Supplementary Information

Identification of Functionally Conserved Regions in the Structure of the Chaperone/CenH3/H4 Complex

Jingjun Hong, Hanqiao Feng, Zheng Zhou, Rodolfo Ghirlando, Yawen Bai

Materials and methods

Gene cloning, expression, and purification of histones and chaperones

The cloning, expression and purification of His6-KK-Cse4$^{Sc}_{151-229}$, H4$^{Sc}_{1-103}$, and Scm3$^{Sc}_{84-169}$ were described in our earlier work (11).

Reconstitution of protein complexes

The heterotrimeric complex Scm3$^{Sc}_{84-169}$/His6-KK-Cse4$^{Sc}_{151-229}$/H4$^{Sc}_{1-103}$ in 50 mM MES (pH 5.6) was prepared as described in our earlier work (11). To make the complex containing His6-KK-Cse4$^{Sc}_{151-229}$ and H4$^{Sc}_{1-103}$, lyophilized proteins were denatured in 6 M GudHCl and dialyze d overnight against 10 mM Tris-HCl, 500 mM NaCl (pH 7.4) at 4°C. After purification by gel filtration under the same conditions, the sample was exchanged to different buffers: 50 mM MES (pH 5.6), 50 mM MES (pH 5.6) and 150 mM NaCl, 50 mM phosphate (pH 6.5) and 150 mM NaCl, 50 mM phosphate (pH 7.4) and 150 mM NaCl, respectively.

Analytical ultracentrifugation experiments

The sample containing GS-HJURP$^{Hs}_{2-81}$, GS-CENP-A$^{Hs}_{60-140}$-LVPR and H4$^{Hs}_{1-103}$ was studied in 50 mM MES (pH 5.6), 5 mM βME at a loading concentration of 0.68 A_{280} (31.5 μM), whereas the sample containing GS-HJURP$^{Hs}_{2-81}$, GS-CENP-A$^{Hs}_{60-140}$-LVPR and GS-H4$^{Hs}_{22-103}$ was studied in 150 mM NaCl, 50 mM phosphate (pH 7.0), 5 mM βME at a loading concentration of 1.00 A_{280} (46.5 μM). Moreover, the sample containing GS-CENP-A$^{Hs}_{60-140}$-LVPR and GS-H4$^{Hs}_{22-103}$ was studied in 50 mM MES (pH 5.6) and 5 mM βME and in 50 mM NaAc (pH 4.5) and 5 mM βME at a dimer concentration of 77 μM. Samples containing His6-

KK-Cse4$^{Sc}_{151-229}$ and H4$^{Sc}_{1-103}$ were studied at a dimer concentration of 46 μM in different buffers: 50 mM MES (pH 5.6); 50 mM MES (pH 5.6) and 150 mM NaCl; 50 mM phosphate (pH 6.5) and 150 mM NaCl; 50 mM phosphate (pH 7.4) and 150 mM NaCl.

Sedimentation velocity experiments were conducted at 20.0°C on a Beckman Optima XL-A or Beckman Coulter ProteomeLab XL-I analytical ultracentrifuge. 400 μl of each of the samples were loaded in 2-channel centerpiece cells and analyzed at a rotor speed of 50 krpm with scans collected at approximately 7.5 minute intervals. Data were collected using the absorbance optical system in continuous mode at 280 nm using radial step size of 0.003 cm. Data were also collected using the Rayleigh interference detection system when possible. All data were analyzed in SEDFIT 12.7 (28) in terms of a continuous *c(s)* distribution of Lamm equation solutions covering a maximal $s_{20,w}$ range of 0 - 10 with a resolution of 20 points per S and a confidence level of 0.68. Excellent fits were obtained with RMSD values ranging from 0.0031 – 0.0054 absorbance units. Solution densities ρ and viscosities η were measured as described (11), or calculated in SEDNTERP 1.09 (29). Sedimentation coefficients *s* were corrected to $s_{20,w}$.

NMR experiments

NMR experiments were performed on Bruker 500 and 700 MHz spectrometers at 25°C (human HJURP/CENP-A/H4, *Sc* Scm3/Cse4/H4, *Kl* Cse4/H4). The following experiments were recorded: [^{1}H, ^{15}N]-HSQC, 3D CBCANH,

CBCA(CO)NH, HNCO, HN(CA)CO, HNCA, HN(CO)CA; TROSY version: [^{1}H, ^{15}N]-TROSY, 3D HNCACB, HN(CO)CACB, HNCA, HN(CO)CA, HNCO, HN(CA)CO.

Table S1. A list of constructs used in this study

---

1. His6-Cse4$^{KI}_{106-183}$-LVPRGS-H4$^{KI}_{30-103}$
2. His6-Cse4$^{KI}_{106-183}$ (L158G/L159G/L161G/H162G)-LVPRGS-H4$^{KI}_{30-103}$
3. H3$^{KI}_{61-136}$
4. H3$^{KI}_{61-136}$ (S88M/G91M/S96A/V97S)
5. H4$^{KI}_{1-103}$
6. His6-SSGLVPRGS-Scm3$^{KI}_{43-143}$
7. His6-KK-Cse4$^{Sc}_{151-207}$-LVPRGS-H4$^{Sc}_{45-103}$
8. His6-KK-Cse4$^{Sc}_{151-229}$
9. H4$^{Sc}_{1-103}$
10. Scm3$^{Sc}_{84-169}$
11. His6-SSGLVPRGS-CENP-A$^{Hs}_{60-140}$-LVPRGS-HJURP$^{Hs}_{2-81}$
12. H3.1$^{Hs}_{1-136}$-His6
13. H3.1$^{Hs}_{1-136}$ (S88Q/G103H)-His6
14. H4$^{Hs}_{1-103}$
15. H4$^{Hs}_{1-15}$-LVPRGS-H4$^{Hs}_{22-103}$

Table S2. Protein samples used in this work

NMR samples

1. GS-Scm3$^{KI}_{43-143}$ / His6-Cse4$^{KI}_{106-183}$-LVPR / GS-H4$^{KI}_{30-103}$

2. His6-KK-Cse4$^{Sc}_{151-207}$-LVPRGS-H4$^{Sc}_{45-103}$

3. His6-KK-Cse4$^{Sc}_{151-207}$-LVPR / GS-H4$^{Sc}_{45-103}$

4. Scm3$^{Sc}_{84-169}$ / His6-KK-Cse4$^{Sc}_{151-229}$ / H4$^{Sc}_{1-103}$

5. GS-HJURP$^{Hs}_{2-81}$ / GS-CENP-A$^{Hs}_{60-140}$-LVPR / H4$^{Hs}_{1-103}$

6. GS-HJURP$^{Hs}_{2-81}$ / GS-CENP-A$^{Hs}_{60-140}$-LVPR / GS-H4$^{Hs}_{22-103}$

7. (GS-CENP-A$^{Hs}_{60-140}$-LVPR / GS-H4$^{Hs}_{22-103}$)$_2$

Sedimentation velocity samples

1. GS-Scm3$^{KI}_{43-143}$ / His6-Cse4$^{KI}_{106-183}$ (L158G/L159G/L161G/H162G)-LVPR / GS-H4$^{KI}_{30-103}$

2. His6-Cse4$^{KI}_{106-183}$-LVPR / GS-H4$^{KI}_{30-103}$

3. His6-KK-Cse4$^{Sc}_{151-229}$ / H4$^{Sc}_{1-103}$

4. GS-HJURP$^{Hs}_{2-81}$ / GS-CENP-A$^{Hs}_{60-140}$-LVPR / H4$^{Hs}_{1-103}$

5. GS-HJURP$^{Hs}_{2-81}$ / GS-CENP-A$^{Hs}_{60-140}$-LVPR / GS-H4$^{Hs}_{22-103}$

6. GS-CENP-A$^{Hs}_{60-140}$-LVPR / GS-H4$^{Hs}_{22-103}$

The proteins described in the text are shown in blue.

Table S3. Parameters for the folded regions ($Cse4_{157-197}$ and $H4_{51-94}$) in the single-chain model (pdb ID: 2LY8).

NMR distance and dihedral constraints	
Total NOE	3382
Intra residues NOE	912
Sequential NOE (|i-j| = 1)	853
Medium range NOE (|i-j| < 5)	1151
Long range NOE (|i-j| > 4)	466
H-bonds	118
Dihedral angles (Φ, Ψ) (TALOS)	190
Structure statistics	
Violations (mean ± s.d.)	
Distance constraints (Å)	0.065 ± 0.003
Dihedral constraints (Å)	0.441 ± 0.02
Deviation from idealized geometry	
Bonds (Å)	0.005 ± 0.0001
Angles (^{o})	0.650 ± 0.025
Impropers (^{o})	0.463 ± 0.030
Average pairwise r.m.s deviation (Å)	
Backbone atoms	0.20
All heavy atoms	0.78
Procheck statistics	
Residues in	
most favored regions	85.6%
additional allowed regions	11.4%
generously allowed regions	2.2%
disallowed regions	0.8%

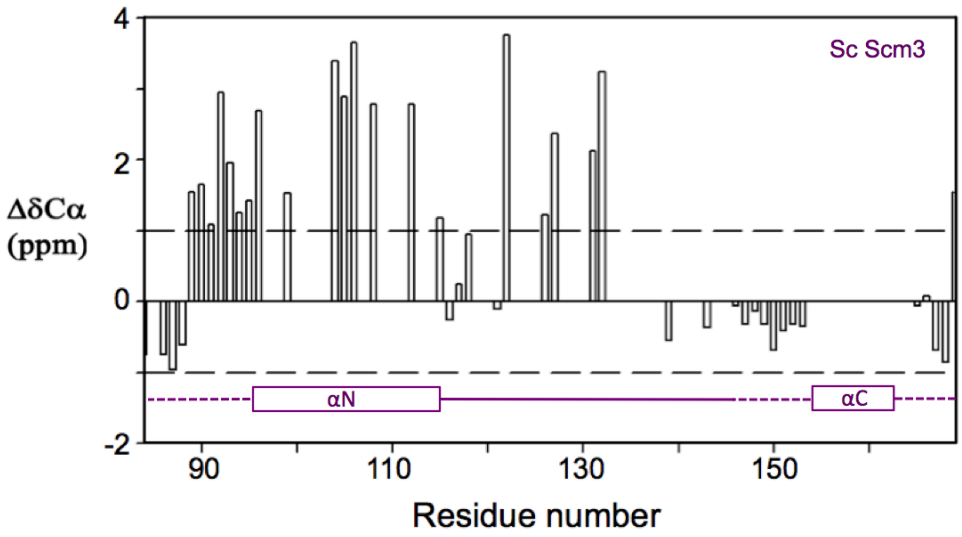

Fig. S1. Scm3 forms a longer helix at lower temperature. Deviations of the Cα chemical shifts of the *Sc* Scm3 from random coil values in the three-chain Scm3/Cse4/H4 complex at 25°C. The structured regions of Scm3 in the NMR structure of the single-chain Sc Scm3/Cse4/H4 complex are shown in solid lines and the unstructured regions are shown in dashed lines.

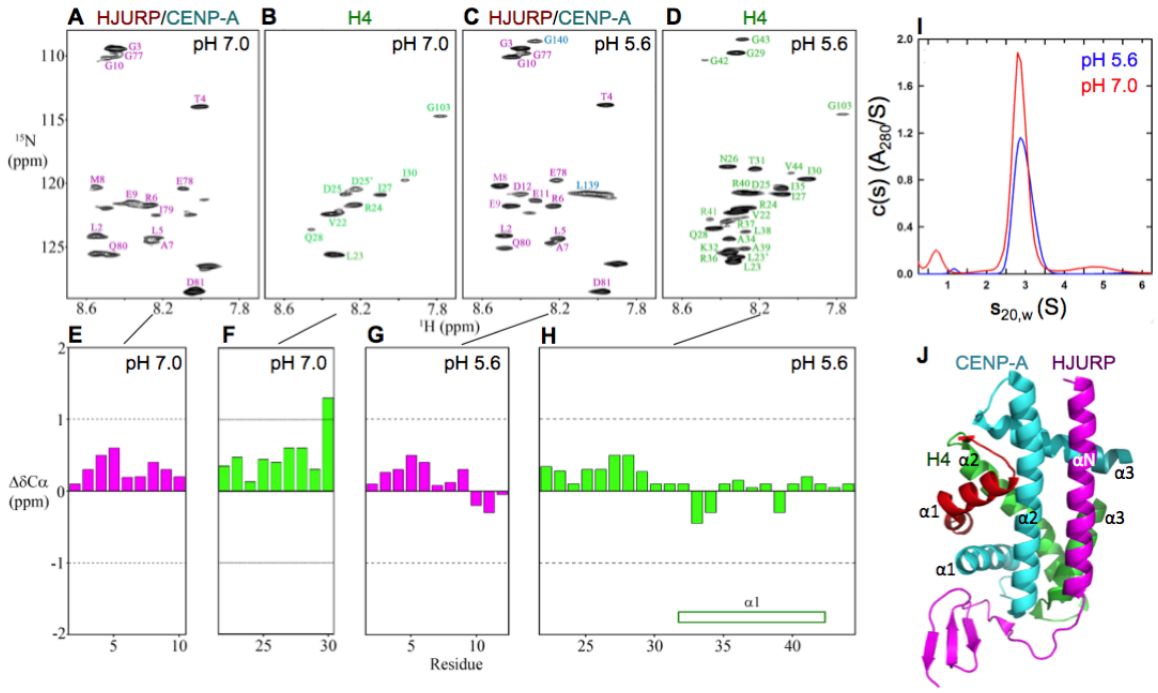

Fig. S2. pH-dependent structure of the HJURP/CENP-A/H4 complex as revealed by NMR. (*A, B*) ^{1}H-^{15}N HSQC spectra of ^{15}N/^{13}C-labeled HJURP/CENP-A in complex with non-labeled H4 (*A*) and ^{15}N/^{13}C-labeled H4 in complex with non-labeled HJURP/CENP-A (*B*) at pH 7.0. (*C, D*) ^{1}H-^{15}N HSQC spectra of ^{15}N/^{13}C-labeled HJURP/CENP-A in complex with non-labeled H4 (*C*) and ^{15}N/^{13}C-labeled H4 in complex with non-labeled HJURP/CENP-A (*D*) at pH 5.6. (*E, F*) Deviations of the Cα Chemical shifts of the N-terminal regions of HJURP (*E*) and H4 (*F*) from random coil values at pH 7.0. (*G, H*) Deviations of the Cα Chemical shifts of the N-terminal region of HJURP (*G*) and the α1 helix of H4 (*H*) from random coil values at pH 5.6. (*I*) Analytical ultracentrifugation on the HJURP/CENP-A/H4 complex at pH 5.6 and 7.0 revealing the presence of a major species at 2.85 *S* with a molar mass ranging from 29 kDa (pH 7.0) to 33 kDa (pH 5.6) indicative of a heterotrimer. (*J*) Mapping of the unfolded α1 helix (red) on the crystal structure of the HJURP/CENP-A/H4 complex. (Note that at pH 7.0, the *KI* Scm3/Cse4/H4 does not produce workable NMR spectra).

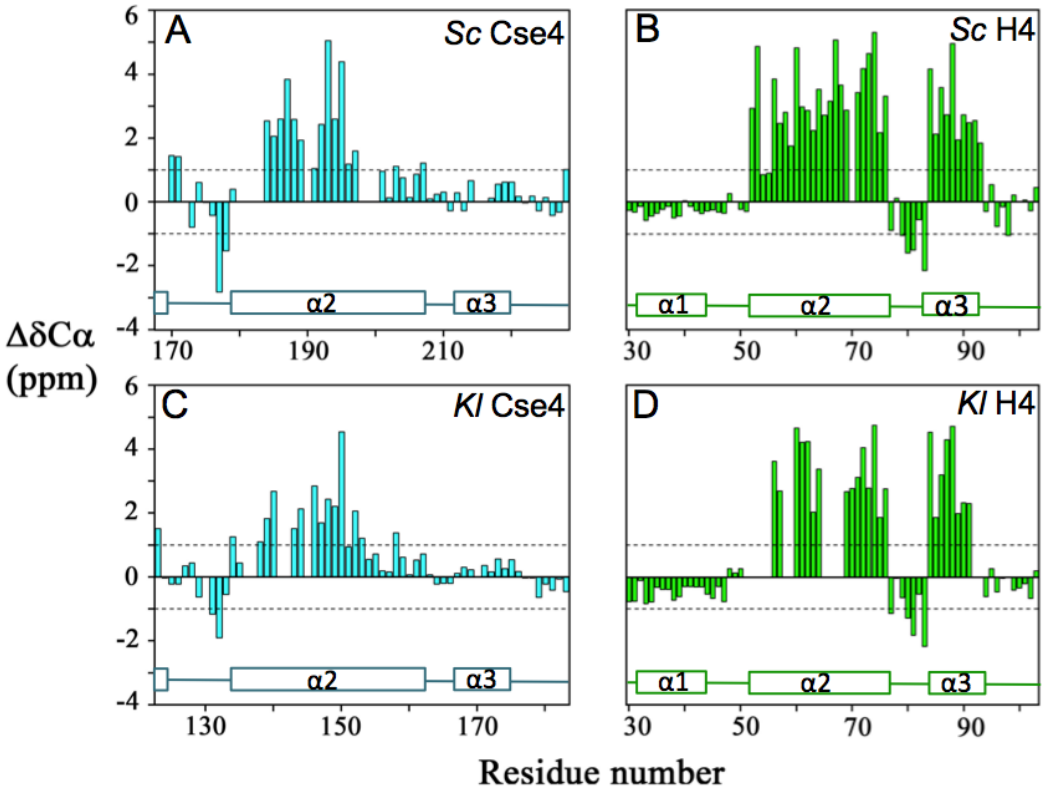

Fig. S3. The C-terminal region of the α2 helix and the α3 helix of Cse4 and the α1 helix of H4 are unstructured in the Cse4/H4 dimer. Deviations of the Cα chemical shifts of the Cse4/H4 dimer from random coil values. (*A*) *Sc* Cse4 (11). (*B*) Sc H4 (11). (*C*) *Kl* Cse4. (*D*) *Kl* H4. The α1 helix region of *Kl* Cse4 was not assigned.

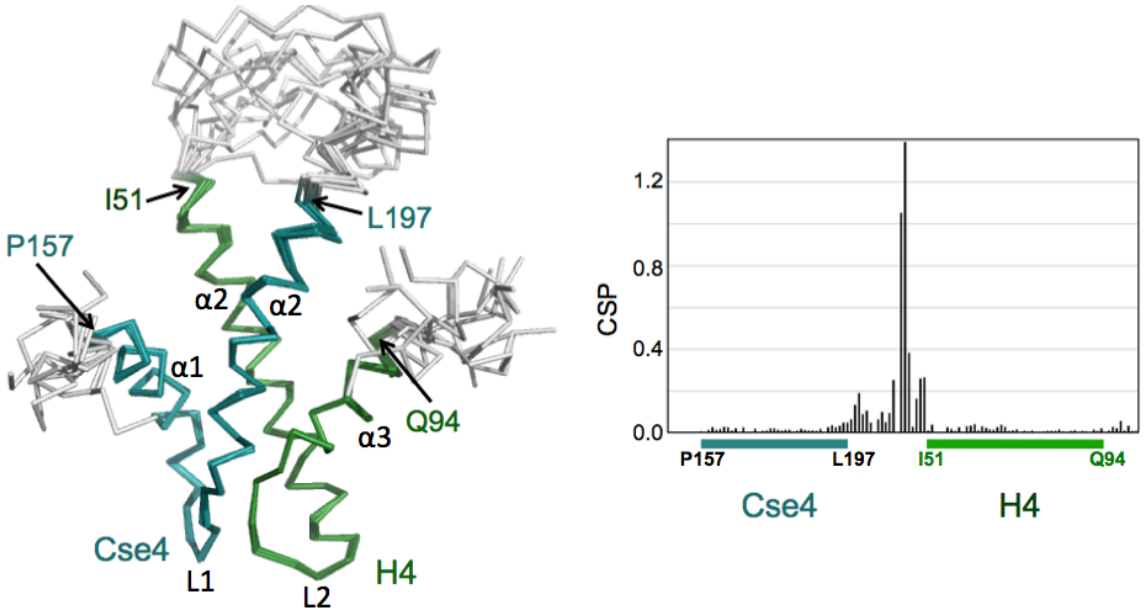

Fig. S4. Structure of the single-chain *Sc* Cse4/H4. (*A*) Overlay of the 10 calculated structures with the lowest energy on the folded core (P157-L197 of Cse4 and I51-Q94 of H4). (*B*) The chemical shift perturbation (CSP) between the single-chain model and two-chain complex that is obtained by cutting the linker with thrombin protease. CSP is calculated with $\{(\Delta\delta^{1}H)^{2} + (\Delta\delta^{15}N/5)^{2}\}^{½}$ for each residue. There are only small chemical shift changes in the folded region of the Cse4/H4 dimer, indicating that the single-chain model does not perturb the folded core of the Cse4/H4 dimer. Note that there are 15 unfolded residues linking the α2 helices of Cse4 and H4.

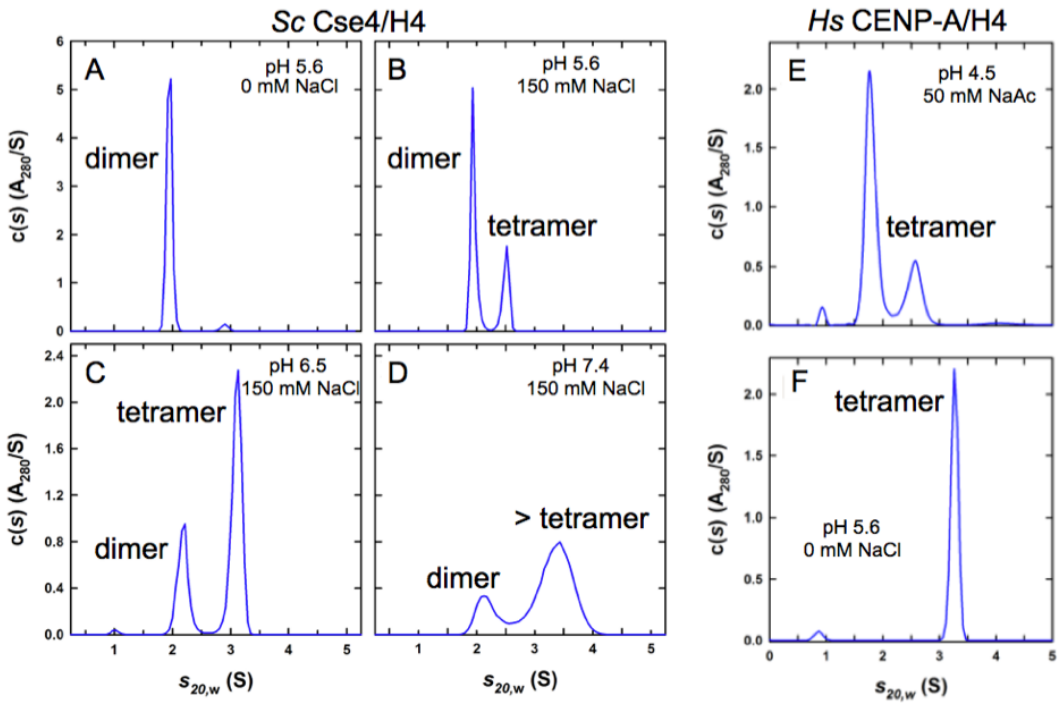

Fig. S5. pH- and salt-dependent equilibria of the CenH3/H4 dimer and the $(CenH3/H4)_2$ tetramer. (*A*-*D*) *Sc* Cse4/H4 dimer and $(Cse4/H4)_2$ tetramer. (*E*-*F*) Human CENP-A/H4 dimer and $(CENP\text{-}A/H4)_2$ tetramer.