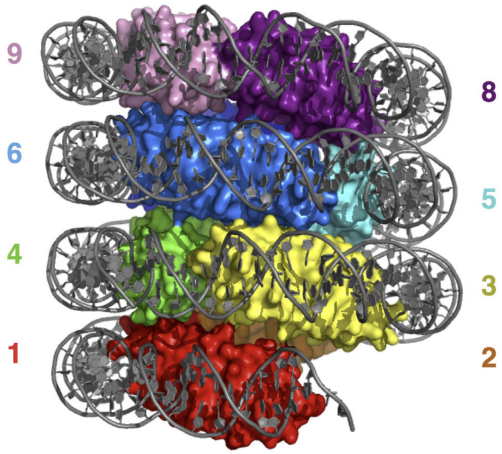
Hypernucleosome formation of the HTkB-DNA complex



THESIS

submitted in partial fulfillment of the
requirements for the degree of

BACHELOR OF SCIENCE
in
PHYSICS

Author :	Benjamin Sluijter
Student ID :	s2354276
Supervisor :	John van Noort
2^{nd} corrector :	Thomas Schmidt

Leiden, The Netherlands, June 16, 2022

Hypernucleosome formation of the HTkB-DNA complex

Benjamin Sluijter

Huygens-Kamerlingh Onnes Laboratory, Leiden University
P.O. Box 9500, 2300 RA Leiden, The Netherlands

June 16, 2022

Abstract

In the eukaryotic cell nucleus, DNA is compacted by histone proteins which form nucleosomes. The amount of compaction influences the cell function in the form of for example gene expression. Homologues of eukaryotic histones have been found in Archaea. Interestingly, in some Archaea, these histones form multimeric complexes, called hypernucleosomes, rather than the octameric nucleosomes in eukaryotes. This has until now only been shown for the HMfA and HMfB histones of *Methanothermes fervidus*. Here we give evidence for hypernucleosome formation of the HTkB histone of *Thermococcus kodakarensis*. This result shows that hypernucleosome formation of HMf is not unique and it is thus likely that we will encounter hypernucleosomes in even more Archaea. More of these assays on different Archaeic histones may shine a light on how the eukaryotic nucleosome descended from Archaeic (hyper)nucleosomes.

Chapter 1

Introduction

All of life on earth can be divided into three groups: the Bacteria, the Archaea and the eukaryotes, to which we humans belong. In the past 40 years, thanks to the progress in genetics research, more and more similarities between eukaryotes and Archaea have been found. This has led scientists to believe that we eukaryotes have evolved in some way from Archaea [1]. Especially the genetic machinery of Archaea has been found to ressemble that of eukaryotes. For example, RNA polymerases in Archaea are much more similar to eukaryotic RNA polymerases than to those in Bacteria [2]. Also, many Archaea have histones that are homologous to eukaryotic histones [3]. Histones are the proteins that bind to DNA in the form of nucleosomes, organizing it into a compacted chromatin structure. They play a key role in the genome function, as they help to regulate the accessibility of genes. In Archaea, the histones also bind to DNA. However, in the Archaea *Thermococcus kodakarensis*, MNase digestion was reported to yield chromatin particles of varying size [4]. This is different from eukaryotes, where MNase digestion produces chromatin particles which are all the same size and correspond to nucleosomes. Mattiroli et al. showed the crystal structure of the complex of HMfB histones of the Archaea *Methanothermes fervidus* with 80 bp DNA fragments, revealing that HMfB dimers form multimers which wrap DNA in a quasi-continuous superhelix [5] (figure 1.1). This structure is also called the hypernucleosome [6].

The formation of the hypernucleosome and its mechanical properties were further examined by Henneman et al. by means of force spectroscopy with magnetic tweezers [7]. In magnetic tweezers, DNA molecules are tethered between microscopic paramagnetic beads and a glass surface. With a magnet, the force on the beads is controlled and with a camera

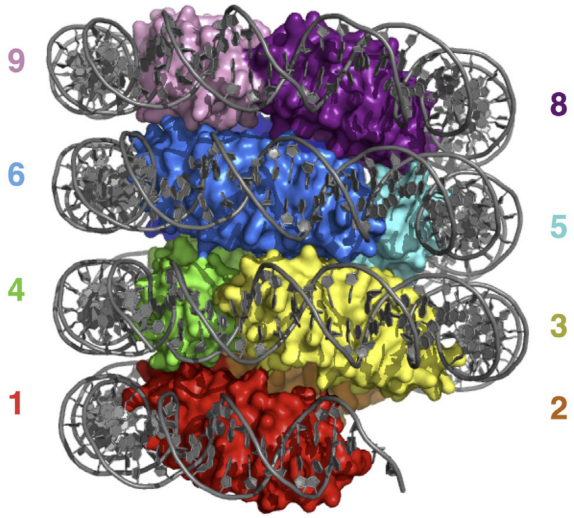


Figure 1.1: *Structure of the hypernucleosome. The histone dimers numbered 1-9 form a multimer which wraps the DNA in a quasi-continuous superhelical shape.*

the height of the beads is measured, giving force-extension curves of the studied DNA. Henneman et al. discovered that the HMfA/HMfB hypernucleosome unstacks as a function of force through two transitions. In this research they developed a statistical mechanics model, accurately describing the force-extension curves and returning different 'stacking energies' for HMfA and HMfB. This gives further insight in the genomic compaction in Methanothermes fervidus, which may use different ratios of its two histones to regulate hypernucleosome size.
In other work, Henneman et al. use the crystal structure from Mattiroli et al. and amino acid sequences to predict for a diversity of Archaea from different phylogenic branches whether they form hypernucleosomes [6]. If the used computational predictions were correct, that would suggest that hypernucleosomes are formed in a variety of different Archaea. However, at the moment, only for *Methanothermes fervidus* the formation of the hypernucleosome has been validated experimentally.
In this thesis we will look at the force-extension behaviour of another earlier mentioned Archaea, *Thermococcus kodakarensis*, which is predicted by Henneman to also form hypernucleosomes. We will do this in a very similar way to Henneman et al., with magnetic tweezers and fitting the same statistical physics model. For the magnetic tweezers, we will also compare two techniques of attaching the DNA to the glass surface (tethering mechanisms).
We reveal that the investigated *Thermococcus kodakarensis* histone, HTkB,

indeed forms hypernucleosomes, as predicted by Henneman et al. We thereby contribute experimental evidence on archaeal genome compaction, hopefully taking us one step closer to finding out where genome compaction of eukaryotes originated from.

Chapter 2

Theory

2.1 The worm like chain model

In 1994, Bustamante et al. showed that the force-extension curve of bare DNA is well described by the worm like chain model, which is used to describe semi-flexible polymers. Unfortunately, this model does not have an analytical solution. In literature, the following approximation is often used [8]:

$$\frac{fA}{k_BT} = \frac{1}{4}((1-\frac{z}{L})^{-2}-1)+\frac{z}{L} \tag{2.1}$$

Where f is the force, A the persistence length of the DNA, k_B the Boltzmann constant, T the temperature, L the contour length of the DNA and z the extension.
However, in this thesis we will use a different approximation, which is an extensible worm like chain formula, as opposed to equation 2.1, which is an inextensible one [9, 10]:

$$z_{WLC}(f) = L(1-\frac{1}{2}\sqrt{\frac{k_BT}{fA}}+\frac{f}{S}) \tag{2.2}$$

Where S is the stretch modulus. The extensible worm like chain takes enthalpic stretching, which happens for higher forces, into account, while the inextensible WLC only accounts for entropic stretching.
There are two more reasons for using this approximation. The first reason is that in magnetic tweezers, the extension is measured as a function of force, as opposed to the more common optical tweezers, where force is measured as function of extension. Therefore it is more useful to have a

formula in which extension is a function of force instead of the other way around.
The second reason is that we can add extensions linearly. For example, imagine we have a DNA string which consists of one part with a persistence length a and a second part with persistence length b. For the extensions of the two parts, we then have $z_{part1} = z_{WLC}(A = a)$ and $z_{part2} = z_{WLC}(A = b)$. In such cases, we can simply add to get the extension of the total DNA string: $z_{total} = z_{WLC}(A = a) + z_{WLC}(A = b)$. We will see in the next section that this property is of great convenience for our model.

2.2 Statistical physics model for the hypernucleosome

The DNA in some Archaea is found to be wrapped around so-called hypernucleosomes. In earlier mentioned research by Henneman et al. on HMfA and HMfB hypernucleosomes, a model was proposed to explain the force-extension curves of DNA folded into hypernucleosomes [7]. Here we will explain this model, as we will be testing this model for HTkB.
Histon proteins, HmfA, HmfB or HtkB, form dimers when bound to DNA. These dimer-DNA complexes fold in three different conformations. In the first one, they are stacked on top of each other, to form a hypernucleosome. In the second conformation, the dimers no longer interact, but they do still interact with the DNA, bending it into a kinked 'beads-on-a-string' structure. In the last conformation, the dimers are still attached to the DNA, but no longer bend it, hence the force-extension relation follows that of bare DNA.
Now, for a complete statistical mechanics model, we consider all possible states the DNA-histone complex can find itself in. With a 'state', we mean a specific distribution of all dimers over the three conformations. We then calculate for all possible states j the extension $z_j(f)$ and the free energy $G_j(f)$. The total extension is the Boltzmann-weighted mean of all states:

$$\langle z_{total}(f) \rangle = \frac{\sum_j z_j(f) exp(-G_j(f)/k_B T)}{\sum_j exp(-G_j(f)/k_B T)} - z_0 \tag{2.3}$$

For $z_j(f)$, we can add the extensions of each of the different conformations. We split the extension up into two parts: the part of the DNA-histone complex that is in hypernucleosome conformation and the rest of the DNA-histone complex.

Version of June 16, 2022– Created June 16, 2022 - 14:57

The hypernucleosome behaves like a Hookean spring and its extension goes linearly with force. However, we use the Freely jointed chain model (FJC) for this extension, instead of the formula for a Hookean spring. At low forces, this model describes the desired Hookean spring and at higher forces, at which hypernucleosome conformation is unstable and all dimers have transitioned into less compact conformations, the model converges to an asymptote. This finite length makes it more realistic than the simple Hookean spring formula which would allow for unphysically large extensions in the force regime where no hypernucleosome is present.
The FJC formula is the following [11]:

$$z_{FJC}(f) = N_{HN} \cdot z_{dimer} \cdot (coth(\frac{3f}{z_{dimer}k}) - \frac{z_{dimer}k}{3f}) \tag{2.4}$$

where N_{HN} is the number of dimers in hypernucleosome conformation, z_{dimer} the height of a single dimer, k the stiffness of the hypernucleosome and f the force.
The remainder of the histone-DNA complex, that is either in the beads-on-a-string conformation or in the unbend conformation, follows the WLC model. The persistence length depends on the number of dimers in the beads-on-a-string conformation and is given by the following equation [12]:

$$A = \frac{A_0}{1 + A_0 \cdot N_{BoS}^2 \cdot 8[1 - cos(\frac{\alpha}{4})]/L)^2} \tag{2.5}$$

where A_0 is the persistence length of bare DNA, N_{BoS} is the number of dimers in the beads-on-a-string conformation, α is the DNA deflection angle induced by the dimer and L is the contour length of the DNA that is either in the beads-on-a-string conformation or in the bare conformation. This contour length is calculated by taking the contour length of the entire DNA string and subtracting L_{dimer} for each dimer that is in hypernucleosome conformation. Here, L_{dimer} is the length of DNA that is wrapped by a single dimer in hypernucleosome conformation.
The extension of the part of the DNA-histone complex that is not in hypernucleosome conformation is then given by equation 2.2, using the just described contour length and using equation 2.5 for the persistence length.

The free energy is calculated similarly; we calculate the hypernucleosome free energy and the free energy of the rest of the DNA separately. We then add these and subtract the work W done by the bead to get the total free energy $G_j(f)$.

Version of June 16, 2022– Created June 16, 2022 - 14:57

The free energy in the hypernucleosome part of the DNA is calculated as follows:

$$G_{HN}(f) = \int z_{FJC}(f)\, df - N_{HN}(g_{wrap} + g_{stack}) \tag{2.6}$$

Where $z_{FJC}(f)$ is the hypernucleosome extension described in eq. 2.4, N_{HN} is the number of dimers in hypernucleosome conformation, g_{wrap} is the interaction energy between DNA and a wrapped dimer and g_{stack} the interaction energy between two stacked dimers.

The free energy in the rest of the DNA is calculated as follows:

$$G_{DNA}(f) = \int z_{WLC}(f)\, df - N_{BoS} \cdot g_{wrap} \tag{2.7}$$

Where $z_{WLC}(f)$ is the worm like chain formula described in equation 2.2 (with persistence length calculated using equation 2.5), N_{BoS} the number of dimers in beads-on-a-string conformation and g_{wrap} the same interaction energy as in equation 2.6.

Finally, the work W done by the bead is simply $f \cdot z$, with z the height of the bead.

We then have

$$G_j(f) = G_{HN}(f) + G_{DNA}(f) - W(f) \tag{2.8}$$

2.2.1 Hypernucleosome model for HMfB

As mentioned earlier, the hypernucleosome model was developed by Henneman et al. to describe the force-extension curves of HMfA and HMfB complexes. These are at the moment still the only archaeal histones for which this model has proved to be working.
The curves found by Henneman et al. are shown in figure 2.1. They fitted equation 2.3 with contour length = 3646 bp, persistence length = 50 nm, dimer footprint = 30 bp and dimer height = 4 nm as fixed parameters. The results of the fitted parameters are displayed in table 2.1. The fit results are shown in table.

2.3 Bead offset

In the magnetic tweezer setup, we can not measure DNA extension directly. What we measure is the bead height Z. The bead height and the

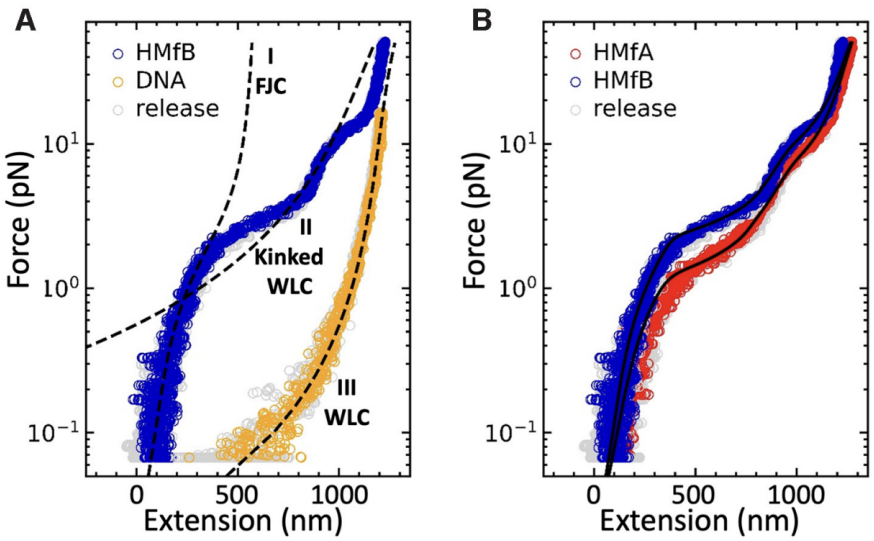


Figure 2.1: *HMfA, HMfB and bare DNA curves found by Henneman et al. Figure taken from [7].*

	HMfA	HMfB
G1 (k_BT)	1.4 ± 0.4	3.5 ± 0.6
G2 (k_BT)	5.7 ± 1.1	5.0 ± 1.1
k (pN/nm)	0.45 ± 0.11	0.83 ± 0.24
angle (deg)	14.9 ± 1.7	11.1 ± 1.0
Ndimer (-)	105 ± 11	109 ± 8

Table 2.1: *Mean and standard deviations of fitted parameters with equation 2.3 to HMfA-DNA complex (N=82) and HMfB-DNA complex (N=56), by Henneman et al. [7]*

DNA extension differ by an offset, which is due to the position of the tether on the bead. This is not always on the lowest point of the bead, as the bead has a magnetic moment that aligns with the horizontal magnetic field of the magnet. The offset is therefore a value between zero and the radius of the bead, which is 1.4 μm in our case. To get the extension of the DNA tether, we add this offset to the measured bead height.

Furthermore, it is important to note that this offset causes the observed extension to deviate from the theoretical extension, as the observed extension can not be smaller than the bead offset. When the observed extension equals the bead offset, the bead height is zero. The bead then touches the flowcell surface and can not go any lower. We will encounter this further on in this thesis in the results section.

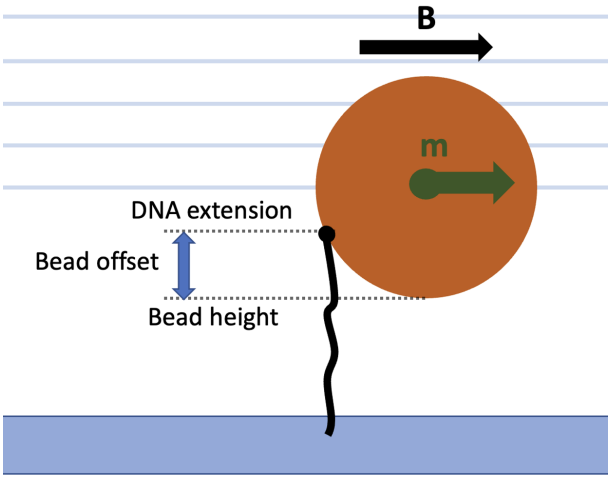


Figure 2.2: *The magnetic bead is shown in brown, the DNA tether in black and the flowcell surface in blue. The direction of the magnetic field is depicted with the blue lines and the black arrow. The magnetic moment of the bead, which aligns with the magnetic field, is shown with the green arrow. As is shown, the difference between the DNA extension and the bead height is the bead offset.*

2.4 Tethering mechanisms

In magnetic tweezers the investigated DNA molecule is attached between a paramagnetic bead on one side and a glass flowcell surface on the other side, tethering the bead to the surface. The most common way to do this is to label one side of the DNA molecule with Biotin, which binds to Streptavidin on the bead, and the other side with Digoxigenin, a flower steroid. You then coat the flowcell surface with an antibody to this steroid, called anti-Digoxigenin.
A more recently developed technique is using so-called click-chemistry [13, 14]. With click-chemistry, a covalent bond is formed between an azide molecule on the flowcell surface and a DBCO molecule, labeled on the DNA. With click-chemistry, one can apply much higher forces without breaking the bond, which is its primary advantage compared to Digoxigenin. However, click-chemistry is slower and has not been used as much as Digoxigenin. In this thesis we will try both methods.

Chapter 3

Materials and methods

3.1 Microscope

The materials used in the microscope setup are described in [15]. The only difference is the objective. We used a Nikon Plan Apo Lambda 20x Microscope Objective.

3.2 Sample preparation

3.2.1 DNA preparation

In this thesis we used three DNA samples. All three were prepared by doing two restrictions and two Klenow reactions. In one Klenow reaction one end of the DNA strand was labeled with Biotin, allowing the DNA to bind to the streptavidin on the beads. The other Klenow reaction labeled the other end of the DNA with either DBCO or Digoxigenin to bind to the respectively Azide or Anti-Digoxigenin functionalized flowcell surface.
From here on, we will refer to the two samples that were labeled with Digoxigenin as 'Dig DNA 1/2' and to the DBCO labeled DNA as 'DBCO DNA'. We will describe the Dig DNA 1 protocol in detail to give an idea of the procedure. For the other two DNA samples we will describe the most important differences with the first protocol.

Dig DNA 1

For the Dig DNA 1 sample we started out with a 5867 bp plasmid, called pCP130, containing 16 times the Widom 601 sequence [16]. A 5217 bp

part of the plasmid containing the Widom 601 sequences was cut out and labeled with the following steps:

1. For the first digestion, we combined the reagents in table 3.1 in an Eppendorf tube. We incubated this overnight at 37°C and 350 rpm in an Eppendorf ThermoMixer.

pCP130 (1.6 $\mu g/\mu L$)	20 μL	32 μg
CutSmart buffer 10 x	5 μL	-
BsaI (10 $units/\mu L$)	3 μL	30 $units$
RNAse free water	22 μL	-
Total volume	50	

Table 3.1: *Digestion with restriction enzyme BsaI*

2. To inactivate the BsaI, we changed the temperature to 65 °C and incubated for another 20 minutes.
3. Then we purified the sample with a Promega Wizard SV & PCR Clean-up system kit and checked the yield with a Biodrop spectrophotometer, which gave a concentration of 380 $ng/\mu L$
4. To label the DNA on one end with biotin by a Klenow reaction, we combined the reagents in table 3.2 in an Eppendorf tube. We incubated this for one hour at 37°C and 0 rpm.

BsaI digested pCP130 (380 $ng/\mu L$)	40 μL	15.2 μg
Klenow buffer 10 x	6.5 μL	-
Bio-11-ddUTP (1 mM)	1.3 μL	20 μM
dCTP (100 mM)	1.3 μL	2 mM
dGTP (100 mM)	1.3 μL	2 mM
Klenow (10 $units/\mu L$)	3 μL	30 $units$
RNAse free water	11.6 μL	-
Total volume	65 μL	

Table 3.2: *Labeling the DNA with Biotin*

5. To inactivate the Klenow fragments, we changed the temperature to 75°C and incubated for another 10 minutes.
6. Then we did another purification with the Promega Wizard kit, this time yielding a concentration of 322 $ng/\mu L$.

7. For the second digestion, we combined the reagents in table 3.3 in an Eppendorf tube. We incubated this overnight at 37°C and 350 rpm in ThermoMixer.

Biotin-labeled pCP130 (322 $ng/\mu L$)	50 μL	16.1 μg
10 x NEBuffer 3.1	6 μL	-
BseYI (5 $units/\mu L$)	4 μL	20 $units$
Total volume	60 μL	

Table 3.3: *Digestion with restriction enzyme BseYI*

8. To functionalize the second end of the DNA string with Digoxigenin with a Klenow reaction, we combined the reagents in table 3.4 in an Eppendorf tube. We incubated this for one hour at 37°C and 0 rpm.

BseYI digested Biotin-labeled pCP130 (322 $ng/\mu L$)	50 μL	16.1 μg
Dig-11-dUTP (1 mM)	1.2 μL	22 μM
dCTP (1 mM)	1.2 μL	22 μM
Klenow (10 $units/\mu L$)	2 μL	20 $units$
Total volume	54.4 μL	

Table 3.4: *Digoxigenin labeling*

9. To inactivate the BseYI, we changed the temperature to 75°C and incubated for another 10 minutes.

10. Then we did again a purification step with the Promega Wizard kit, with a yield of 117 $ng/\mu L$.

11. Finally, to get rid of the 650 bp plasmid leftover, we put 35 μL of our sample on a gel and ran it for 30 minutes at 120 Volt, after which we cut out the 5217 bp part. This was then purified with a QIAquick gel extraction kit, which had us ending up with 70 μL of 16 ng/μL DNA sample.

DBCO DNA

The DBCO DNA was made with almost the same protocol as Dig DNA 1, from the same pCP130 plasmid. For DBCO DNA, the purification in step 3 was skipped. Furthermore, the dCTP and dGTP concentrations in step 4 were $\sim$100 times lower for DBCO DNA. Finally, for the elution in

the purification in step 10 we used our own RNAse free water instead of the nuclease free water from the Promega Wizard kit that was used for Dig DNA 1. The nuclease free water from the kit caused the DNA sample to not sink well to the bottom of the well in the gel electrophoresis. This might suggest that this particular bottle of water was contaminated in some way.

Dig DNA 2

For Dig DNA 2 we started with a 3269 bp linear DNA substrate, also containing 16 times the Widom 601 sequence. This DNA was first digested with NdeI instead of BsaI. It was labeled with Digoxigenin after the first digestion and with Biotin after the BseYI digestion, instead of the other way around.
The purification in step 3 was skipped, just like for DBCO DNA. For this sample, also step 11 was not necessary, as we did not have a large plasmid leftover like for the other two DNA's. The remaining purifications (steps 6 and 10) were done with a QIAquick Gel Extraction kit instead of the Promega Wizard kit.
The end product was 3246 bp long.

3.2.2 Flowcell preparation

Flowcell assembly

1. We cut a piece of Parafilm M with a die cutting machine, using the cutting shape in figure 3.1. As seen in the figure, the shape includes an up- and downside. After cutting, the two sides of parafilm were folded on top of each other.

2. We cleaned a 60 x 25 mm glass microscope slide and a 40 x 25 mm glass microscope slide by picking them up with tweezers and dipping them a few times in a glass beaker filled with Milli-Q water, and drying them with a N_2 spray gun.

3. We placed all the constituents of the flowcell on an analog heatblock, that was heated up to about 80°C, in the order (from bottom to top): glass microscope slide 60 x 25 mm - Parafilm - laser-cut frame - glass microscope slide 40 x 25 mm. By applying pressure to the frame and the top microscope slide, all constituents stuck to the heated parafilm. After letting the construct cool down, the flowcell was finished.

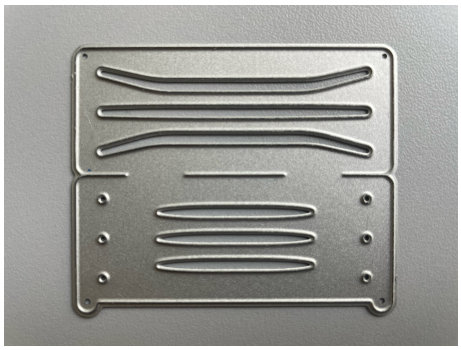


Figure 3.1: *Parafilm cutting shape for the flowcell assembly*

Protocol A. Azide-DBCO click chemistry

1. We flushed each channel with 200 μL HPLC grade water to clean the channels.

2. Then we flushed each channel with 60 μL Azide functionalized BSA (150 μM BSA and 1.5 mM Azido-PEG4-NHS-ester (Jena Bioscience) in PBS, incubated overnight at 37°C), both to passivate the flowcell surface and to functionalize it with Azide for DNA tethering. We let this incubate for an hour at room temperature. .

3. We flushed with 60 μL PBS to remove remaining Azide-BSA.

4. Then we flushed with 60 μL of DBCO DNA, diluted in ESB buffer (10 mM HEPES, 0.1 M KCl, 0.1% Tween 20, 0.2% BSA) to 9.4 $ng/\mu L$. We let this incubate for three days.

5. We flushed with 100 μL ESB buffer to get rid of remaining DNA.

6. We then flushed 60 μL of beads, diluted in ESB to a concentration of 33 $\mu g/mL$, which attached to the Biotin-labeled end of the DNA molecules.

7. Finally, we flushed with 100 μL ESB buffer to remove remaining beads

Protocol B. Anti-Digoxigenin - Digoxigenin

1. We flushed each channel with 200 μL HPLC water to clean the channels.

2. Then we functionalized the flowcell surface with 60 μL Anti-Digoxigenin (10 ng/μL in PBS) for DNA tethering. We let this incubate for an hour at room temperature.

3. We flushed with 100 μL PBS to remove remaining Anti-Digoxigenin

4. Then we passivated the flowcell surface with 60 μL BSA 4% in PBS. We let this incubate for an hour at room temperature.

5. We flushed with 100 μL ESB to remove remaining BSA.

6. Then we injected 60 μL of DNA-beads mix (For Dig DNA 1: 20 pg/μL DNA and 1:300 Dynabeads M-270 Streptavidin (Thermo Fisher). For Dig DNA 2: 23 pg/μL DNA and 1:200 Dynabeads (same). Incubated for 30 minutes at room temperature in rotator.) We let this incubate for 15 minutes at room temperature.

7. We flushed with 100 μL ESB buffer to remove remaining DNA-beads

For HtkB measurement the next steps were taken

8. We incubated the flowcell for ten minutes with 60 μL HTkB, diluted in ESB buffer. HTkB was a kind gift of Ilias Zarguit and Nico van der Vis.

9. We flushed again with 60 μL of the same HTkB solution, as a bit of concentration is lost due to the binding to DNA.

3.3 Measurement methods

The measurement protocol we used is largely described in section 2.3.6 of [15]: Initial bead selection and height calibration. After the protocol described there, we did a quick magnet trajectory, in order to remove non-tethered beads from the flowcell surface. After that we did the measurement.
For each measurement, we first let the magnet approach the sample with constant velocity and then moved it up again with the same constant velocity. This means the force is increased and decreased in an exponential manner, as the force depends exponentially on the magnet height.

3.4 Analysis methods

We corrected all the results for drift. For each measurement, we fitted a linear drift to each trace for low force (the highest ten percentiles of the magnet position). We then took the median (instead of mean because of outliers) of the drifts of all traces and then subtracted that amount of drift from all traces.

Checking for sticking

The first thing we wanted to see for the different sample preparation methods is how much sticking there was. With sticking we mean that the bead or the DNA tether is stuck to the flowcell surface. We quantified this with parameter dZ, which is the difference between the 10th and the 90th percentile of the bead height. Sticking beads are then observed as a peak around zero, while the non sticking beads form a distribution between zero and the contour length of the DNA.

Fitting z_{WLC} curves

For the non-sticking traces we wanted to know whether they follow the worm like chain model. We know that bare DNA should follow this curve, with a persistence length of approximately 50 nm.
So, for each measurement, we fitted the force-extension curves with equation 2.2 with the right contour length (3248 bp for Dig DNA 2 and 5217 bp for DBCO DNA), with the persistence length and bead offset as free parameters.

Fitting the hypernucleosome model

Apart from testing the different tethering mechanisms, we wanted to know if the hypernucleosome model from Henneman et al. [7] also applies to HTkB and if so, what the parameter values are for HTkB.
In order to do this, we fitted the hypernucleosome model to the force-extension traces of the HTkB samples. We only took the traces with a dZ between 0.8 μm and 1.4 μm, to exclude loose beads, stuck beads and beads with a large bead offset.
Because of large non-linear drift that came with long measurement times, a measurement time of ten seconds was used. This short measurement time did lead to hysteresis between the stretch- and release curve (Suppl.

figure 6.1), meaning that at least one of the curves did not happen in thermodynamic equilibrium. We used the stretch curves for the fitting, as the stretching dynamics are expected to be faster than the releasing dynamics, and should thus happen closer to equilibrium.
Furthermore, the fit was done only on the points with F$>$0.4 pN to reduce the influence of the bead offset on the fit.
We did the HTkB-DNA measurements in two different HTkB concentrations ($\sim$100 nM and $\sim$400 nM). However, according to previous TPM experiments executed by Ilias Zarguit [personal communication], the smaller concentration was already enough to saturate the DNA. Also in our experiments the curves looked similar and the fit results agreed well within standard deviations. Therefore we took the results for the two concentrations together in the results section.

Version of June 16, 2022– Created June 16, 2022 - 14:57

Chapter 4

Results

4.1 Bare DNA

4.1.1 Digoxigenin tethering

For the Digoxigenin-anti-Digoxigenin method, two different DNA's were investigated with magnetic tweezers force spectroscopy.
To check for sticking, the dZ values (see section 3.4: sticking) were calculated for all beads. For the Dig DNA 2 (figure 4.1a.), a gaussian-like distribution around a dZ of about 0.8 μm is observed, with no apparent peak around zero. From this we can conclude that there is no sticking occurring. For the Dig DNA 1, almost all dZs were lower than 0.2 μm (Suppl. figure 6.2), which tells us that practically all beads were stuck.
For the non-sticking DNA, Dig DNA 2, the beads with 0.5 $\mu m <$ dZ $<$ 1.4 μm, were fitted with equation 2.2, to see whether the DNA behaves as we expect from bare DNA. The fitting was done on the release curves on the force interval from 0.4 pN to 16 pN. A typical example is shown in figure 4.1b. with a persistence length of 12 nm and an R-squared of 0.97. The trace is adequately described by equation 2.2, down until an extension of around 0.2-0.3 um, which can be explained by the bead offset (section 2.3). As seen from the R-squared values in figure 4.1c., most traces were well described by the WLC.
For the beads with an R-squared greater than 0.93, the persistence length is shown in figure 4.1d. Remarkably, the peak is clearly observed around 10 nm, instead of the expected value around 50 nm. Between 40 nm and 60 nm only $\sim$3% of the beads is found.

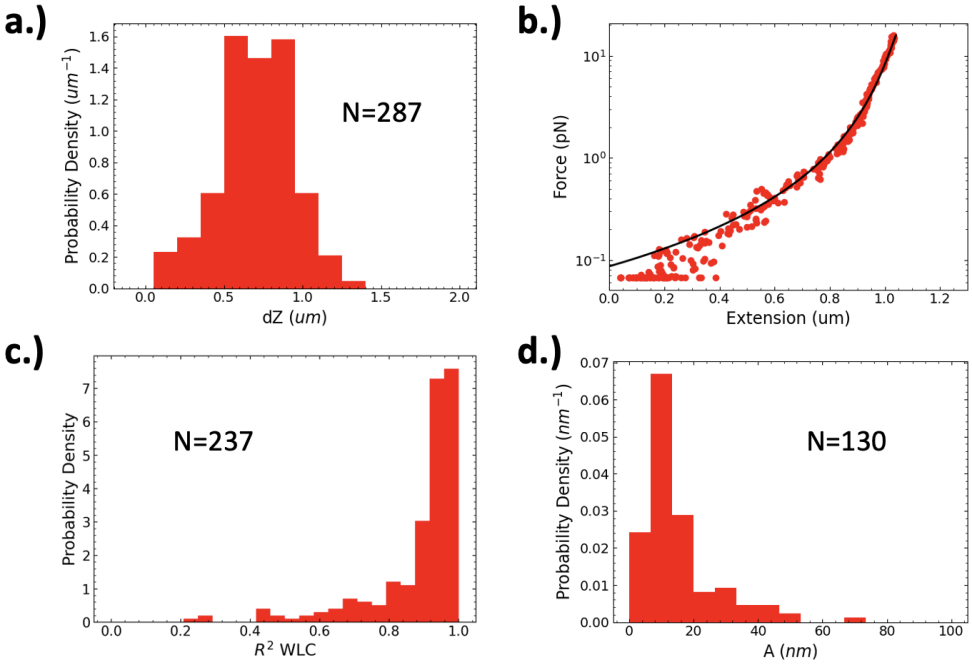


Figure 4.1: *Force-extension curves of Dig DNA 2 are well described by the Worm Like Chain model with very low persistence lengths.* ***a.)*** *Histogram of all dZ values of the measured beads in the Dig DNA 2 experiment (see section 3.4: sticking).* ***b.)*** *Typical Dig DNA 2 curve. We fitted equation 2.2 with an R-squared of 0.97, which is shown by the black line. The fixed parameter is the contour length = 3246 bp. Fitted parameters are the persistence length = (11.9 $\pm$ 0.4) nm and the bead offset = (0.217 $\pm$ 0.004) μm.* ***c.)*** *Histogram of the R-squared values for all Dig DNA 2 curves with a dZ between 0.5 μm and 1.4 μm.* ***d.)*** *Histogram of fitted persistence lengths of all Dig DNA 2 curves with a dZ between 0.5 μm and 1.4 μm and an R-squared of 0.93 or higher.*

4.1.2 DBCO-azide click chemistry

In addition to the more standard Digoxigenin tethering mechanism, we also investigated DBCO-azide click chemistry as a method for magnetic tweezers.
In the dZ histogram in figure 4.2a. it is seen from the first bin that for this method, we do have some sticking. However, it is only about 10%, meaning we do not loose that many beads. For the beads with 0.8 $um < dZ <$ 2.0 um, we fitted a WLC on the release curve for 0.4 $pN < F <$ 16 pN. In figure 4.2b. a typical curve with WLC fit is shown, which again described the trace adequately down until the offset value, which is 0.5 um in this

case. In figure 4.2c we see that $>$ 80% of the traces are fitted with an R-squared above 0.9, meaning most traces are well described with the WLC. Finally, for all traces with an R-squared above 0.93, the persistence length is plotted. This time there seem to be two distributions. One centered around 20 nm and one around 60 nm. In both regimes, the traces seem to be described equally well by the WLC (figure 4.2b. and Suppl. figure 6.3) . Also the mean fit error is on the order of magnitude of 2 nm, hence these different distributions are not expected to be caused by fitting artifacts. The first distribution being centered at about half the persistence length of the second distribution, could make one think of double tethered beads. However, the same measurement with a 5 times lower DNA concentration is shown in Suppl. figure 6.4. In this measurement the persistence length results did not look much better, which you would expect for lower DNA concentration in the case of double tethers, making this hypothesis quite unlikely. Moreover, the tether density was almost 5 times lower (13 ± 5 mm^{-2}), showing that we did not have a great abundance of DNA for the higher DNA concentration measurement. The lower DNA concentration measurement did however have small statistics (N=24), so it might take more measurements to fully discard the double tether hypothesis.

4.2 HTkB-DNA complex

After the analysis of the different bare DNA's, we added HTkB to the Dig DNA 2, to study the HTkB-DNA complex. Results looked very familiar to HMfB, which can be seen in figure 4.3a. Three regimes can be distinguished, one described by an FJC (equation 2.4, one by a WLC with an adjusted persistence length (equation 2.5 and one normal WLC, corresponding respectively to the hypernucleosome conformation, the beads-on-a-string conformation and the bare DNA conformation (I, II and III respectively in figure 4.3c.). The transition from regime 1 to 2 starts around 4 pN and the transition from 2 to 3 around 20 pN. The last regime does deviate from the expected WLC. For bare Dig DNA 2 however, the curve also did not follow the expected WLC (but rather one with a much shorter persistence length), which may explain this.
The black line shows the fit with the hypernucleosome model, which describes the curve quite well, including the transitions between the different regimes.
Not all curves looked like figure 4.3a. About 75% of all traces with dZ $>$ 0.8 μ did not show three clearly distinguishable regimes (Suppl. figure

6.7). They did show a kink, characteristic for the unstacking transition, but for most of these traces this kink happened at a force of 7 pN and higher instead of around 3 or 4 pN. As this is around twice the expected rupture force, it could well be, that these were double tethered beads. This could also explain the missing regimes, as you would then expect the second transition at 40 pN instead of 20 pN and our measurements only go until 35 pN.
We selected the beads with three clearly distinguishable regimes and with an unstacking kink in the range 2-5 pN.
For these beads, we fitted the hypernucleosome model (equation 2.3). The fixed parameters were the contour length $L = 3248$ bp and the persistence length $A = 50$ nm. We fitted the dimer height z_{dimer}, the DNA deflection *angle*, the dimer footprint L_{dimer}, the stacking energy $G1$, the wrapping energy $G2$, hypernucleosome stiffness k, the number of dimers N_{dimer} and the bead offset. The results are plotted as histograms in figure 4.3b., and the mean values and standard deviations are listed in table 4.1.

G1 (k_BT)	4.6 ± 1.4
G2 (k_BT)	6 ± 3
k (pN/nm)	1.0 ± 0.9
angle (deg)	13 ± 3
Ndimer (-)	105 ± 21
Ldimer (bp)	29 ± 3
Zdimer (nm)	3.5 ± 3.3

Table 4.1: *Mean and standard deviations of fitted parameters with equation 2.3, N = 80*

The stacking energy seems to be a bit higher for HTkB than for HMfB and certainly higher than for HMfA. The wrapping energy has a large standard deviation but seems to be close to the wrapping energies of both HMfA and HMfB. The fitted deflection angle also gives a value comparable to both HMfA and HMfB.
We found Ldimer values of 29 ± 3 bp. For our 3248 bp DNA strand, this would allow for 112 ± 12 dimers to bind. This agrees with our found Ndimer value of 105 ± 21. However, with these standard deviations it can not be said whether the strands were fully saturated.
The found fit standard errors were very high. For parameters in some fits the standard errors were an order of magnitude higher than the fitted values themselves (see for example the fits in Suppl. figure 6.5 and Suppl.

figure 6.6). This might explain the rather large standard deviations in the results..
The k and Zdimer values that we found are particularly uncertain. The (relative) fit standard errors do not seem to be significantly higher for these parameters than for the other ones. However, k and Zdimer are the values that determine the stretching behaviour of the hypernucleosome conformation and thus the lowest force/extension regime of the force-extension curve. In this regime, factors like surface interactions and the bead offset may play a roll. Also, any small deviation in the force, due to for example a not entirely homogeneous magnetic field, will have a relatively high impact in this small force regime.
Especially the Zdimer values seem to be almost randomly distributed. A clear difference is observed when comparing curves with low Zdimer (Suppl. figure 6.5), to those with high Zdimer (Suppl. figure 6.6). The higher Zdimer values reflect a more bended shape in the FJC regime of the curve, where the lower Zdimer values reflect a much straighter shape. The reason of this different behaviour remains unknown.
However, apart from Zdimer, the fit results give us a good idea of the HTkB-DNA complex characteristic values.

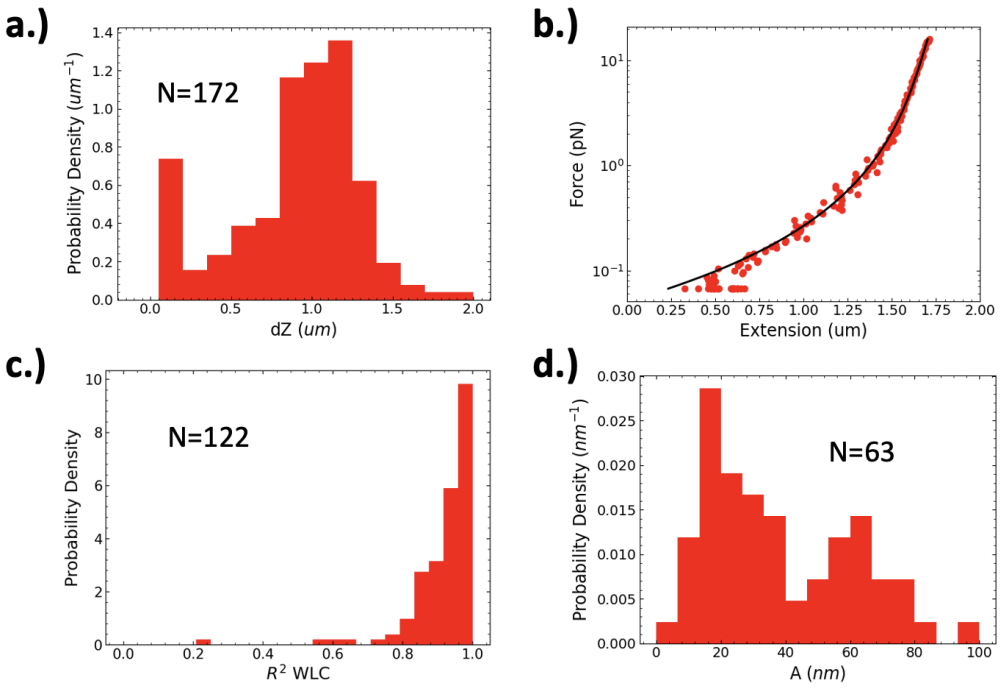


Figure 4.2: *Fitted Worm Like Chains for DBCO DNA reveal two different distributions of persistence lengths.* ***a.)*** *Histogram of all dZ values of the measured beads in the DBCO DNA experiment (Section 3.4: sticking).* ***b.)*** *Typical DBCO DNA curve. We fitted equation 2.2 with an R-squared of 0.98, which is shown by the black line. The fixed parameter is the contour length = 5217 bp. Fitted parameters are the persistence length = (20.3 $\pm$ 0.7) nm and the bead offset = (0.522 $\pm$ 0.006) μm.* ***c.)*** *Histogram of the R-squared values for all Dig DNA 2 curves with a dZ between 0.7 μm and 2.0 μm.* ***d.)*** *Histogram of fitted persistence lengths of all DBCO DNA curves with a dZ between 0.7 μm and 2.0 μm and an R-squared of 0.93 or higher.*

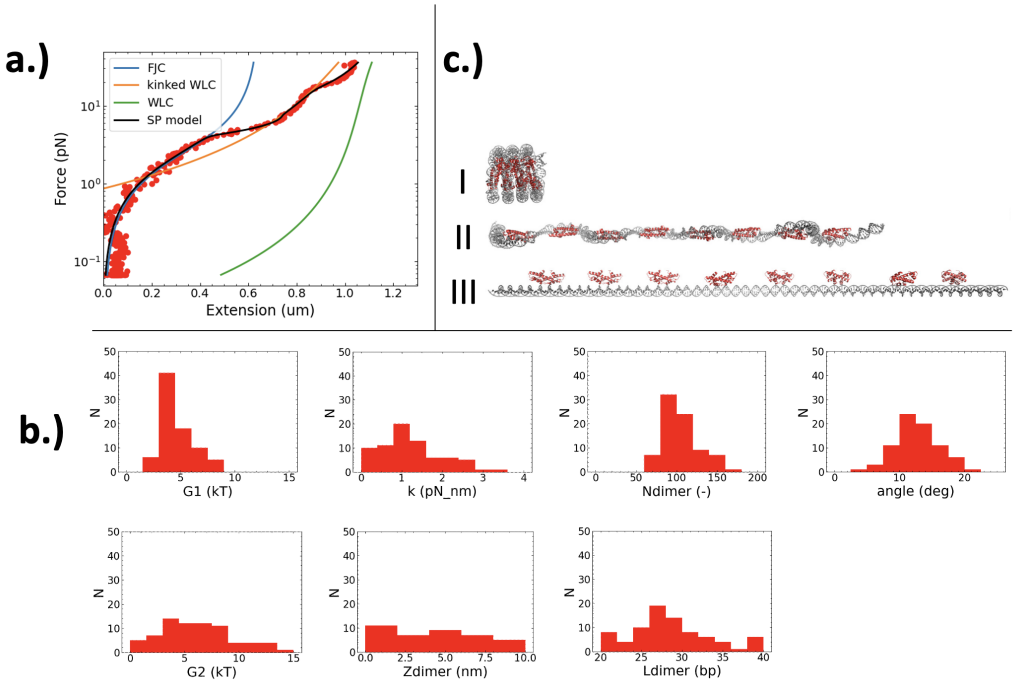


Figure 4.3: *The HTkB-DNA complex seems to be well described by the hypernucleosome model.* ***a.)*** *Typical HTkB-DNA complex force-extension curve. We fitted the hypernucleosome statistical physics model to this curve with an R-squared of 0.995, which is shown as the black line. The fixed parameters were the contour length = 3246 bp and the persistence length = 50 nm. The fitted parameters were the stacking energy G1 = (6 ± 4) k_BT, the wrapping energy G2 = (8 ± 5) k_BT, the deflection angle = (12 ± 8) deg, the dimer height Zdimer = (6 ± 4) nm, the dimer DNA footprint Ldimer = (30 ± 20) bp, the number of dimers Ndimer = (108 ± 73) and the hypernucleosome fiber stifness k = (0.7 ± 0.5) pN/nm. The other three lines use the same fitted parameters to describe the three different regimes of the curve as explained in 2.2.* ***b.)*** *Histograms of the fitted parameters to equation 2.3, N = 80. Statistics are shown in table 4.1.* ***c.)*** *Illustration of the three conformations of the hypernucleosome model, corresponding to the three regimes, mentioned in a.): I the hypernucleosome conformation, II the beads on a string conformation and III the bare DNA conformation.*

Chapter 5

Discussion and Conclusion

For both DBCO-Azide click chemistry tethering and with Digoxigenin body-antibody tethering we achieved force-extension curves that are well captured by the WLC model. The fitted persistence lengths however, were not as expected. For two of our three DNA's we found there to be no or almost no sticking. For the Dig DNA 1 however, all beads were found to be stuck. We found the HTkB-DNA complex to be well described by the Henneman et al. hypernucleosome model. All fitted parameters were found to be in the neighbourhood or at least at the order of magnitude of the HMfA and HMfB values. The most uncertain parameters were the parameters that influence the hypernucleosome regime, being the dimer height and the fiber stifness. Around 75% of the HTkB-DNA tethers showed different behaviour than the rest, with higher dimer unstacking transitions and did not show three clearly distinguishable regimes.
The first result that we will discuss shortly is the sticking for the Dig DNA 1, which we found rather remarkable. Namely, this DNA should be identical to the DBCO DNA apart from the Digoxigenin group instead of the DBCO group. You would thus think that the sticking is caused by the Dig or the Dig tethering mechanism. However, the other Dig DNA did not stick, so apparently that is not the reason either. That leaves us with the conclusion that Dig DNA 1 was not identical to the DBCO DNA and we have to look at the differences between the production protocols. The reason could for example be that in this batch the used nuclease free water from the Promega Wizard kit was contaminated. However, we will not go into this further as follow-up experiments can just use Dig DNA 2 or DBCO DNA, which do not stick.
More important results, are the found persistence lengths, which deviated from the expected persistence length of DNA. The unexpected values of

the fitted persistence lengths can mean one of two things: either there was something wrong with the sample or there was something wrong with the measurement setup/protocol. Both of these causes affect the HTkB results, as those were obtained with the same sample (only with added HTkB) and measured with the same setup.
The simplest explanation would be the case of multiple tethers. For DBCO DNA, we saw two distributions, possibly representing double tethered beads and single tethered beads. For Dig we saw the vast majority of the beads to be at an even lower persistence length. Perhaps there was such an abundance of DNA in this sample that the vast majority had even more than two tethers, leading to persistence lengths even shorter than half the expected length. For the HTkB measurements, which should have the same distribution of single/multiple tethers as the Dig 2 measurements, we also saw a majority with behaviour that might be explained by multiple tethers. The HTkB traces that we selected for the fitting could then correspond to the small percentage of tethers in the Dig results that were found with an expected persistence length of 50 nm. However, this small percentage of normal persistence lengths was significantly lower than the percentage of selected beads in the HTkB measurements, making this hypothesis not too credible.
Apart from a cause in our biological sample, we could also be dealing with an error in the measurement setup/protocol. For example, a wrong magnet trajectory would give you wrong force values, which would then result in a different persistence length. In such cases, you would then expect the same, wrong, persistence lengths for an entire measurement. This is not the case for the DBCO DNA measurements, where traces from both the high persistence length distribution and from the low distribution are found within the same measurement. Also for the HTkB measurements, the 75% different looking curves are found within the same measurements as the normal looking curves. This points at causes within the sample rather than in the setup. Therefore, in a follow-up experiment, the first step would be to check the multiple tethers hypothesis by decreasing the DNA concentrations.
So, as the bare DNA measurements without HTkB already gave strange outcomes, we should not trust blindly on our HTkB results. However, the shape of the curves was correct (WLC shape). Adding HTkB changed the traces from WLC shaped curves to curves that strongly ressemble those found for HMfA and HMfB, and are well fitted with the hypernucleosome model. Regardless of the complications in the bare DNA measurements, this still provides compelling evidence for hypernucleosome formation in the HTkB-DNA complex.

The fitted parameters mostly agreed with those reported for HMfB or obtained from different experiments. For example, Ldimer = 29 $\pm$ 3 bp is consistent with earlier chromatin sequencing experiments on *Thermococcus kodakarensis*, where chromatin particles are found to have variable footprints in multiples of 30 bp [4]. We now know that these chromatin particles are hypernucleosomes consisting of dimers with a footprint of 30 bp.
Also, the found stacking energy is what we could expect based on the predicted stacking interactions by Henneman et al. HTkB is predicted to have 4 stacking interactions (E30-K61, E34-K65, K14-Q48, K26-E58) [6]. HMfA has a stacking energy of 1.4 $\pm$ 0.4 k_BT with two stacking interactions (K31-E62, E35-K62) and HMfB has a stacking energy of 3.5 $\pm$ 0.6 k_BT with three stacking interactions (K30-E61, E34-R65, D14-R48)[7]. The value of 4.6 $\pm$ 1.4 k_BT that we found for HTkB thus fits in the expected trend of higher stacking energy for more (predicted) stacking interactions.
It therefore supports Henneman et al. their proposed criteria for the formation of hypernucleosomes, used in computational analyses of proposed histone structures. To further confirm this, more Archaeal histones should be tested in this way. In particular, ones that have more sequential difference from HMfA, HMfB and HTkB. Our experiments are a first step to uncover common mechanisms of Archaeal genome organisation.

Bibliography

[1] Eme, L., Spang, A., Lombard, J., Stairs, C. W., & Ettema, T. J. G. (2017). Archaea and the origin of eukaryotes. In *Nature Reviews Microbiology* (Vol. 15, Issue 12, pp. 711-723). Nature Publishing Group. https://doi.org/10.1038/nrmicro.2017.133

[2] Baumann, P., Qureshi, S. A., & Jackson, S. P. (1995). Transcription: new insights from studies on Archaea. In *Trends in Genetics* (Vol. 11, Issue 7). https://doi.org/10.1016/S0168-9525(00)89075-7

[3] Luijsterburg, M. S., White, M. F., van Driel, R., & Th. Dame, R. (2008). The major architects of chromatin: Architectural proteins in bacteria, archaea and eukaryotes. In *Critical Reviews in Biochemistry and Molecular Biology* (Vol. 43, Issue 6). https://doi.org/10.1080/10409230802528488

[4] Maruyama, H., Harwood, J. C., Moore, K. M., Paszkiewicz, K., Durley, S. C., Fukushima, H., Atomi, H., Takeyasu, K., & Kent, N. A. (2013). An alternative beads-on-a-string chromatin architecture in Thermococcus kodakarensis. *EMBO Reports*, *14*(8), 711-717. https://doi.org/10.1038/embor.2013.94

[5] Mattiroli, F., Bhattacharyya, S., Dyer, P. N., White, A. E., Sandman, K., Burkhart, B. W., Byrne, K. R., Lee, T., Ahn, N. G., Santangelo, T. J., Reeve, J. N., & Luger, K. (2017). Structure of histone-based chromatin in Archaea. *Science*, *357*(6351), 609-612. https://doi.org/10.1126/science.aaj1849

[6] Henneman, B., van Emmerik, C., van Ingen, H., & Dame, R. T. (2018). Structure and function of archaeal histones. In *PLoS Genetics* (Vol. 14, Issue 9). https://doi.org/10.1371/journal.pgen.1007582

[7] Henneman, B., Brouwer, T. B., Erkelens, A. M., Kuijntjes, G. J., van Emmerik, C., van der Valk, R. A., Timmer, M., Kirolos, N. C. S., van Ingen, H., van Noort, J., & Dame, R. T. (2021). Mechanical and structural properties of archaeal hypernucleosomes. *Nucleic Acids Research*, *49*(8). https://doi.org/10.1093/nar/gkaa1196

[8] Bustamante, C., Marko, J. F., Siggia, E. D., & Smith, S. (1994). Entropic elasticity of Î»-phage DNA. In *Science* (Vol. 265, Issue 5178). https://doi.org/10.1126/science.8079175

[9] Bustamante, C., Smith, S. B., Liphardt, J., & Smith, D. (2000). Single-molecule studies of DNA mechanics. In *Current Opinion in Structural Biology* (Vol. 10, Issue 3). https://doi.org/10.1016/S0959-440X(00)00085-3

[10] Meng, H., Andresen, K., & van Noort, J. (2015). Quantitative analysis of single-molecule force spectroscopy on folded chromatin fibers. *Nucleic Acids Research*, *43*(7). https://doi.org/10.1093/nar/gkv215

[11] Strick, T., Allemand, J. F., Croquette, V., & Bensimon, D. (2000). Twisting and stretching single DNA molecules. In *Progress in Biophysics and Molecular Biology* (Vol. 74, Issues 1-2). https://doi.org/10.1016/S0079-6107(00)00018-3

[12] Kulic, I. M., Mohrbach, H., Thaokar, R., & Schiessel, H. (2007). Equation of state of looped DNA. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, *75*(1). https://doi.org/10.1103/PhysRevE.75.011913

[13] Kim, E., & Koo, H. (2019). Biomedical applications of copper-free click chemistry: In vitro, in vivo, and ex vivo. *Chemical Science*, *10*(34). https://doi.org/10.1039/c9sc03368h

[14] Lin, Y. T., Vermaas, R., Yan, J., de Jong, A. M., & Prins, M. W. J. (2021). Click-Coupling to Electrostatically Grafted Polymers Greatly Improves the Stability of a Continuous Monitoring Sensor with Single-Molecule Resolution. *ACS Sensors*, *6*(5). https://doi.org/10.1021/acssensors.1c00564

[15] Brouwer, T. B., (2020). The Role of Linker DNA in Chromatin Fibers [PhD thesis, Leiden University]. https://hdl.handle.net/1887/138082

[16] Lowary, P. T., & Widom, J. (1998). New DNA Sequence Rules for High Affinity Binding to Histone Octamer and Sequence-directed Nucleosome Positioning

Chapter 6

Supplemental Figures

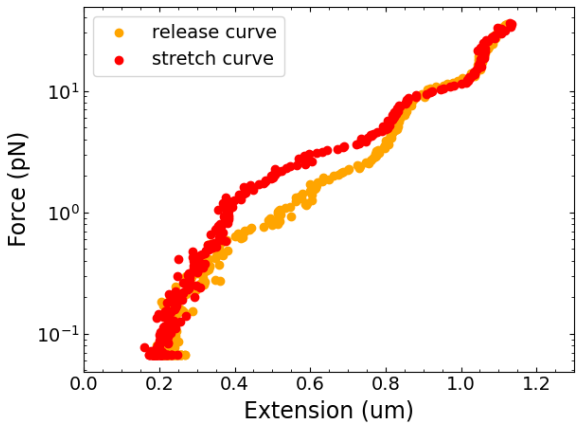


Figure 6.1: *Stretch- and releasecurve of a HTkB-DNA complex on Dig DNA 2. Hysteresis is observed between unstacking and restacking of the hypernucleosome, which in equilibrium, should happen at the same force.*

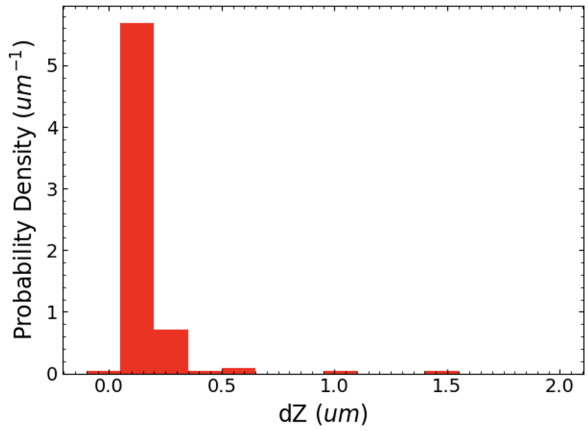


Figure 6.2: *dZ histogram of Dig DNA 1. Barely any tethers are stretched more than 0.3 μm, showing that practically all beads were stuck.*

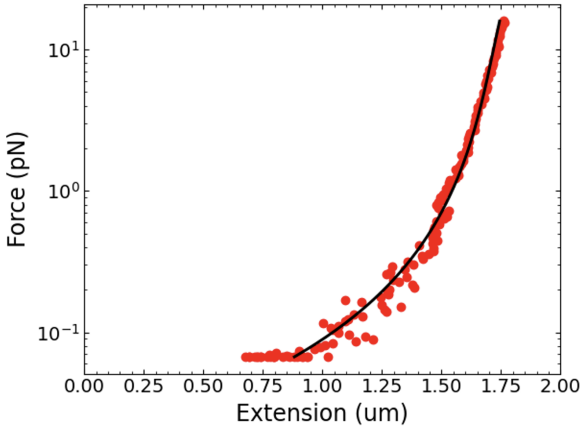


Figure 6.3: *Example of a DBCO DNA force-extension curve with a large persistence length. The fit to equation 2.2 is shown by the black line and gave a persistence length of (61 $\pm$ 3) nm and a bead offset of (0.831 $\pm$ 0.004) μm, with an R-squared of 0.97.*

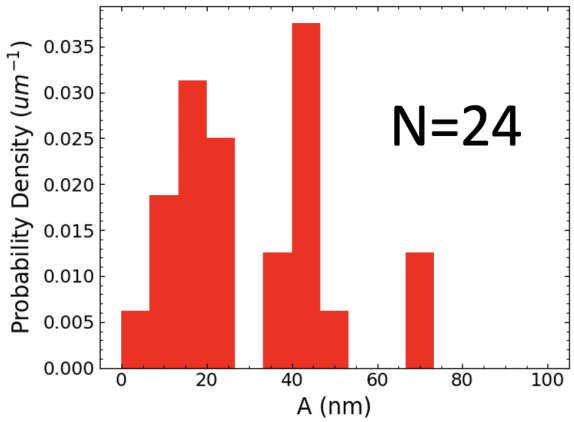


Figure 6.4: *Histogram of fitted persistence lengths of DBCO DNA curves in measurements with a 5 times lower DNA concentration than the concentration in the main DBCO DNA measurements, of which the results are shown in figure 4.2. The histogram only shows the tethers with a dZ between 0.7 μm and 2.0 μm and an R-squared of 0.93 or higher.*

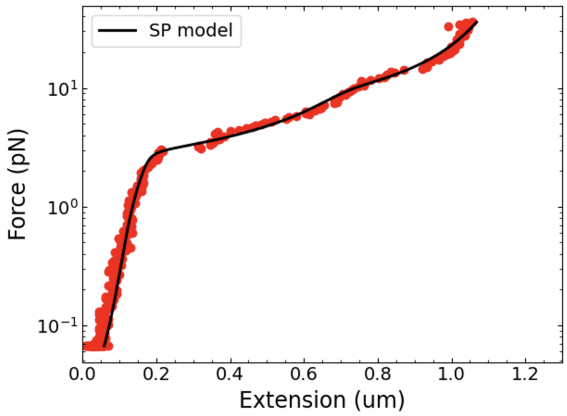


Figure 6.5: *Example of a HTkB force-extension curve that is fitted with a low Zdimer of (0.8 $\pm$ 4.5) nm. The hypernucleosome regime of the curve has a very straight shape. The other fitted parameters of the curve shown as the black line are: k = (6 $\pm$ 29) pN/nm, deflection angle = (10 $\pm$ 49 deg, Ndimer = (144 $\pm$ 712), Ldimer = (20 $\pm$ 99) bp, stacking energy = (3 $\pm$ 16) k_BT, wrapping energy = (6 $\pm$ 30) k_BT and bead offset = (33 $\pm$ 7) nm, with an R-squared of 0.996.*

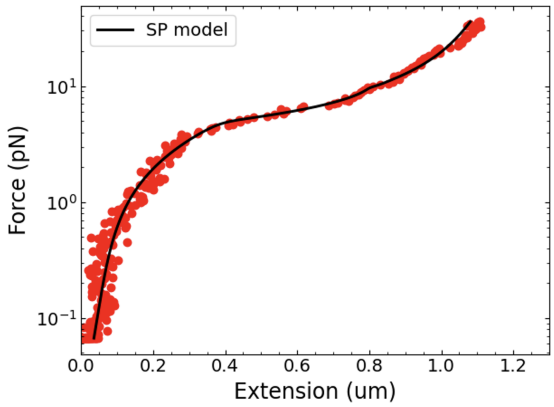


Figure 6.6: *Example of a HTkB force-extension curve that is fitted with a high Zdimer (8 ± 73 nm). The hypernucleosome regime of the curve has a much more bended shape. The other fitted parameters of the curve shown as the black line are: k = (1.4 ± 13.2) pN/nm, deflection angle = (12 ± 111) deg, Ndimer = (102 ± 949), Ldimer = (30 ± 275) bp, stacking energy = (7 ± 65) k_BT, wrapping energy = (6 ± 57) k_BT and bead offset = (0 ± 8) nm, with an R-squared of 0.996.*

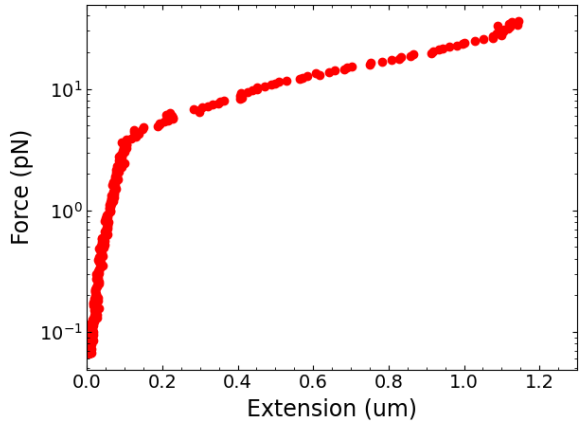


Figure 6.7: *Example of a HTkB curve that does not show three distinguishable regimes. Curves like these were not selected for the further analysis of the HTkB curves.*

Version of June 16, 2022– Created June 16, 2022 - 14:57