G–Quadruplex structure in double–stranded DNA studied by a combination of FRET and Magnetic Tweezers

THESIS

submitted in partial fulfillment of the
requirements for the degree of

MASTER
in
PHYSICS

Author :	N.M. Fennet
Student ID :	0934895
Supervisor :	B.E. de Jong, MSc
	Prof.Dr.ir. S.J.T. van Noort
2^{nd} corrector :	Dr. S. Semrau

Leiden, The Netherlands, November 1, 2017

G–Quadruplex structure in double–stranded DNA studied by a combination of FRET and Magnetic Tweezers

N.M. Fennet

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

November 1, 2017

Abstract

Next to its well–known helix structure, double stranded DNA can form alternative structures that might have biological importance. For example, in guanine–rich DNA sites of the *c–MYC* promotor a second order structure called a G–Quadruplex has been found. In the G–Quadruplex, one strand of the DNA forms a stack of 4 interacting guanines. In this thesis we study the formation of G–Quadruplexes in double–stranded DNA using a combination of Föster Resonance Energy Transfer (FRET) and multiplex Magnetic Tweezers (MT). Moreover, a two–state model was developed which describes the probability to form a G–Quadruplex in double–stranded DNA. Using this model we calculated how the extension and the FRET efficiency depends on force, twist and the sequence of the DNA. Because the synthesis of double–stranded DNA containing a G–Quadruplex site proved challenging, the experimental data could not be compared to the outcomes of the two–state model. Based on simulations we conclude that adding a 3–bp mismatch to the DNA tether next to the G4 site is required for the formation of a G–Quadruplex in dsDNA. Our findings may be relevant for understanding a link with transcription and/or replication.

Acknowledgment

I would like to thank Babette de Jong, John van Noort, Artur Kaczmarczyk, Nicolaas Hermans, Thomas Brouwer, Christine Martens, Redmar Vlieg and Chi Pham for stimulating discussions and helpful suggestions. A special thanks to Babette for all her help during the project and with this thesis.

Contents

1 Introduction **7**

2 Materials and Methods **15**
2.1 DNA construct synthesis 15
2.2 Flow cell preparation 16
2.3 Magnetic tweezers 17
2.4 Experimental procedure 18
2.5 Data analysis 18
2.6 Simulations 19

3 Theory **21**
3.1 Three–state model 22
3.2 Two–state model including G4 23
3.3 Stretch and twist energy 24
3.4 Folding of G4 26
3.5 Measurable parameters 27
3.6 Application of the theory 28

4 Results **33**
4.1 DNA synthesis 33
4.2 Force-extension and twist experiments 37
4.3 Simulations 38

5 Discussion **43**
5.1 DNA synthesis 43
5.2 Force–extension and twist experiments 44
5.3 Simulations 45

6 Conclusion **49**

Chapter 1

Introduction

Since the first discovery of deoxyribonucleic acid (DNA) by Miescher and Schmiedeberg in 1869 [1] and the description of the chemical structure by Watson, Crick, Wilkins and Franklin it has been a topic of extensive research[2]. DNA plays a vital role in our lives because it governs all living cells; determining things from the color of our eyes to the moment a cell dies.
In eukaryotic cells, DNA is in the nucleus where it forms multiple chromosomes. When zoomed in to a smaller scale than the chromosome, DNA is wrapped around nucleosomes. DNA is made of two strands consisting out of nucleotides that interact with each other. A nucleotide consists out of a base, a sugar group and a phosphate group. These nucleotides join via multiple hydrogen bonds to form base pairs (bp), who then form a double helix structure, the natural shape of double–stranded DNA (dsDNA).
There are four different bases in DNA; cytosine **C**, guanine **G**, adenine **A** and thymine **T**. Because of their structure there are two pairs of bases that form a match: **AT** and **CG**. In absence of a complementary strand, a string of the bases forms a single strand of DNA (ssDNA), which occurs at the end of DNA (telomeres) or during DNA replication.
In ssDNA, second–order structures can form. An example of such a structure is a G–Quadruplex (G4). In guanine–rich regions where there are repeats of $[G_{3-4}N_{3-4}]_4$, G4s are formed. Here N denotes a **C, A** or **T**. A G4 consists out of four **G** bases folded into a plateau via Hoogensteen hydrogen bonding in a inter– or intramolecular way. The **G**s are at the four corners of the plateau with in the center a monovalent metallic ion (Figure 1.1A). Multiple plateaus stack on top of each other to form a cubical structure of 0.66x8x8 nm (Figure 1.1B &C) [3, 4]. The **G**s making up the G4 do not necessarily have to be in consecutive order: the non–guanine

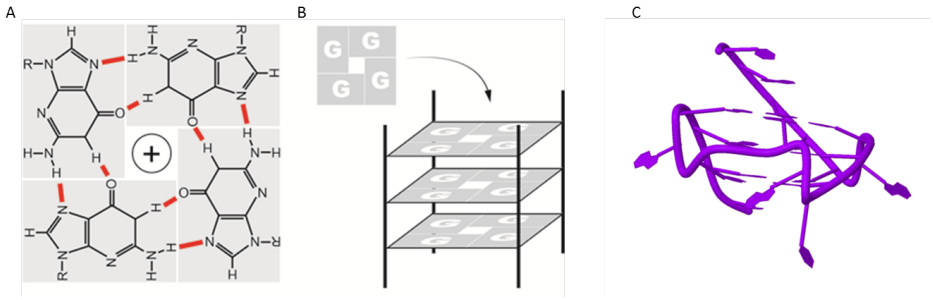

Figure 1.1: Fig. A) and B) modified from Rhodes et al. 2015. Four guanines form a plateau around a positively charged ion and three plateaus stack on top of each other to from a G4. *A) the molecular structure of G4. B) A schematic overview of a folded G4, where the ion is not displayed and the black rods represent the DNA back bone. C) Crystal structure of a G4.*

nucleotides (N) will form loops that stick out of the cubic structure.
The folded G4 structure is known to be stable on large time scales [5, 6]. This is due to the ion in the center which greatly influenced the forming and stability of G4s [7–9]. This ion is usually magnesium, potassium or sodium. However, if there are too many nucleotides N in between the **G**s that form the plateaus of the G4, the loops destabilize the structure, thereby unfolding the G4 [10–13].
Back in 1910, Bang reported that a concentrated solution of guanylic acid formed a gel [14]. Later on in the sixties Gellert *et al* found that these gels contained G–Quadruplex structures [15] . Since then, their properties and especially their function have been extensively studied. Most of these studies where performed *in vitro* [16, 17]. As of recent, research has shown that G4s also form *in vivo* [18, 19]. For example, they can be found in telomeric overhangs [19, 20] and gene–regulatory regions [21, 22]. They can thus form in single–stranded as well as in double–stranded DNA. Furthermore, G4s actually play a role in gene regulation [23–27]. It is no surprise then that G4s have been linked to cancer development [28–33].
One of the first known gene–regulatory regions to contain a G4 is the *c–MYC* promotor. [35–39]. This gene encodes for a transcription factor that activates expression of many genes and is highly conserved. Studies have shown that several different G4 conformations can form in the *c–MYC* promotor, which can be categorized into two groups based on their orientation: *parallel* and *anti–parallel* (Figure 1.2) [40–44]. Figure 1.2A) shows the parallel orientation, here the 5′ and 3′ ends exist the G4 structure at the same side to form a duplex tail. In B) a anti–parallel orientation is shown where the 5′ and 3′ ends exit the G4 structure on opposite sides and do not

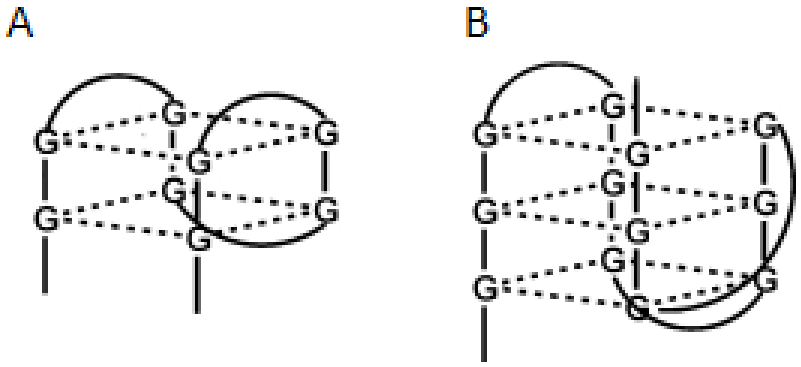

Figure 1.2: The two main configuration groups are parallel folding and anti–parallel folding. *A) An example of the even folding where the 5′ and 3′ ends are in the same side of the G4 structure. B) An example of the odd folding, where the 3′ end exists the G4 structure at the top. Modified from Juskowiak et al[34]*

interact with each other. The different G4 conformations are discriminated based on the order of the base pairs. The order in turn determines the folding energy. Recent studies have shown that the anti-parallel conformation is the most common [45, 46].
For the G4 structure to form in dsDNA, the G4 site must be in a single strand [47]. This means that the base pairs need to be broken. In a cell, this could be done by, for instance, a helicase wich generates a force and/or a negative twist. The negative twist will induce a torque in the dsDNA molecule. When the torque is larger than the energy keeping the base pairs together, the base pairs will unwrap to relieve the torque, thereby forming two ssDNA molecules. Such an experiment has been done with dsDNA *in vitro* at an unwrapping rate of 80 bp s^{-1}, which is comparable to the average transcription speed observed in bacteria *in vivo* [48, 49].
In the studies by Selvam *et al*, it is suggested that a folded G4 can block the transcription and replication of DNA (Figure 1.3) [47, 48]. The G4 does not always block these processes: it depends on its specific configuration. The studies show that relatively weak G4 structures will not stop the transcription and replication processes, but will unfold the G4.
The formation of a G4 structure is not likely because of the annealing of base pairs into dsDNA is energetically much more favorable. However, it is possible for the G4 to fold under certain salt conditions [47, 48]. In these studies, force and twist experiments were also described. However, the force ranges were relatively high and the twist was rather low. These experiments indicated that the force and twist also have an effect on the formation of G4 structures. This raises the question of what the effects of

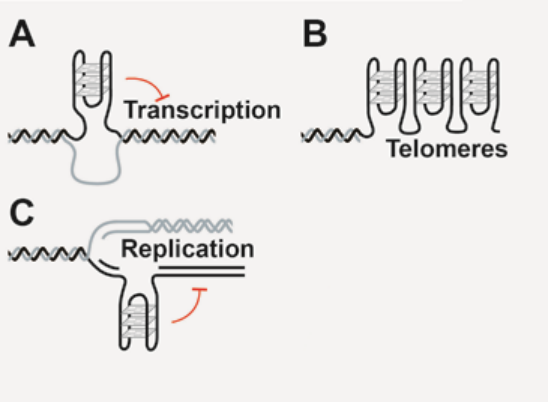

Figure 1.3: G4 in dsDNA plays a prominent role in gene regulation. *In A) and C) a schematic representation of a folded G4 is shown, blocking transcription/replication of that specific region. B) In telomeres, which are G-rich, multiple G4s can form in an array obstructing other proteins in there function. Modified from [50]*

force and torque are and how we can research them.

Magnetic Tweezers (MT) are a highly suitable technique to research G4 folding at the single–molecule level [51, 52]. In this technique, a DNA molecule is immobilized between a glass surface and a paramagnetic bead. A force can be applied and the DNA tether will extend. Moreover, a twist can be applied to induce a torque. The extension is measured and can be described by the Wormlike Chain (WLC) model [53]. To detect the extension, the height of the bead is tracked in a microscope by analysis of an interference pattern. Previous studies have shown that it is possible to detect the folding of G4 in ssDNA using a MT setup. In one study, two dsDNA molecules were ligated to a ssDNA molecule containing a G4 [46]. A force clamp was applied and the extension of the DNA tether detected. The study reports that small fluctuations in the extension of the DNA tether were caused by the (un)folding of the G4 structure. These fluctuations were in the order of 1-10 nm.

Föster resonance energy transfer (FRET) is an other technique that has been used to study the formation of a G4 structure. A donor and an acceptor fluorophore are attached to a DNA molecule. If they are close enough together the donor will excite the acceptor. The efficiency of the excitation is inversely dependent on the distance between the two fluorescent molecules to the sixth power. The strong distance dependency makes FRET a good technique to determine changes on a nanometer scale. In a particular study a donor was placed at an outer plane of the G4 in ssDNA[54]. Proteins that only bind to a specific G4 folding configuration were labeled with an acceptor. The proteins were introduced to the ssDNA and different configurations of G4 were detected. The conclusion was that FRET is an ideal way to map the different configurations of G4 structures. The study also suggested that, when the configuration of a given G4 is known, FRET is a unique technique that can be used for experiments on a

G4 substrate.

There are a few examples of studies that have combined the two techniques on G4 structures in ssDNA or used optical Tweezers (OT) instead of MT on dsDNA. In a recent study, an integrated fluorescence and magnetic tweezers setup was used to probe the unfolding of G4 in human telomeric DNA [55]. In this setup, a Total Internal Reflection Fluorescence (TIRF) Microscopy was used. Contrary to a conventional widefield epifluorescence microscope, where a flow cell is illuminated, TIRF induces an evanescent field of 200 - 300 nanometer at the bottom of a flow cell. The evanescence field decreases exponentially. Light rays enter the flow cell at the bottom under an angle and excite the donor fluorophore. The intensity of the donor and the acceptor fluorophore are detected and the FRET efficiency can be determined. Because the paramagnetic beads are fluorescent, TIRF is higly preferable to reduce undesired fluorescence of the beads.
A single strand containing a G4 was attached to two pieces of dsDNA and labeled with a FRET pair. The FRET pair was at the end of the dsDNA so it would detect the folding of G4. A small force, ranging from 0.7 to 2.5 pN, was applied and the FRET efficiency was simultaneously detected. This experiment made it possible to determine structural rearrangements and force-dependent equilibrium and rate constants [55]. They report that long range interactions between DNA and proteins have a critical effect on the global stability of the folded G4, suggesting that only a few base pairs have to be disrupted to destabilize the G4 structure. This study shows that FRET is an excellent technique to detect the folding of G4 at a resolution that can not be reached in a MT setup.

A study on the effects on G4s in dsDNA using magneto–optical tweezers to induce twist was reported by Selvam *et al* [48]. A dsDNA molecule containing a g-rich region in the center was immobilized between a magnetic and a polystyrene bead. Both beads were placed in different laser traps and by adjusting the distance between the traps a force was induced in the DNA. Using a magnet to apply a twist on the magnetic bead, the twist of the DNA molecule was altered and a force–extension curve was obtained. It was reported that hysteris in the folding/unfolding curves occurred due to the (un)folding of the G4 and that this effect increases as more negative twist is induced. The conclusion was that negative twist showed the second best results to induce G4 folding. The best results were obtained by changing the buffer conditions to high salt. Interestingly, the opposite strand, containing a C–rich region formed an i–motif, compara-

ble in size (number of base pairs en length) and function to a G4. This study shows that applying a negative twist via MT enhances the folding of G4 in dsDNA

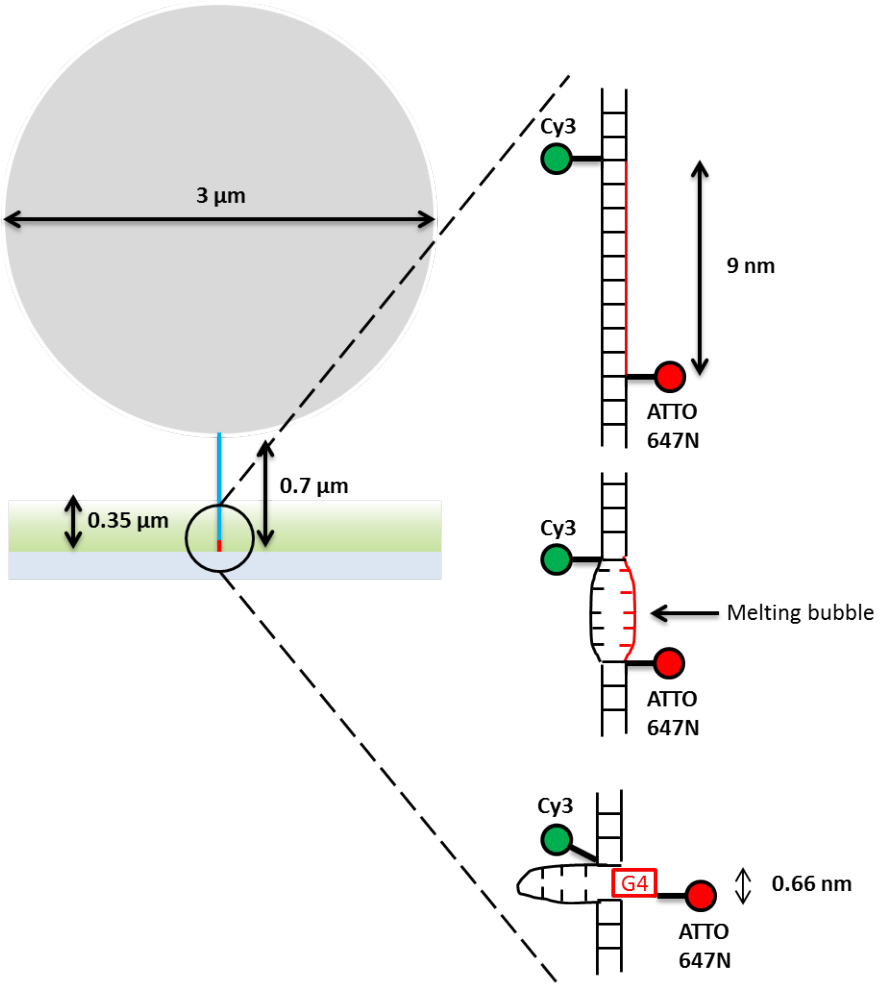

Figure 1.4: Schematic overview of the experiment with typical lengths scales denoted.The G4 is in the range of the TIRF field and if the melting bubble is the size of the G4 sequence FRET can be detected. *Applied twist results in the formation of a melting bubble and when the melting bubble includes the total G4 sequence it is possible for the G4 to fold.*

Our goal is to combine MT with FRET on dsDNA containing a single G4 sequence to research G4 folding and to probe the effects on the extension and stability of the DNA molecule. As of yet, this has not been done. We propose an experiment where a torsionally constrained dsDNA

molecule is tethered between the bottom of a flow cell and a paramagnetic bead in a MT (Figure 1.4). A donor and acceptor (FRET pair) are placed surrounding the G4 region, which is located at the bottom of the DNA tether (red) in the TIRF field. If the paramagnetic bead is in the TIRF field, it will give a bright response and the FRET signal is not detectable. To prevent this, a spacer is introduced to add length to the DNA tether so that the bead is out of the TIRF field upon extension (blue).
When a negative twist is applied to the DNA tether, torque will increase and a melting bubble was formed. A melting bubble consists of consecutive base pairs that are in a single–strand state. When the G4 region is completely melted, it can fold and the distance between the FRET pair decreases, increasing the FRET efficiency. During twisting the extension of the molecule is monitored. We can couple the change in extension to the increase in FRET efficiency and thus accurately detect G4 folding in dsDNA.
In this thesis we show the synthesis of a dsDNA construct containing one G4 site and a Cy3–ATTO647N FRET pair. The DNA construct was used in pulling and twisting experiments and the fluorescence signal from the labels was tested. To further understand the G4 folding, a theoretical model was developed. From the simulations, the probability for the G4 site to melt and to fold were determined, as well as the extension and FRET efficiency. The MT and fluorescence label measurements may contribute to a further understanding of the formation of G4, and their effect on biological processes, in dsDNA.

Chapter 2

Materials and Methods

2.1 DNA construct synthesis

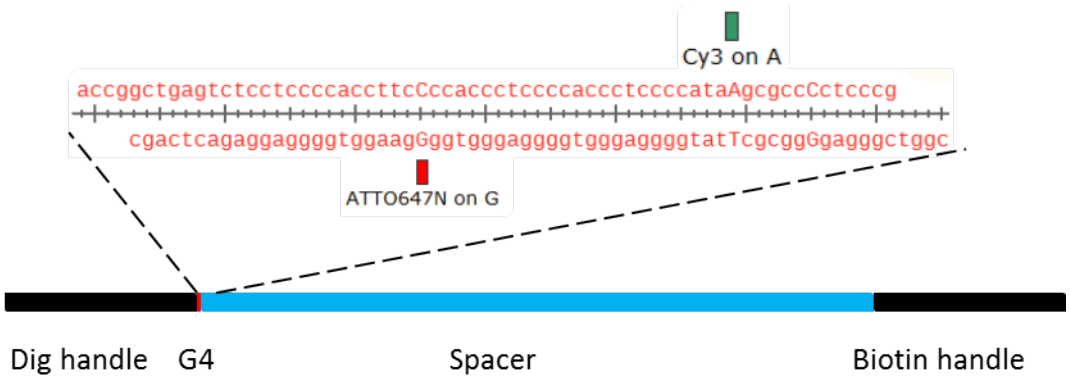

Figure 2.1: The G4 sequence (red) is placed near the DIG handle that will attach to the bottom of the flowcell. *The different components are all on scale. Fluorescent labels are spaced 25 base pairs and surround the G4 site.*

The studied DNA construct contained a G4 site that consists out of two oligos (both at a concentration of 100 μM) with a 4bp overhang:

5′ ACCGGCTGAGTCTCCTCCCCACCTTCCCCACCCTCCCCACCCTCCCCATAaGCGCCCCTCCCGACCG 3′

where the small a was labeled with Cy3 and

5′ CGGTCGGGAGGGGCGCGTTATGGGGAGGGTGGGGAGGGTGGgGAAGGTGGG
3′

where the small g was labeled with ATTO647N. The 2060 bp spacer was restricted from pU27 plasmid with BsaI and BseYI. The handles containing either digoxygenin– or biotin–modified U in a ratio of 1:20 to unmodified T were constructed by PCR from pUC57 plasmid and subsequently digested with BsaI and BseYI. The handles, spacer and the annealed oligos were ligated in one reaction using a T4 ligase. The final product was tested using gel electrophoresis in a 1% 1x TBE agarose gel and imaged using (BioRad, # 731BRO2035) after Ethidium Bromide staining. Cy3 was imaged using green LED illumination and 650/50 nm detection filter, ATTO647N was imaged using red LED illumination and 695/55 nm detection filter, and FRET was imaged using green LED illumination and 695/55 nm detection filter. Enhancement of G4 formation in response to salt was tested on a 1% 1x TBE agarose gel containing a ladder (2 µl sm0331 ladder + 2 µl 6x LD blue), the final product (20 µl G4) and the annealing product (4 µl) at a 90 V for 2 hours in the absence of salt. The gel was imaged with Cy3 and FRET channels. The same gel was soaked for an hour in a high salt bath (50 mM Tris, 50 mM NaCl, 150 mM KCl, 2 mM MgCl based on [52]) and again imaged with Cy3 and FRET channels.

2.2 Flow cell preparation

A clean coverslip was mounted on a polydimethylsiloxane (PDMS) flowcell containing a 2x40x0.4 mm flow channel and flushed with 1 mL MilliQ. Next, the coverslip was coated by flushing in 10 µL anti–digoxygenin (0.25 µg/µL) in 300 µL MilliQ solution and incubated overnight in the fridge. Then the flow cell was passivated with 1 mL of passivation buffer (950 µL 4% bovine serum albumin (BSA) + 50 µL 2% Tween20) and incubated for 2 hours at room temperature. Then the flow cell was rinsed with 1 mL measurement buffer (MB) (10 mM HEPES, 100 mM KCl, 10 mM NaN_3, 0.1%Tween20). 5 µL DNA stock (approximately 1 ng mL^{-1}), 0.5 µL streptavadin–coated paramagnetic bead (diameter 2.8 µm, M270, Invitrogen) and 50 µL MB where mixed together for 10 minutes. Next, 250 µL MB was added and the volume was flowed into the flowcell and incubated for 10 minutes at room temperature. Finally, the flowcell was gently rinsed with 500 µL MB.

2.3 Magnetic tweezers

The MT are shown schematically in Figure 2.2. The extension of the DNA tether was imaged with a home–build multiplexed magnetic tweezers using a collimated LED ($\lambda = 595$ nm), a 40x oil objective (f = 25 mm, NA 1.3) and a CMOS camera (5120x5120 pixel frame rate 30 Hz, CMOS Vision Condor). The LED was placed on top to illuminate the bead. The resulting interference pattern of the bead is used to detect the height z and the center of the bead in the xy–plane. Computer–generated images of the bead containing exactly one spatial frequency per image were cross correlated with the experimental data. The maximum of the cross correlation is the center of the bead (xy–position). At the center of the cc–image the phase is determined. The phase at each z height is corded priot to an experiment to create a look–up table. During an experiment, the shift in phase at each image is used to obtain the height z. Time traces of the tether height were obtained.
The magnets positions was controlled via a stepper motor–based translation stage (M-126, Physik Instrumente) and the height dependent force was determined using [56]:

$$F(z) = F_{max}\left((1-\alpha)e^{\frac{-z}{L_1}} + \alpha e^{\frac{-z}{L_2}}\right) + F_0. \tag{2.1}$$

Here α is 0.3, L_1 and L_2 are the first and second decay length (1.4 and 0.8 mm respectively), F_0 is 0.01 pN and F_{max} depends on the bead size, for 2.8µm–beads it is 85 pN. Using a rotation–motor, the magnets can also rotate in the xy plane.
The dsDNA tether was immobilized between a 2.8 µm streptavadin–coated paramagnetic bead (M270, Invitrogen) and an anti–digoxygenin–coated (sheep, Sigma–Aldrich) clean glass coverslip.

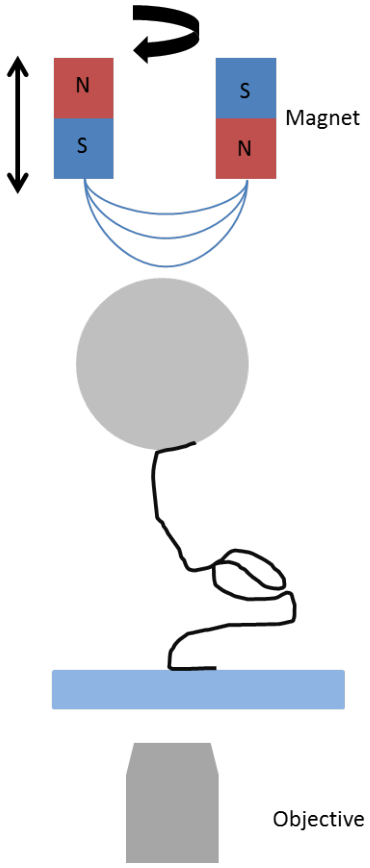

Figure 2.2: Schematic overview of the MT. *The magnets can rotate and move up and down during the detection of the bead.*

2.4 Experimental procedure

For the force–extension experiments the magnet was moved twice from 10 mm (0.1 pN) to 2 mm (15.9 pN) at a rate of 0.5 mm s^{-1} with a 5 second pause in between. For the twist experiments first a postive twist of 40 clockwise (positive) turns ($\sigma = 0.196$) was applied, then to 80 anti–clockwise (negative) turns ($\sigma = -0.196$), back 80 clockwise turns ($\sigma = 0.196$) and finally 40 anti–clockwise turns ($\sigma = 0$) at a rate of 1 turn s^{-1} (0.005 σ s^{-1}). When the stepper motor changes twist orientation there was a 2 second pause.

2.5 Data analysis

Force–extension curves were generated for each tracked tether using in–house written LabVIEW software (National Instruments, Texas). The re-

sulting force–extension curves where fitted using the WLC model. Plots of the force extension and the twist experiments were created using OriginPro 8.5.1.

2.6 Simulations

The in–house written simulations were done using Python 2.7.10 with the NumPY, MatPlotLib, time, csv, collections and itertools packages on a 16 core 2.93 Ghz 4.00 GB RAM pc (Intel). Per force, typical time scales were 12 hours for the simulations without initial melting bubble and 3 hours for with an initial melting bubble.

Chapter 3

Theory

To study the role of G4 in dsDNA a theory was developed based on the three–state model [57]. We used statistical mechanics to calculate the full partition function that describes the DNA as a function of force and change in linking number density. The forming of a G4 and the effects are included. The newly developed theory predicts the probability of G4 folding, the number of melted base pairs and their location, the Boltzmann–weighted extension and the Boltzmann–weighted FRET efficiency.

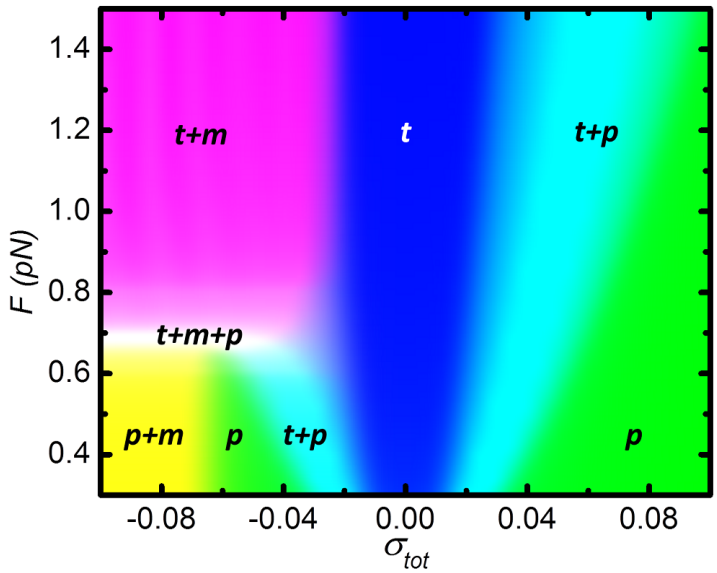

***Figure 3.1: Phase diagram of dsDNA**. For negative twist the plectonemic phase does not occur if the applied force is above 0.7 pN. Modified from Meng et al 2014*

3.1 Three–state model

We will first explain the three–state model and its implications. The three–state model describes the mechanical properties per base pair of a dsDNA molecule based on the coexistence of three states when a twist is applied. The three states are: twisted (t), melted (m) and plectonemic (p). The twisted state is the natural state of a dsDNA molecule when no twist is applied. Here, every 10.4 bp one helical turn is completed. One helical turn is denoted as $Lk_0(= 10.4bp)$. The three–state model predicts the extension of the DNA molecule, the torque and the distribution of the base pairs over the three states. The first state (twisted), the DNA is still double stranded but it will absorb negative or positive twist. The second state is the melted state, the base pair bond is broken and the dsDNA is converted into two single strands. The breaking of the base pairs requires a certain amount of energy denoted as the melting energy ε_m. This energy originates from the torque in the DNA which is induced by twisting. If the torque is too large base pairs will melt to release the tension. The third state is the plectonemic state: dsDNA wraps around itself when a twist is applied. Loops are formed (the plectonemes) and the extension is severely reduced. In the p state the DNA is still in a double–stranded form. The states can coexist simultaneously in a DNA molecule, but a base pair can only be in exactly one of the states. Adding the number of base pairs per state results in the total number of base pairs in the DNA molecule:

$$N = n_t + n_m + n_p. \tag{3.1}$$

When a dsDNA molecule is torsionally constrained, every twist applied to the molecule (ΔLk) will be distributed among the three states. The resulting torque is the same everywhere in the DNA molecule, regardless of the state of the base pairs. However, the response of the base pairs to the torque depends on the state and the twist is distributed based on this property. This means that the twist distribution can not be determined prior to the calculation. But we know that the twist is conserved in the molecule:

$$\Delta Lk = \Delta Lk_t + \Delta Lk_m + \Delta Lk_p. \tag{3.2}$$

The change in twist can be described using the linking number density $\sigma = \Delta Lk/Lk_0$:

$$\sigma_{tot} = \frac{n_t}{N}\sigma_t + \frac{n_m}{N}\sigma_m + \frac{n_p}{N}\sigma_p. \tag{3.3}$$

Here, σ_i denotes the linking number density in each state. Please note that $\sigma = -1$ means that all the original twist is removed from the DNA

molecule. The advantage of σ is that we can compare different lengths of DNA to each other. Instead of the number of twists, we can use the change in linking number density.
The total free energy G can now be described using the free energy per state G_i:

$$G = \sum_{i=t,m,p} n_i G_i. \tag{3.4}$$

The general form of G_i is described by:

$$G_i(F, \sigma_i) = -g_{i,stretch}(F) + \frac{g_{i,twist}(F)}{2}(\sigma_i + \sigma_{0,i})^2 + \varepsilon_i. \tag{3.5}$$

Here, $g_{i,stretch}$ is *the stretching free energy*, $g_{i,twist}$ the *twist energy* and $\sigma_{0,i}$ the degree of twist in absence of torque. For t and p $\sigma_{0,i}$ is zero, but for m it is 1. The term ε_i is the melting energy per base pair, thus zero for t and p and non–zero for m. The three–state model has an average melting energy per base pair, ε_m = 1.6 k_BT. Using the total free energy the probability to be in a conformation can be determined with Boltzmann statistics:

$$P(F, \sigma_{tot}) = Z^{-1} e^{\frac{-G(F, \sigma_{tot}, \sigma_t, \sigma_p, n_t, n_p)}{k_B T}}. \tag{3.6}$$

Here, Z is the partition function defined as:

$$Z = \sum_{\sigma_p=-\infty}^{\infty} \sum_{\sigma_t=-\infty}^{\infty} \sum_{n_p=N-n_s}^{N} \sum_{n_s=0}^{N} e^{\frac{-G(F, \sigma_{tot}, \sigma_t, \sigma_p, n_t, n_p)}{k_B T}}. \tag{3.7}$$

Using these formulas the distribution of the base pairs over the three states can be calculated. Each state has its own characteristic extension and response to force and applied twist, making it possible to predict the mechanical response on DNA in MT.

3.2 Two–state model including G4

Based on the three–state model we developed our own model where we have added the G4 forming. To do so, the melting energy in the model is made sequence depended, so we do not assume an average melting energy per base pair. In this way we can compute the mechanical properties of each individual base pair as long as we know the exact DNA sequence.

Consecutive bases	**Melting energy ε_m (k_BT)**
T A	-0.20
T G or C A	-1.30
C G	-2.40
A G or C T	-2.15
A A or T T	-1.73
A T	-2.12
G A or T C	-2.77
C C or G G	-3.28
A C or G T	-3.40
G C	-4.50

Table 3.1: The melting energy id dsDNA depends on the its di–nucleotide sequence. Modified from www.wikipedia.nl

For the G4 to fold the dsDNA molecule needs to be in the melted state so that the two single strands are detached. Only when all the base pairs forming the G4 sequence are in the melted state the G4 can fold. Therefore in our model we only use negative twist at forces above 0.7 pN, excluding the plectonemic state and only including the twisted and melted state. This model is denoted as the *two–state model*.
An average melting energy as used in the three–state model is not applicable in the two–state model. In fact, ε_m needs to be sequence dependent to give a correct value. Because an **AT** base pair has two hydrogen bonds and a **GC** has three, the **GC** base pairs (and thus the G4) will require a higher melting energy compared to **AT**. In the two–state model the melting energy is determined based on its di–nucleotide sequence (Table 3.2). The melting energy is only included in the melted state and we will adjust the formulas for this state to include sequence dependency.

3.3 Stretch and twist energy

The free energy per state (3.5) consists out of multiple energy terms (Table 3.2). The stretch energy $g_{i,stretch}$ is described by the free energy worm–like chain and is a function of force and persistence length. The stretch energy is multiplied with the extension z_i, given by the worm–like chain

model (WLC) to be:

$$z_i = L\left(1 - \frac{1}{2}\sqrt{\frac{k_B T}{P_i F}}\right). \tag{3.8}$$

Here, L is the contour length per base pair which is 0.33 nm. Because the persistence length differs for double–stranded and single–stranded DNA, the extension for the two states is different and due to twist the extension of the total DNA tether will change. In the melted state there are two strands, thus the force should be divided by two. But we know from other studies that the used persistence length for the melted state is too high [58–60]. Therefore, to calculate the free energy we have multiplied the free energy terms with L_i in stead of z_i, which are the values for dsDNA and overstretched DNA.
The twist energy $g_{i,twist}$ is described in a similar way as the free elastic energy of a spring. The free elastic energy described by the integrated Hooks Law, $\frac{1}{2}kx^2$. In our model the place x is given by the σ and the spring constant is replaced by $g_{i,twist}$. $g_{i,twist}$ can thus be regarded as a twist constant. The two states have a different value for the twist modulus C_i and t depends on the force. There is a small offset of -1 in the melted state: because ssDNA (m) has no intrinsic twist it will become more favorable to be in the melted state when a negative twist is applied.
To compute the corresponding ε_m the base pair index needs to be considered. For this reason a location dependent melting bubble is introduced. A melting bubble consists out of one ore more base pair(s) in the melted state. The energy penalty to create two or more melting bubbles is $\frac{3}{2}k_B T \ln(2n_m)$ per bubble [61]. This suppresses spontaneous melting and we assume that no more than one melting bubble will form per dsDNA molecule. However, there are numerous different possibilities for the melting bubble to form, defined by the number of melted base pairs n_m and their location. Each of these possibilities with its corresponding ε_m, including the case where there is no melting bubble and ε_m is zero, are called *configurations*. Because the total number of base pairs N is preset, the number of base pairs in the twisted state n_t is also fixed (Equation 3.1). So, using the melting bubbles (ε_m) and the stretch and twist energy, the energy per state can be determined (Equation 3.5) and the total free energy per configuration G_{tot} can be computed (Equation 3.4).

	$g_{i,stretch}$ (pN nm)	$g_{i,twist}$ (pN nm)	z_i (nm)	P_i(nm)	L_i(nm)	C_i(nm)
t	$z_t\left(F-\sqrt{\frac{k_BTF}{P_t}}\right)$	$z_t\omega_0^2k_BTC_t\left(1-\frac{C_tk_BT}{4P_t^2F}\right)$	$L\left(1-\frac{1}{2}\sqrt{\frac{k_BT}{P_tF}}\right)$	50	0.34	100
m	$z_m\left(F-\sqrt{\frac{k_BTF}{P_m}}\right)$	$z_m\omega_0^2k_BTC_m$	$L\left(1-\frac{1}{2}\sqrt{\frac{k_BT}{P_mF}}\right)$	4	0.55	28

Table 3.2: The stretch and twist energies for the two states. *P_i is the persistence length, L_i the contour length of a base pair, C_i twist modulus and ω_0 the inverse of the pitch of the double helix with a value of $2\pi/3.6$ nm = 1.75 nm^{-1}*

3.4 Folding of G4

The final part of the theory is to add the forming energy of the G4 structure ε_{G4} to the free energy G_i. ε_{G4} is given by:

$$\varepsilon_{G4}(F) = F\Delta z - \varepsilon_{int} \tag{3.9}$$

and is a function of force. The first term is the work required to decrease the extension of one of the ssDNA strands in the melting bubble. This would be force dependent as $\Delta z(F) = z_{G4}(F) - z_{ss}(F)$. However, literature reports that for force between 1 and 20 pN the decrease in extension is constant and for a G4 structure it is equal to 8 nm [40, 51, 62]. The interaction energy between the different plateaus of the G4 as well as the interaction energy with the central ion yields:

$$\varepsilon_{int} = -17k_BT. \tag{3.10}$$

This is an average number based on [63–66]. Note that the signs of the terms are opposite and that ε_{G4} is approximately zero for $F = 8.5$ pN. The folding energy is added to the free energy G_{tot} when the base pairs in the G4 sequence are all in the melted state.
The probability for a G4 structure to form depends on the free energy of the DNA molecule and the state of the base pairs of the G4 region:

$$P_{fold,G4} = P_{m,G4}\frac{e^{-\varepsilon_{G4}(F)/k_BT}}{1+e^{-\varepsilon_{G4}(F)/k_BT}}. \tag{3.11}$$

Here, $P_{m,G4}$ represents the probability that the base pairs in the G4 region (n_{G4}) are in the melted state according to:

$$P_{m,G4} = \prod_{n_{G4begin}}^{n_{G4end}} P(F,\sigma_{tot}). \tag{3.12}$$

3.5 Measurable parameters

We can not measure the state of each base pair in an experiment. However, the results from the two–state model can be compared to experimental data to derive the number of base pairs per state. Using 3.11 the Boltzmann–weighted extension ($< z >$) of the DNA tether was calculated. The base pairs can be in the melted or the twisted state, each with its corresponding extension shown in Table 3.2. When the G4 structure forms the extension decreases because the folded G4 structure is a much more compact structure. The number of melted base pairs also decreases because the base pairs in the G4 structure (which must be in the melted state for the G4 to fold) do no longer contribute to the extension of the melted state. However, the size of the folded G4 will contribute to the extension and needs to be included. We distinguish two different situations for the extension of the DNA tether: with (z_{G4}) *and* without (z_{noG4}) the forming of the G4 structure. They are given by:

$$\begin{aligned} z_{noG4} &= n_t z_t + n_m z_m \\ z_{G4} &= n_t z_t + (n_m - n_{G4}) z_m + L_{G4}. \end{aligned} \tag{3.13}$$

Here, L_{G4} is the length of a folded G4 (0.66 nm[62]). Combining these two extensions with the probability that a G4 structure does or does not form (Equation 3.11) results in $< z >$:

$$< z >= z_{noG4}(1 - P_{fold,G4}) + z_{G4} P_{fold,G4}. \tag{3.14}$$

To compare the extension to the theory, the detected extension needs to be averaged.
Since we know the state for each base pair we can calculate the extension of a specific part of the DNA tether to determine the the Boltzmann–weighted FRET efficiency ($< E >$). The FRET efficiency depends on the distance in between the two labels r and is in general given by:

$$E = \frac{1}{1 + (r/R)^6}. \tag{3.15}$$

Here, R is the Föster distance of the donor–acceptor pair. Because r depends on the state of the base pairs and their corresponding extensions, we can distinguish the same two situations as with $< z >$: folded and not folded. The difference is that we now only need to consider the state of the base pairs in between the FRET pair, denoted as $n_{i,FRET}$. This results

in the FRET efficiencies E_{noG4} and E_{G4}:

$$E_{noG4} = \frac{1}{1 + \left(\frac{n_{t,FRET} z_t + n_{m,FRET} z_m}{R}\right)^6}$$
$$E_{G4} = \frac{1}{1 + \left(\frac{n_{t,FRET} z_t + (n_{m,FRET} - n_{G4}) z_m + L_{G4}}{R}\right)^6}. \tag{3.16}$$

Multiplying the two FRET efficiencies of Equation 3.16 with their corresponding probabilities results in $< E >$:

$$< E >= E_{noG4}(1 - P_{fold,G4}) + E_{G4} P_{fold,G4} \tag{3.17}$$

Finally, using the size of the melting bubble n_b and the probability for each melting bubble to occur (Equation 3.6), the average number of melted base pairs and the squared average number of melted base pairs were computed. These numbers are a function of force and σ and given by:

$$< n_m > = \sum_i P_i(F, \sigma) n_b \tag{3.18}$$
$$< n_m^2 > = \sum_i P_i(F, \sigma) n_b^2. \tag{3.19}$$

The squared standard deviation (STD) of the average number of melted base pairs can be calculated using the above formulas:

$$\sigma^2_{STDmelt} =< n_m^2 > - < n_m >^2 . \tag{3.20}$$

Every change in linking number density ΔLk applied to a DNA tether adds torque to the system. If the melting energy would be equal for each base pair the average number of melted base pairs is a linear function because the torque is distributed evenly over the base pairs. However, the melting energy is dependent on the sequence and not equal for each base pair, so for a certain amount of torque the number of melted base pairs can differ. Therefore, the average number of melted base pairs will fluctuate severely when a negative twist is applied.

3.6 Application of the theory

To predict the outcomes of the aforementioned theory, numerical simulations were done in which the probability of all possible configurations

were computed. In order to reduce the number of computations the range for σ_i was set from zero to σ_{min}. Here, σ_{min} is an arbitrarily number, we choose -1. The model is described in the following four steps and these steps are shown in pseudo–code on pages 31.

Step 1a: Compute an array containing all configurations

All possible configurations were stored in an array (A). This array contained the start position of each melting bubble, the end position and the length. If the melting bubble contains the complete G4 sequence the configuration was added twice to include both the folded and the not folded G4. The free energy for each configuration containing the sum of the melting energy ε_m and ε_{G4} were stored in a second array.

Step 1b: Iterate total twist and distribution of the twist

The work terms corresponding to a chosen force and ΔLk_{max} were added to the free energy. We iterated over ΔLk and ΔLk_m, where ΔLk_t is given by: $\Delta Lk_t = \Delta Lk - \Delta Lk_m$. Using the length of the melting bubble and the corresponding energy, the free energy of each configuration was computed with Equation 3.4 and Equation 3.9. This results was an array containing all the free energies per configuration for a given ΔLk distribution.

Step 2a: Compute partition function

Since we are only interested in energy differences, the lowest energy was subtracted. Using Equation 3.7 the partition function Equation 3.7 and the probability for each state Equation 3.6 were computed. The probabilities were then coupled to the original array containing the position of the melting bubble. Next we computed the probability for each specific base pair to be in the melted state (A_m).

Step 2b: Compute number of melted bases and mean of parameters

The final step was to compute $< n_m >$ and $< n_m^2 >$ (Equation 3.18) to determine the average melting number and standard deviation for a given twist and force. They were computed using the number of base pairs in a melting bubble n_b and the probability for that configuration. From these quantities, we were able to compute $< z >$ of the DNA tether and $< E >$

using Equation 3.14 and Equation 3.17, respectively.

Algorithm 1 Step 1a: Compute an array containing all configurations

Input: The DNA sequence, variables from Table 3.1, pulling force
Goal: Compute all possible melting bubbles configurations and their corresponding melting energy
A = []
for n_{begin} = 0, N-1 **do**
  for $n_{end} = n_{begin}, N$ **do**
    Compute $\varepsilon_m(n_{begin}, n_{end})$
    append (A[n_{begin},n_{end}, 0, $\varepsilon_m(n_{begin}, n_{end})$])
    if $n_{begin} < n_{G4begin}$&$n_{end} < n_{G4end}$ **then**
      append (A[n_{begin},n_{end}, 0, $\varepsilon_m(n_{begin}, n_{end}) + \varepsilon_{G4}$])
    end if
  end for
end for
Output: A

Algorithm 2 Step 1b: Iterate total twist and distribution of the twist

Compute optimal distribution of twist
Input:$n_{begin}, n_{end}, N, F, \Delta Lk$
$G_{mech} = \infty$
$\Delta Lk_{bubble} = 0$
$n_m = n_{end} - n_{begin}$
$n_t = N - n_m$
for $\Delta Lk = 0, \Delta Lk$ **do**
  $\sigma_m = \Delta Lk_m / n_m$
  $\sigma_t = (\Delta Lk - \Delta Lk_m) / n_t$
  $G = n_m g_m(F, \sigma_m) + n_t g_t(F, \sigma_t)$
  if $G < G_{mech}$ **then**
    $G_{mech} = G$
    $n_{bubble} = n_m$
  end if
end for
Output: G_{mech}, n_m

Algorithm 3 Step 2a: Compute partition function

Input:G_{mech}, A
$G_{conf} = []$
$P = []$
for i in A **do**
  $G_{configi} = G_{mech}(i, F, \Delta Lk) + \varepsilon_m(i) + \varepsilon_{G4}(i)$
  append[G_{conf}]
end for
$G_{conf} = G_{conf} - min(G_{conf})$
$Z = \sum_i exp(-G_{conf}(i))$
$P = exp(-G_{conf})/Z$
Output: P

Algorithm 4 Step 2b: Compute number of melted bases and mean of parameters

Input: A, P
$< n_m > = \sum_i \left(n_{begin} - n_{end}\right) P(i)$
$< n_m^2 > = \sum_i \left(n_{begin} - n_{end}\right)^2 P(i)$
Compute $< z >, < E >$
Output $< n_m >, < n_m^2 >, < z >$ and$< E >$

Chapter 4

Results

4.1 DNA synthesis

To create a DNA substrate that can be used in the combined FRET/MT, we needed to prepare a dsDNA molecule with a G4 site and a FRET pair. A 66 DNA substrate containing a 22 bp G4 site (Pu27 from *c–MYC*) and a FRET pair was created by annealing two oligos. One oligo contained Cy3 and one contained ATTO647N. Oligo annealing was checked by electrophoresis on an agarose gel (Figure 4.1). After ethidium bromide staining, the band corresponding to the annealed product was detected between 50 and 75 bp, as expected of a product with a length of 66 bp (Figure 4.1A). Direct visualization of the product using the fluorescent labels showed overlap of the Cy3 and ATTO647N signal, as expected, as well as the presence of a second product (Figure 4.1B). This second product was very faintly visible in the ethidium bromide gel just below 150 bp and likely stems from the second product present in both oligo stocks.
Despite the presence of a side product, we proceeded with the ligation of the G4 dsDNA substrate to the immobilization handles and a spacer. The spacers added length to the construct, removing the bead from the TIRF field. Multi–DIG and multi–biotin handles were synthesized using PCR, while the spacer was restricted from the pUC57 plasmid. After ligation, the expected 3444 bp product was purified from gel. The purified product was checked by electrophoresis on an agarose gel (Figure 4.3) where it was compared to the 66 bp dsDNA substrate containing the G4 sequence. Because the 3444 bp product is much longer than the 66 bp dsDNA substrate, the construct is expected to remain in the well and not to run trough the agarose gel. Visualization of the ligation product using the fluorescent labels showed no overlap of Cy3 and ATTO647N, indicating that there are

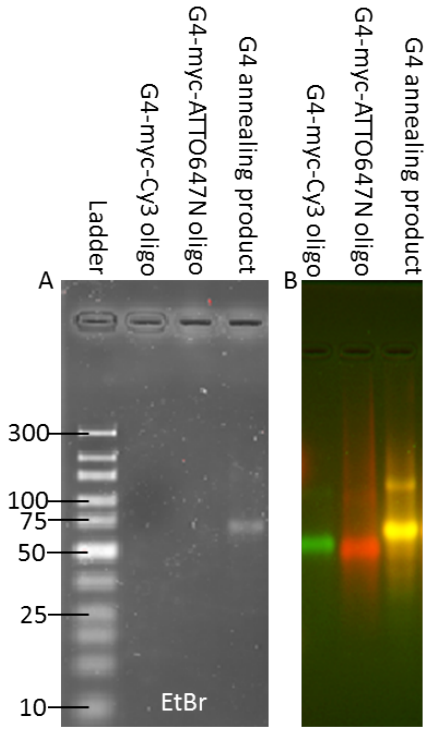

Figure 4.1: ***The G4 annealing product is double stranded and contains a FRET pair.*** *A) The length of the annealing product is between 50 and 75 bp. B) Imaged with Cy3 (green) and ATTO647N (red) filters. The two oligos response for the corresponding filter and FRET is detected in the annealing product. A small second bend is detected in the annealing product lane at around 130 bp.*

no labels in the construct and thus no G4 site. The low concentration of end product can cause experimental difficulties. The field of view in the combined FRET/MT is small which can make it problematic to find DNA tethers. Nevertheless, we continued with this product stock.
To determine the length and the concentration of the ligation product, the G4 construct was loaded onto a agarose gel. For comparison a construct of known concentration (DAXh2) was also loaded onto the same gel. The DAXh2 construct is a 198-bp sequence containing a Widom 601 sequence and a Cy3b-ATTO647N FRET pair spaced 80 bp apart (50 ng μL^{-1}). After Ethidium bromide staining the band for the G4 (red arrow Figure 4.2A) appeared between 3400 and 3500 bp. The expected length was 3444 bp, so the end product was of the correct length. The intensity of the DAXh2 signal was much more intense compared to the G4. We estimated that the concentration of the ligation product to be less than 1 ng mL^{-1}. Visualization of the two constructs using the fluorescent labels shows overlap in the DAXh2 channel and no detectable signal in the G4 construct channel.
To induce the folding of the G4 in the ligation product, the same agarose gel was placed in a high salt bath (2 mM MgCl, 50 mM NaCl, 150 mM KCl and 50 mM tris) for an hour. These are extreme conditions, and because the G4 structure strongly depends on the presence of salt we expected to

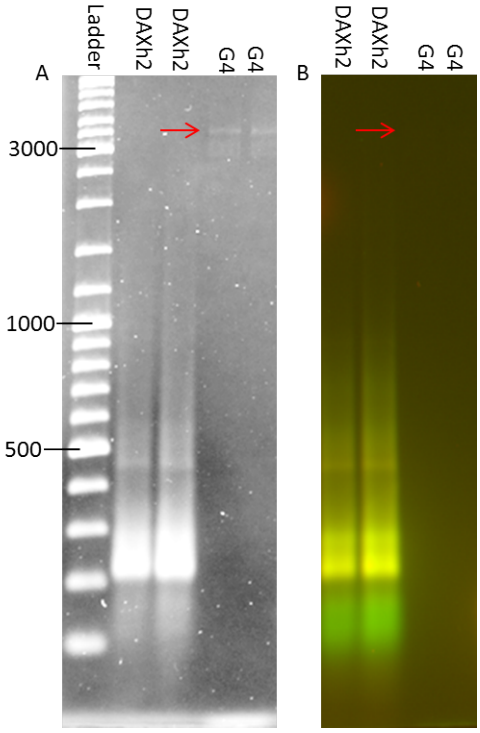

Figure 4.2: The G4 construct was detected at the expected length, but at low concentration and no detectable FRET signal. *In the DAXh2 lanes the sample is a 200 bp long DAXh2 molecule with 2 fluorophore labels at a concentration of 50* $\mathrm{ng\,mL^{-1}}$. *A) The concentration of the ladder was 10* $\mathrm{ng\,mL^{-1}}$ *and when the G4 construct was compared we expect that its concentration was not higher than 1* $\mathrm{ng\,mL^{-1}}$. *B) The gel imaged with Cy3 + ATTO647N filters. No FRET is detected in the G4 lanes. The DAXh2 was clearly visible in both channels because of spectral leakage, the FRET pair is too far apart to have a detectable intensity.*

detect a higher FRET efficiency. The ligation product was visualized using the fluorescent labels(Figure 4.3 C & D). Upon comparison of the intensities before and after the salt bath no difference in intensity was detected. We concluded that the G4 did not fold under the high salt conditions. This indicated that during the ligation process the G4 annealing product was not ligated to the anti-DIG handles and the spacer. Despite that the concentration of the G4 construct was less then 1 $\mathrm{ng\,mL^{-1}}$, it was the proper length and we could continue with force spectroscopy.

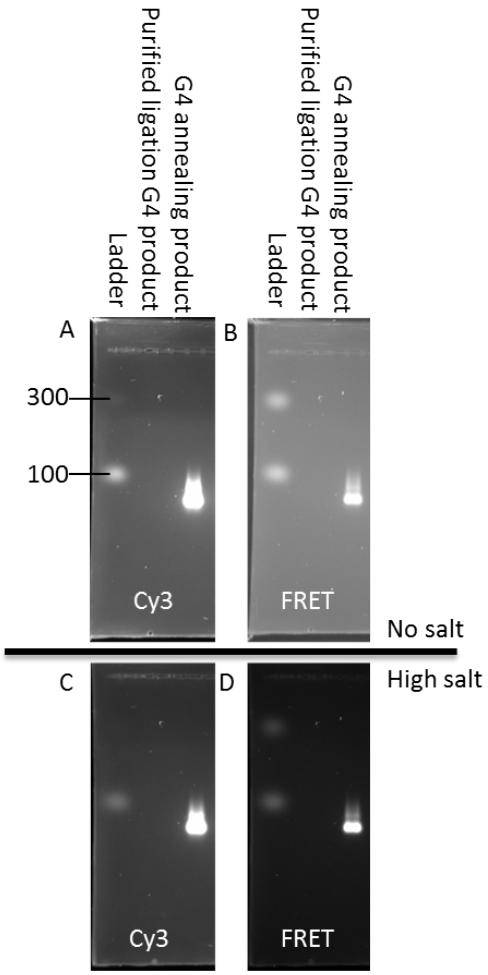

Figure 4.3: The ligation product has a low intensity compared to the annealing product and soaking the gel in a high salt bath does not affect the intensity. *The ladder on the left contains 2* mL *of 100* $\mathrm{ng\,mL^{-1}}$*. A high concentration of annealing product was present, larger than 100* $\mathrm{ng\,mL^{-1}}$*. A) Imaged with Cy3 filter; a Cy3 fluorophore was detected in the annealing product, but not in the ligation product. B) No overlap of the Cy3 and ATTO647N signal detected for the ligation product and spectral leakage was present. C) and D): the gel was soaked for an hour in a high salt bath. There was no detectable difference between A) & C) and B) & D). Please note that the intensity of figures A and C can not be compared to figures B and D respectively due to different exposure times and filters.*

4.2 Force-extension and twist experiments

To further test our G4 construct, we performed force–extension and twist experiments. DNA molecules were tethered to the glass surface via an anti-DIG and to the paramagnetic bead using streptavadin. Per field of view (500 µm by 500 µm) we obtained an average of 30 tethers (Figure 4.4A) with a total of 120 tethers. In experiments done on a 172x12 DNA molecule using the same protocol (data from 172x12 not shown in this thesis) with a concentration of 1 ng mL^{-1}typically approximately 200 tethers where detected per field of view, indicating that the concentration of the G4 construct is 10 times less and in the order of 100 pg mL^{-1}.

To check the length of the tethered DNA, we obtained force–extension curves and fitted them with the wormlike chain model (Equation 3.8, Figure 4.4 B). The fit parameters were the contour length L, z_0 and dz/dt and the fixed parameters the persistence length $P = 45$ nm and stretch modulus $S = 900$ pN. Out of the 120, tethers 32 could be fitted and the distribution of the fitted contour length had a peak at 3450 bp (Figure 4.4C). The other 88 tethers where discarded because either the bead was ruptured from the DNA molecule, the tether was stuck to the surface or the bead went out of focus.

After force extension, we subjected the same 120 tethers to twist experiments to determine whether thy were torsionally constrained. The extension was expected to have an asymmetric response to over- and undertwisting [57], but this was not detected (Figure 4.4D). The trajectory of the motor is depicted in red. During the experiments performed at different forces a linear drift of the sample was detected of 1.5 nms^{-1} (blue line). Comparing the extension of the negative and positive twist, no considerable change was detected indicating that the G4 construct was not torsionally constrained but nicked. This applied to all tethers.

The ligation of the handles and the spacer is the most likely the origin of the nicks. In this process the ends were probably not correctly attached, and this a nick is created. In the ligation process the ligation product was filtered. The concentration of the end product was rather low, supporting our conclusion that it is due to the ligation process. Because the G4 construct is not torsionally constrained it is not possible to build up torque and thus to induce melting of base pairs.

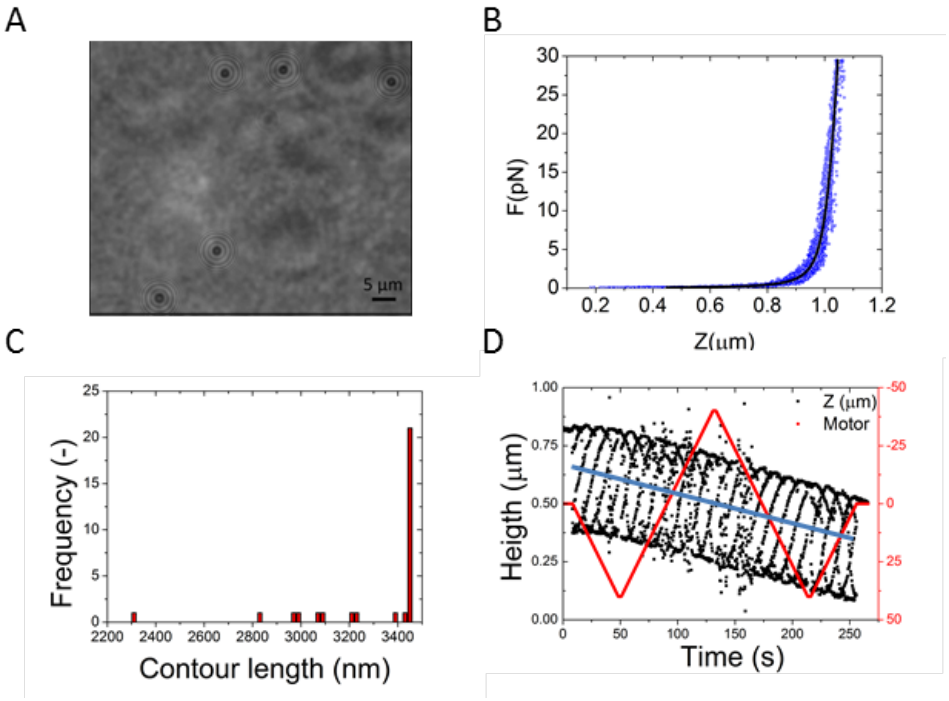

Figure 4.4: Force extension on G4 construct yields proper length and an estimated concentration of G4 lower than 1 $\mathrm{ng\,mL^{-1}}$***, twist experiments did not show torsional constraints, indicating nicks.*** *A) Typical field of view from the Magnetic Tweezers, the black dots are magnetic beads (diameter of 2.8* µm*.B) Typical force-extension curve (blue dots) with a WLC fit (black line). The fixed parameters are: S = 900 pN, P = 45 nm and fit parameters are L,z_0 and dz/dt. Compared to other Magnetic Tweezers experiments the bead yield is 10 times less indicating an G4 concentration lower than 1* $\mathrm{ng\,mL^{-1}}$*. C) Distribution of the lengths of a WLC fitted to the observed data. The peak distribution was at 3450 bp which corresponds to the length of the G4 construct. D) Typical time trace of the height of a tether during negative and positive rotation. There was no significant change in length, indicating that the DNA molecule was not torsional constrained. The blue line has a slope of 1.5* $\mathrm{nms^{-1}}$ *which is the drift of the flow cell during an experiment. In red the trajectory of the twisting motor was shown.*

4.3 Simulations

Because the G4 construct did not qualify our requirements, we could not perform the intended twist and force–extension experiments. For this reason we have performed simulations as described in Chapter 2. In the simulations we have computed the probability for each base pair to be in the melted state, using the original DNA sequence. From the melting proba-

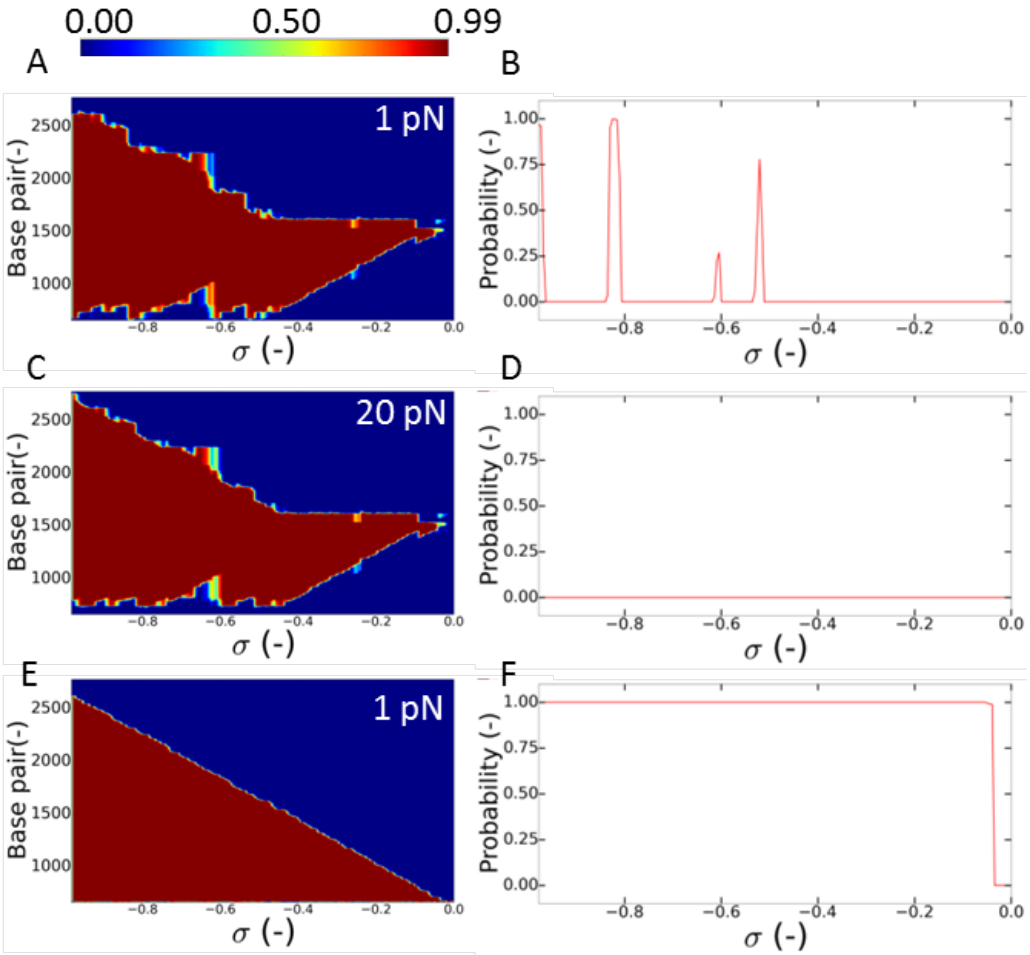

Figure 4.5: The probability for the base pairs that form the G4 sequence (bp 681-703) to melt is nearly zero without an initial mismatch of base pairs near the G4. *The simulations show that the base pairs containing the G4 complex (between base pair index 681 and 703) are not likely to be melted (A, C, E). If the pulling force is increased from 1 pN (B) to 20 pN (D) the probability is zero to form. When a mismatch is introduced at base pair 671,672 and 673 the probability for the base pairs to be in the melted state increase. Due to the nearest neighbor interaction energy the number of base pairs that melt during the twisting fluctuates severely.*

bility we could compute $< z >$ and $< E >$ as well as the probability that a G4 structure can form. The results of the probability to be in a melted state were presented as a phase diagram for the complete construct, but excluding the handles since they were fixed and did not contribute to the dsDNA molecule. We also included a profile of the G4 region to be in the melted state, using Equation 3.6 (Figure 4.5).
In the first simulation a force of 1 pN was applied and a melting bubble was formed in the center of the construct (an **AT** rich region). Figure 4.5A shows that for increasing negative twist the bubble size increased. Due to the variations in melting energy between the base pairs (Table 3.2) the increase of the bubble size was not a continuous process, indicated by the non–smooth profile. When the twist density was nearly minus one the melting bubble contained nearly all the base pairs in the DNA tether. These twist density values are extreme and the model does not include the effects that could occur when the DNA is in the total melted state. For less negative σ, the melting bubble only contained the G4 region for a small range of σ. The melting bubble was not very stable in the G4 region according to the simulations, indicating that the G4 site was not likely to melt nor that the G4 structure could have formed.
In Figure 4.5B the probability for the G4 sequence to be in the melted state was showed for a force of 1 pN, supporting our claim that the G4 could hardly form. For high twist density (σ = -0.5, -0.6 and -0.8) the probability for the G4 sequence to melt was not zero. This indicates that for these specific twist densities the G4 could fold. However when σ further decreases the base pairs will return in the twisted state. For larger forces the probability for the G4 sequence to be in the melted state was zero regardless of the twist density (Figure 4.5 C & D). We concluded two things from Figure 4.5A-D. Firstly, that only for low forces the probability for the G4 sequence to be in the melted state was not zero, but that we would have to apply a relatively large negative σ. Secondly, that the window of experiments was limited to large twist densities and low force regimes.
In an effort to target melting bubble formation to the G4 site, a 3-bp mismatch was introduced in between the G4 and the lower handle at base pairs 671,672 and 673 [67]. Due to the fact that the handle is fixed to the surface, it could not extend to the bottom. When the melting bubble extends it would first encounter the G4 site. Therefore, we expected that the probability for the G4 sequence to be in the melted state would increase, even for lower σ and higher forces. For a pulling force of 1 pN the probability for the G4 sequence to melt was nearly one at σ = 0.05 (Figure 4.5 E &F). Higher forces showed a similar pattern (not shown here), indicating that the initial mismatch strongly enhanced the melting of the G4 site.

When the G4 was in the melted state the probability to be *folded* could be computed using Equation 3.11. Because the results for the G4 without mismatch indicated that it was not likely to melt the G4 region, subsequent simulations where done with the G4 construct with a 3–bp mismatch. Simulation for forces ranging from 1 to 20 pN indicated that the folding probability decreases as the force increases(Figure 4.6A). The turning point was approximately 8.5 pN, where ε_{G4} was equal to zero indicating that ε_{int} and the work W were in balance resulting in a the probability of 50 per cent.
We expected that the extension $< z >$ decreases abruptly when a G4 structure forms because of the size difference. In Figure 4.6B $< z >$ is shown using Equation 3.14. For low forces we detected a decrease of extension when the G4 folds and the size of the extension was in the order of a few nanometers. This should be detectable using MT. The extension decreases as the change in linking number density increases. This was due to the base pair distribution, that shifted from twisted to melted. The extension of a melted base pair was smaller for these forces compared to the twisted state. A second kink was detected for 1 pN at $\sigma = 0.065$. Here, an **AT** rich site was melted. Because this required less melting energy compared to the other di–nucleotide pairs, multiple base pairs changed state and the extension decreases. As the force increases the extension of the base pairs got more similar and this kink was no longer detected.
As described in Chapter two, we can also compute the FRET efficiency $< E >$. The donar and acceptor pair are spaced by 25 base pairs and the extension of these base pairs was computed using Equation 3.17 (Figure 4.6C). When the G4 structure (22 bp) folded, the FRET efficiency increased for 1 and 5 pN. In the pre–folded region the FRET efficiency for 1 pN was larger than for 5 pN. This was due to the extension of the base pairs, for 1 pN it is smaller and the FRET pair was closer together. It is not likely that we can detect this value in a FRET/MT setup due to the system noise which has a value approximately 0.3. However, when the G4 structure forms we can detected the FRET efficiency and compare it to our simulations.
The simulations showed that the initial DNA molecule was not likely to form a G4 structure, even if it was torsionally constrained. The change in linking number density would be rather large and other effects, not included in the two–state model, could occur. The 3–bp mismatch located near the G4 site enhanced the forming strongly and would made it possible to examine the formation of a G4 in dsDNA. The extension and FRET efficiency were in a range were they could be detected in a combined FRET/MT.

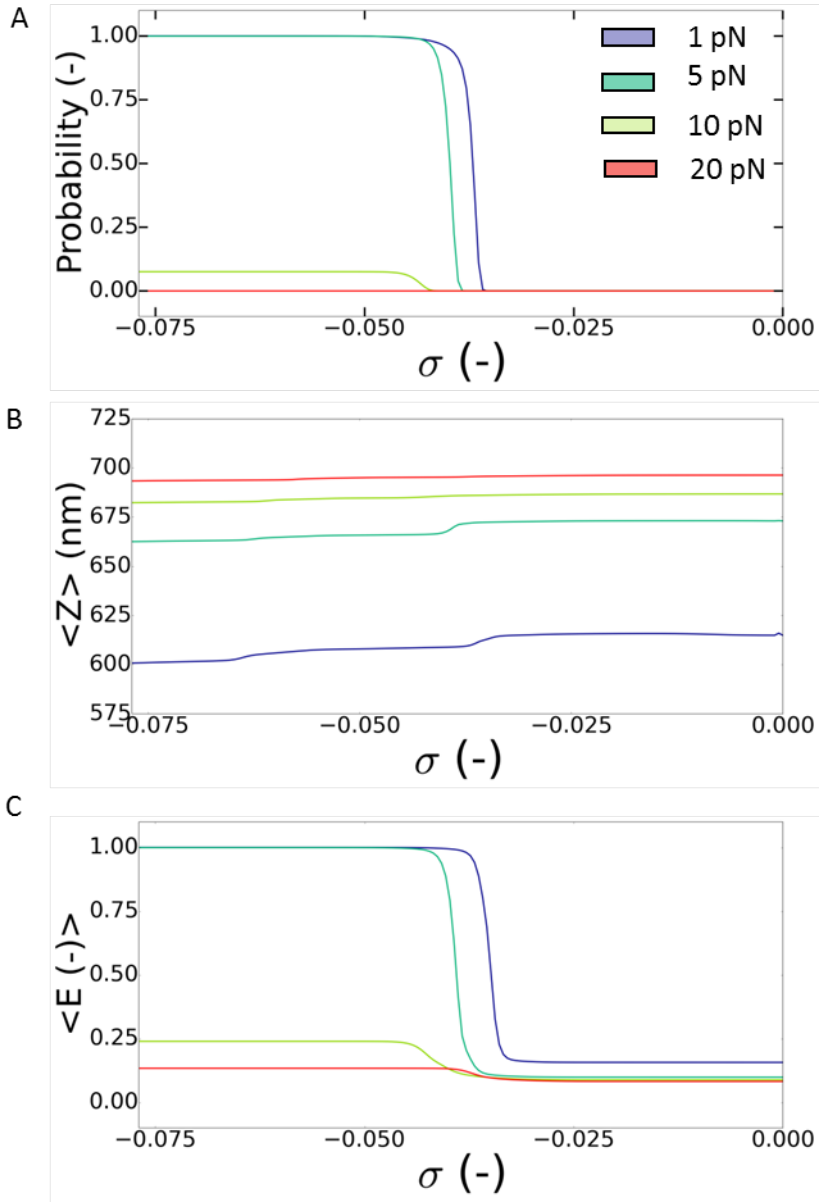

Figure 4.6: The probability for the G4 to fold decreases as the force increases. When the G4 folds the change in FRET efficiency and in length can be detected. *A) At higher forces the G4 is less favorable to form, even if the base pairs are melted. B) The length slightly decreases as more base pairs are melted. When a G4 forms there is a jump of order 10 nm, which can be detected in MT. C) After the folding of the G4 the efficiency is maximal.*

Chapter 5

Discussion

5.1 DNA synthesis

In this study we aimed to study the formation of a single G4 structure in dsDNA using FRET and MT. Despite the successful annealing, we were unable to form a ligation product that met our requirements. We suspect that there are two reasons why this has happened. First, the low success rate of the ligation was caused by the overhangs which were not unique: both handles, the spacer and the annealing product contained a BsaI and BseYI overhang. In the ligation process all four DNA pieces were introduced simultaneously likely given rise to many side products. Because the overhangs were not unique, the side products also included the conformations where the spacer was ligated to itself. The absence of the fluorescence signal in the end product can be explained by the absence of the G4 annealing product, only the two handles and the spacer were ligated. Second, we think it is possible that intermolecular G4s have formed in the ligation buffer (LB). Here, not one single strand formed an intramolecular G4 structure but two single strands were required [18]. LB contained 10 mM Mg^{2+}, and studies have shown that this ion strongly enhances the folding of G4 [48]. When an intermolecular G4 was formed, the overhangs of the annealing product were blocked and the multi–DIG handle and spacer were not able to ligate to the annealing product. This would also explain the absence of fluorescence signal in the end product because G4 sites could not ligate.

5.2 Force–extension and twist experiments

Despite the poor yield, we were still able to perform force spectroscopy on the ligation product. From force-extension (FE) curves the contour length of the end product was determined to be between 3400 and 3500 bp. We expected a contour length of 3444 bp and concluded that the DNA tethers were of the correct length. Results from gel supported this conclusion. In the FE curves a maximal force of 30 pN was applied and, between the pulling and relaxing curves, no hysteresis was detected. This indicates that the base pairs in the G4 site do not melt at these forces. From other studies we know that DNA overstretches at a force of 65 pN; the state of the base pairs goes from twisted to melted and the extension increases [68, 69]. When this happens, the G4 could form and hysterese would occur. However, at these high forces it is unlikely for the G4 to fold. We concluded that only pulling was not sufficient and that we need to twist the DNA tether to melt the G4 site for lower forces.
From the FE curves we also concluded that the concentration of the end product was rather low. For MT this was not a problem since the field of view (FoV) is 500 by 500 µm and an average of 200 tethers per FoV were detected in other studies on the same setup [70]. However, the FoV of the FRET/MT setup is 75 by 75 µm and the low concentration can cause an experimental difficulty. It will take more time to find suitable tethers and DNA tethers are known to deteriorate over time. Therefore, a higher concentration in the order of 1 – 10 ng mL^{-1}would be desirable.
The conducted twist experiments showed that the DNA tether was not torsionally constrained probably due to nicks in the tether. The nicks were probably formed during the ligation process: the annealing product overhang may not have been accessible and only one strand ligated. This prevented us to use this stock of DNA for twist experiments. Also, a linear drift of 1.5 nm s^{-1} was detected in the experiment, probably originating from the movement of the XY–stage. In the time scales of our experiments this could result in a total drift of a few 100 nm, but because it is linear drift the obtained force–extension data can be corrected for it afterwards.

From the force–extension and twist experiments with a dsDNA tether we expected to be able to detect the folding of a G4. This would be a great addition to the studies on ssDNA in MT [51] and on dsDNA in OT [48]. Unfortunately, due to the problems during the ligation reaction, we were not able to create a torsionally constrained FRET–labeled dsDNA construct containing one G4 site. Other studies have described the technical challenges involved with G4 experiments [52, 71]. The two main difficulties

were the salt conditions during the ligation process, making the G4 fold and the location of the fluorophores. On the other hand, ligation buffer requires Mg^{2+}, which induces G4 folding. G4 folding, intra– or inter–molecular can impede ligation. The introduction of fluorophores decrease the stability of the G4 and makes it more difficult to fold. Based on these studies and our own experience, we make two suggestions to improve the DNA synthesis. First, make the overhangs of the four different DNA pieces unique. They should not be able to ligate on to themselves, only on the targeted overhang. Second, adjust the ligation buffer so that it will no longer contain Mg^{2+} and the forming of intermolecular G4s is prevented. A particular study showed that other ions metals like zinc, manganese, zinc, cadmium or calcium also suffice [72].

5.3 Simulations

To provide more insight into the stability of G4s, simulations were performed based on the two–state model. Based on these simulations, we concluded that the G4 site in a DNA construct with a random flanking sequence is unlikely to melt, because there are always **AT**–rich regions that would more easily melt. With a 3–bp mismatch it is possible to direct the bubble to the G4 region. When the melting bubble included the G4 site, the probability for the G4 to fold decreased as the force increased. We can not measure the state of the base pairs directly, but we can detect the changes in extension and FRET signal. We concluded that when the G4 folds, this induced a decrease in the extension in the order of 10 nanometer and an increase in the FRET efficiency. Both effects can be measured in a FRET/MT setup.
The simulations were based on the two–state model, where the model depends on the twist, force and sequence of the DNA. Due to sequence dependence, we were able to calculate the probability for each state and the behavior of the DNA molecule in much more detail than the previously published three-state model that does not take sequence into account. Because the two–state model is based on Boltzmann statistics, there is no history in the simulation: each σ is calculated as if the DNA would reach that point without passing previous intermediate states. In practice, we expect to twist slow enough for the DNA tether to be in equilibrium all the time. This means that when the G4 site is in the melting bubble and the G4 folds, further twisting will not remove the folded G4.
Using MT and FRET we could detect the folding in three ways. First, the

extension decreases, as we observed in the expected extension $< z >$. Second, the extension will not change when σ becomes more negative. The opposing strand (not folded) will absorb torque and thus the distribution of the states will not change. And finally, we would detect an increase of the FRET signal as the two fluorescent labels are brought closer together.
The model predicts when the G4 site melts and if the G4 structure would form. We concluded that for high forces, in the range of 20 pN the folding was unlikely. This was based on Equation 3.11, where the most important parameter ε_{int} is. The value of ε_{int} was an average of multiple studies [63–66], and should be determined with high precision as it determines the folding probability and the likeliness of the G4 site to be in the melted bubble. For instance, it could depend on the salt conditions. This was not included in the model and this is, besides the melting energy, the only variable depending on external conditions. Or, it could be possible to determine the likeliness of different folding configurations based on ε_{int}. To determine this value better one needs to closely examine the exact position of the bases in a G4 configuration and their interaction with the central ion and the interaction between the plateaus and guanine bases.
The simulations on the DNA construct with a 3–bp mismatch showed that the mismatch would greatly enhance the melting probability of the G4 region. Incorporating this in our DNA construct will make it much more likely to induce and detect the folding of G4. When a twisting experiment is performed for a range of force clamps, the turning point from the folded to the unfolded state can be determined. Because at this point $F\Delta z = \varepsilon_{int}$, the value for ε_{int} can be determined more precise experimentally.

Based on the outcomes of the simulation, we concluded that the two–state model can describe the effects of a G4 in dsDNA. However, the model can still be improved. The force–dependent extension of the base pairs is not taken into account when the melt probabilities are calculated. Also, we did not include a torque threshold: others studies have shown that there is a threshold to overcome prior to the melting of base pairs [57, 73, 74]. In the three–state model, a threshold of 3 $k_BT$is used. Below a torque of 3 $k_BT$the base pairs would not melt and no change in state is induced. The expected sigma at which the G4 site is completely melted will thus be larger than based on our simulations.
Also, the plectonemic state should be included when the model is expanded for positive twist. When positive twist is included, we can detect the hysterese as we twist with and without a folded G4. We expect that when the G4 is folded, the decrease in extension will occur at a lower σ and that the extension at zero twist is lower compared to a DNA molecule

of equal contour length without a folded G4. The two–state models gives a good description of a DNA containing a G4 site, but it can be regarded in a much broader vision. We can describe all sort of second order structures, based on the state of the individual base pairs. To do so, only the change in contour length and interaction energy (here, ε_{G4}) have to be adapted. This is all possible due to the sequence dependency of the model, which makes it possible to study each base pair individually.

Chapter 6

Conclusion

Our goal was to use a combined FRET/MT setup to examine the G4 folding in dsDNA. Because of technical problems we were not able to use the combined setup and performed our experiments on MT and a gel imager. The results showed that the DNA construct we originally had in mind was too difficult to construct. We made three suggestions for improvements to the design and ligation process: unique overhangs, remove Mg^{2+} from the ligation buffer and an initial 3–bp mismatch between the G4 site and the multi–DIG handle. We also made some suggestions to improve and expand the two–state model. Although we were unsuccessful in the experiments, simulations based on the two–state model showed that the new design will meet our requirements.
We illustrated, using simulations, that MT is an excellent technique to deform DNA with twist while simultaneously detect the resulting response. When MT is combined with FRET, in theory we would be able to study the formation of G4 in much more detail. The simulations also showed that the folding of a G4 is strongly dependent on force, twist and the interaction energy. For a force of 8.5 pN the probability to fold is 50% when the G4 site is in the melting bubble. We also concluded from the simulations that the FRET efficiency increases and the extension decreases when a G4 folds and that this can be measured in a FRET/MT setup.
We were unable to perform twist experiments due to the DNA tether not being torsional constrained. Fortunately, the developed two–state model showed some promising outcomes. The new design for the DNA substrate can be tested on a combined FRET/MT setup and these results can be compared to the outcomes of the simulations. Our work helps to understand the effect of force and torque on the formation of a G4 in dsDNA in addition to previous work [47, 51]. With the new design for the DNA molecule

the forming energy ε_{G4} can be determined, opening up the possibility to study different folding conformations. Also, by adding fluorescence labeled proteins during the experiment, it can be determined what the effect of the different folding conformations is on the functionality of the proteins. And this should not be limited to one G4 site; a DNA molecule with multiple G4 sites could have a different response compared to a single G4 site.
A deeper understanding of the dynamics and stability of G4 will help us to unravel the role it plays in gene regulation sites. This knowledge can be used, for example, for a better understanding of cancer development, and eventually a cure. Therefore, G4 plays a vital role in unraveling yet another piece of our DNA and, ultimately, our life.

References

[1] F. Miescher and O. Schmiedeberg, *Physiologisch-chemische Untersuchungen über die Lachsmilch*, Naunyn-Schmiedeberg's Archives of Pharmacology **37**, 100 (1869).

[2] J. D. Watson et al., *Molecular structure of nucleic acids*, Nature **171**, 737 (1953).

[3] A. Baumketner, A. Jewett, and J. Shea, *Effects of Confinement in Chaperonin Assisted Protein Folding: Rate Enhancement by Decreasing the Roughness of the Folding Energy Landscape*, Journal of Molecular Biology **332**, 701 (2003).

[4] S. L. Forman, J. C. Fettinger, S. Pieraccini, G. Gottarelli, and J. T. Davis, *Toward artificial ion channels, a lipophilic G-quadruplex*, Journal of the American Chemical Society **122**, 4060 (2000).

[5] L. Petraccone, C. Spink, J. O. Trent, N. C. Garbett, C. S. Mekmaysy, C. Giancola, and J. B. Chaires, *Structure and stability of higher-order human telomeric quadruplexes*, Journal of the American Chemical Society **133**, 20951 (2011).

[6] S. Rankin, A. P. Reszka, J. Huppert, M. Zloh, G. N. Parkinson, A. K. Todd, S. Ladame, S. Balasubramanian, and S. Neidle, *Putative DNA quadruplex formation within the human c-kit oncogene*, Journal of the American Chemical Society **127**, 10584 (2005).

[7] Y. Wang and D. J. Patel, *Guanine residues in d (T2AG3) and d (T2G4) form parallel-stranded potassium cation stabilized G-quadruplexes with anti glycosidic torsion angles in solution.*, Biochemistry **31**, 8112 (1992).

[8] J. Dai, C. Punchihewa, A. Ambrus, D. Chen, R. A. Jones, and D. Yang, *Structure of the intramolecular human telomeric G-quadruplex in potassium solution: a novel adenine triple formation*, Nucleic acids research **35**, 2440 (2007).

[9] P. Akhshi, N. J. Mosey, and G. Wu, *Free-Energy Landscapes of Ion Movement through a G-Quadruplex DNA Channel*, Angewandte Chemie International Edition **51**, 2850 (2012).

[10] A. Guédin, J. Gros, P. Alberti, and J.-L. Mergny, *How long is too long? Effects of loop size on G-quadruplex stability*, Nucleic Acids Research **38**, 7858 (2010).

[11] S. Burge, G. N. Parkinson, P. Hazel, A. K. Todd, and S. Neidle, *Quadruplex DNA: sequence, topology and structure*, Nucleic acids research **34**, 5402 (2006).

[12] P. Hazel, J. Huppert, S. Balasubramanian, and S. Neidle, *Loop-Length-Dependent Folding of G-Quadruplexes*, Journal of the American Chemical Society **126**, 16405 (2004).

[13] P. Hazel, J. Huppert, S. Balasubramanian, and S. Neidle, *Loop-length-dependent folding of G-quadruplexes*, Journal of the American Chemical Society **126**, 16405 (2004).

[14] I. Bang, *Untersuchungen über die Guanylsäure*, Biochemische Zeitschrift **26**, 293 (1910).

[15] M. Gellert, M. N. Lipsett, and D. R. Davies, *Helix formation by guanylic acid*, Proceedings of the National Academy of Sciences **48**, 2013 (1962).

[16] N. H. Campbell, G. N. Parkinson, A. P. Reszka, and S. Neidle, *Structural basis of DNA quadruplex recognition by an acridine drug*, Journal of the American Chemical Society **130**, 6722 (2008).

[17] A. M. Burger, F. Dai, C. M. Schultes, A. P. Reszka, M. J. Moore, J. A. Double, and S. Neidle, *The G-quadruplex-interactive molecule BRACO-19 inhibits tumor growth, consistent with telomere targeting and interference with telomerase function*, Cancer Research **65**, 1489 (2005).

[18] H. J. Lipps and D. Rhodes, *G-quadruplex structures: in vivo evidence and function*, Trends in cell biology **19**, 414 (2009).

[19] K. Paeschke, T. Simonsson, J. Postberg, D. Rhodes, and H. J. Lipps, *Telomere end-binding proteins control the formation of G-quadruplex DNA structures in vivo*, Nature structural & molecular biology **12**, 847 (2005).

[20] H. Tahara, K. Shin-Ya, H. Seimiya, H. Yamada, T. Tsuruo, and T. Ide, *G-Quadruplex stabilization by telomestatin induces TRF2 protein dissociation from telomeres and anaphase bridge formation accompanied by loss of the 3 telomeric overhang in cancer cells.*, Oncogene **25** (2006).

[21] S. Balasubramanian, L. H. Hurley, and S. Neidle, *Targeting G-quadruplexes in gene promoters: a novel anticancer strategy?*, Nature reviews. Drug discovery **10**, 261 (2011).

[22] J. L. Huppert and S. Balasubramanian, *G-quadruplexes in promoters throughout the human genome*, Nucleic acids research **35**, 406 (2006).

[23] G. Biffi, D. Tannahill, J. McCafferty, and S. Balasubramanian, *Quantitative visualization of DNA G-quadruplex structures in human cells*, Nature Chemistry **182**, 16405 (2013).

[24] S. Haider, G. N. Parkinson, and S. Neidle, *Crystal Structure of the Potassium Form of an Oxytricha nova G-quadruplex*, Journal of Molecular Biology **320**, 189 (2002).

[25] D. Monchaud and M.-P. Teulade-Fichou, *A hitchhiker's guide to G-quadruplex ligands*, Organic & biomolecular chemistry **6**, 627 (2008).

[26] M. L. Bochman, K. Paeschke, and V. A. Zakian, *DNA secondary structures: stability and function of G-quadruplex structures*, Nature Reviews Genetics **13**, 770 (2012).

[27] A. De Cian, G. Cristofari, P. Reichenbach, E. De Lemos, D. Monchaud, M.-P. Teulade-Fichou, K. Shin-Ya, L. Lacroix, J. Lingner, and J.-L. Mergny, *Reevaluation of telomerase inhibition by quadruplex ligands and their mechanisms of action*, Proceedings of the National Academy of Sciences **104**, 17347 (2007).

[28] R. Hänsel-Hertsch, M. Di Antonio, and S. Balasubramanian, *DNA G-quadruplexes in the human genome: detection, functions and therapeutic potential*, Nature Reviews Molecular Cell Biology **18**, 279 (2017).

[29] T. A. Brooks, S. Kendrick, and L. Hurley, *Making sense of G-quadruplex and i-motif functions in oncogene promoters*, The FEBS journal **277**, 3459 (2010).

[30] M. Sette, P. D'Addabbo, G. Kelly, A. Cicconi, E. Micheli, S. Cacchione, A. Poma, C. Gargioli, V. Giambra, and D. Frezza, *Evidence for a quadruplex structure in the polymorphic hs1. 2 enhancer of the immunoglobulin heavy chain 3-regulatory regions and its conservation in mammals*, Biopolymers **105**, 768 (2016).

[31] R. Hänsel-Hertsch et al., *G-quadruplex structures mark human regulatory chromatin*, Nature genetics **48**, 1267 (2016).

[32] L. Oganesian and T. M. Bryan, *Physiological relevance of telomeric G-quadruplex formation: a potential drug target*, Bioessays **29**, 155 (2007).

[33] H. Hwang, *Structural dynamics of telomeric overhang accessibility*, PhD thesis, University of Illinois at Urbana-Champaign, 2014.

[34] B. Juskowiak, *Nucleic acid-based fluorescent probes and their analytical potential*, Analytical and bioanalytical chemistry **399**, 3157 (2011).

[35] A. Siddiqui-Jain, C. L. Grand, D. J. Bearss, and L. H. Hurley, *Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription*, Proceedings of the National Academy of Sciences **99**, 11593 (2002).

[36] T.-M. Ou et al., *Stabilization of G-quadruplex DNA and down-regulation of oncogene c-myc by quindoline derivatives*, Journal of medicinal chemistry **50**, 1465 (2007).

[37] A. Ambrus, D. Chen, J. Dai, R. A. Jones, and D. Yang, *Solution structure of the biologically relevant G-quadruplex element in the human c-MYC promoter. Implications for G-quadruplex stabilization*, Biochemistry **44**, 2048 (2005).

[38] A. Rangan, O. Y. Fedoroff, and L. H. Hurley, *Induction of duplex to G-quadruplex transition in the c-myc promoter region by a small molecule*, Journal of Biological Chemistry **276**, 4640 (2001).

[39] A. T. Phan, Y. S. Modi, and D. J. Patel, *Propeller-type parallel-stranded G-quadruplexes in the human c-myc promoter*, Journal of the American Chemical Society **126**, 8710 (2004).

[40] H. You, J. Wu, F. Shao, and J. Yan, *Stability and kinetics of c-MYC promoter G-quadruplexes studied by single-molecule manipulation*, Journal of the American Chemical Society **137**, 2424 (2015).

[41] K. N. Luu, A. T. Phan, V. Kuryavyi, L. Lacroix, and D. J. Patel, *Structure of the human telomere in K+ solution: an intramolecular (3+ 1) G-quadruplex scaffold*, Journal of the American Chemical Society **128**, 9963 (2006).

[42] J. Dai, T. S. Dexheimer, D. Chen, M. Carver, A. Ambrus, R. A. Jones, and D. Yang, *An intramolecular G-quadruplex structure with mixed parallel/antiparallel G-strands formed in the human BCL-2 promoter region in solution*, Journal of the American Chemical Society **128**, 1096 (2006).

[43] A. T. Phan, V. Kuryavyi, S. Burge, S. Neidle, and D. J. Patel, *Structure of an unprecedented G-quadruplex scaffold in the human c-kit promoter*, Journal of the American Chemical Society **129**, 4386 (2007).

[44] A. T. Phan, V. Kuryavyi, H. Y. Gaw, and D. J. Patel, *Small-molecule interaction with a five-guanine-tract G-quadruplex structure from the human MYC promoter*, Nature chemical biology **1**, 167 (2005).

[45] Y. Sannohe, K. Sato, A. Matsugami, K.-i. Shinohara, T. Mashimo, M. Katahira, and H. Sugiyama, *The orientation of the ends of G-quadruplex structures investigated using end-extended oligonucleotides*, Bioorganic & medicinal chemistry **17**, 1870 (2009).

[46] W. Li, X.-M. Hou, P.-Y. Wang, X.-G. Xi, and M. Li, *Direct measurement of sequential folding pathway and energy landscape of human telomeric G-quadruplex structures*, Journal of the American Chemical Society **135**, 6423 (2013).

[47] S. Selvam, D. Koirala, Z. Yu, and H. Mao, *Quantification of topological coupling between DNA superhelicity and G-quadruplex formation*, Journal of the American Chemical Society **136**, 13967 (2014).

[48] S. Selvam, S. Mandal, and H. Mao, *Quantification of Chemical and Mechanical Effects on the Formation of the G-Quadruplex and i-Motif in Duplex DNA*, Biochemistry **56**, 4616 (2017).

[49] S. L. Gotta, O. Miller, and S. L. French, *rRNA transcription rate in Escherichia coli.*, Journal of bacteriology **173**, 6647 (1991).

[50] D. Rhodes and H. J. Lipps, *G-quadruplexes and their regulatory roles in biology*, Nucleic acids research **43**, 8627 (2015).

[51] H. You, S. Lattmann, D. Rhodes, and J. Yan, *RHAU helicase stabilizes G4 in its nucleotide-free state and destabilizes G4 upon ATP hydrolysis*, Nucleic acids research **45**, 206 (2017).

[52] P. Maleki, J. B. Budhathoki, W. A. Roy, and H. Balci, *A practical guide to studying G-quadruplex structures using single-molecule FRET*, Molecular Genetics and Genomics **3**, 483 (2017).

[53] J. F. Marko and E. D. Siggia, *Stretching dna*, Macromolecules **28**, 8759 (1995).

[54] C. Allain, D. Monchaud, and M.-P. Teulade-Fichou, *FRET templated by G-quadruplex DNA: a specific ternary interaction using an original pair of donor/acceptor partners*, Journal of the American Chemical Society **128**, 11890 (2006).

[55] X. Long, J. W. Parks, C. R. Bagshaw, and M. D. Stone, *Mechanical unfolding of human telomere G-quadruplex DNA probed by integrated fluorescence and magnetic tweezers spectroscopy*, Nucleic Acids Research **41**, 2746 (2013).

[56] I. Vilfan, J. Lipfert, D. Koster, S. Lemay, and N. Dekker, *Magnetic tweezers for single-molecule experiments*, in *Handbook of single-molecule biophysics*, pages 371–395, Springer, 2009.

[57] H. Meng et al., *Structural changes in single chromatin fibers induced by tension and torsion*, Leiden Institute of Physics (LION), Faculty of Science, Leiden University, 2014.

[58] B. Tinland, A. Pluen, J. Sturm, and G. Weill, *Persistence length of single-stranded DNA*, Macromolecules **30**, 5763 (1997).

[59] S. B. Smith and A. J. Bendich, *Electrophoretic charge density and persistence length of DNA as measured by fluorescence microscopy*, Biopolymers **29**, 1167 (1990).

[60] J. Bednar, P. Furrer, V. Katritch, A. Stasiak, J. Dubochet, and A. Stasiak, *Determination of DNA persistence length by cryo-electron microscopy. Separation of the static and dynamic contributions to the apparent persistence length of DNA*, Journal of molecular biology **254**, 579 (1995).

[61] J. Noort van et al., *Syllabus for Advance Biophysics*, Leiden University.

[62] P. Shrestha, S. Jonchhe, T. Emura, K. Hidaka, M. Endo, H. Sugiyama, and H. Mao, *Confined space facilitates G-quadruplex formation*, Nature Nanotechnology **12**, 582 (2017).

[63] A. N. Lane, J. B. Chaires, R. D. Gray, and J. O. Trent, *Stability and kinetics of G-quadruplex structures*, Nucleic Acids Research **36**, 5482 (2008).

[64] A. Bugaut and P. Alberti, *Understanding the stability of DNA G-quadruplex units in long human telomeric strands*, Biochimie **113**, 125 (2015).

[65] A. Ramamoorthy, *NMR Spectroscopy of Biological Solids*, Practical Spectroscopy, CRC Press, 2005.

[66] J.-A. Mondragon-Sanchez, E. Mendieta-Fernandez, R. Garduno-Juarez, and G. Sanchez-Gonzalez, *Metadynamics Study of the Free Energy Surface of a G-Quadruplex DNA Structure"*, Biophysical Journal **98** (2009).

[67] F. C. Oberstrass, L. E. Fernandes, and Z. Bryant, *Torque measurements reveal sequence-specific cooperative transitions in supercoiled DNA*, Proceedings of the National Academy of Sciences **109**, 6106 (2012).

[68] S. B. Smith, Y. Cui, and C. Bustamante, *Overstretching B-DNA: the elastic response of individual double-stranded and single-stranded DNA molecules*, Science , 795 (1996).

[69] C. Bustamante, Z. Bryant, and S. B. Smith, *Ten years of tension: single-molecule DNA mechanics*, Nature **421**, 423 (2003).

[70] A. Kaczmarczyk, A. Allahverdi, T. B. Brouwer, L. Nordenskiöld, N. H. Dekker, and J. van Noort, *Single-molecule force spectroscopy on histone H4 tail-cross-linked chromatin reveals fiber folding*, Journal of Biological Chemistry **292**, 17506 (2017).

[71] L. Ying, J. J. Green, H. Li, D. Klenerman, and S. Balasubramanian, *Studies on the structure and dynamics of the human telomeric G quadruplex by single-molecule fluorescence resonance energy transfer*, Proceedings of the National Academy of Sciences **100**, 14629 (2003).

[72] M. R. Taylor, *The Role of Divalent Metal Ions in Enzymatic DNA Ligation*, (2014).

[73] J. F. Marko, *Torque and dynamics of linking number relaxation in stretched supercoiled DNA*, Phys. Rev. E **76**, 021926 (2007).

[74] P. Gross, N. Laurens, L. B. Oddershede, U. Bockelmann, E. J. Peterman, and G. J. Wuite, *Quantifying how DNA stretches, melts and changes twist under tension*, Nature Physics **7**, 731 (2011).