Molecular Microbiology (2010) 76(3), 749–759 ■
doi:10.1111/j.1365-2958.2010.07136.x
First published online 13 April 2010

Structural insights into the interaction of the crenarchaeal chromatin protein Cren7 with DNA

Zhenfeng Zhang,[1†] Yong Gong,[2†] Li Guo,[1] Tao Jiang[2] and Li Huang[1*]

[1]*State Key Laboratory of Microbial Resources, Institute of Microbiology, Chinese Academy of Sciences, No. 1 West Beichen Road, Chaoyang District, Beijing 100101, China.*

[2]*National Key Laboratory of Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, 15 Datun Road, Chaoyang District, Beijing 100101, China.*

Summary

Cren7, a newly found chromatin protein, is highly conserved in the Crenarchaeota. The protein shows higher affinity for double-stranded DNA than for single-stranded DNA, constrains negative DNA supercoils *in vitro* and is associated with genomic DNA *in vivo*. Here we report the crystal structures of the Cren7 protein from *Sulfolobus solfataricus* in complex with two DNA sequences. Cren7 binds in the minor groove of DNA and causes a single-step sharp kink in DNA (~53°) through the intercalation of the hydrophobic side chain of Leu28. Loop β3-β4 of Cren7 undergoes a significant conformational change upon binding of the protein to DNA, suggesting its critical role in the stabilization of the protein–DNA complex. The roles of DNA-contacting amino acid residues in stabilizing the Cren7–DNA interaction were examined by mutational analysis. Structural comparison of Cren7-DNA complexes with Sac7d-DNA complexes reveals significant differences between the two proteins in DNA binding surface, suggesting that Cren7 and Sul7d serve distinct functions in chromosomal organization.

Introduction

Chromatin proteins are essential in all cellular organisms. They function both to package and to regulate the availability of the genome. In Eukarya, DNA is wrapped around the histone core of nucleosome, the basic structural unit for DNA packaging in the chromatin (Luger *et al.*, 1997). Bacteria possess many DNA binding proteins (e.g. HU, IHF, H-NS, Fis and Lrp), and the chromosomal DNA is folded into a compact structure called nucleoid (Wu, 2004; Dame, 2005). However, the situation is intriguingly complex in Archaea. The Euryarchaeota contain archaeal histones that share a common ancestry with the histone fold regions of the eukaryotic nucleosome core histones (Reeve, 2003; Sandman and Reeve, 2005). Interestingly, archaeal histones are absent from the Crenarchaeota, with the exception of some species, e.g. *Thermofilum pendens* Hrk5 (Anderson *et al.*, 2008). A number of small, basic DNA binding proteins have been found in crenarchaea (e.g. Alba, Sul7d, CC1 and Cren7) (Grote *et al.*, 1986; Bell *et al.*, 2002; Luo *et al.*, 2007). However, Sul7d is restricted to the order Sulfolobales, and CC1 exists only in the order Thermoproteales and Desulfurococcales (Luo *et al.*, 2007). Alba is highly conserved among both archaeal kingdoms but appears to bind both DNA and RNA *in vivo* (Guo *et al.*, 2003; Marsh *et al.*, 2005). So the physiological role of this protein family remains to be understood. Cren7, a newly identified small and basic DNA binding protein, is highly conserved in the Crenarchaeota (Guo *et al.*, 2008). This 6.6 kDa protein is abundant in *Sulfolobus solfataricus* (~1% of total cellular protein), binds preferentially to dsDNA over ssDNA and is associated with genomic DNA *in vivo*. Cren7 binds to DNA without apparent sequence preference and significantly increases the melting temperature of DNA. Topology assays show that Cren7 is more efficient than Sul7d in constraining DNA in negative supercoils (Baumann *et al.*, 1994; Mai *et al.*, 1998). These results are consistent with a proposed role of Cren7 as a chromatin protein in crenarchaea.

The solution structure of Cren7 has been determined (Guo *et al.*, 2008). The protein adopts an SH3-like fold and is believed to interact with duplex DNA through a triple-stranded β-sheet and a long flexible loop. Cren7 resembles Sul7d in both biochemical properties and structure although the two proteins share no significant similarity at the amino acid sequence level. However, several notable structural differences exist between Cren7 and Sul7d. The C-terminal segment of Sul7d is 15 residues longer than that of Cren7 and forms a helix that is absent in Cren7. In addition, at the DNA binding surface, Cren7 has a long and highly flexible loop (Loop

Accepted 10 March, 2010. *For correspondence. E-mail huangl@sun.im.ac.cn; Tel. (+86) 10 64807430; Fax (+86) 10 64807429. †These authors contributed equally to this work.

Table 1. Data collection and refinement statistics.

	Cren7–GTAATTAC	Cren7–GCGATCGC
Data collection		
Space group	C2221	C2221
Cell dimensions		
a, *b*, *c* (Å)	49.7, 52.4, 88.6	77.4, 77.2, 105.0
α, *β*, *γ* (°)	90.0, 90.0, 90.0	90.0, 90.0, 90.0
Resolution (Å)	30–1.90 (1.98–1.90)	30–2.30 (2.38–2.30)
R_{merge}	0.05 (0.34)	0.06 (0.46)
$I/\sigma I$	44.0 (8.7)	39.6 (4.8)
Completeness (%)	99.5 (100.0)	99.9 (100.0)
Redundancy	7.7 (7.8)	8.1 (8.3)
Refinement		
Resolution (Å)	30–1.90	30–2.30
No. reflections	8918	13 592
R_{work}/R_{free}	0.216/0.246	0.203/0.234
No. atoms		
Protein	463	914
DNA	322	644
Water	78	115
B-factors		
Protein	41.7	45.5
DNA	37.9	44.0
Water	46.0	40.9
r.m.s.d.		
Bond lengths (Å)	0.008	0.008
Bond angles (°)	1.4	1.4
Ramachandran analysis		
Most favoured (%)	95.5	94.2
Additional allowed (%)	4.5	5.8
Generously allowed (%)	0	0
Disallowed (%)	0	0

β3-β4), whereas only a small hinge exists at the corresponding position in Sul7d (Guo *et al*., 2008).

In the present work, we report two crystal structures of Cren7–DNA complexes. We show that Cren7 binds in the minor groove of DNA and causes a single-step sharp kink in DNA (~50° and ~53° in the two complexes, respectively) through the intercalation of the hydrophobic side chain of Leu28. Loop β3-β4 of Cren7 undergoes a significant conformational change upon binding of the protein to DNA. The roles of DNA-interacting residues in the binding of Cren7 to DNA were investigated by mutational analysis. Structural comparison of Cren7–DNA complexes with Sac7d–DNA complexes reveals significant differences in DNA binding surface, suggesting distinct roles of the two proteins in chromosomal organization.

Results

Overall structures of Cren7–DNA complexes

To understand the structural details of the interaction between Cren7 and DNA, the crystal structures of Cren7–d(GCGATCGC)$_2$ [Protein Data Bank (PDB) code 3lwi] and Cren7-d(GTAATTAC)$_2$ (PDB code 3lwh) have been solved and well refined at 2.3 Å and 1.9 Å respectively (Table 1). The two complexes have different crystal packing interactions. One protein–DNA complex is found in the asymmetric unit of the crystal of Cren7–GTAATTAC. Two molecules are in the asymmetric unit of the Cren7–GCGATCGC complex. They are related by a perfect non-crystallographic twofold symmetry axis and the root mean square deviation (r.m.s.d.) value between them is 0.011 Å.

The two Cren7–DNA complexes share a similar overall structure (Fig. 1A–C). Cren7 binds to the DNA duplex as a monomer. The protein adopts an SH3-like fold and has a 'β-barrel' structure including two anti-parallel β-sheets. The triple-stranded β-sheet and loop β3-β4 bind on the minor groove surface of DNA and contribute all of the contacts with DNA (Fig. 2A). In both complexes, Cren7 covers almost all eight base pairs. The DNA in the complex remains in the B form except for the A3pA4 step, where the DNA minor groove is drastically widened. The binding sites of Cren7 are sharply kinked by the intercalation of a single amino acid residue (Fig. 3A).

Protein–DNA contacts

In the Cren7–GTAATTAC complex, a surface area of 1530 Å^2 is buried in Cren7 (Fig. 2A). A nearly identical interaction surface was also found in the Cren7–GCGATCGC complex. The side chains of many surface

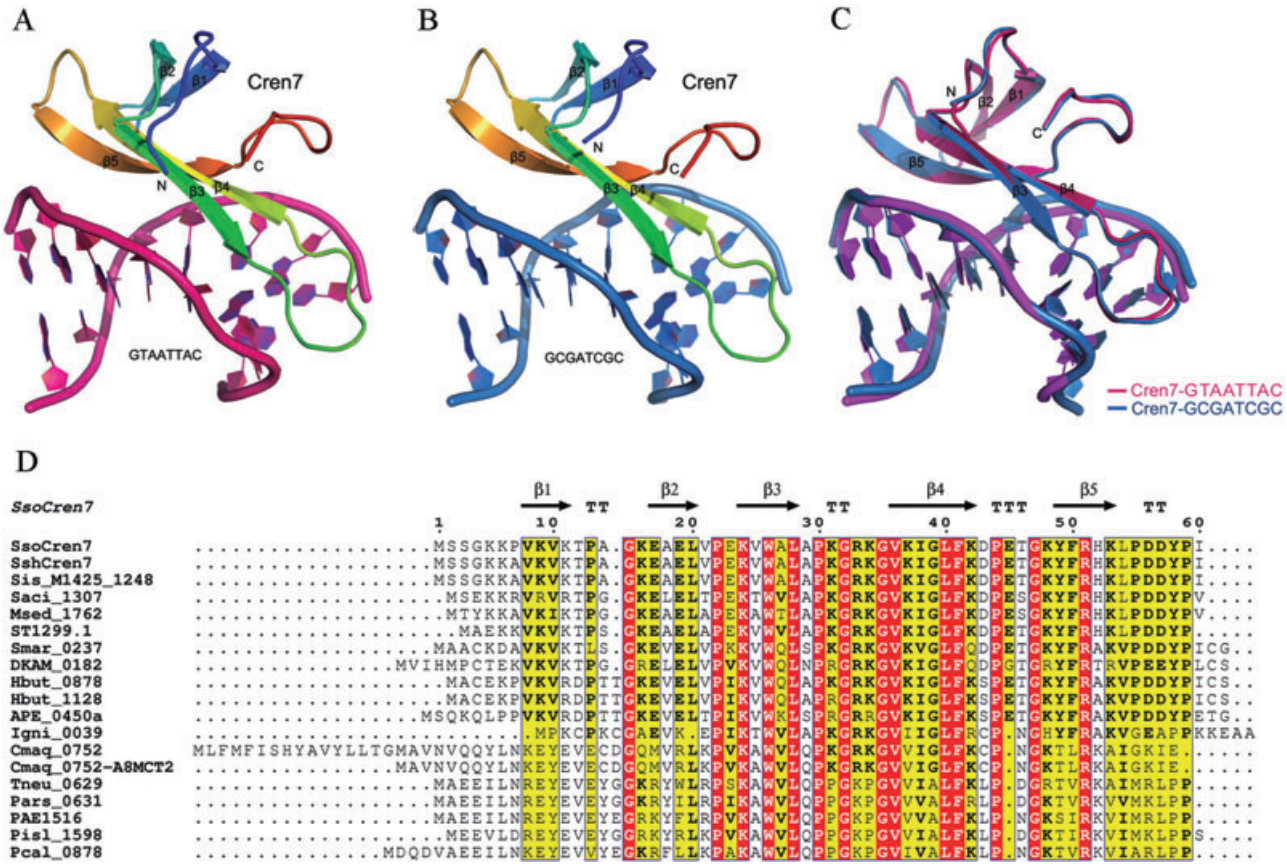

Fig. 1. Overall structures of Cren7–DNA complexes.
A. Ribbon diagram of the Cren7–GTAATTAC complex.
B. Ribbon diagram of the Cren7–GCGATCGC complex.
C. Superposition of the two Cren7–DNA complexes, fitted by common protein main-chain atoms.
D. Sequence alignment of Cren7 homologues in crenarchaea.

amino acid residues (Lys24, Trp26, Leu28, Pro30, Lys31, Arg33, Val36, Ile38, Leu40, Tyr49, Arg51 and Lys53), most of which are conserved in the Cren7 family (Fig. 1D), are in contact with the DNA (Fig. 2B).

Among the most notable structural features in the binding surface of Cren7 is the long β3-β4 loop containing seven amino acid residues (Ala29-Gly35). Several of these residues (Pro30, Lys31 and Arg33) are highly con-

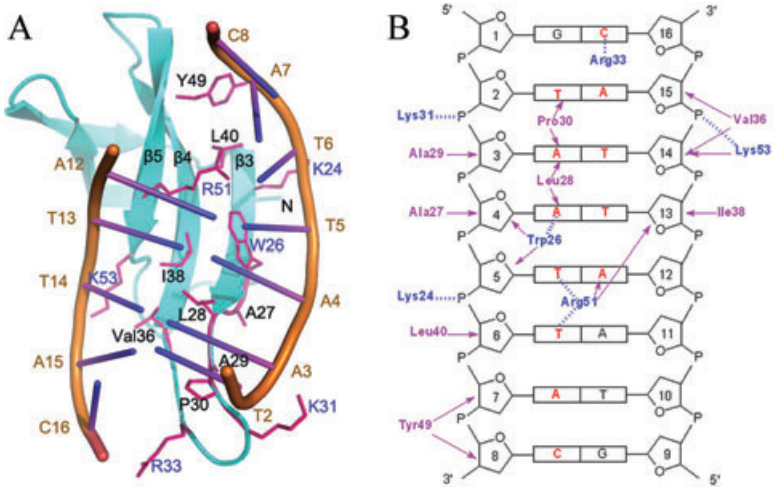

Fig. 2. Detailed view of local structures at the protein–DNA interface of the Cren7–GTAATTAC complex.
A. Selected side chains of Cren7 are shown. Amino acid residues that contact DNA through hydrogen bonds (blue) or hydrophobic interaction (black) are indicated.
B. Schematic diagram summarizing all of the important Cren7–DNA contacts. Hydrogen bonds and hydrophobic interactions are shown with blue dashed lines and pink arrows respectively.

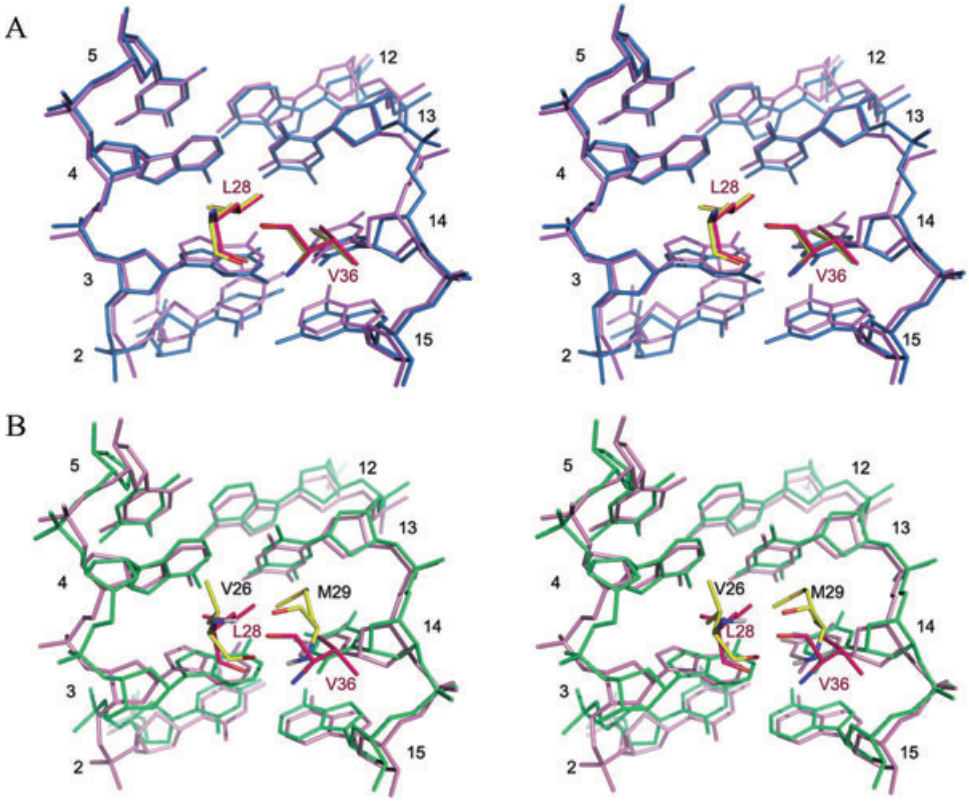

Fig. 3. Stereoscopic view of the sites of intercalation.
A. Superposition of the local structures of the two Cren7–DNA complexes. The DNA octamer is kinked by 53° at the A3pA4 step in the Cren7–GTAATTAC complex and 50° at the G3pA4 step in the Cren7–GCGATCGC complex. The conformations of the intercalating side chain of Leu28 in the two complexes are shown. Cren7–GTAATTAC is shown in red and pink, and Cren7–GCGATCGC in yellow and blue.
B. Superposition of the local structures of Cren7–GTAATTAC and Sac7d–GTAATTAC (PDB code: 1AZQ). Cren7–GTAATTAC is shown in red and pink, and Sac7d–GTAATTAC in yellow and green.

served (Fig. 1D). In the Cren7–GTAATTAC complex, the NH2 group of Arg33 forms a hydrogen bond (3.31 Å) to the O3′ of C16. The N atom of Lys31 is hydrogen-bonded (2.97 Å) to the O2P of A3. In addition, both Ala29 and Pro30 are in close contact with the DNA. These interactions position the loop in the minor groove of DNA in a span of about two base pairs (T2-A15 and A3-T14). The ability of loop β3-β4 to adopt a drastic conformational change in establishing close contact with DNA following the binding of the protein to the DNA points to the importance of the loop in stabilizing the Cren7–DNA interaction.

In the two crystals, the side chain of β3-Leu28 is intercalated between the G3-C14 and A4-T13 base pairs in GCGATCGC and between the A3-T14 and A4-T13 base pairs in GTAATTAC, resulting in a sharp kink in the DNAs. The hydrophobic interactions between Leu28 and the bases of A3 and A4 in the Cren7–GTAATTAC complex bend the DNA in the major groove direction. The intercalation of Leu28 is stabilized by Val36 through its interaction with the sugar moiety of T14 and A15 (Figs 2B and 3A). Several other conserved residues contribute greatly to the intermolecular interaction (Fig. 2B). The NH2 group of Lys24 forms a hydrogen bond (3.11 Å) to the O1P of T6. Arg51 is uniquely positioned to contact both strands of the DNA duplex. The residue forms hydrophobic interactions with T5, T6, A12 and T13 and is hydrogen bonded through its NH2 group and NH1 group to the O2 of T6 (2.69 Å) and the O2 of T5 (3.04 Å) respectively. Like Trp24 in Sso7d, Trp26 in Cren7 plays several roles. First, its bulky ring fills up the space between Cren7 and the DNA. Second, it is in close contact with A4 and T5. Third, its indole NH group forms a hydrogen bond (3.03 Å) to the N3 of A4.

Most of the DNA-interacting residues in Cren7 contact one DNA strand. In fact, only four residues (Arg33, Val36, Ile38 and Lys53), in addition to Arg51, which contacts both DNA strands, are found to interact with the other strand (Fig. 2B). Val36 and Ile38 bind to the backbone sugar through hydrophobic interactions. Lys53 is in close contact with the DNA as its NH2 group forms a specific hydrogen bond (2.79 Å) to the O1P of A15. Despite the asymmetry of the binding of Cren7 to duplex DNA, the protein appears to show much greater affinity for dsDNA than for ssDNA (Guo *et al.*, 2008).

Table 2. Kinetic analysis of the binding of wild-type and mutant Cren7 proteins to dsDNA.

Protein	Binding affinity, K_D (M)	Association rate, k_a (M^{-1} s^{-1})	Dissociation rate, k_d (s^{-1})
Cren7-WT	1.23×10^{-7}	6.36×10^{3}	7.85×10^{-4}
K24A	2.84×10^{-7}	9.17×10^{4}	2.60×10^{-2}
W26A	3.63×10^{-6}	5.47×10^{3}	1.99×10^{-2}
L28A	6.66×10^{-6}	8.27×10^{3}	5.51×10^{-2}
P30A	4.87×10^{-7}	6.84×10^{3}	3.33×10^{-3}
K31A	6.14×10^{-7}	1.63×10^{4}	1.00×10^{-2}
R33A	1.86×10^{-7}	7.90×10^{4}	1.47×10^{-2}
V36A	9.54×10^{-7}	6.49×10^{3}	6.19×10^{-3}
Y49A	4.67×10^{-7}	8.17×10^{3}	3.82×10^{-3}
R51A	2.87×10^{-6}	1.13×10^{4}	3.25×10^{-2}
K53A	3.00×10^{-6}	9.61×10^{3}	2.88×10^{-2}

To learn more about the roles of the DNA-interacting amino acid residues of Cren7, ten single-point mutant proteins containing an alanine substitution at the position of each of these residues were constructed (Table 2). Binding of the mutant proteins to a 30 bp dsDNA fragment was determined by surface plasmon resonance (SPR) assays (Table 2 and Fig. S1). All mutants showed reduced binding affinity for the DNA. Substitution of Ala28 for Leu28 resulted in the largest loss (54-fold) in DNA binding affinity, supporting a critical role for this residue in DNA binding. Trp26, Arg51 and Lys53 are also important in DNA binding as mutation of these residues was accompanied by 30-, 23- and 24-fold decrease in affinity respectively. Depending on their association and dissociation rates in protein–DNA complex formation, these mutants appear to fall into three groups (Fig. 4). K24A and R33A belong to the first group, which is characterized by high rates of association and dissociation, and a modest reduction in DNA binding affinity. The possibility exists that a change in surface electronic potential in the mutant proteins facilitated DNA binding but destabilized the resulting protein–DNA complexes. Second group consists of four mutants (P30A, K31A, V36A and Y49A) with a markedly increased dissociation rate and a slightly increased association rate. The third group of four mutants (W26A, L28A, R51A and K53A) exhibits the lowest affinity for DNA, apparently as a result of the drastically increased dissociation rates. These four residues appear to play a critical role in stabilizing the protein–DNA complex.

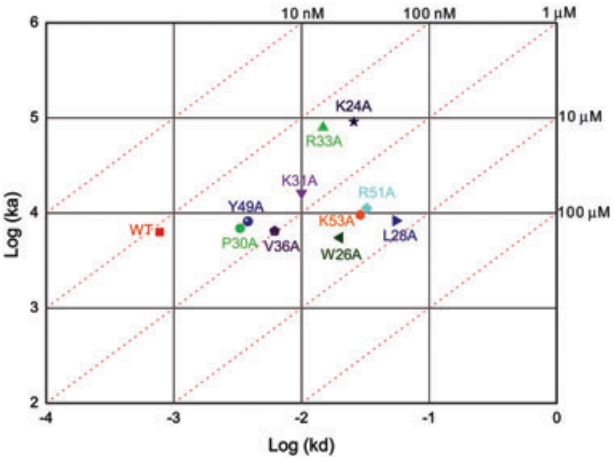

Fig. 4. Kinetic characterization of the binding of wild-type and mutant Cren7 proteins to DNA by SPR. Wild-type or mutant Cren7 was injected over immobilized DNA. The k_a (recognition) is plotted against the k_d (complex stability) on logarithmic scales. The diagonal lines (pink dotted lines) are isoaffinity lines representing the affinity $K_D = k_d/k_a$.

Conformational changes of Cren7 in complexes

The structures of the bound Cren7 in the two complexes are similar with an r.m.s.d. of 1.12 Å (using Cα atoms of residues 3–60). Superposition of free (PDB code: 2JTM) and DNA-bound Cren7 proteins, fitted by their α-carbons (an r.m.s.d. of 2.07 Å), reveals the flexibility of loop β3-β4 in an otherwise similar overall structures (Fig. 5A). The orientation of the loop becomes fixed upon binding of the protein to the DNA. A similar observation was made in our previous steady-state ^{1}H-^{15}N NOE experiments, in which ^{1}H-^{15}N NOE values for residues in loop β3-β4 increased significantly following the binding of Cren7 to DNA (Guo *et al.*, 2008). On the other hand, while the side chains of the core of Cren7 are highly conserved in conformation (Fig. 5B), the surface side chains of the protein, especially those involved in DNA binding, appear to be variable (Fig. 5C). Therefore, we conclude that Cren7 maintains a largely unchanged overall structure with significant conformational adjustments occurring in loop β3-β4 and the side chains of the surface amino acids upon binding to DNA.

DNA deformation

The DNA bound by Cren7 shows similar global conformation to that bound by Sso7d/Sac7d except for an obvious distortion of the backbone of one DNA strand in the T2pA3 region (Fig. 6C). The distortion is due to the translocation of the phosphate group, which may be caused by the formation of a hydrogen bond to Lys31 (Fig. 2B).

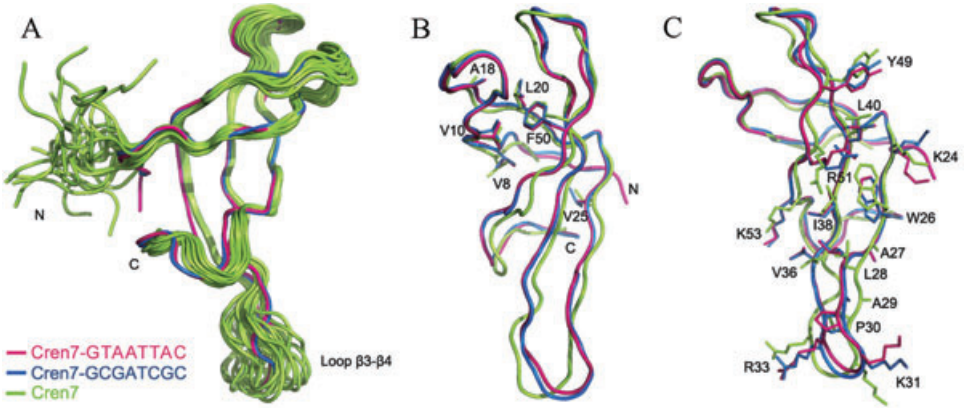

Fig. 5. Superposition of three Cren7 structures from the Cren7–GTAATTAC complex, the Cren7–GCGATCGC complex and the NMR solution structures (PDB code: 2JTM), fitted by common protein main-chain atoms.
A. The wire diagram of the three protein backbones.
B. Core residues that are packed within the 'β-barrel' structure.
C. Residues that are involved in the DNA binding surface.

The major groove of DNA is not bound by Cren7 and therefore entirely accessible. The deformation of DNA is caused by the binding of Cren7 to the DNA from the minor groove. The minor groove is dramatically widened to about 12 Å at the intercalation site indicating a change of DNA conformation from the B form into the A form. In addition, significant positive rolls are found at the base pair steps around the intercalation site. A sharp kink of ~53° occurs at the A3pA4 step in the GTAATTAC duplex and a smaller kink of ~50° at the G3pA4 step in the GCGATCGC duplex. The two kinks are smaller than those introduced by Sac7d in the same DNA sequences (~60° for GTAATTAC and ~61° for GCGATCGC) (Gao *et al*., 1998; Robinson *et al*., 1998). There are two additional rolls at steps adjacent to the site of intercalation (8.0° at T2pA3 step and 11.0° at T5pT6 step in Cren7–GTAATTAC; 8.9° at C2pG3 step and 11.8C° at T5pC6 step in Cren7–GCGATCGC) (Table 3). Rolls at these steps were not found in the Sso7d–DNA complexes. The roll at T2pA3 step seems to result from the hydrophobic interactions of Pro30 with both bases of T2 and A3, while the roll at T5pT6 step is likely introduced by Arg51, which makes hydrogen bonds with T5 and T6. The average roll angle of GTAATTAC (*c*. 10.9°) is larger than that of GCGATCGC (*c*. 9.8°).

Binding by Cren7 also introduces a significant change in the twist of the DNA helix. The base pairs at the site of intercalation show the smallest twist. The average twist for a single base pair is ~29.3° in the Cren7–GTAATTAC complex and ~29.4° in the Cren7–GCGATCGC complex. These values translate into approximately 12.3 base pairs per helical turn in the bound DNA. Given a binding size of ~8 bp, binding of ~9 Cren7 molecules constrains one negative supercoil, a stoichiometry consistent with that determined previously in nick-closure assays (Guo *et al*., 2008).

Structural comparison of Cren7–DNA complexes with Sac7d–DNA complexes

Despite the similarity between Cren7 and Sac7d in overall structure, the two proteins differ drastically in DNA binding surface (Fig. 6B). First, Cren7 shows a substantially larger binding site (~8 bp) (Fig. 2) than Sac7d (4 bp) as the former possesses a long loop between β3 and β4 in the DNA binding surface whereas the latter has a small hinge in the corresponding position. Loop β3-β4 contains four DNA-interacting residues (Ala29, Pro30, Lys31 and Arg33) and covers ~2 bp by itself. Second, the β-sheet composed of β1- and β2-strands in Cren7 does not interact with DNA. By contrast, Lys7, Tyr8 and Lys9 in the corresponding region in Sac7d interact with the backbone deoxyribose of the DNA. Third, upon binding to duplex DNA, a single amino acid residue (Leu28) in Cren7 is intercalated between the base pairs of the bound DNA, generating a sharp DNA kink. Similar DNA deformation is caused by intercalation of two residues (Val26 and Met29) in Sac7d in the Sac7d–DNA complexes (Gao *et al*., 1998; Robinson *et al*., 1998). Leu28 and Val36, two conserved hydrophobic residues in Cren7, are counterparts of Val26 and Met29 in Sac7d respectively (Figs 3B and 6A). However, the side chain of Val36 in Cren7 contacts the deoxyribose of the DNA through hydrophobic interaction instead of intercalating into the DNA.

Discussion

Cren7 from *S. solfataricus* has been characterized as a chromatin protein. Because all genome-sequenced crenarchaea except for *T. pendens* Hrk5 encode a homologue of Cren7, it is suggested that crenarchaea may employ a Cren7-based DNA packaging strategy (Guo

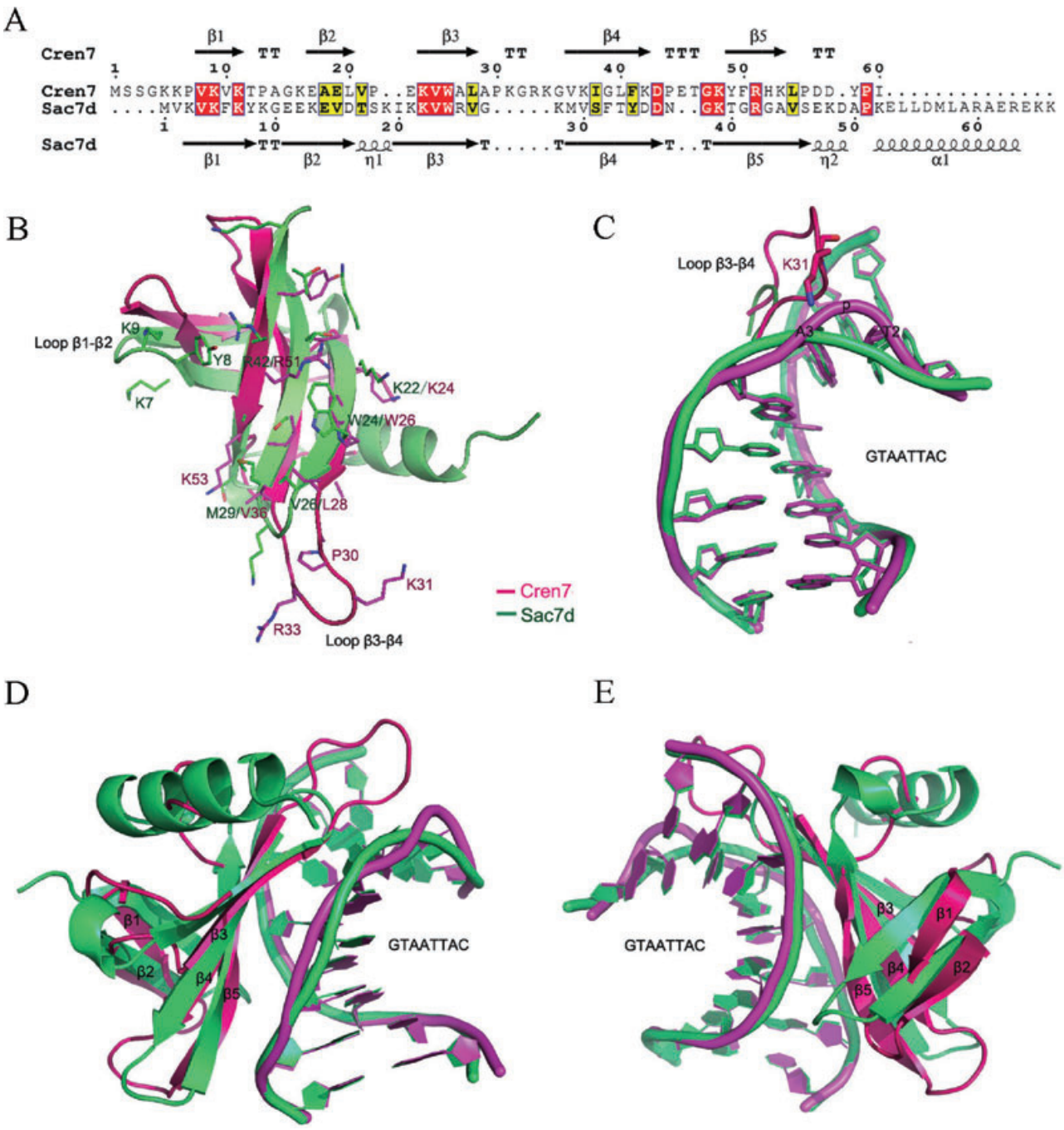

Fig. 6. Comparison of structures of Cren7 and Sac7d in complex with DNA.
A. Structure-based sequence alignment of Cren7 and Sac7d. The secondary structural elements of the two proteins are also shown.
B. The DNA binding surfaces of Cren7 and Sac7d.
C. Comparison of the DNA structures (GTAATTAC) in the Cren7–DNA (purple) complex and the Sac7d–DNA (green; PDB code: 1AZQ) complex. The T2pA3 step is marked, and loop β3-β4 and Lys31 of Cren7 are also shown.
D. Superposition of the Cren7–GTAATTAC and Sac7d–GTAATTAC complexes fitted by main-chain atoms of one of the DNA strands.
E. The back view of the structure in (D). The Cren7–DNA complex is shown in red and purple, as in (B and C); and the Sac7d–DNA complex is shown in green.

et al., 2008). Intriguingly, *T. pendens* Hrk5 synthesizes an archaeal histone (Anderson *et al.*, 2008). Therefore, it is tempting to speculate that Cren7 plays a role in crenarchaea similar to that of archaeal histones in euryarchaea.

Structural analysis of Cren7–DNA complexes will provide insights into the atomic basis of the chromosomal organization in crenarchaea. In the present work, two crystal structures of Cren7–DNA complexes have been

Table 3. DNA helical parameters of the Cren7–DNA complexes and the Sac7d–DNA complexes.

GTAATTAC				GCGATCGC			
Step	Rise (Å)	Roll (°)	Twist (°)	Step	Rise (Å)	Roll (°)	Twist (°)
Cren7–DNA				Cren7–DNA			
GT/AC	3.09	2.80	30.80	GC/GC	3.19	0.89	33.73
TA/TA	3.14	7.95	28.74	CG/CG	3.01	8.92	22.71
AA/TT	5.62	53.00	18.06	GA/TC	6.06	50.09	20.68
AT/AT	2.88	5.88	20.94	AT/AT	3.13	4.22	22.80
TT/AA	3.20	11.04	23.41	TC/GA	3.24	11.84	24.11
TA/TA	3.44	–3.18	49.22	CG/CG	3.28	–4.93	44.25
AC/GT	3.30	–1.15	34.12	GC/GC	3.37	–2.11	37.57
Average	3.52	10.91	29.33	Average	3.61	9.85	29.41
Sac7d–DNA[a]				Sac7d–DNA[b]			
GT/AC	3.17	–2.32	31.78	GC/GC	2.89	2.01	27.65
TA/TA	3.23	3.73	29.21	CG/CG	5.30	61.30	19.92
AA/TT	5.95	59.77	24.11	GA/TC	2.92	8.31	22.12
AT/AT	3.17	5.63	18.14	AT/AT	3.12	3.95	25.69
TT/AA	3.23	6.44	29.49	TC/GA	3.40	1.74	37.57
TA/TA	3.17	–3.75	44.65	CG/CG	3.30	2.71	34.63
AC/GT	3.53	–1.85	29.23	GC/GC	3.22	–2.74	31.52
Average	3.63	9.66	29.52	Average	3.45	11.04	28.44

a. PDB code: 1AZQ.
b. PDB code: 1AZP.

determined. Cren7 binds in the minor groove of DNA, covering approximately eight base pairs. The DNA is sharply kinked at the binding site (~53°) and the minor groove is dramatically widened. The remarkable ability of Cren7 to distort DNA conformation is consistent with a role in DNA supercoiling and compaction.

Cren7 resembles Sul7d in biochemical properties. The global conformation of DNA in Cren7–DNA complexes is similar to that in Sso7d–DNA complexes. Both Cren7 and Sso7d are synthesized in large quantity in *S. solfataricus* P2. A recent analysis of the primary transcriptome of *S. solfataricus* P2 also showed high levels of transcription for both Cren7 and Sso7d (Wurtzel *et al.*, 2010). These findings suggest that the two proteins are both functional *in vivo*. Therefore, a question arises as to why they are both needed. A possibility exists that both Cren7 and Sul7d serve important roles in chromosomal organization but their functions are not entirely redundant. Comparison of the Cren7–DNA and Sac7d–DNA complexes reveals the structural basis for the possible functional differences between the two proteins. In the Cren7–DNA complexes, the protein binds asymmetrically to the duplex DNA. More protein–DNA contacts are found on one of the double strands. This may be related to the observation that the duplex DNA bound by Cren7 was more readily denatured than that bound by Sac7d (Mcafee *et al.*, 1995; Guo *et al.*, 2008). It has also been reported that Cren7 was significantly more efficient than Sul7d in constraining negative DNA supercoils (Guo *et al.*, 2008). This has led to the suggestion that the former induces greater helix unwinding than the latter. The suggestion appears to be consistent with the difference in DNA distortion induced by the two proteins. The average twist of the DNA in the Sac7d–GTAATTAC complex (29.5°) is nearly identical to that in Cren7–GTAATTAC complex (29.3°). However, it should be noted that Sac7d/Sso7d only covers four base pairs of the 8 bp DNA duplex (Gao *et al.*, 1998; Robinson *et al.*, 1998), whereas Cren7 has a binding size of ~8 bp. In a complex where two molecules of Sso7d bound to a 12 bp DNA duplex, covering eight base pairs, the average twist of the DNA (27.7°) remained largely unchanged (Agback *et al.*, 1998). In other words, binding of one Cren7 molecule or two Sac7d molecules to 8 bp of DNA would result in similar deficits in the helical twist of the DNA. Therefore, Cren7 is about twice as efficient as Sac7d/Sso7d in constraining DNA supercoils, as estimated biochemically (Mai *et al.*, 1998; Guo *et al.*, 2008).

Cren7 also differs significantly from Sac7d in DNA binding surface. Loop β3-β4 is a notable feature of Cren7. The crystal structures of Cren7–DNA complexes reveal that four residues in the loop directly contact the DNA, suggesting its critical role in the interaction between Cren7 and DNA. Residue Lys31 in the loop forms a hydrogen bond to the DNA backbone. Substitution of this residue for an alanine reduced the affinity of Cren7 for DNA by approximately fivefold. The reduction in the binding affinity resulted from a drastically increased dissociation rate (~13-fold) against a slightly increased association rate (~2.6-fold). Therefore, Lys31 plays a key role in stabilizing the Cren7–DNA complex. The DNA binding surfaces of the two proteins also differ in the region around loop β1-β2. Three residues in this region in the

Sac7d–DNA complexes are in close contact with the DNA but none in the Cren7–DNA complexes. The double-stranded β-sheet of Cren7 is extended away from the DNA (Fig. 6D and E), permitting ready access to the DNA by other proteins. Residue Lys31 on loop β3-β4 apparently undergoes reversible methylation (Guo *et al*., 2008). In addition, the N-terminal tail of Cren7 contains two serine and two lysine residues on loop β1-β2, which are often sites of phosphorylation and methylation or acetylation in eukaryotic histones (Grant, 2001; Nowak and Corces, 2004; Hansen *et al*., 2006). Therefore, the presence of the N-terminal tail and long and flexible loop β3-β4 in Cren7, but not in Sac7d, suggests that the interaction of the two proteins with DNA may be regulated differently. Taken together, the pronounced differences between Cren7 and Sul7d in DNA binding surface as well as binding pattern suggest that the two proteins may serve distinct roles in chromosomal organization.

Experiment procedures

Protein expression and purification

The recombinant Cren7 protein was expressed and purified by using a modification of the method described previously (Guo *et al*., 2008). Cell lysates were prepared by sonication in 20 mM Tris-Cl, pH 6.8 and 1 mM EDTA (buffer A). After centrifugation, the supernatant was heat-treated at 80°C for 20 min and centrifuged again. The sample was applied to a HiTrap SP Sepharose XL column (5 ml, GE) equilibrated in buffer A and eluted successively with 200 and 500 mM KCl. Proteins eluted in 500 mM KCl were dialysed to 0.1 M NH_4Ac (pH 5.2), then loaded onto a Superdex 75 column (10/300, GE) equilibrated in the same buffer. Fractions containing Cren7 were lyophilized and resuspended in buffer A. The recombinant Cren7 protein contained no post-translational modifications, as verified by mass spectrometry.

Preparation of oligonucleotides

Oligonucleotides 5′-GTAATTAC and 5′-GCGATCGC (Chen *et al*., 2005), synthesized at Sango BioTech (Shanghai, China), were dissolved in 20 mM Tris-Cl, pH 6.8 and 100 mM NaCl. Duplex DNA fragments were prepared by allowing each oligonucleotide to self-anneal in a process involving heating at 95°C for 5 min and subsequent cooling to room temperature in 2 h.

Crystallization, data collection and structure determination

Crystals of Cren7 in complex with $d(GTAATTAC)_2$ or $d(GCGATCGC)_2$ duplexes were grown at 20°C using sitting-drop vapour diffusion. The protein and the duplex DNA were mixed at an equimolar ratio to the final concentration of 1.7 mM each. The sample (1 μl) was mixed with the reservoir solution (1 μl) and equilibrated with 30% PEG1500 at 20°C. The crystals were mounted on nylon loops and immediately frozen in liquid nitrogen. The data for both complexes were collected to 1.9 and 2.3 Å at the Shanghai Synchrotron Radiation Facility (SSRF, China) at 100 K, and the wavelength was 0.97947 Å. Two sets of data were integrated and scaled with DENZO and SCALEPACK (Otwinowski and Minor, 1997).

The structures were determined by molecular replacement using the programme Phaser (Mccoy *et al*., 2007) from the CCP4 programme suite (Collaborative Computational Project, 1994) with both the solution structure of Cren7 (PDB code: 2JTM) (Guo *et al*., 2008) and the DNA structure of Sac7d–DNA complex [PDB code: 1AZP (Robinson *et al*., 1998) and 1WTO (Chen *et al*., 2005)] as the initial search models. The software packages Refmac (Murshudov *et al*., 1997) and Coot (Emsley and Cowtan, 2004) were used to complete the models, and models were refined to final R_{work}/R_{free} of 0.216/0.246 for Cren7–GTAATTAC and 0.203/0.234 for Cren7–GCGATCGC respectively. Statistics for data collection and refinement are given in Table 1. The qualities of the final structures were evaluated using PROCHECK (Laskowski *et al*., 1993). DNA conformations were analysed using the programme X3DNA (Lu and Olson, 2008). The atomic coordinates of the two Cren7–DNA complexes have been deposited in the RCSB Protein Data Bank. All images were prepared using Pymol (Delano, 2002) and ESPript (Gouet *et al*., 1999).

Kinetic analysis of the interaction of Cren7 with DNA by SPR

Surface plasmon resonance experiments were carried out at 25°C using the BIAcore 3000 instrument (BIAcore AB, Uppsala, Sweden). The running buffer contained 50 mM Tris-Cl, pH 7.6, 100 mM NaCl and 0.005% (v/v) Tween 20. Complementary oligonucleotides, including one with a biotin-label (5′-TTTCTACCCTTTGGTGCTAATGCCCAT ACT) were captured on SA sensor chip (91–97 response units). A blank flow cell was used as reference to correct for instrumental and concentration effects. The wild-type or mutant Cren7 protein at a concentration within a range spanning the K_D value for the interaction of the protein with the DNA was injected over the DNA surface and the blank flow cell for 2 min at a flow rate of 30 μl min^{-1}. After the dissociation phase (2–4 min), bound protein was removed with a 30 s wash with 0.01% SDS, followed by a 60 s buffer injection. Measurement was repeated once at each protein concentration. Equilibrium and kinetic constants were calculated by a global fit to 1:1 Langmuir binding model (BIA evaluation 4.1 software).

Acknowledgements

We are grateful to Yuanyuan Chen for her assistance in SPR analysis. We also thank Peng Cao and Yi Han for their assistance in data collection and structure refinement. This work was supported by grants 30730003 and 30621005 from the National Natural Science Foundation of China and 2004CB719603 from the National Basic Research Program of China to L. H.; and grant 30900023 from the National Natural Science Foundation of China to Z. Z. Funding to pay the Open Access publication charges for this article was provided by grant 30730003 from the National Natural Science Foundation of China.

References

Agback, P., Baumann, H., Knapp, S., Ladenstein, R., and Hard, T. (1998) Architecture of nonspecific protein-DNA interactions in the Sso7D-DNA complex. *Nat Struct Biol* **5:** 579–584.

Anderson, I., Rodriguez, J., Susanti, D., Porat, I., Reich, C., Ulrich, L., *et al*. (2008) Genome sequence of *Thermofilum pendens* reveals an exceptional loss of biosynthetic pathways without genome reduction. *J Bacterial* **190:** 2957–2965.

Baumann, H., Knapp, S., Lundback, T., Ladenstein, R., and Hard, T. (1994) Solution structure and DNA-binding properties of a thermostable protein from the archaeon *Sulfolobus solfataricus*. *Nat Struct Biol* **1:** 808–819.

Bell, S.D., Botting, C.H., Wardleworth, B.N., Jackson, S.P., and White, M.F. (2002) The interaction of Alba, a conserved archaeal, chromatin protein, with Sir2 and its regulation by acetylation. *Science* **296:** 148–151.

Chen, C.Y., Ko, T.P., Lin, T.W., Chou, C.C., Chen, C.J., and Wang, A.H. (2005) Probing the DNA kink structure induced by the hyperthermophilic chromosomal protein Sac7d. *Nucleic Acids Res* **33:** 430–438.

Collaborative Computational Project (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* **50:** 760–763.

Dame, R.T. (2005) The role of nucleoid-associated proteins in the organization and compaction of bacterial chromatin. *Mol Microbiol* **56:** 858–870.

Delano, W.L. (2002) *The PyMOL Molecular Graphics System*, DeLano Scientific, San Carlos, CA, USA.

Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60:** 2126–2132.

Gao, Y.G., Su, S.Y., Robinson, H., Padmanabhan, S., Lim, L., McCrary, B.S., *et al*. (1998) The crystal structure of the hyperthermophile chromosomal protein Sso7d bound to DNA. *Nat Struct Biol* **5:** 782–786.

Gouet, P., Courcelle, E., Stuart, D.I., and Metoz, F. (1999) ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics* **15:** 305–308.

Grant, P.A. (2001) A tale of histone modifications. *Genome Biol* **2:** REVIEWS0003.

Grote, M., Dijk, J., and Reinhardt, R. (1986) Ribosomal and DNA binding proteins of the thermoacidophilic archaebacterium *Sulfolobus acidocaldarius*. *Biochim Biophys Acta* **873:** 405–413.

Guo, L., Feng, Y.G., Zhang, Z.F., Yao, H.W., Luo, Y.M., Wang, J.F., and Huang, L. (2008) Biochemical and structural characterization of Cren7, a novel chromatin protein conserved among Crenarchaea. *Nucleic Acids Res* **36:** 1129–1137.

Guo, R., Xue, H., and Huang, L. (2003) Ssh10b, a conserved thermophilic archaeal protein, binds RNA in vivo. *Mol Microbiol* **50:** 1605–1615.

Hansen, J.C., Lu, X., Ross, E.D., and Woody, R.W. (2006) Intrinsic protein disorder, amino acid composition, and histone terminal domains. *J Biol Chem* **281:** 1853–1856.

Laskowski, R.A., MacArthur, M.W., Moss, D.S., and Thornton, J.M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *Appl Crystallogr* **26:** 283–291.

Lu, X., and Olson, W.K. (2008) 3DNA: a versatile, integrated software system for the analysis, rebuilding and visualization of three-dimensional nucleic-acid structures. *Nat Protoc* **3:** 1213–1227.

Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997) Crystal structure of the nucleosome core particle at 2.8 angstrom resolution. *Nature* **389:** 251–260.

Luo, X., Schwarz-Linek, U., Botting, C.H., Hensel, R., Siebers, B., and White, M.F. (2007) CC1, a novel crenarchaeal DNA binding protein. *J Bacterial* **189:** 403–409.

McAfee, J.G., Edmondson, S.P., Datta, P.K., Shriver, J.W., and Gupta, R. (1995) Gene cloning, expression, and characterization of the Sac7 proteins from the hyperthermophile *Sulfolobus acidocaldarius*. *Biochemistry* **34:** 10063–10077.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007) Phaser crystallographic software. *J Appl Crystallogr* **40:** 658–674.

Mai, V.Q., Chen, X., Hong, R., and Huang, L. (1998) Small abundant DNA binding proteins from the thermoacidophilic archaeon *Sulfolobus shibatae* constrain negative DNA supercoils. *J Bacterial* **180:** 2560–2563.

Marsh, V.L., Peak-Chew, S.Y., and Bell, S.D. (2005) Sir2 and the acetyltransferase, Pat, regulate the archaeal chromatin protein, Alba. *J Biol Chem* **280:** 21122–21128.

Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* **53:** 240–255.

Nowak, S.J., and Corces, V.G. (2004) Phosphorylation of histone H3: a balancing act between chromosome condensation and transcriptional activation. *Trends Genet* **20:** 214–220.

Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* **276:** 307–326.

Reeve, J.N. (2003) Archaeal chromatin and transcription. *Mol Microbiol* **48:** 587–598.

Robinson, H., Gao, Y.G., McCrary, B.S., Edmondson, S.P., Shriver, J.W., and Wang, A.H. (1998) The hyperthermophile chromosomal protein Sac7d sharply kinks DNA. *Nature* **392:** 202–205.

Sandman, K., and Reeve, J.N. (2005) Archaeal chromatin proteins: different structures but common function? *Curr Opin Microbiol* **8:** 656–661.

Wu, L.J. (2004) Structure and segregation of the bacterial nucleoid. *Curr Opin Genet Dev* **14:** 126–132.
Wurtzel, O., Sapra, R., Chen, F., Zhu, Y., Simmons, B.A., and Sorek, R. (2010) A single-base resolution map of an archaeal transcriptome. *Genome Res* **20:** 133–141.

Supporting information

Additional supporting information may be found in the online version of this article.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.