CHROMOSOME STRUCTURE

Structure of histone-based chromatin in Archaea

Francesca Mattiroli,[1]* Sudipta Bhattacharyya,[2]*† Pamela N. Dyer,[1] Alison E. White,[1] Kathleen Sandman,[3] Brett W. Burkhart,[2] Kyle R. Byrne,[2] Thomas Lee,[1] Natalie G. Ahn,[1] Thomas J. Santangelo,[2,4] John N. Reeve,[3] Karolin Luger[1,4,5]‡

Small basic proteins present in most Archaea share a common ancestor with the eukaryotic core histones. We report the crystal structure of an archaeal histone-DNA complex. DNA wraps around an extended polymer, formed by archaeal histone homodimers, in a quasi-continuous superhelix with the same geometry as DNA in the eukaryotic nucleosome. Substitutions of a conserved glycine at the interface of adjacent protein layers destabilize archaeal chromatin, reduce growth rate, and impair transcription regulation, confirming the biological importance of the polymeric structure. Our data establish that the histone-based mechanism of DNA compaction predates the nucleosome, illuminating the origin of the nucleosome.

The nucleosome consists of two (H2A-H2B) and two (H3-H4) histone heterodimers assembled as an octamer that wraps 147 base pairs (bp) of DNA in 1.65 negative superhelical turns (1). Histones, the most conserved proteins known, all have a central "histone fold" (HF) dimerization motif formed by three α helices separated by two short loops (fig. S1A). Small HF-containing proteins, present in most Archaea, likely share a common ancestor with the eukaryotic histones (2–4). Hundreds of different archaeal histone sequences are now known [fig. S1B; (5, 6)]. Most are 70 ± 5 amino acids long and lack HF extensions and the basic histone tails, which are the segments specific to each eukaryotic histone that contribute to nucleosome stability and gene regulation [fig. S1A; (3, 7)]. Unlike the mandatory eukaryotic histone heterodimer partnerships, archaeal histones homodimerize and form heterodimers with related paralogs. Here we report the structure of archaeal histone-based chromatin and its participation in gene expression.

To obtain crystals, we used a DNA sequence to which homodimers of histone B from *Methanothermus fervidus* [(HMfB)$_2$] bind at defined locations (8, 9). In the 4 Å crystal structure (table S1), this 90-bp DNA wraps around three (HMfB)$_2$ dimers (Fig. 1A) that are virtually identical when compared to each other, to (HMfB)$_2$ dimers in the absence of DNA [(10); root mean square deviation (RMSD) 0.36 Å], and to the HFs of eukaryotic (H3-H4) and (H2A-H2B) heterodimers (RMSD ~1.7 Å; Fig. 1, A and B,

and fig. S2A). Each HF dimer (HFD) interacts with the DNA in a very similar fashion to the eukaryotic HFDs, with fully conserved amino acid side-chain interactions [RT pair and RD clamp (R, arginine; T, threonine; D, aspartic acid); Fig. 1 and fig. S2, A and B] that mutagenesis studies have confirmed are essential for DNA binding by HMfB (11, 12). Intramolecular hydrogen bonds between the two histones in the (HMfB)$_2$ dimer position the α1 helices and N termini for optimal interaction with DNA and direct an N-terminal extension appropriately through the gyres of the surrounding DNA, as seen in H2A and H3 in the nucleosome (fig. S2C) (7).

The (HMfB)$_2$ dimers are symmetric and, in the crystal lattice, polymerize through identical four–α helix bundles (4HBs; Fig. 2A) to form a continuous helical ramp (Fig. 2B). The geometry of the 4HB is conserved between HMfB-HMfB′, H3-H3′, and H4-H2B (Fig. 2A), and therefore, the arrangement of any four consecutive archaeal HFDs in the crystal structure is markedly similar to the assembly of the four HFDs in the nucleosome octamer (RMSD 2.0 Å, Fig. 2C). The surface of the complex formed by archaeal histones has, however, less positive charge (Fig. 2D).

In the crystal lattice, DNA wraps around the HMfB protein assembly in a quasi-continuous superhelix, through annealing of the 2–nucleotide (nt) 5′ overhangs (Fig. 2E). The geometry, diameter, pitch, and writhe of this superhelix, and the spacing between gyres, strongly resemble the nucleosomal DNA arrangement (Fig. 2F). Consequently, the alignment of DNA grooves (80 bp apart on linear DNA) across two gyres of DNA, termed nucleosomal "supergrooves" (13), is also conserved (arrows in Fig. 2E). The ability of archaeal histones to form polymers was also validated in solution by using (HMfB)$_2$ dimers and (HTkA)$_2$ dimers of histone A from *Thermococcus kodakarensis*, confirming that this arrangement is not a crystallographic artifact (fig. S3 and table S2). Both (HMfB)$_2$ and (HTkA)$_2$ form complexes that protect 60-, 90-, 120-, 150-, and 180-bp fragments from micrococcal nuclease (MNase) digestion (fig. S3A), consistent with previous reports

(8, 14, 15, 16). Ultracentrifugation further confirmed that the complexes formed on 147- and 207-bp DNA molecules contain the predicted number of archaeal histone dimers needed to saturate these DNAs (fig. S3B and table S2). By contrast, the polymerization of eukaryotic histone dimers is limited by their asymmetry to an octamer (Fig. 2A, right panel). Notably, the interactions within the archaeal superhelix do not resemble any of the nucleosome-nucleosome stacking interactions reported so far [reviewed in (17); see also (18, 19)].

To investigate if the extended polymerization has functional importance, we sought to destabilize the superhelix in vivo without compromising the DNA-binding ability of the archaeal histone. Apart from the 4HBs, the only region of close contact between the adjacent layers of the archaeal histone polymer is where the L1 loops of dimers 1 and 4 meet (arrow in Fig. 2B and fig. S4A), a position almost always occupied by a glycine (G16 in HMfB, G17 in HTkA; fig. S1B). To determine if the absence of a side chain facilitates this close packing, we generated *T. kodakarensis* strains isogenic except for G17 substitutions in HTkA, the single histone present and essential for *T. kodakarensis* strain TS600 viability [(18, 20); fig. S5]. Cells with wild-type HTkA, transferred from a sulfur (S°)–containing to a S°-free medium (pyruvate), restart growth after ~4 hours, during which time they reprogram gene expression [Fig. 3A; (21)]. The otherwise isogenic strains with HTkA G17H, G17D, G17N, G17L, or G17S also grew normally in S° but took longer to restart growth when transferred to medium lacking S°, and some also grew slower (Fig. 3A and table S3) (H, histidine; N, asparagine; L, leucine; S, serine). Given the delayed response to nutrient change, we investigated transcription of the media-dependent membrane-bound hydrogenase (MBH)–encoding operon (comprising genes TK2080 to TK2093). As previously established (21), transcription of this operon was elevated in *T. kodakarensis* TS600 with wild-type HTkA when grown in the absence of S°, but this was not the case for TS621, the strain with HTkA G17L (Fig. 3B), indicating a deregulated transcriptional program.

To determine if the negative effects of the G17 substitutions on MBH expression correlated with changes in chromatin structure, chromatin isolated from strains containing HTkA (TS600), HTkA G17L (TS621), and HTkA G17D (TS620), grown with or without S°, was subjected to MNase digestion. As previously observed, chromatin from TS600 protected fragments ranging from 60 to ~300 bp, in increments of ~30 bp (14); the most prominent band was 120 bp, corresponding to protection by four (HTkA)$_2$ dimers (Fig. 3C and fig. S4B). By contrast, digestion of chromatin from TS621 and TS620 generated only ~60- and ~90-bp protected fragments (Fig. 3C and figs. S4B and S6A). Both SDS–polyacrylamide gel electrophoresis and liquid chromatography–tandem mass spectrometry confirmed that the intracellular concentrations of HTkA, HTkA G17L, and HTkA G17D were similar (fig. S6B and table S4). This is consistent with the differences in MNase protection

[1]Department of Chemistry and Biochemistry, University of Colorado Boulder, Boulder, CO 80309, USA. [2]Department of Biochemistry and Molecular Biology, Colorado State University, Fort Collins, CO 80523, USA. [3]Department of Microbiology, Ohio State University, Columbus, OH 43210, USA. [4]Institute for Genome Architecture and Function, Colorado State University, Fort Collins, CO 80523, USA. [5]Howard Hughes Medical Institute, University of Colorado Boulder, Boulder, CO 80309, USA.
*These authors contributed equally to this work.
†Present address: Department of Molecular Biosciences, University of Texas at Austin, Austin, TX 78712, USA.
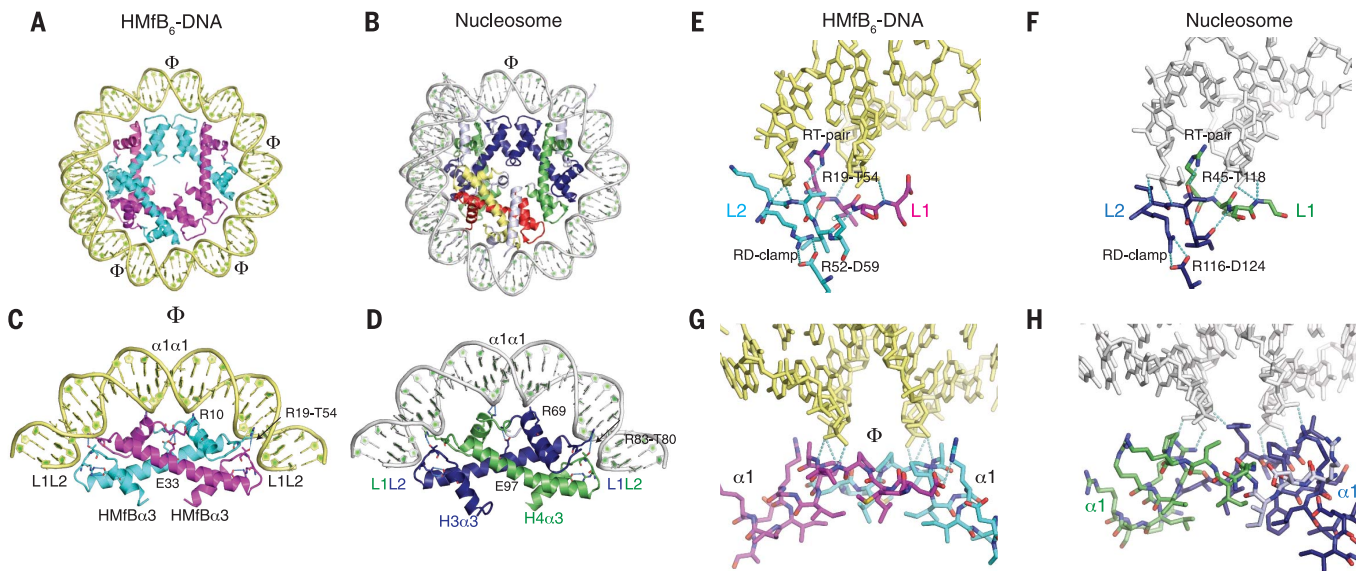‡Corresponding author. Email: karolin.luger@colorado.edu

Fig. 1. DNA binding is conserved between archaeal and eukaryotic histones. (**A**) The structure of three (HMfB)$_2$ dimers bound to a 90-bp SELEX DNA is highly similar to the (**B**) nucleosome hexasome, shown by removing one H2A-H2B heterodimer and the histone tails from the published nucleosome structure (1AOI). The axes of symmetry in both protein assemblies are indicated (Φ). SELEX, systematic evolution of ligands by exponential enrichment. (**C**) HFs of an (HMfB)$_2$ dimer and (**D**) (H3-H4) heterodimer shown in the same orientation with associated DNA. E, glutamic acid. (**E**) The L1L2 interface of an (HMfB)$_2$ dimer and (**F**) an (H3-H4) dimer is shown with conserved interactions with DNA. (**G**) The α1α1 interface in an (HMfB)$_2$ dimer and (**H**) in an (H3-H4) dimer. Further comparisons of the structures formed by HMfB and eukaryotic histones with DNA are shown in fig. S2. In all figures, although identical, the two HMfB monomers in an (HMfB)$_2$ dimer are colored in cyan and magenta; H3 is blue; H4 is green; H2A is yellow; H2B is red. Regions of core histones that are not part of the histone fold are shown in white. DNA organized by HMfB is pale yellow; nucleosomal DNA is gray.
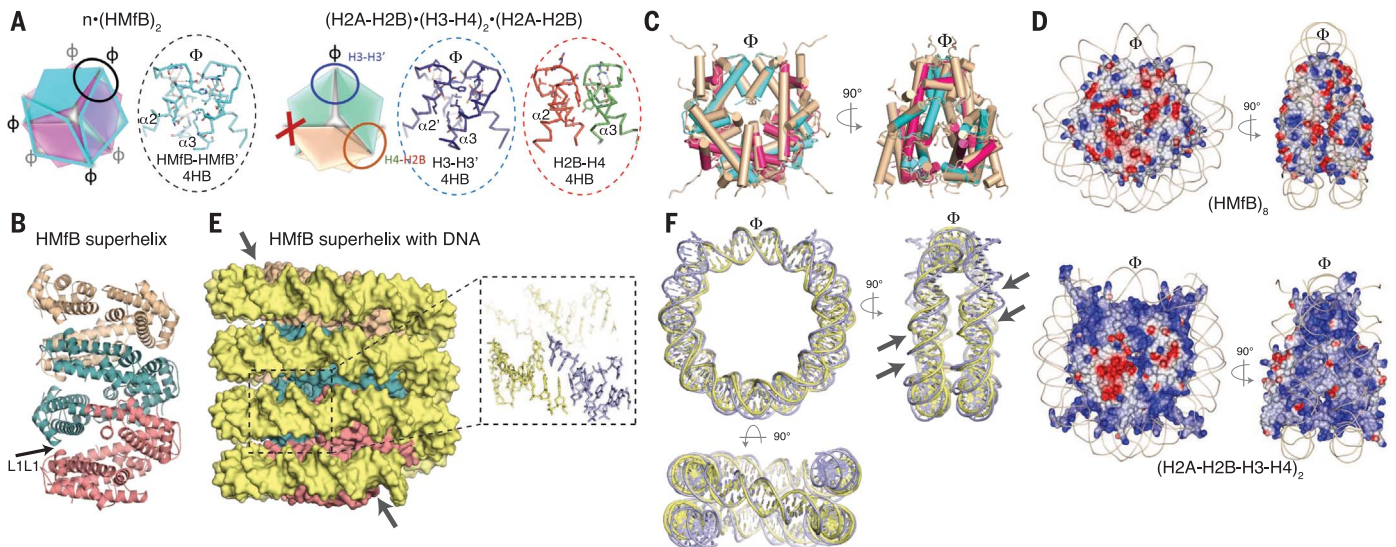

Fig. 2. Archaeal histones form a continuous superhelical ramp.
(**A**) Archaeal (HMfB)$_2$ dimers and eukaryotic core-histone heterodimers polymerize through the assembly of 4HBs involving the C termini of α2 and α3 of the HF. Although the symmetric (HMfB)$_2$ dimers can continue to polymerize, forming a protein fiber with consecutive, identical 4HB bundles (oval and inset), the asymmetry of eukaryotic core-histone dimers prevents continued polymerization at the site marked by the red X. (**B**) Nine (HMfB)$_2$ dimers are shown forming a continuous protein superhelix via 4HB interactions, with groups of three consecutive dimers shown in pink, teal, and tan. Modeling confirmed that the superhelix can also be formed by (HMfA)$_2$ homodimers and by HMfA-HMfB heterodimers. The arrow shows the location of the G16-G16 interaction (L1L1). (**C**) An octamer of archaeal HFs superimposes closely with the eukaryotic histone octamer (tan) in the nucleosome. Helices are shown as tubes with the archaeal histones colored magenta and cyan. (**D**) Archaeal HMfB octamer (top panel) and eukaryotic histone octamer (bottom panel) differ in their charge distribution, with a more positively charged helical ramp on the surface of the histone core (the basic histone tails are excluded for clarity). Electrostatic surfaces are calculated in the *CCP4mg* program and displayed from −0.5 V (red) to 0.5 V (blue). The DNA backbone is shown as a line. (**E**) DNA (shown in space-filling mode) wrapped around the HMfB superhelix shown in the same orientation as in (B). Inset shows a close-up view of the annealed 2-nt 5′ extensions. One supergroove is indicated by two arrows. (**F**) Superposition of 120 bp of DNA organized by four (HMfB)$_2$ HFDs with 146-bp nucleosomal DNA, shown in three orthogonal orientations; the top two orientations are identical to the orientations shown in (C). Two supergrooves (minor and major) are indicated by arrows.
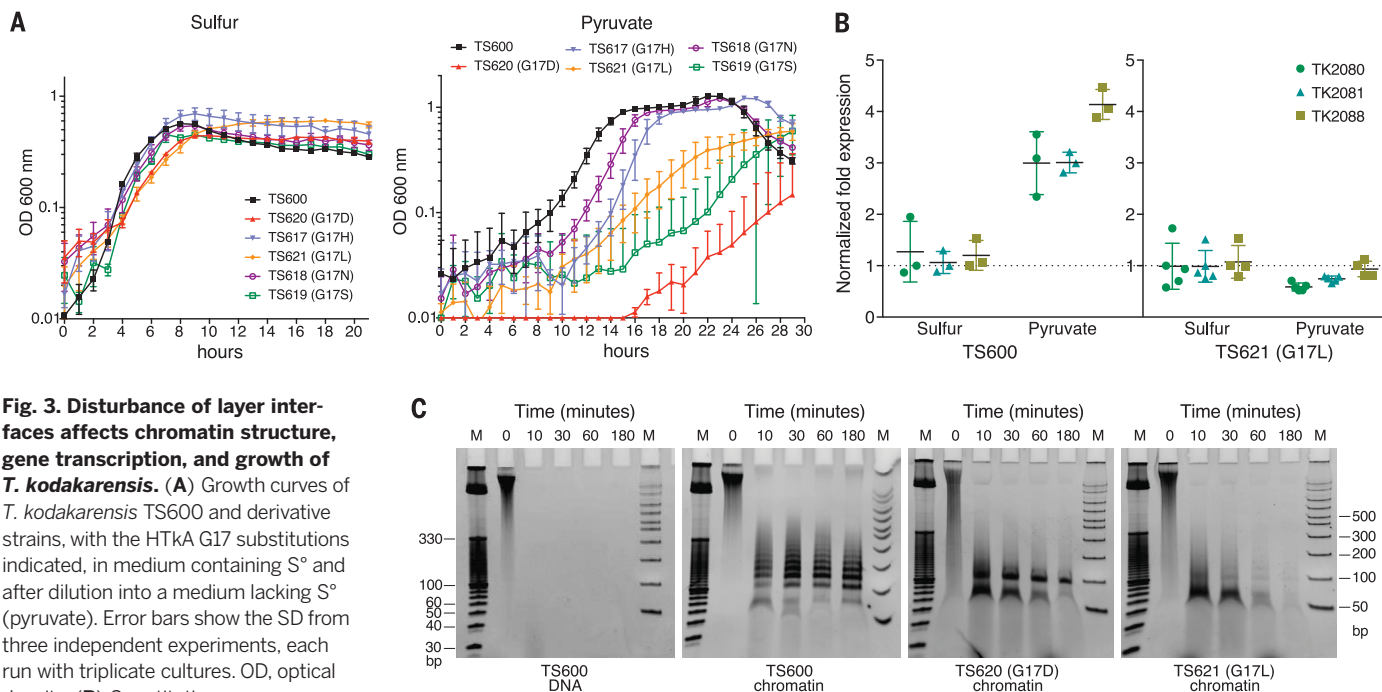
Fig. 3. Disturbance of layer interfaces affects chromatin structure, gene transcription, and growth of *T. kodakarensis*. (**A**) Growth curves of *T. kodakarensis* TS600 and derivative strains, with the HTkA G17 substitutions indicated, in medium containing S° and after dilution into a medium lacking S° (pyruvate). Error bars show the SD from three independent experiments, each run with triplicate cultures. OD, optical density. (**B**) Quantitative reverse transcription–polymerase chain reaction (qRT-PCR) of transcripts of three genes in the hydrogenase operon (TK2080, TK2081, TK2088) present in *T. kodakarensis* cells containing HTkA or HTkA G17L grown in the presence or absence of S°. Transcripts of TK0895, TK1431, and TK1311 were quantified as constitutively expressed reference genes. Shown is the fold change of the hydrogenase transcripts in cells after dilution into a medium lacking S° (pyruvate). (**C**) DNA fragments generated by MNase digestion of chromatin isolated from *T. kodakarensis* TS600 and derivative TS620 (G17D) and TS621 (G17L). DNA stripped from histones prior to MNase digestion is shown as a control (TS600). Size standards are in lanes labeled M.

resulting from the inability of the HTkA variants to form a stable extended superhelix. Apparently, substitution of leucine or aspartate for G17 prevents the close adjacent assembly of more than three (HTkA)$_2$ dimers on DNA.

Our data establish that most features of eukaryotic DNA compaction into nucleosomes are conserved in archaeal histone-based chromatin. The histone-mediated DNA geometry within these assemblies is exactly the same. However, archaeal histone-DNA complexes are not limited to one discrete structure. Unlike the defined nucleosome, archaeal histones can form complexes with variable numbers of histone dimers assembled along the DNA (*15*), and the resulting extended structure plays a role in gene regulation.

Why was the more flexible, variable-length archaeal chromatin structure replaced by a defined nucleosome consisting of four distinct histones very early in eukaryotic evolution? Possibly, with increasing genome size, it was necessary to limit histone assembly to defined nucleosomes to allow further compaction into precisely organized but still readily accessible higher-order chromatin. With diversification into four distinct histones and numerous histone variants, plus the addition of HF extensions and tails, eukaryotes further gained the ability to selectively position nucleosomes, have a conserved chromatin architecture recognizable by regulatory proteins, and develop elaborate epigenetic regulation through post-translational modification of histone tails. Intriguingly, some recently identified archaeal

histone sequences do have histone tails, hinting at the beginnings of this diversification (fig. S1B). However, to date, there is no evidence for archaeal functional homologs, and thus determining the ancestry of eukaryotic histone chaperones, chromatin remodelers, and posttranslational histone regulators remains a challenge (*22*).

REFERENCES AND NOTES

1. K. Luger, A. W. Mäder, R. K. Richmond, D. F. Sargent, T. J. Richmond, *Nature* **389**, 251–260 (1997).
2. K. Sandman, J. N. Reeve, *Curr. Opin. Microbiol.* **9**, 520–525 (2006).
3. H. S. Malik, S. Henikoff, *Nat. Struct. Biol.* **10**, 882–891 (2003).
4. P. B. Talbert, S. Henikoff, *Nat. Rev. Mol. Cell Biol.* **11**, 264–275 (2010).
5. K. Zaremba-Niedzwiedzka et al., *Nature* **541**, 353–358 (2017).
6. H. Nishida, T. Oshima, *J. Gen. Appl. Microbiol.* **63**, 28–35 (2017).
7. K. Luger, T. J. Richmond, *Curr. Opin. Genet. Dev.* **8**, 140–146 (1998).
8. K. A. Bailey, S. L. Pereira, J. Widom, J. N. Reeve, *J. Mol. Biol.* **303**, 25–34 (2000).
9. K. Sandman, D. Soares, J. N. Reeve, *Biochimie* **83**, 277–281 (2001).
10. K. Decanniere, A. M. Babu, K. Sandman, J. N. Reeve, U. Heinemann, *J. Mol. Biol.* **303**, 35–47 (2000).
11. K. Sandman, H. Louvel, R. Y. Samson, S. L. Pereira, J. N. Reeve, *Extremophiles* **12**, 811–817 (2008).
12. D. J. Soares, K. Sandman, J. N. Reeve, *J. Mol. Biol.* **297**, 39–47 (2000).
13. R. S. Edayathumangalam, P. Weyermann, J. M. Gottesfeld, P. B. Dervan, K. Luger, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 6864–6869 (2004).
14. N. Nalabothula et al., *BMC Genomics* **14**, 391 (2013).
15. H. Maruyama et al., *EMBO Rep.* **14**, 711–717 (2013).
16. M. Tomschik, M. A. Karymov, J. Zlatanova, S. H. Leuba, *Structure* **9**, 1201–1211 (2001).
17. S. Tan, C. A. Davey, *Curr. Opin. Struct. Biol.* **21**, 128–136 (2011).
18. D. Kato et al., *Science* **356**, 205–208 (2017).
19. F. Song et al., *Science* **344**, 376–380 (2014).
20. T. H. Hileman, T. J. Santangelo, *Front. Microbiol.* **3**, 195 (2012).
21. T. J. Santangelo, L. Cuboňová, J. N. Reeve, *Mol. Microbiol.* **81**, 897–911 (2011).
22. L. Aravind, A. M. Burroughs, D. Zhang, L. M. Iyer, *Cold Spring Harb. Perspect. Biol.* **6**, a016063 (2014).

Science

Structure of histone-based chromatin in Archaea

Francesca Mattiroli, Sudipta Bhattacharyya, Pamela N. Dyer, Alison E. White, Kathleen Sandman, Brett W. Burkhart, Kyle R. Byrne, Thomas Lee, Natalie G. Ahn, Thomas J. Santangelo, John N. Reeve and Karolin Luger

Origin of DNA compaction

As a repeating unit in eukaryotic chromatin, a nucleosome wraps DNA in superhelical turns around a histone octamer. Mattiroli *et al.* present the crystal structure of an archaeal histone-DNA complex in which the histone-mediated DNA geometry is exactly the same as that in the nucleosome. Comparing features of archaeal and eukaryotic chromatin structures offers important insights into the evolution of eukaryotic nucleosomes.

Science, this issue p. 609

ARTICLE TOOLS	http://science.sciencemag.org/content/357/6351/609
SUPPLEMENTARY MATERIALS	http://science.sciencemag.org/content/suppl/2017/08/09/357.6351.609.DC1
REFERENCES	This article cites 47 articles, 7 of which you can access for free http://science.sciencemag.org/content/357/6351/609#BIBL
PERMISSIONS	http://www.sciencemag.org/help/reprints-and-permissions

Use of this article is subject to the Terms of Service