemical shift and sparse NOE data based Rosetta protein folding, the solution structure of the H2A.B-H2B dimer was determined. We find that H2A.B-H2B folds into the canonical histone fold with a well-defined core region and disordered, highly flexible N- and C-terminal extensions, including part of the docking domain. The H2A.B α2 helix is bent in solution as was observed in the crystal. The structure further shows that the interaction surface with H3-H4 has an altered electrostatic potential that may contribute to their looser association in addition to the truncation. Surprisingly, we find that the variant dimer has significantly increased thermal stability compared to the canonical dimer. Through a mutational analysis we show that reduced electrostatic repulsion between H2A.B and H2B monomers causes enhanced stability, which may contribute to the increased exchange dynamics for H2A.B.

Materials and methods

Histone protein production and refolding. *Drosophila melanogaster* (*Dm.*) canonical histones H2A (Uniprot-id: P84051) and H2B (Uniprot-id: P02283) and *Homo sapiens* (*Hs.*) H2A.B (Uniprot: P0C5Z0) 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, 1 mM EDTA, 150 mM NaCl, pH 7.5) using a Superdex 75 column (GE) and ion exchange with a salt gradient from buffer A to buffer A supplemented with 1 M NaCl.[27] Histones used for NMR studies were produced in D_2O or H_2O-based M9 minimal medium containing desired isotopes. Histones used for thermal shift assays were produced in LB medium. Histone dimers were refolded from equimolar mixes of denatured purified histones by dialysis to 2 M NaCl at room temperature and subsequent purification using size-exclusion chromatography over a Superdex 200 column (GE) in 2M NaCl buffer.[27] Purified dimers were stored at 4 °C before buffer exchange to lower salt concentration for NMR or thermal stability studies.

Site-directed mutagenesis. Site-directed mutagenesis was performed on the pET-21b plasmid containing the *Drosophila* H2A gene or on the pET-3a plasmid containing human H2A.B gene. For each mutation, a pair of complementary mutagenic primers was designed and used to amplify the entire plasmid in a thermocycling reaction using in-house produced Pfu DNA polymerase. The template plasmid was digested with Dpn1 after reaction, and the generated nicked circular DNA was used to transform *E.coli* DH5α cells. Successful transformations were used to extract plasmids after which correctness of mutations was verified by sequencing.

NMR spectroscopy. All NMR experiments were performed at 35 °C on a Bruker Avance III spectrometer operating at 20.0 (21.1) T corresponding to 850 (900) MHz ^{1}H Larmor frequency, equipped with a cryo-probe. Samples for H2A.B and H2B backbone assignment contained ~0.15 mM H2A.B-H2B refolded from either fractionally deuterated, uniformly ^{15}N/^{13}C-labeled H2A.B and unlabeled H2B or

vice versa. H2A.B data was collected in NMR buffer 1 (20 mM MES, 50 mM NaCl, pH 6, 0.1 mM PMSF, 5 mM DTT), while H2B data was collected at higher salt levels to promote stability of the complex (NMR buffer 2: 20 mM $NaPO_4$, 200 mM NaCl, pH 6.5, 5 mM β-mercaptoethanol). ^{1}H-^{15}N HSQC fingerprint spectra showed negligible chemical shift differences between these conditions. Samples for measurement of ^{15}N T_1 and T_2 relaxation times contained 0.15-0.3 mM dimer with either ^{15}N-labeled H2A.B or H2B.

Backbone resonances for H2A.B and H2B in the H2A.B-H2B heterodimer were assigned using standard sequential backbone assignment approach based on the set of 3D TROSY versions of HNCO, HN(CA)CO, HNCACB, HNCA, HNCB, HN(CO)CA experiments. In total, 89% and 91% of the H_N, N, C_α, C_β, C' backbone resonances were assigned for the entire sequence of H2A.B and H2B respectively. Backbone resonances for the core regions of H2A.B (V31-F101) and H2B (Y34-K122) were assigned to 99% and 93% completion, respectively. No chemical shift assignment could be obtained for the first 8 residues (PRRRRRRG) and C19 in the N-terminal tail together with S28, F29 of H2A.B, and P1, S5, G6 in the N-terminal tail together with S53 in the core and T119, S120 in the C-terminus of H2B, either due to extensive overlap or missing connectivity information.

A 3D ^{15}N-edited NOESY with 200 ms mixing time was recorded on a sample containing 0.38 mM dimer in NMR buffer 2. The dimer was refolded from perdeuterated ^{15}N-labeled H2B and unlabeled H2A.B to selectively record intermolecular NOEs.

All NMR data were processed using Bruker Topspin, or NMRPipe[28] and analyzed using NMRFAM-Sparky[29]. Secondary structure of H2A.B-H2B dimer was predicted by TALOS-N using H_N, N, C_α, C_β, C' chemical shifts.[30]

Structure calculations. Calculation of the structure of the H2A.B-H2B dimer was carried out in a two-step approach. First, an initial model based only on backbone chemical shifts was obtained using CS-Rosetta.[31,32] Then, this model was used to assign intermolecular NOEs between H2A.B and H2B. These NOEs were subsequently used as additional constraints in a final CS-Rosetta structure calculation, after which modeling of the N- and C-terminal tails yielded the final solution

structure of H2A.B-H2B. For the first step, the H_N, N, C_α, C_β, C' backbone chemical shifts of the core regions of H2A.B (V31-F101) and H2B (Y34-K122) were used to calculate 3000 structures of the H2A.B-H2B histone fold core using the CS-Rosetta webserver (https://csrosetta.bmrb.wisc.edu/csrosetta/submit). To allow Rosetta to fold the dimeric core, the two proteins were connected by a random coil $(Gly)_8$ poly-glycine linker into a single chain. From the 20 lowest energy models, the 10 structures with the lowest C_α RMSD to the lowest energy model were selected as the best models.

The initial CS-Rosetta ensemble was used to semi-automatically assign intermolecular NOEs in the ^{15}N-edited NOESY recorded on a H2A.B-[U-^{2}H,^{15}N]-H2B sample using an in-house written python script. All H2B H_N protons were assigned based on backbone (H_N, N) chemical shifts. One intermolecular NOE could be assigned unambiguously based on the backbone assignments as an H_N-H_N contact. For all other observed NOEs, we relied on chemical shifts statistics, correspondence to the initial structural model, and networking support to assign the H2A.B proton, in a procedure akin to that implemented in CYANA.[33] Briefly, candidate assignments were compiled based on correspondence to the atom and residue specific average chemical shift as deposited in the BMRB using a chemical shift tolerance of 0.4 ppm. These candidate assignments were then filtered according to their match to the structural model, requiring that the candidate NOE was satisfied in at least 60% of models in the initial CS-Rosetta ensemble using a distance cutoff of 7 Å for intense peaks (S/N ≥ 10) and 8 Å for weaker peaks (S/N < 10). This resulted in 35 unambiguously assigned NOEs in which there was only one assignment possibility. These assignments were used as anchors to extend the assignments in a network approach in which cross peaks with ^{1}H chemical shifts within 0.01 ppm were assigned to same H2A.B proton wherever structurally supported, resulting in 44 assigned intermolecular NOEs in total. For 20 intense cross peaks (S/N ≥ 10) multiple assignment possibilities were obtained and could thus not be assigned.

The NOEs were converted to distance constraints based on a calibration of H2B backbone H_N-H_N NOEs and average helical *i*, *i*+1, *i*+2, *i*+3 distances. Each NOE was converted into one of five categories of upper distance restraints: <= 2.4, 2.4–3, 3–4, 4–5 and 5–7 Å.

Intensities of NOEs involving methyl protons were divided by three per methyl group before calibration.

In final step of the structure calculation, the distance restraints derived from the intermolecular NOEs were used together with the chemical-shift generated fragments to calculate the structure of the H2A.B-H2B core region (H2A.B V31-F101 and H2B Y34-K122 connected by a poly-glycine linker) using the Rosetta *AbInitio* protocol as described previously[34] (www.rosettacommons.org/demos/latest/public/abinitio_w_chemicalshift_noe/README). In total, 8000 structures were generated and scored on the basis of their the full-atom and distance restraint energy (*E*) by Rosetta. For each model, backbone chemical shifts of $^{13}C_\alpha$, $^{13}C_\beta$, ^{13}C, $^{1}H_N$ and ^{15}N were predicted by SPARTA+.[35] The correspondence between predicted and experimental chemical shifts was used to rescore the initial Rosetta energy to a rescored energy (*E'*) according to:

$$E' = E + c \times \chi^2_{CS}$$

where *c* is a weighting factor set to 0.25, and

$$\chi^2_{CS} = \sum_i \sum_j (\delta^{exp}_{i,j} - \delta^{pred}_{i,j})^2 / \sigma^2_{i,j}$$

where $\delta^{exp}_{i,j}$ is the backbone chemical shift observed from NMR experiments for each atom type i ($^{13}C_\alpha$, $^{13}C_\beta$, ^{13}C, $^{1}H_N$, ^{15}N) for a given residue j; and $\delta^{pred}_{i,j}$ is the backbone chemical shift predicted by SPARTA+; and $\sigma^2_{i,j}$ is the uncertainty of $\delta^{pred}_{i,j}$. The model with the lowest *E'* was selected as the reference model, and C_α-RMSD of each model to the reference model was calculated. The 10 models with the lowest C_α-RMSD from the 20 lowest energy models were selected as the final structure ensemble for the histone fold core. Finally, the poly-glycine linker was removed and disordered N-and C-terminal tails of H2A.B and H2B were added using MODELLER[36] to result in the final solution structure of H2A.B-H2B.

Electrostatics were calculated using the adaptive Poisson-Boltzmann solver[37] using the lowest-energy structure. All structure images were created using open source PyMOL (The PyMOL Molecular Graphics System, Version 1.7, Schrödinger, LLC).

Thermal stability assays. A buffer series with pH range of 6 to 8 and salt concentrations from 50 mM to 250 mM NaCl was made to test thermo-stability of H2A.B-H2B dimer. For each well on 96-well plate, final dimer concentration was 5 μM (corresponding to a 30-fold dilution from the stock) with 1000-fold diluted Sypro Orange (Sigma-Aldrich) in a 25 μL final volume. Thermo-stability of mutant dimers (~10-20-fold diluted from stock) were measured with same combinations in 20 mM phosphate buffer at pH 6.5 with 200 mM NaCl. Measurements were performed in duplicate with the Biorad CFX96 Real-Time System. The temperature gradient was increased from 20°C to 80°C at 1°C/ minute and the fluorescence signal was recorded using FRET channel. The melting temperature was derived from the maximum of the first derivative of the absorbance signal using the Bio-Rad CFX manager program.

Results

H2A.B-H2B heterodimer has the canonical histone-fold core in solution. To study structure and dynamics of H2A.B-H2B in solution, we refolded the heterodimer using either isotope-labeled H2A.B with unlabeled H2B or vice versa and studied these using NMR spectroscopy. Of note, in this study we used *Drosophila* H2B that carries four conservative substitutions in the core region compared to the human sequence (Figure 3.1b). Both H2A.B and H2B give rise to high-quality TROSY fingerprint spectra, confirming that the dimers are well-folded (Figure S3.1). Near-complete backbone assignment could be obtained for both histones. Chemical shift indices (CSI) obtained from the experimental Cα and Cβ chemical shifts confirm the presence of the characteristic histone fold, a central long α-helix flanked by two shorter helices, for both histones (Figure 3.1c). In addition, the C-terminal αC helix of each histone is present together with the two β-strands that form two intermolecular β-sheets. The CSI further indicates that the N-terminal tail of both H2A.B and H2B and C-terminal end of H2A.B docking domain (F101-D114) are disordered in solution. Overall, the secondary structure matches well to the crystal structure of the variant H2A.B-H2B heterodimer[25]. Interestingly, in solution the dimer lacks the N-terminal αN-helix that is observed in

the H2A.B-H2B crystal. Similarly, the αN-helix is also unfolded in the canonical H2A-H2B heterodimer in solution, while it is a defined structural element in crystal structures of the nucleosome[23,26,38]. Secondary structure prediction by TALOS-N[30] based on the experimental backbone chemical shift values confirmed the observations from CSI, albeit that in both H2A.B and H2B lower propensities for the presence of the β2 strand were found (Figure 3.1c). Overall, these data indicate that the H2A.B-H2B dimer forms a well-folded histone fold domain with disordered tails. It further suggests the H2A.B αN helix is a labile secondary structure element that can be stabilized by contacts to the DNA, as seen for canonical H2A, or by crystal contacts as seen in the H2A.B-H2B crystal structure.

Backbone chemical shifts do not suffice to define the H2A.B-H2B core structure. Initially, we followed the approach used by Moriwaki *et al.*[26] to calculate the structure of the dimer based on backbone chemical shifts using CS-Rosetta. This approach relies on the Rosetta protein folding engine, supplemented by chemical shift based selection of protein fragments that are assembled into a 3D fold and followed by chemical shift based scoring of the final solutions. Using the experimental H_N, N, Cα, Cβ, and C' backbone chemical shifts, 3000 models for the H2A.B-H2B core were generated using the BMRB CS-Rosetta web server. While the calculation converged, superposition of the 10 models with the lowest C_α-RMSD from the lowest energy 20 models showed that the H2A.B α1 and H2B αC helices are ill-defined within the otherwise reasonably defined histone fold-core (Figure S3.2). While such structural heterogeneity could be intrinsic feature of the H2A.B-H2B dimer, it could also be the result of limitations in the modeling approach and input data.
We thus decided to refine this initial model with sparse distance restraints obtained from an intermolecular NOESY experiment. Exploiting the *in vitro* reconstitution of the H2A.B-H2B complex, we refolded dimers from unlabeled H2A.B and perdeuterated and ^{15}N-labeled H2B. This labeling scheme ensures that only intermolecular NOEs will be observed in a standard ^{15}N-NOESY experiment, with the obvious exception of intramolecular H_N-H_N NOEs. To assign these intermolecular NOEs to their corresponding ^{1}H atoms in H2A.B, we took advantage of the initial CS-Rosetta model as a structural reference

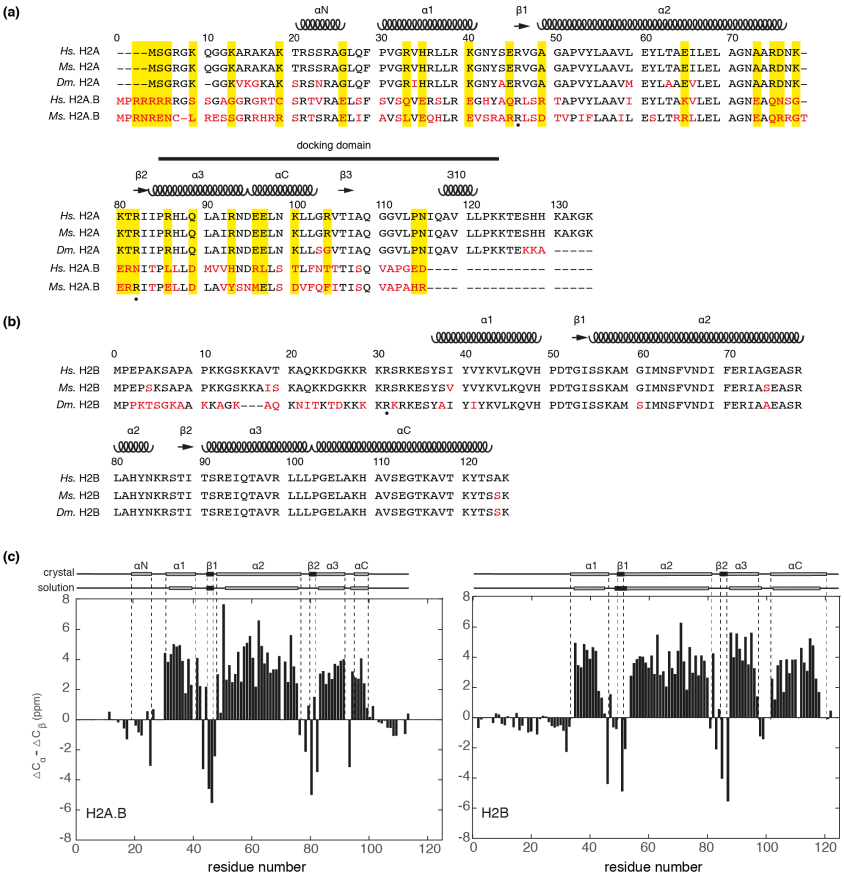

Figure 3.1. Sequence analysis and secondary structure of the H2A.B-H2B heterodimer. (a, b) Sequence alignment of human, mouse, and drosophila H2A, H2A.B (a), and H2B (b) sequences; *Hs.* – *Homo sapiens*; *Mm.* – *Mus musculus*; *Dm.* – *Drosophila melanogaster*. Sequence differences from *Hs.* are highlighted in red. Secondary structure in the human canonical nucleosome (PDB-id: 3afa) is indicated above the sequence. Minor groove arginines are marked with a black dot. In (a) substitutions involving charged residues are highlighted in yellow. **(c)**. Chemical shift indices based on Cα and Cβ chemical shifts for H2A.B and H2B in the H2A.B-H2B heterodimer, confirming the presence of the histone-fold in solution. Secondary structure elements in the crystal structure of the H2A.B-H2B dimer (PDB-id 6a7u) and as predicted by TALOS-N are indicated above the bar graph and labeled. Dashed lines are drawn to indicate the beginnings and ends of each secondary structure element observed in the crystal structure.

in a semi-automatic assignment procedure akin to that implemented in CYANA[33] (see Materials and Methods).
This allowed the unambiguous assignment of 44 intermolecular NOEs, among which 8 involve the H2B αC helix, and 36 involve the converged core region of the dimer. For example, clear intermolecular H_N-H_α and H_N-H_β contacts are observed for H2B residues G50 and S52, indicating formation of β-sheet between the H2B β1-strand and H2A.B (Figure 3.2a,b). Of interest, several clear intermolecular NOEs were observed for residues in the H2B αC helix, including H106 and A107, indicating the structure of the H2B αC helix is well-defined within the H2A.B-H2B heterodimer (Figure 3.2c,d). Unfortunately, no intermolecular NOEs to residues in the H2A.B α1 helix could be established, likely because it is positioned relatively far from H2B H_N atoms. The extent of intermolecular NOEs observed for the H2B αC helix demonstrate that it has a well-defined position in the core of the dimer. We thus conclude that backbone chemical shifts do not suffice to properly define the H2A.B-H2B structure.

Solution structure of the H2A.B-H2B heterodimer highlights altered electrostatic potential. The intermolecular NOEs and backbone chemical shifts were combined in a Rosetta calculation according to established protocols (see Materials and Methods). The calculation converged well (Figure S3.3), resulting in a final ensemble of 10 structures of the H2A.B-H2B core (Figure 3.3a). Structural statistics are reported in Supplementary Table S3.1. The histone-fold core as well as the orientation of the H2B αC-helix are well-defined (heavy atom backbone RMSD 1.1 ± 0.2 Å). Inclusion of the NOE data resulted in both local and overall better definition of the structure compared to using backbone chemical shifts only (Figure S3.4). The position of H2A.B α1-helix within the core is somewhat less well-defined. Since backbone dynamics of this helix, as judged from ^{15}N T_1/T_2 ratios, are not significantly different from the rest of the core (Figure S3.5), this is most likely due to the lack of intermolecular NOE data. Notably, these data further support that the H2A.B docking domain is flexible and unstructured after the αC helix.

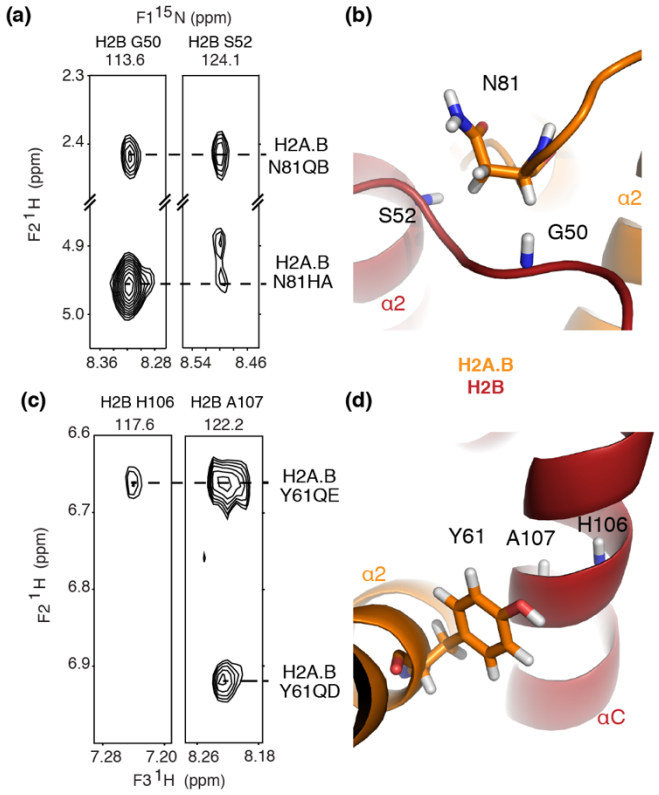

Figure 3.2. Intermolecular NOEs between H2A.B and H2B within the histone core support a well-defined conformation. (a, c) Strips from the 3D ^{15}N-edited NOESY recorded on heterodimers with perdeuterated ^{15}N-labeled H2B and unlabeled H2A.B showing intermolecular NOEs between H2B H_N and H2A.B protons. Assignment of H2B H_N and H2A.B resonances are indicated. **(b, d)** Zoom on the lowest-energy CS-Rosetta model corresponding to the intermolecular contact shown in the NOESY spectrum. NOEs shown in (a) restrain the H2A.B β2- H2B β1 β-strand, NOEs in (c) restrain the H2B αC-helix position.

The solution structure agrees well to the recent crystal structure of the H2A.B-H2B core with backbone RMSD of 1.7 Å (Figure 3.3b). The H2A.B α2 helix is slightly bent (average bend angle 25 ± 8° over the whole ensemble), as was observed also in the crystal structure (20° bend angle). While Dai *et al.* hypothesized the bend may be due to the crystal packing, its observation here demonstrates it is an intrinsic property of the dimer. Dai *et al.* have pointed out that this bend is less prominent in the canonical H2A structure in the nucleosomal

conformation, causing a slight change in the position of the DNA-binding L2 loop. Careful analysis shows however that, while the α2-helix in canonical H2A is indeed less bent (17° bend angle), slight changes in the relative position of the α2 helix within the dimer core also contribute to this difference in position. This is also apparent from comparison of the solution H2A.B-H2B structure with the canonical H2A-H2B structure in the nucleosome (Figure 3.3c). Interestingly, the solution structure of canonical H2A-H2B dimer also shows an increased bend angle for the H2A α2 helix (32 ± 9°), suggesting that the quaternary state of the H2A-H2B dimer, whether it is an isolated dimer or incorporated into nucleosomes, is a greater influence on the structure than the H2A variant type.

The most striking difference between the solution and crystal structure of H2A.B-H2B is the absence of the H2A.B αN helix in solution (see Figure 3.3b). Compared to canonical H2A, most amino acid substitutions occur on the outward facing surface of the core. Notably, one of the major structural differences between the canonical and variant dimer was shown by Dai *et al.* to be the substitution of E92 in H2A to L96 in H2A.B, which disrupts the hydrogen bonding to H2B residues E105 and L106.[25] A changed hydrogen bonding in solution is supported by the large amide backbone chemical shifts perturbations that are observed for the corresponding H2B residues between canonical and variant dimer (Figure S3.6). In particular for the glutamate a large upfield shift of the $^{1}H_{N}$ resonance is observed, from 10.12 ppm in the canonical dimer to 9.15 ppm in the variant dimer, supporting reduced hydrogen bonding.

For sake of completeness, we modelled the disordered tails to the structure of the folded core, highlighting the length of these extension compared to the globular core (Figure 3.4). Backbone dynamics derived from ^{15}N relaxation experiments support the highly flexible and unstructured nature of the tails compared to the relatively rigid folded-core (Figure S3.5).

Dai *et al.* highlighted the reduced electropositive charge at the DNA binding surface of the H2A.B-H2B core in their crystal structure[25], which is also apparent in solution (Figure 3.5a). In addition, a reduced negative potential is obvious for the region between the α2, α3 helix of H2A.B and αC of H2B, which forms the acidic patch on the canonical dimer (Figure 3.5b). Several charge neutralizing or charge

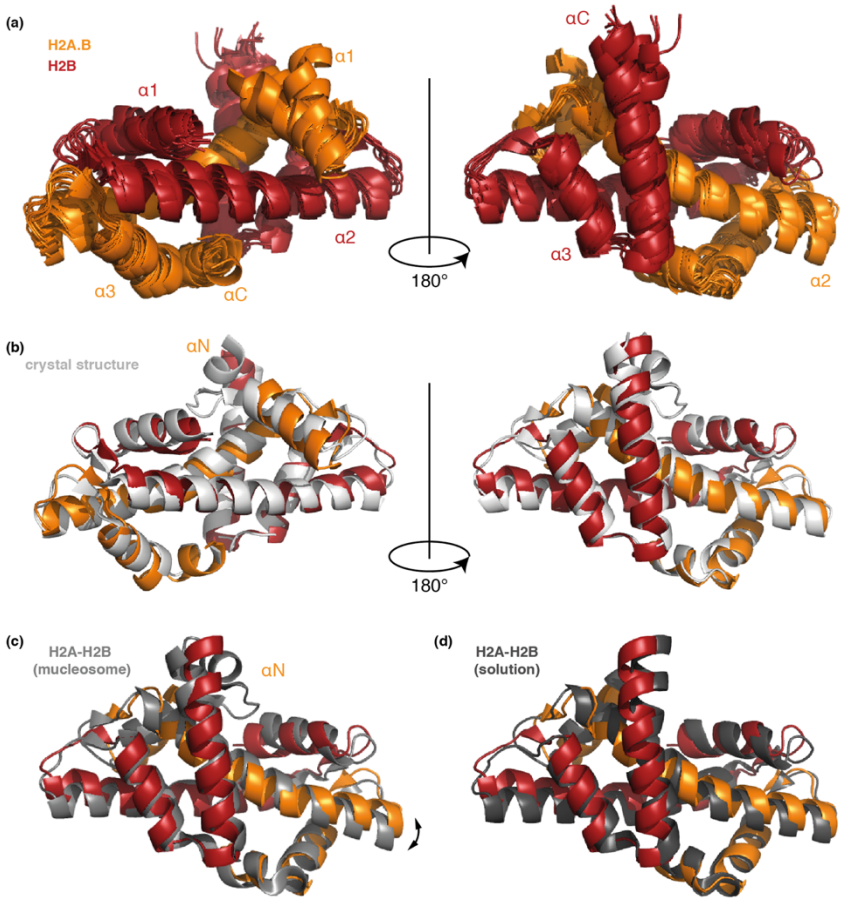

Figure 3.3. Solution structure of the H2A.B-H2B core. (a) Superposition of the final ensemble of 10 structural models calculated by NOE-supplemented CS-Rosetta. Color coding indicated in the figure; most prominent secondary structure elements are labeled. **(b)** Comparison of the H2A.B-H2B solution and crystal structure (PDB-id 6a7u). The bent H2A.B α2 helix is indicated. **(c)** Comparison of the variant dimer solution structure and the canonical H2A-H2B dimer from nucleosome crystal structure (PDB-id 3afa) highlighting the difference in H2A.B α2 orientation. **(d)** Comparison of the H2A.B-H2B and H2A-H2B solution structures (PDB-id 2rvq). Comparisons in (b), (c), and (d) were made using the closest-to-mean structure as the best representative structure of the final H2A.B-H2B ensemble. Structures were superimposed on heavy backbone atoms in the core region of the dimer H2A.B V31-L100 and H2B Y34-S121. Heavy atom backbone RMSDs for the superpositions shown in (b), (c) and (d) are 1.7, 1.8 and 1.9 Å, respectively.

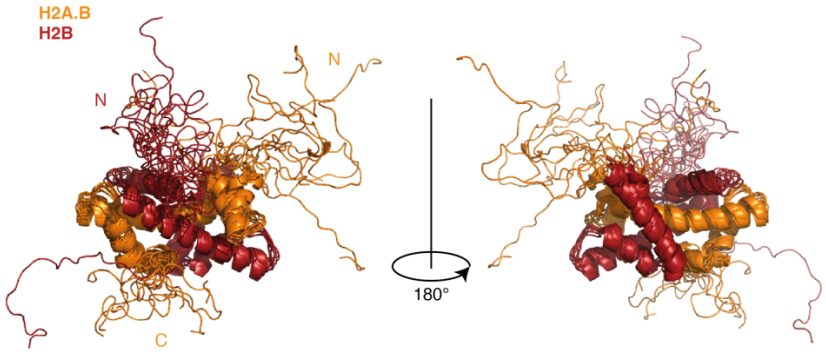

Figure 3.4. Solution structure of the H2A.B-H2B heterodimer including N- and C-terminal tails. Color coding as in Figure 3.3. The position of the N- and C-terminus is indicated.

reversal substitution in H2A.B (H2A E61, E91, E92 to H2A.B K65, R95, L96) significantly reduce the net charge of this patch. Closer inspection of the surface potential reveals that the H2A.B-H2B dimer exhibits a unique acidic surface on its H3-H4 interacting surface, which is positively charged in the canonical dimer. The change to a net negative charge is the result of substitution of Q84 and R88 in the H2A α3 helix with D88 and H92 in H2A.B. Within the canonical octamer, H2A R88 is involved in a long-range electrostatic interaction with H3 E94 and E105 on the adjacent H3-H4 dimer (Figure 3.5c). Within an H2A.B-nucleosome, this electrostatic interaction would either be neutralized or transformed into a repulsive interaction, depending on the charge-state of H92 (Figure 3.5d). These changes may thus contribute to the destabilization of the H2A.B containing octamers and nucleosomes. Previous studies using H2A and H2A.B domain-swapped proteins showed that a H2A chimera including the H2A.B α3 helix and the rest of the docking domain does not allow refolding of the histone octamer using salt dialysis[12].

Decreased net positive charge of the H2A.B-H2B core results in higher thermostability. We established that the H2A.B-H2B dimer retains a well-folded core in solution and has distinctly altered electrostatics compared to the canonical dimer. Such changes in surface electrostatics have been shown to impact protein stability even

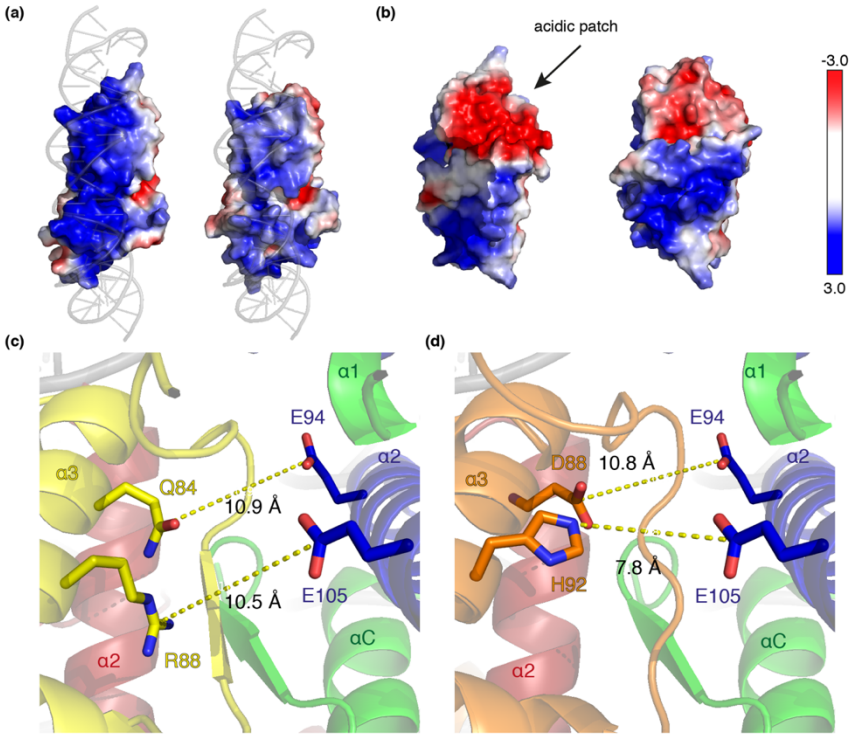

Figure 3.5. Electrostatic surface of H2A.B-H2B dimer compared to canonical H2A-H2B dimer. (a, b) View on the DNA binding surface (a) and acidic patch region (b) of H2A-H2B (left) and the H2A.B-H2B solution structure (right) with colour coding according to the electrostatic potential highlighting the reduced positive charge on the variant. The nucleosomal DNA is traced in light grey. Colour bar indicates the electro potential scale. **(c)** Zoom on the canonical nucleosome structure (PDB-id 3afa) focussing on an attractive electrostatic interaction in the interface between H2A-H2B and H3-H4. Selected side chains are shown as sticks. Colour coding: H2A – yellow; H2B – red; H3 – blue; H4 – green. **(d)** Same view as in (c) on a model of a H2A.B-containing nucleosome. H2A.B shown in orange. The docking domain C-terminal extension of H2A.B was modelled in the nucleosomal conformation of canonical H2A.

when these do not affect packing of the hydrophobic core[39]. Given that the structure and packing of the hydrophobic core of the variant dimer is highly similar to the canonical dimer, we hypothesized that the altered electrostatics may impact the stability of the dimer. Variant and canonical heterodimer stability was assessed by measuring their unfolding temperatures in a thermostability assay (TSA). Over a wide range of buffer conditions with varying pH and ionic strengths, we

observed that the variant dimer unfolds at significantly elevated temperatures compared to the canonical dimer (Figure 3.6a,b). H2A.B-H2B dimer unfolds at around 63 °C and H2A-H2B dimer unfolds at around 50 °C. While both types of dimers prefer higher salt concentrations (250 mM NaCl) over lower salt concentration (50 mM NaCl), H2A-H2B shows a stronger salt-induced stabilization effect than H2A.B-H2B. The canonical H2A-H2B dimer is the most stable in pH 7 buffers while the variant dimer prefers pH 6. For canonical H2A-H2B the melting temperatures and their buffer dependence compare well to previous data obtained from dimers isolated from chicken erythrocytes[40].

As is apparent from the electrostatic potentials displayed in Figure 3.5a, the core region of H2A.B carries overall fewer positively charged residues than H2A (7 vs. 12, respectively). The net charge of the core region is –1 for H2A.B and +4 for H2A, while the net charge for H2B is +6. This suggests that electrostatic repulsion between canonical H2A and H2B histone cores may result in a lower dimer thermostability compared to the variant in which there is slight overall electrostatic attraction.

To verify this hypothesis, we performed TSA assays on several H2A-H2B and H2A.B-H2B mutant dimers in which charged residues in H2A were substituted for the corresponding residues in H2A.B or vice versa. Two sites of charge reversal were selected to reduce the net charge of H2A maximally by a single mutation. Dimers refolded from H2A mutant K35E and K74E with net charge of +2 showed a significant increase in thermostability, which rescued half of the difference in melting temperatures between the canonical and variant dimers (Figure 3.6c). Dimers refolded from the reciprocal H2A.B mutants E40K and E79K with net charge +1 showed reduced thermostability compared to H2A.B-H2B dimer, thus qualitatively confirming our expectations. An H2A H30E mutant protein showed a melting temperature similar to that of wild-type protein, likely because at pH of the measurements (pH 6.5) the histidine is not fully protonated. Increase of the net positive charge of the canonical dimer by removal of glutamic acid E90 or E91 in the acidic patch H2A, either by substitution to arginine (E90R, net charge +6) or to leucine (E91L, net charge +5) from H2A.B, results in decreased melting temperatures. Overall these data reveal a clear trend of increasing unfolding

temperature with decreasing net charge of the dimer core. We thus conclude that the reduced net charge of H2A.B core and the resulting reduced electrostatic repulsion with H2B is responsible for the increased thermostability of the variant dimer.

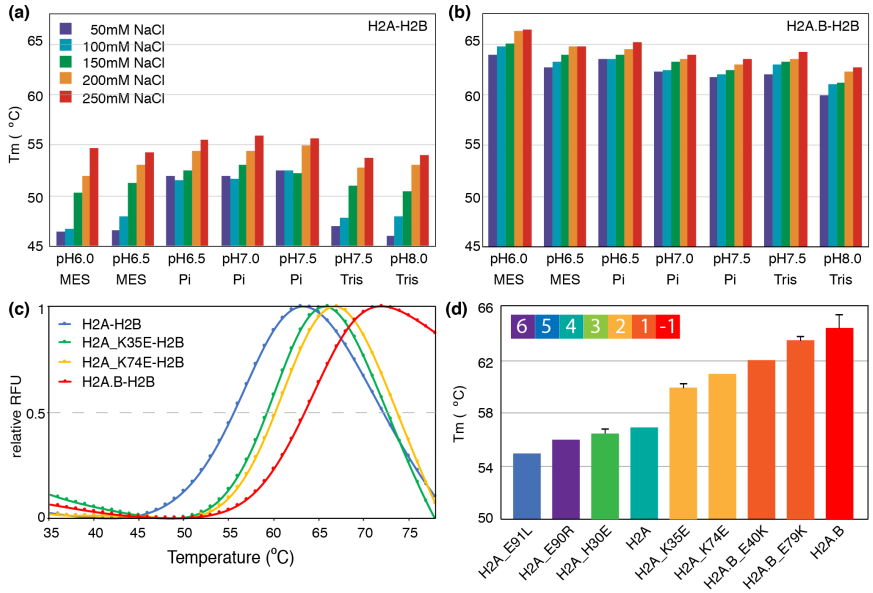

Figure 3.6. Thermostability of variant H2A.B-H2B and canonical H2A-H2B heterodimer. (a, b) Melting temperatures of H2A-H2B dimer (a) and H2A.B-H2B (b) across various pH and ionic strength conditions. Colour coding indicated. **(c)** Melting curves of H2A.B-H2B, H2A-H2B and selected mutant dimers as measured in thermal shift assay. Colour coding indicated. **(d)** Melting temperatures of H2A-H2B and H2A.B-H2B mutant dimers. The type of H2A protein incorporated in the dimer is shown on the x-axis. The net charge of the core region of H2A histones is indicated. Error bars based on duplicate experiments. Melting curves for (c, d) were measured in 20 mM NaPi buffer at pH 6.5 with 200 mM NaCl.

Discussion

Histone variant H2A.B is one of the most divergent histone variants that have evolved. Despite the specialized function and distinct properties of H2A.B-containing chromatin and nucleosomes, our understanding of the underlying molecular basis is yet limited. High resolution structural information on H2A.B-containing nucleosomes is

lacking, likely due to the destabilizing effect of H2A.B on the interactions to DNA and H3-H4. Here, we used solution-based method to study the structure, dynamics and stability of the H2A.B-H2B dimer. We find that H2A.B-H2B is structurally highly comparable to the free canonical dimer, yet with a markedly higher thermostability. Mutational analysis showed that the stability is correlated to the net charge of the dimer core, strongly suggesting that reduced electrostatic repulsion between the histone-fold domains of H2A.B and H2B is responsible for its increased thermostability. Since the change in charge is conserved across most of the 15 reported H2A.B sequences, including the key H30E, K35E, E91L and K74E substitutions, it is likely that the stability differences between canonical and variant dimers observed in this study are conserved across the H2A.B family.

To obtain the solution structure of H2A.B-H2B heterodimer we made use of a CS-Rosetta[41] driven approach, similar to the structure determination of the canonical H2A-H2B dimer[26]. Since addition of sparse data, i.e. either from RDC and/or NOE data[34,42,43] or evolutionary sequence information[44] can critically improve performance, we here included intermolecular NOEs as additional constraints in the calculations. This resulted in an overall better defined structural ensemble compared to using backbone chemical shift data only, in particular with respect to the position of the H2B αC helix. For the canonical H2A-H2B structure that was based on chemical shift data only, variable positions of the H2B αC helix and H2A α1 helix were taken as sign of intrinsic flexibility[26]. To what extent inclusion of NOE data would improve the definition of helices position for the canonical dimer remains to be established.

The solution structure of the isolated H2A.B-H2B heterodimer features a well-defined histone-fold core with disordered N- and C-terminal histones tails. The H2A.B αN-helix and the C-terminal extension of the docking domain are unfolded and likely only fold upon incorporation into a nucleosome. The chemical shift data do now show indications of transient folding for these regions, pointing to disorder-to-order transition upon binding to either DNA or H3-H4 in the nucleosome. Notably, these regions, in particular the αN helix, likely retain increased flexibility within the nucleosome, as was recently shown for canonical H2A in a solid-state NMR investigation of sedimented nucleosomes[45]. Furthermore, the L2 loop arginine (R80) in

H2A.B is shifted by one residue compared to H2A (see the alignment in Figure 3.1a). Since the H2A R76 at L2 loop inserts DNA minor groove at SHL 5.5[23], a change in its position could directly influences DNA opening. From our structure it would be expected that this arginine clashes with the nucleosomal DNA, thus potentially forcing the DNA in a more open conformation. Whether the L2 loop and thus the R80 position, is maintained within the H2A.B-nucleosome, or will adapt to the nucleosomal DNA remains to be explored. Finally, we have highlighted here the increased negative charge in the H2A.B α3 helix that faces an acidic surface in the H3-H4 tetramer in the context of the histone octamer. The resulting electrostatic repulsion could further contribute to destabilization of H2A.B containing nucleosome and to the failure in refolding the H2A.B containing histone octamer by salt-dialysis[27].

We have found that the isolated H2A.B-H2B dimer has significantly higher thermostability than the canonical dimer (13 °C degree higher Tm). For most proteins a higher thermostability can be interpreted as a sign of overall increased thermodynamic stability[46]. Indeed, stability measurements of the canonical dimer at constant temperature (25 °C) have been in reasonable agreement with thermal melting data[40,47]. This thus further supports an increased thermodynamic stability of the H2A.B-H2B dimer. Strikingly, the H2A.B-containing nucleosome was recently shown to also have a slightly higher thermostability (5 °C higher Tm) compared to the canonical nucleosome[48]. Here, it seems unlikely to directly translate to increase stability at ambient temperature given the more complex nature of this assembly and the extensive experimental evidence for reduced intermolecular interactions within the H2A.B-nucleosome.[12,13,14,19,25] We thus hypothesize that the relative free energy of H2A.B-nucleosome is decreased on one hand by the reduced interaction to DNA and H3-H4, and on the other hand by the increased stability of the free H2A.B-H2B dimer. Our results suggest that these effects are at least in part interrelated since both originate from the lower positive charge of H2A.B. The net effect of these changes is to promote a shift in equilibrium towards dissociated, isolated H2A.B-H2B dimers and thus to promote opening up of the nucleosome and chromatin. Indeed, the free histone pool contains a relatively large amount of highly diffusive isolated H2A.B-H2B dimers[21,19]. In addition, these free energy changes

likely also result in decreased energy barriers between states, facilitating increased exchange dynamics from nucleosomes to the free histone pool[17,49].
In summary, we presented the solution structure of the H2A variant H2A.B-H2B dimer, highlighting extensive changes in its surface electrostatics within an otherwise regular histone-fold, and showed that H2A.B-H2B has higher thermostability than the canonical dimer. This structure provides an important building block for a further detailed structural characterization of the assembly and exchange of the H2A.B nucleosome and provides new insights in its impact on chromatin structure and function.

Accession codes

The ensemble of solution structures is deposited in the Protein Data Bank under accession codes 6SMM. The NMR chemical shift assignment and NOE distance restraints have been deposited in the Biological Magnetic Resonance Data Bank under accession code 34426. Proteins used in this study are *Dm.* H2A (Uniprot-id: P84051) and *Dm.* H2B (Uniprot-id: P02283) and *Hs.* H2A.B (Uniprot: P0C5Z0).

Acknowledgement

This research was financially supported by The Dutch Research Council NWO (VIDI grant 723.013.010 to HvI). We thank Dr. Karolin Luger (University of Colorado) for providing the H2A.B plasmid, Alma Svatos for her help in the chemical shift assignment, Dr. Jorgen Schaarschmidt and prof. Alexandre Bonvin for help in setting up the Rosetta calculations, Bart van den Berg, Renske Snoeks, Kerven Dammers and Kim Fernandez for their help in making some of the mutants used in this study, Anneloes Blok and Mark Daniels for technical assistance, and the NMR group members for their stimulating support.

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. Singh, R.; Bassett, E.; Chakravarti, A.; Parthun, M. R., Replication-dependent histone isoforms: a new source of complexity in chromatin structure and function. *Nucleic Acids Res* **2018,** *46* (17), 8665-8678.

3. 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**.

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

5. Cheema, M. S.; Ausio, J., The Structural Determinants behind the Epigenetic Role of Histone Variants. *Genes (Basel)* **2015,** *6* (3), 685-713.

6. Zink, L. M.; Hake, S. B., Histone variants: nuclear function and disease. *Curr Opin Genet Dev* **2016,** *37*, 82-89.

7. Palmer, D. K.; O'Day, K.; Wener, M. H.; Andrews, B. S.; Margolis, R. L., A 17-kD centromere protein (CENP-A) copurifies with nucleosome core particles and with histones. *J Cell Biol* **1987,** *104* (4), 805-15.

8. Rogakou, E. P.; Pilch, D. R.; Orr, A. H.; Ivanova, V. S.; Bonner, W. M., DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139. *J Biol Chem* **1998,** *273* (10), 5858-68.

9. 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.

10. Gonzalez-Romero, R.; Mendez, J.; Ausio, J.; Eirin-Lopez, J. M., Quickly evolving histones, nucleosome stability and chromatin folding: all about histone H2A.Bbd. *Gene* **2008,** *413* (1-2), 1-7.

11. 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.

12. 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.

13. 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.

14. 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.

15. 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.

16. 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.

17. 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.

18. 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.

19. 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.

20. 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.

21. 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.

22. McGinty, R. K.; Tan, S., Nucleosome structure and function. *Chem Rev* **2015,** *115* (6), 2255-73.

23. 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.

24. 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.

25. 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.

26. Moriwaki, Y.; Yamane, T.; Ohtomo, H.; Ikeguchi, M.; Kurita, J.; Sato, M.; Nagadoi, A.; Shimojo, H.; Nishimura, Y., Solution structure of the isolated histone H2A-H2B heterodimer. *Sci Rep* **2016,** *6*, 24999.

27. Luger, K.; Rechsteiner, T. J.; Richmond, T. J., Preparation of nucleosome core particle from recombinant histones. *Methods Enzymol* **1999,** *304*, 3-19.

28. Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A., NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* **1995,** *6* (3), 277-93.

29. Lee, W.; Tonelli, M.; Markley, J. L., NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. *Bioinformatics* **2015,** *31* (8), 1325-7.

30. Shen, Y.; Bax, A., Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *J Biomol NMR* **2013,** *56* (3), 227-41.

31. Shen, Y.; Vernon, R.; Baker, D.; Bax, A., De novo protein structure generation from incomplete chemical shift assignments. *J Biomol NMR* **2009,** *43* (2), 63-78.

32. Shen, Y.; Bryan, P. N.; He, Y.; Orban, J.; Baker, D.; Bax, A., De novo structure generation using chemical shifts for proteins with high-sequence identity but different folds. *Protein Sci* **2010,** *19* (2), 349-56.

33. Guntert, P.; Buchner, L., Combined automated NOE assignment and structure calculation with CYANA. *J Biomol NMR* **2015,** *62* (4), 453-71.

34. Lange, O. F.; Rossi, P.; Sgourakis, N. G.; Song, Y.; Lee, H. W.; Aramini, J. M.; Ertekin, A.; Xiao, R.; Acton, T. B.; Montelione, G. T.; Baker, D., Determination of

solution structures of proteins up to 40 kDa using CS-Rosetta with sparse NMR data from deuterated samples. *Proc Natl Acad Sci U S A* **2012,** *109* (27), 10873-8.

35. Shen, Y.; Bax, A., SPARTA+: a modest improvement in empirical NMR chemical shift prediction by means of an artificial neural network. *J Biomol NMR* **2010,** *48* (1), 13-22.

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

37. Dolinsky, T. J.; Nielsen, J. E.; McCammon, J. A.; Baker, N. A., PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res* **2004,** *32* (Web Server issue), W665-7.

38. Tachiwana, H.; Kagawa, W.; Osakabe, A.; Kawaguchi, K.; Shiga, T.; Hayashi-Takanaka, Y.; Kimura, H.; Kurumizaka, H., Structural basis of instability of the nucleosome containing a testis-specific histone variant, human H3T. *Proc Natl Acad Sci U S A* **2010,** *107* (23), 10454-9.

39. 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.

40. Karantza, V.; Baxevanis, A. D.; Freire, E.; Moudrianakis, E. N., Thermodynamic studies of the core histones: ionic strength and pH dependence of H2A-H2B dimer stability. *Biochemistry* **1995,** *34* (17), 5988-96.

41. Shen, Y.; Lange, O.; Delaglio, F.; Rossi, P.; Aramini, J. M.; Liu, G.; Eletsky, A.; Wu, Y.; Singarapu, K. K.; Lemak, A.; Ignatchenko, A.; Arrowsmith, C. H.; Szyperski, T.; Montelione, G. T.; Baker, D.; Bax, A., Consistent blind protein structure generation from NMR chemical shift data. *Proc Natl Acad Sci U S A* **2008,** *105* (12), 4685-90.

42. Raman, S.; Lange, O. F.; Rossi, P.; Tyka, M.; Wang, X.; Aramini, J.; Liu, G.; Ramelot, T. A.; Eletsky, A.; Szyperski, T.; Kennedy, M. A.; Prestegard, J.; Montelione, G. T.; Baker, D., NMR structure determination for larger proteins using backbone-only data. *Science* **2010,** *327* (5968), 1014-8.

43. ElGamacy, M.; Riss, M.; Zhu, H.; Truffault, V.; Coles, M., Mapping Local Conformational Landscapes of Proteins in Solution. *Structure* **2019,** *27* (5), 853-865 e5.

44. Tang, Y.; Huang, Y. J.; Hopf, T. A.; Sander, C.; Marks, D. S.; Montelione, G. T., Protein structure determination by combining sparse NMR data with evolutionary couplings. *Nat Methods* **2015,** *12* (8), 751-4.

45. 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.

46. Rees, D. C.; Robertson, A. D., Some thermodynamic implications for the thermostability of proteins. *Protein Sci* **2001,** *10* (6), 1187-94.

47. Gloss, L. M.; Placek, B. J., The effect of salts on the stability of the H2A-H2B histone dimer. *Biochemistry* **2002,** *41* (50), 14951-9.

48. 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.

49. 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.

Supplements

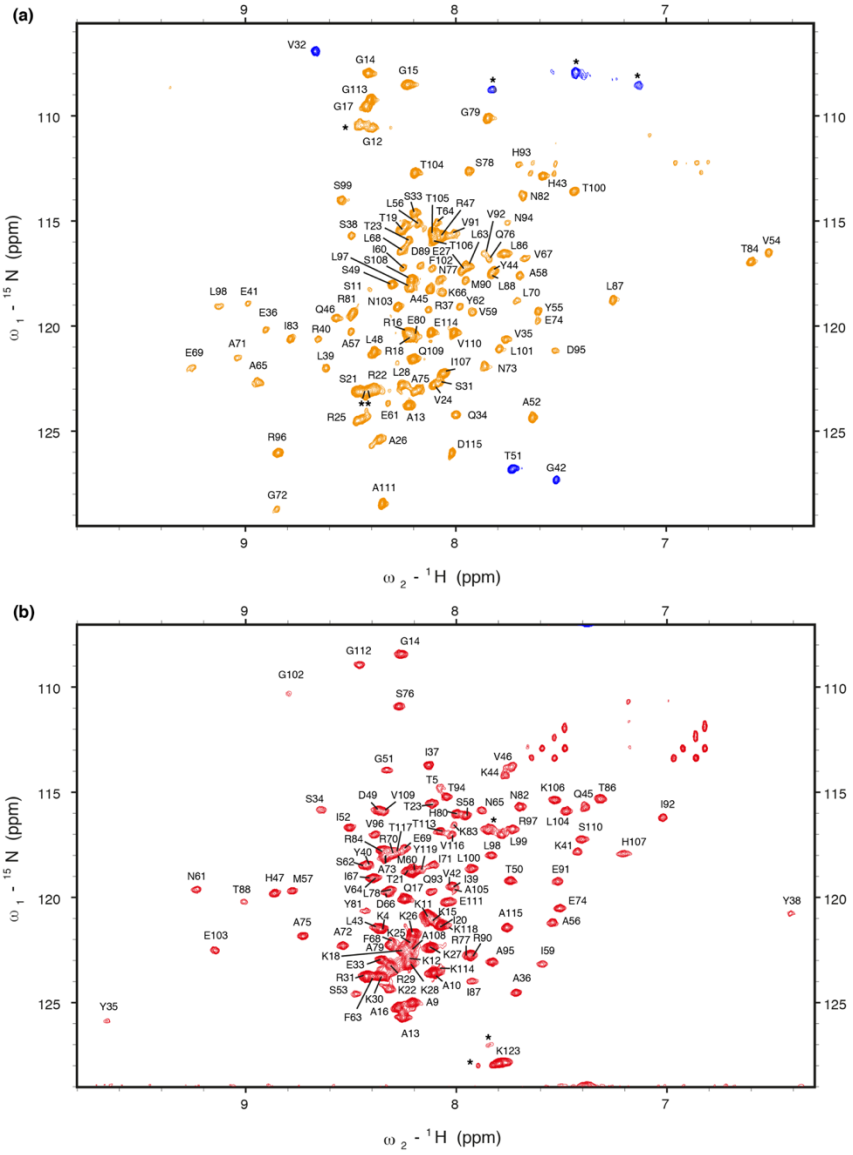

Figure S3.1. 1H-15N TROSY spectra of (a) H2A.B and (b) H2B within the H2A.B-H2B heterodimer. Assignments are indicated, unassigned backbone resonances are indicated with a *. Negative folded peaks are shown in blue.

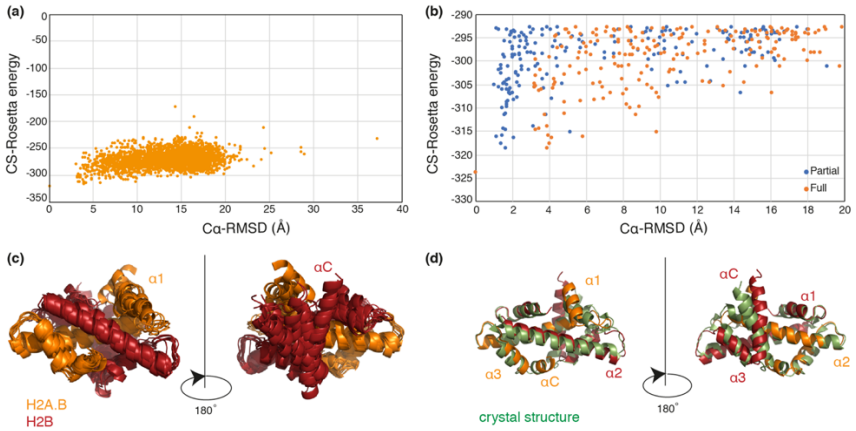

Figure S3.2. CS-Rosetta solution structure of H2A.B-H2B. (a,b) CS-Rosetta energy plotted versus Cα-RMSD to the lowest energy structure. In (a) Cα-RMSD was calculated over the complete histone core shown for all 3000 calculated solutions. In (b) a zoom of the plot in (a) is shown with Cα-RMSD calculated over the complete histone core in orange, and RMSD excluding the H2B αC and H2A.B α1 helices in blue. **(c)** Superposition of the 10 CS-Rosetta models with the lowest C_α-RMSD from the lowest energy 20 models, showing ill-defined positions of H2B αC and H2A.B α1-helices within an otherwise converged structure. **(d)** Structural comparison between lowest-energy CS-Rosetta model and the H2A.B-H2B crystal structure (PDB-id 6a7u).

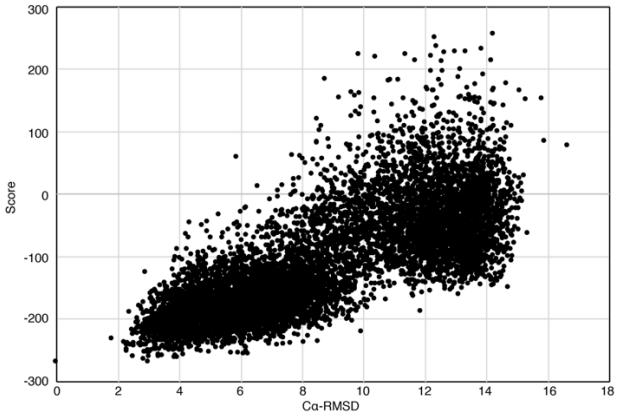

Figure S3.3. Final re-scored Rosetta energy including both chemical shift and NOE correspondence plotted versus Cα-RMSD to the lowest-energy structure of the H2A.B-H2B core. In total 8000 solutions were calculated. The RMSD is calculated over the complete histone-core.

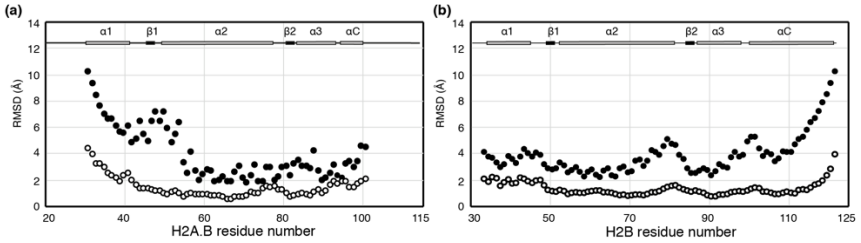

Figure S3.4. Impact of inclusion of intermolecular NOE data on CS-Rosetta calculated structure. (a, b) Average per-residue heavy atom backbone RMSD, averaged over the 10 models in either the final ensemble calculated without/with (filled/blank circles) NOE restraints. Average per-residues RMSDs are plotted for H2A.B (a) and H2B (b) core regions. Secondary structure elements in the final solution structure are plotted on top.

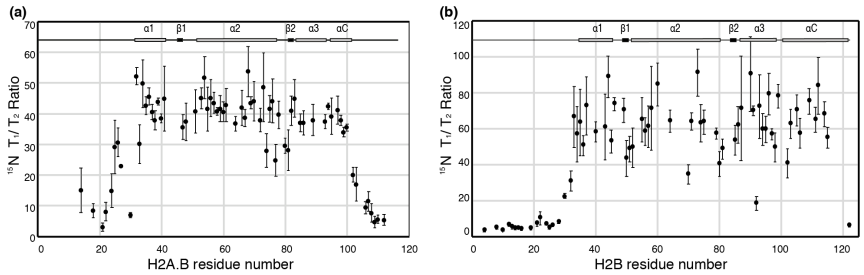

Figure S3.5. Backbone dynamics of H2A.B-H2B dimer. (a, b) Ratio of backbone amide ^{15}N-T_1 and ^{15}N-T_2 values for H2A.B (a) and H2B (b), illustrating the high flexibility of the H2A.B N- and C-terminal tails, as well as the H2B N-terminal tail. The average T_1/T_2-ratio for residues the folded core of H2B are elevated compared to those in the folded parts of H2A.B, most likely due the higher salt concentration in the sample used for measuring H2B dynamics. Secondary structure elements in the final solution structure are plotted on top.

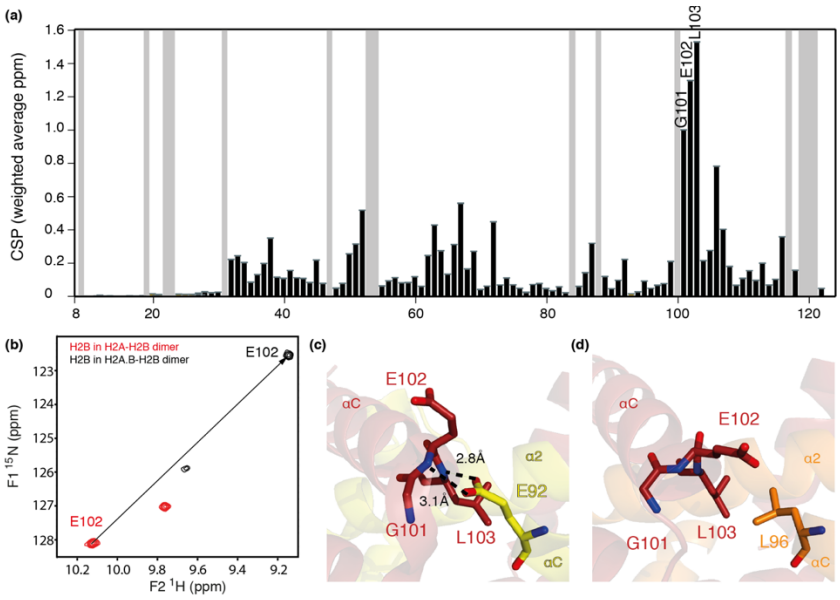

Figure S3.6. Chemical shift perturbations of H2B residues in canonical and variant dimers caused by sequence variance between H2A and H2A.B. (a) Chemical shift perturbations (CSPs) for H2B residues between two canonical and variant dimers. Unassigned residues shown in grey. Residues with CSP larger than 1ppm are labeled. **(b)** Overlay of H2B TROSY spectra within canonical H2A-H2B dimer (red) and variant H2A.B-H2B dimer (black), zooming in on the resonance of E102 in H2B. **(c,d)** Comparison of solution structure of H2A.B-H2B (d) and crystal structure of H2A-H2B (c) zooming in on the loss of a hydrogen-bond due to substitution of H2A E92 to H2A.B L96.

Table S3.1. Structural statistics for the core region of isolated H2A.B-H2B heterodimer

A. Restraint information	
number of intermolecular NOEs	44
B. Average RMS deviation from experimental restraints	
All experimental distance restraints (Å)	0.08 ± 0.24
C. Coordinate RMS deviation[a] (Å)	
Average RMSD to mean	
Ordered backbone atoms	1.08 ± 0.24
Ordered heavy atoms	1.54 ± 0.25
Global backbone atoms	1.48 ± 0.25
Global all heavy atoms	1.92 ± 0.33
Average Pairwise RMSD	
Ordered backbone atoms	1.61 ± 0.35
Ordered all heavy atoms	2.30 ± 0.36
Global backbone atoms	2.21± 0.38
Global all heavy atoms	2.91 ± 0.41

[a] Ordered regions correspond to residues V31-F101 of H2A.B and Y34-K122 of H2B.