Supplementary Materials for

Histone H2B ubiquitylation disrupts local and higher order chromatin compaction

Beat Fierz[1], Champak Chatterjee[1,3], Robert K. McGinty[1], Maya Bar-Dagan[1],

Daniel P. Raleigh[2] and Tom W. Muir[1,*]

[1]Laboratory of Synthetic Protein Chemistry, The Rockefeller University, 1230 York Avenue, New York, NY 10065

[2]Department of Chemistry, SUNY Stony Brook, Stony Brook, NY 11794-3400

[3]Present address: Department of Chemistry, University of Washington, Seattle, WA 98195-1700

[*]To whom correspondence should be addressed. E-mail: muirt@rockefeller.edu

Supplementary methods

Computational modeling of chromatin fibers

To understand how sedimentation coefficient and homo-RET are related, we implemented a nucleosomal array model [1] in Matlab (The MathWorks, Natick, MA). From such a three dimensional model of a chromatin fibre, SSA [2] and $S_{20,w}$ values can be calculated. A single nucleosome was approximated as a sphere with 4.5 nm radius from which the DNA strands project with an exit angle α. The DNA linking two nucleosomes was twisted by a torsion angle β (**Supplementary Fig. 6a**). Chromophores were placed at both faces of the nucleosome at a distance of 2.7 nm from the center as estimated from the nucleosome structure [3]. A DNA linker length of 10 nm was assumed to produce chromatin chains with different conformations dependent on the values chosen for α and β. The basic parameters and the configuration of compact chromatin (α = 41° and β = 77°) were chosen to approximate a model for the 30 nm fiber as a two-start helix (**Supplementary Fig. 6b**), based on the structure of a tetranucleosome [4] and in vitro experiments [5]. To generate fully extended chains, randomized torsion angles around 100° and exit-angles around 120° were chosen and a large ensemble of structures was generated. Chains of intermediate compaction were generated using angles from either the fully folded or the fully unfolded ensemble or linear combination of both for each nucleosome in the chain, depending on the folding model (see below). From these generated structures the sedimentation coefficient, $S_{20,w}$, was approximated using the formula [6-8]

$$S_{20,w} = S_1\left(1+\frac{2R}{N}\sum_{i}^{N}\sum_{j>i}^{N}\frac{1}{R_{ij}}\right)$$

where S_1 represents the sedimentation coefficient of the mono-nucleosome (11.1 S) [9], R is its hydrodynamic radius (5.46 nm) [8], N is the number of nucleosomes in the array (12) and R_{ij} are the distances between nucleosomes i and j. For simplicity, the contribution from linker DNA to $S_{20,w}$ was neglected.

The total emission anisotropy from all interacting chromophores was determined by solving the rate equations for energy transfer [2,10]:

$$\frac{d\rho_j(t)}{dt} = -\frac{\rho_j(t)}{\tau} - \sum_{k\neq j}^{N} F_{jk}\cdot\rho_j(t) + \sum_{m\neq j}^{N} F_{mj}\cdot\rho_m(t)$$

where $\rho_j(t)$ is the probability that the molecule j is excited at time t, F_{ij} denote the Förster transfer rates which depend on the intermolecular distances between chromophore *i* and *j*

$$F_{ij} = \frac{1}{\tau}\left(\frac{R_0}{R_{ij}}\right)^6$$

using R_0 = 4.73 nm (the Förster radius, **Supplementary Fig. 5e**), and τ is the average fluorescence lifetime of the excited chromophore (4.2 ns for fluorescein coupled to chromatin arrays, see **Supplementary Fig. 5f**).

The quantum efficiency of emission from the initially excited molecule is then given by

$$\phi_j = 1/\tau_0 \cdot \int_0^{\infty} \rho_j(t)dt$$

where τ_0 is the natural lifetime of the chromophore. Assuming full depolarization after one energy transfer event the total anisotropy can be calculated by averaging over all chromophores

$$r_{tot} = 1/N \cdot \sum_{i=1}^{N} r_0 \frac{\phi_i}{\phi_{tot}}; \quad \phi_{tot} = \tau / \tau_0$$

using r_0 = 0.236 as the anisotropy of a single fluorescein chromophore bound to nucleosomes in an array (**Supplementary Fig. 4b**). The differential equations were solved for the different fiber models applying a matrix approach as outlined in ref. 2,10.

We tested two models for compaction, compatible with measured $S_{20,w}$ values: 1) A *homogenous* compaction model, where throughout the folding process internucleosomal distances are gradually reduced and 2) a local interaction or *heterogenous* model, where an increasing number of nucleosomes i and i+2 form an interaction resulting in nucleosome stacking, until the fully compacted fiber is formed (**Supplementary Fig. 6c**). For structures of the homogenous compaction model, linear combinations of angles of fully folded or unfolded chains were used. In contrast, for the local interaction model randomly picked nucleosomes were assigned fully folded angle values producing local clusters of stacked nucleosomes. These clusters were subsequently extended to form fully folded fibers. For each folding mechanism, 500 different oligo-nucleosome conformations were generated in different compaction states and their $S_{20,w}$ and SSA were calculated. Using the measured dependence of $S_{20,w}$ on [Mg^{2+}] (**Fig. 1c**), calculated $S_{20,w}$ values were mapped to magnesium concentrations and thus SSA traces as a function of [Mg^{2+}] were generated (**Fig. 3d**). Due to *i)* the increased number of fluorophores within $2*R_0$ of each other, i.e. within range for RET (**Supplementary Fig. 6d**), and *ii)* the R^6 dependence of RET efficiency these two folding pathways resulted in very different SSA values for

conformations with similar $S_{20,w}$ values. Based on the measured SSA values, the homogenous compaction model for chromatin folding can be readily excluded. The *heterogenous* model however reproduced the general shape of SSA loss upon compaction (**Fig. 3d**).

General laboratory methods

Amino acid derivatives, coupling reagents and resins were purchased from Novabiochem (Merck, NJ). All commonly used chemical reagents and solvents were purchased from Sigma-Aldrich Chemical Company (Milwaukee, WI) or Fischer Scientific (Pittsburgh, PA). [^{3}H]-S-adenosyl methionine, Amplify solution and Sephacryl S-200 resin were obtained from GE Healthcare (Waukesha, WI). Chemically competent DH5α and BL21(DE3)pLysS cells were purchased from Novagen (Madison, WI). T4 DNA ligase and restriction enzymes were obtained from New England BioLabs (Ipswitch, MA). Primer synthesis and gene sequencing were performed by Integrated DNA Technologies (Coralville, IA) and Genewiz (South Plainfeld, NJ), respectively. Gene mutagenesis was achieved using a QuikChange Site-Directed Mutagenesis kit (Stratagene, La Jolla, CA). Criterion 15% Tris-HCl and 5% TBE gels were purchased from BioRad (Hercules, CA). Centrifugal filtration units were from Sartorius (Goettingen, Germany) and Slide-A-Lyzer dialysis cassettes and MINI dialysis units were from Pierce (Rockford, IL). Qiafilter Plasmid Giga, PCR purification and gel extraction kits were purchased from Qiagen (Valencia, CA). Antibodies for uH2B and H4 K16ac were purchased from Medimabs (Royalmount, Quebec, Canada) and Active Motif (Carlsbad, CA), respectively. Micrococcal nuclease was purchased from Worthington (Lakewood, NJ). Size exclusion

chromatography was performed on an AKTA FPLC system from GE Healthcare equipped with a P-920 pump and UPC-900 monitor. Analytical reversed-phase HPLC (RP-HPLC) was performed on a Hewlett-Packard 1100 series instrument with a Vydac C18 column (5 micron, 4 x 150 mm), employing 0.1% TFA in water (HPLC solvent A), and 90% acetonitrile, 0.1% TFA in water (HPLC solvent B), as the mobile phases. Typical analytical gradients were 30-70% solvent B over 30 min at a flow rate of 1 mL/min. Preparative scale purifications were conducted on a Waters DeltaPrep 4000 system equipped with a Waters 486 tunable detector. A Vydac C18 process column (15-20 micron, 50 x 250 mm) or a semi-preparative column (12 micron, 10 mm x 250 mm) was employed at a flow rate of 30 mL/min or 4 mL/min, respectively. ESI-MS analysis was conducted on a Sciex API-100 single quadruple spectrometer. LC-MS was performed at the Proteomics Resource Center, Rockefeller University, New York employing a Dionex U3000 capillary/nano-HPLC system (Dionex, Sunnyvale, CA) directly interfaced with a ThermoFisher LTQ-Orbitrap mass spectrometer (ThermoFisher, San Jose, CA). All protein starting materials and ligation products were analyzed by C18 analytical RP-HPLC and ESI-MS.

Analysis of native chromatin

NIH/3T3 fibroblasts were grown to near confluence and harvested by scraping. The cells were washed twice with phosphate buffered saline. In the second wash, protease inhibitors (Complete, EDTA-free Protease Inhibitor Cocktail, Roche Applied Science, Indianapolis, IN), 400 nM Trichostatin A (TSA, Sigma) and 20 mM NEM (N-ethylmaleimide, Sigma) were added. Cells were lysed following ref. 11 with changes as

described as follows. The cell pellet was resuspended in lysis buffer (10 mM HEPES pH 7, 10 KCl, 2 mM $MgCl_2$, 340 mM sucrose, 10% glycerol (v/v), 50 μg/ml bovine serum albumin (BSA), 1 mM TCEP (tris(2-carboxyethyl)phosphine), 20 mM NEM, Complete Protease Inhibitor Cocktail). An equal volume of lysis buffer containing 0.2% Triton X-100 was added, and the cells were lysed for 10 min on ice with gentle agitation. Nuclei were harvested by centrifugation (5 min, 1300 x g, 4º C), resuspended in lysis buffer and further purified through a sucrose cushion (10 mM HEPES pH 7, 30% sucrose (w/v), 2 mM $MgCl_2$, 50 μg/ml bovine serum albumin (BSA), 1 mM TCEP (tris(2-carboxyethyl)phosphine), 400 nM TSA, 20 mM NEM, Complete Protease Inhibitor Cocktail) by centrifugation for 10 min at 1300x g at 4º C [12]. The purified nuclei were re-suspended in lysis buffer without NEM at a concentration of 10^8 nuclei/ml. $CaCl_2$ was added to a concentration of 1 mM and the suspension was pre-warmed to 37º C for *in nucleo* digestion with micrococcal nuclease. 60 U micrococcal nuclease (Worthington) were added per ml suspension, followed by incubation at 37º C for 10 min. The digestion was stopped by addition of 5 mM ethylene glycol tetraacetic acid (EGTA) and placing on ice (total nuclei, fraction 1). Centrifugation at 1300x g allowed to separate supernatant containing mainly mononucleosomes (fraction 2, 13% of total chromatin) from digested nuclei. Chromatin was extracted following ref. 13 with some changes. The nuclei were resuspended in extraction buffer (10mM HEPES pH 7.5, 2 mM MgCl2, 2 mM EGTA, 0.1% Triton X-100, 400 nM TSA, 20 mM NEM, Complete Protease Inhibitor Cocktail) containing 80 mM or 150 mM NaCl. The suspensions were agitated for 2 h at 4º C, followed by centrifugation at 1300x g for 10 min. The supernatants containing fragments of active chromatin (fractions 3 and 4, 10% and 21% of total chromatin) were saved and

the pellets were further extracted with extraction buffer containing 600 mM NaCl. After agitation for 2 h at 4º C followed by centrifugation at 1300x g for 10 min, the supernatants containing the bulk of the chromatin (fraction 5, 60% of total chromatin) were separated from the pellets containing insoluble chromatin (fraction 6, 9% of total chromatin). DNA was extracted from all fractions using QIAQuick spin columns. Histones were obtained by acid extraction for 16 h with 0.4 N H_2SO_4 followed by ethanol precipitation. Precipitated proteins were resuspended in 5% acetic acid and lyophilized. The lyophilized powder was dissolved in ddH2O and the protein concentration was determined by BCA assay (Pierce, Rockford, IL). Equal amounts of histones were separated by 15% SDS-PAGE and analyzed by western blotting using antibodies against uH2B (Medimabs, Royalmount, Quebec, Canada) and H4 K16ac (Active Motif, Carlsbad, CA). Western blots were imaged using the Odyssey Infrared Imaging System (LI-COR Biotechnology, Lincoln, NE).

Mutagenesis of H2A(N110C)

Using the *Xenopus* H2A expression plasmid [14] as a template, the point mutation asparagine 110 to cysteine was generated by site-directed mutagenesis employing a QuikChange Site-Directed Mutagenesis kit, using the following primers: forward primer, 5'-CGGGGTCCTGCCCTGCATCCAGTCCGTG-3', reverse primer, 5'-CACGGACTGGATGCAGGGCAGGACCCCG-3'. The correct sequence was confirmed by gene sequencing.

Expression of recombinant histones

Recombinant *Xenopus* histones H2A, H2A(N110C), H2B, H3(C110A) [15] and H4 were expressed as previously described [14,16]. In short, E. coli BL21(DE3)pLysS cells were transfected with histone expression plasmids, grown in 6 L LB medium at 37°C until OD_{600} 0.6 and protein expression was induced by addition of 0.5 mM IPTG for 3 h. Cells were harvested by centrifugation at 7000 x g and lysed by 5 x passage through a French Press. The insoluble pellets were washed twice with 40 mL wash buffer (20 mM Tris pH 7.5, 200 mM NaCl, 1 mM EDTA, 1 mM 2-mercaptoethanol) containing 1% triton X-100, and once with wash buffer. The pellets were extracted with extraction buffer (7 M guanidinium hydrochloride, 20 mM Tris pH 7.5, 200 mM NaCl, 1 mM EDTA, 1 mM 2-mercaptoethanol) and purified over a Sephacryl S-200 column (1 L bed volume). After dialysis of the histone-containing fractions into water containing 1 mM DTT, the fractions were lyophilized. The histones were then further purified using preparative C-18 RP-HPLC and a gradient of 30-70% HPLC solvent B. The recombinant histones were analyzed by analytical HPLC and ESI-MS (H2B: calculated, $[M+H]^+$ 13,818 Da, observed 13,820±4 Da, H4: calculated, $[M+H]^+$ 11,237 Da, observed 11,238±4 Da, H2A: calculated, $[M+H]^+$ 13,951 Da, observed 13,952±4 Da, H2A(N110C): calculated, $[M+H]^+$ 13,940 Da, observed 13,941±4 Da, H3(C110A): calculated, $[M+H]^+$ 15,240 Da, observed 15,241±4 Da).

Labeling of H2A(N110C) with fluorescein-5-maleimide

In a typical reaction, 8.5 mg HPLC purified H2A(N110C) (0.6 μmol) was dissolved in 4 mL labeling buffer (20 mM Tris, pH 7.8, 6 M guanidinium hydrochloride, 200 μM

TCEP) and 0.7 mg fluorescein-5-maleimide (1.6 μmol, 2.7 equiv.) was added in 50 μl N,N-dimethylformamide (DMF) in the dark. The mixture was stirred for 1 h and quenched by addition of 1 mM 2-mercaptoethanol. Fluorescein labeled H2A (fH2A) was subsequently purified by using semi preparative C-18 RP-HPLC and a gradient of 30-70% HPLC solvent B. The protein was characterized by analytical HPLC and ESI-MS ($(M + H)^+$ observed: 11490±4 Da; expected 11488 Da) (**Supplementary Fig. 1a, b**).

Synthesis of $uH2B_{SS}$ and hub1-$H2B_{SS}$

$uH2B_{SS}$ hub1-$H2B_{SS}$ were produced according to the protocols in ref. 15. $uH2B_{SS}$ was characterized by ESI-MS ($(M + H)^+$ observed: 22420±5 Da; expected 22415 Da) and analytical HPLC (**Supplementary Fig. 2c, d**). Hub1-$H2B_{SS}$ was characterized by ESI-MS ($(M + H)^+$ observed: 22375±6 Da; expected 22376 Da) and analytical HPLC (**Supplementary Fig. 1g, h**).

Synthesis of acetylated H4 α-thioester peptide - acH4(1-37)-COSR

N^{α}-^{t}butoxycarbonyl-leucyl-3-mercaptopropionamide-MBHA resin was prepared according to ref. 17. The sequence corresponding to residues 1-37 of *Xenopus* H4 was synthesized on the afore mentioned resin on a 0.1 mmol scale with a ^{t}butoxycarbonyl (Boc) N^{α} protection strategy and using an in situ neutralization / 2-(1H-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU) amino-acid activation protocol [18]. The following side-chain protection groups were used: Arg(Tos), Lys(2-Cl-Z), Asn(Xan), Asp(OcHx) and His(Bom), acetylated lysines (residues 5, 8, 12, 16 and 20) were introduced as Boc-Lys(Ac)-OH. After the synthesis was completed the N-terminus was acety-

lated by treatment with a mixture of acetic anhydride/DIEA/DMF 15/15/70 for 10 min. Cleavage off the resin was achieved by treatment with anhydrous hydrogen fluoride in the presence of *p*-cresol. Subsequently, 40 mg of the crude acH4(1-37)-COSR thioester peptide were purified by RP-HPLC on a preparative scale using a 20–45% B gradient over 60 min, yielding 10 mg pure peptide. The acH4(1-37)-COSR thioester peptide was characterized by ESI-MS ($(M + H)^+$ observed: 4279±2 Da; expected: 4278 Da)

Preparation of recombinant H4 N-terminal fragment – H4(37-102, A38C)

Using the *Xenopus* H4 expression plasmid [14] as a template, a H4 gene was constructed, containing an internal his-tag at position 19 in the tail region and a tobacco etch virus (TEV) protease cleavage site (ENLYFQC) preceding arginine 39, which will generate an N-terminal cysteine residue upon cleavage with TEV protease. This was achieved by two successive site-directed mutagenesis steps employing the QuikChange Site-Directed Mutagenesis kit as per manufacturer's protocols and mutagenic primers as follows: Internal his-tag: forward primer, 5'-GCTAAACGTCACCATCACCATCACCACCGTAAAGTTCTGC-3', reverse primer 5'- GCAGAACTTTACGGTGGTGATGGTGATGGTGACGTTTAGC-3'. TEV cleavage site: forward primer, 5'- CGTCGTCTGGCTGAAAACCTGTACTTCCAGTGCCGTCGTGGTGGT-3', reverse primer, 5'- ACCACCACGACGGCACTGGAAGTACAGGTTTTCAGCCAGACGACG-3'.

For protein expression, E. coli BL21(DE3)pLysS cells which were transformed with the afore mentioned H4 plasmid were grown in 6 L LB medium at 37°C until OD_{600} 0.6 at which time protein expression was induced by the addition of 0.5 mM IPTG for 2h. Cells were harvested by centrifugation and lysed by passage through a French Press,

followed by short sonication. Pelleted inclusion bodies were washed twice with wash buffer (20 mM Tris, 200 mM NaCl, 1 mM EDTA, 1 mM 2-mercaptoethanol, pH 7.5) containing 1% triton X-100 and once with wash buffer (20 mM Tris, 200, mM NaCl, 1 mM EDTA, 1 mM 2-mercaptoethanol, 1% triton, pH 7.5). Inclusion bodies were extracted with extraction buffer (6 M guanidinium chloride, 20 mM Tris pH 7.5 and 1 mM 2-mercaptoethanol) and applied to Ni:NTA resin. The resin was washed with 10 column volumes extraction buffer containing 25 mM imidazole and the protein was eluted with 3 x 1 column volume extraction buffer containing 400 mM imidazole. The protein concentration was adjusted to 1 mg / mL and the solution was dialyzed into water containing 1 mM DTT. For cleavage with TEV protease, 10x cleavage buffer was added to reach final concentrations of 50 mM Tris, pH 6.9, 1 mM EDTA, 10 mM DTT and 10 mM cysteine. Approximately 0.4 mg TEV protease, expressed using the plasmid pRK793 [19] and purified using a Ni affinity protocol, was added per 10 mg H4 protein and the cleavage was allowed to proceed for 6 h, at which point the sample was dialyzed back to extraction buffer and applied to Ni:NTA resin to remove the his-tagged N-terminal tail, uncleaved protein and the TEV protease. The N-terminally truncated H4 protein, H4(37-102, A38C), was further concentrated using Vivaspin 15 centrifugal filter units (molecular weight cutoff 5000 Da) and purified by semi preparative C-18 RP-HPLC using a 40-60% HPLC solvent B gradient. The protein was characterized by ESI-MS ($(M + H)^+$ observed: 7351±2 Da; expected 7349 Da). Typical yields were 6 mg pure H4(37-102, A38C) per liter cell culture.

Generation of acetylated H4(A38C)

In a typical ligation reaction, 2 mg of H4(37-102, A38C) protein (0.27 μmol) and 2.4 mg of acH4(1-37)-COSR thioester peptide (0.56 μmol, 2 equiv.) were dissolved in 140 μL ligation buffer (300 mM phosphate pH 7.8, 6 M guanidine hydrochloride) in the presence of 100 mM 4-mercaptophenylacetic acid and 50 mM tris(2-carboxyethyl)phosphine (TCEP). The ligation proceeded rapidly and yielded full conversion within 2 h, at which time the reaction was quenched by addition of an equal volume of HPLC solvent B. The product was purified by semi preparative C-18 RP-HPLC using a 40-60% HPLC solvent B gradient over 45 min yielding 1 mg of pure acH4(A38C). The product was characterized by ESI-MS ($(M + H)^+$ observed: 11520±3 Da; expected 11520 Da).

Desulfurization of acH4(A38C) to acH4

We applied a free-radical based approach to desulfurize cysteine 38 to the native alanine residue [20,21]. In a typical reaction 1 mg acH4(A38C) was dissolved in 100 μL solubilization buffer (200 mM phosphate, pH 7 and 6 M guanidinium chloride). 2 μL ethanethiol, 10 μL 2-methylpropane thiol and 150 μL of 0.5 M TCEP, pH 7 in solubilization buffer were added followed by addition of 0.2 M VA-061 (Wako Chemicals) in methanol and the mixture was incubated at 37°C for 24 h. After complete desulfurization as determined by ESI-MS, acH4 was purified by semi preparative C-18 RP-HPLC over a 0-70% HPLC solvent B gradient yielding 0.5 mg protein. The product was characterized by analytical HPLC and ESI-MS ($(M + H)^+$ observed: 11490±4 Da; expected 11488 Da) (**Supplementary Fig. 1e, f**).

Octamer formation

Histone octamers were formed as previously described [14,16] with minor modifications. Briefly, recombinant histone proteins were dissolved in unfolding buffer (20 mM Tris, pH 7.5, 6 M guanidinium hydrochloride), combined in equimolar amounts and the total histone concentration was adjusted to 1 mg / mL. The mixture was dialyzed into refolding buffer (10 mM Tris, pH 7.5, 200 mM NaCl, 1 mM EDTA) for 2 x 4 h and a final dialysis step over 16 h. The refolded histone octamers were concentrated using Vivaspin 500 centrifugal filter units (5000 Da molecular weight cutoff) and purified from aggregates and free histone monomers or dimers by size-exclusion chromatography using a Superdex 200 10/300 column. Fractions containing pure octamers (i.e. equivalent amounts of all histones) were pooled and concentrated using Vivaspin 500 centrifugal filter units to an octamer concentration of approx. 50 μM. After addition of 50% glycerol (v/v) the refolded octamers were quantified by UV spectroscopy and stored at -20°C.

DNA preparation

For array formation, a plasmid containing 12 copies of a 177 base pair repeat of the 601 nucleosome positioning sequence (12_177_601) [22,23] was purified from a 6 L culture of DH5α cells using a QIAGEN Qiafilter Plasmid Giga kit. The 12_177_601 sequence was obtained by preparative digestion of the plasmid with EcoRV followed by selective precipitation of the fragment with 6% polyethylene glycol-6000 on ice followed by centrifugation at 26,000 g for 30 min. After phenol extraction and ethanol precipitation the DNA

was redissolved in TE buffer (10 mM Tris pH 7.5, 1 mM EDTA), quantified by UV spectroscopy and stored in aliquots at -20°C.

For formation of mono-nucleosomes a 153 base pair segment of 601 DNA (153_601) was produced. A plasmid containing 8 copies of the 147 base pair of the 601 DNA flanked by EcoRV sites was assembled following the general protocol in ref. 24. The plasmid was purified using a QIAGEN Qiafilter Plasmid Giga kit and digested with EcoRV releasing the 153_601 sequence. After precipitation of the vector with 9% polyethylene glycol-6000, the 153_601 fragment was ethanol precipitated, redissolved in TE buffer, quantified by UV spectroscopy and stored at -20°C.

Mono-nucleosome / nucleosome array reconstitution

Mono-nucleosomes were formed as previously described [14] with some modifications. In a typical reconstitution, appropriate histone octamers and DNA (153_601) were combined at a concentration of 2 μM in 75 μL reconstitution buffer (2 M KCl, 10 mM Tris pH 7.8, 0.1 mM EDTA). The mixture was transferred into a Slide-A-Lyzer MINI dialysis unit and dialyzed at 4°C against reconstitution buffer containing 1.4 M KCl, 1.2 M KCl, 1 M KCl, 0.8 M KCl, 0.5 M KCl and 10 mM KCl for 90 min each, followed by a final dialysis step against reconstitution buffer containing 10 mM KCl. The nucleosomes were concentrated using Vivaspin 500 centrifugal filter units (5000 Da molecular weight cutoff) and the quality of the assembly was assessed by separation on a Criterion 5% TBE gel run in 0.5x TBE buffer, followed by staining with ethidium bromide.

Nucleosome arrays were formed as described in ref. [23] with slight changes. In a typical reconstitution, appropriate histone octamers and DNA (12_177_601) were com-

bined at concentrations of 2 μM for the octamers and 2 μM of 601 sites for the DNA in 75 μL reconstitution buffer (2 M KCl, 10 mM Tris pH 7.8, 0.1 mM EDTA). In addition, 1 mM of 147 bp fragments of the weaker binding MMTV A DNA [25] were added to prevent overloading of the array with octamers. The mixture was transferred into a Slide-A-Lyzer MINI dialysis unit and stepwise dialysis was performed as described above. After assembly, free mono-nucleosomes and DNA fragments were removed by selective precipitation of the arrays with $MgCl_2$ (the concentration depended on the octamer type). In detail, arrays were incubated in the presence of 2.5 – 7 mM $MgCl_2$ on ice for 10 min followed by centrifugation at 16,000 g for 10 min. After removal of the supernatant, the arrays were resuspended in TEK buffer, dialyzed against fresh TEK buffer followed by concentration determination by UV spectroscopy. The quality of the assembly was determined by native 1% agarose, 1% polyacrylamide gel electrophoresis (APAGE) separation followed by staining with ethidium bromide (**Supplementary Figs. 2a, 8a** and **9a**). Mono-nucleosomes and nucleosomal arrays were stored on ice for no longer than 4 days.

Full array saturation was confirmed by digestion of 2 pmol of nucleosomes in arrays with 10 U of ScaI restriction enzyme in NEB buffer 3 at room temperature for 12 h, followed by separation on a Criterion 5% TBE gel run in 0.5x TBE buffer and staining with ethidium bromide. The presence of a nucleosome band as well as the absence of free DNA and higher running species indicated full array occupancy. DTT in the digestion and restriction enzyme buffer can lead to partial loss of the ubiquitin moiety in $uH2B_{SS}$ containing arrays (**Supplementary Figs. 2b, 8b** and **9b**). Note that small differences in array occupancy may lead to changes in the observed anisotropy values.

Presence of 12 nucleosomes per array was further confirmed by partial digestion with micrococcal nuclease. 2 pmol of nucleosomes in arrays were digested with 0.2 U of micrococcal nuclease for 60 s on ice. The reaction was stopped by addition of 0.2% (w/v) SDS, 20 mM EDTA followed by DNA purification using a Qiagen PCR purification kit. The DNA fragments were separated on a Criterion 5% TBE gel run in 0.5x TBE buffer and visualized with ethidium bromide (**Supplementary Fig. 2c**).

Determination of R_0 of fH2A

The R_0 of fluorescein-5-maleimide coupled to H2A (in histone octamers) of 47.3 Å was determined using the formula $R_0 = 0.211\left(\kappa^2 n^{-4} Q_D J(\lambda)\right)^{1/6}$ [26], where the overlap integral $J(\lambda)$ was calculated from UV and fluorescence emission spectra (**Supplementary Fig. 5e**) with an extinction coefficient for fluorescein-5-maleimide of 83000 $cm^{-1}M^{-1}$, the orientation factor was assumed $\kappa^2 = 2/3$, the refractive index was taken as $n = 1.34$, and the donor quantum yield Q_D of fluorescently labeled H2A was determined as 0.534 relative to 5(6)-carboxyfluorescein as reference.

Determination of fluorescence lifetime in fH2A containing arrays

Fluorescence lifetime measurements were performed using a Fluorolog-3 fluorescence spectrometer (HORIBA Jobin Yvon, Edison NJ) in a T-channel configuration and a pulsed laser excitation source at 445 nm, with 700 ps pulse length. Emission was recorded at 520 nm. The samples were prepared at the same concentration and buffer conditions as for anisotropy measurements. Data was analyzed by fitting to a single exponential decay using the DAS6 analysis software (HORIBA Jobin Yvon, Edison NJ).

Mass spectrometric analysis of H3 K79 methylation

For mass spectrometric analysis, hDot1L assays were performed using non-radioactive S-adenosyl methionine. The reactions were separated by SDS-PAGE on a Criterion 15% Tris-glycine gel followed by visualization by Coomassie G-250 staining. The H3 containing band was excised and analyzed on a ThermoFisher LTQ-Orbitrap mass spectrometer after in-gel trypsin digestion (**Supplementary Fig. 3c, d**).

References

1. Woodcock, C. L., Grigoryev, S. A., Horowitz, R. A. & Whitaker, N. A chromatin folding model that incorporates linker variability generates fibers resembling the native structures. *Proc Natl Acad Sci U S A* **90**, 9021-5 (1993).
2. Runnels, L. W. & Scarlata, S. F. Theory and application of fluorescence homotransfer to melittin oligomerization. *Biophys J* **69**, 1569-83 (1995).
3. 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* **389**, 251-60 (1997).
4. Schalch, T., Duda, S., Sargent, D. F. & Richmond, T. J. X-ray structure of a tetranucleosome and its implications for the chromatin fibre. *Nature* **436**, 138-41 (2005).
5. Dorigo, B. et al. Nucleosome arrays reveal the two-start organization of the chromatin fiber. *Science* **306**, 1571-3 (2004).
6. Kirkwood, J. G. The general theory of irreversible processes in solutions of macromolecules. *J Polym Sci: Part B* **12**, 1-14 (1954).
7. Bloomfield, V., Dalton, W. O. & Van Holde, K. E. Frictional coefficients of multisubunit structures. I. Theory. *Biopolymers* **5**, 135-48 (1967).
8. Arya, G., Zhang, Q. & Schlick, T. Flexible histone tails in a new mesoscopic oligonucleosome model. *Biophys J* **91**, 133-50 (2006).
9. van Holde, K. *Chromatin* (ed. Rich, A.) (Springer, New York, 1989).
10. Craver, F. & Knox, R. Theory of polarization quenching by excitation transfer II. Anisotropy and second-neighbour considerations. *Mol Phys* **22**, 385-402 (1971).
11. Mendez, J. & Stillman, B. Chromatin association of human origin recognition complex, cdc6, and minichromosome maintenance proteins during the cell cycle: assembly of prereplication complexes in late mitosis. *Mol Cell Biol* **20**, 8602-12 (2000).
12. O'Neill, L. P. & Turner, B. M. Immunoprecipitation of native chromatin: NChIP. *Methods* **31**, 76-82 (2003).

13. Henikoff, S., Henikoff, J. G., Sakai, A., Loeb, G. B. & Ahmad, K. Genome-wide profiling of salt fractions maps physical properties of chromatin. *Genome Res* **19**, 460-9 (2009).
14. Luger, K., Rechsteiner, T. J., Flaus, A. J., Waye, M. M. & Richmond, T. J. Characterization of nucleosome core particles containing histone proteins made in bacteria. *J Mol Biol* **272**, 301-11 (1997).
15. McGinty, R. K. et al. Structure activity analysis of semisynthetic nucleosomes: Mechanistic insights into the stimulation of Dot1L by ubiquitylated histone H2B. *ACS Chem Biol* (2009).
16. McGinty, R. K., Kim, J., Chatterjee, C., Roeder, R. G. & Muir, T. W. Chemically ubiquitylated histone H2B stimulates hDot1L-mediated intranucleosomal methylation. *Nature* **453**, 812-6 (2008).
17. Camarero, J. A. & Muir, T. W. Native chemical ligation of polypeptides. *Curr Protoc Protein Sci* **Chapter 18**, Unit18 4 (2001).
18. Schnolzer, M., Alewood, P., Jones, A., Alewood, D. & Kent, S. B. In situ neutralization in Boc-chemistry solid phase peptide synthesis. Rapid, high yield assembly of difficult sequences. *Int J Pept Protein Res* **40**, 180-93 (1992).
19. Kapust, R. B. et al. Tobacco etch virus protease: mechanism of autolysis and rational design of stable mutants with wild-type catalytic proficiency. *Protein Eng* **14**, 993-1000 (2001).
20. Wan, Q. & Danishefsky, S. J. Free-radical-based, specific desulfurization of cysteine: a powerful advance in the synthesis of polypeptides and glycopolypeptides. *Angew Chem Int Ed Engl* **46**, 9248-52 (2007).
21. Chiang, K. P., Jensen, M. S., McGinty, R. K. & Muir, T. W. A semisynthetic strategy to generate phosphorylated and acetylated histone H2B. *Chembiochem* **10**, 2182-7 (2009).
22. Lowary, P. T. & Widom, J. New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. *J Mol Biol* **276**, 19-42 (1998).
23. Dorigo, B., Schalch, T., Bystricky, K. & Richmond, T. J. Chromatin fiber folding: requirement for the histone H4 N-terminal tail. *J Mol Biol* **327**, 85-96 (2003).
24. Dyer, P. N. et al. Reconstitution of nucleosome core particles from recombinant histones and DNA. *Methods Enzymol* **375**, 23-44 (2004).
25. Flaus, A. & Richmond, T. J. Positioning and stability of nucleosomes on MMTV 3'LTR sequences. *J Mol Biol* **275**, 427-41 (1998).
26. Lakowicz, J. R. *Principles of Fluorescence Spectroscopy* (Plenum Press, New York, 1983).

Supplementary Results

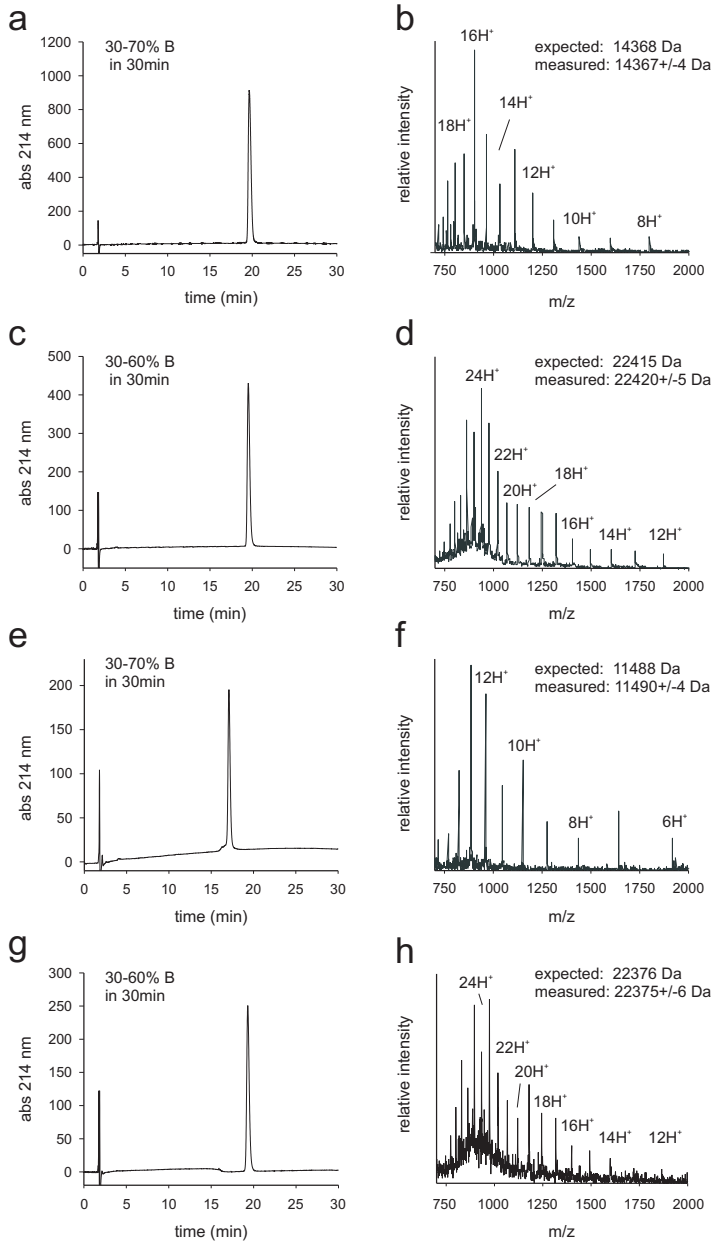

Figure S1. | Characterization of fluorescently labeled H2A (fH2A), disulfide-linked ubiquitylated H2B (uH2B$_{SS}$), semi-synthetic tail-acetylated H4 (acH4), and disulfide linked Hub1-H2B (hub1-H2B$_{SS}$) a, C18 analytical RP-HPLC chromatogram of purified fH2A. **b**, ESI-MS of purified fH2A. **c**, C18 analytical RP-HPLC chromatogram of purified uH2B$_{SS}$. **d**, ESI-MS of purified disulfide-linked uH2B$_{SS}$. **e**, C18 analytical RP-HPLC chromatogram of purified H4 bearing an acetylated N-terminus and acetylated lysines 5,8,12,16 and 20. **f**, ESI-MS of purified acH4. **g**, C18 analytical RP-HPLC chromatogram of purified hub1-H2B$_{SS}$. **h**, ESI-MS of purified disulfide-linked hub1-H2B$_{SS}$.

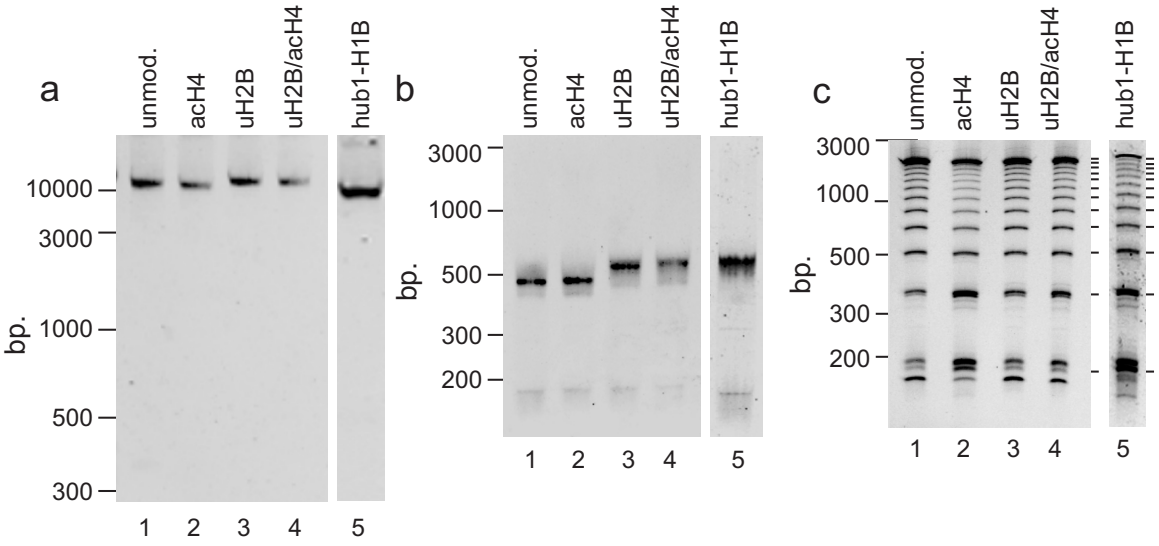

Figure S2. | Reconstitution of modified chromatin arrays. a, Native 1% agarose, 1% polyacrylamide gel electrophoresis (APAGE) analysis of reconstituted 12-mer nucleosomes arrays, containing unmodified histones, acH4, uH2B$_{SS}$, acH4/uH2B$_{SS}$ or hub1-H2B$_{SS}$. 2 pmol of nucleosomes in reconstituted arrays were loaded in 10% sucrose, run for 2 h at 200 V, 4 °C and stained with ethidium bromide. **b**, Ethidium bromide stained 5% TBE gel of ScaI restriction enzyme digests of reconstituted arrays containing unmodified histones, acH4, uH2B$_{SS}$, acH4/uH2B$_{SS}$ or hub1-H2B$_{SS}$ showing full nucleosome occupancy. **c**, Ethidium bromide stained 5% TBE gels of micrococcal nuclease (MNase) digests of reconstituted arrays containing unmodified histones, acH4, uH2B, acH4/uH2B$_{SS}$ or hub1-H2B$_{SS}$ showing similar digestion patterns for all arrays.

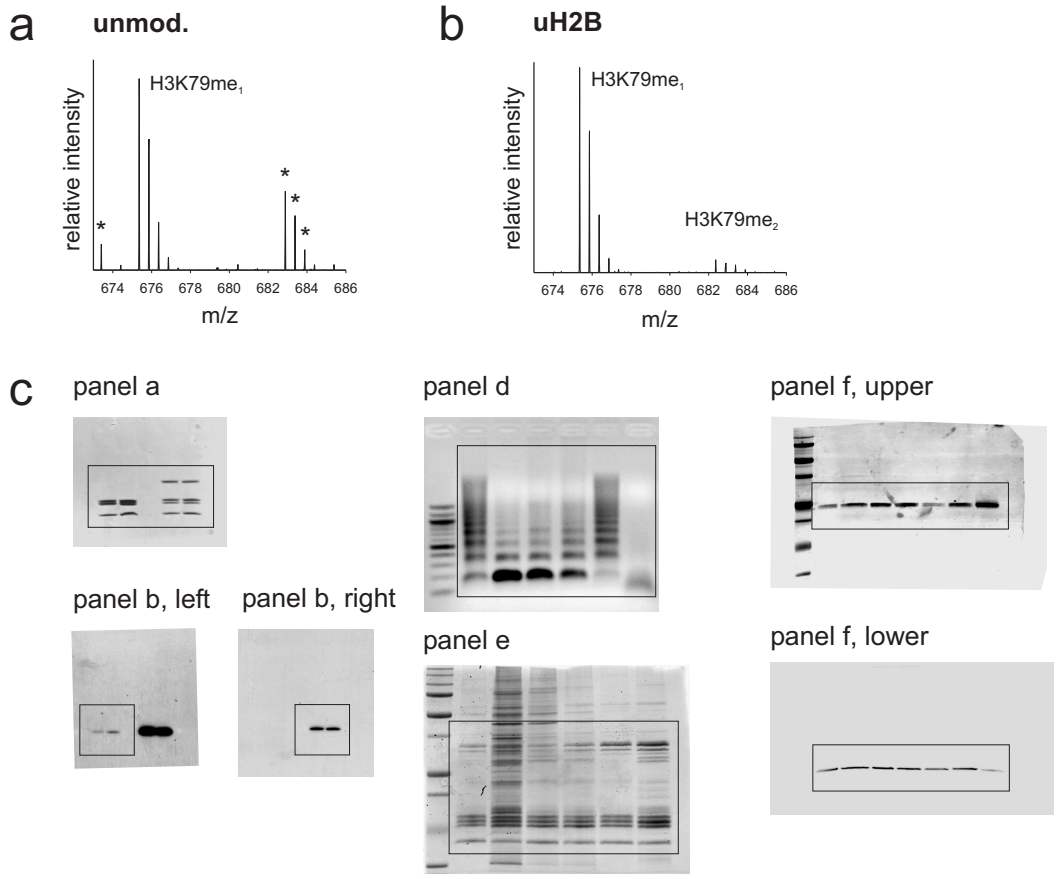

Figure S3. | hDot1L mono- and dimethylates H3 K79. a,b Mass spectrometric analysis of hDot1L activity on modified chromatin arrays confirms the specificity for H3 K79. In unmodified arrays (**a**) only mono-methylated lysine 79 is detected (doubly charged, m/z = 675.35), whereas for $uH2B_{SS}$ containing arrays (**b**) also di-methylated lysine 79 is observed (doubly charged, m/z = 682.37). An asterisk denotes ion intensity arising from peptides unrelated to H3. Tri-methylated lysine 79 is not observed. **c**, Full gels and Western blots of Fig. 2. The boxes indicate the part displayed in Fig. 2.

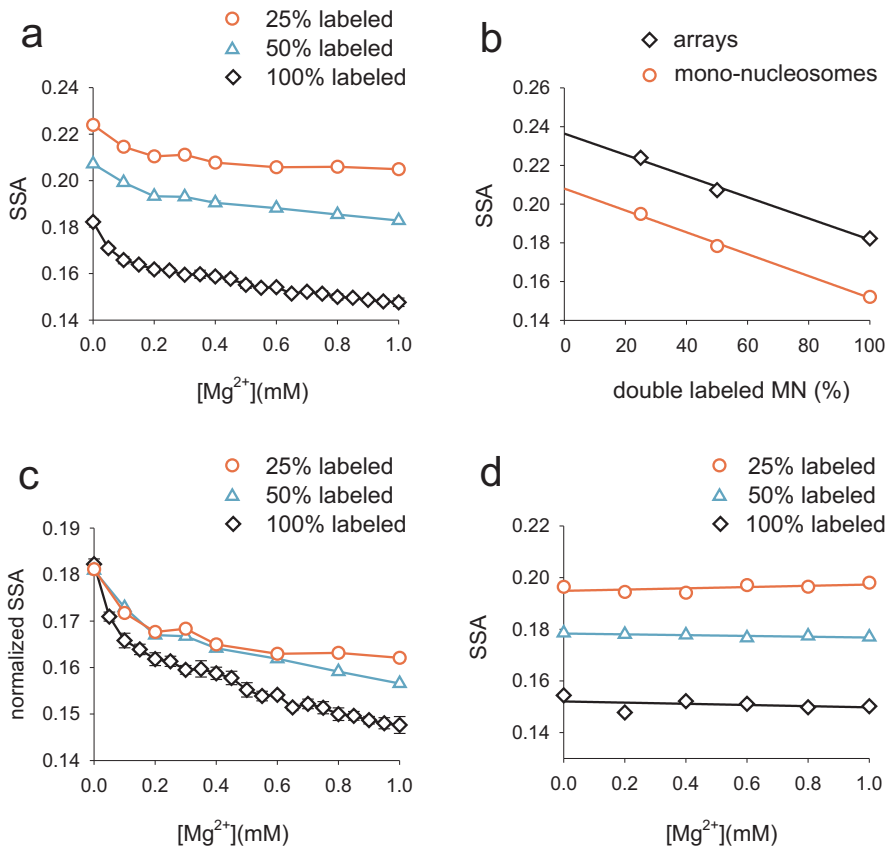

Figure S4. | Homo-RET is the major mechanism of fluorescence depolarization upon array compaction. a, SSA of nucleosomal arrays containing either 100% fluorescently labeled H2A N110C (fH2A) (black), 50% fH2A (blue) or 25% fH2A (red) at different Mg^{2+} concentrations. **b**, SSA values at 0 mM Mg^{2+} are plotted against the fraction of doubly labeled nucleosomes for mono-nucleosomes (red) and nucleosomal arrays (black). The solid lines are linear fits to the data. Due to intranucleosomal homo-RET, SSA rises linearly with decreasing amount of doubly labeled nucleosomes. The y-axis intersection represents the SSA of a singly labeled mono-nucleosome or nucleosome in an array. **c**, SSA values from arrays are normalized to similar values at 0 mM Mg^{2+} using the linear extrapolation shown in panel b. The decreased dynamic range upon decreasing degree of labeling indicates that homo-RET is the dominant mechanism of depolarization. **d**, SSA of mono-nucleosomes containing either 100% fH2A (black), 50% fH2A (blue) or 25% fH2A (red) at different Mg^{2+} concentrations. The solid lines are linear fits to the data, showing no significant dependence of SSA on Mg^{2+} concentration. **a, c**. Error bars, standard deviation (n=2-3).

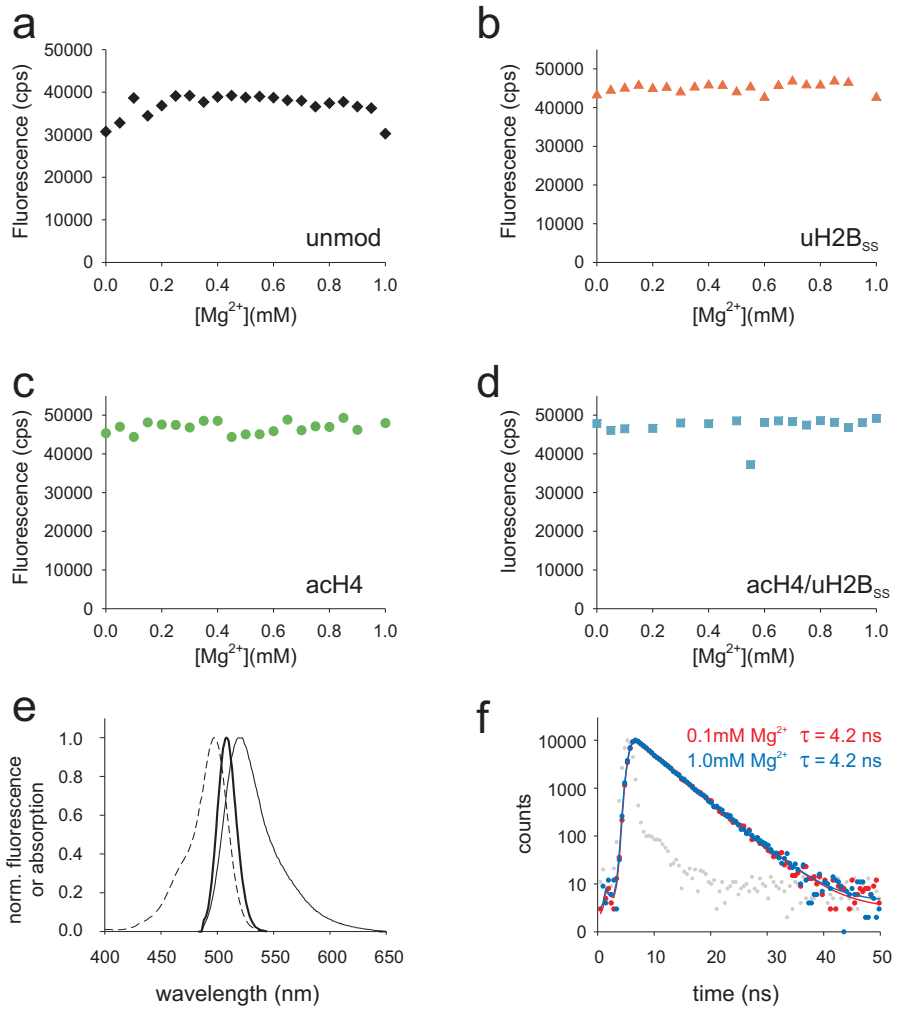

Figure S5. | Fluorescence emission of arrays does not change with increasing Mg^{2+} concentration. Vertically polarized fluorescence emission of unmodified (**a**, black), uH2B containing (**b**, red), acH4 (**c**, green) and uH2B/acH4 doubly modified (**d**, blue) arrays. Excitation is achieved with vertically polarized light. **e**, Spectral properties of fH2A and determination of R_0. Dashed line: normalized absorption spectrum of fH2A, solid line: normalized fluorescence emission spectrum of fH2A, bold line: normalized overlap integral used to calculate the Förster radius R_0 as 47.3 Å (see supplementary text). **f**, Fluorescence lifetime of fluorescein attached to chromatin arrays at 0.1mM Mg^{2+} (red) or 1.0mM Mg^{2+} (blue). The solid lines are fits to the data, exhibiting a single exponential decay with time constants of 4.21±0.02 ns (0.1mM Mg^{2+}) and 4.22±0.02 ns (1.0mM Mg^{2+}). The excitation laser pulse is shown in grey. Spectra and kinetic traces were recorded in 10 mM Tris buffer, pH 7.8 and 10 mM KCl at 22.5°C.

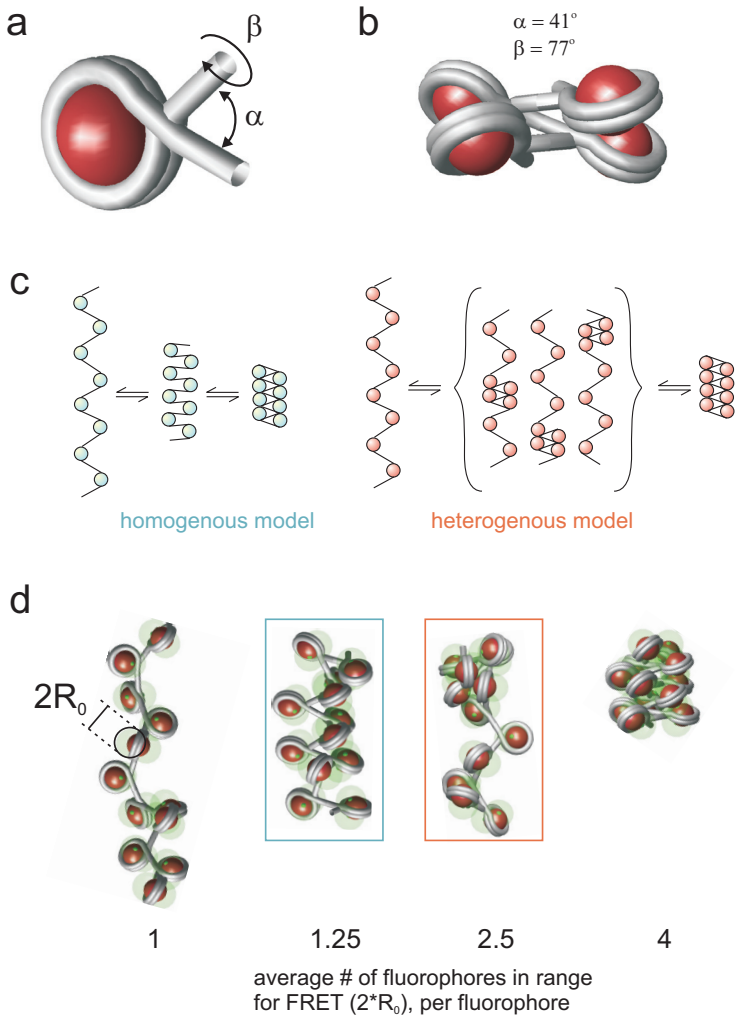

Figure S6. | Folding of chromatin arrays proceeds through close interactions between nucleosomes. a, Two angle model of a nucleosomal array, based on the model of Woodcock *et al.* [1]. For details see supplementary text. **b**, Conformation of four consecutive nucleosomes in the compacted state. **c**, Two different models of compaction: In the *homogenous* compaction model, nucleosomes approach gradually throughout the folding transition until the final, fully compacted state is reached. In the *heterogenous* model, clusters of stacked nucleosomes form whereas the remaining nucleosomes are unconstrained. The compacted clusters increase in size throughout the transition until the fiber is fully folded. **d**, Representative structures of either fully extended (sedimentation coefficient: 36 S, SSA: 0.181), intermediately compacted following the homogenous (blue box, 42 S, SSA: 0.180) and the heterogeneous model (red box, 42 S, SSA: 0.162), or fully compacted (52 S, SSA: 0.147) array models are displayed. The fluorophore attachment sites are indicated by small green dots. The transparent green spheres with radius R_0 allow an estimate of the distance range within which homo-RET will occur (< 2*R_0). The average calculated number of fluorophores within the distance of 2*R_0 from a given fluorophore is indicated.

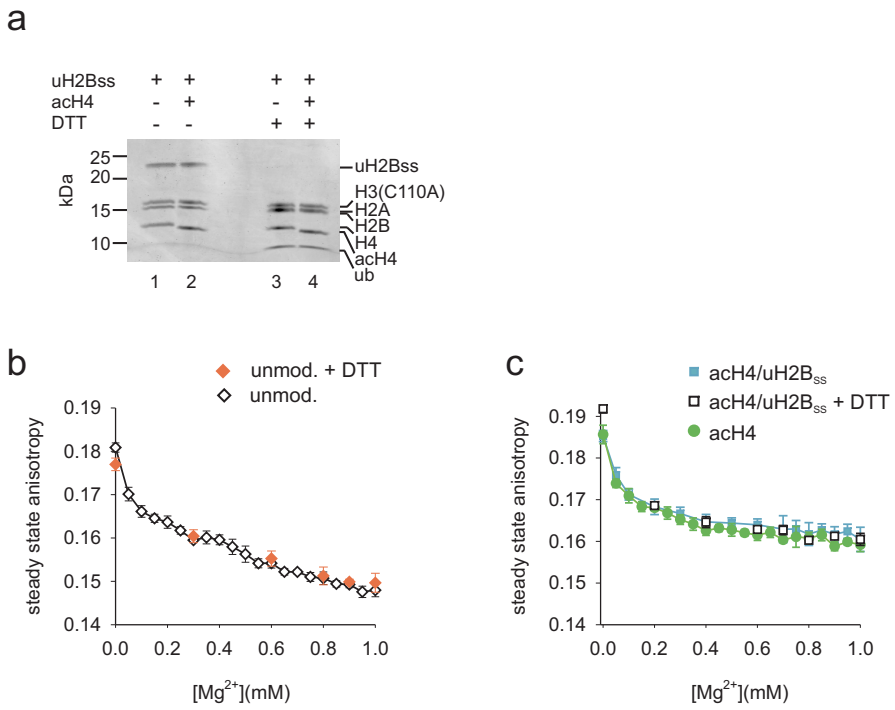

Figure S7. | Reductive deubiquitylation with DTT and the effect on array folding. a, Nucleosomal arrays containing either $uH2B_{SS}$ or both $uH2B_{SS}$ and acH4 are incubated in the absence or presence of 10 mM DTT for 30 min on ice, followed by non-reducing SDS-PAGE and staining with Coomassie Brilliant Blue. Full reduction of the ubiquitin-H2B disulfide bond is observed. **b**, DTT has no effect on compaction of non-ubiquitylated arrays. Compaction of unmodified arrays is measured in absence (black) or presence (red) of 10 mM DTT. **c**, Chemical deubiquitylation has no effect on compaction of doubly modified arrays. Compaction of uH2B/acH4 arrays is measured in absence (blue) or presence (black) of 10 mM DTT. The data of acH4 containing arrays (green) is included for comparison. **b, c**. Error bars, standard deviation (n=2-3).

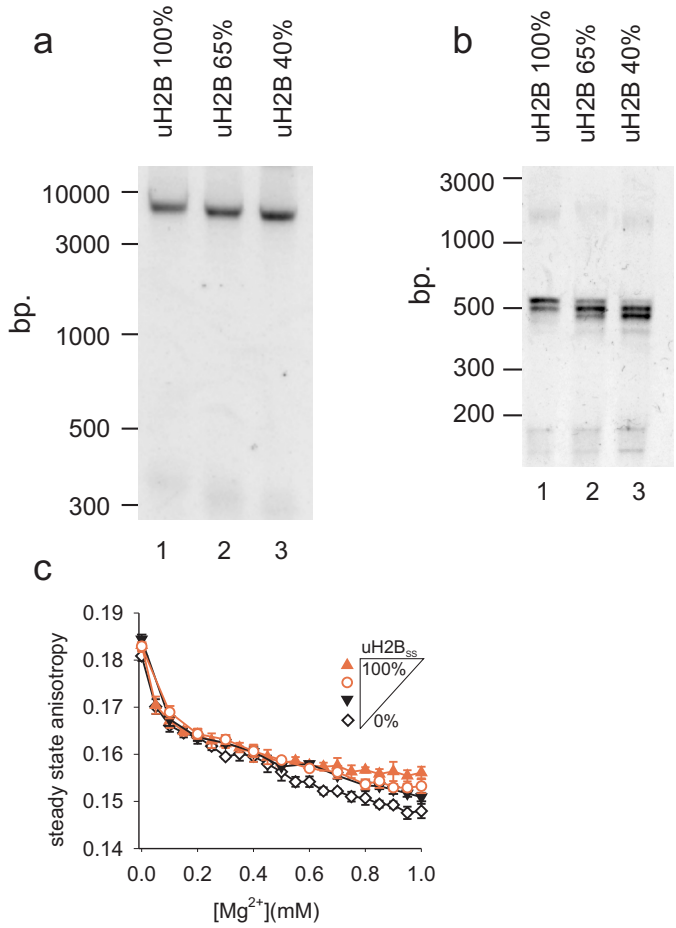

Figure S8. | uH2B decompacts chromatin structure in a dosage dependent manner. **a**, Native APAGE of reconstituted arrays containing 100%, 65% or 40% $uH2B_{SS}$. **b**, Ethidium bromide stained 5% TBE gels of ScaI restriction enzyme digests of reconstituted arrays containing 100%, 65% or 40% $uH2B_{SS}$. The double band in the 100% $uH2B_{SS}$ containing preparation arises from partial reduction of the ubiquitin-H2B disulfide bond due to DTT in the restriction enzyme buffer. **c**, Ubiquitylated H2B gradually opens up chromatin structure. Chromatin folding upon Mg^{2+} addition was determined by measuring SSA for arrays containing unmodified (black, open symbols), 40% (black, closed symbols), 65% (red, open symbols) and 100% (red, closed symbols) ubiquitylated H2B.

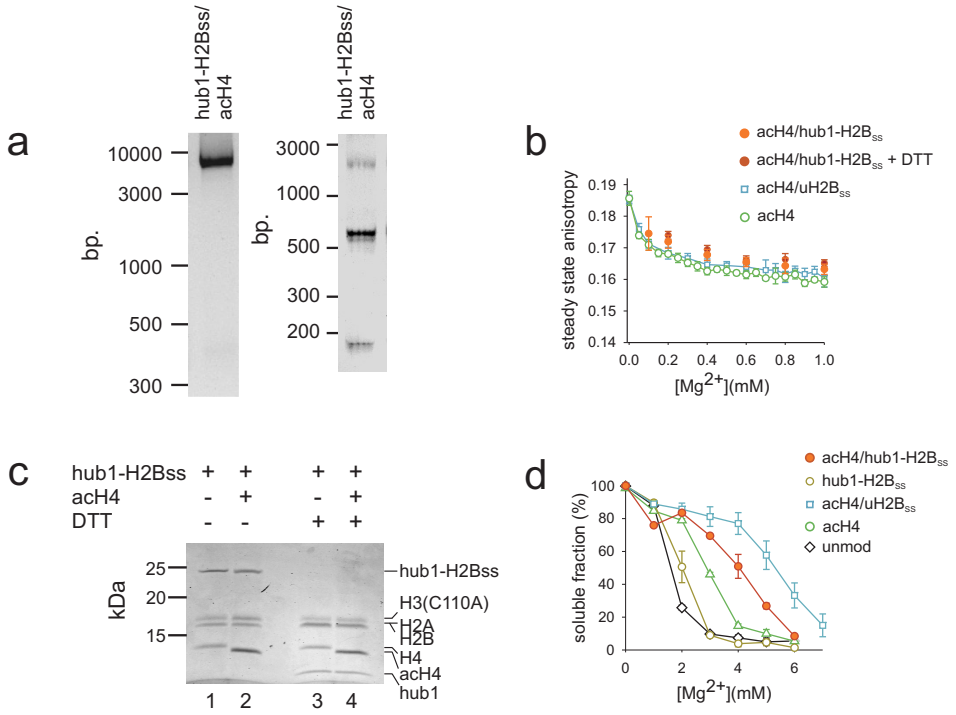

Figure S9. | The effect of hub1-H2B$_{SS}$ incorporation on compaction and oligomerization of fibers containing acH4. a, Left panel, native 1% agarose, 1% polyacrylamide gel electrophoresis (APAGE) analysis of reconstituted arrays, containing both hub1-H2B$_{SS}$ and acH4, stained with ethidium bromide. Right panel, ethidium bromide stained 5% TBE gel of ScaI restriction enzyme digests of reconstituted arrays, containing both hub1-H2B$_{SS}$ and acH4. **b,** Chemical deubiquitylation has no effect on compaction of arrays containing acH4 and hub1-H2B$_{SS}$. Compaction of hub1-H2B$_{SS}$ is measured in absence (orange) or presence (brown) of 10 mM DTT. The data of acH4 containing arrays (green) and uH2B/acH4 arrays (blue) is included for comparison. Error bars, standard deviation (n=2-3). **c**, Nucleosomal arrays containing either hub1-H2B$_{SS}$ or both hub1-H2B$_{SS}$ and acH4 are incubated in the absence or presence of 10 mM DTT for 30 min on ice, followed by non-reducing SDS-PAGE and staining with Coomassie Brilliant Blue. Full reduction of the hub1-H2B disulfide bond is observed. **d**, Arrays containing acH4 and hub1-H2B$_{SS}$ (orange) were incubated in the presence of indicated concentrations of Mg^{2+} and oligomers were removed by centrifugation. The amount of arrays remaining in solution was determined by UV absorption. Data from unmodified, hub1-modified (yellow), ubiquitylated (red), H4 acetylated (green) and doubly acetylated/ubiquitylated (blue) are displayed for comparison. Errors bars, SEM (n=2-4).

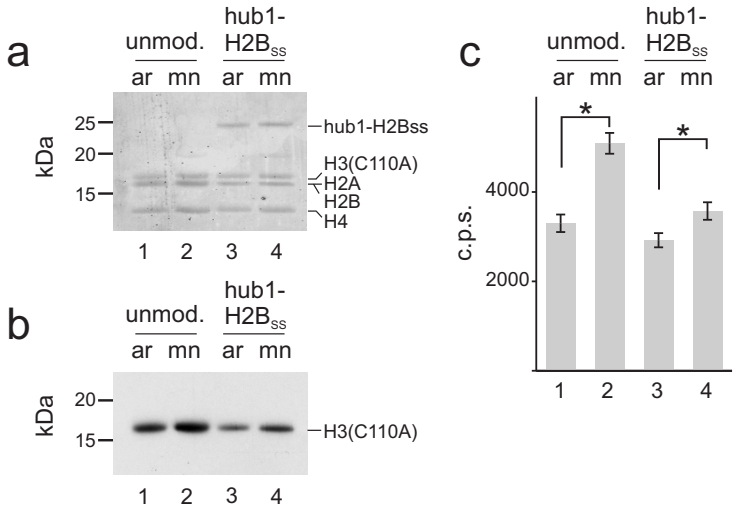

Figure S10. | Hub1-H2B$_{SS}$ incorporation does not allow full hDot1L access. Unmodified arrays (lane 1), unmodified mono-nucleosomes (lane 2), hub1-H2B$_{SS}$ containing arrays (land 3) and hub1-H2B$_{SS}$ containing mono-nucleosomes (lane 4) were used as substrates for methyltransferase assays with ^{3}H-SAM at 1 mM Mg^{2+}. Assay samples were subjected to SDS-PAGE and stained with Coomassie Brilliant Blue (**a**) before probing for 3H-methyl incorporation by fluorography (**b**). Quantification of methylation was performed by p81 filter binding followed by liquid scintillation counting (**c**), Student's two tailed t-test: *: $p = 7 \cdot 10^{-6}$ for unmodified, $p = 0.02$ for hub1-H2B$_{SS}$ containing samples. Error bars, SEM (n=12).

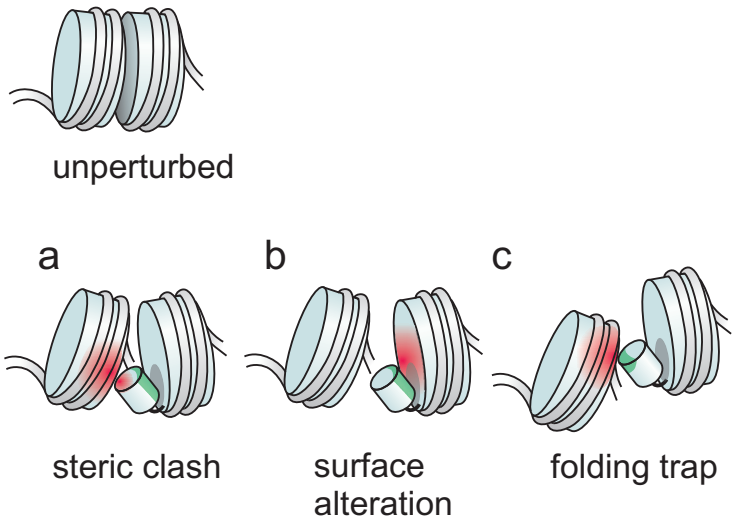

Figure S11. | Models for the effect of uH2B on chromatin compaction. Based on our results we can conceive different non-exclusive mechanisms through which the ubiquitin moiety could impair internucleosomal interactions and thereby fiber compaction. The important surfaces on ubiquitin leading to these effects are indicated in green, whereas the perturbation is indicated in red. For simplicity, only a single ubiquitin moiety is shown. **a**, Ubiquitin moieties properly positioned on the nucleosomal surface lead to steric clashes preventing tight nucleosomal stacking. **b**, Specific interactions of ubiquitin with the nucleosomal surface lead to structural changes in the nucleosome which reduce interactions with neighboring nucleosomes. **c**, Interactions of the ubiquitin moiety with neighboring nucleosomes or DNA result in a nucleosome array incompetent of forming a 30 nm fiber.