Histone-DNA assemblies in archaea : shaping the genome on the edge of life

Henneman, B.

CHAPTER 6

Specific DNA binding of archaeal histones

Abstract

In archaea, histones play a role in genome compaction and are involved in transcription regulation. Although no specific sequence is required for DNA binding, archaeal histones prefer binding to DNA containing repeats of alternating A/T and G/C motifs. These motifs are also present on an artificial sequence Clone20, a high affinity model sequence for binding of the histones from *Methanothermus fervidus*. Here, we investigate binding of HMfA and HMfB to Clone20 DNA. We show that specific binding at low protein concentrations (<22 nM) modestly compacts DNA, whereas nonspecific binding strongly compacts DNA. We propose mechanisms by which high affinity sequences on the genome may interplay with other components that are likely involved in archaeal transcription regulation, such as histone variants and the hypernucleosome.

Archaea and eukaryotes express histones, which are involved in genome compaction and transcription regulation (177, 237). Together with other architectural proteins, such as Alba and MC1, archaeal histones are hypothesized to function as transcription regulators, although the mechanisms employed are currently unknown (117). The hypernucleosome, a rod-like structure consisting of a protein core of an archaeal histone multimer (multiple associated dimers), with DNA wrapped around compacts DNA, and may play a key role in archaeal transcription regulation (134, 197) (*see also* CHAPTER 4). Histones from species throughout the archaeal domain of life are predicted to be able to assemble into hypernucleosomes, although some genomes also (sometimes exclusively) encode histones that likely cannot multimerize (134). Archaeal histones have also been proposed to have defined roles as dimers or tetramers (196, 238). The dimer is the smallest functional unit of archaeal histones, and may be formed by monomers of the same or of different histone variants. Homo- and heterodimerization of histones could be key to modulating DNA compaction and transcription (251). Archaeal genomes harbor up to 11 histone variants, which allows archaea to vary the composition of histone dimers, tetramers and hypernucleosomes. Each form may have different functionality, with its formation dependent on growth phase or environmental cues (171, 251, 305).

Histones, like most archaeal chromatin proteins, bind DNA in a nonspecific way. However, histones have been suggested to be positioned at specific locations on the genome (196, 347). Also, histones are depleted from transcription start sites (TSSs), which is common in most unicellular organisms (221, 246). These TSSs are AT-rich, which disfavors nucleosome formation in both archaea and eukaryotes (221, 348). HMfB, one of the two histones from the archaeon *Methanothermus fervidus*, was reported to preferentially bind GC-rich sequences with alternating GC and AT motifs (164, 349). Specifically, (T)3(A)3GCCG repeat sequences exhibit a very high affinity towards *M. fervidus* histones (350). In *Haloferax volcanii*, patterns of GC and AT were shown to position histone tetramers in a similar way as reported for eukaryotes (220, 351). Sequences to which HMfB binds with high affinity were selected in early studies by Reeve and co-workers using a systematic evolution of ligands by exponential enrichment (SELEX) approach (164). Especially Clone20, a sequence resulting from this SELEX experiment, was shown to have a very high binding affinity for HMfB tetramers (an increase of two orders of magnitude compared to control DNA for HMfB, while the increase in affinity of HMfA was less than twofold, as measured via electrophoretic mobility shift assay [EMSA])(160). Clone20 contains repeats of alternating A/T- and

G/C-rich regions (*see* **Materials & Methods**). Clone20 can be regarded as the archaeal counterpart of the nucleosome positioning sequence 601, a sequence that energetically favors nucleosome formation, which is used in numerous studies on eukaryotic histones (320, 352-354). DNA sequences with high similarity to Clone20 and 601 have thus far not been identified in genomes, but high affinity sequences may be present on archaeal genomes (163, 221). High affinity sequences may be involved in positioning histones at or close to regulatory DNA elements, by means of which they could have an effect on transcription regulation.

Nonspecific DNA is cooperatively bound and compacted by HMfA and HMfB. Although affinity of HMfA and HMfB to Clone20 is high, it is unclear whether high affinity sequences play specific roles in DNA compaction, for instance as nucleation sites. Here, we use tethered particle motion (TPM) to investigate the role of the Clone20 sequence in DNA compaction by *M. fervidus* histones. The experiments were carried out as described elsewhere (195), using a linear DNA substrate with the Clone20 sequence at its center (*see* **Materials and Methods**).

Our experiments show that both HMfA and HMfB compact the Clone20 substrate, indicated by reduction of the root mean square displacement (RMS) of the DNA tether in TPM (**figure 6.1A** and **B**). Compaction occurs in two steps. The strong decrease in RMS at high protein concentrations (at >18 nM for HMfA and >22 nM for HMfB) is similar to that observed for strong cooperative compaction of nonspecific DNA into a hypernucleosome, as has been shown in chapter 4 of this thesis. (**figure S6.1**). Therefore we conclude that specific binding to the Clone20 sequence at concentrations <18 nM does not increase the effective binding affinity of the hypernucleosome, disfavoring a mechanism of site-specific nucleation, followed by propagation. Compaction is both somewhat more cooperative (reflected in a steeper slope) and stronger for HMfB than HMfA on the Clone20 substrate (an RMS of ~77 vs ~72 nm, **figure 6.1C**). Similar behavior was observed for nonspecific compaction on a DNA substrate without high affinity sequence (**figure S6.1**). The compaction at low protein concentrations (5-10 nM for HMfA and 10-22 nM for HMfB) was not observed for nonspecific DNA, and is attributed to specific binding of HMfA and HMfB to the Clone20 sequence. In this concentration regime, two populations of tethers with RMS values of ~155 nm and ~134 nm can be observed (**figure 6.1A** and **B**). The population sizes were used to calculate Clone20 sequence occupancy as a function of protein concentration, and a Hill equation was fit (**figure 6.1D**). This yielded a significantly lower binding constant (K_D) for binding to this sequence for HMfA than HMfB (5.7 ± 1.6 nM and 14.6 ± 1.6 nM, respectively). The higher affinity of HMfA for the Clone20 sequence compared to HMfB is unexpected as the sequence was

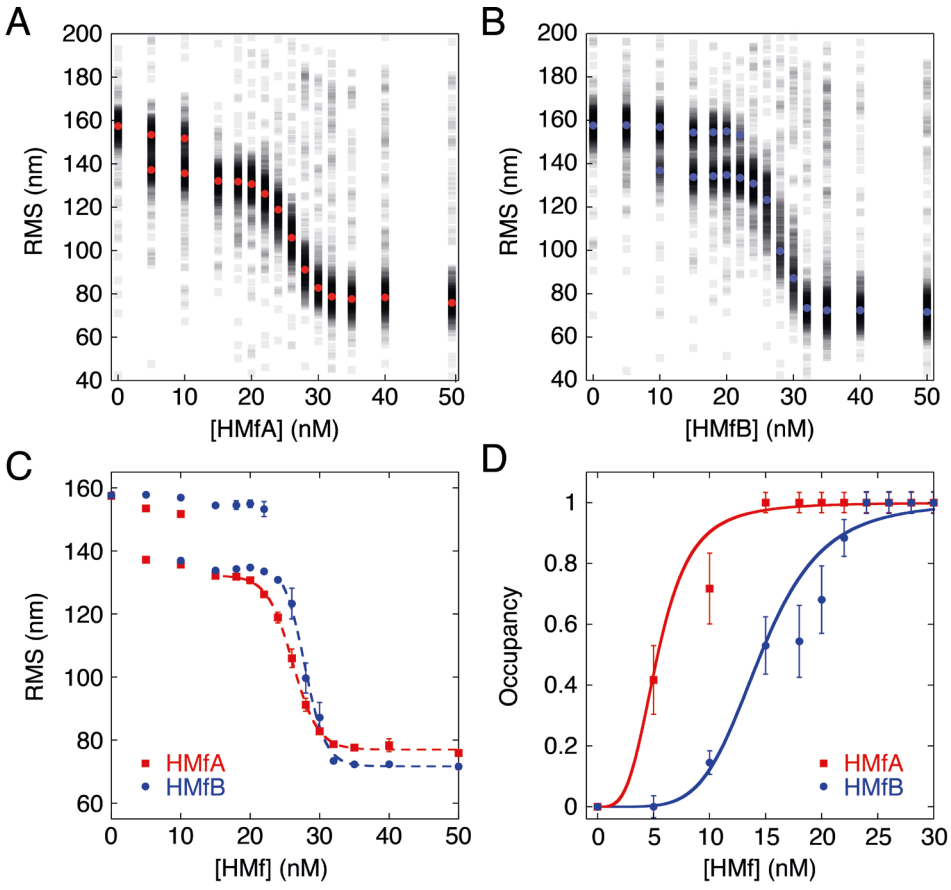

Figure 6.1: Compaction of a DNA substrate containing the Clone20 sequence by HMfA and HMfB using tethered particle motion (TPM). A) Root mean square displacement (RMS) values of all DNA tethers incubated with HMfA and B) with HMfB. Fitted mean RMS values are represented by red and blue dots, respectively. Grey squares represent the individual RMS values. C) Mean RMS values of DNA tethers with HMfA and HMfB. Error bars represent standard error of the mean of the replicates. Dashed lines are to guide the eye. D) Binding curve for specific binding of HMfA and HMfB to the Clone20 substrate. The data points were fit using the Hill binding model. The binding affinity (K_D) resulting from the fit is 5.7 ± 1.6 nM and 14.6 ± 1.6 nM and the Hill binding coefficient (n) is 3.6 ± 0.8 and 5.2 ± 2.0 for HMfA and HMfB, respectively. Error bars represent the standard error of the mean of the replicates, and propagated error for data points at saturation.

SELEX optimized for HMfB and is in disagreement with results from EMSA experiments (160). The difference may be caused by a different pH (7.0 in our experiments, 8.0 in the experiments of Bailey *et al.*). The difference in measured affinity could also be attributed to differences in methodology, with EMSA measuring mobility of protein-DNA complexes in a gel matrix and TPM measuring the conformation of protein-DNA complexes in solution attached to a glass surface. The observed reduction in RMS at concentrations <22 nM is attributed to for-

mation of a tetrameric histone-DNA complex (160, 171). This complex does not function as a nucleation site, as the concentration at which the hypernucleosome is formed, was not affected by this complex.

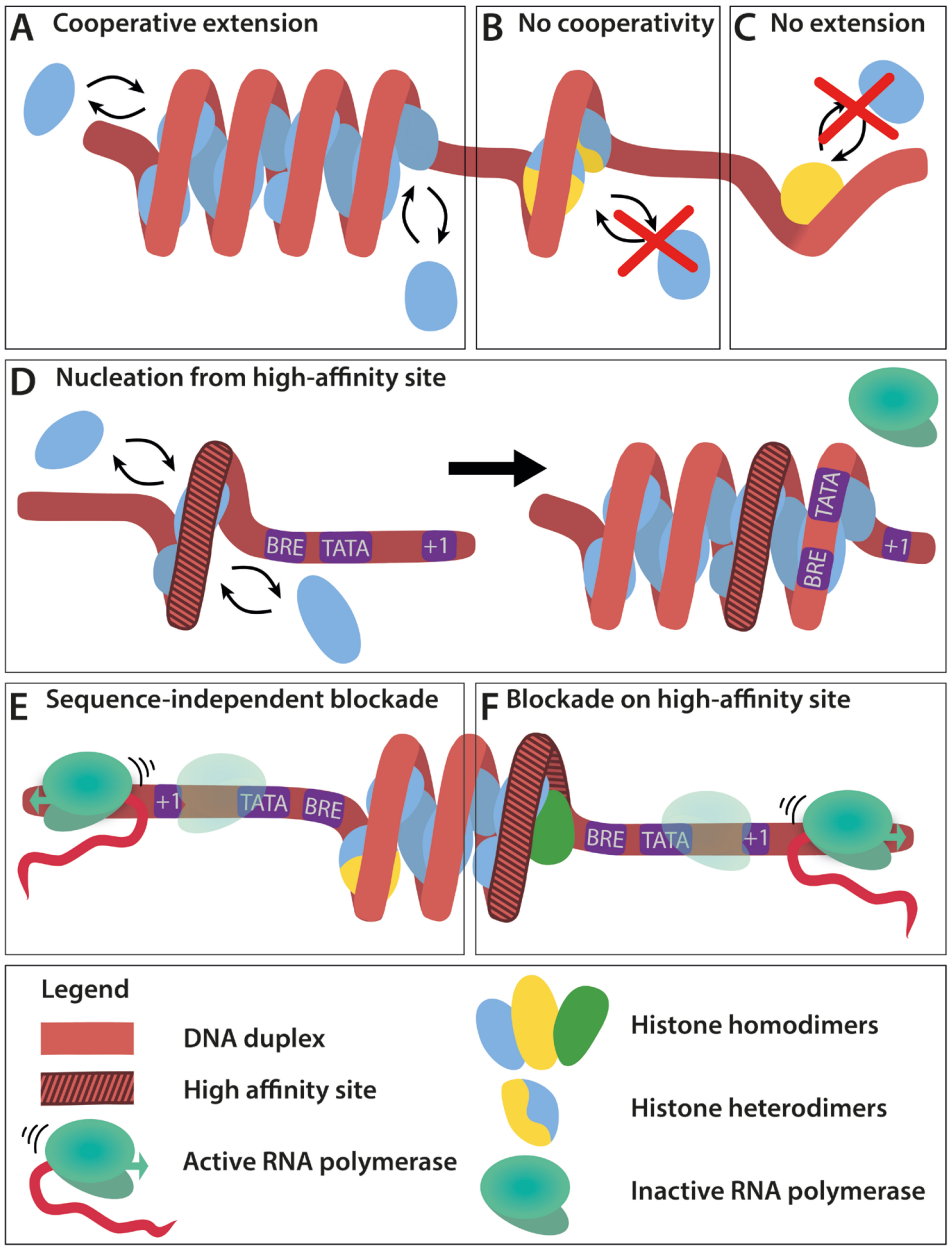

Figure 6.2: Possible mechanisms for involvement of archaeal histones in transcription regulation. Histones cooperatively bind and compact DNA in hypernucleosomes, which can be modulated via extension or dissociation from both ends (A). As tetramers (B) or dimers (C),

Conclusion

A DNA substrate containing the artificial high affinity sequence Clone20 is compacted by *M. fervidus* histones in two distinct steps, indicating two distinct complexes. At low concentrations, a histone assembly that is likely a tetramer, modestly compacts DNA. Although no sequences with a binding affinity similar to that of the Clone20 sequence have been identified on the *M. fervidus* genome thus far, it suggests a functional role for high affinity sequences. HMfA has a lower K_D than HMfB on Clone20. However, HMfB multimerizes more cooperatively into a hypernucleosome, and also nonspecifically compacts DNA more than HMfA. This may reflect a distinct cellular function, with HMfA being able to better position tetramers and/or hypernucleosomes at specific locations on the genome, and HMfB forming hypernucleosomes that are more stable than those formed by HMfA. Differential expression of HMfA and HMfB is in accordance with the hypothesis that histone variants play a role in transcription regulation, and these variants may concur with high affinity sequences and different forms of histone multimers in archaeal transcription regulation.

(Figure 6.2, continued) histones are also able to non-cooperatively bind DNA, thereby inducing less compaction than a hypernucleosome. Histones can function as transcription factors, possibly binding to specific sequences on the genome (D). These specific sequences may function as a positioning sequence for the hypernucleosome. Histone variant assemblies that are not able to multimerize may be able to nonspecifically (E) or specifically (F) prevent hypernucleosome formation in the promoter region, thereby leaving the promoter accessible to RNA polymerase.

Materials and Methods

DNA substrate preparation

For the TPM DNA subtrate, the Clone20 sequence (GCACAGTTGAGCGATCAAAAACGCCGTAGAACGCTTTAATTGATAATCAAAGGCCGCAGA, (164)) was cloned into pBR322, as described previously (127), resulting in plasmid pRD120. We used PCR to generate and amplify a 685 bp linear substrate containing the cloned sequence, using digoxygenin- and biotin-labeled oligonucleotides (195).

Tethered particle motion

The tethered particle motion experiments, data analysis and representation of results were performed as previously described (322). To select single-tethered beads, we used a standard deviation cut-off of 8% and an anisotropic ratio cut-off of 1.3. As measurement buffer we used 50 mM Tris-HCl pH 7 and 75 mM KCl in water. HMfA and HMfB were kindly provided by John Reeve and Kathleen Sandman (20). We verified the identity of both proteins using mass spectrometry.

Supplemental figure

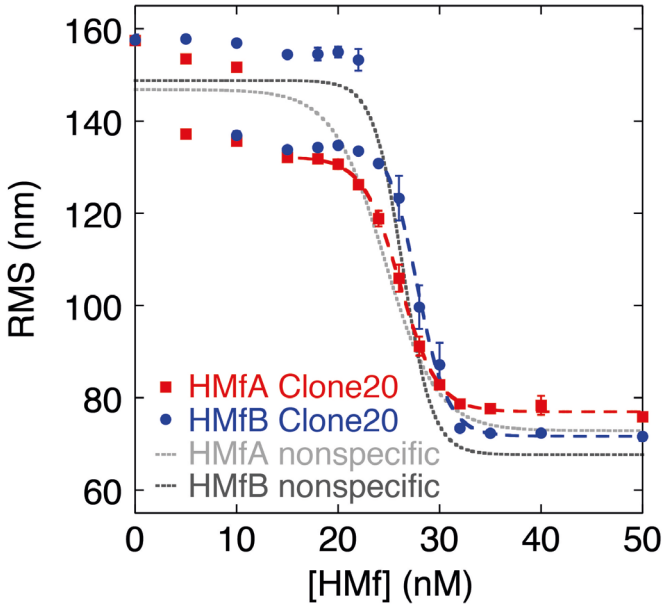

Figure S6.1: Compaction of HMfA and HMfB on a specific (Clone20) and nonspecific DNA substrate. Data of HMfA and HMfB on the Clone20 substrate is from this chapter (dashed lines are to guide the eye, error bars represent standard error of the mean), data of HMfA and HMfB on the nonspecific substrate is reproduced from CHAPTER 4 and depicted as a line to guide the eye.