Cover Page

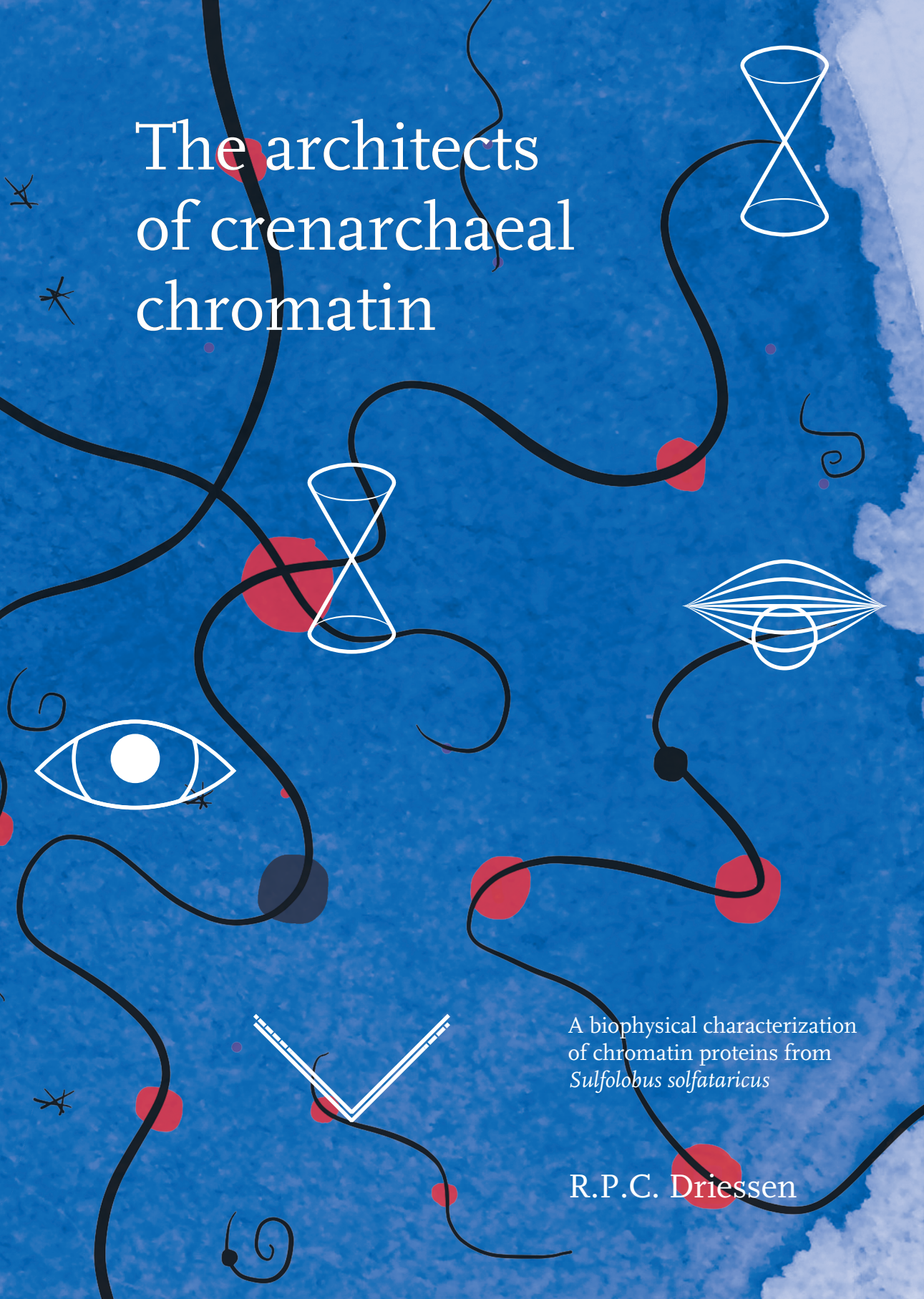
Universiteit Leiden

Leiden University Repository

The handle http://hdl.handle.net/1887/25579 holds various files of this Leiden University dissertation

Author: Driessen, R. P. C.
Title: The architects of crenarchaeal chromatin : a biophysical characterization of chromatin proteins from *Sulfolobus solfataricus*
Issue Date: 2014-05-06

The architects of crenarchaeal chromatin



A biophysical characterization of chromatin proteins from *Sulfolobus solfataricus*

R.P.C. Driessen

The architects of crenarchaeal chromatin

A biophysical characterization
of chromatin proteins from
Sulfolobus solfataricus

R. P. C. Driessen

The architects of crenarchaeal chromatin

A biophysical characterization of chromatin proteins from *Sulfolobus solfataricus*

PROEFSCHRIFT

ter verkrijging van

de graad van Doctor aan de Universiteit Leiden,

op gezag van Rector Magnificus prof. mr. dr. C.J.J.M. Stolker

volgens besluit van het College van Promoties

te verdedigen op dinsdag 6 mei 2014

klokke 15:00 uur

door

Rosalie Paula Catharina Driessen

geboren te Utrecht, Nederland

16 juni 1983

PROMOTIECOMMISSIE

Promotor	Prof. dr. Jaap Brouwer
Copromotor	dr. Remus Th. Dame
Overige leden	Prof. dr. Malcolm White (St Andrews University, UK)
	Prof. dr. John van der Oost (Universiteit Wageningen)
	dr. John van Noort
	Prof. dr. Mathieu Noteborn

Cover image and book design by Marcel Schouwenaar

Cover image inspired by the work of Joan Miró and Sirvan Ugur

Published by: Uitgeverij BOXPress, 's-Hertogenbosch

Copyright © 2014 by R.P.C. Driessen

ISBN: 978-90-8891-866-7

TABLE OF CONTENTS

1. Introduction 7

2. Crenarchaeal chromatin proteins Cren7 and Sul7 compact DNA by inducing rigid bends 39

3. Tethered Particle Motion reveals how temperature affects DNA flexibility and protein-DNA interactions..... 67

4. Alba shapes the archaeal genome using a delicate balance of bridging and stiffening the DNA..... 89

5. Diverse architectural properties of Sso10a proteins: evidence for a role in chromatin compaction and organization..... 115

6. Summary and Conclusions 143

Nederlandse samenvatting..... 149

References 155

Curriculum Vitae..... 181

List of publications 183

Chapter 1

Introduction

Parts of this chapter are based on the following articles:

Driessen, R.P.C. and Dame, R.T. (2011) Nucleoid-associated proteins in Crenarchaea. *Biochem. Soc. Trans.*, **39**, 116-121.

Driessen, R.P.C. and Dame, R.T. (2013) Structure and dynamics of the crenarchaeal nucleoid. *Biochem Soc Trans*, **41**, 321-325.

Archaea represent one of the three domains of life. The discovery of archaea came only in 1977 when Carl Woese identified a separate lineage for archaea, originally thought to belong to the same phylogenetic lineage as bacteria: the prokaryotes (1). Based on his findings he later renamed this classification into the three domains of life as we know it today: Bacteria, Archaea and Eukarya (2) (see Figure 1.1). The archaeal domain is currently divided into five phyla: the Crenarchaea, Euryarchaea (2), Thaumarchaea (3), the Korarchaea (4) and the Nanoarchaea (5). Most of the archaeal species identified so far belong to either the euryarchaeal or crenarchaeal branch. Archaeal species are notorious due to their ability to thrive in extreme environments (6) with high temperatures (thermophiles), high salt concentrations (halophiles), high pressure (in deep ocean, piezophiles), or extremely low pH (acidophiles). However, archaea are also found in much milder environments such as soil, oceans and lakes and constitute at least 20% of the total biomass on earth (7). Even in human niches archaeal species are present and they might play a role in human health and disease (8). Details about their interactions with human cells or other microorganisms within the human body are still unrevealed and whether archaea exist as human pathogens is to date unclear (9).

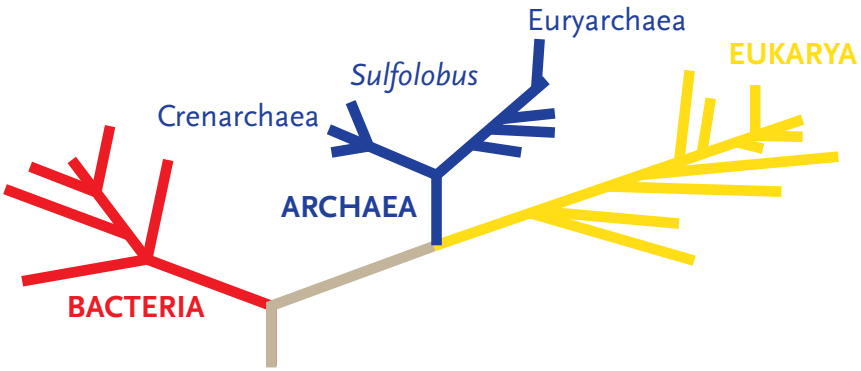


Figure 1.1 Phylogenetic tree showing the three domains of life: Bacteria, Archaea and Eukarya. The two main archaeal phyla are the Euryarchaea and the Crenarchaea. *Sulfolobus* species are part of the crenarchaeal branch.

Ever since the discovery of archaea, a wide diversity of research areas has shown an increasing interest in archaea due to their unique structural and biochemical features (10). The morphology of arch aea strongly resembles that of bacteria, as they are micron-sized unicellular organisms that lack a nucleus. In contrast, many of the cellular processes such as replication, repair and transcription, are based on molecular machineries that reminisce their eukaryotic counterparts (11-13).

This thesis focuses on chromatin organization and dynamics in one of the two main archaeal branches: the Crenarchaea. In particular, different chromatin proteins from the crenarchaeal model organism *Sulfolobus solfataricus* (see Box 1) are characterized using a combination of biochemical and biophysical approaches (see Chapter 2-5).

CHROMATIN COMPACTION AND ORGANIZATION

Organisms of all domains of life need to organize and compact their relatively stiff genomic DNA to reduce its dimensions below that of a cell or a nucleus (see Box 2). Yet the genetic material needs to be flexibly and dynamically organized to permit DNA-based processes such as transcription, replication and repair. Key compaction mechanisms employed by cells are DNA supercoiling, macromolecular crowding and DNA-binding proteins (chromatin proteins) that structure the genomic DNA (14).

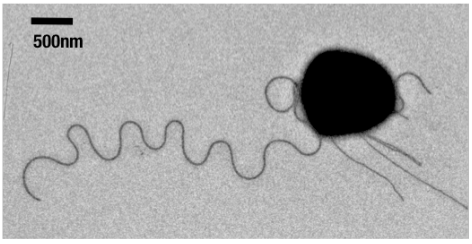
DNA supercoiling

Supercoiling refers to the under- or overwounding of DNA around its own axis. Proteins that bind or move along the DNA can induce supercoiling by changing the torsional stress of the DNA. As a result so-called plectonemes are formed to release the torsional stress from the under- or overwound DNA. As plectonemes are further distributed into branches, this will lead to a significant reduction (about 10-fold) of the volume occupied by the DNA (15,16). The supercoiling state of the chromosomal DNA is dynamically

maintained by a balanced action of enzymes (topoisomerases) introducing negative or positive supercoiling and affects global chromatin conformation and local DNA structure. Negative supercoiling can for instance destabilize the DNA duplex, which facilitates processes such as transcription or replication.

Box 1 – Sulfolobus: the crenarchaeal model organism

Sulfolobus species serve as model organisms to study cellular processes such as transcription, replication and chromatin organization within the crenarchaeal branch. The *Sulfolobus* genus was identified and characterized for the first time in 1972 by Brock and coworkers (17). Thermal soils and acidic hot springs are the natural habitat of *Sulfolobus* species, which are classified as thermoacidophiles. They grow optimally at high temperatures of 75 – 80 °C and in acidic environments of pH 2 – 3. Nevertheless, they are able to keep their cytoplasm close to neutral, around pH 6.5 (18). Since the availability of the complete genome sequences of different *Sulfolobus* species (19-22), many genetic tools have been developed and improved (23). The circular genome of *Sulfolobus* is about 3 Mbp long with an average GC content of ~ 36% and contains about three thousand genes. Different extrachromosomal elements have been identified in *Sulfolobus*, including several plasmids and viruses (24).



Electron microscope image of a *Sulfolobus acidocaldarius* cell. Image credit: Kerstin Lassak.

While the supercoiling state of DNA from bacteria, eukarya and mesophilic archaea is typically negative, the DNA from thermophilic archaeal species is found to be relaxed to positively supercoiled (25-27). Since positively supercoiled DNA is less susceptible to thermal denaturation of DNA, the supercoiling state of the genomic DNA could be an important mechanism to maintain DNA integrity in thermophilic organisms. Indeed reverse gyrase – the only topoisomerase that induces positive supercoiling – has been identified in bacterial and archaeal species that live in thermophilic habitats (28) and has been proposed to play an important role in maintaining genome stability

Macromolecular crowding

The cytoplasm of the cell contains a high concentration of proteins and other macromolecules that occupy up to 20 – 30% of the total cell volume (29). Macromolecules cause a depletion attraction that drives compaction of the chromatin. This force is entropy-driven due to the gain in available solvent for the particles that surround the genomic DNA (30). Besides the direct compaction by this depletion attraction, macromolecular crowding enhances the association of chromatin proteins to DNA, which causes further compaction (31). On the other hand, DNA that is (partly) covered by nucleoid-associated proteins appears to be more sensitive to compaction by macromolecular crowding (16,32). Thus, the synergy between chromatin proteins and macromolecular crowding is likely to be an important factor in achieving compaction of the chromatin.

Chromatin proteins

Chromatin proteins modulate the spatial path of the DNA by wrapping, bridging or bending it (see Figure 1.2). These architectural properties, rather than sequence homology, appear to be conserved among all domains of life (14). Not only do the architectural properties of these proteins explain compaction, they also provide a framework for understanding how different proteins can work together in providing a dynamic genome. The delicate

interplay between architectural proteins permits them to modulate the degree of compaction and thereby regulate the accessibility of the genomic DNA for different DNA-based processes. As the negative charge of the DNA backbone leads to repulsive forces between DNA segments, shielding of this charge by positively charged chromatin proteins is likely to help DNA compaction as well.

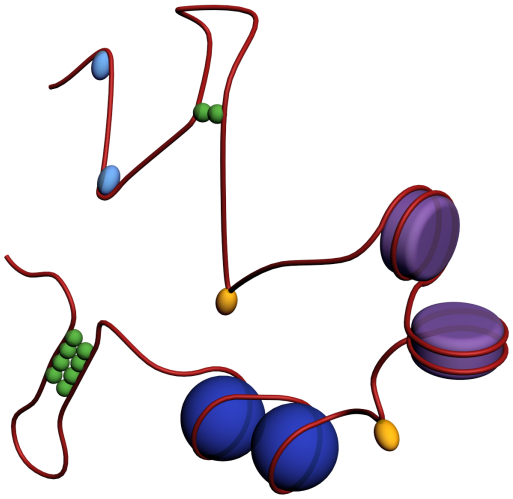


Figure 1.2 Schematic overview of the three classes of architectural proteins: DNA benders (*yellow and light blue*), DNA bridgers (*green*) and DNA wrappers (*purple and dark blue*).

Box 2 – Physical properties and dimensions of genomic DNA

The genetic code of all organisms is encoded in double-helical DNA. Double-stranded DNA (dsDNA) is a semi-flexible polymer, which is stiff on a short length scale of about 150 basepairs or 50 nm, the persistence length, where bending is energetically unfavorable. On longer length scales however, DNA can smoothly bend and behaves as a flexible polymer, which takes the conformation of a random coil. The conformation of a semi-flexible polymer, such as dsDNA, can be described by the so-called worm-like-chain (WLC) model, which describes the average end-to-end distance (R) of the polymer as a function of its persistence length (L_p) and contour length (L_0) (33):

$$\left\langle R^2 \right\rangle = 2L_pL_0\left(1-\frac{L_p}{L_0}\left(1-e^{-\frac{L_0}{L_p}}\right)\right) \qquad (1.1)$$

From this relation, the typical size of a random coil can be calculated as the radius of gyration R_g, which is the average distance of the polymer segments from the center of mass:

$$R_g \approx \sqrt{\frac{L_0L_p}{3}} \qquad (1.2)$$

As the genomic DNA is relatively long (e.g. ~ 2 m for the human genome and ~ 1 mm for bacterial or archaeal cells), the typical volume the genome would occupy in solution is much larger than the volume of the cell or cell nucleus. For example, the genome of *Sulfolobus solfataricus* is about 3 Mb in length which corresponds to a contour length of $L_0 \approx 1$ mm. As the persistence length of DNA is $L_p \approx 50$ nm, coil in solution would have an $R_g \approx 4$ µm. This corresponds to a volume of $V \approx 300$ µm^3, which is about 600 times the typical volume of a *Sulfolobus* cell $V_{cell} \approx 0.5$ µm^3. Therefore, the genomic DNA needs to be compacted extensively to fit into the cell. This compaction is achieved by compaction mechanisms such as macromolecular crowding, supercoiling and chromatin proteins, which are described in the main text.

In eukaryotes, the highly conserved histone proteins are the main architectural proteins. Histone octamers wrap ~ 150 bp of DNA around their surface, yielding nucleosomes (34). Arrays of nucleosomes are subjected to further higher-order compaction with the aid of other (smaller) architectural proteins (35) and nucleosomal interactions. Bacteria lack histone homologues and rely on small chromatin proteins that bend or bridge their genomic DNA into a compact nucleoid structure (36). For instance, in *Escherichia coli* the genome is dynamically organized by the nucleoid-associated proteins HU (histone like protein from *E. coli* strain U93), H-NS (histone-like nucleoid structuring protein), Fis (factor for inversion stimulation) and IHF (integration host factor) (37,38). The situation in archaea is hybrid as archaeal species employ a variety of architectural proteins, functionally homologous to either eukaryotic or bacterial chromatin proteins, to organize and compact their genome (39-43).

Archaeal chromatin organization

Similar to bacteria, archaea are unicellular organisms, which lack a cell nucleus. Their genome, which is in the order of several megabasepairs in size, needs to be compacted into the micrometer-sized dimensions of the cell. Archaea employ a diverse array of chromatin proteins to organize their genome into a compact nucleoid structure. None of these proteins are conserved among all archaeal species, and most of the archaeal chromatin proteins identified to date have a limited phylogenetic distribution (see Figure 1.3). The two main phyla of Archaea, the Euryarchaea and the Crenarchaea, employ different mechanisms to achieve compaction and organization of their genome. While Euryarchaea synthesize true histone homologues, which form tetrameric nucleosome-like structures (44), Crenarchaea lack histone homologues (with some rare exceptions (see Figure 1.3)) and rely on a set of other small proteins (40).

Euryarchaeal histones

Most euryarchaeal species encode two homologues of eukaryotic core histones. Whereas eukaryotic histones assemble as octamers from H2A-H2B and H3-H4 heterodimers, archaeal histones are arranged into tetrameric structures composed of either hetero- or homodimers of H3 and H4 homologues (45). The archaeal tetrameric nucleosome wraps ~ 60 bp DNA around its surface (44), in contrast with the ~ 150 bp found in eukaryotic nucleosomes (34). Archaeal histones lack the N- and C-terminal tails that are subjected to post-translational modifications in their eukaryotic counterparts (46,47). Such modifications are key to modulating chromatin organization in eukaryotes in relation to gene expression. *In vitro* reconstitution of nucleosomes from archaeal histone protein HMf from *Methanothermus fervidus* revealed a chromatin fiber structure similar to the 'bead-on-a-string' structures observed in eukaryotic chromatin fibers (48). Visualization by atomic force microscopy (AFM) of isolated chromatin fibers from *Thermococcus kodakarensis* also showed a bead-on-a-string like structure (49).

The positioning of archaeal nucleosomes *in vivo* has been investigated in several recent studies. Analysis of the nucleosome occupancy in *Haloferax volcanii* showed that these nucleosomes cover ~ 60 bp of DNA, consistent with the assembly of tetrameric nucleosomes (50). In addition, nucleosomes were found to be depleted at transcription start sites, flanked by a higher nucleosome occupancy both upstream and downstream from this region. Similar results were reported on the nucleosome positioning *in vivo* in *T. kodakarensis,* where nucleosomes – covering ~ 60 bp of DNA – are positioned near the promoter (51). In contrast, a similar study on *T. kodakarensis* showed that histones can assemble into chromatin particles of variable sizes covering DNA stretches ranging from 30 bp to > 450 bp, consisting of subunits that cover 30 bp of DNA (the length occupied by a dimer) (52). This coverage indicates that histones can bind as dimers, probably introducing bending of the DNA, or when binding cooperatively

these adjacent dimers could form filaments covering long stretches of DNA. Such a heterogeneous binding pattern indicates a more dynamic beads-on-a-string chromatin structure in *T. kodakarensis,* in contrast to the specific nucleosome positioning observed in eukaryotes and *H. volcanii*.

A few other DNA-binding proteins identified in the euryarchaeal domain have been classified as chromatin proteins. The MC1 protein, expressed by species of the Methanosarcinales order, is a small (~ 10 kDa) basic protein that binds to DNA in a monomeric form (53). It shows some structural similarities to crenarchaeal chromatin proteins Sul7 and Cren7 as it contains two anti-parallel β-sheets (54). Binding of MC1 to dsDNA results in a sharp bending of about 120° (55).

Another small DNA-binding protein that has been characterized is exclusively expressed by *Methanopyrus kandleri*: 7kMK (56). 7kMK is a 7 kDa dimeric protein that has been shown to form bends and possibly small loops when binding to dsDNA. Although the exact role of 7kMK remains unclear, similarities with ribbon-helix-helix (RHH) transcription regulators suggest that this protein might play a role in transcription regulation.

Thermoplasma species express chromatin proteins which are true homologues of the bacterial DNA-bending HU protein (57). Interestingly, these species are the only euryarchaeal species that do not express true histone homologues (see Figure 1.3), suggesting that HU carries out a similar task as histones.

Another novel chromatin protein that has been reported in *T. kodakarensis* is TK047 (58). This protein is about 32 kDa in size and can form thick stiff filaments when binding DNA. Besides a general role in chromatin organization, this protein might play an additional role as a global transcription repressor, as it was shown that it represses transcription when bound to the promoter region.

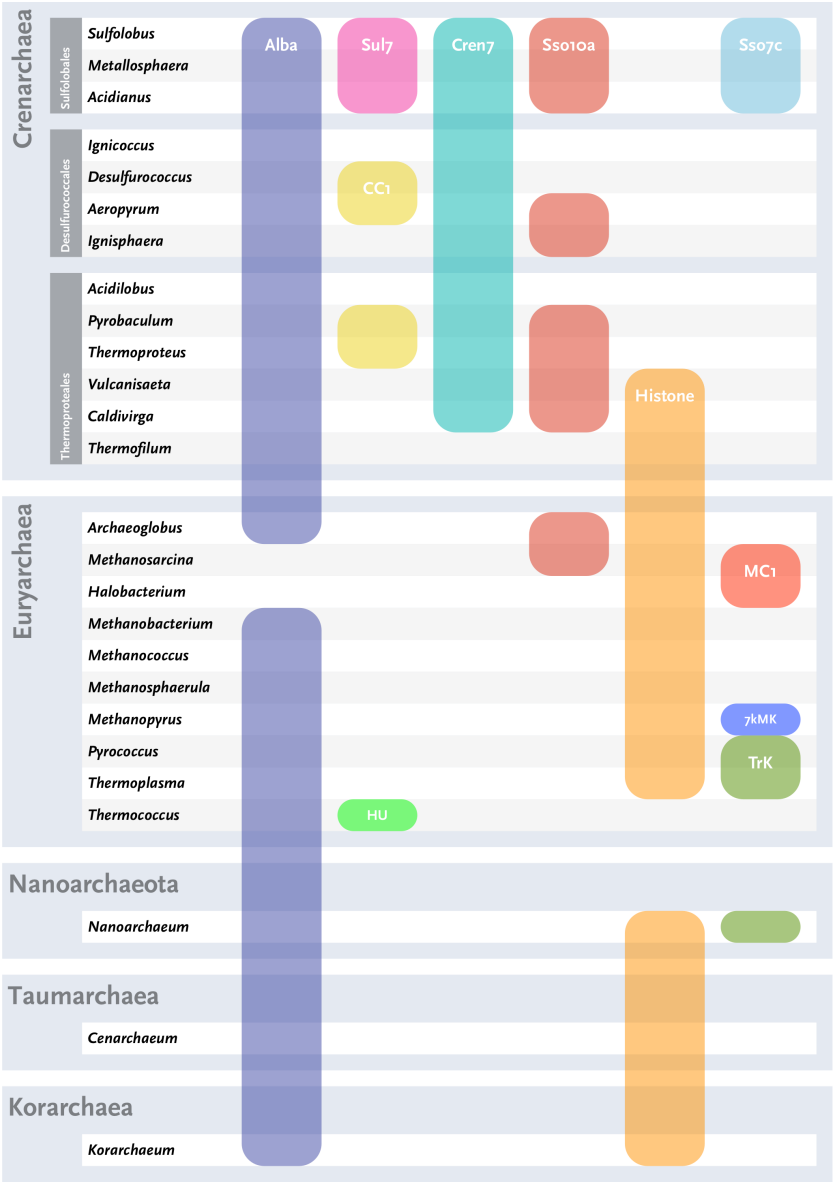


Figure 1.3 Distribution of archaeal chromatin proteins among different archaeal lineages.

Crenarchaeal chromatin proteins

The majority of crenarchaea do not express histone proteins (with the exception of some *Thermoproteales* species), but instead encode a variety of small, basic, abundant architectural proteins (39,40). Although these proteins are not universally conserved among all crenarchaea, each species contains at least two different nucleoid-associated proteins (generally a bender and a bridger) to jointly compact the genomic DNA and modulate the accessibility of the genome. Several potential chromatin proteins have been identified and characterized in the crenarchaeal model organism *S. solfataricus* (59): Alba (previously referred to as Sso10b) (60), Sul7 (referred to as Sso7d in *S. solfataricus*) (61,62), Cren7 (63), Sso10a (64,65) and Sso7c (66,67). Although a general consistent nomenclature would be useful, these proteins are commonly named after the organism from which they originate and their size. All these proteins are small (7 – 10 kDa), highly abundant in the cell, basic and bind to dsDNA with no apparent sequence specificity.

Alba

The most widely distributed chromatin protein in crenarchaea is Alba, which has been identified in all crenarchaeal species sequenced so far. In addition, homologues of the Alba family are found in most euryarchaeal species. This protein was originally identified in *S. acidocaldarius* and *S. solfataricus* as Sac10b and Sso10b, respectively, but later renamed to Alba (Acetylation Lowers Binding Affinity). Alba is a small (10 kDa), basic protein and highly abundant in these species comprising ~ 5% of all cellular protein, corresponding roughly to 1 Alba dimer per 5 bp. Alba exists as a dimer (see Figure 1.4) in solution and binds cooperatively without apparent specificity to DNA. Alba proteins can constrain negative supercoiling (68) and protect dsDNA at high temperatures by increasing its melting temperature (69).

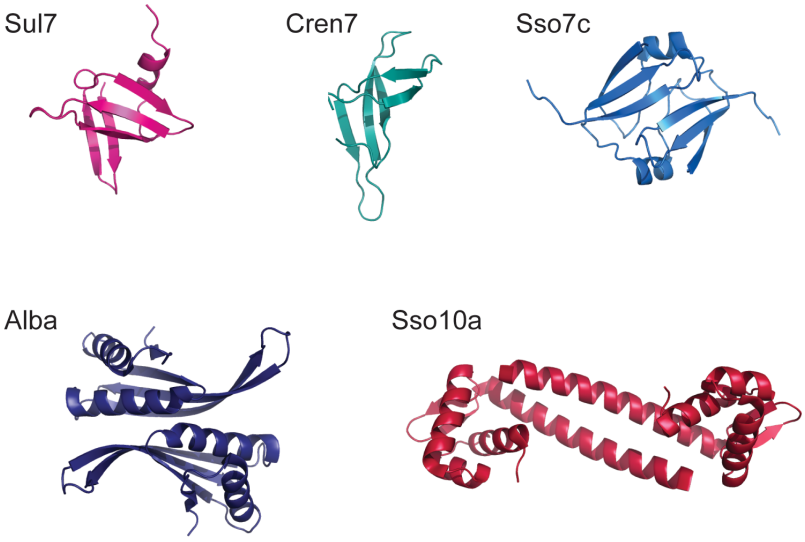


Figure 1.4 Crystal/solution structures of crenarchaeal chromatin proteins Sul7, Cren7, Sso7c, Alba and Sso10a (PDB codes 1SSO (70), 2JTM (63), 2L66 (66), 1HOY (71) and 1RJ7 (64) respectively).

The structure of Alba has been known for several years and has been solved for homologues from different archaeal species (71-79). Recently, also a co-crystal structure of the Alba homologue from *Aeropyrum pernix* bound to DNA has been solved (see Figure 1.5 A) (80). This structure revealed that dimer-dimer interactions of Alba enable bridging of two DNA duplexes. Electron microscopy (EM) and AFM studies have shown the formation of two structurally different types of Alba-DNA complexes depending on the concentration of the protein (76,81)(see also Chapter 4). At low and intermediate concentrations binding of Alba results in compaction due to the formation of looped structures. The ability of Alba to bridge two DNA fragments, as observed in the protein-DNA co-crystal structure (80), could underlie the observed loop formation. At high protein concentrations the

protein fully coats DNA giving rise to stiff filaments, which at first sight do not result in compaction. Interestingly, both binding modes rely on dimer-dimer interactions (76,80).

Binding of Alba to DNA can be modulated by its acetylation/deacetylation at a single residue, Lys16, situated at the DNA binding surface (hence the name Alba: Acetylation Lowers Binding Affinity). A number of archaeal species encode two Alba homologues. For instance, *S. solfataricus* expresses Alba1, which is highly conserved among different archaeal species, and Alba2, which is less conserved. Alba2 is expressed at levels 20 times lower than Alba1 (76) and found uniquely as heterodimer with Alba1 at physiological conditions (76). Differences in DNA binding observed for the Alba heterodimer and homodimer can be attributed to changes in dimer-dimer interactions.

Several studies have shown that Alba proteins can bind to RNA, suggesting that Alba may have divergent functions (82). Alba proteins from *S. solfataricus* were found to bind both to DNA and RNA *in vivo* (83), pointing to an additional function in regulation besides chromatin organization. Another Alba homologue, Sso10b from *Sulfolobus shibatae*, showed similar binding affinities to dsDNA, ssDNA and RNA *in vitro* but was only found to be associated to RNA *in vivo* (84). These findings, together with the structural similarity between Alba and RNA-binding proteins, have led to the suggestion that Alba is involved in RNA metabolism *in vivo* (85). A few Alba homologues that lack the key residue (Lys16) in modulating Alba binding to DNA via acetylation have been identified. Ape10b2 from *A. pernix* and Mma10b from *Methanococcus maripaludis* are the only examples characterized (86,87). Due to the absence of Lys16 the DNA-binding affinity of these Alba homologues cannot be modulated by changing the acetylation status at this site. Interestingly, Mma10b binds with sequence specificity to DNA; its binding site corresponds to an AT-rich palindrome of 18 bp (87). It is less abundant (~ 0,01% of total cellular proteins) than other architectural proteins, and has a low affinity for generic DNA. It has been shown in

chromatin immunoprecipitation (ChIP) studies that Mma10b associates with coding regions of genes. Deletion mutants of Mma10b in this species show up/down-regulation of some genes, similar to Mvo10b deletion mutants in *Methanococcus voltae* (88). The relatively low expression levels of Mma10b compared to Alba, together with its sequence-specific binding suggest that Mma10b has evolved towards to a more specific role in transcription regulation deviating from the primary role of Alba proteins in genome compaction. Yet another Alba homologue, Mth10b from the euryarchaeon *Methanobacterium thermoautotrophicum*, was shown to be an acidic protein and did not show binding to either DNA or RNA *in vitro*, suggesting a different physiological function of this protein (89).

Sul7 and Cren7

Sulfolobus species express a number of highly conserved architectural proteins of about 7 kDa, among which the Cren7 and Sul7 proteins. Both proteins are expressed at high levels (~ 1% of the total cellular protein for Cren7 and up to 5% for Sul7). *Sulfolobus* species express one up to four Sul7 proteins, which are almost identical, and Cren7 is encoded by one up to five almost identical genes in different crenarchaeal species. Both proteins are highly basic and bind to DNA without apparent sequence specificity. Although Sul7 is limited to Sulfolobales species, Cren7 is found in almost all crenarchaeal species (see Figure 1.3).

Cren7 and Sul7 are completely different at the level of amino-acid sequence, yet very similar in structure and – not unexpectedly – in biochemical and architectural properties (61-63,70,90-94)(see also Chapter 2). Both proteins are folded similarly, containing two antiparallel β-sheets. The main difference is an extended flexible loop in Cren7 (63) and a C-terminal α-helix of Sul7 (see Figure 1.4) (70,91). Cren7 and Sul7 exist as monomers in solution and the monomer is also the form in which they bind to DNA. Both proteins bind dsDNA by intercalation of hydrophobic side chains into the minor groove, which results in bending of DNA (see Figure 1.5 A) (95-101). Binding of Sul7 to DNA is facilitated by insertion of one of the β-

sheets into the minor groove with concomitant intercalation of residues Val26 and Met29 (95,97,98,100). Co-crystal structures have shown that the binding of the protein results in a DNA bend of about 65° (see Figure 1.5 A) (97,100). Cren7 binds to DNA through one of the β-sheets and the flexible loop (96,101). Its DNA binding is similar to that of Sul7 with the β-sheet inserted into the minor groove and residues Leu28 and Val36 intercalating. As a consequence, binding of Cren7 to DNA induces a bend of about 53° as is shown in the co-crystal structures (see Figure 1.5 A). Molecular dynamics (MD) simulations suggest that the bending angles of these proteins are actually very similar, with a bending angle of 46° (90) (see also Chapter 2). The binding of these proteins affects DNA topology as well, as the distortion of the DNA results in the constraint of negative supercoiling (63,94).

Cren7 and Sul7 are posttranslationally modified by methylation of several lysine residues in the N-termini of both proteins, the C-terminal part of Sul7 and the looped domain of Cren7 (61,63). In contrast with Alba, these modifications do not change the DNA-binding affinity of the protein. Instead, they might help to protect the protein from thermal denaturation as Sul7 methylation is found to increase after heat-shock (62). Molecular dynamic simulations have shown that Sul7 enhances thermal stability of dsDNA (102,103), which was also proposed for Cren7 (104). As a result, DNA-Sul7 and DNA-Cren7 complexes have an increased melting temperature compared to bare DNA (63,105). Such protection of dsDNA against denaturation is relevant in the light of the hyperthermophilic natural habitat of *Sulfolobus* species (55 – 95 °C).

Some additional functions have been proposed for Sul7. Sul7 has been shown to ensure integrity and activity of other proteins *in vitro* by disaggregating proteins, which have been denatured by high temperatures in an ATP hydrolysis dependent manner (106,107). Sul7 has also been shown to be able to repair UV-induced damage *in vitro*. Upon exposure to UV, when bound to a thymidine dimer, a conserved tryptophan residue at the DNA-binding interface is oxidized and acts as an electron donor. This

electron can subsequently be transferred to a thymidine dimer, which leads to the reversal and thus repair of the thymidine dimer (108). It remains to be shown whether these aforementioned activities are relevant *in vivo*.

Sso10a

Sso10a homologues are widely distributed among crenarchaea and also encoded in some euryarchaeal species. Interestingly, *Sulfolobus* species possess at least two and up to six genes encoding for Sso10a homologues. Little is known about Sso10a at the biochemical level and about its contributes to genome compaction. However, some characteristic features are evident in the structure of the protein. The protein exists in solution as homodimer and its dimerization occurs via an anti-parallel coiled-coil (see Figure 1.4). The DNA-binding winged helix structures are positioned on opposite sides of the coiled-coil (64,65,109). Chen *et al.* have proposed a model, based on analogy to the structure of a γδ-resolvase, where an Sso10a dimer binds to DNA with the two DNA-binding domains, which results in bending of the DNA. In addition, this protein has been shown to bring DNA duplexes together in EM studies (81), possibly due to bridging. A detailed characterization of two Sso10a homologues from *S. solfataricus* is described in Chapter 5 of this thesis.

Sso7c

Sso7c is yet another 7 kDa protein from *S. solfataricus* that has been suggested to be involved in chromatin organization (66,67,81). Two (almost identical) homologous versions of this protein are expressed in all species within the Sulfolobales order. These homologues only differ in a few amino acids, with the most noticeable difference being an extra C-terminal tryptophan in one of them. Nothing is known about Ssoc7 abundance within the cell, but the transcriptome of *S. solfataricus* reveals that at least one of the two homologues is transcribed at a level comparable to that of Cren7 (110). Sso7c forms a dimer in solution (see Figure 1.4) and binds non-specifically to the major groove of DNA. A model according to which Sso7c bends DNA has been proposed (66). However, whether and how this

protein functions in chromatin organization is currently unclear. At the level of secondary structure, it exhibits homology with several bacterial and archaeal proteins described as transcription factors, which might suggest a similar role for Sso7c.

CC1

CC1 (crenarchaeal chromatin protein 1) is a small (6 kDa), abundant, monomeric, basic DNA-binding protein found in a limited number of species: *Thermoproteus tenax, Pyrobaculum aerophilum* and *A. pernix*. It was discovered in a search for single-stranded DNA binding proteins in archaeal species and found to bind single-stranded and double-stranded DNA with equal affinities (111). The only structural information available is that CC1 is rich in β-sheet structure. It is unknown how this protein binds to DNA, but it constrains negative supercoiling and increases the melting temperature of duplex DNA (86).

CRENARCHAEAL CHROMATIN ORGANIZATION ON HIGHER-ORDER LEVELS

Structural and single-molecule studies have provided useful insights into the protein-DNA interactions of Sul7, Cren7 and Alba on the scale of 10 – 100 basepairs. However, how these proteins cooperate on a megabasepair scale to fold a genome of about 1 mm in length into a compact and organized nucleoid of about 1 μm^3 is still unknown. In bacteria it has been shown that the genome is organized in looped domains, which might be due to bridges formed by H-NS (112,113). Interestingly, Alba exhibits DNA-binding modes (*cis* and *trans*) similar to those of H-NS (114-116), which could permit this protein to act analogous to H-NS. Alba might thus facilitate the formation of higher-order structured loops in archaea. Other proteins that could be involved in structuring looped domains are structural maintenance of chromosomes (SMC) proteins (e.g. cohesin in eukaryotes (117) and MukBEF in *E. coli* (118)). SMC-like proteins have been identified

throughout all archaeal branches (119) (including *Sulfolobus* species (120)) and are thus likely to be involved in the higher-order chromatin organization of archaea. Looped structures – formed by either Alba or SMC proteins – could be further organized and compacted by the action of other chromatin proteins such as Sul7 and Cren7 (see Figure 1.5 B).

Differential expression of chromatin proteins is a possible mechanism to dynamically change global chromatin structure and adapt to growth phase or environmental conditions. Bacteria exhibit different expression patterns of chromatin proteins depending on growth phase (121). For instance, Fis levels are high during fast growth but they sharply drop upon transition to stationary phase, giving rise to high expression of other chromatin proteins, CbpA and DPS (121,122). High expression of these proteins may be related to the observed drastic changes in chromatin structure in stationary phase, highly condensing DNA into a crystalline structure (123). Differential expression of chromatin proteins could play an important role in dynamically shaping the genome in archaea, as well. In fact, euryarchaea synthesize different homologous histone proteins that are able to form tetramers with varying DNA binding affinities depending on its subunit composition (124). The changes in expression patterns of these subunits could therefore lead to a change in histone coverage and hence be involved in regulating chromatin structure and/or gene expression. The transcription of several genes encoding chromatin proteins in *S. acidocaldarius* has also been shown to be cell-cycle dependent (125). While the transcription levels of Cren7 and Sul7 vary significantly at different stages during the cell cycle, the expression level of Alba proteins is relatively stable. It has also been shown that the degree of compaction of the nucleoid of *Sulfolobus*, as seen by light microscopy, changes during different growth phases and appears to be less compact in the stationary phase of growth (126), which is likely related to differential expression of chromatin proteins. Surprisingly, this is opposite to the situation in *E. coli*, which expresses specific chromatin proteins at stationary growth phase to achieve a highly condensed nucleoid.

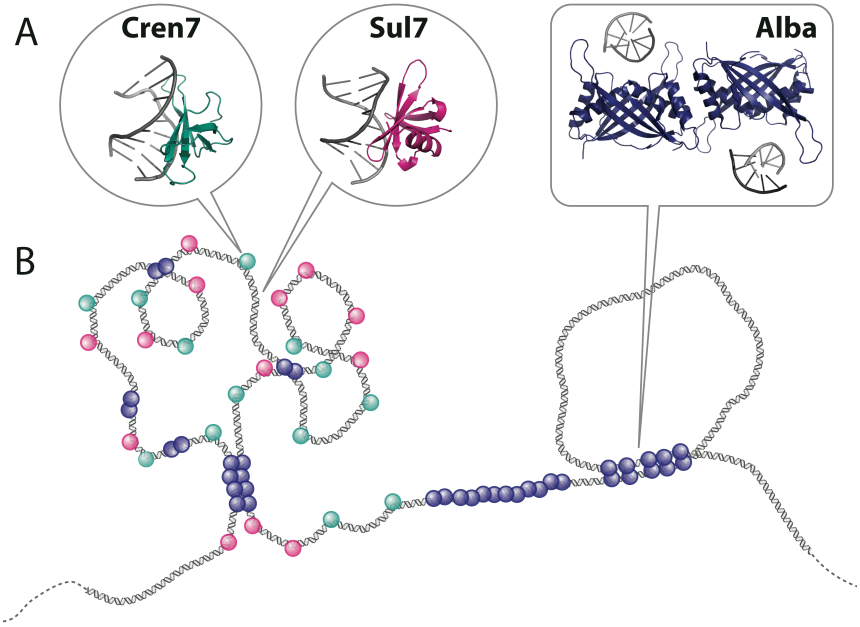


Figure 1.5 Model of the action of DNA benders Cren7 and Sul7 and DNA bridger Alba in higher-order chromatin organization of the circular crenarchaeal genome at different scales **A.** Co-crystal structures (1 – 10 nm scale) show that Cren7 and Sul7 bend DNA by ~ 50° and Alba bridges two DNA duplexes by dimer-dimer interactions forming a small looped structure (PDB codes 3LWH (101), 1BNZ (97) and 3U6Y (80) respectively). **B.** Alba forms looped structures on the order of thousands of base pairs by holding distant positions along the DNA duplex together (100 – 1000 nm scale). The looped domain could be compacted further by the action of Alba, Sul7 and Cren7 at the shorter scales described in (A). At the 10 – 100 nm scale, Alba forms filamentous patches by binding closely side-by-side. Both Alba-DNA bridges and Alba-coated filaments could lead to gene repression by blocking transcription (initiation).

REGULATION OF ARCHAEAL GENE EXPRESSION

The archaeal transcription machinery

Archaeal transcription shows an interesting mix of eukaryotic and bacterial features (41,127,128). While the archaeal basal transcription machinery closely resembles that of eukarya, most transcription regulators are related to bacterial regulatory proteins (129,130). The core component of the archaeal transcription machinery, the RNA polymerase (RNAP), is composed of 11 – 12 subunits that are functional and structural homologues of the eukaryotic RNAPII subunits (11,131). Two general transcription factors are required for reconstitution of transcription *in vitro* (132,133): TATA-box binding protein (TBP) and transcription factor B (TFB), both closely related to their eukaryotic counterparts. Transcription is initiated by binding of TBP to the TATA-box sequence upstream of the transcription start site (134,135). TFB binds to the TBP-promoter complex and recognizes a purine-rich element, the B recognition element (BRE), located upstream of the TATA-box (136). Given the symmetry of TBP, the interaction between TFB and the BRE element is crucial in the determination of the orientation of transcription (137). The final step of transcription initiation is accomplished by recruitment of RNAP by TFB (138) (see Figure 1.6). A third basal transcription factor, transcription factor E (TFE), is not essential for transcription *in vitro*. It does however exhibit a stimulatory effect on transcription *in vitro* under some conditions most probably by facilitating or stabilizing TBP binding to the TATA-box (139). Interestingly, TFE complements a TBP mutant which is unable to recruit RNAP, pointing to a synergistic action between the two factors during transcription initiation (140). TFE plays an additional role in stabilization of the transcription bubble by stimulating DNA melting (141). Some archaea encode multiple homologues of TBP and/or TFB, which suggests a possible global gene regulation mechanism by differential binding of these transcription factors to specific promoters (127,142).

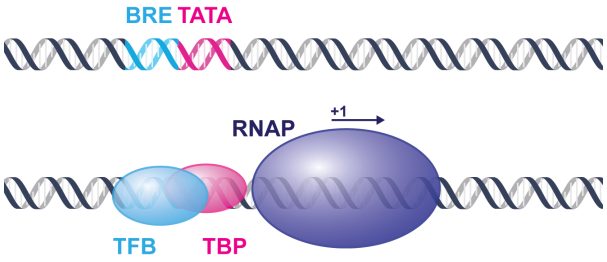


Figure 1.6 Schematic representation of archaeal transcription initiation, showing the minimal preinitiation complex required for *in vitro* transcription. Binding of the transcription factors TBP and TFB to the TATA-box and BRE respectively leads to recruitment of the RNAP, which will initiate transcription.

Mechanisms of transcription regulation

Most putative transcriptional regulators identified in archaeal genomes up to date are homologues of bacterial regulators (130,143), possessing a helix-turn-helix (HTH) DNA-binding domain (129). Common mechanisms of transcription regulation are 1) repression of transcription by inhibition of transcription initiation by promoter occlusion of TBP/TFB or RNAP binding or 2) transcription activation by stimulating recruitment of general transcription factors close to the promoter site or the RNAP to the promoter site itself. Molecular details of the regulatory mechanisms of only a few archaeal regulators have been identified so far. The most widespread transcriptional regulators found in archaea belong to the Lrp/AsnC family of regulators (144). Members of the Lrp (Leucine-responsive Regulatory Protein) family are important regulators in bacteria that respond to specific amino acid levels, of which the prototype Lrp in *E. coli* is a global regulator known to affect the expression of many genes (up to 10%) (145). In general, archaeal Lrp-like regulators bind in close proximity to the promoter site and

– depending on their exact positioning – function either as transcriptional repressors or activators (143,146). An example of a transcriptional repressor is the LrpA protein from *P. furiosus*, which binds to the transcription start site of its own promoter (147). Binding of LrpA to the promoter site inhibits recruitment of the RNAP by sterically shielding the binding site of RNAP (148). Several other archaeal repressors have been identified and shown to be auto-regulators. Lrs14 from *S. solfataricus* inhibits transcription of its own gene *in vitro* by affecting the accessibility of the TATA-box and BRE site for the transcription factors TBP and TFB (149,150). A similar repression mechanism is proposed for transcriptional repressors Ss-Lrp from *S. solfataricus* (151), Sa-Lrp from *S. acidocaldarius* (152), FL11 from *Pyrococcus OT3* (153) and LrpA1 from *H. salinarium* (154) as they all inhibit transcription initiation by binding to their own promoter region. However, it has not been shown experimentally that the observed transcriptional repression is obtained by occlusion of the basal transcription factors TBP and TFB. Transcription regulators that act as activators generally affect the transcription initiation complex by stimulating recruitment of basal transcription factors, either by direct protein-protein interactions or via a change in DNA configuration at the promoter site. For example, the activator Ptr2 from *Methanococcus jannaschi*i stimulates transcription through recruitment of TBP (155). LysM from *S. solfataricus* has been proposed to function as an activator as it binds upstream of the target sites for TBP, TFB and RNAP (156). However, the mechanism by which activation occurs has not been revealed. A special case is Ss-LrpB from *S. solfataricus*, which functions either as repressor or as activator depending on its concentration (157). At low concentration binding of a single dimer to a specific binding site can lead to a stimulatory effect on transcription, while at high protein concentrations multiple dimers bind cooperatively, which leads to wrapping of the DNA around the oligomeric protein. Formation of such a nucleoprotein complex could result in repression of transcription (158).

Regulation of gene expression by chromatin proteins

Although transcription regulators generally bind to specific sites close to promoter regions, they can also bind non-specifically to affect DNA structure on a more global scale. On the other hand, chromatin proteins can play an important role in gene regulation, either on a local or global scale. Both transcription regulators and chromatin proteins are thus involved in chromatin organization and gene regulation, and the functions of these two classes of proteins overlap. Since (archaeal) chromatin proteins are abundant and bind with little or no apparent sequence specificity it can be envisioned that these proteins have an effect on transcription on a global scale, analogous to the chromatin proteins in bacteria (38,159). Based on the analogy with bacteria, some mechanisms are discussed that could be involved in gene regulation in crenarchaeal species.

Most chromatin proteins are able to constrain supercoils, which can locally affect the supercoiling state of the genome. Supercoiling and global gene expression are both responsive to changes in environmental conditions, which suggests that supercoiling can affect gene expression on a global scale (160,161). Indeed supercoiling was found to influence gene expression in *E. coli* by affecting local DNA structure (162,163). The interplay between chromatin proteins, topoisomerases and RNAPs determines the superhelical state of the bacterial or archaeal chromatin, which could possibly affect the expression of many genes sensitive to supercoiling.

Bacterial chromatin proteins such as IHF, Fis and H-NS can act as transcription regulators by binding at preferred sites adjacent to or overlapping with promoters (164). Binding of a chromatin or transcription regulatory protein or a combination of several proteins at the promoter region can lead to activation of transcription by interaction with basal transcription factors or direct recruitment of the RNA polymerase. Enhancement of transcription can also be achieved by inducing a conformational change of a promoter. Whether and how chromatin

proteins in crenarchaea act as transcriptional regulators at specific promoters is currently not known. Only one example is known of chromatin proteins binding to a specific region near a promoter: chromatin proteins Sul7 and Alba as well as transcription regulator Lrs14 were found to bind to a specific site upstream of the gene which encodes an alcohol dehydrogenase (165). How these proteins exactly affect the transcription of this gene was however not revealed.

Direct repression of transcription by occlusion of RNA polymerase from promoter regions by chromatin proteins is a simple mechanism employed in bacteria. For instance in Gram-negative bacteria H-NS is known to act as a global regulator by binding specifically to AT-rich promoter regions overlapping with promoters (166). Like H-NS proteins, Alba is able to form protein-DNA filaments (see Figure 1.5 B), which could interfere with transcription and other cellular processes. Indeed, Alba has been shown to be able to modulate DNA accessibility *in vitro* (167,168). Specificity in gene repression by Alba could be achieved via a DNA sequence preference in binding of itself or a synergistic action of Alba combined with another chromatin protein or transcription regulator, analogous to H-NS.

In many cases the interplay between different chromatin proteins with specific architectural features and/or transcription regulators influences transcription. Chromatin immunoprecipitation experiments have shown that co-localization of architectural proteins along the genome is frequent in *E. coli* (164). Chromatin proteins may act in concert, stabilizing each others action, or counteract each other (14). For example, a bender may stabilize a bridged loop by binding within the loop. Such synergistic nucleo-protein structures have been shown for IHF, Fis and H-NS in bacteria (169). Many repressory complexes in bacteria involve DNA looping on short scales, which requires sharp DNA bending. Indeed, HU is involved in stabilization of many repression loops in *E. coli* (170,171). For example, in the bacterial Gal repressosome, two GalR repressors form a small loop facilitated by binding of a HU dimer in between (172). On the other hand, co-localization

of different architectural proteins can also counteract repression, which was shown for the action of DNA-bender IHF that counteracts repression by DNA-bridger H-NS (173). It is likely that architectural proteins in crenarchaea also act in concert or interact with specific regulators to affect transcription. A bender (Sul7 or Cren7) could bind adjacent to Alba and either support repression by stabilizing a bridged loop, or relieve repression by Alba coated filaments by interfering with the cooperative side-by-side binding of Alba. Genome-wide association studies addressing localization of these chromatin proteins *in vivo* followed by more detailed *in vivo/vitro* studies of specific genes could provide information on whether such mechanisms exist in crenarchaea.

Two mechanisms have been described that could fine-tune repression by chromatin proteins as described above. First, the DNA-binding properties of chromatin proteins can be altered by post-translational modifications (e.g. acetylation or methylation). Such modifications constitute an important mechanism in relation to gene regulation in eukaryotes in which posttranslational modifications modulate the DNA-binding properties of histones and interactions between adjacent nucleosomes. While there is little evidence to date of posttranslational modification of chromatin proteins in bacteria, proteins in archaeal organisms are generally extensively post-translationally modified (174). For instance, Alba-DNA interactions are altered by the acetylation of a single lysine residue in the DNA-binding interface of Alba (71,167), which reduces DNA binding affinity. Acetylation of Alba by acetyltransferase Pat (83) – reducing the positive charge of the binding surface – decreases the binding affinity to DNA about 30-fold (71,167). The action of Sir2 leads to deacetylation of this residue and restores the DNA binding affinity of Alba. Switching of the acetylation status at Lys16 could provide a mechanism to regulate the binding affinity of the protein and thus its ability to repress transcription. Indeed, non-acetylated Alba represses transcription *in vitro* by cooperatively covering the DNA, which can be relieved by acetylation of Alba (83,167). In contrast, Sul7 had no effect on transcription *in vitro* at similar concentrations (167). In

addition to hindering transcription, binding of Alba to DNA likely affects any other type of DNA-based processes. For instance, it has been shown that progression of the helicase MCM (mini-chromosome maintenance) is impeded on Alba-DNA complexes *in vitro*. Reduction of binding affinity of Alba by acetylation could make the DNA more accessible and therefore allow MCM to unwind DNA (168). The switching of the acetylation status of Alba thus emerges as an important mechanism in modulating DNA accessibility.

Lysine methylation is another common posttranslational modification found in archaea, especially in thermophilic crenarchaea (175). Cren7 and Sul7 are both known to be methylated at several lysine residues (62,63). Recently the methyltransferase aKMT4, responsible for lysine methylation, has been identified (176). Lysine methylation of Sul7, but not that of Cren7, is significantly enhanced by the presence of DNA (177). However, since methylation does not alter the DNA binding affinity *in vitro*, it remains unclear whether and how these modifications alter the function of Sul7 and Cren7. Besides a proposed role in enhancing protein thermostability, a plausible function of these modifications could be that they affect protein-protein interactions with either other chromatin proteins or transcription regulators.

A second mechanism to alter the DNA-binding properties of chromatin proteins is expression of multiple homologues of a specific chromatin protein. For instance, *Sulfolobus* species express two Alba homologues, two up to six Sso10a homologues but only one Cren7 homologue (based on the available *Sulfolobus* genome sequences (19-22)). Sul7 is encoded in two up to four genes which are either identical or differ in only a few amino acids. A clear example of this mechanism is the interplay between the Alba proteins Alba1 and Alba2, which changes the structural effects of Alba *in vitro* and might be linked to gene regulation (76). Differential expression of Alba1 and Alba2 would change the ratio of Alba homo- and heterodimers within the cell. Since heterodimers exhibit weaker dimer-dimer interactions

as compared to Alba homodimers (76), this could make the DNA more accessible when the level of heterodimers increases. The relative levels of Alba1 and Alba2 have indeed been shown to affect the balance between bridged and stiffened DNA *in vitro* (178) (see also Chapter 4).

ARCHITECTURAL PROPERTIES OF CHROMATIN PROTEINS: A SINGLE-MOLECULE APPROACH

This thesis describes the characterization of chromatin proteins from *Sulfolobus solfataricus* using a range of single-molecule techniques. Biophysical single-molecule techniques have been extensively used over the past decade to get a detailed understanding on the binding properties of these proteins and how they alter the mechanical properties of DNA (179).

Micromanipulation techniques such as magnetic and optical tweezers probe the force response of a single DNA molecule and the effect of protein-DNA binding (see Figure 1.7 A and B). The typical force range which can be applied on a DNA molecule depends on the used technique and is about 0.1 – 100 pN for optical tweezers and 0.001 – 100 pN for magnetic tweezers (180). These techniques are often used to quantify the architectural properties of chromatin proteins as the force response gives a typical signature of generic properties such as bending, bridging or wrapping (181). Mechanical parameters such as the persistence length and contour length can be extracted from a force-extension curve obtained from such an experiment, by fitting to a worm-like-chain (WLC) model. For example, proteins that bend the DNA will reduce the apparent persistence length of DNA (see Figure 1.7 A), while DNA stiffening will result in an increase of apparent persistence length. DNA bridging proteins often reduce the effective contour length of a single DNA molecule due to the formation of small loops. Exerting a force on the DNA can lead to disruption of the loops, coinciding with peaks in the force response (see Figure 1.7 B, *red curve*).

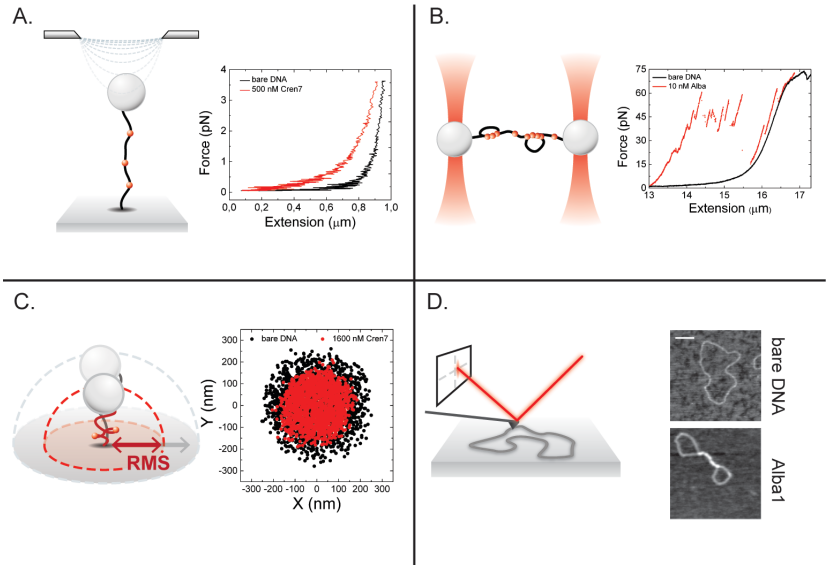


Figure 1.7 Overview of single-molecule techniques used to characterize architectural properties of chromatin proteins and representative data obtained using these techniques. **A. Magnetic tweezers**. A DNA molecule is attached to a surface on one end and a paramagnetic bead on the other end. Force is applied by a magnetic field and altered by changing the height of the magnets. *Right*: Force extension curves of a bare DNA molecule (*black*) and a DNA molecule covered with a DNA bending protein (e.g. Cren7) (*red*), which reduces the apparent persistence length. **B. Optical tweezers**. A single DNA molecule is attached to polystyrene beads at each end, which are trapped in two optical traps. Disruption of loops formed by a DNA bridging protein (e.g. Alba) results in peaks in the force response (*red curve*). Extension of a bare DNA molecule is represented by the black curve. **C. Tethered particle motion** (TPM). The excursion of a bead tethered to a single DNA molecule attached to a surface is monitored and quantified by the root-mean-square distance (RMS). Compaction of the DNA molecule by a chromatin protein (e.g. Cren7, *red data points*) results in a decrease of the RMS. **D. Atomic force microscopy** (AFM). A flexible cantilever with a sharp tip scans the sample surface of DNA molecules immobilized on a flat mica surface. The deflection of the cantilever is probed by a laser spot reflected by the tip of the cantilever onto a quadrant photodiode. This results in a high-resolution image of single DNA molecules or protein-DNA complexes (*right*, scale bar represents 100 nm).

Tethered particle motion (TPM) is another simpler technique, which monitors the excursion of a bead tethered to a single DNA molecule, which is attached to the surface of a flow cell (see Figure 1.7 C). Although no external force is applied on the DNA molecule, surface interactions of the tethered bead lead to small force in the order femtonewtons (182). As the DNA molecule constrains the motion of the bead, changes in the DNA configuration will lead to a change in the excursion of the bead, which is expressed in terms of the root-mean-square distance (RMS). For instance, compaction of the DNA molecule will lead to a reduction of RMS by reducing the effective contour length or DNA stiffness. Details on the compaction mode (whether it is due to bending, bridging or wrapping) can however not be revealed using this technique. An advantage of this technique is that it is relatively easy and less time-consuming compared to micromanipulation experiments. This enables to measure large numbers of individual DNA molecules, which improves statistics and can also reveal different populations of protein-DNA complexes.

To visualize single DNA molecules or single protein-DNA complexes, atomic force microscopy (AFM) is a powerful technique. Atomic force microscopy permits imaging of single DNA molecules deposited on a flat (mica) surface up to sub-nanometer resolution (see Figure 1.7 D). Visualization of protein-DNA complexes can give direct insight on the architectural properties of the chromatin protein (see Figure 1.7 D). If the size of the protein is large enough to be distinguished, this allows determination of the location of the protein along the DNA and its structural effect on the DNA (e.g. bending angle).

These single-molecule techniques, complemented with biochemical tools and molecular dynamics simulations provided a comprehensive framework to the studies described in this thesis, characterizing the DNA-binding and architectural properties of crenarchaeal chromatin proteins. Clearly our understanding on how crenarchaeal chromatin is dynamically organized and regulated is limited. The studies described in this thesis have revealed

many of the architectural properties of chromatin proteins *in vitro*, but little is known about how these proteins act *in vivo* and how the interplay between these proteins contributes to modulating the structure of the genomic material and gene expression. Further studies on the architectural properties and functions of crenarchaeal chromatin proteins – both *in vivo* and *in vitro* – will help to expand and refine our current model of chromatin organization in archaea.

CHAPTER 2

Crenarchaeal chromatin proteins Cren7 and Sul7 compact DNA by inducing rigid bends

This chapter has been published as:

Rosalie P.C. Driessen, He Meng, Gorle Suresh, Rajesh Shahapure, Giovanni Lanzani, U Deva Priyakumar, Malcolm F. White, Helmut Schiessel, John van Noort and Remus Th. Dame, *Nucleic Acids Research*, **41**, 196-205, 2013

ABSTRACT

Archaeal chromatin proteins share molecular and functional similarities with both bacterial and eukaryotic chromatin proteins. These proteins play an important role in functionally organizing the genomic DNA into a compact nucleoid. Cren7 and Sul7 are two crenarchaeal nucleoid-associated proteins, which are structurally homologous, but not conserved at the sequence level. Co-crystal structures have shown that these two proteins induce a sharp bend on binding to DNA. In this study, we have investigated the architectural properties of these proteins using atomic force microscopy, molecular dynamics simulations and magnetic tweezers. We demonstrate that Cren7 and Sul7 both compact DNA molecules to a similar extent. Using a theoretical model, we quantify the number of individual proteins bound to the DNA as a function of protein concentration and show that forces up to 3.5 pN do not affect this binding. Moreover, we investigate the flexibility of the bending angle induced by Cren7 and Sul7 and show that the protein–DNA complexes differ in flexibility from analogous bacterial and eukaryotic DNA-bending proteins.

INTRODUCTION

Organisms in all three domains of life need to compact and functionally organize their genomic DNA into the relatively small volume of a nucleus or a cell. Architectural proteins (histones and other chromatin proteins in eukaryotes and nucleoid-associated proteins in bacteria and archaea) play an important role both in compaction and functional organization of DNA, thus affecting DNA transactions as diverse as transcription, repair and replication (14).

In eukaryotes, DNA is wrapped around histone octamers, forming nucleosomes. With the aid of other chromatin proteins, fibres with nucleosomes are folded into higher-order structures, obtaining multiple levels of organization. Bacteria organize their genomic DNA into a nucleoid, which is shaped by the action of small chromatin proteins that bend or bridge the DNA (e.g. HU, IHF, Fis and H-NS) (36,37). Archaea, constituting the third domain of life, synthesize numerous nucleoid-associated proteins with molecular and functional similarities to both bacterial and eukaryotic chromatin proteins (39,183). The two main archaeal phyla, Euryarchaea and Crenarchaea, express different sets of chromatin proteins, none of which is conserved throughout the whole archaeal domain. Euryarchaea synthesize true tetrameric histone homologues, which form nucleosomes similar to eukaryotic tetrasomes (44,45). Crenarchaea, on the other hand, do not synthesize histone homologues, although there are some rare exceptions (184). Each crenarchaeal species encodes at least two different small chromatin proteins and several paralogues, which may act in concert to compact the genomic DNA and to regulate its accessibility (40).

In the crenarchaeum *S. solfataricus*, four nucleoid-associated proteins have been identified so far, Alba, Sso10a, Cren7 and Sso7d (the Sul7 family of chromatin proteins). Alba forms dimers in solution (71) and has been shown to bridge DNA at low concentrations, providing a means of organizing and compacting DNA (81,178). Sso10a homologues (65,109) also

exist as dimers in solution and have been shown to form protein-DNA complexes, similar in appearance to Alba-DNA complexes (81). Otherwise, little is known about the role of Sso10a in compacting and organizing genomic DNA.

Cren7 and Sul7 are both small (~ 7 kDa) basic monomeric proteins, which bind to DNA with no apparent sequence specificity. Although they share no similarities at the amino acid level, their tertiary structures and their biochemical properties are very similar (61-63,92-94,105). Both proteins are folded similarly, containing two antiparallel β-sheets. They mainly differ in the presence of an extended loop located in between the two β-sheets in Cren7 (63) and an additional C-terminal α-helix in Sul7 (70). Protein-DNA co-crystal structures showed that they bind to DNA by intercalation of hydrophobic side chains into the minor groove (see inset Figure 2.1), which results in bending of DNA by up to 50 – 60° (95-97,100,101).

The transcriptome of *S. solfataricus* P2 shows high levels of transcription of both Cren7 and Sul7 (110), and, consequently, both proteins are abundant in the cell (~ 1% and up to 5% of the total cellular protein for Cren7 and Sul7 respectively (63,92)). Native Cren7 and Sul7 proteins are post-translationally modified by methylation of several lysine residues. Nevertheless, the function of methylation remains unclear, as it does not change the binding affinity to DNA *in vitro* (62,63), in contrast with the acetylation of Alba (167).

Proteins that bend DNA by sequence non-specific binding into the minor groove play an important role in DNA compaction and chromatin organization throughout all domains of life (14,185-188). In eukaryotes, high-mobility group (HMG) proteins form moderately flexible bends in DNA by intercalating into the minor groove (188), enhancing the overall flexibility of the DNA, and they are likely to play a supplementary role in chromatin organization (187). In bacteria, DNA-bending proteins are crucial players in chromatin organization, as bacteria lack histone homologues.

That Cren7 and Sul7 function similarly to this group of bacterial and eukaryotic DNA-bending proteins is suggested by the co-crystal structures, but has not yet been investigated. A functional difference between the two proteins is expected as they coexist in species from the Sulfolobales order, but previously reported biochemical assays have not yet revealed a significant difference in protein-DNA interactions between Cren7 and Sul7. To obtain a detailed understanding of the architectural properties of the two proteins, we use a set of different techniques including atomic force microscopy (AFM), molecular dynamics (MD) simulations and magnetic tweezers. Our AFM studies yield insight in the overall compaction properties of the two proteins. The MD simulations using X-ray co-crystal structures yield an ensemble of equilibrated 'solution conformations' of the complexes. This gives information on the DNA bending angle induced and the rigidity of this bend. With magnetic tweezers we investigate the mechanical properties of single protein-DNA complexes. Elaborate analysis of measured force-extension curves enables us to quantify the number of proteins bound on single DNA molecules and to get a detailed understanding on the force-dependency of protein binding and the flexibility of the bending angle.

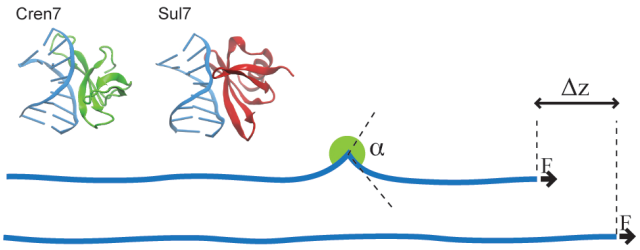


Figure 2.1 Schematic illustrating the reduction of end-to-end distance due to binding of a DNA-bending protein. Binding of a protein that induces a bend of deflection angle α reduces the end-to-end distance with $\Delta z = -1/2 \cdot \lambda \cdot C(\alpha)$. Inset: Average conformation of the Cren7-DNA and Sul7-DNA complexes obtained from MD simulations. Both proteins bend the DNA by intercalating in the minor groove. Images are generated using visual molecular dynamics (VMD) software (189).

RESULTS AND DISCUSSION

To investigate the architectural properties of Cren7 and Sul7, we used a combination of different single-molecule techniques. We used AFM to visualize individual protein-DNA complexes, MD simulations to evaluate the bending angle of the protein-DNA complexes in solution and magnetic tweezers experiments to quantify the physical properties of single protein-DNA complexes and their force dependency.

Effect of Cren7 and Sul7 on the conformation of single DNA molecules

To investigate the effect of Cren7 and Sul7 on the conformation of single DNA molecules, we visualized protein-DNA complexes using AFM. As can be seen in Figure 2.2, addition of either Cren7 or Sul7 results in DNA conformations that are more compact than bare DNA molecules. An increasing amount of protein results in a higher degree of compaction. It is likely that the observed compaction is caused by bends induced by the proteins. However, the small size of the proteins (~ 7 kDa) limits the unambiguous identification and quantitative analysis of individual proteins on the DNA. To gain more insights into the structure and dynamics of the bends induced by the two proteins, we carried out MD simulations of protein-DNA complexes in solution.

Bending angle determination by MD simulations

The co-crystal structures of Cren7-DNA and Sul7-DNA complexes showed that both proteins induce a bending in the DNA of ~ 50 – 60°. However, a complex in a crystal is fixed in a single conformation and may reflect a rare conformation favoured by crystal packing. To evaluate whether the bending angles observed in the co-crystal structures reflect the average conformation in solution, we performed MD simulations in solution using the co-crystal structures as initial coordinates. The MD simulations yielded stable structures except for a small deviation in the Sul7-DNA complex at the end

of the simulation run because of rigid body motions of the C-terminal region (see Figure S2.1 for the root mean square deviations). The overall flexibility of the DNA in all three complexes examined using root mean square deviations was found to be similar. Additionally, the extent of DNA bending was investigated by calculating the roll angle (190). Figure 2.3 shows the probability distributions of the roll angles during the simulation. Fitting the distributions to a Gaussian we find that both protein-DNA complexes exhibit similar bending angles of $\alpha_{Cren7} = 47.5 \pm 5.6°$ and $\alpha_{Sul7} = 45.3 \pm 6.7°$, which slightly deviate from the bending angle found in the co-crystal structures. It is important to note that the width of the roll angle distributions (which reflects the flexibility of the complex) of both protein-DNA complexes is small and comparable with the width of the roll angle distribution of bare DNA ($\alpha_{bareDNA} = 7.3 \pm 5.7°$).

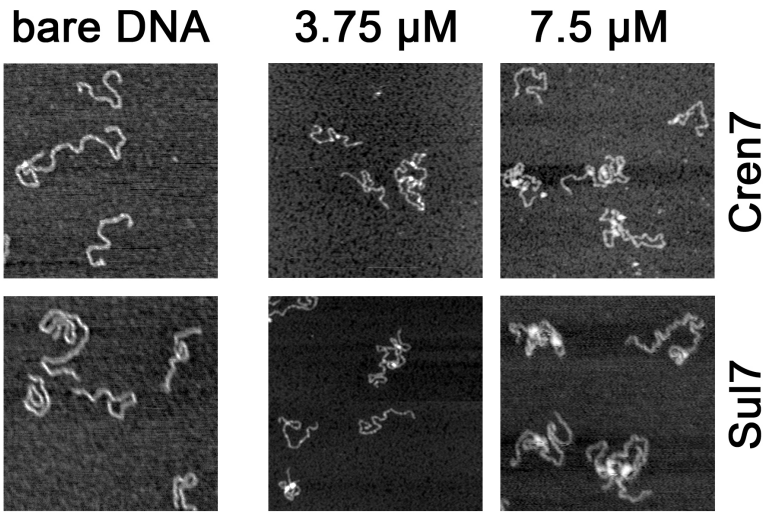


Figure 2.2 AFM images of bare DNA, Cren7-DNA complexes and Sul7-DNA complexes. The protein-DNA complexes are more compact than the bare DNA molecules. Increasing the protein concentration results in more compaction. The DNA is an equimolar mixture of both 1841 and 845 bp linear fragments. Images are 500 x 500 nm in size.

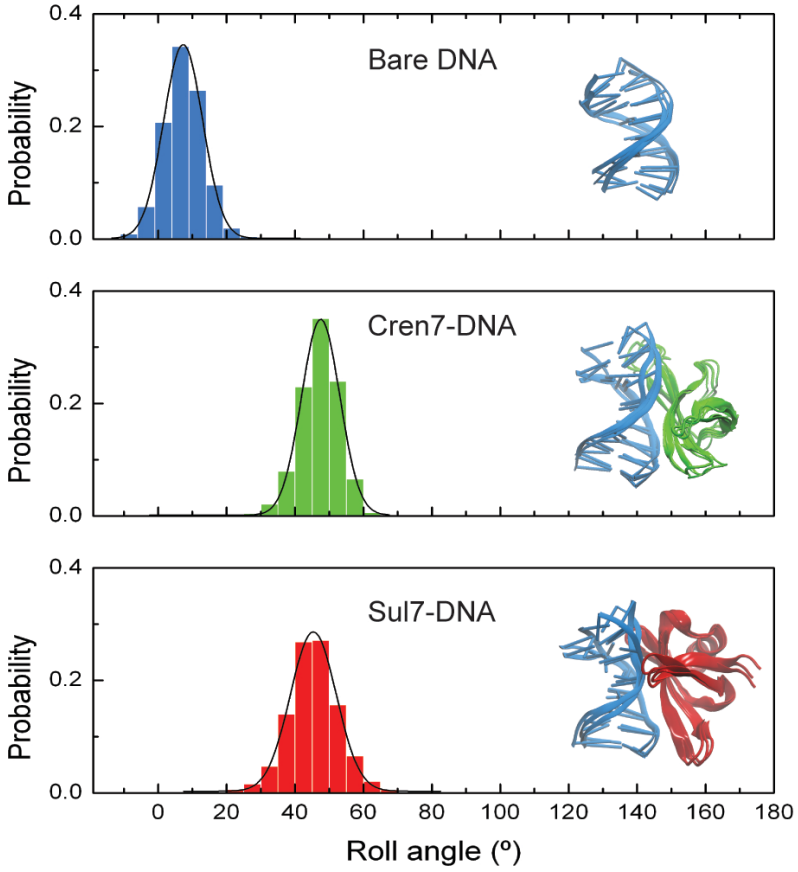


Figure 2.3 Probability distribution of the roll angle of the kinked base pair step in the bare DNA (*blue*), the DNA in the Cren7-DNA complex (*green*) and the DNA in the Sul7-DNA complex (*red*). Distributions are obtained from the final 90 ns of the MD simulations and fitted to a Gaussian ($\alpha_{bareDNA}$ = 7.3 ± 5.7°, α_{Cren7} = 47.5 ± 5.6° and α_{Sul7} = 45.3 ± 6.7°). Insets: The average conformations of the bare DNA the Cren7-DNA and Sul7-DNA complexes are shown and aligned with two extreme conformations to illustrate the flexibility of the complexes. Images are generated using visual molecular dynamics (VMD) software (189).

Quantification of DNA compaction by Cren7 and Sul7

To quantify the molecular mechanisms underlying the observed DNA compaction, we carried out single-molecule micromanipulation experiments. In these experiments, we measured the force-extension relation of single double stranded DNA (dsDNA) molecules over a range of concentrations for both Cren7 and Sul7 (0, 40, 80, 200, 400 and 800 nM) (Figure 2.4 A and B). In both cases, the addition of protein results in a reduction of z (compaction), which decreases progressively upon raising the protein concentration. No significant difference is observed between Cren7 and Sul7 as they compact the DNA to a similar extent, which confirms the compaction observed in our AFM studies. The reduction of the end-to-end distance (z) saturates at a protein concentration of ~ 400 nM (Figure 2.4 C and D). At the highest concentration (800 nM), the end-to-end distance increases slightly, which could point to interactions between adjacent proteins, causing a mild extension of the compacted DNA molecules. However, incubation with higher protein concentrations (up to 1600 nM) did not induce a further increase in end-to-end distance (*data not shown*).

To quantify the effects of the protein on the mechanical properties of DNA, the WLC model [Equation (2.2)] was fitted to the force-extension curves. While the contour length stays constant (L_0 = 1.0 ± 0.05 μm), the apparent persistence length (191,192) decreases drastically (Figure 2.5). At high concentrations the apparent persistence length reaches a minimum of L_p = 8.8 ± 1.3 nm for Cren7 and L_p = 9.0 ± 1.6 nm for Sul7, which is more than four times smaller than the measured persistence length of bare DNA, L_p = 40.5 ± 11.2 nm. Thus, binding of these proteins results in compaction of DNA by reducing the apparent persistence length.

It is interesting to note that higher protein concentrations do not result in stiffening of the DNA. Previous studies of similar DNA-bending proteins (e.g. the bacterial proteins HU and IHF and eukaryotic HMG proteins) have shown that these proteins compact DNA by forming bends at low protein

concentrations (193-196). However, HU and HMG proteins enter a second binding regime at high concentrations, forming stiff filaments due to close adjacent binding. It has been suggested that IHF behaves similarly (37) when binding non-specifically along DNA (193,197), but this has not yet been experimentally confirmed. The bimodal binding of HU, causing a transition from a softening to a stiffening mode, has been shown to be salt dependent and was only observed at NaCl concentrations < 150 mM (198).

As such salt-dependent effects might exist also for Cren7 and Sul7, we measured force-extension curves under different ionic conditions. Previous studies have already shown that the binding affinity of Sul7 decreases with increasing ionic strength (199). Measurements in a buffer containing a lower salt concentration (25 mM NaCl) showed that the decrease in apparent persistence length occurs at much lower protein concentration (Figure 2.5 B), which suggests an increased binding affinity, because of enhanced electrostatic protein-DNA interactions at low ionic strength, as expected. Again at the highest protein concentration the persistence length did not increase and a stiffening mode was not observed. These findings suggest that Cren7 and Sul7 do not stiffen DNA in a second binding mode, even at protein concentrations where high DNA coverage is expected. It has been shown for bacterial H-NS-like proteins that these proteins exhibit two binding modes (116,200,201), which can be switched by the addition of $MgCl_2$ (115,202). To test whether divalent ions have an effect on the mode of binding of Cren7 and Sul7 we performed measurements in the presence of $MgCl_2$. These experiments did not reveal an effect on the binding behaviour of Cren7 and Sul7, other than a decrease in binding affinity (see Figure S2.2). This reinforces that the binding is dominated by electrostatic interactions, which are reduced by increasing the ionic strength.

Quantification of the number of proteins bound

The number of proteins bound to the DNA within the range of measured force and protein concentrations is determined by applying Equation (2.5) to

the measured force-extension curves (Figure 2.4 A and B) using the bending angle α determined from the MD simulations (α_{Cren7} = 47.5° and α_{Sul7} = 45.3°, see Figure 2.3). As expected, the number of proteins bound on the DNA (N_b) increases with increasing protein concentration (see Figure 2.6).

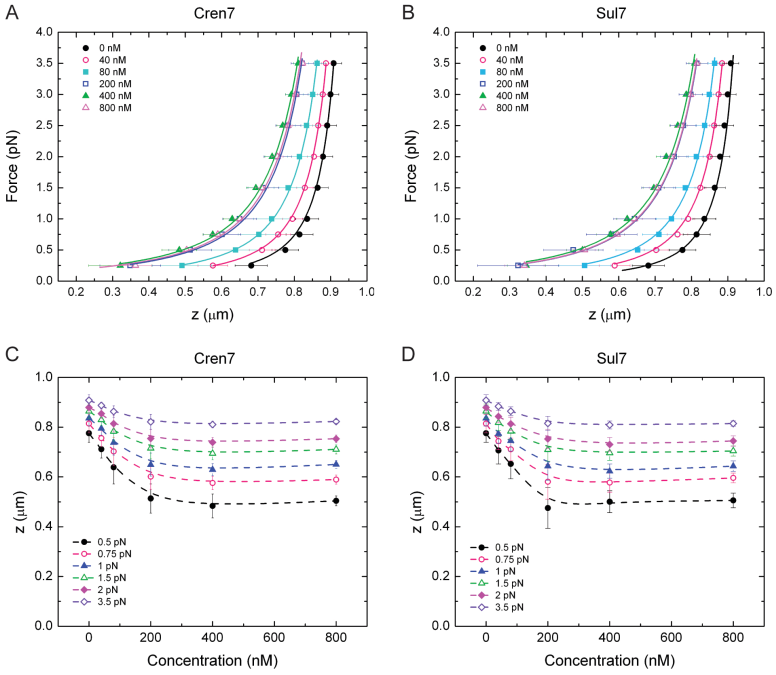


Figure 2.4 Force-extension curves of protein-DNA complexes. Each data point represents the average of a number of DNA molecules measured in buffer containing 100 mM NaCl (N = 26 for bare DNA molecules and N = 6 – 14 for DNA-protein complexes). Data is fitted with the WLC model [Equation (2.2)]. Error bars represent the standard deviation. **A and B.** Averaged force-extension curves of DNA at different protein concentrations of Cren7 and Sul7. **C and D.** End-to-end distance as a function of bulk protein concentration represented at constant forces. Progressive addition of protein results in a decrease of the end-to-end distance z, and thus compaction of DNA molecules.

Considering the condition for which our model is valid $N_b \ll L_0 / \lambda$ (see Materials and Methods section), we note that this condition holds for the 40 nM and 80 nM measurements, where proteins are still bound distant enough from each other. For the lowest concentrations the number of bound proteins is found to be constant throughout the measured force range (Figure 2.6), from which we conclude that forces up to 3.5 pN do not affect the number of proteins on the DNA. At higher protein concentrations, the high occupancy may introduce a systematic error in the calculated N_b, but as in the case of lower concentrations we still find a force independent value for N_b.

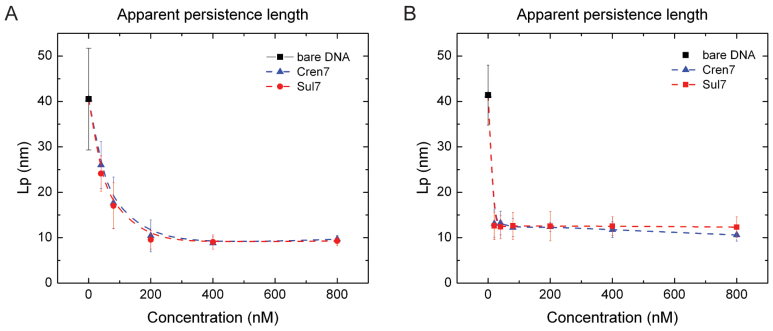


Figure 2.5 Apparent persistence length (L_p) decreases as a function of protein concentration. The persistence length is determined by fitting individual force-extension curves to the WLC model [Equation (2.2)] and averaged for each concentration point. Error bars represent the standard deviation. **A.** Average apparent persistence length from F-z curves measured in buffer containing 100 mM NaCl. **B.** Average apparent persistence length from F-z curves measured in buffer containing 25 mM NaCl.

This is a surprising observation, as one would expect a significant decrease in binding affinity with applied force. The energy it costs to form a bend is force-dependent and for the highest force (3.5 pN) this would be $E_{kink} = \sim k_BT$ (see Supplementary Information). As the dissociation constant K_d is related to E_{kink} and the binding free energy ΔG_0 as follows:

$$K_d(F) \sim \exp\left(\frac{\Delta G_0 + E_{kink}(F)}{k_B T} \right) \quad (2.1)$$

one would expect the dissociation constant to increase with a factor 2.7 for both Cren7 and Sul7 when the force is increased from 0 to 3.5 pN, resulting in a lower protein occupancy at higher forces. Apparently, these proteins bind stably enough to withstand such forces on our experimental timescale (~ 30 s). Note that measurements were reversible for all measured protein concentrations, which could be explained by off-rates much larger than the experimental timeframe (see Figure S2.3). Even holding the protein-DNA complexes at 3.5 pN for 10 minutes did not result in an increase in z or a different retraction curve showing that the proteins stay stably bound even on longer timescales (see Figure S2.4). Unexpectedly, low off-rates of proteins bound to DNA have also been observed in previous studies on HU and HMG protein-DNA complexes in protein free solution (198,203). Apparently, there is a large energy barrier for the proteins to dissociate which is not reached by applying a force of 3.5 pN. Such strong binding could be crucial for maintaining the integrity of the DNA in extreme environmental conditions of the organism's natural habitat. Cellular processes that apply local tension on the DNA such as transcription, replication and repair act on timescales that are much shorter compared with our experimental timescale. This suggests that these proteins are relatively stably bound to the DNA *in vivo*. However, cellular machineries involved in replication and transcription such as helicases and polymerases may directly drive the proteins to dissociate (204) allowing access to the DNA track.

Determining the number of proteins bound to DNA and how this is affected by force on the DNA has been a topic of interest in various recent studies (205-208). It is important in the light of understanding the contribution of the individual bound proteins to overall genome organization and how forces induced by cellular processes affect this.

Theoretical models have described the number of proteins bound to single DNA molecules in relation to applied force. These models showed that the force response of the bound proteins is highly dependent on the flexibility of the protein-DNA complex if the binding is relatively strong (207). With the use of thermodynamic Maxwell relations (208), the change in the number of bound proteins can be determined, given the condition that the protein-DNA interactions are in thermodynamic equilibrium. Analysis with the thermodynamic Maxwell relations of single-DNA force-extension data with DNA-bending proteins from *E. coli*, HU and Fis, suggested that these proteins are driven off the DNA by an applied force, even at relatively small forces < 1 pN (206). It should be noted, however, that this approach does not take into account that the degree of protein induced DNA bending could be altered by DNA tension. Applying force to the protein-DNA complexes could reduce the degree of bending, which would result in a net mechanical response, similar to that of protein unbinding (196). Thus, the proposed unbinding of proteins in those studies might actually be (in part) caused by a decrease of bending angle rather than protein unbinding.

Do Cren7 and Sul7 induce flexible or static bends?

Previous studies of bacterial chromatin protein HU have shown that bends of HU-DNA complexes are highly flexible (209) [varying from 0° to 180°(210)]. Although a similar flexible hinge model was initially proposed for eukaryotic HMG domains (194), further experiments showed a more narrow angle distribution of HMGB-DNA complexes (188). Non-specific binding of IHF may induce flexible bends analogous to HU, whereas the site-specific binding of IHF is stabilized by interaction between the DNA flanking the kink sites and the body of the protein, making this bend static (209). Based on the analogy of Cren7 and Sul7 with those non-specific DNA-bending proteins, one would anticipate that these proteins induce bends with a certain degree of flexibility. Applying a force on such flexible bends would decrease the bending angle. Equation (2.5) yields an increase of the extension of the DNA for both a reduction in the bending angle and a

reduction of N_b, which are indistinguishable from each other. Surprisingly, the number of bound proteins N_b is unaffected by the applied force, which implies that the bending angles are independent of forces ≤ 3.5 pN. The MD simulations of the protein-DNA complexes in solution at zero force show that the distributions of roll angles are narrow (Figure 2.3) (α_{Cren7} = 47.5 ± 5.6° and α_{Sul7} = 45.3 ± 6.7°) and comparable in width with the roll angle distribution of bare DNA ($\alpha_{bareDNA}$ = 7.3 ± 5.7°). Both observations show that Cren7 and Sul7 bind with a rigid bend, in contrast with flexible hinges induced by analogous DNA-bending proteins. If a flexible angle is essential in permitting tight packing of proteins this could explain the absence of a second binding mode in which the DNA is stiffened.

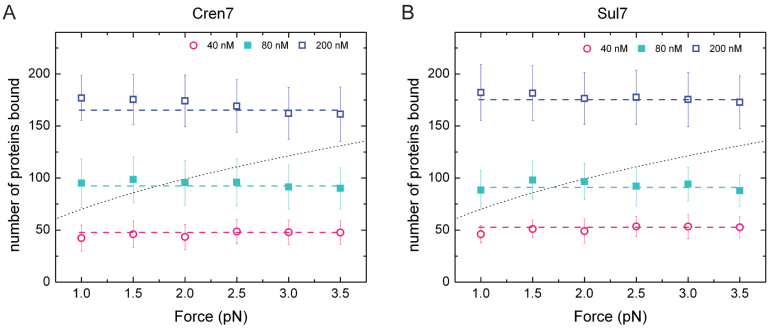


Figure 2.6 Number of bound proteins (N_b) on a single DNA molecule (2947 bp) as a function of force, calculated from the force-extension curves according to Equation (2.5). The number of proteins increases with increasing protein concentration, but is unaffected by forces up to 3.5 pN. Lines represent N_b calculated from the apparent persistence length. Black dashed line represents the condition $N_b \ll L_0 / \lambda$.

CONCLUSIONS

The protein-DNA interactions of Cren7 and Sul7 are very similar. Like in previous biochemical assays we found that they bind with comparable affinity. Here we have demonstrated that they induce a similar degree of

compaction of DNA, bind independent of forces up to 3.5 pN and bend DNA with a non-flexible bend. Whether the bending and binding of Cren7 and Sul7 also resists higher forces remains to be seen and could be assessed with other single-molecule techniques that permit DNA-stretching at higher forces. Visualization of fluorescent proteins in combination with such techniques (211) could confirm our findings in the low force regime. Although no difference has yet been observed, given the fact that Cren7 and Sul7 coexist in the same organisms one would still expect them to have a functional difference *in vivo*. As mentioned earlier, Cren7 and Sul7 can both be methylated at several lysine residues. Such post-translational modifications of nucleoid-associated proteins can play an important role in regulation of interactions between individual proteins on the DNA, as well as the interaction with other proteins, such as transcription factors or other chromatin proteins. How this affects the role of Cren7 and Sul7 in chromatin organization, and possibly gene regulation, remains to be investigated.

MATERIALS AND METHODS

Protein purification

Cren7

The Cren7 protein was purified from *Escherichia col*i strain BL21 (DE3) containing plasmid pET30a, including the gene encoding Cren7 (gene SSO6901) from *S. solfataricus* (63). Cells were grown in LB medium up to $OD_{600} \approx 0.4$ and expression was induced using 0.5 mM IPTG at 37 °C. Two hours after induction, cells were harvested by centrifugation, washed with physiological saline (0.9% w/v NaCl), and resuspended in buffer A [50 mM Tris-HCl (pH 8.0), 10 mM ethylenediaminetetraacetic acid (EDTA), 10% Glycerol, 10 mM β-mercaptoethanol]. Cells were lysed by sonication, and the cell lysate was centrifuged at 37.000 r.p.m. for 30 min at 4°C. The supernatant was applied to an SP column (GE Healthcare) and was eluted with buffer A containing 200 mM NaCl. The eluted protein was heated at

65 °C for 30 min and centrifuged at 37.000 r.p.m. for 15 min. The supernatant was diluted to 100 mM NaCl in buffer A, loaded on a heparin column (GE Healthcare) and eluted with a linear gradient of 0.1 – 1 M NaCl in buffer A. The protein was eluted at ~ 270 mM NaCl.

Sul7

A synthetic gene encoding the *S. solfataricus* Sul7 (Ss07d) protein (gene SSO10610) was constructed in expression vector pJexpress411 (supplied by DNA2.0). The DNA sequence used is available from the authors on request. The plasmid was transformed into *E. coli* strain BL21 (DE3) Rosetta cells, grown in LB medium, and expression was induced using 0.2 mM IPTG at 37 °C for 3 h. Following cell lysis by sonication, the Sul7 protein was loaded onto a heparin column (GE Healthcare) and eluted with a linear gradient of 0 – 1 M NaCl in buffer B [20 mM Tris-HCl (pH 7.4), 1 mM DTT, 1 mM EDTA]. Fractions containing the Sul7 protein were pooled, concentrated and applied to a Superdex 200 gel filtration column (GE Healthcare) and were eluted by isocratic flow with buffer B supplemented with 200 mM NaCl.

Cren7 and Sul7 proteins were dialysed at 4°C against a storage buffer [20 mM Hepes (pH 7.5), 100mM NaCl, 10% Glycerol, 10 mM β-mercaptoethanol], and stored at -80°C until required. Protein concentrations were determined by ultraviolet absorbance at 280 nm, using a molar extinction coefficient ε_{Cren7} = 8250 $M^{-1}cm^{-1}$ (63) and ε_{Sul7} = 8300 $M^{-1}cm^{-1}$ (70).

DNA constructs

AFM

pUC18 was propagated in *E. coli* strain XL10 and was purified (Qiagen plasmid midi kit). Digestion with EcoRI and XmnI resulted in a mixture of 845-bp and 1841-bp linear fragments.

Magnetic tweezers

pBluescript II KS+ (Stratagene) was digested with SalI and HindIII enzymes to generate a 2947-bp linear fragment. The cohesive end at the SalI digestion site was labelled with dUTP-digoxigenin (DIG), using the Klenow fragment (Fermentas). Subsequently the fragments were purified, labelled with dUTP-biotin at the extremity resulting from HindIII digestion using dATP, dCTP and dGTP and the Klenow fragment and were purified again.

Molecular dynamics simulations

MD simulations of 100 ns were performed in an explicit solvent environment as previously described (103) using initial coordinates from the co-crystal structures of Cren7-DNA and Sul7-DNA (PDB codes 3LWH and 1BNZ respectively) and an ideal B-DNA duplex with identical sequence was used for the bare DNA simulation. A detailed description can be found in the supplementary materials and methods (see Supplementary Information).

Atomic Force Microscopy experiments

Freshly cleaved mica was incubated with 10 µl of 0.05% poly-L-lysine, rinsed with MilliQ water and dried with nitrogen gas. Protein-DNA complexes were formed by incubating 50 ng of DNA with varying concentrations of protein in 10 µl buffer containing 10 mM HEPES (pH 7.5) and 100 mM NaCl for 10 min at room temperature (~ 23 °C). After incubation, this mixture was diluted 5-fold in water and was directly deposited onto poly-L-lysine coated mica, rinsed with MilliQ water and dried with nitrogen gas. Images were collected on a NanoScopeIII AFM (Digital Instruments, Santa Barbara, CA) using tapping mode in air (micro cantilevers, Olympus MCL-AC240TS-W2, resonance frequency 70 kHz, spring constant 2 N/m) at a frequency of 2 Hz and were flattened using Nanoscope software (Veeco Instruments).

Magnetic tweezers experiments

Magnetic tweezers experiments were performed on a custom-built instrument described earlier (212). Images were acquired with a CCD camera (Pulnix TM-6710CL) at 60 Hz and real time image processing was done using custom-developed LabView software (National Instruments). Magnets were positioned in the optical axis of the microscope and were controlled by a stepper motor-based translational stage (M-126, Physik Instrumente). Force was calibrated by measuring fluctuations of a tethered DNA-bead construct and was calculated according to the equipartition theorem (213).

The bottom slide of the flow cell was pre-coated with 1% (w/v) polystyrene in toluene. Subsequently, the flow cell was filled with 1 µg/ml anti-DIG antibodies (Roche) and was incubated for 2 h at 4 °C. Next, the flow cell was flushed with 2% (w/v) bovine serum albumine (BSA) and 1% (v/v) Tween-20 solution and was incubated overnight at 4 °C. The flow cell was flushed with buffer I [10 mM HEPES (pH 7.6), 100 mM KAc, 0.2% (w/v) BSA, 0.1% (v/v) Tween-20], filled with 20 ng/ml DNA (functionalized with biotin and DIG) in buffer I and incubated for 10 min. One microlitre of streptavidin coated superparamagnetic beads with a diameter of 1 µm (DYNAL Myone) was washed three times in 1x TE [10 mM Tris-HCl (pH 8.0), 1 mM EDTA], resuspended in 500 µl buffer I, flushed into the flow cell and incubated for 10 minutes. Before measurements, the flow cell was washed thoroughly (~ 10 times the volume of a flowcell) with buffer II [10 mM HEPES (pH 7.5), 0.2% (w/v) BSA, 0.1% (v/v) Tween-20] with addition of either 25 mM or 100 mM NaCl (with or without 10 mM $MgCl_2$ added). Cren7 and Sul7 stock solutions were diluted in buffer II to the desired protein concentration (40 – 800 nM), and incubated for 10 minutes before starting measurements. Force-extension curves of single double-stranded DNA (dsDNA) molecules and protein-DNA complexes were measured by increasing the force to 3.5 pN during 30 s and subsequently reducing the

force to zero in 10 s. All measurements were performed at room temperature (~ 23 °C).

Data analysis

Force-extension curves of single dsDNA molecules and protein-DNA complexes were quantified using the worm-like chain model (WLC). This model describes the mechanical properties of DNA and is defined by its intrinsic properties: the persistence length L_p and the contour length L_0 (214):

$$F(z) = \left(\frac{k_B T}{L_p} \right) \left(\frac{1}{4(1 - z / L_0)^2} - \frac{1}{4} + \frac{z}{L_0} \right) \tag{2.2}$$

Here F denotes the applied force, z the end-to-end distance of the DNA molecule, k_B the Boltzmann constant and T the absolute temperature. The effect of protein-DNA interactions on the DNA can be quantified by the fitting parameters L_p and L_0 from F-z curves of protein-DNA complexes.

The number of proteins that is bound to the DNA can be directly related to the change in end-to-end distance of the DNA molecule. Intercalation of a protein in the minor groove of the DNA that induces a bend in the DNA affects the end-to-end distance with a change in length Δz (Figure 2.1). This change in end-to-end distance is related to the force-dependent deflection length $\lambda(F) = \sqrt{k_B T \cdot L_p / F}$ as follows (191,192):

$$\Delta z = -\frac{1}{2} \lambda \cdot C(\alpha) \tag{2.3}$$

where L_p is the persistence length of bare DNA and C(α) is a geometric coefficient depending on the bending angle α:

$$C(\alpha) = 8\left[1 - \cos\left(\frac{\alpha}{4}\right)\right] \tag{2.4}$$

Implementing the reduction of length Δz [Equation (2.3)] into the WLC [Equation (2.2)] and noting that $1/(4(1-z/L_0)^2) \gg z/L_0 - 1/4$ in our force regime, leads to a relation where the end-to-end distance (z) is directly related to the average number of proteins bound N_b (191,192):

$$\frac{z(F,N_b)}{L_0} \approx 1 - \frac{1}{2}\frac{\lambda(F)}{L_p}\left[1 + N_b \cdot C(\alpha)\frac{L_p}{L_0}\right] \tag{2.5}$$

This leads to an expression for the apparent persistence length, normalized by the number of proteins bound:

$$L_p^{app}(N_b) = \frac{L_p}{\left(1 + N_b \cdot \frac{L_p}{L_0} C(\alpha)\right)^2} \tag{2.6}$$

The number of proteins bound along single DNA molecules can now be determined by a model describing the mechanical response of DNA with multiple independent bends [Equation (2.6)]. It should be noted that this model only holds when the DNA is sufficiently stretched, such that the average distance between bound proteins is larger than the force-dependent deflection length, that is, $N_b \ll L_0/\lambda$. In this force regime, bound proteins are distant enough to prevent the induced bends from interacting with each other. The aforementioned condition follows from an exact calculation that we performed for arbitrary linker lengths for a special case of protein arrangement (zigzag geometry). As for shorter linker lengths the end-to-end distance depends on the rotational setting of adjacent proteins, our aforementioned theory can lead to under- or overestimation of the number of bound proteins outside the allowed regime.

SUPPLEMENTARY INFORMATION

Geometry and energy of a kink formation

Consider a DNA chain under tension F with a protein adsorbed that causes a deflection α (see Figure 2.1). The energy of this chain is given by the sum of the bending and the potential energy:

$$E_{kink} = \int_0^L \left(\frac{A}{2}\dot{\theta} - F\cos\theta \right) ds \tag{2.7}$$

Here $A = L_p k_B T$ is the bending modulus of the DNA chain, L_0 is its contour length, s gives the position along its contour, $0 \le s \le L_0$, and $\theta(s)$ is the angle between the tangent at position s and the force direction. The shape of the two sections of DNA chain to the left and to the right of the kink follow from the Euler-Lagrange equation $\ddot{\theta} = \lambda^{-2}\sin\theta$ with $\lambda = \sqrt{A/F}$.

In the following we extend the chain to $s = \mp\infty$ which introduces a negligible small error in the results as long as $\lambda << L_0$. The solution with the proper boundary conditions at the ends, namely $\theta(-\infty) = 0$ and $\theta(+\infty) = 2\pi$, is given by

$$\cos\theta(s) = 1 - \frac{2}{\cosh^2(s/\lambda)} \tag{2.8}$$

This solution features a single loop around $s = 0$. We cut out a segment of the solution from $-s_0 < s < s_0$ to obtain the geometry shown in Figure 2.1. The boundary condition at the protein-induced kink for e.g. the right arm is given by $\theta(s_0) = -\alpha/2$. The length loss due to the kink follows from

$$\Delta z = \int_{s_0}^{\infty} \left(\cos\theta(s) - 1\right) ds = 4\lambda\left(\tanh(s_0/\lambda) - 1\right) \tag{2.9}$$

Using $\theta(s_0) = \alpha / 2$ and trigonometric identities this can be rewritten as

$$\Delta z = -4\lambda\left(1 - \cos(\alpha / 4)\right) \tag{2.10}$$

From this follows directly Equation (2.5) and Equation (2.6). This result has already been reported earlier .

In addition, we calculate the energy associated with the formation of the kink, which is given by:

$$E_{\text{kink}} = A\int_{s_0}^{\infty} \dot{\theta}^2(s)\,ds - F\Delta z = \frac{4A}{\lambda^2}\int_{s_0}^{\infty} \frac{ds}{\cosh^2(s/\lambda)} - F\Delta z = 8\lambda F\left(1 - \cos(\alpha/4)\right) \tag{2.11}$$

For the highest force, $F_{\max} = 3.5$ pN , and the bending angle obtained from molecular dynamics (MD) simulations (α = ~ 46°, see Figure 2.3) this leads to the energy for kink formation $E_{kink} = 0.99 \cdot k_B T$.

Molecular dynamics simulations

The initial coordinates for protein-DNA complexes of Sul7 and Cren7 were taken from the co-crystal structures (PDB codes: 1BNZ for Sul7 (97) and 3LWH for Cren7 (101)). The SYBYL7.2 (Tripos Inc) generated ideal B-DNA duplex coordinates with identical sequence and length of the DNA form the co-crystal structures is used as the initial coordinates for the bare DNA simulation. All MD simulations were performed using CHARMM biomolecular simulation program (215) employing CHARMM22 all atom protein force field with CMAP corrections for proteins, and CHARMM27 all atom nucleic acid force field for nucleic acids (216-219). Missing hydrogens were added using HBUILD utility of CHARMM and allowed for 500 step steepest decent (SD) minimization on having harmonic constraints on heavy atoms. Later these systems were immersed in a cubic box with an edge length of 58 Å, treating the water using TIP3P model (220) and the water molecules within 2.4 Å distance of solute were deleted. All the crystal

water molecules were retained during the simulations. Sodium ions were placed randomly to make the systems electrically neutral. The bonds involving hydrogens were restrained by using SHAKE algorithm (221). Next, the systems were minimized using SD and Adopted Basis Newton Raphson (ABNR) methods up to 500 steps each with harmonic constraints on heavy atoms of protein and DNA. The systems were equilibrated for 100 ps with the same conditions. Particle mesh ewald summation method (222,223) was used to treat long-range interactions. CRYSTAL module, implemented in CHARMM, is used for periodic boundary conditions. All the simulations were performed in isobaric-isothermal (NPT) ensemble using the Nose-Hoover thermostat (224) for temperature control at 300 K and the Langevin piston algorithm (225) for constant pressure. The production simulations have been extended to 0.1 μs with an integration time step of 2 fs using the Leapfrog integrator and the coordinates were saved for every 2 ps for further analysis.

The final 90 ns of the simulation trajectories were used for analysis using CHARMM and Curves+ (190) programs. All the translational and rotational motions were removed and only the systems with protein and DNA were considered by deleting the water and ions for this analysis. To assess the extent of bending in the DNA, roll parameters at the kink site region were calculated using Curves + (190).

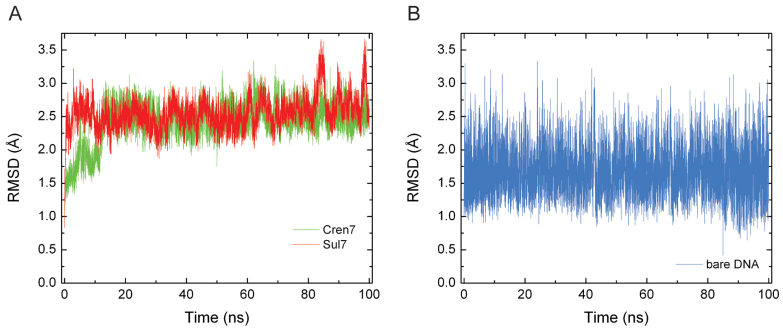


Figure S2.1 Root mean square deviations of MD simulations of **A.** Sul7-DNA and Cren7-DNA complexes and **B.** bare DNA

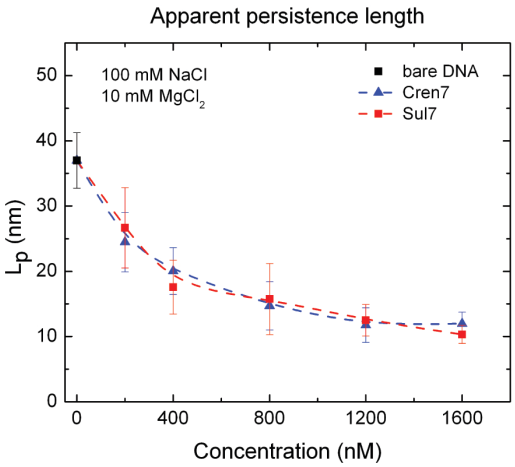


Figure S2.2 Apparent persistence length as a function of protein concentration in buffer containing mono- and divalent ions (100 mM NaCl and 10 mM $MgCl_2$). The persistence length decreases as a function of protein concentration at a higher protein concentration compared to experiments performed in buffer containing only 100 mM NaCl (see Figure 2.5 A).

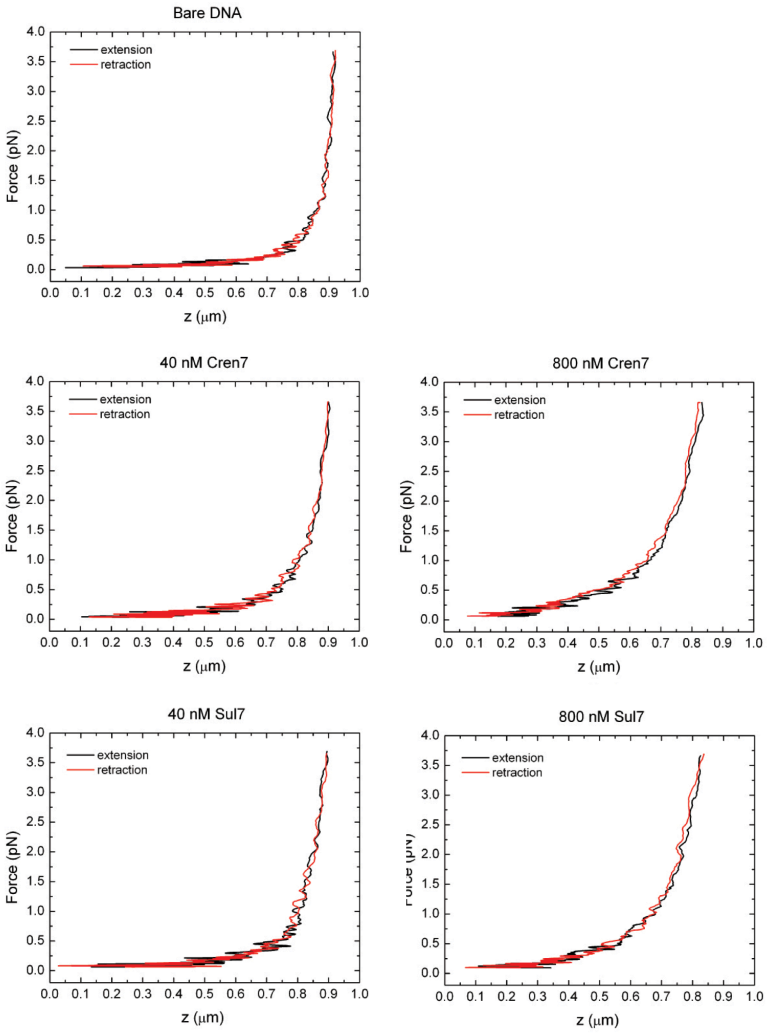


Figure S2.3 Force extension curves of bare DNA and protein-DNA complexes. The overlap of the extension and retraction curves shows that force extension curves are reversible, independent of the protein concentration.

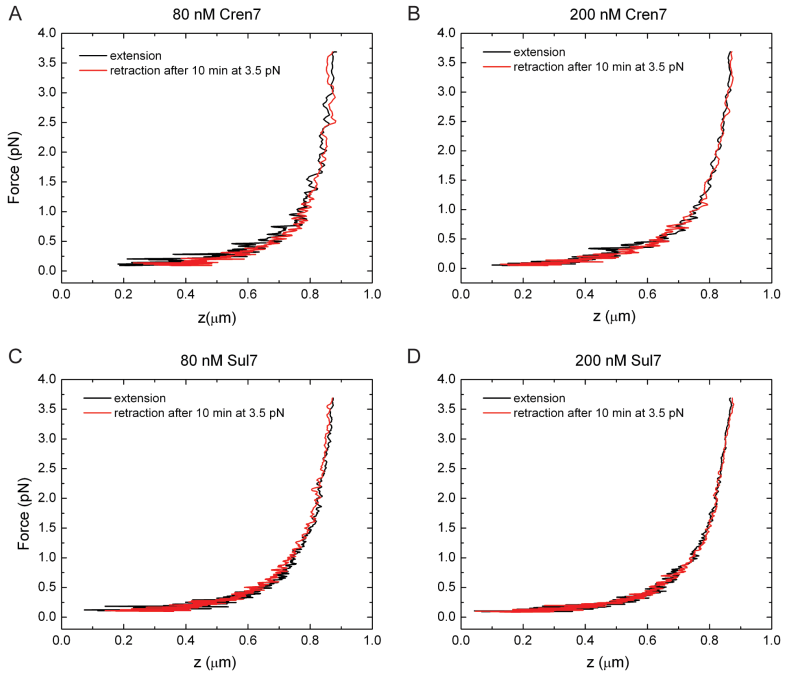


Figure S2.4 Force extension curves of protein-DNA complexes of long timescale experiments. The force extension curves are reversible even when holding the protein-DNA complexes at 3.5 pN for 10 minutes. This confirms that the proteins are bound stable to the DNA and do not dissociate within this timescale.

Chapter 3

Tethered Particle Motion reveals how temperature affects DNA flexibility and protein-DNA interactions

This chapter has been submitted for publication

Rosalie P.C. Driessen, Gerrit Sitters, Niels Laurens, Geri F. Moolenaar, Gijs J. L. Wuite, Nora Goosen, Remus Th. Dame

ABSTRACT

The helical structure of double stranded DNA is destabilized by increasing temperature. Above a critical temperature (the melting temperature) the two strands in duplex DNA become fully separated. Below this temperature the structural effects are localized. Using Tethered Particle Motion in a temperature-controlled sample chamber, we systematically investigate the effect of increasing temperature on DNA structure and the interplay between this effect and protein binding. Our measurements reveal 1) that increasing temperature enhances DNA flexibility, effectively leading to more compact folding of the double stranded DNA chain and 2) that this effect on DNA structure differentially affects different types of DNA-bending chromatin proteins from mesophilic and thermophilic unicellular organisms. Thus, our findings are key to understanding genome organization in organisms thriving at moderate as well as extreme temperatures. Moreover, our results underscore the importance of carefully controlling and measuring temperature in single-molecule DNA (micromanipulation) experiments.

INTRODUCTION

Double stranded DNA (dsDNA) is a semi-flexible polymer (226). On short length scales (in the order of its persistence length: ~ 150 basepairs or ~ 50 nm) stiffness dominates its conformation and bending is energetically unfavourable. However, on the larger scale dsDNA acts as a flexible polymer and forms a random coil. Although the exact cause of DNA rigidity remains unclear, it has been proposed that both basepair stacking and the electrostatic repulsions of the negatively charged phosphate backbone contribute to the local stiffness of the DNA (226,227). In order to facilitate genome-based processes such as transcription, DNA repair and replication, the relatively stiff DNA is sharply bent on local scales (228). Despite its global flexibility genomic DNA also needs to be bent substantially to fit into the volume of a cell or cell nucleus.

Across all domains of life, cells employ small architectural proteins that bend the DNA in order to compact the genome (14). Eukaryotes and many archaeal species express histone proteins that sharply bend and wrap DNA into compacted nucleosomes (34,45,52). Other nucleoid-associated proteins have been shown to induce compaction by local DNA bending. Examples include the eukaryotic high mobility group (HMG) proteins (194), the bacterial DNA bending proteins HU, IHF and Fis (36,37), and the crenarchaeal proteins Sul7 and Cren7 (229). Besides the role of DNA-bending proteins in genome compaction and organization, the bending induced by these proteins is also crucial in regulatory processes (230). Regulatory complexes *in vivo* often involve the formation of small DNA loops (231) which require sharp bending of DNA. A mechanism to enhance the formation and stability of these loops is the binding of a DNA bending protein within the loop. For example, HU was shown to stabilize various kinds of repression loops (among which loops mediated by the bacterial Lac and Gal repressor) by enhancing DNA flexibility (170,171). In the Gal repressosome two GalR dimers form a small loop facilitated by the binding of an HU dimer at the apex of the loop (172). Other studies have shown that

sharp bending of DNA can occur spontaneously, yielding a higher intrinsic bendability of DNA on short length scales compared to long length scales (232,233). Such strong bending of short dsDNA is caused either by the formation of small melting bubbles or transient kinks (234-238). Based on the observation that sharp bending of DNA occurs spontaneously it was suggested that DNA flexibility itself acts as a factor affecting the conformation and stability of looped regulatory complexes *in vivo* (239).

It is likely that the flexibility and bendability of DNA affect spatial genome folding and functioning. Indeed it has been shown in biochemical ensemble measurements that temperature directly affects DNA structure by changing its persistence length (240). Besides such a direct effect on DNA structure, temperature might also influence the interactions between DNA and architectural proteins, and hence chromatin structure. Growth temperature and fluctuations in this temperature are thus expected to have a strong impact on *in vivo* DNA organization and gene regulation. Indeed, the transcription of many genes changes following moderate temperature shifts in both mesophilic *E. coli* (241) as well as thermophilic *S. solfataricus* (242). In addition to being associated with a general stress response (243), such effects on transcription could in part be mediated by global changes in chromatin structure due to the change in temperature. Single-molecule experiments concerning DNA flexibility and protein-DNA interactions are generally conducted at room temperature. However, the large majority of organisms lives at temperatures different from room temperature: psychrophiles thrive at temperatures around 0 °C, mesophiles at 20 – 45 °C, while thermophilic organisms live at temperatures up to ~ 100 °C. In this study, we investigate the effect of temperature on DNA structure and flexibility at temperatures ranging from 23 – 52 °C using a temperature controlled tethered particle motion (TPM) set-up. Moreover, we investigate the effect of temperature on the binding and bending behaviour of the architectural proteins Cren7 and Sul7 from the thermophilic organism *S. solfataricus* (living at temperatures between 40 and 90 °C) and HU from mesophilic *E. coli* (living at temperatures between 10 and 40 °C).

RESULTS

Increasing temperature enhances DNA flexibility

The motivation of our studies is to study protein-DNA interactions at temperatures relevant for the organisms encoding these proteins. Such interactions might be affected by changes in structure and DNA flexibility occurring by changing temperature. To investigate the direct effect of temperature on DNA structure and flexibility we performed Tethered Particle Motion (TPM) experiments over a range of temperatures. Since the physical properties of dsDNA are dependent on the basepair composition (244), we measured three different DNA constructs of identical length (685 bps) but different average GC content (32%, 53% and 70%). Using bulk DNA melting measurements we first determined the melting temperatures (T_m) to be $T_{m,32\%}$ = 76.8 ± 0.5 °C, $T_{m,53\%}$ = 84.9 ± 0.5 °C and $T_{m,70\%}$ = 91.2 ± 0.5 °C for the substrate with 32%, 53% and 70% GC basepairs respectively (see Figure S3.1).

In our TPM experiments we focus on the local effects of temperature on global DNA conformation by measuring at temperatures below the melting temperature of all three DNA substrates (T = 23 – 52 °C). We probed the conformational state of many individual DNA molecules in this range of temperatures by tracking the excursion of the attached bead (quantified by the root mean square (RMS) value of its excursion). If the length of a tether is not changed, the RMS of the attached bead is a measure of its apparent persistence length; a tether attains a more compact formation if it becomes more flexible or if it is locally bent by a ligand (see Figure 3.1 A). At 23 °C, the measured RMS values ranged from RMS = 157.2 ± 0.5 nm (error represents standard error of the mean (SE)) for the 32% GC substrate to RMS = 161.4 ± 0.7 nm for the 70% GC substrate. DNA with a higher GC content yielded a slightly higher RMS value, due to an on average somewhat less flexible DNA substrate. Increasing the temperature caused the RMS values of all DNA tethers to decrease (see Figure 3.2 A), pointing to an increase in flexibility.

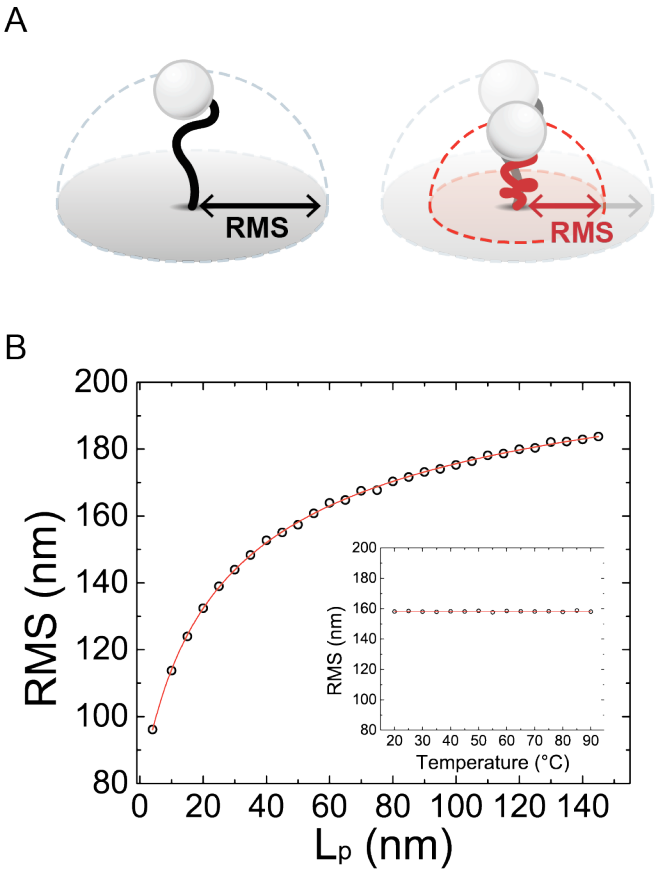


Figure 3.1 RMS is a measure for DNA persistence length. **A.** Schematic representation of a TPM experiment. *Left*: The RMS value quantifies the excursion of a bead attached to a single DNA molecule, tethered to the surface of the flow cell. *Right:* As the persistence length decreases, the DNA molecule will be in a more compact configuration (*red DNA molecule*), which leads to a smaller RMS value. **B.** RMS dependence on persistence length L_p obtained from bead movement simulations (L_0 = 233 nm and d_{bead} = 460 nm) fitted with Equation (3.1). Inset shows that the bead movement (RMS) itself is not affected by temperature when DNA parameters are kept constant (L_0 = 233 nm, L_p = 50 nm and d_{bead} = 460 nm).

To quantitatively relate the measured RMS values to the physical properties of the DNA, we performed numerical simulations describing the movement of the bead as a function of DNA persistence length L_p. The RMS of the simulated tethered beads was calculated from the x and y positions obtained from the simulations (see Materials and Methods). Figure 3.1 B shows the relation between the persistence length and the measured RMS value for a tether with contour length L_0 = 233 nm (685 bps) and bead diameter d_{bead} = 460 nm. By fitting the results with a hyperbola function, we obtained the following empirical relation between L_p and RMS:

$$\mathrm{RMS} = 233 - \frac{156}{(1 + 0.08 \cdot L_p)^{0.45}} \qquad (3.1)$$

To validate our approach and to ensure that the measured change in RMS is exclusively due to changes in flexibility of the DNA, we tested if the temperature has an effect on the bead motion itself. In our simulations a change in temperature may affect the movement of the bead by influencing the viscosity of the solution and the thermal fluctuations on the bead and the DNA. Despite these temperature effects, the bead movement (L_0 = 233 nm, L_p = 50 nm and d_{bead} = 460 nm) yielded a constant RMS value at different temperatures (see inset Figure 3.1 B), which indicates that the RMS itself is not influenced by temperature within a range of 20 – 90 °C. This confirms that the measured RMS is dependent only on the physical properties of the DNA, which we can describe as a function of apparent persistence length according to Equation (3.1). The apparent persistence length of the DNA molecule is a measure for the average flexibility of a heterogeneous chain (192,196).

Using Equation (3.1) we calculated the apparent persistence length of the DNA substrates within the measured temperature range. Figure 3.2 B shows that the apparent persistence length scales linearly with temperature within this temperature range. Linear fitting of the data ($L_p = L_p^0 - C \cdot T$) yielded a temperature dependence of the persistence length (C), which is

slightly dependent on the GC percentage of the DNA substrate (see Table 3.1). L_p^0 denotes the apparent persistence length at T = 0 °C, but as water will freeze around this temperature, this linear dependence is not valid at temperatures close to T = 0 °C. The generic effect of temperature on DNA flexibility is in agreement with recent magnetic tweezer studies (245,246). The observed linear dependence of apparent persistence length on temperature is in good qualitative agreement with the work of Geggier *et al.* (240), in which cyclization of short DNA fragments was investigated in the range of 5 – 42 °C in ensemble measurements.

Table 3.1 Temperature dependent DNA properties of DNA substrates (685 bp) with different GC content.

GC content	C (nm/ °C)	L_p^0 (nm)	T_m (°C)
32%	0.66 ± 0.05	62.7 ± 1.9	76.8 ± 0.5
53%	0.79 ± 0.05	71.2 ± 2.3	84.9 ± 0.5
70%	0.82 ± 0.03	74.9 ± 1.2	91.2 ± 0.5

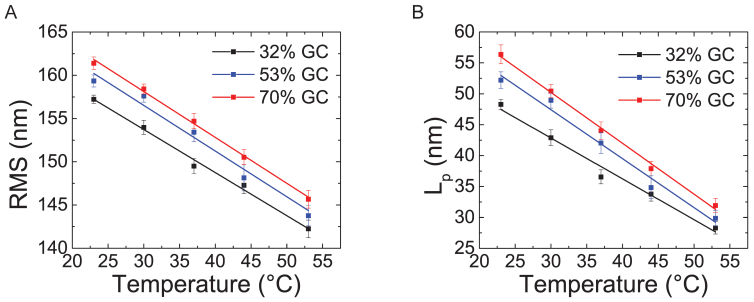


Figure 3.2 Temperature dependent DNA flexibility **A.** RMS distances of dsDNA molecules (685 bp) with a GC percentage of 32%, 53% and 70% as a function of temperature. Error bars represent the standard error of the mean (N = 58 – 352) **B.** Apparent persistence length of all three substrates as a function of temperature. L_p is calculated from data showed in A. according to Equation (3.1). Lines show a linear fit (see Table 3.1 for fitting parameters).

Influence of temperature on the binding and bending of DNA bending proteins

A higher flexibility of DNA at higher temperatures could be an important element contributing to DNA organization and dynamics in organisms with high growth temperatures. As the apparent persistence length of DNA is reduced by increasing temperature, the conformation of the molecule becomes more compact independent of architectural protein binding. In addition to such a direct effect of temperature on the intrinsic properties of DNA, the interaction between architectural proteins and DNA might also be affected by temperature. However, single-molecule experiments concerning protein-DNA interactions of architectural proteins have been traditionally conducted at room temperature.

To determine the architectural properties of such proteins under conditions more relevant for the *in vivo* situation, we investigated the effects of the DNA bending proteins Cren7 and Sul7 from thermophilic *S. solfataricus* and HU from mesophilic *E. coli* on DNA structure at physiological temperatures. *Sulfolobus* species (40,229) live at temperatures in the range of 40 – 90 °C (17). The Cren7 and Sul7 proteins encoded by these organisms are structural homologues, both inducing rigid bends upon binding to DNA (90) by intercalating into the minor groove (97,101). Our earlier single-molecule micromanipulation experiments performed at room temperature (~ 23 °C) revealed that this bending results in a decrease of apparent persistence length and thus compaction of DNA molecules (90). DNA melting experiments revealed that both Cren7 and Sul7 increase the melting temperature (63,70), which might imply an important role for these proteins in maintaining DNA integrity at high temperatures. To determine how the interaction of Cren7 and Sul7 with DNA is affected by temperature in the sub-melting regime in which DNA becomes more flexible, we performed TPM experiments at both 23 °C and 52 °C. For these studies we used the 32% GC substrate, which is representative for the average GC content of the *S. solfataricus* genome (22). The persistence length L_p for

protein concentrations of 0 – 1600 nM for both Cren7 and Sul7 at 23 °C and 52 °C, calculated from the measured RMS values using Equation (3.1) is shown in Figure 3.3 A and B. At 23 °C both proteins reach a minimum persistence length of ~ 10 nm, which is consistent with our earlier measurements using magnetic tweezers conducted at this temperature (90). This agreement further validates our approach of converting RMS values into values for persistence length L_p. At 52 °C (*red data points*) compaction sets in at lower protein concentrations than at 23 °C (*black data points*), but the maximum level of compaction achieved at both temperatures is comparable. At high protein concentrations, when the DNA is saturated with protein the change in temperature does not measurably influence the conformation of the protein-DNA complexes.

To quantitatively compare the binding affinities of the proteins at the different temperatures, we calculated the fractional coverage (υ) of the DNA as follows (178):

$$\upsilon = \frac{\sqrt{1/L_p} - \sqrt{1/L_{p,\,bare}}}{\sqrt{1/L_{p,\,saturated}} - \sqrt{1/L_{p,\,bare}}} \qquad (3.2)$$

where L_p represents the measured persistence length, $L_{p,saturated}$ the minimum persistence length at saturation and $L_{p,bare}$ the persistence length of bare DNA. In this approach it is assumed that each bound protein has an equal contribution to the decrease in DNA stiffness. Figure 3.3 C and D show the fractional coverage as a function of protein concentration.

To calculate binding affinities under the different conditions, the fractional coverage was fitted to the McGhee and von Hippel theory (247). This model describes protein binding to DNA in terms of the association constant (K), a cooperativity parameter (ω) and the protein binding site (n). Using a footprint of n = 8 bp for both proteins (95,101), we obtained fitting

parameters for K and ω. Both Cren7 and Sul7 show an increase in binding affinity at 52 °C compared to 23 °C. The association constant found for Cren7 was $K_{Cren,23}$ = (2.5 ± 0.2) · 10^5 nM^{-1} at room temperature (23 °C) and increased to $K_{Cren,52}$ = (7.4 ± 1.4) · 10^5 nM^{-1} at 52 °C. The cooperativity factor of $\omega_{Cren,23}$ = 18.8 ± 3.2, indicates low cooperativity in binding, which is somewhat affected by the increase of temperature as $\omega_{Cren,52}$ = 10.4 ± 3.6. Sul7 exhibits a lower DNA binding affinity compared to Cren7 as $K_{Sul,23}$ = (1.3 ± 0.3) · 10^5 nM^{-1} at 23 °C and $K_{Sul,52}$ = (1.8 ± 0.2) · 10^5 nM^{-1} at 52 °C, yielding a slightly higher binding affinity at 52 °C. Similar to Cren7 the cooperativity in binding of Sul7 to DNA is essential independent of temperature: $\omega_{Sul,23}$ = 8.4 ± 2.4 and $\omega_{Sul,52}$ = 8.1 ± 2.0.

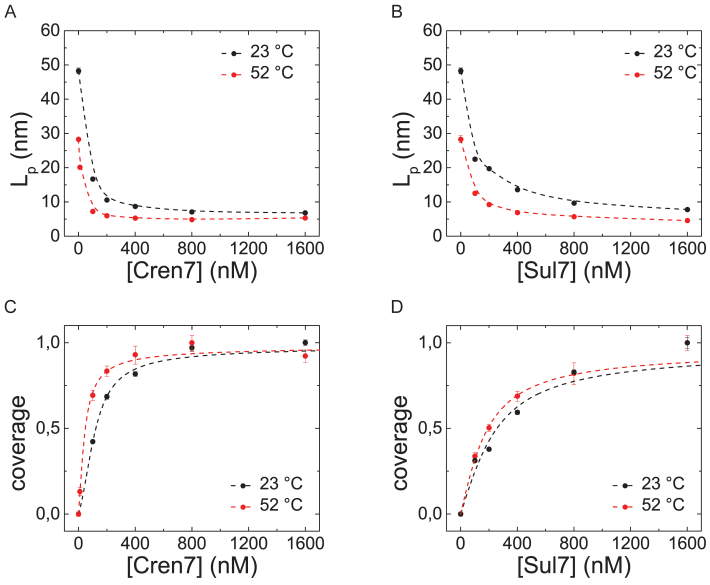


Figure 3.3 Apparent persistence length as a function of protein concentration of **A.** Cren7 and **B.** Sul7 at 23 °C and 52 °C. **C and D.** Fractional coverage as a function of protein concentration, calculated according to Equation (3.2) from data showed in A and B. Fitting the coverage to the McGhee- von Hippel theory showed that binding affinity is increased at 52 °C: $K_{Cren,23}$ = (2.5 ± 0.2) · 10^5 nM^{-1} and $K_{Cren,52}$ = (7.4 ± 1.4) · 10^5 nM^{-1} and $K_{Sul,23}$ = (1.3 ± 0.3) · 10^5 nM^{-1} and $K_{Sul,52}$ = (1.8 ± 0.2) · 10^5 nM^{-1} for Sul7. Error bars represent the standard error of the mean (N = 69 – 352).

To investigate whether the effect of temperature on DNA binding shown for Cren7 and Sul7 is generic for DNA bending proteins, we performed similar experiments with the bacterial chromatin protein HU, the most abundant DNA bending protein in *E. coli*. It exhibits two different DNA binding modes depending on the stoichiometry: DNA bending and compaction at low protein concentrations and DNA stiffening at high protein concentrations (195,196,248). TPM measurements performed at 23 °C show that HU induces compaction by reducing the apparent persistence length to a minimum of L_p = 12.9 ± 0.2 nm at a protein concentration of 400 nM (see Figure 3.4). At HU concentrations > 400 nM the persistence length increases up to L_p = 20.1 ± 0.4 nm at a protein concentration of 1600 nM reflecting the transition into the stiffening mode. In this concentration regime, HU proteins bind side-by-side forming filaments, which stabilizes the DNA helix (249). We did not observe an increase of the persistence length above that of bare DNA as observed in previous micromanipulation experiments of single HU-DNA complexes (195,196). However, previously reported TPM experiments exhibited a similar trend: the persistence length increases at high HU concentrations, but does not exceed the persistence length of bare DNA molecules (248). Possibly, the force applied in the previous DNA micromanipulation experiments facilitates the transition into the stiffening mode of HU, giving rise to the apparent discrepancy in observed persistence lengths.

Next, we carried out TPM measurements for the same protein concentration range at 37 °C, the optimum growth temperature of *E. coli*. Surprisingly, a temperature increase of only 14 °C had a large effect on the measured apparent persistence length. Although the maximum level of compaction was achieved at a similar protein concentration (400 nM), the persistence length at this concentration was reduced significantly to L_p = 6.1 ± 0.2 nm at 37 °C. In the stiffening regime (> 400 nM) the persistence length increased slightly to L_p = 10.7 ± 0.3 nm at 1600 nM, indicating that the DNA stiffening still occurs in this regime at 37 °C. As the maximum level of compaction is achieved at the same protein concentration, the

affinity of HU for DNA is not affected by the change in temperature from 23 to 37 °C. The change in temperature did however significantly enhance the degree of compaction induced by HU, suggesting that the degree of bending induced by HU is increased at 37 °C.

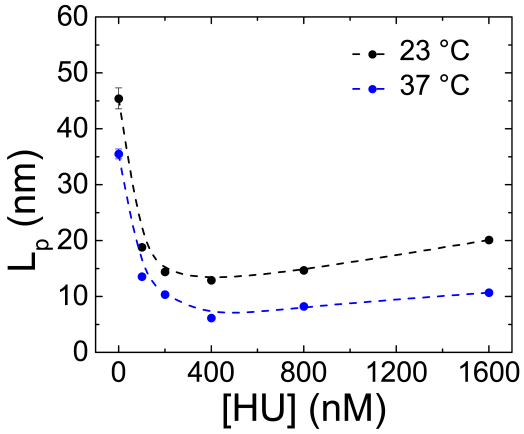


Figure 3.4 Apparent persistence length as a function of HU concentration at 23 °C (*black*) and 37 °C (*blue*). Compaction is increased at 37 °C. Error bars represent the standard error of the mean (N = 52 – 236).

DISCUSSION

Our single-molecule TPM experiments reveal that the flexibility of DNA strongly depends on temperature in the range of 23 – 52 °C. The temperature coefficient of the persistence length (C) is slightly dependent on GC content of the DNA substrate and ranges from C = 0.66 ± 0.05 nm/°C for the DNA substrate with an average GC content of 32% to C = 0.82 ± 0.03 nm/°C for the substrate with 70% GC content. Although the temperature effect on the apparent persistence length observed in our study is much more pronounced than in previous work by Geggier *et al.* (240), the observed trend is similar: increased temperature lowers the resistance to bending of the DNA. What causes such a strong temperature dependence

on the L_p remains unclear. Thermal stability of the dsDNA helix involves base pair interactions between bases of complementary strands and base-stacking interactions between adjacent bases. Although Yakovchuk *et al.* (250) found that the base-stacking interactions are the dominant factor in stabilizing the double-stranded helix, they did not show how this relates to the bending rigidity of the DNA. A recent study addressed this question by measuring mechanical properties of chemically modified DNA molecules, with altered charge or base stacking interactions (251). This revealed that stacking interactions do not straightforwardly correlate with the mechanical properties of DNA, such as the persistence length. It is thus unlikely that stacking interactions alone cause such a pronounced temperature dependent persistence length. Spontaneous sharp bending of dsDNA has been explained in different models involving kinks (caused by unstacking of adjacent basepairs) or small melting bubbles (disrupted base pairing and/or base stacking). Probably increasing temperature enhances both of these processes, strongly influencing the apparent persistence length (as observed in our experiments). A recent theoretical study indeed suggests a model that describes the experimentally observed temperature dependent DNA persistence length as a function of both an isotropic temperature dependent bending stiffness of the dsDNA helix and small local melting bubbles (252).

The fact that the flexibility of dsDNA strongly depends on temperature has important implications for genome conformation *in vivo*. Organisms will experience different mechanical properties of DNA depending on the temperature of their natural habitat. As temperature increases the bendability of dsDNA, organisms living at elevated temperatures could benefit from this increased bendability in the light of genome compaction. The persistence length directly relates to the size of an unconstrained dsDNA molecule in solution, typically quantified by the so-called radius of gyration (R_g). For instance, the genome of an archaeal or bacterial cell is in the order of millimeters in length which corresponds to $R_g \approx 4\ \mu$m at T = 25 °C (L_p = 50 nm). A change in temperature to 52 °C would reduce the radius of gyration of a genome of that length to $R_g \approx 3.2\ \mu$m (a reduction in

effective volume by about 25%). Temperature itself could thus serve as a mechanism to modulate the compaction levels of genomic DNA. It is therefore an important aspect to consider when studying *in vivo* chromatin organization and compaction of thermophilic organisms and psychrophilic organisms living at temperatures that differ from room temperature.

Our results show that temperature not only affects the intrinsic properties of DNA, but also influences protein-DNA interactions of DNA-bending proteins. Interestingly, this temperature dependence is not shared among the different DNA-bending proteins investigated; it seems to depend on the nature of the protein-DNA interactions. Both Cren7 and Sul7 from thermophilic *S. solfataricus* exhibit an increased binding affinity at an elevated temperature of 52 °C. Increased binding affinity at higher temperatures was shown before for Sul7 by isothermal titration calorimetry (ITC) (253). In these ITC studies the binding affinity for poly(dGdC) increased about 5 fold when increasing the temperature from 25 °C to 80 °C. Molecular dynamics simulations of Cren7 also showed an increased affinity at elevated temperatures (104). The increase in binding affinity could be attributed to the increased flexibility of DNA at higher temperatures, which lowers the energy barrier to induce DNA bending. Another mechanism, that could contribute to an increased affinity, is that transient melting bubbles make the DNA more accessible for binding of Cren7 and Sul7, as these proteins intercalate into the minor groove of DNA.

In contrast with Cren7 and Sul7, the DNA-bending protein HU from mesophilic *E.coli* did not show an increased binding affinity upon a temperature increase of $\Delta T = 14$ °C. Instead, HU binding at 37 °C resulted in significantly enhanced compaction. Previous studies, performed at room temperature, have shown that HU induces flexible bends, which can range from 0 – 180° with equal probability (196,209). An increase in temperature could bias the bending angle distribution towards higher bending angles as the energy barrier required to bend the DNA is decreased. The observed increase in compaction could thus be a result of enhanced bending at

higher temperature. At 37 °C the bimodal behavior of HU still persisted: the observed compaction reached a maximum at 400 nM and at concentrations > 400 nM the apparent persistence length increased. The stiffening regime is however less pronounced compared to 23 °C as the apparent persistence length reached a maximum of L_p = 10.7 ± 0.3 nm at 1600 nM (compared to L_p = 20.1 ± 0.4 nm at 23 °C). As the molecular mechanism underlying the stiffening mode is not understood, it is difficult to explain what causes the difference in stiffening at different temperatures. Possibly, the increased flexibility of the DNA at 37 °C effectively counteracts the stiffening effect of the HU filaments. Also, protein-protein interactions could be affected by temperature, influencing the side-by-side binding of HU proteins, which causes the observed stiffening. A smaller distribution of bending angles could also cause a less tight packing of the proteins on the DNA within the stiffening regime, causing a less pronounced stiffening effect, as observed. Indeed it has been suggested that a flexible bending angle is needed for tight packing of proteins along the DNA (90).

The increase in flexibility of DNA at increased physiologically relevant temperatures is important when aiming to understand chromatin organization *in vivo*. It does not only lead to a more compact configuration of the genomic DNA, but can also have an important effect on cellular processes such as gene regulation. As DNA structure and topology are temperature dependent this may affect gene expression on a global level (254). DNA supercoiling is proposed to act as a global regulator as it changes in response to environmental conditions and affects the expression of many genes (162). Analogous to supercoiling, temperature dependent DNA flexibility (and associated local changes in twist) and global compaction could provide a mechanism for temperature sensitive gene regulation.

MATERIALS AND METHODS

Protein purification

Cren7 was purified as described before (90). Sul7 was purified from *E. coli* strain BL21-CodonPlus(DE3) containing plasmid pRD26 (a pET11a derivative containing the gene encoding Sul7 (gene SSO10610) from *S. solfataricus*). Cells were grown in LB medium up to $OD_{600} \approx 0.4$, and expression was induced using 0.5 mM IPTG at 37 °C. Two hours after induction, cells were harvested by centrifugation and resuspended in 20 ml buffer A [50 mM Tris–HCl (pH 8.0), 2 mM $MgCl_2$, 0.1% Triton X100, 386 µg/ml benzamidine hydrochloride, 10 mM β-mercaptoethanol]. Cells were lysed by sonication, 1000 units OmniCleave Endonuclease (Westburg) per gram cells was added and the cell lysate was incubated for 30 min at room temperature. After heating the cell lysate for 40 min at 70 °C, 1 ml of 0.5 M EDTA was added. The cell lysate was centrifuged for 30 min at 37.000 r.p.m. and filtered through a 0.45 µm membrane filter (MiliPore). The supernatant was applied to a HiTrap-S column (GE Healthcare), equilibrated in buffer B [10 mM KPO_4 (pH 7.0), 10% glycerol]. Protein was eluted with a linear gradient of 0 – 1 M NaCl in buffer B. Cren7 and Sul7 proteins were dialyzed at 4 °C against a storage buffer [20 mM HEPES (pH 7.5), 100 mM NaCl, 10% Glycerol, 10 mM β-mercaptoethanol], and stored at -80 °C until required. Protein concentrations were determined using a Bicinchoninic Acid (BCA) protein Assay (Thermo Scientific). HU protein was purified as described before (255).

DNA constructs

End-labeled DNA constructs of 685 bp with different GC content (32%, 53% and 70% GC) were generated by PCR using plasmids pRD118, pNP83 (256) and pBTH154 as templates and biotin- and digoxygenin-labeled primers (designed to specifically yield a product of the desired length based on the sequences cloned into these plasmids). pRD118 was constructed by inserting a 685 bp long fragment from the *S. solfataricus* P2 genome (22)

into the NdeI and BamHI site of pET3-his (257). pBTH154 was constructed by inserting the dasR gene (SCO5231) of the *Streptomyces coelicolor* A3 (2) (258) into the XbaI and XmaI site of pUT18C (Euromedex). All PCR products were purified using a GenElute PCR Clean-up kit (Sigma-Aldrich). DNA substrates were checked for local GC percentage and predicted curvature (see Figure S3.2) to ensure the variation is small and to prevent local curvature dominating the global flexibility of the substrates. DNA substrates used for the bulk melting experiments were obtained following the same approach with unlabeled primers.

Bulk melting curves

Melting curves of the three DNA substrates with different GC-content (see above) were measured using a Varian Cary300Bio UV-Vis spectrophotometer, measuring ultraviolet absorbance at λ = 260 nm. DNA was diluted to a final concentration of 4 µg/ml in a buffer containing 10 mM HEPES (pH 7.5) and 100 mM NaCl. Temperature was increased with a rate of 1 °C/min within a range of 25 – 95 °C. The melting temperature T_m is defined as the temperature at which half of the dsDNA is dissociated into ssDNA, which equals the temperature at which the slope of the melting curve is at its maximum. To determine the T_m the first derivative was calculated and the peak position (corresponding to the T_m) was determined by fitting a Gaussian distribution.

Tethered particle motion experiments

Flow cells (volume ~ 30 µl) were incubated with 20 µg/ml anti-DIG antibodies (Roche) and for 5 min. Passivation of the surface was achieved by flushing the flow cell with 0.4% (w/v) Blotting Grade Blocker (BGB) (Bio-Rad) in buffer I [10 mM Tris (pH 7.5), 150 mM NaCl, 1 mM EDTA, 1 mM DTT, 3% glycerol, 100 µg/ml acetylated BSA (Ambion)] and incubating for 15 min at the desired temperature (23 – 52 °C). The flow cell was flushed with buffer I, filled with 200 pM DNA solution (functionalized with biotin and DIG) and incubated for 10 min. and flushed again with buffer I.

Streptavidin coated polystyrene beads with a diameter of 0.46 μm (1% w/v; G Kisker) were diluted 300 times in buffer I, flushed into the flow cell and incubated for 10 min to allow binding to the biotin-ends of the DNA. The flow cell was washed with protein diluted in buffer II (10 mM HEPES (pH 7.5), 100 mM NaCl, 0.2% (w/v) BGB) for Cren7 and Sul7 or buffer III (20 mM HEPES (pH 7.9), 60 mM KCl, 0.2% (w/v) BGB) for HU. After flowing in the final incubation solution with or without protein the flow cell was closed and incubated at the desired temperature (23 – 52 °C) for 10 min before starting measurements. Experiments without protein were all done in buffer II.

Tethered particle motion experiments were performed on an inverted Nikon microscope (Diaphot 300), using a 100x oil-immersion objective (NA = 1.25). To control the temperature of the flow cell, a custom-built temperature control system was implemented by placing heating elements around the objective and inside the flow cell holder. Using a calibrated temperature probe the accuracy of the temperature set on the controller was verified in the area directly above the objective. Next, a feedback system was used to secure constant temperature within ±1 °C. The sample stage was isolated to avoid temperature drift within the sample. Images were acquired using a CMOS camera (Thorlabs) at 25 Hz. The x- and y- coordinates of individual beads were tracked real-time by custom-developed LabView software (National Instruments) as described previously (259).

Data Analysis

The root mean square (RMS) value of the excursion of each individual bead was calculated from x- and y-coordinates of a 40 s time trace (corrected for linear drift) by $\mathrm{RMS} = \sqrt{\left\langle (x-\overline{x})^2 + (y-\overline{y})^2 \right\rangle}$, where $\overline{x}$ and $\overline{y}$ are averaged over the full time trace. Symmetry of the excursion of the tethered beads was evaluated by calculating the anisotropic ratio $a = l_{\mathrm{major}} / l_{\mathrm{minor}}$ from the xy-scatter plots, where l_{major} and l_{minor} represent the major and minor axis of the xy-scatter plot respectively. Fluctuations of RMS in time were quantified

by calculating the relative standard deviation of the smoothed RMS $\sigma_{rel} = \sigma_{smoothed} / < \mathrm{RMS} >_{smoothed}$, where $\sigma_{smoothed}$ represents the standard deviation of $\mathrm{RMS}_{smoothed}$, which represents the RMS smoothed over a 2 s time window. Only tethers with a high symmetry $a \leq 1.14$ and small RMS fluctuations $\sigma_{rel} \leq 0.06$ where qualified as good tethers and selected for further analysis. For each measured condition RMS values were obtained by fitting a single Gaussian to the histogram of the RMS values of individual tethers (N = 52 – 352).

Bead movement simulations

Bead movement is simulated numerically by solving the Langevin equation for a tethered bead. This is done for both translation and rotation (260):

$$\begin{aligned} x_{n+1} &= x_n + (\gamma - 1) \cdot F_{ext} + F_{brown} \Delta t \\ \varphi_{n+1} &= \varphi_n + (\beta - 1) \cdot \Theta_{DNA} + \Theta_{brown} \Delta t \end{aligned} \qquad (3.3)$$

where x_n and x_{n+1} are the bead's three dimensional position at step n and n+1 of the simulation respectively. The angles φ_n and φ_{n+1} denote the bead's orientation. γ and β are the effective translational and rotational drag coefficient vectors, which are calculated using Faxen's law. F_{ext} is the external force working on the bead caused by gravity, buoyancy, the surface and the DNA. Θ_{DNA} is the torque exerted on the bead caused by the DNA. The DNA is modeled by the finite WLC model assuming fully constraint boundary conditions (261). F_{brown} and Θ_{brown} are the fluctuating thermal force and torque that are described by the fluctuation-dissipation theorem and are Gaussian distributed (262). Δt is the time step of the simulation and is typically set to 1 μs. Decreasing Δt did not affect the statistical properties of the simulated traces.

SUPPLEMENTARY INFORMATION

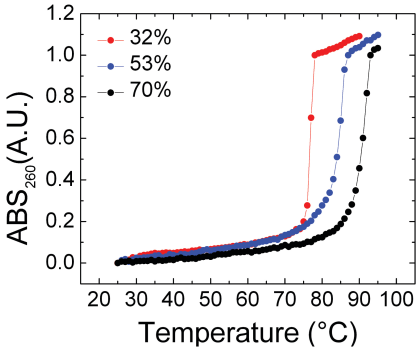


Figure S3.1 Bulk dsDNA melting curves of 685 bp fragments containing 32%, 53% and 70% GC basepairs showed melting temperatures of $T_{m,32\%}$ = 76.8 ± 0.5 °C, $T_{m,53\%}$ = 84.9 ± 0.5 °C and $T_{m,70\%}$ = 91.2 ± 0.5 °C respectively.

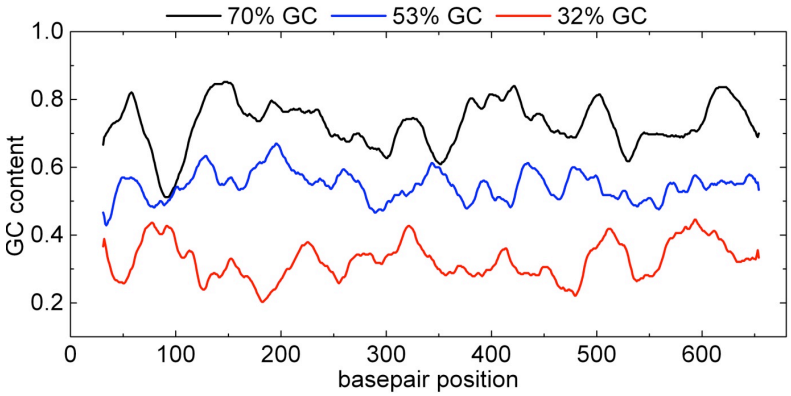


Figure S3.2 Local GC content of DNA substrates used in TPM experiments. Values were obtained using the bend.it server (263) with a window size of 31 and smoothed by adjacent averaging over 10 datapoints.

Chapter 4

Alba shapes the archaeal genome using a delicate balance of bridging and stiffening the DNA

This chapter has been published as:

Niels Laurens*, Rosalie P.C. Driessen*, Iddo Heller, Daan Vorselen, Maarten C. Noom, Felix J.H. Hol, Malcolm F. White, Gijs J.L. Wuite†, Remus T. Dame†, *Nature Communications*, **3**, 1328 (2012)

---

* † These authors contributed equally to this work

ABSTRACT

Architectural proteins play an important role in shaping the genome and act as global regulators of gene expression. How these proteins jointly modulate genome plasticity is largely unknown. In archaea one of the most abundant proteins, Alba, is considered to play a key role in organizing the genome. Here we characterize the multimodal architectural properties and interplay of the Alba1 and Alba2 proteins using single-molecule imaging and manipulation techniques. We demonstrate that the two paralogues can bridge and rigidify DNA and that the interplay between the two proteins influences the balance between these effects. Our data yield a structural model that explains the multimodal behaviour of Alba proteins and its impact on genome folding.

INTRODUCTION

The genome inside organisms from all domains of life is very long compared to the size of the cell or compartment to which it is confined. Hence genomic DNA needs to be tightly compacted in order to fit. To facilitate compaction cells utilize generic physical mechanisms such as molecular crowding and supercoiling (264). In addition, cells synthesize architectural proteins that shape the structure of the genome (14). Despite a general lack of homology between architectural proteins in organisms from different domains of life, the way they shape the genome seems highly conserved (14). In bacteria architectural proteins such as HU and H-NS, respectively, shape the genome by inducing bends into DNA (196) or promoting loop formation between distant sites along the DNA (201). All these proteins act jointly to organize and compact the genomic DNA into a structure referred to as the nucleoid. In eukaryotes the genome is organized by wrapping the DNA around histones, yielding a nucleosomal fiber (34,265). This structure folds into a chromatin fiber, which is further organized into loops mediated by proteins such as CTCF and cohesin (117,266,267). Mitochondria also harbor their own genome. The mitochondrial genome is not organized by histones, but instead by a set of small architectural proteins, analogous to the bacterial chromatin (268,269). Relatively little is known about the architectural properties of the proteins that shape the genome in archaea (39,42). Archaea are unicellular organisms that share the DNA replication, transcription and translation machinery with eukaryotes (12) and are therefore often used as a model for the more complicated eukaryotic systems. Analogous to bacteria, archaea lack a nucleus and their genome is organized into a nucleoid, using strategies of both eukaryotes and bacteria. Several archaeal branches exist of which the main two are the Euryarchaea and the Crenarchaea. Euryarchaea encode legitimate homologues of eukaryotic histone proteins (270), while Crenarchaea generally lack histones and encode architectural proteins that use mechanisms as observed in bacteria (39,40).

Alba (Acetylation Lowers Binding Affinity) is one of the most abundant and highly conserved non-specific double stranded DNA binding proteins found throughout the archaeal domain (42,60,81) (~ 4% of the cellular protein in *S. shibatae* (68)), making it a key candidate to be involved in chromatin organization. Alba is a small protein (~ 10 kDa per subunit) that forms dimers in solution; larger multimeric forms of the protein have not been reported (59,76). The crystal structure of the protein has been known for over a decade (71), while an Alba-DNA co-crystal structure was resolved only recently (80). Alba has been suggested to have two distinct structural effects on DNA depending on the protein:DNA stoichiometry. At low stoichiometry (about 1 dimer per 15 bp) Alba is able to bring two duplexes of DNA together, while under saturating conditions (effectively 1 dimer per 5 bp) it binds cooperatively along the DNA (71,81,271). Both types of effects rely on dimer-dimer interactions. Dimer-dimer interactions promote the cooperative side-by-side binding along DNA and have a direct impact on the ability to form a filamentous structure (271). On the other hand, dimer-dimer interactions also appear to be responsible for bringing two DNA duplexes together (80). A striking observation is that cooperative side-by-side binding along the DNA is affected by a mutation of the F60 residue (271), which is also involved in the bridging dimer-dimer interactions (80).

S. solfataricus encodes two distinct Alba proteins denoted Alba1 and Alba2 (76) (36% sequence identity and 63% sequence similarity). Alba2 has a 10-fold lower transcription level compared to Alba1 (110) and is expressed at ~ 5% of the Alba1 level in stationary phase (76). Furthermore, Alba2 lacks the F60 residue that is responsible for the cooperative side-by-side binding of Alba1. Alba1 and Alba2 form obligate heterodimers, which have DNA binding properties different from Alba1 homodimers. Based on this notion it has been proposed that tuning the fraction of heterodimers is a mechanism to modulate Alba-mediated genome organization *in vivo* (73,76). Nevertheless, a comprehensive molecular model of the different modes of Alba binding, and the modulatory interplay between Alba1 and Alba2 in relation to genome organization is still unavailable.

The action of architectural proteins has been studied *in vitro* for many years using conventional biochemical assays as well as with single-molecule techniques more recently (179,181). These techniques allow the visualization and manipulation of single DNA molecules interacting with single proteins and reveal the kinetics and dynamics of these interactions. Here we use a set of complementary single-molecule techniques to study the action and interplay of Alba1 and Alba2 on DNA. We characterize the structural changes of DNA upon Alba binding and investigate the physical properties of Alba-DNA complexes. Based on our results we propose a model that explains the concentration dependent behaviour of Alba1 and its modulation by Alba2.

RESULTS

Visualizing Alba-DNA interactions

In order to investigate the architectural properties of Alba proteins, Alba-DNA complexes were visualized using Atomic Force Microscopy (AFM) at different protein:DNA stoichiometries. A series of representative images for each stoichiometry is shown in Figure 4.1. The complexes were classified by visual inspection of the AFM images into 'open', 'bridged' or 'condensed' (see Materials and Methods) and counted for different stoichiometries (see Figure S4.1).

Alba1

At protein-DNA ratios of 1:60 and 1:30 dimer:bp Alba1 binding resulted in the formation of intramolecular bridges. At 1:30 dimer:bp most Alba1-DNA complexes are bridged (60 %), of which most contained a single protein patch extending over part of the DNA molecule (Figure 4.1 C, left), and even some of the DNA molecules were found to be completely bridged (14 % of all molecules) (Figure 4.1 C, right). Both features suggest DNA-induced cooperative bridging, which is enhanced by the formation of stable bridges. At increased protein-DNA ratios (1:15 dimer:bp) multiple sites of bridging

are commonly seen within the same molecule, causing the DNA molecule to condense (Figure 4.1 D). At even higher protein-DNA ratios (1:7.5 dimer:bp) DNA molecules attained an open appearance (Figure 4.1 E), suggesting stiffening of the DNA by cooperative side-by-side binding along the DNA molecule as reported previously (271).

Alba1:Alba2

To investigate the role of the dimer-dimer interface in cooperative side-by-side binding we studied Alba1:Alba2 heterodimers. Since the Alba2 subunit lacks the conserved F60 dimer-dimer interface we would expect cooperativity to be suppressed. Binding of the heterodimers at protein-DNA ratios of 1:60 and 1:30 dimer:bp resulted in intramolecular bridges as seen with the wild-type Alba1 (Figure 4.1 G – H). However, the images show that the heterodimers do not exhibit the same cooperative behaviour as Alba1 homodimers, as none of the molecules were completely bridged. At increased protein:DNA ratios (1:15 dimer:bp) relatively more bridged and condensed molecules were observed (Figure 4.1 I). In stark contrast with Alba1 homodimers, at a concentration of 1:7.5 dimer:bp, all DNA complexes were strongly condensed instead of stiffened (Figure 4.1 J), pointing, as expected, to a reduction in cooperativity between adjacent dimers in comparison with the Alba1 homodimers.

Alba1 F60A

As the F60 residue has been shown to be responsible for cooperative side-by-side binding, we studied DNA binding properties of the Alba1 F60A mutant (271). Binding of the Alba1 F60A mutant at concentrations of 1:60 and 1:30 dimer:bp yielded intramolecular bridges comparable to the Alba1:Alba2 protein-DNA complexes. DNA condensation was observed at increased protein:DNA ratios (Figure 4.1 N and O). Stiff and open molecules as with the wild type protein were never observed. These results confirm the importance of the dimer-dimer interactions for cooperative side-by-side binding and the essential role of residue F60 therein.

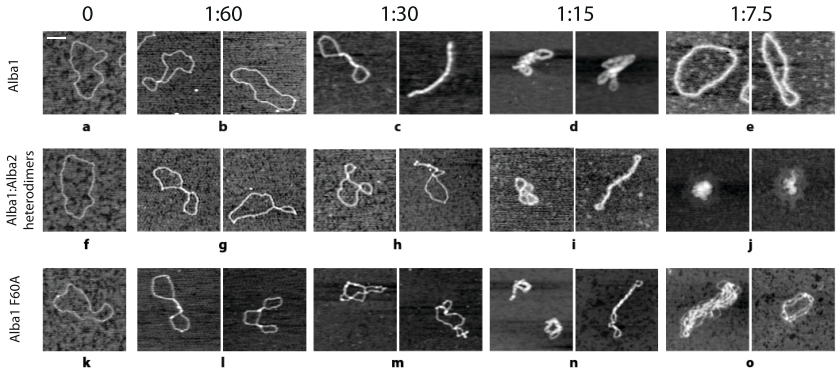


Figure 4.1 Representative images of Alba-DNA complexes visualized by Atomic Force Microscopy (AFM). Nicked pRD24 plasmids are incubated at different stoichiometries (indicated as dimer:bp) with different Alba proteins. Alba1 forms bridges at 1:60 and 1:30 ratios, condenses molecules at a 1:15 ratio, and forms stiff open DNA molecules at 1:7.5 ratio. Alba1:Alba2 heterodimers form bridged protein-DNA complexes at 1:60, 1:30 and 1:15. However, at 1:7.5 ratio DNA molecules are highly condensed, and do not show a stiffened configuration as Alba1. Alba1 F60A dimers are able to from bridges between DNA duplexes at all different concentrations. Alba1 F60A and Alba2 both lack the crucial F60 residue and its equivalent. The sequence identity of the α1-helix responsible for dimer-dimer interactions in both proteins is only 36%, which might explain the differences between Alba1 F60A homodimers and Alba1:Alba2 heterodimers. Scale bar represents 100 nm.

Mechanical properties of Alba1-DNA complexes

First we investigated the effect of Alba1 on the physical properties of single DNA molecules. In these experiments, a DNA molecule is captured between two optically trapped beads and gradually stretched while the force on the molecule is recorded (Figure 4.2 A), resulting in a force-distance (FD) curve. First a FD curve of bare DNA is measured as a reference (Figure 4.2 B, *black curve*). Subsequently, the same molecule is incubated with Alba1, and a new FD curve is recorded to measure the effect of Alba1 on the physical properties of the DNA (Figure 4.2 B, *green and magenta curves*). At

low protein concentrations (1 – 100 nM) intersegmental bridging was observed (as reflected in numerous peaks in force at distances far below the contour length), which could be disrupted by applied force in some cases (Figure 4.2 B, *green curve*). However, in most cases, bridges could resist forces even up to 400 pN (maximum force that can be applied in our instrument), pointing to extremely stable bridging interactions. At concentrations where the DNA molecules were saturated with Alba1 (> 100 nM) no bridging was observed (Figure 4.2 B, *magenta curve*). Instead, at low extension less force was required to stretch the protein-DNA (Figure 4.2 B, *magenta vs. black curve*), which indicates that less entropic energy is present and that the binding of Alba1 stiffens the DNA molecule. To quantify the observed stiffening, FD curves were recorded over a range of protein concentrations (1 – 2000 nM) and fitted with the eWLC model (Equation (4.1)). The contour length (L_0) and the stretching modulus (K_0) were found to be constant over the measured concentration range (see Table S4.1). However, the persistence length (L_p), which is a measure for the flexibility of the DNA, increased up to a five-fold higher value at protein concentrations > 100 nM (L_p = 260 ± 30 nm for 2 mM Alba1 compared to L_p = 49 ± 2 nm for bare DNA, see Figure 4.2 C, *black circles*). The increase in persistence length confirms that protein binding induces stiffening of the DNA as was observed in our AFM experiments (Figure 4.1 E). To investigate the impact of Alba on the twist parameters of DNA (twist-stretch coupling and torsional stiffness) the tWLC model (Equation (4.2)) was fitted to the FD curves measured at 2 mM Alba1, obtaining g_0 = - 560 ± 100 pN· nm, g_1 = 19 ± 6 nm, and F_C = 31 ± 7 pN (N = 10). These values show no deviation within the error from the values of bare DNA (272), indicating that Alba1 does not change the twist-stretch coupling of DNA.

Cooperative binding of Alba1

The DNA binding kinetics of Alba1 on DNA can be extracted from the amount of Alba1 bound to the molecule as a function of concentration. The persistence length of the Alba1 coated DNA molecules reached its

maximum around 100 nM and stayed constant up to 2 mM, indicating that the DNA is saturated ($\upsilon = 1$). The fractional DNA coverage is calculated from the persistence length as a function of the Alba1 concentration using Equation (4.5) (Figure 4.2 D, *black circles*). Using the reported value of the binding site n = 5 bp (71,80), the McGhee-von Hippel model (Equation (4.6)) fits well to the calculated coverage levels of Alba1-DNA binding (Figure 4.2 D, *black line*). The fit yields an association constant $K = (4 \pm 1) \cdot 10^2$ nM^{-1} and a cooperativity factor $\omega = (3 \pm 1) \cdot 10^2$. The value of $\omega >> 1$ confirms the cooperative binding of Alba1 along DNA, as observed in AFM experiments and shown in earlier work (271).

Alba1:Alba2 heterodimers lack cooperative binding behaviour

To investigate the functional differences and interplay between the two homologous Alba proteins we characterized the effect of Alba1:Alba2 heterodimers on the mechanical properties of the DNA. We find that the heterodimers increase the persistence length of DNA (Figure 4.2 C, *magenta triangles*), yet, to a lower extent than Alba1. The impact of the Alba1:Alba2 heterodimer on DNA is similar to that of the Alba1 F60A mutant, indicating that the F60 residue plays an essential role in the cooperative increase of the persistence length (Figure 4.2 C, *green squares*). To quantify this effect we determined the fractional coverage as a function of the protein concentration for both Alba1:Alba2 heterodimers and the Alba F60A mutant using Equation (4.5) (Figure 4.2 D, *magenta triangles and green squares respectively*). To investigate if both binding curves are consistent with a loss in cooperativity we plotted the McGhee-von Hippel model (Equation (4.6)) using the derived binding affinity for Alba1 ($K = 4 \cdot 10^2$ nM^{-1}), the same binding site (n = 5 bp) but without cooperativity ($\omega = 1$). The McGhee-von Hippel binding isotherm using these parameters describes both the Alba1:Alba2 heterodimer and Alba F60A binding curve very well (Figure 4.2 D, *dotted line*), reaffirming that the F60 residue is responsible for side-by-side cooperative dimer-dimer interactions.

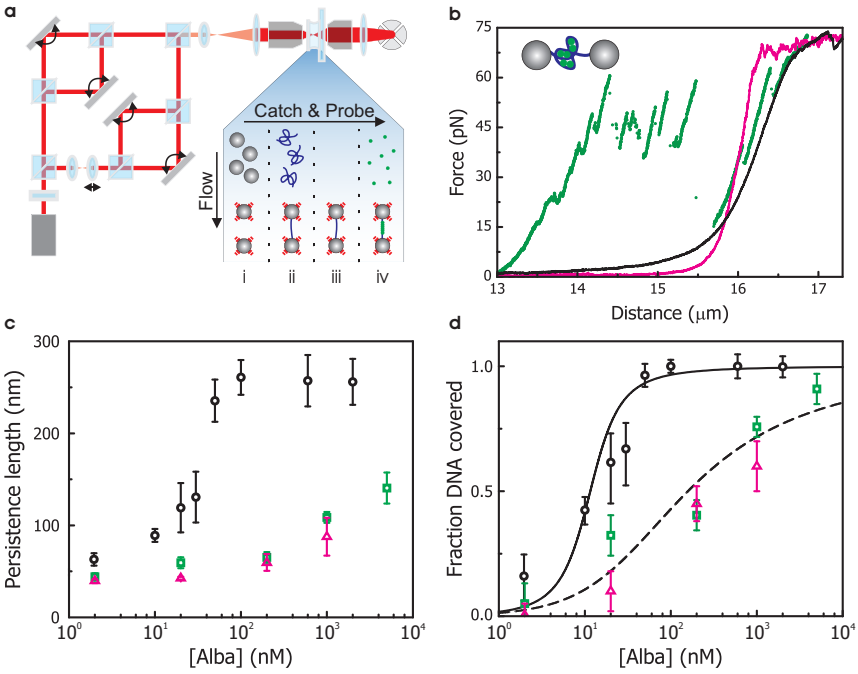


Figure 4.2 Characterization of Alba-DNA binding by DNA micromanipulation experiments. **A.** Schematic of the optical trapping setup used for both single and dual DNA experiments. Four optical traps are generated using a single laser. Each trap is steerable by a piezo-driven mirror. DNA molecules are caught (i and ii), tested (iii) and measured (iv) in a multi-channel flow cell. **B.** Force distance (FD) curves of individual λ-DNA molecules without protein (*black curve*), in the presence of 10 nM (*green curve*) and 2 mM Alba1 (*magenta curve*). At 10 nM Alba1, the FD curve shows several peaks in the force before the molecule is fully extended, indicating that the DNA molecule is bridged (cartoon inset). At a concentration of 2 mM Alba1 the DNA molecule is stiffened (*magenta curve*). **C.** The persistence length (L_P) measured at different protein concentrations of Alba1, Alba1 F60A and Alba1:Alba2 heterodimers (*black circles, green squares and magenta triangles respectively*). The persistence length is obtained by fitting FD curves with the eWLC model (Equation (4.1)). The value for naked DNA was found to be Lp = 49 ± 2 nm. Error bars represent standard error of the mean **D.** The fractional coverage of Alba1, Alba1 F60A and Alba1:Alba2 heterodimers as a function of the concentration (*black circles, green squares and magenta triangles respectively*). The Alba1 data points are fitted by the cooperative binding McGhee-von Hippel model (K = (4 ± 1) · 102 nM-1, ω = (3 ± 1) · 10^2, *solid*

line). The model with similar parameters but $\omega = 1$ (*dashed line*), describes the data points of Alba1 F60A and the Alba1:Alba2 heterodimers well, confirming the lack of cooperativity. Error bars represent the standard error of the mean.

The functional and structural consequences of Alba1:Alba2 heterodimer formation are best quantitatively addressed in the absence of Alba1 and Alba2 homodimers. When mixed at a 1:1 ratio – as in our experiments – heterodimers only are formed (76). We also assessed the influence of Alba1:Alba2 heterodimers within a context of abundant Alba1 homodimers (at a ratio of 1:20), which can be encountered in an *in vivo* situation (76). At this ratio the small amount of Alba2 had no detectable effect on the cooperative formation of the Alba1 filament (*data not shown*).

Dynamics and structure of Alba induced DNA bridging

Alba1 induced intersegmental bridges at relatively low protein concentrations (1 – 100 nM) (Figure 4.2 B, *green curve*). However, at saturating Alba1 concentrations (> 100 nM), bridging was not observed. Intersegmental interactions might be reduced due to stiffening of the DNA molecule, thereby decreasing the chance of two DNA segments within the same molecule to come close enough to form bridges. To test if this is the cause of the absence of bridges at high Alba concentrations, we probed interactions between two individual DNA molecules using a quadruple optical trap configuration. We incubated the DNA molecules in 1 mM Alba1 to fully saturate them and next brought the two Alba1-DNA into close proximity to allow bridge formation. No bridging events were observed showing that DNA coated with Alba1 is not able to interact with other Alba1-DNA filaments. This indicates that the absence of bridge formation is not caused by DNA stiffening, but instead the interaction domains required for bridging are not accessible when proteins have side-by-side interaction with adjacent proteins.

Since Alba1 F60A is capable of forming bridged protein-DNA complexes over a large range of concentrations, the mutant is an ideal candidate to

study Alba induced bridges. Bridging dynamics and structure were investigated using two DNA molecules brought in close proximity that were subsequently incubated with Alba1 F60A (100 nM). After incubation, a force is applied to one end of a DNA molecule, exerting a uniform shearing force along all protein-mediated bridges (Figure 4.3 A, *inset*). Typically the force on the opposite DNA molecule rises, indicating the existence of protein mediated bridges (Figure 4.3 A, *black curve*). Furthermore, sharp drops in the force can be registered as the bridges rupture. However, on most occasions the DNA bridges were able to resist shear forces of at least 400 pN (see Figure S4.2). It is interesting to note that Alba1 bridges are very stable compared to the bacterial DNA bridging protein H-NS which was only able to withstand ~ 25 pN in shearing mode with equal pulling rate (201).

To generate a force exclusively on the first protein-induced bridge, we performed unzipping experiments in which we pull perpendicular to the protein induced DNA bridges (Figure 4.3 B, *inset*). Forces up to 50 pN were required for bridge rupture. When a bridge is ruptured, the total DNA length between the two beads (L) increases, leading to a decrease in the measured force. The experimental data show a clear stepwise increase of L, indicating rupture events of protein-mediated DNA bridges (Figure 4.3 B, *black curve*). The amount of DNA that is released with the rupture of each protein-mediated bridge corresponds to the effective footprint of each bridge. Since bridging the two DNA strands requires the faces of the helices to be aligned, the release in DNA is likely to correspond to an integer number of helical repeats (~ 10.5 bp). To extract the footprint of bridged patches with each rupture event, the histogram of the length (L) of each time trace is plotted and fitted with multiple Gaussians (Figure 4.3 B, *right panel*). The distance between those peaks corresponds to the step size and is plotted in a histogram, which exhibits a distinct peak around ~ 20 bp (see Figure S4.3), corresponding to a distance of two helical repeats.

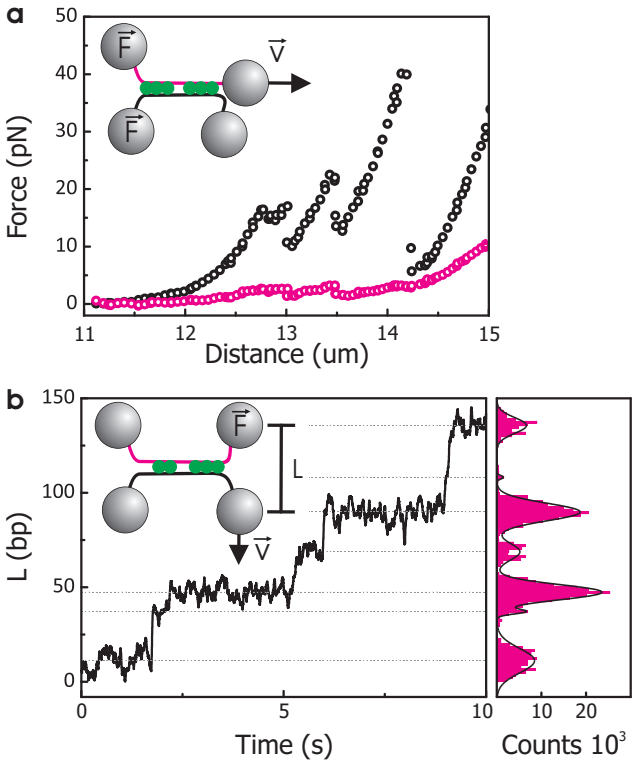


Figure 4.3 Characterization of Alba mediated bridges. Two DNA molecules are crossed and bridged by incubating in 100 nM Alba1 F60A. **A.** Shearing experiment. One of the two DNA molecules is stretched with speed v (*magenta strand*), creating a uniform shearing force (F) over all the formed bridges. The force rises on the opposite molecule (*black data points*) and relaxes as protein-DNA complexes are ruptured. When the magenta DNA molecule is stretched to ~ 14.5 mm, the force also rises on the stretched molecule, possibly caused by bridges within the same molecule. **B.** Unzipping experiment. A force up to 50 pN is built up over the first bridge between the two DNA molecules, by pulling perpendicular to the DNA molecules with speed v. As protein-DNA complexes are ruptured over time the length of DNA between the two beads (L) becomes larger in discrete steps. The histogram of the time trace is fitted with multiple Gaussians to find the step sizes (*right panel*), which represent the distance between two adjacent bridges.

This indicates that two adjacent bridges are less closely packed than found in the co-crystal structure, were the distance between two bridges is found to be a single helical repeat (80).

DISCUSSION

Our experiments have shown that Alba1 homodimers and Alba1:Alba2 heterodimers have distinct DNA binding properties due to their difference in dimer-dimer interactions. McGhee-von Hippel analysis shows that Alba1 binds cooperatively along single DNA molecules, as suggested by previous data (81,167,271). This cooperativity depends on the F60 residue, which is not conserved in the Alba2 paralogue. The fact that Alba1 F60A exhibits identical non-cooperative behaviour as the heterodimers (Figure 4.2 D) underlines the crucial role of the F60 residue in cooperative side-by-side binding.

Besides playing an important role in dimer-dimer stacking, the ability of Alba to bridge two DNA duplexes relies on a dimer-dimer interaction domain, which includes the F60 residue (80). Apparently the dimer-dimer interface is responsible for two distinct structural effects on the DNA: bridging two DNA duplexes and stiffening by cooperative side-by-side binding. How can these two different binding modes be understood in one comprehensible model?

A simple and straightforward generic model that would describe this data assumes that bridges can only be formed at low protein densities when free DNA tracts are available for bridging. At higher protein densities no bridges can be formed. These types of effects are even more pronounced when protein binding is cooperative such that relatively low protein concentrations yield fully covered DNA (filaments). This model has been put forward to explain the DNA binding properties of the H-NS like proteins StpA and MvaT (202,273). H-NS like proteins can bind side-by-side along a DNA molecule leading to DNA stiffening (202,273). In the case of

MvaT it was shown that failure to form nucleoprotein filaments due to disturbed dimer-dimer interactions indeed enhances bridge formation (273). In the case of these proteins no structural information of the proteins bound along DNA (stiffening) or between two DNA molecules (bridging) is available. H-NS and StpA differ from MvaT and the Alba proteins in their functional layout. H-NS and StpA have DNA binding and multimerization domains separated. This is not the case for MvaT, where a point mutation can disrupt multimerization without affecting dimerization and DNA binding (273). It is also not the case for Alba proteins where the same interface is involved in promoting side-by-side interactions along DNA and bridging interactions between DNA molecules (80,271).

Therefore we propose a similar but more detailed model based on our experimental findings and the recent insights into the molecular basis of the dimer-dimer interactions. In this model Alba dimers can bind and alternate between two different orientations on the DNA, facilitating either bridging of two DNA duplexes or side-by-side binding of dimers (Figure 4.4 A). The cooperative side-by-side binding of Alba stiffens the DNA and consequently prevents bridges to form, due to the orientation of the dimer-dimer interaction domain in the direction of adjacent dimers along the DNA. This also explains the fact that two DNA molecules saturated with Alba1 proteins do not show any interaction when brought in close proximity. Since Alba1:Alba2 dimers lack the conserved F60 residue at one side of the dimer, the dimer-dimer interactions along a single DNA duplex are limited to an interaction between two dimers (Figure 4.4 B) and consequently a long patch of closely packed dimers cannot be formed. Heterodimers are likely to be able to form bridges in three different configurations depending on the orientation of the monomeric units within the bridged complex (Figure 4.4 B). The Alba1 F60A mutant confirms the important role of the dimer-dimer interactions on the DNA binding Alba1 dimers, Alba1 F60A dimers and Alba1:Alba2 heterodimers are all able to form bridges at low protein:DNA ratios, which indicates that the F60

residue in the conserved dimer-dimer interface of Alba1 is not crucial for the formation of bridge structures.

Indeed, the co-crystal structure of an Alba1 homologue from *Aeropyrum pernix* K1 (80) shows that the dimer-dimer interactions rely on hydrophobic interactions between the two α1-helices, which are additionally stabilized by the stacking interaction between the F60 residues (see Figure 4.4 D). The Alba1 F60A mutant and the Alba1:Alba2 heterodimer are also able to form bridges, since the ability to bridge is warranted by the interactions of the two α1-helices (see Figure S4.4). Small differences in behaviour of the two types of dimers may be attributed to sequence differences between the α1-helices (only 36% sequence identity) leading to local variations in hydrophobicity along the helix that likely affect the strength of bridging.

At relatively high protein concentrations, cooperative side-by-side binding of Alba1 yields stiff protein-DNA complexes. AFM images show that this results in stiff and open DNA molecules without any intersegmental bridges (Figure 4.1 E). At similar protein:DNA ratios Alba1 F60A and Alba1:Alba2 heterodimers form condensed protein-DNA complexes by multiple bridged patches. In addition, McGhee-von Hippel analysis on the optical tweezers data demonstrated conclusively that the different behaviour of the Alba1:Alba2 heterodimers and the Alba1 F60A mutant is caused by a lack of cooperativity. These observations suggest that the Alba1 F60A mutant and the Alba1:Alba2 heterodimers cannot be packed closely along the DNA to form fully-coated stiff filaments, due to the absence of the F60 residue in both units of Alba1 F60A dimers and in the Alba2 subunit of Alba1:Alba2 heterodimers.

The ability of Alba to bridge two DNA duplexes suggests an important role in shaping archaeal chromatin structure. Architectural proteins that bridge DNA allow the formation of loops, which can functionally organize the genome (112). In *Escherichia coli*, the genome is organized into topologically isolated domains (274,275) and the formation of these domains has been

attributed to bridging/looping by H-NS (113). Whether the archaeal genome is also organized in loops is to date unknown. If such a higher-order structure indeed exists, it could be facilitated by Alba induced bridging.

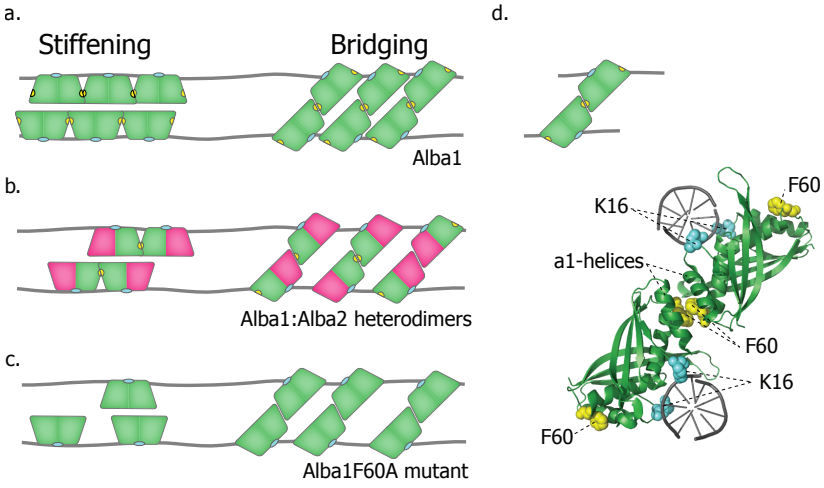


Figure 4.4 Model of Alba mediated stiffening and bridging. **A.** Alba1 dimers can bind to the DNA in two different modes depending on their orientation on the DNA (*cyan depicts DNA-interaction domains*). Dimers can either bind side-by-side along the DNA, stiffening the DNA, or interact with dimers on an adjacent duplex, yielding a bridged structure. The F60 residue (*yellow*) stabilizes dimer-dimer interactions and is responsible for cooperative side-by-side binding. **B.** Alba1:Alba2 heterodimers can bind along the DNA but cooperative side-by-side binding is limited to a maximum of two dimers since only the Alba1 subunit contains the F60 residue. Bridged structures can be formed in three different configurations, depending on the orientation of the dimer in the bridged complex. **C.** The Alba1 F60A mutant does not exhibit cooperative side-by-side binding along the DNA. Bridges can be formed, but are less stable compared to the Alba1 bridges. **D.** Model of Alba1-DNA bridging interaction based on the Alba1 crystal structure (PDB 1H0X) structurally aligned with the co-crystal structure of (Ape10b2)-dsDNA (PDB 2H9U (80)) using PyMol. Alba dimers interact with the minor groove of DNA (*DNA-interacting K16 residues are shown in cyan*). By hydrophobic interactions of the two α-helices and interactions of the F60 residues (*yellow*) two DNA duplexes are bridged properties of Alba, as it shows behaviour similar to that of the Alba1:Alba2 heterodimers (see Figure S4.4 C).

The dual binding mode of Alba can play an important role in shaping chromatin structure *in vivo* by tuning the balance between stiffened and bridged DNA. Such a dual binding mode has been reported for the bacterial chromatin protein H-NS (attributed to *cis* or *trans* binding of dimers (116)) and been suggested to be important for gene regulation (115). Although the molecular details of the binding mode of H-NS are different compared to Alba (the *cis* or *trans* binding is based on the binding mode of dimers instead of tetramers), the functional mechanisms could be very similar. However, whether Alba acts as a global gene regulator like H-NS (166) remains to be investigated. Since Alba2 is expressed only at a few percents of Alba1, the majority of Alba *in vivo* will be in the form of Alba1 homodimers. Depending on the expression level of Alba1 and the resulting (local) Alba1 concentrations, dimers are able to form either stiffened regions of DNA or bridged duplexes. In addition, differential expression of Alba1 and Alba2 can tune the relative amount of Alba1:Alba2 heterodimers and interfere with the cooperative side-by-side binding of Alba1, making the stiffened regions of DNA more accessible. The interplay between Alba homodimers and heterodimers may thus have an important regulatory function in modulating gene accessibility in addition to its role in genome organization.

MATERIALS AND METHODS

DNA and proteins

The pRD24 plasmid (a pUC19 derivative, containing a Bpu10I recognition site inserted into the multiple cloning site) used for the AFM experiments was propagated in *E. coli* strain XL10 and purified (Qiagen plasmid midi kit). pRD24 was nicked by digestion with Bpu10I (Fermentas).

The DNA used in the optical trapping experiments was made by labeling the 12-nucleotide long 5' overhang of bacteriophage λ-DNA (Roche) with dTTP, dGTP, biotin-14-dATP and biotin-14-dCTP using Klenow DNA

polymerase exo⁻ (Fermentas). The DNA was captured in the optical tweezers instrument between two streptavidin coated beads (1.87 μm in diameter, Kisker) in buffer containing 10 mM Tris-HCl (pH 7.5), 200 mM NaCl, 1 mM EDTA, 3.5 mM DTT.

Purification of the Alba1, Alba1 F60A and Alba2 proteins was carried out as described (71,76,271). To obtain Alba1:Alba2 heterodimers the two proteins were mixed in a 1:1 ratio (76). Since Alba2 forms obligate heterodimers with Alba1 (76) this mixture will consist of Alba1:Alba2 heterodimers exclusively.

Atomic Force Microscopy

Alba-DNA complexes were formed by incubating 90 ng of nicked pRD24 with varying amounts of Alba proteins in 10 ml AFM buffer containing 40 mM Hepes (pH 7.5) and 25 mM NaCl for 30 min at room temperature (~ 23 °C). After incubation, this mixture was diluted 10-fold in 1 mM $MgCl_2$ and directly deposited onto freshly cleaved mica. After 40 s, the mica disc was rinsed with HPLC water and dried with nitrogen gas. The AFM images were collected on a NanoScopeIII AFM (Digital Instruments, Santa Barbara, CA) using micro cantilevers (Olympus MCL-AC240TS-W2, resonance frequency 70 kHz, spring constant 2 N/m). Images were obtained using tapping mode in air at 2 Hz and flattened using Nanoscope software (Veeco Instruments).

The protein-DNA complexes were classified by visual inspection of the AFM images. Molecules with single DNA crossings or no crossings (e.g. Figure 4.1 A) were classified as 'open'. Molecules with bridged patches, of which the length is longer than a single crossover of two DNA duplexes, were classified as 'bridged' (e.g. Figure 4.1 G). Molecules containing patches of more than two duplexes bridged (e.g. Figure 4.1 D) or highly compacted regions (e.g. Figure 4.1 J) were classified as 'condensed'.

Optical Tweezers

In order to perform DNA micromanipulation experiments, an instrument was designed and built, capable of manipulating four optical traps in 3 dimensions (Figure 4.2 A). Instead of time sharing three traps using an AOM (276), we chose to use continuous traps (20 W, λ = 1070 nm YLR-20-LP IPG lasers GmbH) and rely on accurate piezo scanning mirrors rather than on motorized lenses (277). The four optical traps enable manipulation and measurements on both a single DNA molecule (dual optical tweezers) as well as two DNA molecules (quadruple optical tweezers). The forces on the DNA molecules were measured by tracking bead displacement, using back focal plane interferometry on the trapping laser and using a second detection laser (140 mW, λ = 980 nm Power Technology Inc.) when using the quadruple optical trap.

The measurements were carried out using a multi-channel laminar flow cell that permits rapid capture, testing and measurement of individual DNA molecules (Figure 4.2 A). After two individual DNA molecules have been caught, they were moved to a channel containing Alba in 10 mM Tris-HCl (pH 7.9), 1 mM EDTA, 3.5 mM DTT, 60 mM KCl, 10 mM $MgCl_2$. In control experiments (bandshift assays) DNA binding behaviour of Alba proteins did not differ in the absence or presence of 10 mM $MgCl_2$ (*data not shown*). The DNA molecules were incubated with Alba without applying tension for ~ 2 min. Force-distance (FD) curves were recorded in ~ 30 seconds in the absence of flow. The resulting FD curves were highly reproducible under these conditions. The procedure of DNA catching and wrapping, switching buffers, incubation and force distance measurement are all automated using a custom-written Labview software suite (National Instruments). All experiments were performed at room temperature (~ 23 °C).

To quantify the properties of the measured DNA molecules and protein-DNA complexes the FD curves are fitted to the extensible worm-like chain model (eWLC) (278):

$$d(F) = L_0\left[1 - \frac{1}{2}\left(\frac{k_B T}{F L_p}\right)^{1/2} + \frac{F}{K_0}\right] \tag{4.1}$$

This model gives a quantitative description of the mechanical properties of DNA in terms of the end-to-end distance (d) and its response to an applied force (F), which is dependent on the total length (contour length L_0), the flexibility (persistence length L_p) and the spring constant (stretching modulus K_0).

To test if the proteins have an effect on the twist-force response of the DNA, force-distance curves were also fitted to the twistable worm-like chain model (tWLC), which describes how much the DNA unwinds as a function of the tension (272):

$$d = L_0\left(1 - \frac{1}{2}\sqrt{\frac{k_B T}{F L_p} + \frac{C}{-g(F)^2 + K_0 C}\cdot F}\right) \tag{4.2}$$

where C represents the twist rigidity and g(F) the force dependent twist-stretch coupling defined by parameters g_0, g_1 and F_C:

$$g(F) = \begin{cases} g_0 + g_1 F_C & F < F_C \\ g_0 + g_1 F & F \geq F_C \end{cases} \tag{4.3}$$

In order to measure FD curves in experiments where bridging occurs, we fully extended the DNA molecule to remove these bridges and recorded the FD curve by decreasing the extension of the DNA molecule. If bridges within a molecule were too stable to be removed this molecule was discarded.

Quantifying DNA coverage and binding kinetics

At relatively low protein concentrations a fraction (υ) of DNA is covered by protein and the complementary fraction (1-υ) consists of naked DNA. Assuming that each protein has an equal contribution to the increase of the DNA stiffness, the measured FD curve will be a linear combination of a FD curve of a saturated DNA molecule ($d_{saturated}(F)$) and a FD curve of a bare DNA molecule ($d_{bare}(F)$).

$$d_{measured}(F) = \upsilon \cdot d_{saturated}(F) + (1-\upsilon) \cdot d_{bare}(F) \tag{4.4}$$

An expression for the fractional coverage as a function of the persistence length is obtained by inserting the eWLC model (Equation (4.1)), given that all parameters (L_0, K_0 and T) are constant for all measured concentrations except for the persistence length (L_p). This expression can be obtained for any other parameter as long as all other parameters are unaffected by protein binding (268).

$$\upsilon = \frac{d_{measured}(F) - d_{naked}(F)}{d_{saturated}(F) - d_{naked}(F)} = \frac{\sqrt{1/L_{p,measured}} - \sqrt{1/L_{p,measured}}}{\sqrt{1/L_{p,saturated}} - \sqrt{1/L_{p,naked}}} \tag{4.5}$$

The fractional coverage (υ) as a function of protein concentration (c) is fitted to the McGhee- von Hippel theory (247), in order to obtain the kinetics on the protein-DNA interactions. In this theory, protein binding to DNA is described by the association constant (K), a cooperativity parameter (ω): the attractive or repulsive interactions between proteins binding to the DNA, and the effective binding site per protein in base pairs (n):

$$\frac{\upsilon}{c} = K \cdot (1-n\upsilon) \cdot \left(\frac{(2\omega+1)(1-n\upsilon) + \upsilon - R}{2(\omega-1)(1-n\upsilon)} \right)^{n-1} \cdot \left(\frac{1-(n+1)\upsilon + R}{2(1-n\upsilon)} \right)^{2} \tag{4.6}$$

where

$$R = \sqrt{(1-(n+1)\upsilon)^2 + 4\omega\upsilon(1-n\upsilon)} \tag{4.7}$$

SUPPLEMENTARY INFORMATION

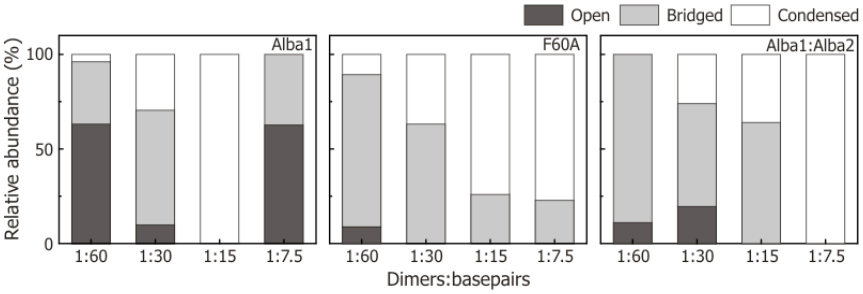


Figure S4.1 AFM results. AFM images are analyzed and divided into 3 classes, open (*dark grey*), bridged (*light grey*) and condensed (*white*) (see Materials and Methods for classification method). Shown are the distributions for different concentrations of Alba1 (N = 548 molecules), Alba1 F60A (N = 337 molecules) and the Alba1:Alba2 heterodimers (N = 134 molecules).

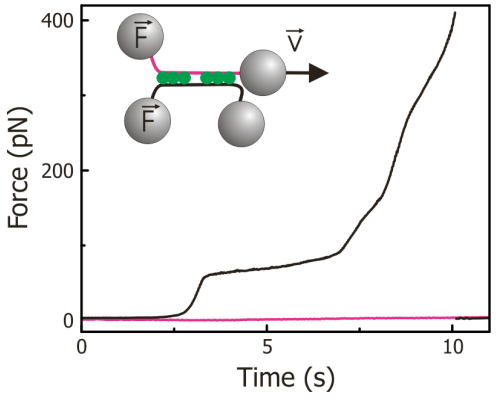


Figure S4.2 Alba bridges are able to withstand high forces. Shown is the force on two DNA molecules bridged by Alba over time, while one of the four beads is moved at a constant speed (v), generating a force on the Alba-DNA complex (*black trace*). As can be seen the forces rise to 400 pN before the bead escapes from the optical trap, relaxing the force.

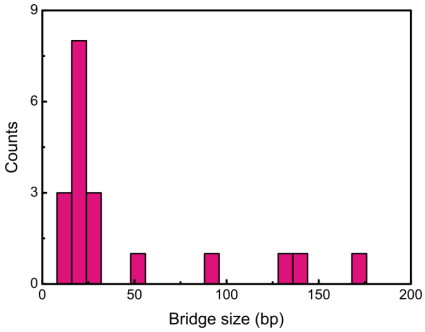


Figure S4.3 Step size distribution of Alba1 F60A bridge ruptures. Histogram of the amount of base pairs released at a rupture event of a bridge in unzipping experiments with the Alba1 F60A mutant (100 nM, see Figure 4.3 B). The distribution shows a peak at ~ 20 bp (N = 19 events, bin size 8 bp).

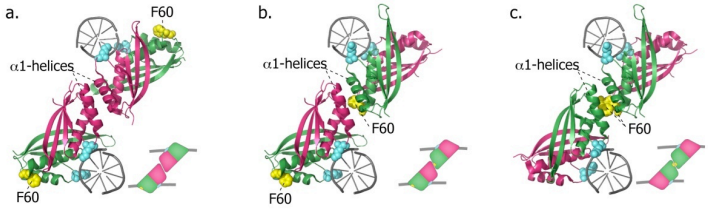


Figure S4.4 Alba1:Alba2 heterodimer bridging interaction. Possible configurations of an Alba1:Alba2 heterodimer bridge as shown in Figure 4.4 B. **A.** The Alba2 subunits (*magenta*) of the heterodimers can interact by hydrophobic interactions of the α1-helices. **B.** The Alba1 subunit (*green*) can interact with the Alba2 unit (*magenta*) of the opposite dimer. In this configuration the F60 residue on the Alba1 subunit (*yellow*) cannot interact with the Alba2 subunit, due to the absence of an equivalent residue in the Alba2 monomer. **C.** The Alba1 units (*green*) can bridge similar to Alba1 homodimers, using both the hydrophobic interactions of the α1-helices and the additional interaction of the two F60 residues. Cyan shows the K16 residue of the Alba1 unit and the K11 residue of the Alba2 unit, which interact with the minor groove of the DNA. The models are obtained by structural alignment of the Alba1:Alba2 crystal structure (PDB 2BKY (76)) with the co-crystal structure of (Ape10b2)-dsDNA (PDB 2H9U (80)) using PyMol.

Table S4.1 eWLC and tWLC fitting parameters. Extensible worm-like chain and twistable worm-like chain fitting parameters as obtained from fitting FD curves to Equation (4.1). Shown for bare DNA (-) and highest measured protein concentrations of Alba1 (2 mM), Alba1 F60A (5 mM), and the Alba1:Alba2 heterodimers (1 mM). Values are averages of N individual DNA molecules and indicated error is given as the standard error of the mean.

Protein	Model	c (mM)	L_p (nm)	L_0 (mm)	K_0 (pN)	g_0 (pN · nm)	g_1 (nm)	F_c (pN)	N
-	eWLC	0	49 ± 2	15.8 ± 0.4	1600 ± 300				27
Alba1	eWLC	2	260 ± 30	16.2 ± 0.2	2400 ± 600				21
Alba1 F60A	eWLC	5	140 ± 20	16.0 ± 0.1	2200 ± 200				15
Alba1:Alba2	eWLC	1	90 ± 20	16.4 ± 0.1	1800 ± 200				8
Alba1	tWLC	2	240 ± 30	16.3 ± 0.1	2100 ± 300	560 ± 100	19 ± 6	31 ± 7	10

Chapter 5

Diverse architectural properties of Sso10a proteins: evidence for a role in chromatin compaction and organization

Rosalie P.C. Driessen, Alson L. H. van der Meulen, Geri F. Moolenaar, Nora Goosen, Remus Th. Dame

ABSTRACT

Sso10a is a small DNA-binding protein expressed by the crenarchaeal model organism *Sulfolobus solfataricus*. Based on the structure of Sso10a, which contains a winged helix-turn-helix motif, it is believed that Sso10a proteins function as sequence-specific transcription factors. Here we show that Sso10a1 and Sso10a2 exhibit a variety of DNA-binding modes in which they bend, bridge and/or stiffen the DNA. While the DNA-bending property is shared between the two proteins, bridging is only observed for Sso10a1 and only Sso10a2 exhibits stiffening of the DNA. The architectural properties of Sso10a proteins indicate that these proteins are involved in chromatin compaction and organization. The interplay between the different binding modes could provide an important mechanism to dynamically shape the crenarchaeal genome and regulate gene expression on a global scale.

INTRODUCTION

Chromatin proteins play an important role in compacting and organizing the genomic DNA throughout all domains of life (14). Besides folding the genome into a compact structure, chromatin proteins are involved in important cellular processes such as transcription, replication and repair. Although these proteins are not conserved throughout all domains of life at the level of amino acid sequence, they appear to be functionally conserved in terms of their architectural properties (14). Chromatin proteins in general modulate the spatial path of the DNA by wrapping, bridging or bending the dsDNA helix. Not only do the architectural properties of these proteins explain compaction, they also provide a framework for understanding how different proteins can work together in providing a dynamic/plastic genome. The delicate interplay between architectural proteins (e.g. related to their differential expression in different conditions) permits them to modulate the degree of local/global compaction and thereby regulate the accessibility of the genomic DNA for different DNA-based processes.

Archaeal species express a wide variety of chromatin proteins that organize their genomic DNA into a compact nucleoid structure. These proteins are hence often referred to as nucleoid-associated proteins. While species from the euryarchaeal branch generally express true histone homologues that form nucleosome-like structures (44), the majority of crenarchaea lack histone homologues and only synthesize other small chromatin proteins (40) (see Figure 1.3), analogous to bacteria. Several chromatin proteins have been identified and characterized in the crenarchaeal model organism *S. solfataricus:* Cren7 (63), Sul7 (61,62) and Alba (60). Cren7 and Sul7 are monomeric proteins that are structural and functional homologues, despite a lack of sequence conservation (95-97,101). Both proteins induce rigid bends in dsDNA, resulting in compaction (90). Alba proteins exist as dimers and exhibit a multimodal behaviour of DNA bridging and DNA stiffening, depending on protein concentration and dimer composition (178). The balance between the two binding modes, bridging and stiffening,

is believed to be an important mechanism in dynamic chromatin organization and gene regulation *in vivo*.

Another small (10 kDa) basic DNA-binding protein, Sso10a, has initially been identified in *S. acidocaldarius* as Sac10a (59). The structure of Sso10a reveals that it forms dimers in solution as a result of dimerization via an anti-parallel coiled-coil structure (64,65,109) (see Figure 5.1). The domain at the end of the coiled-coil is believed to be involved in DNA-binding as it adopts a winged helix-turn-helix (wHTH) motif. This is a subfamily of the helix-turn-helix (HTH) motif that is common in DNA-binding proteins from all domains of life (279,280). In archaea, the helix-turn-helix motif is often found in transcription regulators that bind to specific sequences (129), which suggests a similar role for Sso10a proteins. The wHTH motif generally consists of three α-helices (α1-α3) and a β-sheet wing. The α3-helix is often referred to as the recognition helix as it binds DNA by inserting in the major groove and is thought to be important for sequence specific binding. In many cases the β-sheet wing makes additional protein-DNA contacts by binding along the phosphate backbone (281).

Characterization of Sac10a-DNA complexes using electron microscopy demonstrated that Sac10a brings two DNA duplexes together (81), suggesting that it could play a role chromatin organization and compaction. Little is known about the DNA-binding properties of Sso10a, besides that Sac10a binds preferentially to poly(dAdT) dsDNA (compared to poly(dA)-poly(dT) and DNA isolated from *E. coli*) with high affinity in the nanomolar range (109). Based on analogy with another HTH DNA-binding protein (γδ-resolvase), a DNA binding model has been proposed in which Sso10a binds along the DNA via the two α3-helices that are inserted in the major groove. A possible interaction could exist between the minor groove with the wing and the N-terminus (64). According to this model, binding of Sso10a induces bending of the DNA by ~ 60°. Three Sso10a homologues have been identified in *S. solfataricus*: Sso10a1, Sso10a2 and Sso10a3 (29 – 39 % sequence identity and 55 – 62 % sequence similarity, see Figure 5.2).

Although nothing is known about the abundance of these proteins within the cell, the transcriptome of *S. solfataricus* (110) reveals that the transcription level of the three Sso10a proteins together is similar to that of Cren7 (constituting ~ 1% of the total cellular protein (63)), suggesting that Sso10a is abundant within the cell. A noticeable difference in the amino acid sequence of the three Sso10a proteins is that Sso10a2 contains four extra residues within the winged loop domain in comparison to Sso10a1 and Sso10a3, which are more similar based on the amino acid sequence. The recently resolved structure of Sso10a2 (282) demonstrates that these extra amino acids lead to an extended loop within the wing domain (see Figure 5.1 B). Structural comparison of the Sso10a1 and Sso10a2 structures also revealed a difference in the orientation angle between the α3-helices (see Figure 5.1 C). As the relative position and orientation of the two α3-helices is predicted to have an important influence on the DNA-binding mode (283), this could result in distinctly different DNA-binding properties for Sso10a1 and Sso10a2. (284)

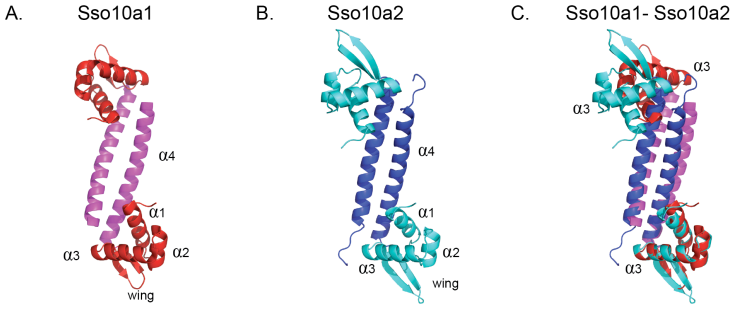


Figure 5.1 Crystal structures of **A.** Sso10a1 (PDB code 1RJ7 (64)) and **B.** Sso10a2 (PDB code 4WH0 (282)). The Sso10a dimer consists of an anti-parallel coiled-coil (*magenta/blue*) with two DNA-binding winged helix-turn-helix motifs (*red/cyan*) on opposite sides of the coiled-coil. The main structural differences between Sso10a1 and Sso10a2 are the length of the β-sheet wing and the relative position of the α3-helices, which are thought to be important for DNA interactions. **C.** Sso10a1 and Sso10a2 aligned on the lower α3-helix, illustrating the difference in relative position and orientation of the two α3-helices.

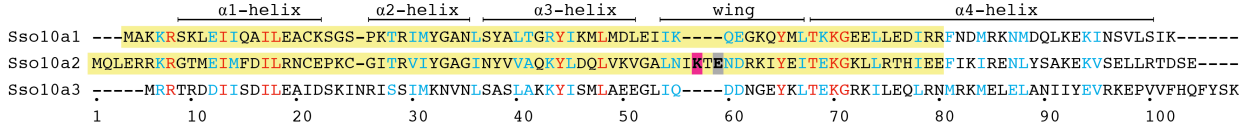


Figure 5.2 Amino acid sequence alignment of Sso10a1 (SSO10449), Sso10a2 (SSO2827) and Sso10a3 (SSO2986). Sequence alignment was performed using the Clustal Omega program (284). Secondary structure elements are indicated based on the Sso10a1 structure. Identical residues are indicated in red and conserved residues are indicated in cyan. The DNA binding domains Sso10a1_DBD and Sso10a2_DBD are highlighted in yellow. Residues K57 and E59 of Sso10a2 are highlighted in magenta and grey respectively.

From the currently available data it is unclear whether Sso10a proteins primarily acts as transcription regulators involved in expression of specific genes or rather acts as chromatin proteins involved in DNA compaction and organization of the archaeal genome. To investigate the function of the Sso10a proteins we have examined the architectural properties of Sso10a1 and Sso10a2 proteins using a set of different single-molecule techniques. We characterized the conformational changes of DNA due to binding of Sso10a1 and Sso10a2 to determine whether these proteins yield different types of protein-DNA complexes. Our data indicate that Sso10a1 and Sso10a2 act as chromatin proteins by employing a set of different DNA-binding modes in which they bend, bridge and/or stiffen the DNA without apparent sequence specificity.

RESULTS

Characterization of DNA compaction by Sso10a1 and Sso10a2

To characterize the architectural properties of Sso10a1 and Sso10a2 we used tethered particle motion (TPM) assays, which enable observation of structural changes in the DNA at a single-molecule level. In this assay, the excursion of a bead tethered to a DNA molecule is expressed in terms of a root mean square distance (RMS). The RMS distance is affected by the conformational state of the DNA molecule, which is found to be 161.8 ± 0.7 nm (error represents standard error of the mean) for bare DNA in our experimental conditions. In these buffer conditions Sso10a1 and Sso10a2 bind DNA with similar affinity ($K_D \approx 1$ nM) in bandshift assays (*data not shown*). Addition of either Sso10a1 or Sso10a2 reduces the RMS in a concentration dependent manner at concentrations ≤ 100 nM (see Figure 5.3) up to a minimum of 115.8 ± 0.6 nm for Sso10a1 and 120.0 ± 0.8 nm for Sso10a2 at a concentration of 50 nM. The reduction of RMS indicates that the DNA is compacted due to protein binding of either Sso10a1 or Sso10a2 to a similar extent, which can be attributed to bending and/or bridging of the DNA.

Binding of Sso10a2 however leads to an increase in RMS for concentrations ≥ 1000 nM up to a maximum of 173.0 ± 3.7 nm at 3000 nM, which is higher than the RMS value found for bare DNA. The increase of RMS at high concentrations of Sso10a2 suggests that, besides bending and/or bridging, Sso10a2 has a second binding mode of DNA stiffening. Interestingly, this mode of binding is salt sensitive. Reduction of the salt concentration (20 mM NaCl instead of 100 mM NaCl) shifts the concentration regime of this binding mode, while it does not affect the intrinsic binding affinity, as the onset of compaction occurs at similar protein concentrations (see Figure 5.4). At this salt concentration, the RMS increased at Sso10a2 concentrations ≥ 200 nM and goes up to a maximum value of 185.8 ± 1.8 nm at 2000 nM. The effect of NaCl concentration on the stiffening mode suggests that this binding mode depends on electrostatic interactions.

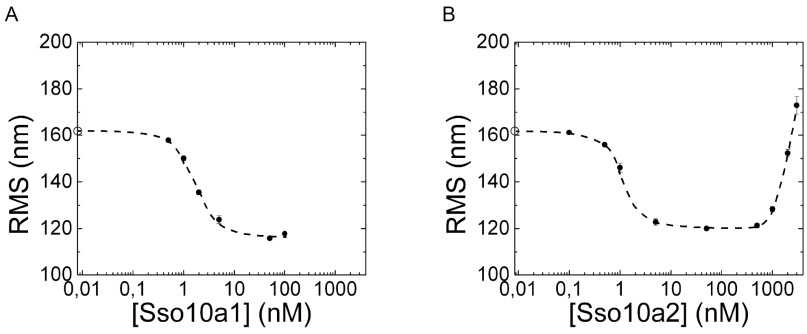


Figure 5.3 Root mean square (RMS) values as a function of protein concentration obtained from tethered particle motion experiments in 100 mM NaCl with **A.** Sso10a1 and **B.** Sso10a2. Both proteins compact DNA (decrease in RMS) and Sso10a2 stiffens DNA at concentration > 1000 nM. Open circle on the y-axis represents the RMS value for bare DNA (0 nM). Error bars represent the standard error of the mean. Dashed lines are a guide to the eye.

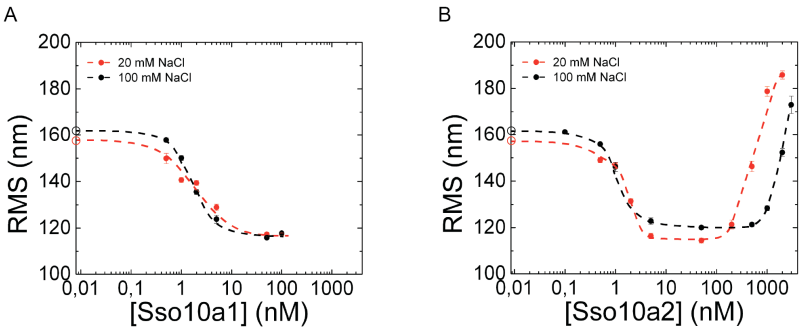


Figure 5.4 Root mean square (RMS) values as a function of protein concentration obtained from tethered particle motion experiments with **A.** Sso10a1 and **B.** Sso10a2 in a buffer containing 20 mM NaCl (*red*) or 100 mM NaCl (*black*). Open circles on the y-axis represent the RMS value for bare DNA (0 nM). Error bars represent the standard error of the mean. Dashed lines are a guide to the eye.

For Sso10a1, concentrations > 100 nM result in all DNA tethers or beads being stuck to the flow cell surface, which makes it impossible to collect data at these concentrations. Possibly, the DNA is compacted so strongly at these concentrations that bead-surface or DNA-surface interactions are promoted, causing the tethers to stick to the surface. Whether Sso10a1 stiffens DNA at these concentrations similar to Sso10a2 can therefore not be determined from these experiments.

To investigate whether the observed compaction is caused by the action of single DNA-binding domains or the joint action of two DNA-binding domains at the ends of the dimer, we have constructed truncated Sso10a proteins, which can no longer dimerize (see Material and Methods). As these proteins (referred to as Sso10a1_DBD and Sso10a2_DBD) only consist of the DNA-binding-domain (see Figure 5.2) they are expected to form monomers in solution. Although there are some small differences in binding affinity, both Sso10a1_DBD and Sso10a2_DBD bind to DNA with a binding affinity in the same order of magnitude as that of the full-length proteins, as could be seen in bandshift assays (*data not shown*).

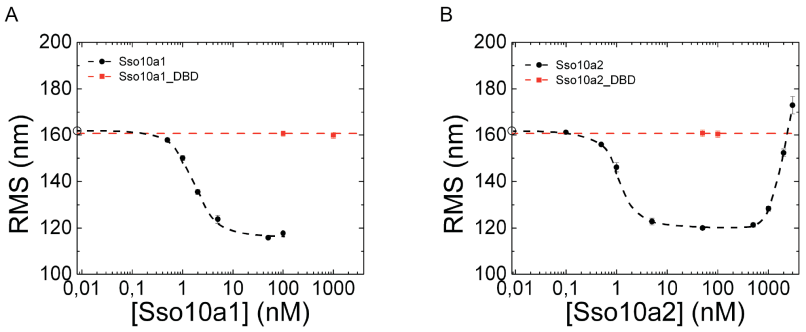


Figure 5.5 Comparison of the full-length Sso10a dimer with the isolated DNA-binding domain obtained from tethered particle motion experiments in 100 mM NaCl with **A.** Sso10a1 (*black*) and Sso10a1_DBD (*red*) and **B.** Sso10a2 (*black*) and Sso10a2_DBD (*red*). Open circles on the y-axis represent the RMS value for bare DNA (0 nM). Error bars represent the standard error of the mean. Dashed lines are a guide to the eye.

However, TPM experiments with Sso10a1_DBD and Sso10a2_DBD do not result in compaction of the DNA, even at protein concentrations two orders of magnitude higher than those that result in maximum compaction for full-length Sso10a1 and Sso10a2 (see Figure 5.5). Instead, the RMS values do not deviate from the RMS measured for bare DNA, suggesting that binding of the DNA-binding domains alone does not change the DNA configuration. Apparently both DNA-binding domains of the Sso10a dimer need to interact with the DNA to result in DNA compaction.

Visualization of protein-DNA complexes

Our TPM experiments showed that Sso10a1 and Sso10a2 behave differently at high protein concentrations. To characterize the different binding modes between Sso10a1 and Sso10a2 in this regime we visualized protein-DNA complexes using atomic force microscopy (AFM). We used a single-nicked circular DNA substrate of 2.7 kilobasepairs (see Materials and Methods). Addition of Sso10a1 (0.5 μM and 1 μM) results in compaction of the DNA molecules (see Figure 5.6). Within individual protein-DNA complexes regions of DNA are aligned, which points to bridging of two DNA duplexes

by the protein. Bending of DNA might also contribute to the observed compaction. However, protein-induced bending cannot be quantitatively analyzed, as the small size of the proteins does not allow the identification of individual proteins on the DNA. Classification of the protein-DNA complexes by visual inspection revealed that the fraction of bridged molecules slightly increases at higher Sso10a1 concentrations from 59% at 0.5 μM to 66% at 1 μM (see Table S5.1 for details). At Sso10a1 concentrations above 1 μM, protein-DNA complexes could not be deposited onto the mica surface. Possibly, high Sso10a1 concentrations lead to very strong compaction, shielding the negative charge of the DNA and preventing protein-DNA complexes to interact with the mica surface. Such high compaction of DNA molecules at high Sso10a1 concentrations could also explain why protein-DNA complexes were stuck on the flow cell surface in our TPM experiments.

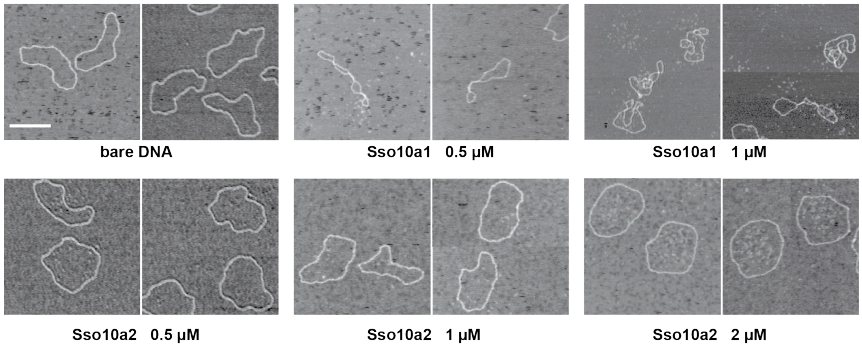


Figure 5.6 Representative images of individual Sso10a-DNA complexes, visualized by atomic force microscopy. Nicked pRD24 plasmids are incubated at different protein concentrations with either Sso10a1 or Sso10a2. Incubation with 0.5 μM and 1 μM Sso10a1 results in compacted protein-DNA complexes. Bridged patches are formed by Sso10a1. Addition of 0.5 μM and 1 μM Sso10a2 does not induce observable changes in DNA configuration. At a concentration of 2 μM Sso10a2, the DNA plasmids obtain a more open conformation, indicating stiffening of the DNA. Scale bar represents 200 nm.

In contrast, addition of Sso10a2 at comparable concentrations yields complexes distinctly different from those observed for Sso10a1. While addition of 0.5 μM and 1 μM Sso10a2 does not induce observable changes in DNA configuration, a protein concentration of 2 μM induces a more open conformation of the plasmids, indicating stiffening of the DNA. These Sso10a2-DNA complexes look very similar to the protein-DNA complexes observed for bacterial chromatin protein HU, which stiffens DNA at high concentrations by cooperative side-by-side binding (285). This observation is consistent with the increased RMS values observed at high Sso10a2 protein concentration in the TPM experiments.

Quantification of the architectural properties of Sso10a1 and Sso10a2

In order to quantitatively dissect the binding modes of the Sso10a proteins, underlying the observed compaction and stiffening, we performed DNA micromanipulation studies. We investigated the mechanical properties of single DNA molecules by recording force distance (FD) curves using optical tweezers. A typical experiment consists of two steps. First a DNA molecule is captured between two beads and a control FD curve is recorded by gradually increasing the distance between the two beads while measuring the force (see Figure 5.7, *black curves*). Hereafter, the molecule is incubated in the absence of force (6 μm between the centers of the two beads) in the presence of protein at the desired concentration for 2 minutes. After incubation of the molecule a new FD curve is recorded to probe the mechanical properties of the protein-DNA complex. Independent of the applied force, the end-to-end distance of the protein-DNA complex is lower than that of the bare DNA (see Figure 5.7, *blue curves*), in agreement with the compaction observed in the TPM experiments. To quantify the structural changes in the DNA, FD curves are fitted to the worm-like-chain (WLC) model, yielding the DNA persistence length (L_p) and DNA contour length (L_0) (see Table S5.2 for WLC fitting parameters). The physical properties of bare DNA obtained from the control FD curves are L_p = 39.7 ±

7.9 nm and L_0 = 16.7 ± 0.1 µm (average values and error represents the standard deviation). At a protein concentration of 100 nM the apparent persistence length is reduced to only a few nanometers for both Sso10a1 and Sso10a2 (L_p = 4.2 ± 2.5 nm for Sso10a1 and L_p = 2.8 ± 1.7 nm for Sso10a2). Moreover, the contour length of the DNA increased by about 10% due to binding of the protein (L_0 = 18.3 ± 0.7 µm for Sso10a1 and L_0 = 18.7 ± 0.6 µm for Sso10a2).

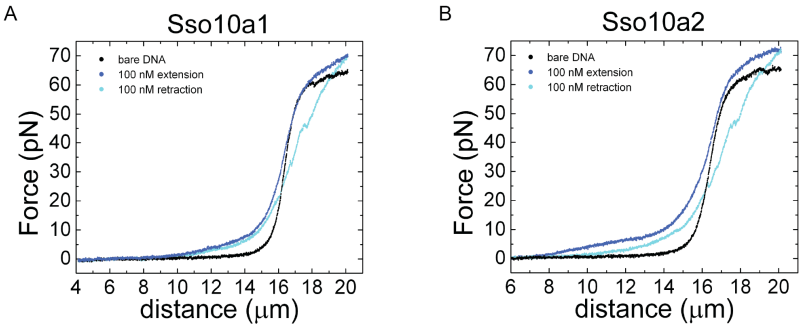


Figure 5.7 Force distance curves of individual DNA molecules incubated with 100 nM Sso10a showing the extension (*blue*) and retraction (*cyan*) of the DNA molecule. The control FD curve of the bare DNA molecule is depicted in black. The average persistence length and contour length of bare DNA are L_p = 39.7 ± 7.9 nm and L_0 = 16.7 ± 0.1 µm (N = 28). Binding of Sso10a1 and Sso10a2 results in a decrease in persistence length and an increase in contour length. **A.** Sso10a1; The persistence length at 100 nM is L_p = 4.2 ± 2.5 nm and the contour length is L_0 = 18.3 ± 0.7 µm (N = 15). **B.** Sso10a2; The persistence length at 100 nM is L_p = 2.8 ± 1.7 nm and the contour length is L_0 = 18.7 ± 0.6 µm (N = 8).

Quantification of the physical properties of the protein-DNA complexes indicates that the compaction observed here, as well as in our TPM experiments, is achieved by bending and slight unwinding of the DNA. In the model proposed by Chen *et al.* (64) the DNA-binding domains of the dimer interact with the major groove of the DNA on a binding site of 25 basepairs, resulting in a bend of about 60 °. This bending angle is of the same order as that of DNA-bending proteins Cren7 and Sul7 (90).

However, Sso10a reduces the persistence length to a much greater extent compared to Cren7 and Sul7 (~ 3 nm for Sso10a compared to ~ 9 nm for Cren7 and Sul7). Moreover, Cren7 and Sul7 have a much smaller binding site (~ 8 bp) (95,101) compared to the 25 bp proposed for Sso10a, which allows more of these proteins to bind and bend the DNA per unit length. According to the proposed DNA bending model one would therefore not expect that Sso10a bending alone would result in such a low apparent persistence length. Indeed, the increase in contour length suggests that partial unwinding of the dsDNA helix occurs, which contributes to a decrease in persistence length.

In a subset of FD curves of 100 nM Sso10a1 (~ 26%) a small peak was observed in the low force regime (< 15 pN). The occurrence of peaks reflects the formation of intersegmental bridges by the protein, which are disrupted by the applied force. Bridges are formed in a concentration dependent manner as the amount of molecules that show bridges increases with the Sso10a1 concentration. Increasing the protein concentration to 1 μM resulted in an increase in the amount of bridged molecules in the low force regime (~ 57%) (see Figure 5.8 A), while L_p and L_0 were similar compared to those of 100 nM. At an even higher protein concentration of 2 μM Sso10a1, all recorded FD curves showed multiple small peaks at distances far below the contour length (see Figure 5.8 B). Most bridges are disrupted at low forces < 15 pN, but in some cases force peaks are observed up to ~ 40 pN (see Figure S5.1).

The length of DNA loops formed by these bridges can be calculated from the relative height of the force peaks at the disruption events and corresponds typically to a few hundred nanometers. The formation of such relatively small loops could be enhanced by the increased DNA flexibility due to DNA-bending by Sso10a1. AFM images of Sso10a1-DNA only showed small patches of bridged DNA. This suggests that DNA bridging by Sso10a1 is not strongly cooperative, as this would lead to large patches of bridged DNA. In contrast, the DNA-bridging protein Alba exhibits DNA-induced

cooperative bridging resulting in long stable bridges (178), which can withstand much higher forces (up to 400 pN) than Sso10a1 (see Figure S4.2). DNA-bridging by Alba was proposed to rely on dimer-dimer interactions (see Figure 4.4), based on the crystal contacts in the Alba-DNA co-crystal structure (80). Whether and how bridging by Sso10a1 depends on dimer-dimer interactions is not evident. The crystal contacts in the Sso10a1 crystal structure (64) do not show potential dimer-dimer interactions that could be identified using the Protein Interfaces, Surfaces and Assemblies (PISA) software (286).

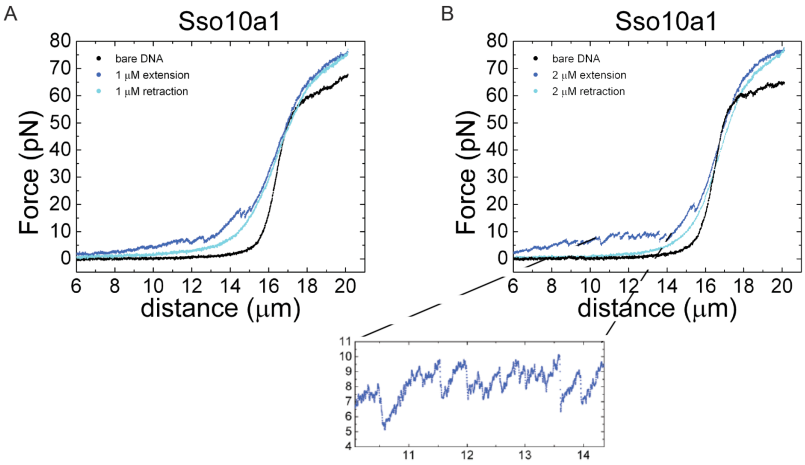


Figure 5.8 Force distance curves of individual DNA molecules incubated with Sso10a1 showing the extension (*blue*) and retraction (*cyan*) of the DNA molecule. The control FD curve of the bare DNA molecule is depicted in black. The force peaks at distance < L_0 indicate intersegmental bridging. **A.** 1 μM Sso10a1 **B.** 2 μM Sso10a1

In contrast to Sso10a1, FD curves of Sso10a2-DNA complexes do not yield any disruption events, suggesting that this protein is not capable of bridging. Instead, the persistence length and contour length are constant at all measured concentrations (0.1 – 3 μM), indicating saturation of the DNA at 100 nM. Both AFM and TPM experiments suggest an increase in persistence length at high Sso10a2 concentrations, but even at 3 μM the

DNA is not stiffened and the persistence length is found to be ~ 3 nm. As TPM and AFM experiments do not apply force on the DNA molecules, in contrast to DNA pulling experiments, this suggests that the DNA stiffening mode is very sensitive to force, possibly due to disruption of protein-protein interactions.

In all FD curves, the overstretching plateau, which is typically ~ 65 pN for bare DNA, was slightly increased upon addition of protein. This suggests that Sso10a proteins are still bound to the DNA at high forces within the DNA overstretching regime and stabilize the dsDNA from force-induced melting. Hysteresis of the retraction curve suggests nevertheless, that a fraction of proteins is dissociated due to the applied force and/or that the protein is differently bound, stabilizing the unwound DNA segments that are formed in the overstretching regime. Control experiments in which Sso10a-DNA complexes are extended up to a smaller maximum distance did not show hysteresis (see Figure S5.2), indicating that the observed hysteresis is associated with structural changes in the DNA within the overstretching regime.

DISCUSSION & CONCLUSIONS

Sso10a1 and Sso10a2 exhibit similar behaviour at low protein concentrations as they compact DNA by changing the physical parameters of the DNA to the same extent. At high protein concentrations however, Sso10a1 and Sso10a2 behave completely different as both proteins reveal a second mode of binding leading to bridging or stiffening of DNA. Based on our observations we propose a model in which we describe the concentration dependent binding modes of Sso10a1 and Sso10a2 (see Figure 5.9).

Sso10a1 and Sso10a2 induce DNA bending

At low protein concentrations, both Sso10a1 and Sso10a2 bend DNA by binding parallel along the DNA (see Figure 5.9). Bending is induced by the

joint action of the DNA-binding domains, as isolated DNA-binding domains do not lead to compaction of DNA molecules. Many wHTH proteins bind DNA by insertion of the α3-helix into the major groove (287,288). This binding permits additional interactions between the wing and the DNA backbone, which stabilize the protein-DNA complex. The joint action of the two DNA-binding domains can result in DNA bending and this bending depends on the relative orientation of the two DNA-binding domains and the distance between them (283). Analogous to these wHTH proteins, Sso10a is likely to bind and bend DNA by interactions between the α3-helix and the wing with the DNA. Alignment of the crystal structures of Sso10a1 and Sso10a2 shows a difference in relative orientation of the DNA-binding domains (i.e. the α3-helix, see Figure 5.1). Based on our findings we expect that the relative orientation of the α3-helix of Sso10a1 and Sso10a2 is rather similar in solution since both proteins bind DNA with the same affinity and bend DNA to the same extent.

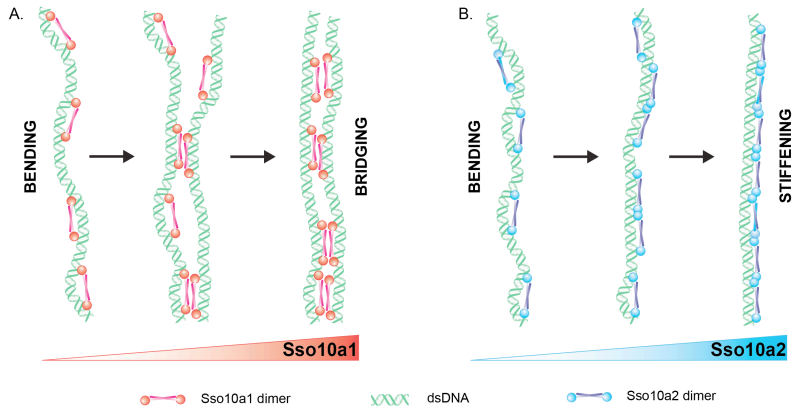


Figure 5.9 Model of the concentration dependent binding modes Sso10a. **A.** At low concentrations, Sso10a1 bends DNA by binding in *cis*. Increased concentrations lead to dimer-dimer interactions, which bridge two DNA segments. **B.** At low concentrations, Sso10a2 bends DNA by binding in *cis*, similar to Sso10a1. At increased protein concentrations adjacent dimers exhibit electrostatic dimer-dimer interactions via the wing, which results in stiffening of the DNA.

Sso10a1 bridges DNA duplexes at high protein concentrations

Upon addition of increased protein concentrations Sso10a1 dimers bridge DNA (see Figure 5.9 A). A plausible model is that bridging is achieved through dimer-dimer interactions via the coiled-coil domains. These domains might be able to form a tetrameric coiled-coil structure, which is a multimeric form that is found for several coiled-coil motifs (289-291). Interactions between dimers that are bound along the DNA can then result in bridged DNA segments (see Figure 5.9 A). In solution multimeric forms of Sso10a1 have however not been observed (64). Possibly such dimer-dimer interactions are only possible between Sso10a1 dimers that are bound to DNA. Interestingly, Sso10a2 does not show DNA bridging at any of the measured concentrations. This could be a consequence of the differences between the amino acid sequence of the α4-helix of Sso10a1 and Sso10a2 (see Figure 5.2), which could favor dimer-dimer interactions via the coiled-coil for Sso10a1 but not for Sso10a2.

Sso10a2 stiffens DNA at high concentrations and zero force

At high Sso10a2 concentrations DNA is stiffened in the absence of force. Stiffening of the DNA could be due to close side-by-side binding of Sso10a2 dimers along the DNA, stabilized by dimer-dimer interactions (see Figure 5.9 B). As the stiffening mode is sensitive for monovalent salt concentration (see Figure 5.4 B), this suggests that the stiffening mode is dependent on electrostatic interactions between adjacent dimers. The wing is a good candidate for interactions between Sso10a2 dimers as the wing is one of the main structural differences between Sso10a1 and Sso10a2. Moreover, the wing is positioned at both ends of the dimer, which would allow interaction between adjacent Sso10a2 dimers, assuming that the dimers bind parallel to the DNA. The Sso10a2 crystal structure indeed reveals possible electrostatic interactions (K57-E59 and E59-K57) between the wings of adjacent dimers (see Figure 5.10) that could underlie the observed salt sensitive stiffening. Interactions between the wings are not present in the crystal structure of

Sso10a1 as it has a different crystal packing and has a smaller wing, lacking these residues (64).

Protein-protein interactions via the wing in a wHTH protein have previously been shown for a eukaryotic heat shock transcription factor (HSF) (292). Interestingly, unlike many other wHTH motif proteins the wing of the DNA-binding domain of HSF is involved in protein-protein interactions rather than interacting with the DNA.

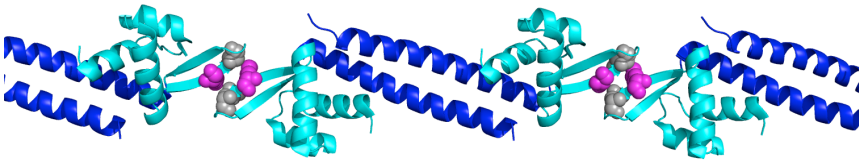


Figure 5.10 Sso10a2 crystal structure showing possible electrostatic interactions between adjacent dimers (PDB 4WH0 (282)). Residues K57 (*magenta*) and E59 (*grey*) within the wing domain of the dimer can form electrostatic dimer-dimer interactions when bound side-by-side along the DNA.

Possibly, the wing of Sso10a2 is capable of mediating both types of interactions and is key to switching between binding modes. At low Sso10a2 concentrations, the DNA occupancy is low and the wing will mostly interact with the DNA, inducing or stabilizing DNA bending. At high concentrations however, the wing can be involved in dimer-dimer interactions with adjacent dimers, which will release the DNA bending and induce DNA stiffening (see Figure 5.9 B). As force disrupts DNA stiffening by Sso10a2, this likely interrupts the wing-wing interactions of adjacent Sso10a2 dimers. Based on this finding it is compelling to speculate that Sso10a2 could act as force-sensor *in vivo*, changing the local DNA configuration in response to forces, which for example arise from processes involved in DNA-based processes such as transcription or replication.

Multiple and different binding modes for Sso10a1 and Sso10a2

Based on our findings we conclude that the Sso10a proteins function as chromatin proteins as they bind DNA with high affinity and no apparent sequence specificity and display different architectural modes of DNA bending, bridging and stiffening. The interplay between the different binding modes of Sso10a provides mechanisms to dynamically shape the chromatin structure and to regulate gene expression *in vivo*.

Such interplay has for example been proposed to be a mechanism by which archaeal chromatin protein Alba can dynamically shape the genomic DNA (178). Alba stiffens and bridges DNA, and the balance between those two binding modes depends on the composition of the Alba dimers (Alba1 homodimers or Alba1:Alba2 heterodimers). Differential expression of Alba1 and Alba2 can thus tune local chromatin structure, which could play an important role in global gene regulation. Analogous to Alba, differential expression of Sso10a homologues could lead to modulation of chromatin conformation and global gene expression. Binding of Sso10a1 could possibly interfere with DNA filaments stiffened by Sso10a2 by disrupting dimer-dimer interactions of adjacent Sso10a2 dimers, making the DNA more accessible for DNA-based processes.

Different architectural modes can also work synergistically. DNA bending can support the formation of small loops by a DNA bridging protein. This way the DNA bending action of Sso10a1 and Sso10a2, which results in an increase in local DNA concentration, can facilitate bridging by Sso10a1 (or another DNA-bridging protein like Alba). Such synergistic action could also occur between Sso10a and other chromatin proteins such as Cren7 and Sul7 that bend DNA and/or Alba that is able to bridge DNA. Genome wide association studies on these architectural proteins could indicate whether and how these proteins co-localize on the crenarchaeal genome and if they form cooperative multi-protein complexes as has been observed in bacteria (164). Altogether, the versatile binding modes of Sso10a provides a

framework to dynamically organize the crenarchaeal chromatin structure and to regulate gene expression on a global scale.

MATERIALS AND METHODS

Protein purification

Sso10a1 and Sso10a2 were overproduced in *E. coli* strain BL21-CodonPlus (DE3) and *E. coli* strain BL21 (DE3) respectively, containing plasmid pRD114 or pRD 111 (pET11a including either the gene encoding Sso10a1 or Sso10a2 (gene SSO10449 and SSO2827, respectively)) from *S. solfataricus*. Cells were grown in LB medium containing 30 µg/ml ampicilin at 37 °C up to $OD_{600} \approx 0.4$ and expression was induced using 0.5 mM IPTG. Two hours after induction, cells were harvested by centrifugation and resuspended in 20 ml buffer A [50 mM Tris–HCl (pH 8.0), 2 mM $MgCl_2$, 0.1% Triton X100, 386 µg/ml benzamidine hydrochloride and 10 mM β-mercaptoethanol]. Cells were lysed by sonication, 1000 units OmniCleave Endonuclease (Westburg) per gram cells were added and the cell lysate was incubated for 30 min at room temperature. After heating the cell lysate for 40 min at 70 °C, EDTA and NaCl were added to a final concentration of 20 mM EDTA and 100 mM NaCl. The cell lysate was centrifuged for 30 min at 37.000 rpm and filtered through a 0.45 µm membrane filter (MiliPore). The supernatant was applied to a HiTrap SP column (GE Healthcare), equilibrated in buffer B [10 mM KPO_4 (pH 7.0), 100 mM NaCl, 10% glycerol and 10 mM β-mercaptoethanol]. Protein was eluted with a linear gradient of 0.1 – 1 M NaCl in buffer B. After dialyzing at 4°C against buffer B the protein was loaded onto a Heparine HP column equilibrated in the same buffer, and eluted with a linear gradient of 0.1 – 1 M NaCl in buffer B. Proteins were dialyzed at 4°C against buffer B containing 200 mM NaCl, and stored at -80 °C until required.

The isolated His-tagged DNA-binding-domains of Sso10a1 and Sso10a2 (Sso10a1_DBD and Sso10a2_DBD) were produced by inserting the part of

the gene that encodes for the amino acids M1-R73 and M1-E80 respectively with an additional His-tag at the C-terminal end into plasmid pET11a (resulting in plasmid pRD122 and pRD119, respectively). Overproduction was performed in *E. coli* strain BL21-CodonPlus (DE3) and *E. coli* strain BL21 respectively. Cells were grown in LB medium containing 30 µg/ml ampicilin at 37 °C up to $OD_{600} \approx 0.4$ and expression was induced using 0.5 mM IPTG. Two hours after induction, cells were harvested by centrifugation and resuspended in 20 ml buffer C [20 mM Tris–HCl (pH 7), 500 mM KCl, 10% glycerol and 10 mM β-mercaptoethanol]. Cells were lysed by sonication, 1000 units OmniCleave Endonuclease (Westburg) per gram cells were added and the cell lysate was incubated for 30 min at room temperature. The cell lysate was centrifuged for 30 min at 37.000 rpm. The supernatant was applied to a His-Trap column (GE Healthcare), equilibrated in buffer D [20 mM Tris–HCl (pH 7), 500 mM KCl and 10% glycerol]. Protein was eluted with a linear gradient of 0 – 250 mM imidazole (Sso10a1_DBD) or 0 – 500 mM imidazole (Sso10a2_DBD) in buffer D. Proteins were dialyzed at 4°C against a storage buffer [20 mM Tris-HCl (pH 7.0), 10% glycerol and 10 mM β-mercaptoethanol] containing either 600 mM NaCl (Sso10a1_DBD) or 200 mM NaCl (Sso10a2_DBD), and stored at -80 °C until required.

Protein concentrations were determined using a Bicinchoninic Acid (BCA) protein Assay (Thermo Scientific) and are indicated as monomer concentrations.

DNA constructs

TPM substrate

An end-labeled DNA construct of 685-bp was obtained by PCR using biotin- and digoxygenin(DIG)-labeled primers and plasmid pRD118. pRD118 was constructed by inserting a 685-bp long fragment from the *S. solfataricus* P2 genome (22) into the NdeI and BamHI site of pET3-his (257). The PCR product was purified using a GenElute PCR Clean-up kit (Sigma-Aldrich).

AFM substrate

The pRD24 plasmid (a pUC19 derivative, containing a Bpu10I recognition site inserted into the multiple cloning site) used for the AFM experiments was propagated in *E. coli* strain XL10 and purified (Qiagen plasmid midi kit) as described before (178). pRD24 was nicked by digestion with Bpu10I (Fermentas).

Optical tweezers substrate

Bacteriophage λ-DNA (Roche) was biotinylated at both ends by labeling the 12-nucleotide long 5' overhangs with dTTP, dGTP, biotin-14-dATP and biotin-14-dCTP using Klenow DNA polymerase exo$^-$ (Fermentas).

Tethered particle motion experiments

Flow cells (volume ~ 30 µl) were incubated with 20 µg/ml anti-DIG antibodies (Roche) and incubated for 5 min. Passivation of the surface was achieved by flushing the flow cell with 0.4% (w/v) Blotting Grade Blocker (BGB) (Bio-Rad) in buffer I [10 mM Tris (pH 7.5), 150 mM NaCl, 1 mM EDTA, 1 mM DTT, 3% glycerol and 100 µg/ml acetylated BSA (Ambion)] and incubating for 15 min at room temperature. The flow cell was flushed with buffer I, filled with 100 pM DNA solution (functionalized with biotin and DIG), incubated for 10 min and flushed again with buffer I. Streptavidin coated polystyrene beads with a diameter of 0.46 µm (1% w/v; G Kisker) were diluted 300 times in buffer I, flushed into the flow cell and incubated for 10 min to allow binding to the biotinylated-ends of the DNA. After washing the cell with buffer II [10 mM KPO_4 (pH 6.0), 100 mM NaCl, 100 µg/ml acetylated BSA, 1 mM DTT and 0.5% (w/v) CHAPS (Sigma-Aldrich)], the desired concentration (0 – 2000 nM) of Sso10a1 or Sso10a2 of protein diluted in buffer II was flushed into the flow cell and incubated for 10 minutes at room temperature.

Tethered particle motion experiments were performed on an inverted Nikon microscope (Diaphot 300), using a 100x oil-immersion objective (NA = 1.25). Images were acquired by a CMOS camera (Thorlabs) at 25 Hz.

The x- and y-coordinates of individual beads were tracked real-time by custom-developed LabView software (National Instruments) as described previously (259). All experiments were performed at room temperature (~ 23 °C).

Data analysis was performed as described before (see Chapter 3). For each measured condition RMS values were obtained by fitting a single Gaussian to the histogram of the RMS values of individual tethers (N = 46 – 359).

Atomic force microscopy experiments

Sso10a-DNA complexes were formed by incubating 100 ng of nicked pRD24 with varying concentrations of Sso10a proteins in 10 µl AFM buffer [40 mM Hepes (pH 7.5) and 25 mM NaCl] for 10 min at room temperature (~ 23 °C). After incubation, this mixture was diluted 10-fold in 1 mM $MgCl_2$ and directly deposited onto freshly cleaved mica. After 40 s, the mica disc was rinsed with HPLC water and dried with nitrogen gas. The AFM images were collected on a NanoScopeIII AFM (Digital Instruments, Santa Barbara, CA) using micro cantilevers (Olympus MCL-AC240TS-W2, resonance frequency 70 kHz, spring constant 2 N/m). Images were obtained using tapping mode in air at 2 Hz and flattened using Nanoscope software (Veeco Instruments).

Optical tweezers experiments

DNA micromanipulation experiments were performed on a JPK NanoTracker™ set-up integrated with a Nikon Eclipse Ti microscope. A multi-channel laminar flow cell (see Figure 4.2 A) was used to capture single DNA molecules between two streptavidin-coated beads (1.87 µm in diameter, Kisker) in a buffer containing 10 mM Tris–HCl (pH 7.5), 200

mM NaCl, 1 mM EDTA and 1 mM DTT. Force distance (FD) curves were recorded in buffer II containing 100 mM NaCl. DNA molecules were incubated with protein in the absence of tension (4.13 µm between both ends of the DNA molecule) for ~ 2 min. FD curves were performed at a pulling speed of v = 2 µm/s. All experiments were performed at room temperature (~ 23 °C). FD curves were analyzed by fitting to the WLC (214) model using JPK data analysis software.

SUPPLEMENTARY INFORMATION

Table S5.1 Qualitative classification of protein-DNA complexes visualized by AFM. Protein-DNA complexes are classified as: '*open*' if segments of the DNA molecule do not overlap, '*cross-over*' if DNA segments overlap at a single crossover and as '*bridged*' if the DNA molecule contains bridged patches of which the length is longer than a single crossover. N denotes the total number of DNA molecules counted for each condition.

Protein	c (μM)	open	cross-over	bridged	N
-	0	87%	4%	9%	169
Sso10a1	0.5	40%	1%	59%	58
Sso10a1	1	29%	5%	66%	104
Sso10a2	0.5	86%	3%	11%	167
Sso10a2	1	86%	3%	11%	84
Sso10a2	2	93%	-	7%	72

Table S5.2 Physical parameters of DNA and protein-DNA complexes. WLC fitting parameters of FD curves are shown for different protein concentrations. L_0 represents the DNA contour length, L_p the persistence length. All values are the average of N molecules and the error is given as the standard deviation. Fitting parameters for Sso10a1 at 2 μM could not be determined (*nd*) due to the formation of intersegmental bridges.

Protein	c (μM)	L_o (μm)	L_p (nm)	N
-	0	16.7 ± 0.1	39.7 ± 7.9	28
Sso10a1	0.1	18.3 ± 0.7	4.2 ± 2.5	15
Sso10a1	1	18.5 ± 0.4	2.2 ± 0.4	6
Sso10a1	2	*nd*	*nd*	*nd*
Sso10a2	0.1	18.7 ± 0.6	2.8 ± 1.7	8
Sso10a2	1	18.5 ± 2.8	2.8 ± 1.0	11
Sso10a2	3	18.2 ± 0.9	3.3 ± 1.0	9

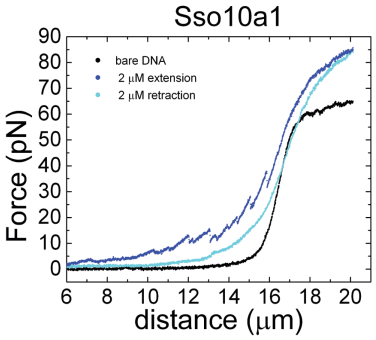


Figure S5.1 Force distance curve of an individual DNA molecule incubated with 2 μM Sso10a1 showing the extension (*blue*) and retraction (*cyan*) of the DNA molecule. The control FD curve of the bare DNA molecule is depicted in black. The force peaks at distance < L_0 indicate intersegmental bridging.

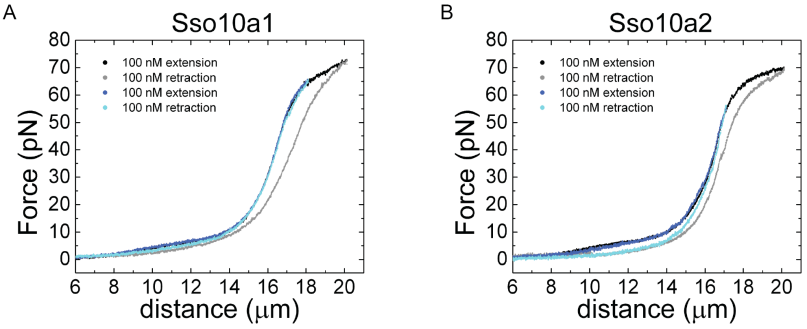


Figure S5.2 FD curves of Sso10a1 and Sso10a2 (100 nM), showing hysteresis between the extension (*black*) and retraction (*grey*) curves. The hysteresis is reduced when the molecule is extended up to forces below the overstretching regime (*blue and cyan curves*).

Chapter 6

Summary and Conclusions

Understanding of chromatin organization and compaction in Archaea is currently limited. The genome of several megabasepairs long is folded by a set of small chromatin proteins to fit into the micron-sized cell. A first step in understanding archaeal chromatin organization is to study the action of individual chromatin proteins on DNA. Characterization of the architectural properties of these proteins is essential to understand how they shape and modulate the archaeal genome. This thesis describes the biophysical characterization of several chromatin proteins from the crenarchaeal model organism *Sulfolobus solfataricus*: Cren7, Sul7, Alba and Sso10a. The architectural properties of these proteins resemble those of their bacterial counterparts, suggesting that they could play a similar role in chromatin organization and global gene regulation.

A characterization of the architectural properties of Cren7 and Sul7 is described in **Chapter 2**. These small monomeric proteins are highly abundant in the cell and bind with no apparent sequence specificity to dsDNA. Cren7 and Sul7 are therefore believed to play an important role in the organization of the genomic DNA of *Sulfolobus* species. Although Cren7 and Sul7 are not conserved at the sequence level they are structural homologues and share biochemical and architectural properties. Both proteins induce a sharp bend upon binding to DNA. This bending angle is investigated using molecular dynamics simulations, which reveal that Cren7 and Sul7 both induce a non-flexible bend of ~ 50°. In depth analysis of DNA micromanipulation experiments confirms that the induced bends are non-flexible. In this respect these proteins differ from their bacterial and eukaryotic counterparts (HU and HMGB), which induce bends with a certain degree of flexibility. Single-molecule experiments reveal that Cren7 and Sul7 compact DNA molecules to a similar extent, by reducing the apparent persistence length. Interestingly, the protein-DNA interactions of Cren7 and Sul7 are unaffected by applied forces up to 3.5 pN as the proteins stay bound to the DNA at this force. This finding suggests that Cren7 and Sul7 are relatively stably bound to the DNA *in vivo*, which could be

important in the light of maintaining DNA integrity under the extreme environmental conditions of the habitat of thermophilic *Sulfolobus* species.

As *Sulfolobus* species live in extreme environments and are subjected to high temperatures and large temperature fluctuations it is important to know how DNA structure and protein-DNA interactions are affected by temperature. To this purpose we have developed a temperature-controlled tethered particle motion (TPM) set-up, which enables to investigate DNA structure and protein-DNA interactions at a single-molecule level at controlled temperatures (**see Chapter 3**). Using this technique we reveal that the intrinsic flexibility of dsDNA strongly depends on temperature in a range well below the DNA melting temperature. In the measured temperature range (23 – 52 °C) the apparent persistence length of dsDNA depends linearly on temperature and is slightly dependent on the composition of the DNA (i.e. the percentage of AT/GC base pairs). Besides the temperature effect on intrinsic DNA flexibility, our experiments reveal that temperature affects protein-DNA interactions of DNA-bending proteins in different manners. Temperature affects the binding affinity of Cren7 and Sul7 (static kink DNA-bending proteins), while it affects the distribution of bending angles induced by HU (a flexible hinge bending protein), rather than its binding affinity. The increase in flexibility of DNA at high – physiological relevant – temperatures and temperature dependent protein-DNA interactions of DNA-bending proteins suggest that high temperature contributes to a more compact configuration of the genomic DNA *in vivo*. In addition, as genome structure and topology are temperature dependent this may affect gene expression on a global level.

Chapter 4 describes a detailed characterization of the crenarchaeal chromatin protein Alba. *Sulfolobus* species express two different Alba proteins: Alba1 and Alba2. Alba2 forms obligate heterodimers with Alba1. The DNA-binding properties of these heterodimers differ significantly from Alba1 homodimers. Alba displays two different architectural modes: DNA-

bridging and DNA stiffening. Furthermore, binding of Alba1 is strongly cooperative due to dimer-dimer interactions, while Alba1:Alba2 heterodimers lack this cooperative binding behaviour. The lack of cooperative DNA binding of the Alba F60A mutant reveals that the F60 residue (which is not conserved in Alba2) plays an essential role in this cooperative binding behaviour. Interestingly, the bridging binding mode of Alba also depends on dimer-dimer interactions, including the F60 residue. Based on these findings we propose a model in which Alba dimers bound to DNA can alternate between two different orientations, facilitating either bridging of two DNA duplexes or side-by-side binding of dimers, resulting in stiffening. As Alba1:Alba2 heterodimers lack the F60 residue at one side of the dimer, they are unable to form long patches of closely packed dimers on the DNA. The ratio between Alba1 homodimers and Alba1:Alba2 heterodimers (controlled by expression levels of Alba1 and Alba2) can therefore tune the balance between stiffened and bridged DNA, which could provide mechanisms to regulate gene expression on a global scale *in vivo*.

The Sso10a protein is yet another DNA-binding protein that has been identified in *Sulfolobus* (**see Chapter 5**). Sso10a proteins exist as dimers with a coiled-coil dimerization domain with two DNA-binding domains at the ends. The DNA-binding domain of Sso10a adopts a helix-turn-helix (HTH) motif, which is often found in archaeal transcription regulators. *S. solfataricus* expresses three homologous Sso10a proteins: Sso10a1, Sso10a2 and Sso10a3, which are – based on transcription levels – expected to be abundant within the cell. Structural comparison of Sso10a1 and Sso10a2 reveals differences in the DNA-binding domains, which are expected to lead to different DNA-binding properties. An investigation on the architectural properties of Sso10a reveals three different binding modes: bending, bridging and stiffening. Both Sso10a1 and Sso10a2 induce bends into DNA by binding *in cis*, which leads to a decrease in the apparent persistence length. At relatively high protein concentrations Sso10a1 exhibits a second binding mode of binding, DNA-bridging, not observed for Sso10a2. Instead, Sso10a2 stiffens DNA at similar protein concentrations, which is likely due

to electrostatic dimer-dimer interactions. The multiple architectural properties of the Sso10a proteins suggest that these proteins function in shaping chromatin structure. The interplay between the different homologues and binding modes could provide mechanisms to dynamically shape the genome and to regulate gene expression *in vivo*.

With the characterization of the architectural properties of Cren7, Sul7, Alba and Sso10a, a first step has been made in understanding how crenarchaeal chromatin is organized. To further our understanding of the organization of the crenarchaeal nucleoid more information is needed on how these proteins act *in vivo* on different length scales within the complex cellular environment. A detailed understanding on the interplay of architectural proteins complemented with genome association studies could shed light on whether and how such interplay exist *in vivo*. Chromatin immunoprecipiation studies in combination with sequencing (CHip-seq) can provide information on whether different architectural proteins co-localize *in vivo* at specific sites along the genome, which could point to regulation mechanisms obtained by the interplay between these proteins. *In vitro* single-molecule studies could provide more detailed structural information on such nucleo-protein complexes formed by the combined action of different architectural proteins (eg. DNA-bending protein Cren7 or Sul7 with DNA-bridging protein Alba).

The numerous analogies between crenarchaeal and bacterial chromatin proteins suggest that higher-order organization of the crenarchaeal nucleoid relies on similar mechanisms as that of the bacterial nucleoid (**see Chapter 1**). Bridging proteins Alba and/or Sso10a could facilitate higher-order looped domains, similar to the bacterial bridging protein H-NS. Such looped domains could be further compacted by the action of DNA-bending proteins Cren7, Sul7 or Sso10a (see Figure 1.5). Information on higher-order structures *in vivo* could be obtained by applying chromosome conformation capture (3C) techniques, which allow to study the threedimensional genome structure at a high resolution. By probing

interactions between distant DNA segments this technique has been successfully applied in studies on *in vivo* genome structure in eukaryotes and bacteria (293). Combining 3C with chromatin immunoprecipitation (ChIP-loop) allows to study specific protein-mediated DNA-DNA interactions. Applying this technique could provide insights on how specific chromatin proteins are involved in higher-order organization of the crenarchaeal nucleoid and thus answer the question whether Alba and/or Sso10a are involved in higher-order organization by bridging distant DNA segments.

Integrating information from *in vitro* and *in vivo* studies at different scales will provide frameworks to obtain a detailed understanding on the structure and dynamics of the crenarchaeal nucleoid. Recent developments on genetic tools for archaea (294) enable *in vivo* studies, which have been unavailable for archaea until recently. Combining knowledge on structural and DNA-binding properties of chromatin proteins with information on genome-wide association and three-dimensional chromatin structure will provide a more detailed model on the organization and dynamics of the crenarchaeal nucleoid.

Nederlandse samenvatting

Het begrip van chromatineorganisatie in archaea is momenteel beperkt. Het genoom van enkele megabaseparen lang wordt opgevouwen door zogenaamde nucleoïde-geassocieerde eiwitten om in de cel met een diameter van ongeveer een micrometer te passen. Het bestuderen van de werking van individuele chromatine-eiwitten op DNA is een eerste stap in naar het begrijpen van archaeale chromatineorganisatie. Karakterisatie van de architecturele eigenschappen van deze eiwitten is essentieel om te doorzien hoe zij het archaeale genoom vormgeven en moduleren. Dit proefschrift beschrijft de biofysische karakterisatie van chromatine-eiwitten uit het crenarchaeale modelorganisme *Sulfolobus solfataricus*: Cren7, Sul7, Alba en Sso10a. De overeenkomsten in architecturele eigenschappen van de crenarchaeale chromatine eiwitten met die van hun bacteriële tegenhangers suggereren dat deze eiwitten een vergelijkbare rol in chromatineorganisatie en genregulatie spelen.

Hoofdstuk 2 beschrijft de karakterisatie van de architecturele eigenschappen van Cren7 en Sul7. Deze kleine monomere eiwitten zijn overvloedig aanwezig in de cel en binden dubbelstrengs DNA (dsDNA) zonder sequentie-specificiteit. Cren7 en Sul7 spelen waarschijnlijk een belangrijke rol in de ruimtelijke organisatie van het genoom van *Sulfolobus*-soorten. Hoewel Cren7 en Sul7 geen overeenkomsten hebben op sequentieniveau is de structuur van de gevouwen eiwitten vergelijkbaar. Deze eiwitten hebben vergelijkbare biochemische en architecturele eigenschappen en induceren een scherpe buiging in DNA. Deze buigingshoek is in detail onderzocht middels moleculaire dynamica-simulaties. Hieruit blijkt dat Cren7 en Sul7 beide een niet-flexibele hoek induceren van ~ 50°. Gedetailleerde analyse van DNA-micromanipulatie

experimenten bevestigt dat Cren7 en Sul7 het DNA buigen onder een niet-flexibele buigingshoek. In dit opzicht verschillen deze eiwitten van hun bacteriële en eukaryote tegenhangers (HU en HMGB), die DNA buigen onder een hoek met een bepaalde flexibiliteit. Experimenten aan individuele DNA moleculen tonen aan dat Cren7 en Sul7 DNA in gelijke mate compact maken doordat ze de effectieve persistentielengte reduceren. Het is opmerkelijk dat de eiwit-DNA interacties van Cren7 en Sul7 niet beïnvloed worden door krachten tot 3.5 pN, aangezien de eiwitten zelfs bij deze kracht aan het DNA gebonden blijven. Deze bevindingen suggereren dat Cren7 en Sul7 relatief stevig aan het DNA gebonden zijn *in vivo*, wat belangrijk zou kunnen zijn in de context van het waarborgen van DNA-integriteit in de extreme omgevingscondities van thermofiele *Sulfolobus*-soorten.

Aangezien *Sulfolobus*-soorten in extreme leefmilieus leven en blootgesteld zijn aan hoge temperaturen en grote temperatuurschommelingen is het belangrijk om te weten hoe DNA-structuur en eiwit-DNA-interacties afhankelijk zijn van temperatuur. Om dit te bestuderen hebben we een 'Tethered Particle Motion' (TPM) opstelling ontwikkeld, waarmee DNA-structuur en eiwit-DNA-interacties bestudeerd kunnen worden op gedefinieerde temperaturen (**zie Hoofdstuk 3**). Met behulp van deze techniek hebben we aangetoond dat de intrinsieke flexibiliteit van dubbelstrengs DNA sterk afhankelijk is van de temperatuur bij temperaturen ver onder de smelttemperatuur van DNA. Binnen het gemeten temperatuurbereik (23 – 52 °C) verhoudt de effectieve persistentielengte van dubbelstrengs DNA zich lineair tot de temperatuur en is enigszins afhankelijk van de samenstelling van het DNA (het percentage AT/GC basenparen). Behalve het effect van temperatuur op de intrinsieke DNA-flexibiliteit laten onze experimenten ook zien dat temperatuur de eiwit-DNA interacties van DNA-buigende eiwitten op verschillende manieren kan beïnvloeden. Temperatuur beïnvloedt de bindingsaffiniteit van Cren7 en Sul7, die DNA buigen onder een niet-flexibele hoek. In tegenstelling tot Cren7 en Sul7 wordt de bindingsaffiniteit van HU, dat DNA buigt onder een flexibele hoek, niet beïnvloed door

temperatuur. In plaats daarvan is de flexibiliteit van de geïnduceerde buiging afhankelijk van temperatuur. De toename van de flexibiliteit van DNA op hoge – fysiologisch relevante – temperaturen en de temperatuurafhankelijke eiwit-DNA-interacties van DNA-buigende eiwitten suggereren dat een hoge temperatuur bijdraagt aan een compacte configuratie van het genoom van thermofiele organismen. Daarnaast zou temperatuur genexpressie op een globaal niveau kunnen beïnvloeden, daar de structuur en topologie van het genoom afhankelijk zijn van temperatuur.

Hoofdstuk 4 beschrijft een gedetailleerde karakterisatie van het crenarchaeale eiwit Alba. *Sulfolobus*-soorten synthetiseren twee verschillende Alba eiwitten: Alba1 en Alba2. Alba2 vormt obligate heterodimeren met Alba1. De DNA-bindingseigenschappen van deze heterodimeren verschillen aanzienlijk ten opzichte van Alba1 homodimeren. Alba heeft twee verschillende bindingsmodi: DNA-brugging en DNA-verstijving. Binding van Alba1 is sterk coöperatief dankzij dimeer-dimeer interacties, terwijl voor Alba1:Alba2-heterodimeren deze coöperatieve binding ontbreekt. Het feit dat de Alba F60A-mutant niet coöperatief bindt aan DNA bevestigt dat het F60 residu (dat afwezig is in Alba2) verantwoordelijk is voor deze coöperatieve DNA-binding. Uit onze studies blijkt dat de DNA-bruggingsmodus van Alba ook afhankelijk is van dimeer-dimeer interacties en dat ook daar het F60-residu een cruciale rol speelt. Gebaseerd op deze bevindingen stellen we een model voor waarin Alba-dimeren, gebonden aan het DNA, van oriëntatie kunnen veranderen. Hierdoor wordt of de vorming van bruggen tussen twee DNA strengen danwel DNA verstijving – veroorzaakt door naast elkaar bindende dimeren – in staat gesteld. Omdat Alba1:Alba2 heterodimeren het F60 residu aan één kant van de dimeer missen kunnen ze geen lange stukken van aaneengesloten dimeren op het DNA vormen. Derhalve kan de verhouding tussen Alba1 homodimeren en Alba1:Alba2 heterodimeren (afhankelijk van de expressie niveaus van Alba1 en Alba2) de balans tussen gebrugd en verstijfd DNA afstemmen, wat een belangrijk mechanisme zou kunnen zijn om *in vivo* genexpressie op een globaal niveau te kunnen reguleren.

Het Sso10a eiwit is een ander overvloedig DNA-bindend eiwit uit *Sulfolobus* -soorten (**zie Hoofdstuk 5**). Sso10a eiwitten bestaan in de vorm van dimeren door dimerisatie via een coiled-coil domein waaraan zich aan het uiteinde een DNA-bindend domein bevindt. Het DNA-bindende domein van Sso10a heeft een karakteristieke 'helix-turn-helix'- structuur, die vaak voorkomt in archaeale transcriptiefactoren. *S. solfataricus* brengt drie homologe Sso10a eiwitten tot expressie: Sso10a1, Sso10a2 en Sso10a3, waarvan verondersteld wordt – op basis van de mate van transcriptie – dat ze overvloedig aanwezig zijn in de cel. Structurele vergelijking van Sso10a1 en Sso10a2 laat verschillen zien tussen de DNA-bindende domeinen, die naar waarschijnlijkheid tot verschillende DNA-bindende eigenschappen leiden. Onderzoek naar de architecturele eigenschappen van Sso10a laat drie verschillende bindingsmodi zien: buiging, brugging en verstijving. Sso10a1 en Sso10a2 buigen DNA bij binding *in cis* en dit leidt tot een afname van de effectieve persistentielengte. Op relatief hoge eiwitconcentraties heeft Sso10a1 een tweede bindingsmodus, DNA-brugging. Deze modus wordt niet waargenomen voor Sso10a2. In plaats daarvan verstijft Sso10a2 DNA bij vergelijkbare eiwitconcentraties, wat waarschijnlijk een gevolg is van elektrostatische interacties tussen naastgelegen dimeren. De multimodale architecturele eigenschappen van de Sso10a eiwitten suggereren dat deze eiwitten een functie vervullen in het vormgeven en aanpassen van chromatinestructuur. Samenspel van verschillende homologen en bindingsmodi zou mechanismen kunnen bieden om *in vivo* het genoom dynamisch vorm te geven en genexpressie te reguleren.

Met de karakterisatie van de architecturele eigenschappen van Cren7, Sul7, Alba en Sso10a is een eerste stap gezet tot het begrijpen van de organisatie van crenarchaeaal chromatine. Om een beter en meer volledig inzicht te verkrijgen in de organisatie van de crenarchaeale nucleoïde is echter meer informatie nodig over hoe deze eiwitten *in vivo* werken op verschillende lengteschalen in de complexe context van de cel. Een gedetailleerd begrip van de interactie van de architecturele eiwitten aangevuld met genoombrede associatiestudies zou licht werpen op de vraag of en hoe dergelijk

samenspel bestaat *in vivo*. Chromatineimmunoprecipiatiestudies in combinatie met sequencing (ChIP-seq) kan informatie verschaffen over de vraag of verschillende architecturele eiwitten co-lokaliseren *in vivo* op specifieke plaatsen langs het genoom, wat zou kunnen wijzen op regulatiemechanismen verkregen door de interactie tussen deze eiwitten. *In vitro* experimenten op het niveau van enkele moleculen zouden meer gedetailleerde structurele informatie kunnen verschaffen over zulke eiwit-DNA complexen, gevormd door de gecombineerde werking van verschillende architecturele eiwitten in (bijv. DNA-buigend eiwit Cren7 of Sul7 met DNA-bruggend eiwit Alba).

Analogie tussen crenarchaeale en bacteriële chromatine-eiwitten suggereert dat hogere-orde organisatie van de crenarchaeale nucleoïde is gebaseerd op dezelfde mechanismen als die van de bacteriële nucleoide (**zie hoofdstuk 1**) . Bruggende eiwitten Alba en/of Sso10a zouden hogere-orde lusvormige domeinen kunnen vormen, vergelijkbaar met het bacteriële bruggende eiwit H-NS. Dergelijke lusvormige domeinen kunnen verder compact worden gemaakt door de werking van DNA-buigende eiwitten Cren7, Sul7 of Sso10a (zie figuur 1.5). Informatie over hogere-orde structuren *in vivo* kan worden verkregen met behulp van 'Chromosome Conformation Capture' (3C) technieken, waarmee de driedimensionale structuur van het genoom op hoge resolutie bestudeerd kan worden. Door middel van het meten van interactiefrequenties tussen DNA-segmenten is deze techniek eerder succesvol toegepast in studies naar de *in vivo* genoomstructuur in eukaryoten en bacteriën (293). De combinatie van 3C met chromatine immunoprecipitatie (ChIP-loop) maakt het mogelijk om DNA-DNA interacties te bestuderen, die door de binding van een specifiek eiwit worden veroorzaakt. Deze techniek kan inzichten bieden in hoe specifieke chromatine-eiwitten zijn betrokken bij hogere-orde organisatie van de crenarchaeale nucleoïde en zou zo kunnen uitwijzen of Alba en/of Sso10a hogere-orde structuren vormen door bruggen te vormen tussen ver van elkaar verwijderde DNA-segmenten.

Het integreren van informatie van *in vitro* en *in vivo* studies op verschillende schalen zal helpen om een gedetailleerd begrip van de structuur en dynamiek van de crenarchaeale nucleoïde te verkrijgen. Recente ontwikkelingen op het gebied van genetische instrumenten voor archaea (294) maken *in vivo* studies mogelijk die tot voor kort niet mogelijk waren voor archaea. Bundeling van kennis over de structurele en DNA-bindende eigenschappen van chromatine-eiwitten met informatie over genoombrede associatie en driedimensionale chromatinestructuur zal een meer gedetailleerd model over de organisatie en de dynamiek van de crenarchaeale nucleoïde bieden.

REFERENCES

1. Woese, C.R. and Fox, G.E. (1977) Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proc Natl Acad Sci U S A*, **74**, 5088-5090.
2. Woese, C.R., Kandler, O. and Wheelis, M.L. (1990) Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proc Natl Acad Sci U S A*, **87**, 4576-4579.
3. Brochier-Armanet, C., Boussau, B., Gribaldo, S. and Forterre, P. (2008) Mesophilic Crenarchaeota: proposal for a third archaeal phylum, the Thaumarchaeota. *Nat Rev Microbiol*, **6**, 245-252.
4. Elkins, J.G., Podar, M., Graham, D.E., Makarova, K.S., Wolf, Y., Randau, L., Hedlund, B.P., Brochier-Armanet, C., Kunin, V., Anderson, I. *et al.* (2008) A korarchaeal genome reveals insights into the evolution of the Archaea. *Proc Natl Acad Sci U S A*, **105**, 8102-8107.
5. Waters, E., Hohn, M.J., Ahel, I., Graham, D.E., Adams, M.D., Barnstead, M., Beeson, K.Y., Bibbs, L., Bolanos, R., Keller, M. *et al.* (2003) The genome of *Nanoarchaeum equitans*: insights into early archaeal evolution and derived parasitism. *Proc Natl Acad Sci U S A*, **100**, 12984-12988.
6. Pikuta, E.V., Hoover, R.B. and Tang, J. (2007) Microbial extremophiles at the limits of life. *Crit Rev Microbiol*, **33**, 183-209.
7. DeLong, E.F. and Pace, N.R. (2001) Environmental diversity of bacteria and archaea. *Syst Biol*, **50**, 470-478.
8. Eckburg, P.B., Lepp, P.W. and Relman, D.A. (2003) Archaea and their potential role in human disease. *Infect Immun*, **71**, 591-596.
9. Reeve, J.N. (1999) Archaebacteria then ... Archaes now (are there really no archaeal pathogens?). *J Bacteriol*, **181**, 3613-3617.
10. Jarrell, K.F., Walters, A.D., Bochiwal, C., Borgia, J.M., Dickinson, T. and Chong, J.P. (2011) Major players on the microbial stage: why archaea are important. *Microbiology*, **157**, 919-936.
11. Grohmann, D. and Werner, F. (2011) Recent advances in the understanding of archaeal transcription. *Curr Opin Microbiol*, **14**, 328-334.

12. Kelman, Z. and White, M.F. (2005) Archaeal DNA replication and repair. *Curr Opin Microbiol*, **8**, 669-676.
13. Lindas, A.C. and Bernander, R. (2013) The cell cycle of archaea. *Nat Rev Microbiol*, **11**, 627-638.
14. Luijsterburg, M.S., White, M.F., van Driel, R. and Dame, R.T. (2008) The major architects of chromatin: architectural proteins in bacteria, archaea and eukaryotes. *Crit Rev Biochem Mol Biol*, **43**, 393-418.
15. Cunha, S., Woldringh, C.L. and Odijk, T. (2001) Polymer-mediated compaction and internal dynamics of isolated *Escherichia coli* nucleoids. *J Struct Biol*, **136**, 53-66.
16. de Vries, R. (2010) DNA condensation in bacteria: Interplay between macromolecular crowding and nucleoid proteins. *Biochimie*, **92**, 1715-1721.
17. Brock, T.D., Brock, K.M., Belly, R.T. and Weiss, R.L. (1972) *Sulfolobus*: a new genus of sulfur-oxidizing bacteria living at low pH and high temperature. *Arch Mikrobiol*, **84**, 54-68.
18. van de Vossenberg, J.L., Ubbink-Kok, T., Elferink, M.G., Driessen, A.J. and Konings, W.N. (1995) Ion permeability of the cytoplasmic membrane limits the maximum growth temperature of bacteria and archaea. *Mol Microbiol*, **18**, 925-932.
19. Chen, L., Brugger, K., Skovgaard, M., Redder, P., She, Q., Torarinsson, E., Greve, B., Awayez, M., Zibat, A., Klenk, H.P. *et al.* (2005) The genome of *Sulfolobus acidocaldarius*, a model organism of the Crenarchaeota. *J Bacteriol*, **187**, 4992-4999.
20. Kawarabayasi, Y., Hino, Y., Horikawa, H., Jin-no, K., Takahashi, M., Sekine, M., Baba, S., Ankai, A., Kosugi, H., Hosoyama, A. *et al.* (2001) Complete genome sequence of an aerobic thermoacidophilic crenarchaeon, *Sulfolobus tokodaii* strain7. *DNA Res*, **8**, 123-140.
21. Reno, M.L., Held, N.L., Fields, C.J., Burke, P.V. and Whitaker, R.J. (2009) Biogeography of the *Sulfolobus islandicus* pan-genome. *Proc Natl Acad Sci U S A*, **106**, 8605-8610.
22. She, Q., Singh, R.K., Confalonieri, F., Zivanovic, Y., Allard, G., Awayez, M.J., Chan-Weiher, C.C., Clausen, I.G., Curtis, B.A., De Moors, A. *et al.* (2001) The complete genome of the crenarchaeon *Sulfolobus solfataricus* P2. *Proc Natl Acad Sci U S A*, **98**, 7835-7840.
23. Berkner, S. and Lipps, G. (2008) Genetic tools for *Sulfolobus* spp.: vectors and first applications. *Arch Microbiol*, **190**, 217-230.

24. Zillig, W., Arnold, H.P., Holz, I., Prangishvili, D., Schweier, A., Stedman, K., She, Q., Phan, H., Garrett, R. and Kristjansson, J.K. (1998) Genetic elements in the extremely thermophilic archaeon *Sulfolobus*. *Extremophiles*, **2**, 131-140.
25. Charbonnier, F. and Forterre, P. (1994) Comparison of plasmid DNA topology among mesophilic and thermophilic eubacteria and archaebacteria. *J Bacteriol*, **176**, 1251-1259.
26. Forterre, P., Bergerat, A. and Lopez-Garcia, P. (1996) The unique DNA topology and DNA topoisomerases of hyperthermophilic archaea. *FEMS Microbiol Rev*, **18**, 237-248.
27. Lopez-Garcia, P. and Forterre, P. (1997) DNA topology in hyperthermophilic archaea: reference states and their variation with growth phase, growth temperature, and temperature stresses. *Mol Microbiol*, **23**, 1267-1279.
28. Nadal, M. (2007) Reverse gyrase: an insight into the role of DNA-topoisomerases. *Biochimie*, **89**, 447-455.
29. Ellis, R.J. (2001) Macromolecular crowding: obvious but underappreciated. *Trends Biochem Sci*, **26**, 597-604.
30. Marenduzzo, D., Finan, K. and Cook, P.R. (2006) The depletion attraction: an underappreciated force driving cellular organization. *J Cell Biol*, **175**, 681-686.
31. Murphy, L.D. and Zimmerman, S.B. (1994) Macromolecular crowding effects on the interaction of DNA with *Escherichia coli* DNA-binding proteins: a model for bacterial nucleoid stabilization. *Biochim Biophys Acta*, **1219**, 277-284.
32. de Vries, R. (2012) DNA compaction by nonbinding macromolecules. *Polym. Sci. Ser. C*, **54**, 30-35.
33. Flory, P.J. (1969) Statistical mechanics of chain molecules. *Interscience Publishers, New York*.
34. 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 A resolution. *Nature*, **389**, 251-260.
35. Luger, K., Dechassa, M.L. and Tremethick, D.J. (2012) New insights into nucleosome and chromatin structure: an ordered state or a disordered affair? *Nat Rev Mol Cell Biol*, **13**, 436-447.
36. Luijsterburg, M.S., Noom, M.C., Wuite, G.J. and Dame, R.T. (2006) The architectural role of nucleoid-associated proteins in the organization of bacterial chromatin: a molecular perspective. *J Struct Biol*, **156**, 262-272.

37. Dame, R.T. (2005) The role of nucleoid-associated proteins in the organization and compaction of bacterial chromatin. *Mol Microbiol*, **56**, 858-870.
38. Dillon, S.C. and Dorman, C.J. (2010) Bacterial nucleoid-associated proteins, nucleoid structure and gene expression. *Nat Rev Microbiol*, **8**, 185-195.
39. Bell, S.D. and White, M.F. (2010) In Dame, R. T. and Dorman, C. J. (eds.), *Bacterial Chromatin*. Springer, pp. 205-217.
40. Driessen, R.P.C. and Dame, R.T. (2011) Nucleoid-associated proteins in Crenarchaea. *Biochem. Soc. Trans.*, **39**, 116-121.
41. Reeve, J.N. (2003) Archaeal chromatin and transcription. *Mol Microbiol*, **48**, 587-598.
42. White, M.F. and Bell, S.D. (2002) Holding it together: chromatin in the Archaea. *Trends Genet*, **18**, 621-626.
43. Zhang, Z., Guo, L. and Huang, L. (2012) Archaeal chromatin proteins. *Sci China Life Sci*, **55**, 377-385.
44. Pereira, S.L., Grayling, R.A., Lurz, R. and Reeve, J.N. (1997) Archaeal nucleosomes. *Proc Natl Acad Sci U S A*, **94**, 12633-12637.
45. Reeve, J.N., Bailey, K.A., Li, W.-t., Marc, F., Sandman, K. and Soares, D.J. (2004) Archaeal histones: structures, stability and DNA binding. *Biochem Soc Trans*, **32**, 227-230.
46. Decanniere, K., Babu, A.M., Sandman, K., Reeve, J.N. and Heinemann, U. (2000) Crystal structures of recombinant histones HMfA and HMfB from the hyperthermophilic archaeon *Methanothermus fervidus*. *J Mol Biol*, **303**, 35-47.
47. Sandman, K. and Reeve, J.N. (2006) Archaeal histones and the origin of the histone fold. *Curr Opin Microbiol*, **9**, 520-525.
48. Tomschik, M., Karymov, M.A., Zlatanova, J. and Leuba, S.H. (2001) The archaeal histone-fold protein HMf organizes DNA into bona fide chromatin fibers. *Structure*, **9**, 1201-1211.
49. Maruyama, H., Shin, M., Oda, T., Matsumi, R., Ohniwa, R.L., Itoh, T., Shirahige, K., Imanaka, T., Atomi, H., Yoshimura, S.H. *et al.* (2010) Histone and TK0471/TrmBL2 form a novel heterogeneous genome architecture in the hyperthermophilic archaeon *Thermococcus kodakarensis*. *Molecular biology of the cell*.
50. Ammar, R., Torti, D., Tsui, K., Gebbia, M., Durbic, T., Bader, G.D., Giaever, G. and Nislow, C. (2012) Chromatin is an ancient innovation conserved between Archaea and Eukarya. *Elife*, **1**, e00078.

51. Nalabothula, N., Xi, L., Bhattacharyya, S., Widom, J., Wang, J.P., Reeve, J.N., Santangelo, T.J. and Fondufe-Mittendorf, Y.N. (2013) Archaeal nucleosome positioning *in vivo* and *in vitro* is directed by primary sequence motifs. *BMC Genomics*, **14**, 391.
52. Maruyama, H., Harwood, J.C., Moore, K.M., Paszkiewicz, K., Durley, S.C., Fukushima, H., Atomi, H., Takeyasu, K. and Kent, N.A. (2013) An alternative beads-on-a-string chromatin architecture in *Thermococcus kodakarensis*. *EMBO Rep*.
53. Culard, F., Laine, B., Sautiere, P. and Maurizot, J.C. (1993) Stoichiometry of the binding of chromosomal protein MC1 from the archaebacterium, *Methanosarcina* spp. CHTI55, to DNA. *FEBS Lett*, **315**, 335-339.
54. Paquet, F., Culard, F., Barbault, F., Maurizot, J.C. and Lancelot, G. (2004) NMR solution structure of the archaebacterial chromosomal protein MC1 reveals a new protein fold. *Biochemistry*, **43**, 14971-14978.
55. Cam, E.L., Culard, F., Larquet, E., Delain, E. and Cognet, J.A. (1999) DNA bending induced by the archaebacterial histone-like protein MC1. *J Mol Biol*, **285**, 1011-1021.
56. Pavlov, N.A., Cherny, D.I., Nazimov, I.V., Slesarev, A.I. and Subramaniam, V. (2002) Identification, cloning and characterization of a new DNA-binding protein from the hyperthermophilic methanogen *Methanopyrus kandleri*. *Nucleic Acids Res*, **30**, 685-694.
57. Stein, D.B. and Searcy, D.G. (1978) Physiologically important stabilization of DNA by a prokaryotic histone-like protein. *Science*, **202**, 219-221.
58. Maruyama, H., Shin, M., Oda, T., Matsumi, R., Ohniwa, R.L., Itoh, T., Shirahige, K., Imanaka, T., Atomi, H., Yoshimura, S.H. *et al.* (2011) Histone and TK0471/TrmBL2 form a novel heterogeneous genome architecture in the hyperthermophilic archaeon *Thermococcus kodakarensis*. *Mol Biol Cell*, **22**, 386-398.
59. Grote, M., Dijk, J. and Reinhardt, R. (1986) Ribosomal and DNA binding proteins of the thermoacidophilic archaebacterium *Sulfolobus acidocaldarius*. *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology*, **873**, 405-413.
60. Forterre, P., Confalonieri, F. and Knapp, S. (1999) Identification of the gene encoding archeal-specific DNA-binding proteins of the Sac10b family. *Mol Microbiol*, **32**, 669-670.

61. Choli, T., Henning, P., Wittmann-Liebold, B. and Reinhardt, R. (1988) Isolation, characterization and microsequence analysis of a small basic methylated DNA-binding protein from the Archaebacterium, *Sulfolobus solfataricus*. *Biochim Biophys Acta*, **950**, 193-203.
62. Edmondson, S.P. and Shriver, J.W. (2001) DNA binding proteins Sac7d and Sso7d from Sulfolobus. *Methods Enzymol*, **334**, 129-145.
63. Guo, L., Feng, Y., Zhang, Z., Yao, H., Luo, Y., Wang, J. and Huang, L. (2008) Biochemical and structural characterization of Cren7, a novel chromatin protein conserved among Crenarchaea. *Nucleic Acids Res*, **36**, 1129-1137.
64. Chen, L., Chen, L.R., Zhou, X.E., Wang, Y., Kahsai, M.A., Clark, A.T., Edmondson, S.P., Liu, Z.J., Rose, J.P., Wang, B.C. *et al.* (2004) The hyperthermophile protein Sso10a is a dimer of winged helix DNA-binding domains linked by an antiparallel coiled coil rod. *J Mol Biol*, **341**, 73-91.
65. Kahsai, M.A., Vogler, B., Clark, A.T., Edmondson, S.P. and Shriver, J.W. (2005) Solution structure, stability, and flexibility of Sso10a: a hyperthermophile coiled-coil DNA-binding protein. *Biochemistry*, **44**, 2822-2832.
66. Hsu, C.H. and Wang, A.H. (2011) The DNA-recognition fold of Sso7c4 suggests a new member of SpoVT-AbrB superfamily from archaea. *Nucleic Acids Res*, **39**, 6764-6774.
67. Oppermann, U.C., Knapp, S., Bonetto, V., Ladenstein, R. and Jornvall, H. (1998) Isolation and structure of repressor-like proteins from the archaeon *Sulfolobus solfataricus*. Co-purification of RNase A with Sso7c. *FEBS Lett*, **432**, 141-144.
68. Xue, H., Guo, R., Wen, Y., Liu, D. and Huang, L. (2000) An abundant DNA binding protein from the hyperthermophilic archaeon *Sulfolobus shibatae* affects DNA supercoiling in a temperature-dependent fashion. *J Bacteriol*, **182**, 3929-3933.
69. Crnigoj, M., Podlesek, Z., Zorko, M., Jerala, R., Anderluh, G. and Ulrih, N.P. (2013) Interactions of archaeal chromatin proteins Alba1 and Alba2 with nucleic acids. *PLoS One*, **8**, e58237.
70. Baumann, H., Knapp, S., Lundbäck, T., Ladenstein, R. and Härd, T. (1994) Solution structure and DNA-binding properties of a thermostable protein from the archaeon *Sulfolobus solfataricus*. *Nat Struct Biol*, **1**, 808-819.
71. Wardleworth, B.N., Russell, R.J., Bell, S.D., Taylor, G.L. and White, M.F. (2002) Structure of Alba: an archaeal chromatin protein modulated by acetylation. *EMBO J*, **21**, 4654-4662.

72. Biyani, K., Kahsai, M.A., Clark, A.T., Armstrong, T.L., Edmondson, S.P. and Shriver, J.W. (2005) Solution structure, stability, and nucleic acid binding of the hyperthermophile protein Sso10b2. *Biochemistry*, **44**, 14217-14230.
73. Chou, C.C., Lin, T.W., Chen, C.Y. and Wang, A.H. (2003) Crystal structure of the hyperthermophilic archaeal DNA-binding protein Sso10b2 at a resolution of 1.85 Angstroms. *J Bacteriol*, **185**, 4066-4073.
74. Cui, Q., Tong, Y., Xue, H., Huang, L., Feng, Y. and Wang, J. (2003) Two conformations of archaeal Ssh10b. The origin of its temperature-dependent interaction with DNA. *J Biol Chem*, **278**, 51015-51022.
75. Hada, K., Nakashima, T., Osawa, T., Shimada, H., Kakuta, Y. and Kimura, M. (2008) Crystal structure and functional analysis of an archaeal chromatin protein Alba from the hyperthermophilic archaeon *Pyrococcus horikoshii* OT3. *Biosci Biotechnol Biochem*, **72**, 749-758.
76. Jelinska, C., Conroy, M.J., Craven, C.J., Hounslow, A.M., Bullough, P.A., Waltho, J.P., Taylor, G.L. and White, M.F. (2005) Obligate heterodimerization of the archaeal Alba2 protein with Alba1 provides a mechanism for control of DNA packaging. *Structure*, **13**, 963-971.
77. Kumarevel, T., Sakamoto, K., Gopinath, S.C., Shinkai, A., Kumar, P.K. and Yokoyama, S. (2008) Crystal structure of an archaeal specific DNA-binding protein (Ape10b2) from *Aeropyrum pernix* K1. *Proteins*, **71**, 1156-1162.
78. Wang, G., Guo, R., Bartlam, M., Yang, H., Xue, H., Liu, Y., Huang, L. and Rao, Z. (2003) Crystal structure of a DNA binding protein from the hyperthermophilic euryarchaeon *Methanococcus jannaschii*. *Protein Sci*, **12**, 2815-2822.
79. Zhao, K., Chai, X. and Marmorstein, R. (2003) Structure of a Sir2 substrate, Alba, reveals a mechanism for deacetylation-induced enhancement of DNA binding. *J Biol Chem*, **278**, 26071-26077.
80. Tanaka, T., Padavattan, S. and Kumarevel, T. (2012) Crystal Structure of Archaeal Chromatin Protein Alba2-Double-stranded DNA Complex from *Aeropyrum pernix* K1. *J Biol Chem*, **287**, 10394-10402.
81. Lurz, R., Grote, M., Dijk, J., Reinhardt, R. and Dobrinski, B. (1986) Electron microscopic study of DNA complexes with proteins from the Archaebacterium *Sulfolobus acidocaldarius*. *EMBO J*, **5**, 3715-3721.

82. Xuan, J. and Feng, Y. (2012) The archaeal Sac10b protein family: conserved proteins with divergent functions. *Curr Protein Pept Sci*, **13**, 258-266.
83. 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.
84. Guo, R., Xue, H. and Huang, L. (2003) Ssh10b, a conserved thermophilic archaeal protein, binds RNA *in vivo*. *Mol Microbiol*, **50**, 1605-1615.
85. Aravind, L., Iyer, L.M. and Anantharaman, V. (2003) The two faces of Alba: the evolutionary connection between proteins participating in chromatin structure and RNA metabolism. *Genome Biol*, **4**, R64.
86. Hardy, C.D. and Martin, P.K. (2008) Biochemical characterization of DNA-binding proteins from *Pyrobaculum aerophilum* and Aeropyrum pernix. *Extremophiles*, **12**, 235-246.
87. Liu, Y., Guo, L., Guo, R., Wong, R.L., Hernandez, H., Hu, J., Chu, Y., Amster, I.J., Whitman, W.B. and Huang, L. (2009) The Sac10b homolog in *Methanococcus maripaludis* binds DNA at specific sites. *J Bacteriol*, **191**, 2315-2329.
88. Heinicke, I., Muller, J., Pittelkow, M. and Klein, A. (2004) Mutational analysis of genes encoding chromatin proteins in the archaeon *Methanococcus* voltae indicates their involvement in the regulation of gene expression. *Mol Genet Genomics*, **272**, 76-87.
89. Liu, Y.F., Zhang, N., Yao, H.W., Pan, X.M. and Ge, M. (2011) Mth10b, a unique member of the Sac10b family, does not bind nucleic acid. *PLoS One*, **6**, e19977.
90. Driessen, R.P.C., Meng, H., Suresh, G., Shahapure, R., Lanzani, G., Priyakumar, U.D., White, M.F., Schiessel, H., van Noort, J. and Dame, R.T. (2013) Crenarchaeal chromatin proteins Cren7 and Sul7 compact DNA by inducing rigid bends. *Nucleic Acids Res*, **41**, 196-205.
91. Edmondson, S.P., Qiu, L. and Shriver, J.W. (1995) Solution structure of the DNA-binding protein Sac7d from the hyperthermophile *Sulfolobus acidocaldarius*. *Biochemistry*, **34**, 13289-13304.
92. 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 Bacteriol*, **180**, 2560-2563.
93. 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.
94. Napoli, A., Zivanovic, Y., Bocs, C., Buhler, C., Rossi, M., Forterre, P. and Ciaramella, M. (2002) DNA bending, compaction and negative supercoiling by the architectural protein Sso7d of *Sulfolobus solfataricus*. *Nucleic Acids Res*, **30**, 2656-2662.
95. 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.
96. Feng, Y., Yao, H. and Wang, J. (2010) Crystal structure of the crenarchaeal conserved chromatin protein Cren7 and double-stranded DNA complex. *Protein Sci*, **19**, 1253-1257.
97. Gao, Y.G., Su, S.Y., Robinson, H., Padmanabhan, S., Lim, L., McCrary, B.S., Edmondson, S.P., Shriver, J.W. and Wang, A.H. (1998) The crystal structure of the hyperthermophile chromosomal protein Sso7d bound to DNA. *Nat Struct Biol*, **5**, 782-786.
98. Ko, T.P., Chu, H.M., Chen, C.Y., Chou, C.C. and Wang, A.H. (2004) Structures of the hyperthermophilic chromosomal protein Sac7d in complex with DNA decamers. *Acta Crystallogr D Biol Crystallogr*, **60**, 1381-1387.
99. Krueger, J.K., McCrary, B.S., Wang, A.H., Shriver, J.W., Trewhella, J. and Edmondson, S.P. (1999) The solution structure of the Sac7d/DNA complex: a small-angle X-ray scattering study. *Biochemistry*, **38**, 10247-10255.
100. 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.
101. Zhang, Z., Gong, Y., Guo, L., Jiang, T. and Huang, L. (2010) Structural insights into the interaction of the crenarchaeal chromatin protein Cren7 with DNA. *Mol Microbiol*, **76**, 749-759.
102. Chen, L., Zheng, Q.C., Yu, L.Y., Chu, W.T., Zhang, J.L., Xue, Q., Zhang, H.X. and Sun, C.C. (2012) Insights into the thermal stabilization and conformational transitions of DNA by hyperthermophile protein Sso7d: molecular dynamics simulations and MM-PBSA analysis. *J Biomol Struct Dyn*, **30**, 716-727.
103. Priyakumar, U.D., Harika, G. and Suresh, G. (2010) Molecular simulations on the thermal stabilization of DNA by hyperthermophilic chromatin protein Sac7d, and associated conformational transitions. *J Phys Chem B*, **114**, 16548-16557.
104. Chen, L., Zhang, J.L., Yu, L.Y., Zheng, Q.C., Chu, W.T., Xue, Q., Zhang, H.X. and Sun, C.C. (2012) Influence of hyperthermophilic

protein Cren7 on the stability and conformation of DNA: insights from molecular dynamics simulation and free energy analysis. *J Phys Chem B*, **116**, 12415-12425.
105. McAfee, J.G., Edmondson, S.P., Zegar, I. and Shriver, J.W. (1996) Equilibrium DNA binding of Sac7d protein from the hyperthermophile *Sulfolobus acidocaldarius*: fluorescence and circular dichroism studies. *Biochemistry*, **35**, 4034-4045.
106. Guagliardi, A., Cerchia, L., Moracci, M. and Rossi, M. (2000) The chromosomal protein sso7d of the crenarchaeon *Sulfolobus solfataricus* rescues aggregated proteins in an ATP hydrolysis-dependent manner. *J Biol Chem*, **275**, 31813-31818.
107. Guagliardi, A., Cerchia, L. and Rossi, M. (2002) The Sso7d protein of *Sulfolobus solfataricus*: *in vitro* relationship among different activities. *Archaea*, **1**, 87-93.
108. Tashiro, R., Wang, A.H. and Sugiyama, H. (2006) Photoreactivation of DNA by an archaeal nucleoprotein Sso7d. *Proc Natl Acad Sci U S A*, **103**, 16655-16659.
109. Edmondson, S.P., Kahsai, M.A., Gupta, R. and Shriver, J.W. (2004) Characterization of Sac10a, a hyperthermophile DNA-binding protein from *Sulfolobus acidocaldarius*. *Biochemistry*, **43**, 13026-13036.
110. 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.
111. 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 Bacteriol*, **189**, 403-409.
112. Dame, R.T., Kalmykowa, O.J. and Grainger, D.C. (2011) Chromosomal macrodomains and associated proteins: implications for DNA organization and replication in gram negative bacteria. *PLoS Genet*, **7**, e1002123.
113. Noom, M.C., Navarre, W.W., Oshima, T., Wuite, G.J. and Dame, R.T. (2007) H-NS promotes looped domain formation in the bacterial chromosome. *Curr Biol*, **17**, R913-914.
114. Amit, R., Oppenheim, A.B. and Stavans, J. (2003) Increased bending rigidity of single DNA molecules by H-NS, a temperature and osmolarity sensor. *Biophys J*, **84**, 2467-2473.
115. Liu, Y., Chen, H., Kenney, L.J. and Yan, J. (2010) A divalent switch drives H-NS/DNA-binding conformations between stiffening and bridging modes. *Genes Dev*, **24**, 339-344.

116. Wiggins, P.A., Dame, R.T., Noom, M.C. and Wuite, G.J. (2009) Protein-mediated molecular bridging: a key mechanism in biopolymer organization. *Biophys J*, **97**, 1997-2003.
117. Dorsett, D. (2011) Cohesin: genomic insights into controlling gene transcription and development. *Curr Opin Genet Dev*, **21**, 199-206.
118. Petrushenko, Z.M., Cui, Y., She, W. and Rybenkov, V.V. (2010) Mechanics of DNA bridging by bacterial condensin MukBEF *in vitro* and *in singulo*. *EMBO J*, **29**, 1126-1135.
119. Soppa, J. (2001) Prokaryotic structural maintenance of chromosomes (SMC) proteins: distribution, phylogeny, and comparison with MukBs and additional prokaryotic and eukaryotic coiled-coil proteins. *Gene*, **278**, 253-264.
120. Elie, C., Baucher, M.F., Fondrat, C. and Forterre, P. (1997) A protein related to eucaryal and bacterial DNA-motor proteins in the hyperthermophilic archaeon *Sulfolobus acidocaldarius*. *J Mol Evol*, **45**, 107-114.
121. Ali Azam, T., Iwata, A., Nishimura, A., Ueda, S. and Ishihama, A. (1999) Growth phase-dependent variation in protein composition of the *Escherichia coli* nucleoid. *J Bacteriol*, **181**, 6361-6370.
122. Cosgriff, S., Chintakayala, K., Chim, Y.T., Chen, X., Allen, S., Lovering, A.L. and Grainger, D.C. (2010) Dimerization and DNA-dependent aggregation of the *Escherichia coli* nucleoid protein and chaperone CbpA. *Mol Microbiol*, **77**, 1289-1300.
123. Frenkiel-Krispin, D., Ben-Avraham, I., Englander, J., Shimoni, E., Wolf, S.G. and Minsky, A. (2004) Nucleoid restructuring in stationary-state bacteria. *Mol Microbiol*, **51**, 395-405.
124. Sandman, K., Grayling, R.A., Dobrinski, B., Lurz, R. and Reeve, J.N. (1994) Growth-phase-dependent synthesis of histones in the archaeon *Methanothermus fervidus*. *Proc Natl Acad Sci U S A*, **91**, 12624-12628.
125. Lundgren, M. and Bernander, R. (2007) Genome-wide transcription map of an archaeal cell cycle. *Proc Natl Acad Sci U S A*, **104**, 2939-2944.
126. Poplawski, A. and Bernander, R. (1997) Nucleoid structure and distribution in thermophilic Archaea. *J Bacteriol*, **179**, 7625-7630.
127. Jun, S.H., Reichlen, M.J., Tajiri, M. and Murakami, K.S. (2011) Archaeal RNA polymerase and transcription regulation. *Crit Rev Biochem Mol Biol*, **46**, 27-40.
128. Thomm, M. (2007) Transcription: Mechanism and Regulation. *In: Cavicchioli R. (ed.) Archaea: Molecular and Cellular Biology.*, 139-157.

129. Aravind, L. and Koonin, E.V. (1999) DNA-binding proteins and evolution of transcription regulation in the archaea. *Nucleic Acids Res*, **27**, 4658-4670.
130. Kyrpides, N.C. and Ouzounis, C.A. (1999) Transcription in archaea. *Proc Natl Acad Sci U S A*, **96**, 8545-8550.
131. Bell, S.D. and Jackson, S.P. (2001) Mechanism and regulation of transcription in archaea. *Curr Opin Microbiol*, **4**, 208-213.
132. Hausner, W., Wettach, J., Hethke, C. and Thomm, M. (1996) Two transcription factors related with the eucaryal transcription factors TATA-binding protein and transcription factor IIB direct promoter recognition by an archaeal RNA polymerase. *J Biol Chem*, **271**, 30144-30148.
133. Qureshi, S.A., Bell, S.D. and Jackson, S.P. (1997) Factor requirements for transcription in the Archaeon *Sulfolobus shibatae*. *EMBO J*, **16**, 2927-2936.
134. Reiter, W.D., Palm, P. and Zillig, W. (1988) Analysis of transcription in the archaebacterium *Sulfolobus* indicates that archaebacterial promoters are homologous to eukaryotic pol II promoters. *Nucleic Acids Res*, **16**, 1-19.
135. Rowlands, T., Baumann, P. and Jackson, S.P. (1994) The TATA-binding protein: a general transcription factor in eukaryotes and archaebacteria. *Science*, **264**, 1326-1329.
136. Soppa, J. (1999) Normalized nucleotide frequencies allow the definition of archaeal promoter elements for different archaeal groups and reveal base-specific TFB contacts upstream of the TATA box. *Mol Microbiol*, **31**, 1589-1592.
137. Bell, S.D., Kosa, P.L., Sigler, P.B. and Jackson, S.P. (1999) Orientation of the transcription preinitiation complex in archaea. *Proc Natl Acad Sci U S A*, **96**, 13662-13667.
138. Bell, S.D. and Jackson, S.P. (2000) The role of transcription factor B in transcription initiation and promoter clearance in the archaeon *Sulfolobus acidocaldarius*. *J Biol Chem*, **275**, 12934-12940.
139. Bell, S.D., Brinkman, A.B., van der Oost, J. and Jackson, S.P. (2001) The archaeal TFIIEalpha homologue facilitates transcription initiation by enhancing TATA-box recognition. *EMBO Rep*, **2**, 133-138.
140. Werner, F. and Weinzierl, R.O. (2005) Direct modulation of RNA polymerase core functions by basal transcription factors. *Mol Cell Biol*, **25**, 8344-8355.

141. Grunberg, S., Bartlett, M.S., Naji, S. and Thomm, M. (2007) Transcription factor E is a part of transcription elongation complexes. *J Biol Chem*, **282**, 35482-35490.
142. Baliga, N.S., Goo, Y.A., Ng, W.V., Hood, L., Daniels, C.J. and DasSarma, S. (2000) Is gene expression in *Halobacterium* NRC-1 regulated by multiple TBP and TFB transcription factors? *Mol Microbiol*, **36**, 1184-1185.
143. Bell, S.D. (2005) Archaeal transcriptional regulation--variation on a bacterial theme? *Trends Microbiol*, **13**, 262-265.
144. Brinkman, A.B., Ettema, T.J., de Vos, W.M. and van der Oost, J. (2003) The Lrp family of transcriptional regulators. *Mol Microbiol*, **48**, 287-294.
145. Tani, T.H., Khodursky, A., Blumenthal, R.M., Brown, P.O. and Matthews, R.G. (2002) Adaptation to famine: a family of stationary-phase genes revealed by microarray analysis. *Proc Natl Acad Sci U S A*, **99**, 13471-13476.
146. Peeters, E. and Charlier, D. (2010) The Lrp family of transcription regulators in archaea. *Archaea*, **2010**, 750457.
147. Brinkman, A.B., Dahlke, I., Tuininga, J.E., Lammers, T., Dumay, V., de Heus, E., Lebbink, J.H., Thomm, M., de Vos, W.M. and van Der Oost, J. (2000) An Lrp-like transcriptional regulator from the archaeon *Pyrococcus furiosus* is negatively autoregulated. *J Biol Chem*, **275**, 38160-38169.
148. Dahlke, I. and Thomm, M. (2002) A *Pyrococcus* homolog of the leucine-responsive regulatory protein, LrpA, inhibits transcription by abrogating RNA polymerase recruitment. *Nucleic Acids Res*, **30**, 701-710.
149. Bell, S.D. and Jackson, S.P. (2000) Mechanism of autoregulation by an archaeal transcriptional repressor. *J Biol Chem*, **275**, 31624-31629.
150. Napoli, A., van der Oost, J., Sensen, C.W., Charlebois, R.L., Rossi, M. and Ciaramella, M. (1999) An Lrp-like protein of the hyperthermophilic archaeon *Sulfolobus solfataricus* which binds to its own promoter. *J Bacteriol*, **181**, 1474-1480.
151. Enoru-Eta, J., Gigot, D., Glansdorff, N. and Charlier, D. (2002) High resolution contact probing of the Lrp-like DNA-binding protein Ss-Lrp from the hyperthermoacidophilic crenarchaeote *Sulfolobus solfataricus* P2. *Mol Microbiol*, **45**, 1541-1555.
152. Enoru-Eta, J., Gigot, D., Thia-Toong, T.L., Glansdorff, N. and Charlier, D. (2000) Purification and characterization of Sa-lrp, a DNA-binding protein from the extreme thermoacidophilic

archaeon *Sulfolobus acidocaldarius* homologous to the bacterial global transcriptional regulator Lrp. *J Bacteriol*, **182**, 3661-3672.
153. Yokoyama, K., Ishijima, S.A., Koike, H., Kurihara, C., Shimowasa, A., Kabasawa, M., Kawashima, T. and Suzuki, M. (2007) Feast/famine regulation by transcription factor FL11 for the survival of the hyperthermophilic archaeon *Pyrococcus* OT3. *Structure*, **15**, 1542-1554.
154. Schwaiger, R., Schwarz, C., Furtwangler, K., Tarasov, V., Wende, A. and Oesterhelt, D. (2010) Transcriptional control by two leucine-responsive regulatory proteins in *Halobacterium salinarum* R1. *BMC Mol Biol*, **11**, 40.
155. Ouhammouch, M., Dewhurst, R.E., Hausner, W., Thomm, M. and Geiduschek, E.P. (2003) Activation of archaeal transcription by recruitment of the TATA-binding protein. *Proc Natl Acad Sci U S A*, **100**, 5097-5102.
156. Brinkman, A.B., Bell, S.D., Lebbink, R.J., de Vos, W.M. and van der Oost, J. (2002) The *Sulfolobus solfataricus* Lrp-like protein LysM regulates lysine biosynthesis in response to lysine availability. *J Biol Chem*, **277**, 29537-29549.
157. Peeters, E., van Oeffelen, L., Nadal, M., Forterre, P. and Charlier, D. (2013) A thermodynamic model of the cooperative interaction between the archaeal transcription factor Ss-LrpB and its tripartite operator DNA. *Gene*, **524**, 330-340.
158. Peeters, E., Willaert, R., Maes, D. and Charlier, D. (2006) Ss-LrpB from *Sulfolobus solfataricus* condenses about 100 base pairs of its own operator DNA into globular nucleoprotein complexes. *J Biol Chem*, **281**, 11721-11728.
159. Browning, D.F., Grainger, D.C. and Busby, S.J. (2010) Effects of nucleoid-associated proteins on bacterial chromosome structure and gene expression. *Curr Opin Microbiol*, **13**, 773-780.
160. Lopez-Garcia, P. and Forterre, P. (1999) Control of DNA topology during thermal stress in hyperthermophilic archaea: DNA topoisomerase levels, activities and induced thermotolerance during heat and cold shock in *Sulfolobus*. *Mol Microbiol*, **33**, 766-777.
161. Dorman, C.J. (1996) Flexible response: DNA supercoiling, transcription and bacterial adaptation to environmental stress. *Trends Microbiol*, **4**, 214-216.
162. Hatfield, G.W. and Benham, C.J. (2002) DNA topology-mediated control of global gene expression in *Escherichia coli*. *Annu Rev Genet*, **36**, 175-203.

163. Peter, B.J., Arsuaga, J., Breier, A.M., Khodursky, A.B., Brown, P.O. and Cozzarelli, N.R. (2004) Genomic transcriptional response to loss of chromosomal supercoiling in *Escherichia coli*. *Genome Biol*, **5**, R87.
164. Grainger, D.C., Hurd, D., Goldberg, M.D. and Busby, S.J. (2006) Association of nucleoid proteins with coding and non-coding segments of the *Escherichia coli* genome. *Nucleic Acids Res*, **34**, 4642-4652.
165. Fiorentino, G., Cannio, R., Rossi, M. and Bartolucci, S. (2003) Transcriptional regulation of the gene encoding an alcohol dehydrogenase in the archaeon *Sulfolobus solfataricus* involves multiple factors and control elements. *J Bacteriol*, **185**, 3926-3934.
166. Dorman, C.J. (2004) H-NS: a universal regulator for a dynamic genome. *Nat Rev Microbiol*, **2**, 391-400.
167. 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.
168. Marsh, V.L., McGeoch, A.T. and Bell, S.D. (2006) Influence of chromatin and single strand binding proteins on the activity of an archaeal MCM. *J Mol Biol*, **357**, 1345-1350.
169. Browning, D.F., Cole, J.A. and Busby, S.J. (2000) Suppression of FNR-dependent transcription activation at the *Escherichia coli* nir promoter by Fis, IHF and H-NS: modulation of transcription initiation by a complex nucleo-protein assembly. *Mol Microbiol*, **37**, 1258-1269.
170. Becker, N.A., Kahn, J.D. and Maher, L.J., 3rd. (2005) Bacterial repression loops require enhanced DNA flexibility. *J Mol Biol*, **349**, 716-730.
171. Becker, N.A., Kahn, J.D. and Maher, L.J., 3rd. (2007) Effects of nucleoid proteins on DNA repression loop formation in *Escherichia coli*. *Nucleic Acids Res*, **35**, 3988-4000.
172. Geanacopoulos, M., Vasmatzis, G., Zhurkin, V.B. and Adhya, S. (2001) Gal repressosome contains an antiparallel DNA loop. *Nat Struct Biol*, **8**, 432-436.
173. van Ulsen, P., Hillebrand, M., Zulianello, L., van de Putte, P. and Goosen, N. (1997) The integration host factor-DNA complex upstream of the early promoter of bacteriophage Mu is functionally symmetric. *J Bacteriol*, **179**, 3073-3075.
174. Eichler, J. and Adams, M.W. (2005) Posttranslational protein modification in Archaea. *Microbiol Mol Biol Rev*, **69**, 393-425.

175. Botting, C.H., Talbot, P., Paytubi, S. and White, M.F. (2010) Extensive lysine methylation in hyperthermophilic crenarchaea: potential implications for protein stability and recombinant enzymes. *Archaea*, **2010**.
176. Chu, Y., Zhang, Z., Wang, Q., Luo, Y. and Huang, L. (2012) Identification and characterization of a highly conserved crenarchaeal protein lysine methyltransferase with broad substrate specificity. *J Bacteriol*, **194**, 6917-6926.
177. Niu, Y., Xia, Y., Wang, S., Li, J., Niu, C., Li, X., Zhao, Y., Xiong, H., Li, Z., Lou, H. *et al.* (2013) A prototypic lysine methyltransferase 4 from archaea with degenerate sequence specificity methylates chromatin proteins Sul7d and Cren7 in different patterns. *J Biol Chem*, **288**, 13728-13740.
178. Laurens, N., Driessen, R.P.C., Heller, I., Vorselen, D., Noom, M.C., Hol, F.J.H., White, M.F., Dame, R.T. and Wuite, G.J.L. (2012) Alba shapes the archaeal genome using a delicate balance of bridging and stiffening the DNA. *Nat Commun*, **3**, 1328.
179. Chaurasiya, K.R., Paramanathan, T., Mccauley, M.J. and Williams, M.C. (2010) Biophysical characterization of DNA binding from single molecule force measurements. *Physics of Life Reviews*, 1-43.
180. Neuman, K.C. and Nagy, A. (2008) Single-molecule force spectroscopy: optical tweezers, magnetic tweezers and atomic force microscopy. *Nat Methods*, **5**, 491-505.
181. Dame, R.T. (2008) Single-molecule micromanipulation studies of DNA and architectural proteins. *Biochem Soc Trans*, **36**, 732-737.
182. Segall, D.E., Nelson, P.C. and Phillips, R. (2006) Volume-exclusion effects in tethered-particle experiments: bead size matters. *Phys Rev Lett*, **96**, 088306.
183. Sandman, K. and Reeve, J.N. (2005) Archaeal chromatin proteins: different structures but common function? *Curr Opin Microbiol*, **8**, 656-661.
184. Cubonova, L., Sandman, K., Hallam, S.J., Delong, E.F. and Reeve, J.N. (2005) Histones in crenarchaea. *J Bacteriol*, **187**, 5482-5485.
185. Bessa, E., Wintraecken, K., Geerling, A. and de Vries, R. (2007) Synergy of DNA-bending nucleoid proteins and macromolecular crowding in condensing DNA. *Biophysical Reviews and Letters*, **2**, 1-7.
186. Swinger, K.K. and Rice, P.A. (2004) IHF and HU: flexible architects of bent DNA. *Curr Opin Struct Biol*, **14**, 28-35.

187. Thomas, J.O. and Travers, A.A. (2001) HMG1 and 2, and related 'architectural' DNA-binding proteins. *Trends Biochem Sci*, **26**, 167-174.
188. Zhang, J., McCauley, M.J., Maher, L.J., 3rd, Williams, M.C. and Israeloff, N.E. (2009) Mechanism of DNA flexibility enhancement by HMGB proteins. *Nucleic Acids Res*, **37**, 1107-1114.
189. Humphrey, W., Dalke, A. and Schulten, K. (1996) VMD: visual molecular dynamics. *J Mol Graph*, **14**, 33-38, 27-38.
190. Lavery, R., Moakher, M., Maddocks, J.H., Petkeviciute, D. and Zakrzewska, K. (2009) Conformational analysis of nucleic acids revisited: Curves+. *Nucleic Acids Res*, **37**, 5917-5929.
191. Kulić, I.M., Mohrbach, H., Thaokar, R. and Schiessel, H. (2007) Equation of state of looped DNA. *Physical Review E*, **75**, 011913.
192. Kulić, I.M., Mohrbach, H., Lobaskin, V., Thaokar, R. and Schiessel, H. (2005) Apparent persistence length renormalization of bent DNA. *Physical Review E*, **72**, 041905.
193. Ali, B.M., Amit, R., Braslavsky, I., Oppenheim, A.B., Gileadi, O. and Stavans, J. (2001) Compaction of single DNA molecules induced by binding of integration host factor (IHF). *Proc Natl Acad Sci U S A*, **98**, 10658-10663.
194. McCauley, M., Hardwidge, P.R., Maher, L.J., 3rd and Williams, M.C. (2005) Dual binding modes for an HMG domain from human HMGB2 on DNA. *Biophys J*, **89**, 353-364.
195. Skoko, D., Wong, B., Johnson, R.C. and Marko, J.F. (2004) Micromechanical analysis of the binding of DNA-bending proteins HMGB1, NHP6A, and HU reveals their ability to form highly stable DNA-protein complexes. *Biochemistry*, **43**, 13867-13874.
196. van Noort, J., Verbrugge, S., Goosen, N., Dekker, C. and Dame, R.T. (2004) Dual architectural roles of HU: formation of flexible hinges and rigid filaments. *Proc Natl Acad Sci USA*, **101**, 6969-6974.
197. Holbrook, J.A., Tsodikov, O.V., Saecker, R.M. and Record, M.T., Jr. (2001) Specific and non-specific interactions of integration host factor with DNA: thermodynamic evidence for disruption of multiple IHF surface salt-bridges coupled to DNA binding. *J Mol Biol*, **310**, 379-401.
198. Xiao, B., Johnson, R.C. and Marko, J.F. (2010) Modulation of HU-DNA interactions by salt concentration and applied force. *Nucleic Acids Res*, **38**, 6176–6185.

199. Lundbäck, T. and Härd, T. (1996) Salt Dependence of the Free Energy, Enthalpy, and Entropy of Nonsequence Specific DNA Binding. *The Journal of Physical Chemistry*, **100**, 17690-17695.
200. Dame, R.T., Wyman, C. and Goosen, N. (2000) H-NS mediated compaction of DNA visualised by atomic force microscopy. *Nucleic Acids Res*, **28**, 3504-3510.
201. Dame, R.T., Noom, M.C. and Wuite, G.J. (2006) Bacterial chromatin organization by H-NS protein unravelled using dual DNA manipulation. *Nature*, **444**, 387-390.
202. Lim, C.J., Whang, Y.R., Kenney, L.J. and Yan, J. (2012) Gene silencing H-NS paralogue StpA forms a rigid protein filament along DNA that blocks DNA accessibility. *Nucleic Acids Res*, **40**, 3316-3328.
203. Graham, J.S., Johnson, R.C. and Marko, J.F. (2010) Concentration-dependent exchange accelerates turnover of proteins bound to double-stranded DNA. *Nucleic Acids Res*.
204. Bustamante, C., Cheng, W. and Mejia, Y.X. (2011) Revisiting the central dogma one molecule at a time. *Cell*, **144**, 480-497.
205. Liebesny, P., Goyal, S., Dunlap, D., Family, F. and Finzi, L. (2010) Determination of the number of proteins bound non-specifically to DNA. *J Phys Condens Matter*, **22**, 414104.
206. Xiao, B., Zhang, H., Johnson, R.C. and Marko, J.F. (2011) Force-driven unbinding of proteins HU and Fis from DNA quantified using a thermodynamic Maxwell relation. *Nucleic Acids Res*, **39**, 5568–5577.
207. Yan, J. and Marko, J.F. (2003) Effects of DNA-distorting proteins on DNA elastic response. *Phys Rev E Stat Nonlin Soft Matter Phys*, **68**, 011905.
208. Zhang, H. and Marko, J.F. (2008) Maxwell relations for single-DNA experiments: Monitoring protein binding and double-helix torque with force-extension measurements. *Physical Review E*, **77**, 031916.
209. Swinger, K.K., Lemberg, K.M., Zhang, Y. and Rice, P.A. (2003) Flexible DNA bending in HU-DNA cocrystal structures. *EMBO J*, **22**, 3749-3760.
210. van Noort, J., Verbrugge, S., Goosen, N., Dekker, C. and Dame, R.T. (2004) Dual architectural roles of HU: formation of flexible hinges and rigid filaments. *Proc Natl Acad Sci U S A*, **101**, 6969-6974.
211. Candelli, A., Wuite, G.J. and Peterman, E.J. (2011) Combining optical trapping, fluorescence microscopy and micro-fluidics for

single molecule studies of DNA-protein interactions. *Phys Chem Chem Phys*, **13**, 7263-7272.
212. Kruithof, M., Chien, F., de Jager, M. and van Noort, J. (2008) Subpiconewton dynamic force spectroscopy using magnetic tweezers. *Biophys J*, **94**, 2343-2348.
213. Strick, T.R., Allemand, J.F., Bensimon, D., Bensimon, A. and Croquette, V. (1996) The elasticity of a single supercoiled DNA molecule. *Science*, **271**, 1835-1837.
214. Marko, J.F. and Siggia, E.D. (1995) Stretching DNA. *Macromolecules*, **28**, 8759-8770.
215. Brooks, B.R., Brooks, C.L., 3rd, Mackerell, A.D., Jr., Nilsson, L., Petrella, R.J., Roux, B., Won, Y., Archontis, G., Bartels, C., Boresch, S. *et al.* (2009) CHARMM: the biomolecular simulation program. *J Comput Chem*, **30**, 1545-1614.
216. Foloppe, N. and MacKerell, J.A.D. (2000) All-atom empirical force field for nucleic acids: I. Parameter optimization based on small molecule and condensed phase macromolecular target data. *Journal of Computational Chemistry*, **21**, 86-104.
217. MacKerell, A.D., Bashford, D., Bellott, Dunbrack, R.L., Evanseck, J.D., Field, M.J., Fischer, S., Gao, J., Guo, H., Ha, S. *et al.* (1998) All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *The Journal of Physical Chemistry B*, **102**, 3586-3616.
218. MacKerell, A.D. and Banavali, N.K. (2000) All-atom empirical force field for nucleic acids: II. Application to molecular dynamics simulations of DNA and RNA in solution. *Journal of Computational Chemistry*, **21**, 105-120.
219. Mackerell, A.D., Jr., Feig, M. and Brooks, C.L., 3rd. (2004) Extending the treatment of backbone energetics in protein force fields: limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *J Comput Chem*, **25**, 1400-1415.
220. Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W. and Klein, M.L. (1983) Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics*, **79**, 926-935.
221. Ryckaert, J.P., Ciccotti, G. and Berendsen, H.J.C. (1977) Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes. *Journal of Computational Physics*, **23**, 327-341.

222. Darden, T., Perera, L., Li, L. and Pedersen, L. (1999) New tricks for modelers from the crystallography toolkit: the particle mesh Ewald algorithm and its use in nucleic acid simulations. *Structure*, **7**, R55-60.
223. Essmann, U., Perera, L., Berkowitz, M., Darden, T., Lee, H. and Pedersen, L. (1995) A smooth particle mesh Ewald method. *The Journal of Chemical Physics*, **103**, 8577-8593.
224. Hoover, W.G. (1985) Canonical dynamics: Equilibrium phase-space distributions. *Phys Rev A*, **31**, 1695-1697.
225. Feller, S.E., Zhang, Y., Pastor, R.W. and Brooks, B.R. (1995) Constant pressure molecular dynamics simulation: The Langevin piston method. *The Journal of Chemical Physics*, **103**, 4613-4621.
226. Peters, J.P., 3rd and Maher, L.J. (2010) DNA curvature and flexibility *in vitro* and *in vivo*. *Q Rev Biophys*, **43**, 23-63.
227. Williams, L.D. and Maher, L.J., 3rd. (2000) Electrostatic mechanisms of DNA deformation. *Annu Rev Biophys Biomol Struct*, **29**, 497-521.
228. Garcia, H.G., Grayson, P., Han, L., Inamdar, M., Kondev, J., Nelson, P.C., Phillips, R., Widom, J. and Wiggins, P.A. (2007) Biological consequences of tightly bent DNA: The other life of a macromolecular celebrity. *Biopolymers*, **85**, 115-130.
229. Driessen, R.P.C. and Dame, R.T. (2013) Structure and dynamics of the crenarchaeal nucleoid. *Biochem Soc Trans*, **41**, 321-325.
230. Perez-Martin, J., Rojo, F. and de Lorenzo, V. (1994) Promoters responsive to DNA bending: a common theme in prokaryotic gene expression. *Microbiol Rev*, **58**, 268-290.
231. Saiz, L. and Vilar, J.M. (2006) DNA looping: the consequences and its control. *Curr Opin Struct Biol*, **16**, 344-350.
232. Cloutier, T.E. and Widom, J. (2004) Spontaneous sharp bending of double-stranded DNA. *Mol Cell*, **14**, 355-362.
233. Vafabakhsh, R. and Ha, T. (2012) Extreme bendability of DNA less than 100 base pairs long revealed by single-molecule cyclization. *Science*, **337**, 1097-1101.
234. Altan-Bonnet, G., Libchaber, A. and Krichevsky, O. (2003) Bubble dynamics in double-stranded DNA. *Phys Rev Lett*, **90**, 138101.
235. Forties, R.A., Bundschuh, R. and Poirier, M.G. (2009) The flexibility of locally melted DNA. *Nucleic Acids Res*, **37**, 4580-4586.
236. Vologodskii, A. and M, D.F.-K. (2013) Strong bending of the DNA double helix. *Nucleic Acids Res*.

237. Wiggins, P.A., Phillips, R. and Nelson, P.C. (2005) Exact theory of kinkable elastic polymers. *Phys Rev E Stat Nonlin Soft Matter Phys*, **71**, 021909.
238. Yan, J. and Marko, J.F. (2004) Localized single-stranded bubble mechanism for cyclization of short double helix DNA. *Phys Rev Lett*, **93**, 108108.
239. Cloutier, T.E. and Widom, J. (2005) DNA twisting flexibility and the formation of sharply looped protein-DNA complexes. *Proc Natl Acad Sci U S A*, **102**, 3645-3650.
240. Geggier, S., Kotlyar, A. and Vologodskii, A. (2011) Temperature dependence of DNA persistence length. *Nucleic Acids Res*, **39**, 1419-1426.
241. Gadgil, M., Kapur, V. and Hu, W.S. (2005) Transcriptional response of *Escherichia coli* to temperature shift. *Biotechnol Prog*, **21**, 689-699.
242. Tachdjian, S. and Kelly, R.M. (2006) Dynamic metabolic adjustments and genome plasticity are implicated in the heat shock response of the extremely thermoacidophilic archaeon *Sulfolobus solfataricus*. *J Bacteriol*, **188**, 4553-4559.
243. Battesti, A., Majdalani, N. and Gottesman, S. (2011) The RpoS-mediated general stress response in *Escherichia coli*. *Annual review of microbiology*, **65**, 189-213.
244. Geggier, S. and Vologodskii, A. (2010) Sequence dependence of DNA bending rigidity. *Proc Natl Acad Sci U S A*, **107**, 15421-15426.
245. Lee, N.K., Park, J.S., Johner, A., Obukhov, S., Hyon, J.Y., Lee, K.J. and Hong, S.C. (2008) Elasticity of cisplatin-bound DNA reveals the degree of cisplatin binding. *Physical review letters*, **101**, 248101.
246. Galburt, E.A., Tomko, E.J., Stump, W.T. and Ruiz Manzano, A. (2014) Force-dependent melting of supercoiled DNA at thermophilic temperatures. *Biophysical chemistry*, **187-188**, 23-28.
247. McGhee, J.D. and von Hippel, P.H. (1974) Theoretical aspects of DNA-protein interactions: co-operative and non-co-operative binding of large ligands to a one-dimensional homogeneous lattice. *J Mol Biol*, **86**, 469-489.
248. Nir, G., Lindner, M., Dietrich, H.R., Girshevitz, O., Vorgias, C.E. and Garini, Y. (2011) HU protein induces incoherent DNA persistence length. *Biophys J*, **100**, 784-790.
249. Dame, R.T., Hall, M.A. and Wang, M.D. (2013) Single-Molecule Unzipping Force Analysis of HU-DNA Complexes. *Chembiochem*, **14**, 1954-1957.

250. Yakovchuk, P., Protozanova, E. and Frank-Kamenetskii, M.D. (2006) Base-stacking and base-pairing contributions into thermal stability of the DNA double helix. *Nucleic Acids Res*, **34**, 564-574.
251. Peters, J.P., Yelgaonkar, S.P., Srivatsan, S.G., Tor, Y. and James Maher, L., 3rd. (2013) Mechanical properties of DNA-like polymers. *Nucleic Acids Res*.
252. Meyer, S., Jost, D., Theodorakopoulos, N., Peyrard, M., Lavery, R. and Everaers, R. (2013) Temperature dependence of the DNA double helix at the nanoscale: structure, elasticity, and fluctuations. *Biophys J*, **105**, 1904-1914.
253. Peters, W.B., Edmondson, S.P. and Shriver, J.W. (2004) Thermodynamics of DNA binding and distortion by the hyperthermophile chromatin protein Sac7d. *J Mol Biol*, **343**, 339-360.
254. Travers, A. (2013) Dynamic DNA underpins chromosome dynamics. *Biophys J*, **105**, 2235-2237.
255. Dame, R.T. (2003) Ph.D. thesis. *Leiden Univ., Leiden, The Netherlands*.
256. Moolenaar, G.F., Visse, R., Ortiz-Buysse, M., Goosen, N. and van de Putte, P. (1994) Helicase motifs V and VI of the *Escherichia coli* UvrB protein of the UvrABC endonuclease are essential for the formation of the preincision complex. *J Mol Biol*, **240**, 294-307.
257. Ramos, C.R., Abreu, P.A., Nascimento, A.L. and Ho, P.L. (2004) A high-copy T7 *Escherichia coli* expression vector for the production of recombinant proteins with a minimal N-terminal His-tagged fusion peptide. *Braz J Med Biol Res*, **37**, 1103-1109.
258. Bentley, S.D., Chater, K.F., Cerdeno-Tarraga, A.M., Challis, G.L., Thomson, N.R., James, K.D., Harris, D.E., Quail, M.A., Kieser, H., Harper, D. *et al.* (2002) Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature*, **417**, 141-147.
259. Laurens, N., Bellamy, S.R., Harms, A.F., Kovacheva, Y.S., Halford, S.E. and Wuite, G.J. (2009) Dissecting protein-induced DNA looping dynamics in real time. *Nucleic Acids Res*, **37**, 5454-5464.
260. Reif, R. (1985) The Fundamentals of Statistical and Thermal Physics. *McGraw-Hill*.
261. Seol, Y., Li, J., Nelson, P.C., Perkins, T.T. and Betterton, M.D. (2007) Elasticity of short DNA molecules: theory and experiment for contour lengths of 0.6-7 microm. *Biophys J*, **93**, 4360-4373.
262. Risken, H. (1989) The Fokker-Planck Equation: Methods of Solution and Applications. *Springer*.

263. Vlahovicek, K., Kajan, L. and Pongor, S. (2003) DNA analysis servers: plot.it, bend.it, model.it and IS. *Nucleic Acids Res*, **31**, 3686-3687.
264. Odijk, T. (1998) Osmotic compaction of supercoiled DNA into a bacterial nucleoid. *Biophys Chem*, **73**, 23-29.
265. Misteli, T. (2010) Higher-order genome organization in human disease. *Cold Spring Harb Perspect Biol*, **2**, a000794.
266. Tark-Dame, M., van Driel, R. and Heermann, D.W. (2011) Chromatin folding--from biology to polymer models and back. *J Cell Sci*, **124**, 839-845.
267. Zlatanova, J. and Caiafa, P. (2009) CCCTC-binding factor: to loop or to bridge. *Cell Mol Life Sci*, **66**, 1647-1660.
268. Farge, G., Laurens, N., Broekmans, O.D., van den Wildenberg, S.M., Dekker, L.C., Gaspari, M., Gustafsson, C.M., Peterman, E.J., Falkenberg, M. and Wuite, G.J. (2012) Protein sliding and DNA denaturation are essential for DNA organization by human mitochondrial transcription factor A. *Nat Commun*, **3**, 1013.
269. Kucej, M. and Butow, R.A. (2007) Evolutionary tinkering with mitochondrial nucleoids. *Trends Cell Biol*, **17**, 586-592.
270. Sandman, K., Krzycki, J.A., Dobrinski, B., Lurz, R. and Reeve, J.N. (1990) HMf, a DNA-binding protein isolated from the hyperthermophilic archaeon *Methanothermus fervidus*, is most closely related to histones. *Proc Natl Acad Sci USA*, **87**, 5788-5791.
271. Jelinska, C., Petrovic-Stojanovska, B., Ingledew, W.J. and White, M.F. (2010) Dimer-dimer stacking interactions are important for nucleic acid binding by the archaeal chromatin protein Alba. *Biochem J*, **427**, 49-55.
272. Gross, P., Laurens, N., Oddershede, L.B., Bockelmann, U., Peterman, E.J.G. and Wuite, G.J.L. (2011) Quantifying how DNA stretches, melts and changes twist under tension. *Nat Phys*, **7**, 731-736.
273. Winardhi, R.S., Fu, W., Castang, S., Li, Y., Dove, S.L. and Yan, J. (2012) Higher order oligomerization is required for H-NS family member MvaT to form gene-silencing nucleoprotein filament. *Nucleic Acids Res*.
274. Higgins, N.P., Yang, X., Fu, Q. and Roth, J.R. (1996) Surveying a supercoil domain by using the gamma delta resolution system in *Salmonella typhimurium*. *J Bacteriol*, **178**, 2825-2835.
275. Postow, L., Hardy, C.D., Arsuaga, J. and Cozzarelli, N.R. (2004) Topological domain structure of the *Escherichia coli* chromosome. *Genes Dev*, **18**, 1766-1779.

276. Noom, M.C., van den Broek, B., van Mameren, J.v. and Wuite, G.J.L. (2007) Visualizing single DNA-bound proteins using DNA as a scanning probe. *Nat Methods*, **4**, 1031-1036.
277. Mangeol, P. and Bockelmann, U. (2008) Interference and crosstalk in double optical tweezers using a single laser source. *Rev Sci Instrum*, **79**, 083103.
278. Odijk, T. (1995) Stiff Chains and Filaments under Tension. *Macromolecules*, **28**, 7016-7018.
279. Brennan, R.G. (1993) The winged-helix DNA-binding motif: another helix-turn-helix takeoff. *Cell*, **74**, 773-776.
280. Gajiwala, K.S. and Burley, S.K. (2000) Winged helix proteins. *Curr Opin Struct Biol*, **10**, 110-116.
281. Harrison, S.C. and Aggarwal, A.K. (1990) DNA recognition by proteins with the helix-turn-helix motif. *Annu Rev Biochem*, **59**, 933-969.
282. Waterreus, W.J. (2013) Ph.D. thesis. *Leiden Univ., Leiden, The Netherlands*.
283. Suzuki, M., Yagi, N. and Gerstein, M. (1995) DNA recognition and superstructure formation by helix-turn-helix proteins. *Protein Eng*, **8**, 329-338.
284. Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Soding, J. *et al.* (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol*, **7**, 539.
285. Dame, R.T. and Goosen, N. (2002) HU: promoting or counteracting DNA compaction? *FEBS Lett*, **529**, 151-156.
286. Krissinel, E. and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol*, **372**, 774-797.
287. Clark, K.L., Halay, E.D., Lai, E. and Burley, S.K. (1993) Co-crystal structure of the HNF-3/fork head DNA-recognition motif resembles histone H5. *Nature*, **364**, 412-420.
288. Kodandapani, R., Pio, F., Ni, C.Z., Piccialli, G., Klemsz, M., McKercher, S., Maki, R.A. and Ely, K.R. (1996) A new pattern for helix-turn-helix recognition revealed by the PU.1 ETS-domain-DNA complex. *Nature*, **380**, 456-460.
289. Harbury, P.B., Zhang, T., Kim, P.S. and Alber, T. (1993) A switch between two-, three-, and four-stranded coiled coils in GCN4 leucine zipper mutants. *Science*, **262**, 1401-1407.
290. Liu, J., Zheng, Q., Deng, Y., Li, Q., Kallenbach, N.R. and Lu, M. (2007) Conformational specificity of the lac repressor coiled-coil tetramerization domain. *Biochemistry*, **46**, 14951-14959.

291. Stetefeld, J., Jenny, M., Schulthess, T., Landwehr, R., Engel, J. and Kammerer, R.A. (2000) Crystal structure of a naturally occurring parallel right-handed coiled coil tetramer. *Nat Struct Biol*, **7**, 772-776.
292. Littlefield, O. and Nelson, H.C. (1999) A new use for the 'wing' of the 'winged' helix-turn-helix motif in the HSF-DNA cocrystal. *Nat Struct Biol*, **6**, 464-470.
293. Simonis, M., Kooren, J. and de Laat, W. (2007) An evaluation of 3C-based methods to capture DNA interactions. *Nat Methods*, **4**, 895-901.
294. Farkas, J.A., Picking, J.W. and Santangelo, T.J. (2013) Genetic techniques for the archaea. *Annu Rev Genet*, **47**, 539-561.

CURRICULUM VITAE

16-06-1983	Born in Utrecht, The Netherlands
1995-2001	High School, Christelijk Gymnasium Utrecht
2002-2009	M.Sc. Applied Physics, Delft University of Technology
2009	Internship JPK Instruments, Berlin, Germany
2010-2014	Ph.D research under supervision of dr. Remus Th. Dame
	"The architects of crenarchaeal chromatin – a biophysical characterization of chromatin proteins from *Sulfolobus solfataricus*"
	Leiden University, The Netherlands

LIST OF PUBLICATIONS

Rajesh Shahapure, **Rosalie P.C. Driessen**, M. Florencia Haurat, Sonja-Verena Albers and Remus Th. Dame, The archaellum: a rotating type IV pilus. Mol Microbiol, 2014

Rosalie P.C. Driessen and Remus Th. Dame, Structure and dynamics of the crenarchaeal nucleoid, Biochem. Soc. Trans., 2013

Niels Laurens*, **Rosalie P.C. Driessen***, Iddo Heller, Daan Vorselen, Maarten C. Noom, Felix J.H. Hol, Malcolm F. White, Remus Th. Dame, Gijs J.L. Wuite, Alba shapes the archaeal genome using a delicate balance of bridging and stiffening the DNA, Nature Comm., 2012

Rosalie P.C. Driessen, He Meng, Gorle Suresh, Rajesh Shahapure, Giovanni Lanzani, U Deva Priyakumar, Malcolm F. White, Helmut Schiessel, John van Noort, and Remus Th. Dame, Crenarchaeal chromatin proteins Cren7 and Sul7 compact DNA by inducing rigid bends, Nucleic Acid Res, 2012.

R.P. C. Driessen, R. Th Dame, Nucleoid-associated proteins in Crenarchaea, Biochem. Soc. Trans., 2011.

J. Männik, **R. Driessen**, P. Galajda, J. E. Keymer & C. Dekker, Bacterial growth and motility in sub-micron constrictions, PNAS, 2009.

* These authors contributed equally