Supplemental Table 1. Data collection and refinement statistics for $(CENP\text{-}A/H4)_2$

Data collection		
Beamline	ALS, 8.2.2	APS, 23-ID-D
Wavelength (Å)	1.000	1.000
Unit cell (Å)	*a*=*b*=62.87, *c*=159.51 α=β=90°, γ=120	*a*=*b*=61.46, *c*=185.62 α=β=90°, γ=120°
Space group	$P6_522$	$P6_522$
No. (CENP-A/H4) / a.u.	1	1
Resolution limit (Å)	54.4 – 2.1	53.2 – 2.5
Measured reflections	95,281	60,994
Unique reflections	11,585	7,136
Completeness (%; overall/ last shell)	97.5 / 87.7 [a]	90.9 / 90.5 [b]
I / σI (overall / last shell)	23.39 / 4.30 [a]	12.15 / 2.44 [b]
R_{sym} [e] (overall / last shell)	5.7 / 53.5 [a]	12.1 / 68.5 [b]
Refinement		
Resolution limit (Å)	20 – 2.10	20 – 2.50
No. reflections (working/free)	10,961 / 550	6,806 / 330
R_{cryst} [f] (overall / last shell)	0.180 / 0.208 [c]	0.237 / 0.272 [d]
R_{free} (overall / last shell)	0.241 / 0.313 [c]	0.308 / 0.341 [d]
No. protein/water/ion res.	143/136/4	141/25/5
rmsd from ideal geometry (Å)		
Bond length	0.014	0.015
Angle distances	1.329	1.638
Estimated coordinate error (Å)	0.133	0.262
Ramachandran plot statistics		
Residues in most favored regions	98.5%	92.2%
Residues in allowed regions	0.8%	7.8%
Residues in gen. allowed regions	0.8%	0.0%
Residues in disallowed regions	0.0%	0.0%
PDB ID	3NQJ	3NQU

[a]last shell = 2.08 - 2.21 Å, [b]last shell =2.50 – 2.64 Å,
[c]last shell = 2.10 - 2.15 Å, [d]last shell = 2.50 - 2.56 Å,
[e]$R_{sym} = \Sigma|I-\langle I\rangle|/ \Sigma I$,
[f]$R_{cryst} = \Sigma||F_{obs}|-|F_{calc}||/ \Sigma|F_{obs}|$, 5 % randomly omitted reflections were used for R_{free}

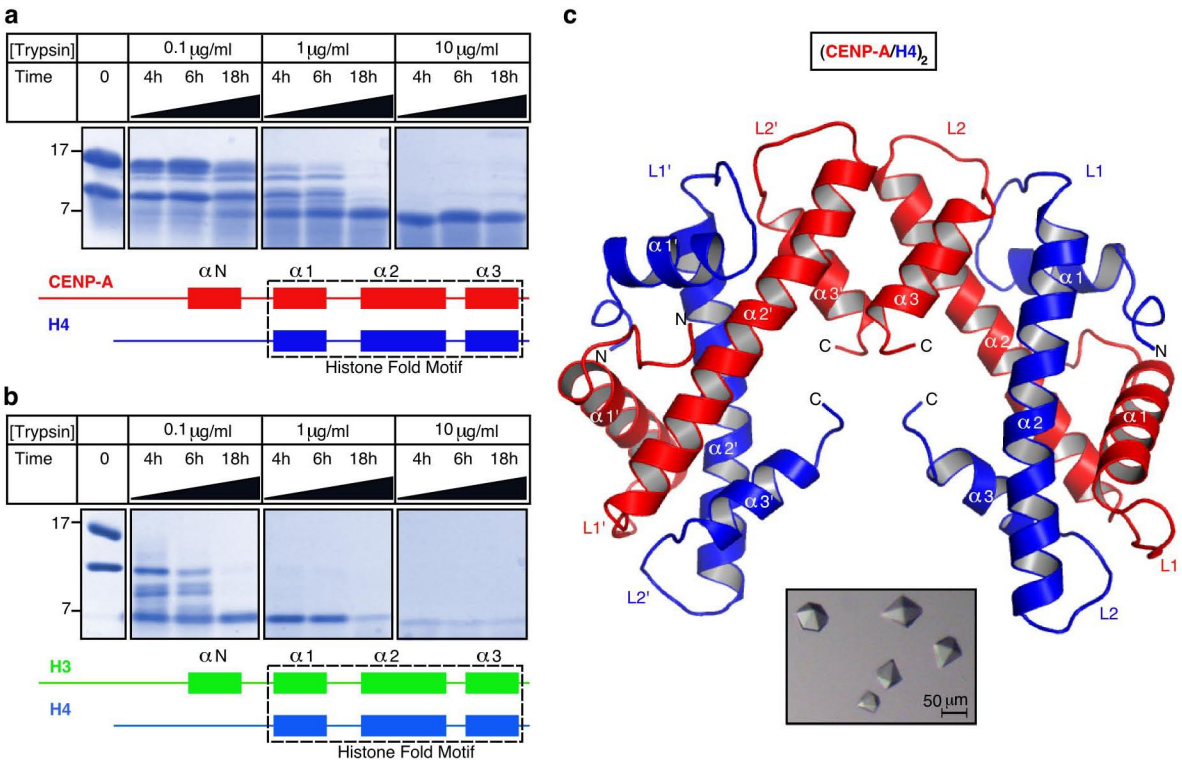

Figure S1 I Crystal structure of the partially trypsinized (CENP-A/H4)$_2$ heterotetramer. a, Partial trypsinization of CENP-A/H4 complex. We reasoned that the removal of the most flexible portions of the CENP-A/H4 complex (regions highly unlikely to provide useful structural information) may ease crystallization since our efforts to generate high quality crystals with the full length proteins were unsuccessful. We found that upon proteolysis by trypsin, CENP-A/H4 complexes strongly protect segments at 6-7 kDa consistent with protection of the CENP-A and H4 histone fold domains (i.e. the α1, α2, and α3 helices and connector loops [L1 and L2] of each histone subunit; the boxed region in the diagram below represents the predicted region of protection within CENP-A and H4). Molecular markers are indicated in kDa. **b,** Trypsinization of the sub-nucleosomal H3/H4 complex. Conventional (H3/H4)$_2$ sub-nucleosomal heterotetramers protect products at 6-7 kDa, as well. We noted that despite the presence of the same number (nine) of consensus trypsinization sites within their histone fold domains relative to the conventional complex, the CENP-A/H4 histone fold domains were resistant to further proteolysis under conditions where the H3/H4 histone fold domains were readily degraded. Furthermore, the chimeric histone H3 containing the CATD, H3^{CATD}, confers similar protection from proteolysis, as does CENP-A (Fig. S10). These data suggest that the conformational rigidity conferred to CENP-A and its binding partner, H4, by the CATD[1] protects CENP-A/H4 complexes from trypsinization under conditions where the more flexible H3/H4 complexes are completely degraded. The increased flexibility provides a likely explanation for a lack, to date, of an atomic resolution crystal structure of isolated (H3/H4)$_2$ heterotetramers. **c**, A ribbon diagram of the (CENP-A/H4)$_2$ heterotetramer. The inset shows (CENP-A/H4)$_2$ heterotetramer crystals forming after partial trypsinization. The final structure was refined to crystallographic R-factor of 0.237 and R_{free} of 0.308 (Supplementary Table 1). The structure, with clear electron density for residues 59-134 of CENP-A and 25-91 of H4, confirms that the region protected from proteolysis consists of the histone fold domains of each CENP-A and H4.

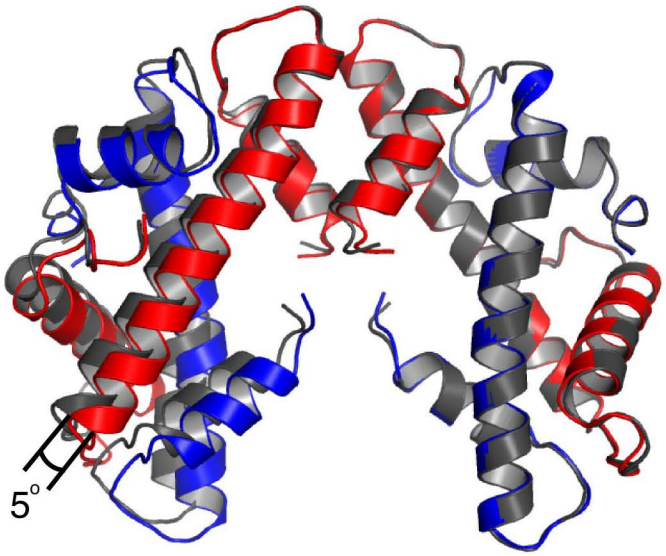

Figure S2 | An overlay of the engineered (CENP-A/H4)$_2$ heterotetramer crystal structure with the structure obtained from partial trypsinization of the full length proteins. Crystal structure obtained with engineered protein is colored as in Fig. 1a, while structure obtained by partial proteolysis (Fig. S1) is in grey. Overlay was done with an SSM algorithm performed with one CENP-A molecule from each complex.

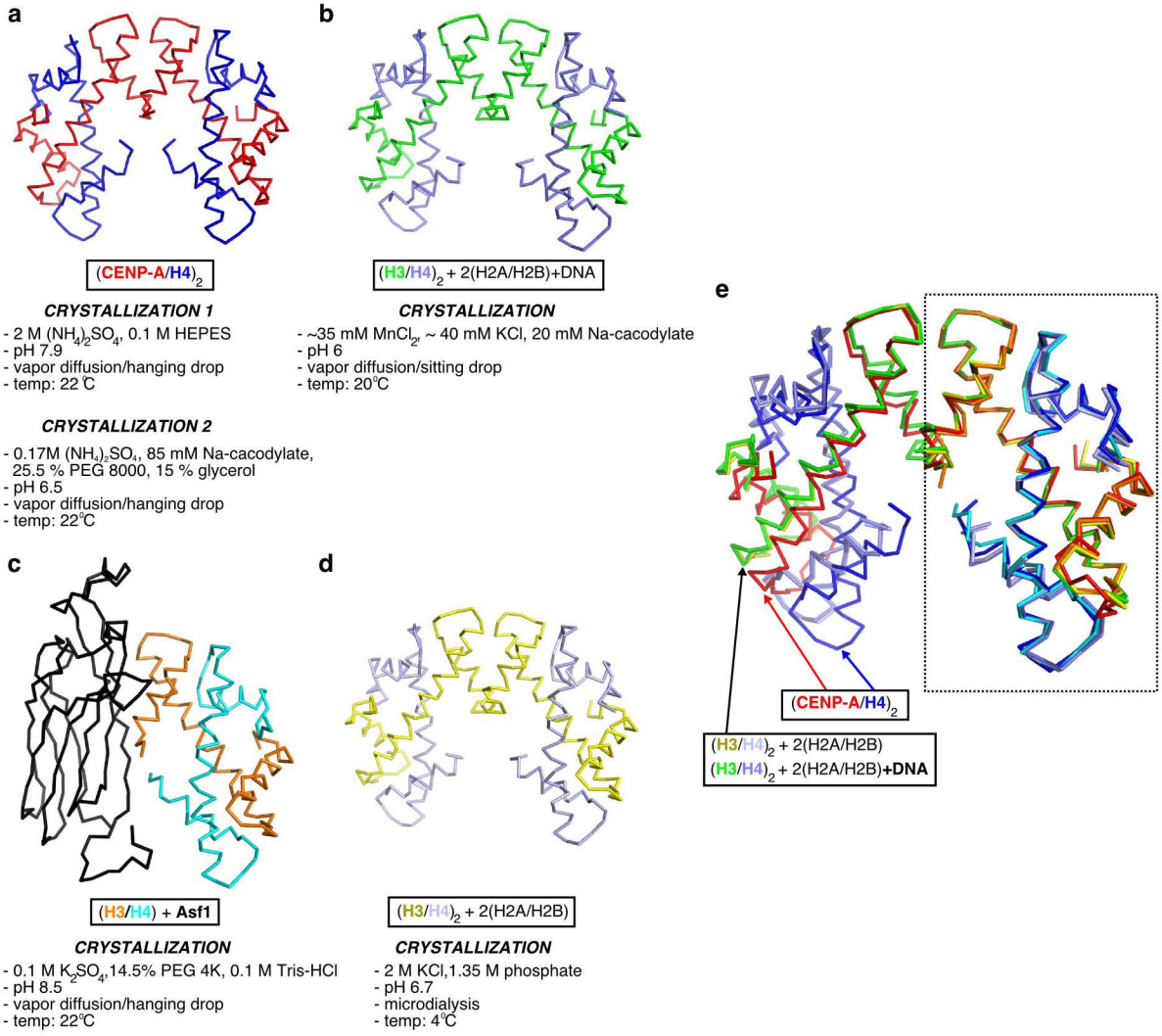

Figure S3 | Structural comparisons of histones in different complexes. a, Heterotetramer of CENP-A and H4. Crystallization 1 corresponds to partially proteolyzed heterotetramer and crystallization 2 to heterotetramer obtained with engineered truncated protein variants. **b**, Tetramer of H3 and H4 from a structure of the nucleosome (PDB ID 1KX5[2]). **c**, Trimer of H3, H4, and Asf1 (PDB ID 2HUE[3]). **d**, Heterotetramer of H3 and H4 from a structure of the histone octamer (H2A/H2B dimers bound but no DNA) (PDB ID 1TZY[4]). For easier comparison all histone structures are truncated to the same length. Crystallization conditions for each complex are as indicated. **e**, Overlay of structures shown in a-d based of right half of the tetramer (boxed). Asf1 (black in c) is omitted for clarity. Note the remarkable identity in the overall fold of histones in all structures in spite of completely different crystallization conditions. Although, only molecules on the left are used for overlay, H3/H4 tetramer overlays perfectly whether structure is coming from within the nucleosome or octamer without DNA.

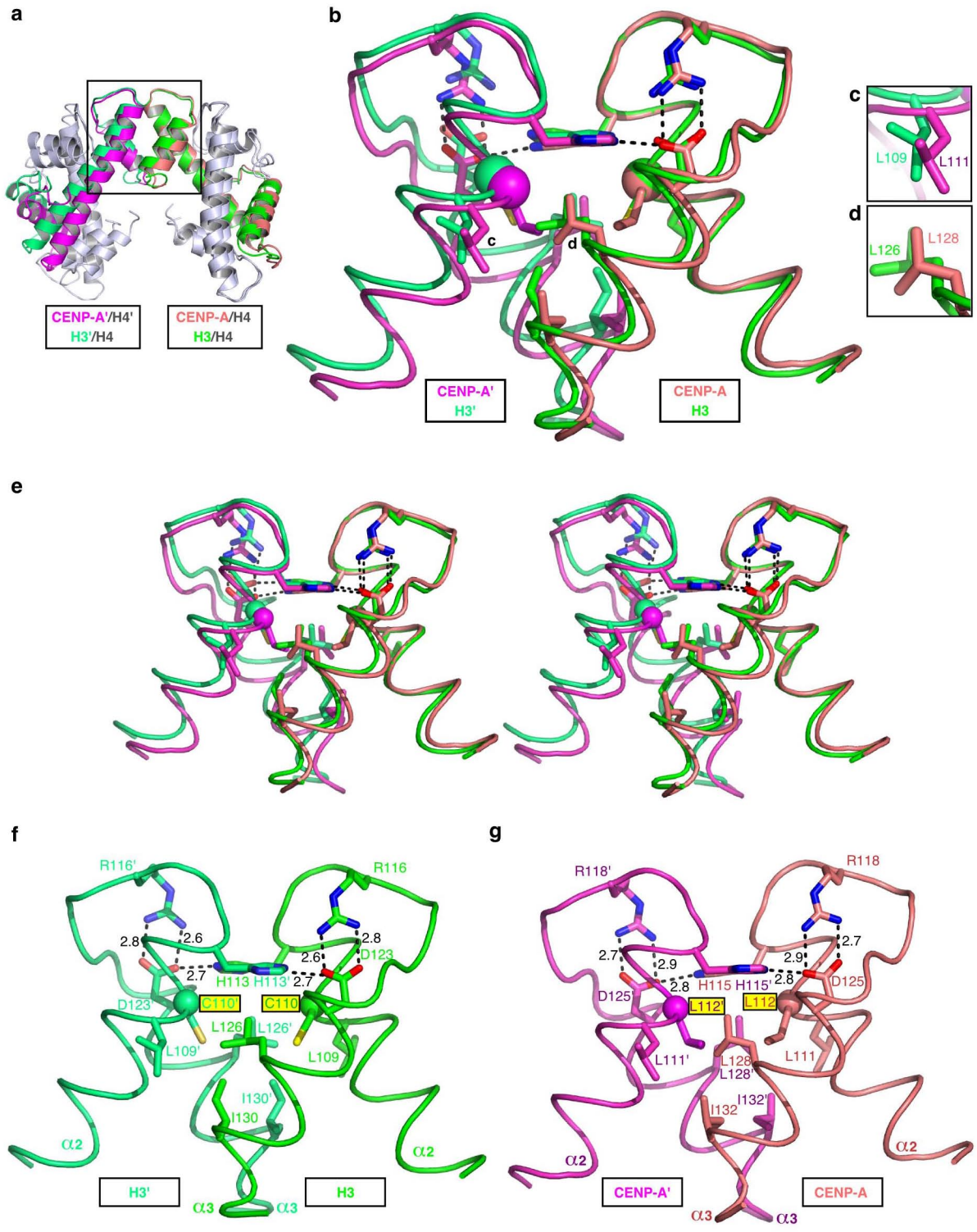

Figure S4 I The tetramerization interface of the (CENP-A/H4)$_2$ heterotetramer. a, Overlay of (CENP-A/H4)$_2$ and (H3/H4)$_2$ with tetramerization interface boxed. Residues 111-113 of CENP-A are overlaid with the analogous positions in H3 (rmsd = 0.13 Å). **b,** Zoom-in view of the overlay in a. α-carbon residues of Cys110 in H3 and Leu112 in CENP-A are shown as spheres. **c** and **d**, Isolated magnified views showing different rotamers of the indicated residues at the CENP-A/CENP-A and H3/H3 interfaces. **e**, Stereo-view of the overlay in b. Isolated views of H3/H3 interface, **f**, and the CENP-A/CENP-A interface, **g**.

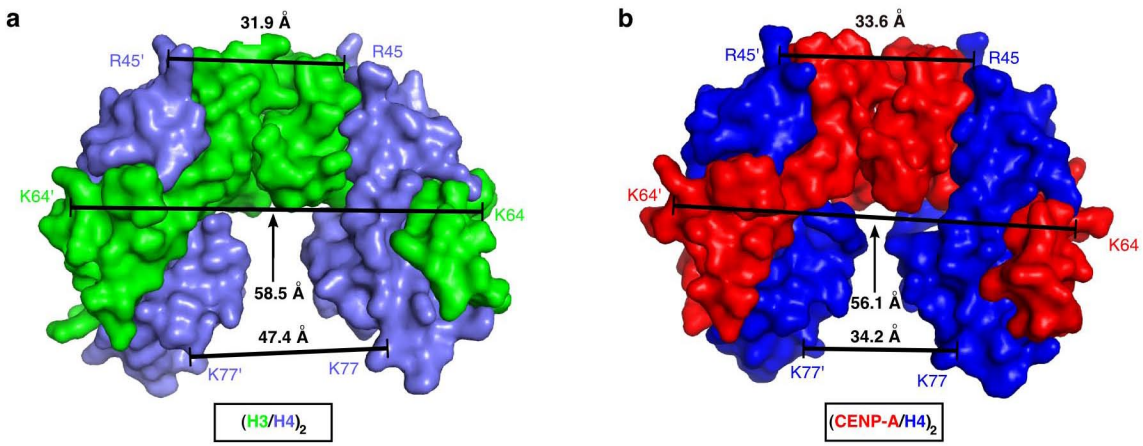

Figure S5 | The $(CENP\text{-}A/H4)_2$ heterotetramer is more compact than $(H3/H4)_2$ heterotetramer within the nucleosome. a and **b,** Surface representations and distances measurements for $(H3/H4)_2$ (a) and $(CENP\text{-}A/H4)_2$ (b).

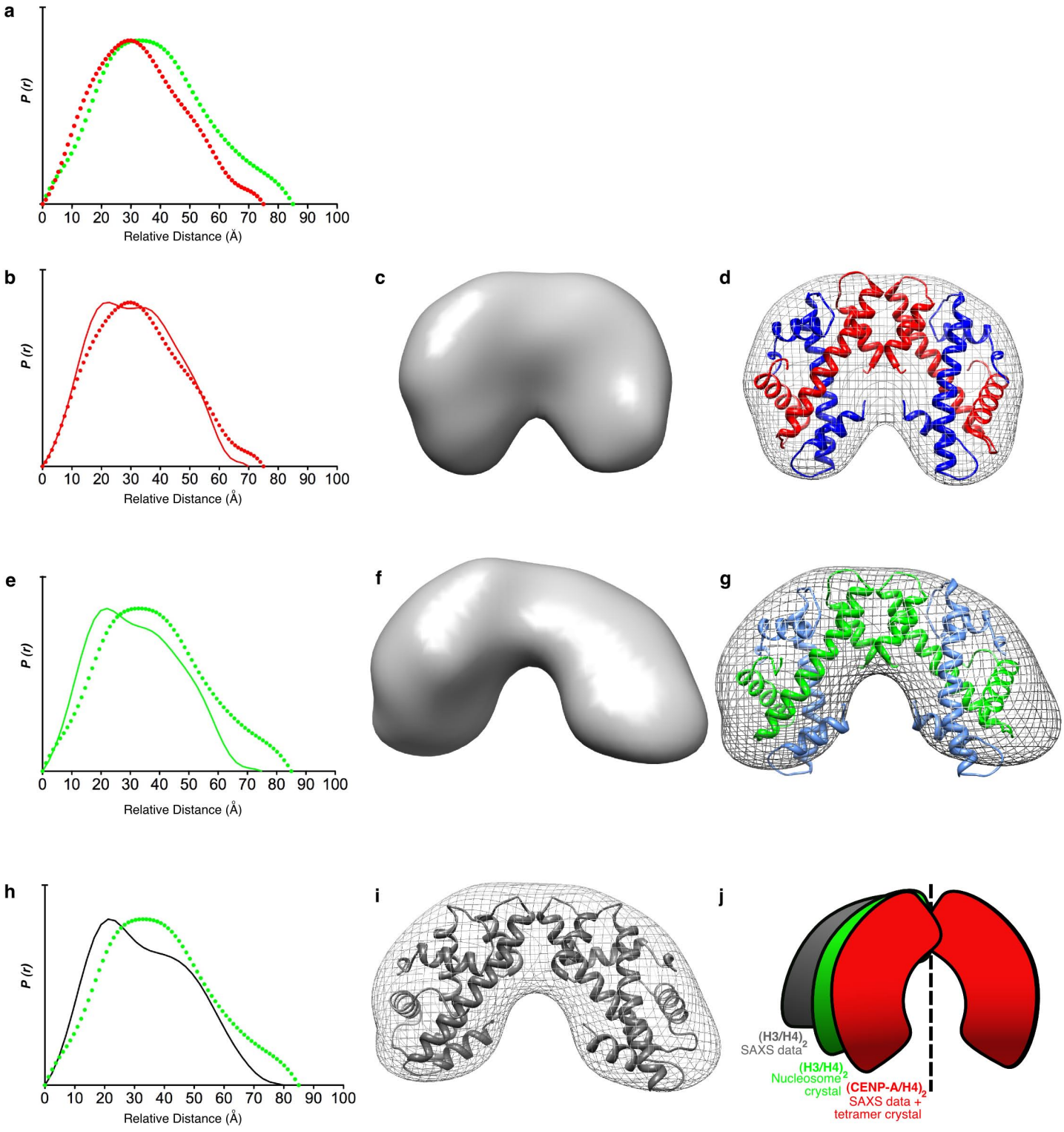

Figure S6 | SAXS measurements of $(CENP\text{-}A/H4)_2$ and $(H3/H4)_2$ heterotetramers. a, Pair-distance distribution (*P(r)*) curves describing the experimentally measured interatomic distances within $(CENP\text{-}A/H4)_2$ (red) or $(H3/H4)_2$ (green) determined by SAXS. **b**, Comparison of *P(r)* curves for $(CENP\text{-}A/H4)_2$ calculated from the crystal structure (solid line) or from the experimental data (solid circles). The slightly larger experimental D_{max} (75 Å) is expected and could be due to several factors, most likely an underestimation of protein shell hydration that is typical of calculations from crystal structures. **c**, SAXS envelope of $(CENP\text{-}A/H4)_2$ determined by averaging the 10 rotational states between the two CENP-A/H4

dimer halves of the tetramer that best fit the scattering data (see Methods for details). **d**, Docking of the $(CENP\text{-}A/H4)_2$ crystal structure onto the SAXS envelope (mesh). **e,** *P(r)* curves calculated from the nucleosomal $(H3/H4)_2$ crystal structure (solid line; from PDB ID 1KX5[2]) or SAXS measurement (solid circles). **f**, $(H3/H4)_2$ SAXS envelope obtained as in panel c. **g,** Docking of the nucleosomal $(H3/H4)_2$ crystal structure onto the SAXS envelope (mesh). We note from these SAXS experiments three interesting features. First, the measured D_{max} of $(CENP\text{-}A/H4)_2$ is indeed smaller than that of $(H3/H4)_2$, as predicted from the crystal structure. Second, the lack of features in the peaks of the measured *P(r)* curves for both tetramers, suggests that there is more flexibility in solution than in the crystal structure for both types of tetramers. Nevertheless, the averaged envelopes clearly show distinct favored states in solution: a closed conformation of $(CENP\text{-}A/H4)_2$, as in the crystal structure, and a much more open $(H3/H4)_2$ conformation. Third, the measurements for $(H3/H4)_2$ suggest that its favored rotated state in solution is an even more extended conformation than what is observed in the nucleosome structure. Note that in panel g, the crystal structure does not fit well into the favored rotational state measured by SAXS. We chose a model obtained directly from SASREF analysis as representative rotated state of $(H3/H4)_2$ heterotetramer that provides a better fit between the experimental curve (green circles) and calculated *P(r)* curve (black line), **h**, as well as into envelope (mesh) as shown in panel **i**. **j**, Cartoon summarizing rotational states of $(H3/H4)_2$ tetramer and $(CENP\text{-}A/H4)_2$ tetramer obtained by SAXS and crystallography.

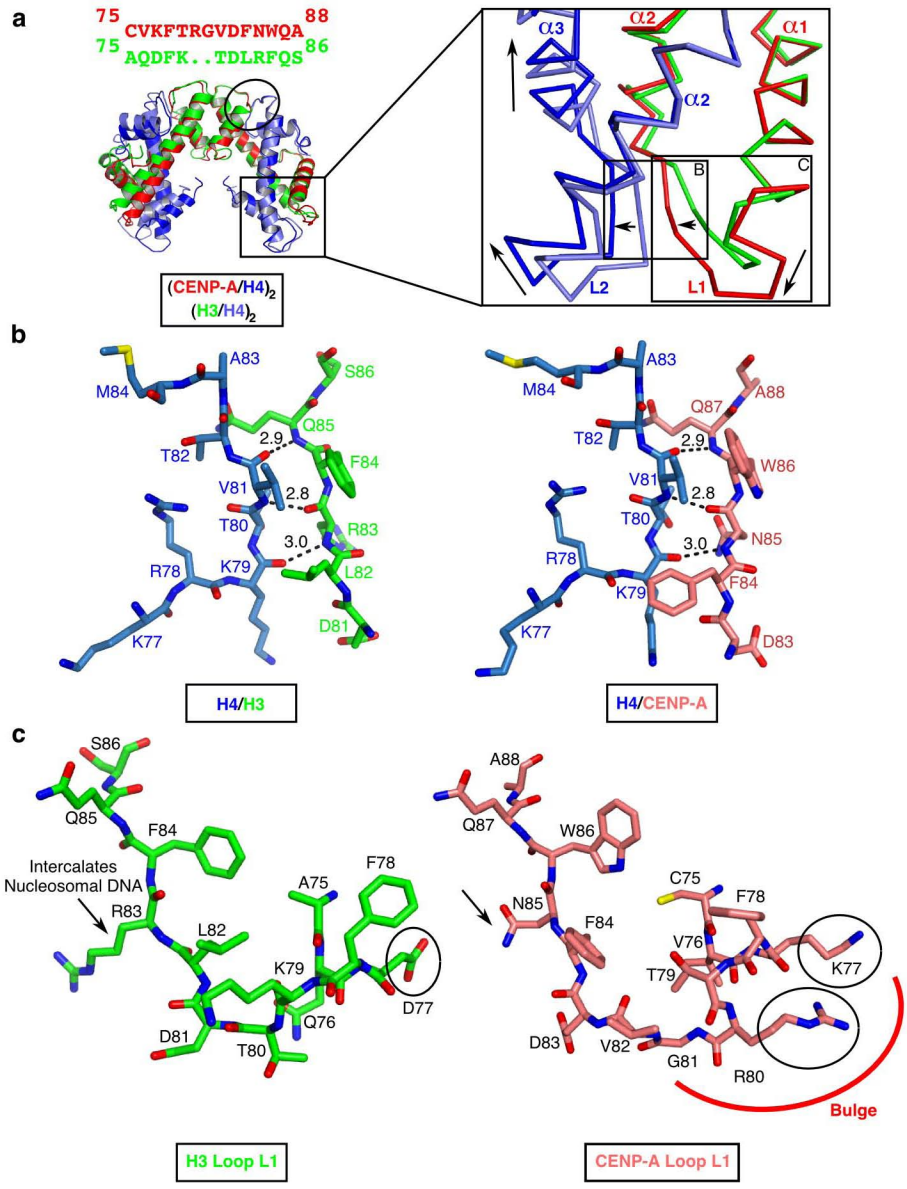

Figure S7 | Loop 1 of CENP-A changes the local charge and shape relative to the corresponding exposed surface in H3-containing nucleosomes. a, The enlargement shows an overlay based on α-carbon residues (positions 70-76 of each CENP-A and H3) within the α1 helix. As is the convention of histone fold domains, the interaction of histone pairs (e.g. H3 and H4) involves reciprocal interchain backbone loop L1/L2 interactions on either end of the antiparallel α2 helices of each subunit. While the interaction of L2 of CENP-A with L1 of H4 appears similar to the corresponding features of H3/H4 (circled region), the L1 of CENP-A contains two additional residues as compared to H3, an insertion that substantially alters local structure. The major consequence of the insertion is a bulge at the N-terminal portion of L1 of CENP-A. H4, when bound to either CENP-A or H3, undergoes displacement of its L2 that culminates in substantial curving of its α2 helix. In the case of CENP-A/H4 pairs, the local displacement of H4 features is slightly askew (1-2 Å difference highlighted by arrows) from that seen in the structure of the nucleosome[5]. The displacement is partially due to the longer L1 of CENP-A, but it should be noted that a similar direction and magnitude of alteration in both L2 and α3 helix of H4 is seen upon H3/H4 binding to the histone chaperone, Asf1[3,6] (Fig. S11). Thus, this minor local skewing in the sub-nucleosomal heterotetramer is a likely consequence of the absence of H2A/H2B dimers that bind to and slightly rotate the C-terminus of H4 in the context of the nucleosome[5]. **b**, Loop L2 of H4 and loop L1 of H3 (left) are connected through three backbone hydrogen bonds that are conserved in the structure of the (CENP-A/H4)$_2$ heterotetramer (right). **c,** Stick representation of the L1 loop of H3 (left) and CENP-A (right). Key

exposed charged residues are circled and Arg83 in H3 and its corresponding counterpart in CENP-A, Asn85, are indicated by arrows. The L1 of H3 also serves as a major DNA contact point in the nucleosome[5], with Arg83 penetrating into the minor groove. In CENP-A, this position is switched to asparagine (Asn85), but its relative position and side-chain orientation is nearly identical to that of H3 Arg83. This strongly implies that the mode and/or strength of DNA binding at the L1 contact site are altered in CENP-A-containing nucleosomes relative to their H3-containing counterparts.

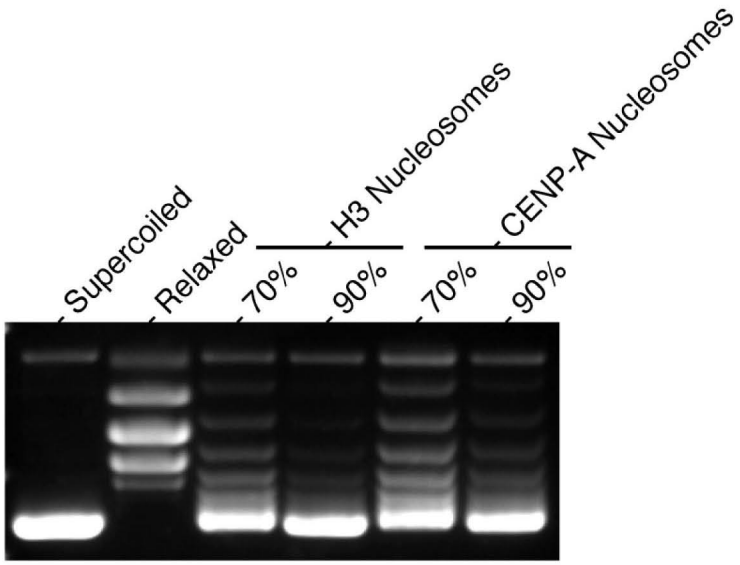

Figure S8 I Assembly of H3- and CENP-A-containing nucleosomes. Nucleosomes were assembled onto pUC19 with targeted nucleosome occupancies of 70 and 90%. The degree of supercoiling is consistent with nearly quantitative loading of nucleosomes. Additional supercoiling is present in assembly reactions targeted at 90%, either for H3- or CENP-A-containing nucleosomes assembled in parallel, indicative of an increase in total number of nucleosomes assembled on each template molecule.

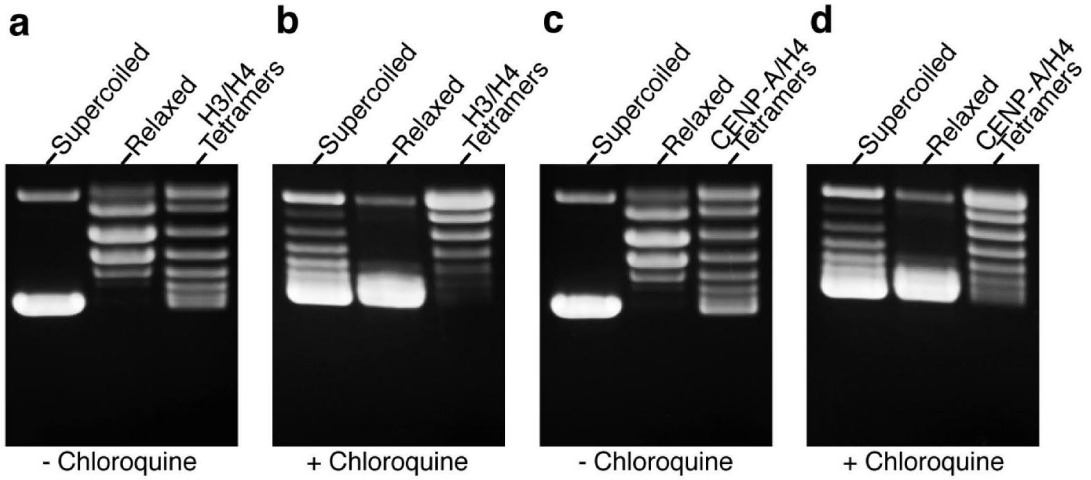

Figure S9. The $(CENP\text{-}A/H4)_2$ heterotetramer directs left-handed DNA wrapping in the absence of H2A/H2B dimers. H3-containing tetrasomes (**a** and **b**) and CENP-A-containing tetrasomes (**c** and **d**) were analyzed in the absence (a and c) or presence (b and d) of chloroquine. 70% $(H3/H4)_2$ targeted occupancy and 90% $(CENP\text{-}A/H4)_2$ targeted occupancy (assuming one tetrasome per 104 base pairs of DNA) yielded similar levels of supercoiling. Note that the reduced torsional constraints generated on the closed circular plasmid per $(CENP\text{-}A/H4)_2$ molecule (relative to $(H3/H4)_2$) is predicted by the substitution of a DNA intercalating arginine with asparagine (Fig. S7c) that weakens slightly the tetrasome entry/exit site. For both CENP-A-containing and H3-containing tetrasomes, chloroquine addition (b and d) relaxes the supercoils that they induce, indicating that the major form for both types of tetrasomes exist with left-handed DNA wrapping.

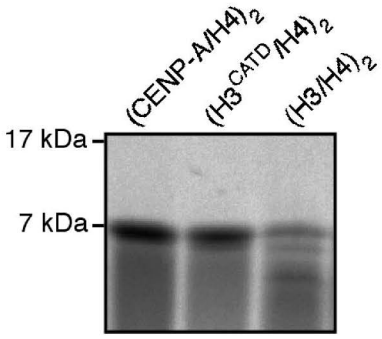

Figure S10 | Partial trypsinization of $H3^{CATD}$/H4 complex. Protein complexes indicated were incubated as in Figure S1 with 10 μg/ml trypsin for 6 hr.

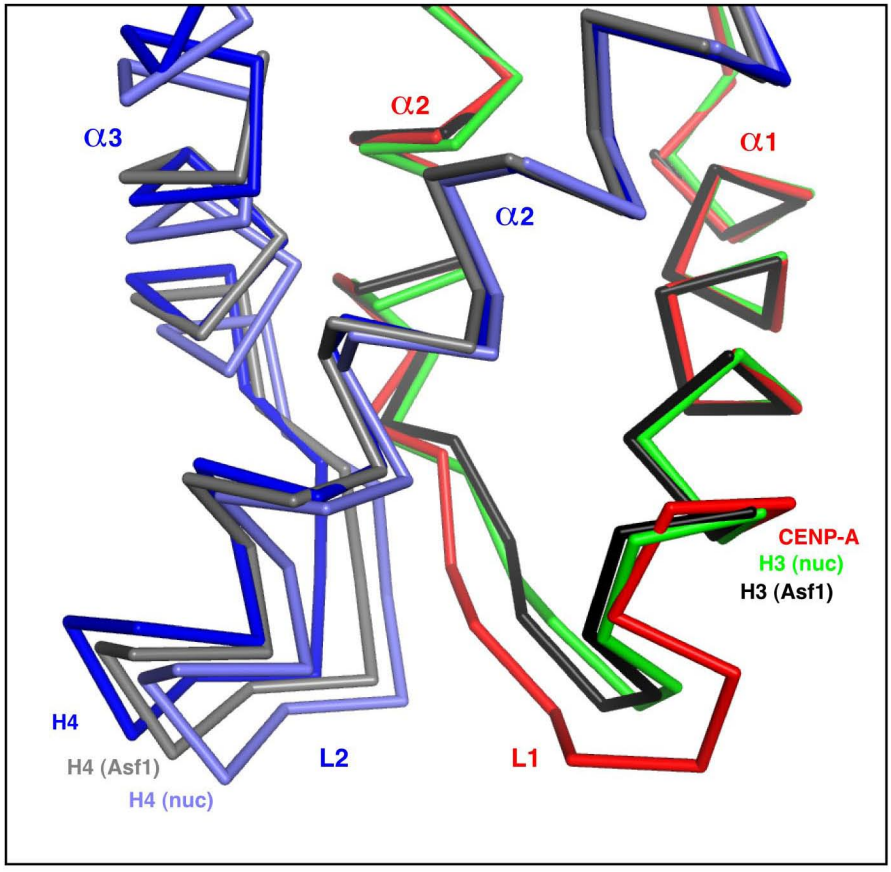

Figure S11 | Comparison of Loop L1 of $(CENP\text{-}A/H4)_2$ to the same loop of H3 either in the context of the nucleosome or in the ternary H3/H4/Asf1 complex. Structures of $(CENP\text{-}A/H4)_2$ tetramer, H3/H4 from within the nucleosome (PDB ID 1KX5[2]) and H3/H4 in complex with histone chaperone Asf1 (PDB ID 2HUE[3]) are overlaid on α-carbons of residues 70-76 (helix α1 of CENP-A or H3). H3 from both structures overlay almost identically in the region of loop L1 in the nucleosome or Asf1 complex. Loop L2 of H4, however is skewed away from H3 L1 in the Asf1 structure relative to the nucleosome. The bulged L1 from CENP-A moves towards the L2 from H4, which is skewed in a similar manner and degree as in the H3/H4/Asf1 complex.

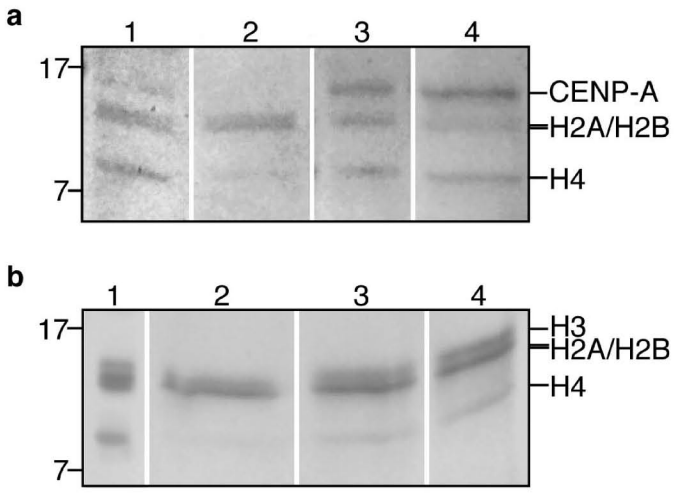

Figure S12. Stability of CENP-A-containing nucleosomes (a) and H3-containing nucleosomes (b) measured by salt dissociation. Nucleosomes were reconstituted as in Figure 4 and then absorbed to hydroxyapatite and eluted with a salt gradient using a standard approach to assess nucleosome stability[7]. Both types of nucleosomes completely dissociate in the range of 1.5-1.8 M NaCl, similar to reports from conventional native chromatin[7], with H2A/H2B dimers eluting prior to tetramers in both cases. Both panels a and b are silver stained SDS-PAGE gels, with lane 1 containing the sample applied to hydroxyapatite and lanes 2-4 showing fractions containing the dissociating histones. Conductance was measured in the fractions and correspond to the following NaCl concentrations: panel a, lane 2 = 1.54 M, lane 3 = 1.64 M, lane 4 = 1.73 M; panel b, lane 2 = 1.59 M, lane 3 = 1.64 M, lane 4 = 1.69 M.

References for Supplementary Information

1 Black, B.E. *et al.*, Structural determinants for generating centromeric chromatin. *Nature* 430 (6999), 578-582 (2004).

2 Davey, C.A., Sargent, D.F., Luger, K., Maeder, A.W., & Richmond, T.J., Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 a resolution. *J Mol Biol* 319 (5), 1097-1113 (2002).

3 English, C.M., Adkins, M.W., Carson, J.J., Churchill, M.E., & Tyler, J.K., Structural basis for the histone chaperone activity of Asf1. *Cell* 127 (3), 495-508 (2006).

4 Wood, C.M. *et al.*, High-resolution structure of the native histone octamer. *Acta Crystallogr Sect F Struct Biol Cryst Commun* 61 (Pt 6), 541-545 (2005).

5 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 (6648), 251-260 (1997).

6 Natsume, R. *et al.*, Structure and function of the histone chaperone CIA/ASF1 complexed with histones H3 and H4. *Nature* 446 (7133), 338-341 (2007).

7 Thambirajah, A.A. *et al.*, H2A.Z stabilizes chromatin in a way that is dependent on core histone acetylation. *J Biol Chem* 281 (29), 20036-20044 (2006).