Utrecht University

Debye Institute

Biophysics

Master Thesis

On the mechanics of novel G-quadruplex DNA structures

Author:
Querijn de Graaf, BSc

Supervisors:
Dr. Gerhard Blab
Dr. Arnout Imhof

June 30, 2016

Abstract

Single particle and molecule manipulation has lead to numerous insights in the areas of physics, biology, chemistry and medicine. Currently the techniques for doing measurements on single particles are expensive and laborious. In this thesis we looked at an alternative, first devised by Halvorsen and Wong in 2010, as a cheaper faster way of doing multiple single particle measurements in parallel using centrifugal force[1]. For this purpose a device was built called a Centrifugal Force Microscope which we used to study the folding and unfolding of a non-standard DNA type called a G-quadruplex. It was known from previous research that the set-up showed various problems with stability. These problems turned out to be much larger than expected, delaying the acquisition of data. However I have shown that centrifugal force microscopy is indeed still a viable alternative to other single particle manipulation techniques, provided improvements are made on its stability. We were able to reach sub pixel accuracies in the 10 nm range, finding a workaround for present artifacts. I was able to fully characterise the set-up and make improvements on the sample creation protocol. Unfortunately not enough useful data on actual G-quadruplexes could be obtained to make hard statements about their response to forces in the femtonewton range.

Contents

1 Introduction

Since the discovery of its double helix structure by Watson and Crick with valuable contributions from Rosalind Franklin, Deoxyribonucleic acid (DNA) has proved of prime interest to scientists in fields ranging from molecular biology to biophysics [2][3]. This is due to the fact that it holds the biological information of the organism to whom the DNA belongs and due to its strong connection with various diseases, including cancer. Through the years, in particular in the eighties, a plethora of other DNA structures were hypothesised and found such as Z-DNA, cruciform DNA and Okazaki fragments [4]. Respectively these are double helix DNA with a different orientation, DNA which has formed a cruciform out of four double strands, and short newly synthesized DNA fragments. Many more of these alternative structures exist and of these structures the one which has garnered the most attention is the G-quadruplex. This is due to two main reasons: its remarkable stability and its potential for the treatment of human cancers and targeted drug-delivery [5][6][7]. Despite the interest, little is known about the slow folding and unfolding kinetics or the many small variations in structure of the quadruplex.

Although biologists and biochemists have acquired a lot of information on G-quadruplexes, they have so far been optically studied mostly through two methods. The first is fluorescence imaging, which is limited in resolution, and unable to tell us anything about the dynamics of a G-quadruplex [8][9]. The second is through imaging methods such as atomic force microscopy and optical traps, which allow for detailed and accurate studying of the forces on the G-quadruplex and its kinetics, but are limited by their single particle nature [10][11]. Furthermore the latter methods are complicated to use and expensive to set up. Due to their promise for a variety of disciplines and research fields a cheaper, easier to use imaging method would open up a realm of possibilities in studying G-quadruplex, and other DNA.

Taking inspiration from a 2010 paper by Halvorsen and Wong a Centrifugal Force Microscope (CFM) was built [1]. Our aim was to use this microscope to determine accurately at what forces the G-quadruplex folds and unfolds. From earlier work done on the CFM it was known that there were several issues with calibration and artifacts in the data. These issues turned out to be much more problematic than was initially expected. Because of this the aim of this research was shifted to solving the problems with the CFM. Although data on G-quadruplex movement was obtained this was only possible at the very end of the research. Regardless in this thesis progress was made on sample creation and a full characterisation of the microscope was achieved.

2 Theory

2.1 G-quadruplex Structure

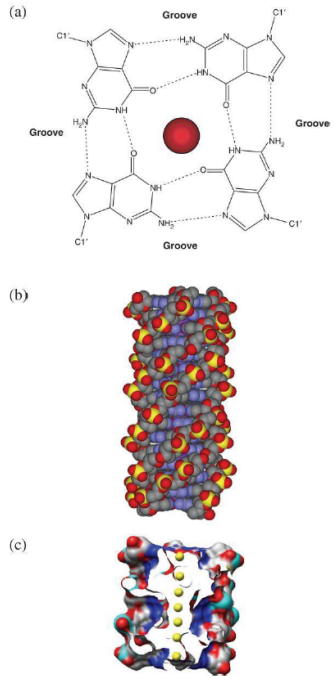

Figure 1: Structure of the G-quadruplex. **A**: Schematic representation of the G-tetrad, **B**: Stack of G-tetrads, **C**: Cross section of G-tetrad stack showing a chain of stabilising monovalent positive ions or cations. [12]

DNA consists of nucleobases and a backbone of alternating sugar and phosphate groups and. Attached to each sugar is one of these nucleobases. The four nucleobases found in DNA are adenine (A), cytosine (C), guanine (G) and thymine (T). In standard double-helix DNA the nucleotides are complementary and held together by hydrogen bonds in so called Watson-Crick base pairs, either adenine and thymine or guanine and cytosine [13].

The G-quadruplex, as well as other nonstandard DNA, is also held together by hydrogen bonds but does not have the double helix structure due to different base pairing. This leads to three or four-strand structures [14]. The G-quadruplex is a four-stranded structure formed from G-rich sequences. These are built around tetrads of hydrogen bonded guanine bases. In figure 1(A) a schematic representation of a G-tetrad is shown. Four guanine nucleobases are arranged in a plane with a stabilising positive monovalent cation in the middle of the plane. The plane is rotationally symmetrical and forms through hydrogen bonds between N_1 - O_6 and N_2 - N_7. A stack of these G-tetrads is a G-quadruplex (figure 1(B)) which is held stable by the monovalent positive ions forming a strand in the middle as well as a π-stacking and a sugar phosphate backbone [15].

DNA also has a direction. With directionality of DNA we mean the biochemical direction. A single strand of DNA or RNA has two ends, one is called a 5'-end and the other the 3'-end [16]. These ends are named after the number of the carbon atom in the sugar-ring it is located at. The 5'-end is the end of the DNA strand that contains the 5th carbon atom in the ring and a phosphate group is attached to this end. The 3'-end is the end of the strand that lies at the hydroxyl group of the 3rd carbon atom in the ring. Through ligation the phosphate group at the 5'-end can bind to the hydroxyl group at the 3'-end to synthesise DNA from loose strands.

G-quadruplexes can form a variety of conformations. In figure 2 some of these possible conformations are shown. They can be classified by the number of DNA strands participating in forming the quadruplex structure. For example in figure 2 the first three conformations are all quadruplexes formed of single DNA strands, they are known as intramolecular G-quadruplexes. The latter two are intermolecular quadruplexes formed from multiple strands. Another way to characterise the quadruplex is to determine whether or not the DNA strands from which it is formed run in the same direction (parallel) or in opposing directions (anti-parallel). Examples of both can be seen in figure 2.

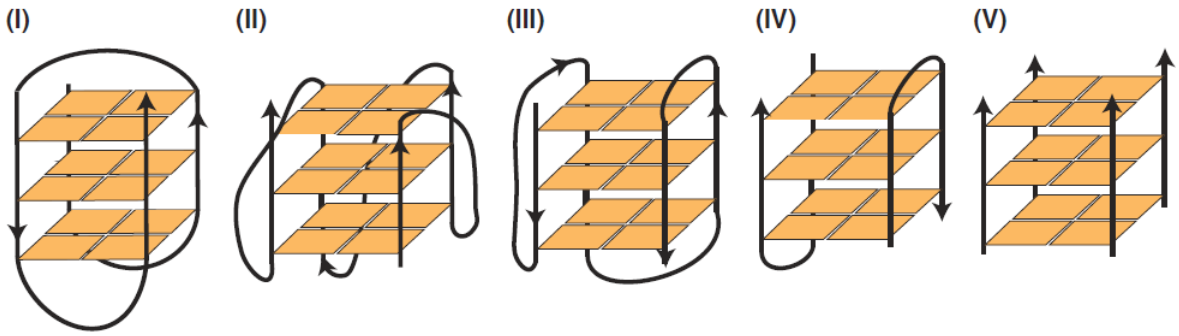

Figure 2: Different possible G-quadruplex conformations. (**I**) monomolecular an- tiparallel G-quadruplex, (**II**) monomolecular parallel G-quadruplex, (**III**) hybrid monomolecular G-quadruplex, (**IV**) bimolecular anti-parallel G-quadruplex, (**V**) intramolecular parallel G-quadruplex. [17]

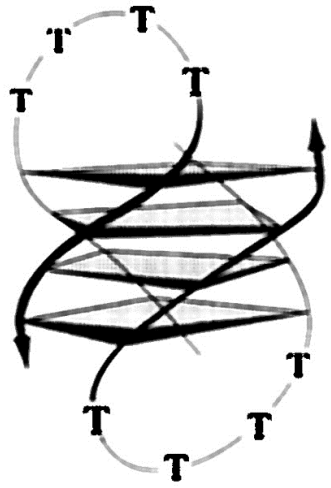

Figure 3: Helical structure of a G-quadruplex. Although the figure shows only Thymine in the loop, this is not necessarily the case for all G-quadruplexes [18].

Although all the images so far have depicted the G-quadruplex as having a backbone that is straight in the vertical direction, this is not the case in reality. Figure 3 shows a more accurate depiction of what a quadruplex looks like. A helical structure similar to that of regular two stranded DNA is visible.

2.2 Polymerase Chain Reaction

To obtain the DNA strands used in my research a process called polymerase chain reaction, or PCR, was used. Building on a template string of DNA one can use enzymatic synthesis, through this process, to create strings of DNA of specific lengths and base pair composition by replicating this template string and attaching the resulting strings to each other [19].

The requirements for this process consist of a template string of DNA, to determine the final product, the enzyme polymerase, a buffer and deoxynucleoside triphosphates (dNTP's) and two oligonucleotide primers [20]. The template is a small amount of DNA which can be any string of DNA of interest. The two primers are complimentary single strands of DNA. These primers can be used to mark the ends of the region we are interested in of the DNA of our template. The DNA polymerase can then fill in a single strand of DNA using the NTP's, this is done at the spot indicated by the two primers. This process can be repeated to create strings of DNA of various lengths.

The PCR process is a set of cycles run generally for around 20 to 40 times. One cycle consists of three temperature changes or steps. The temperature values vary slightly depending on the primers and enzymes used. The times given can also vary slightly depending on the length of the DNA to be copied. The times and temperatures given below are the ones used for the DNA used in this thesis.

- Denaturation or melting: The reaction mixture is heated to 98°C for 30 seconds which causes the hydrogen bonds in the DNA template to be disrupted creating single stranded DNA
- Annealing: The reaction temperature is lowered to 55°C for 45 seconds to allow the primers to attach to the single stranded DNA
- Elongation: The reaction mixture is heated again to 72°C for 30 seconds at which point the DNA polymerases synthesises a complimentary DNA strand

2.3 Fourier Analysis

In this thesis we will often use spectral analysis to obtain information from our images and data. To do this we apply Fourier theory. Fourier theory states that any signal can be decomposed into a number of discrete frequencies or a continuous spectrum of frequencies [21]. To go from a signal to its Fourier spectrum one can use a Fourier transform. In its simplest mathematical form this is given by

$$\mathcal{F}(k) = \int_{-\infty}^{\infty} f(x) e^{-2\pi i x k} dx \tag{1}$$

Where $\mathcal{F}(k)$ is the Fourier transform of the function $f(x)$ for any real number k. One can also use an inverse Fourier transform to retrieve the original function or signal from the Fourier transform. This is given by

$$f(x) = \int_{-\infty}^{\infty} \mathcal{F}(k) e^{2\pi i k x} dk \tag{2}$$

Note that in both these transforms, normalisation is key and this can be done in different ways depending on the precise transform used. One can therefore often find a factor $1/2\pi$ or $1/\sqrt{2\pi}$ in front of the integral.

Many specific types of Fourier transforms exist, which can be roughly divided into discrete and continuous transforms in either the time or spatial domains. As we work with digital images in this

thesis I shall focus on the discrete spatial Fourier transforms in this section.

The discrete Fourier transform of a dataset x of length N is given by [22]

$$\mathcal{X}(k) = \sum_{n=0}^{N-1} x_n e^{-2\pi i k n/N} \tag{3}$$

Here k are frequencies, which are integer multiples of a fundamental frequency, whose corresponding period is the interval at which the dataset was sampled. Note that the transform is normalised by the length of the dataset N. If the input dataset is a set of real numbers, the resultant Fourier transform will be a set of complex numbers of the same length. This means that if x_n is real there will be a relation in the Fourier transform of x_n which reads

$$\mathcal{X}(N-k) = \mathcal{X}(k)^* \tag{4}$$

where $\mathcal{X}(k)^*$ denotes the complex conjugate of $\mathcal{X}(k)$. As a results of this the Fourier spectrum resulting from the transform will be symmetrical, as the positive input signal is split in positive and negative frequencies. There will be two peaks for each present frequency. The exception to this is the peak at frequency 0, which represents the sum of all frequencies.

The absolute value of the Fourier spectrum is called the power spectrum. The power spectrum is a powerful tool in analysing signals and data as the size of a peak in the power spectrum represents the amplitude of the frequency in the signal. This is very useful for analysing the relative amplitudes of various signal components. To obtain the absolute amplitude of a frequency component is a little more difficult. In this case the power spectrum has to be properly normalised. Furthermore the frequency of which one wants to obtain the amplitude has to be an integer multiple of the sampling frequency the data was obtained with. In this thesis we are more interested in the relative amplitudes of our signal components.

2.4 Single Particle Tracking

Obtaining data on the kinetics of something as small as a G-quadruplex, that is to say smaller than visible light which is 400-700 nm in size, can be done using an approach called single particle tracking. Considering the resolution of a light based microscope is limited by the wavelength of light one needs to work around this resolution limit in order to obtain the accuracy required for useful information.

2.4.1 Resolution Limit

When using any optical system it is key to take into account its resolution. With resolution we mean the ability of the system to resolve detail in the image it obtains. This can be defined by the smallest distance between two objects in an image where the system can still distinguish between the two. For any optical system this resolution is limited by two things: aberration and diffraction. Aberrations are caused by imperfections in the optical components of the system and can be minimised with more accurate optics. Diffraction, however, is caused by the wave-like properties of light. The aperture of any optical system is always of a finite size, so any light passing through it will cause a diffraction pattern. This diffraction pattern causes a maximum possible theoretical resolution. One can calculate this theoretical limit through simple wave equations, the result is given by the Rayleigh criterion [23]:

$$R = 0.61\frac{\lambda}{\mathrm{NA}} \tag{5}$$

Where R is the resolution, λ the wavelength of light and NA the numerical aperture of the objective in the microscope. This resolution limit is around 200 nm for visible light.

As the changes in length of the G-quadruplex we want to visualise are smaller than this, ways around this limit for both the x- and y-positions as well as the z-position have to be found. In this thesis several methods were applied to reach sub-pixel accuracies better than the resolution limit of a light microscope.

2.4.2 Localisation & Semi-autocorrelation

The first step to circumvent the resolution limit is through localisation of our measurements. The beads attached to the DNA are one micron in size and hence visible with a standard light based microscope. If we only obtain data, in the form of images, of a region of interest around a bead, where there are no other beads, we are obtaining data on one strand of DNA only. This only holds, of course, when

the sample is correctly prepared so that no bead is attached to multiple strings of DNA. This makes the question of whether or not we can discern between two strands close to each other irrelevant. To now obtain information about the position of the bead smaller than our resolution limit we use a form of cross-correlation we shall call semi-autocorrelation. As correlation functions are not limited by pixel size when used correctly they can give us information on our tethers in the sub-pixel range.

In signal processing an autocorrelation is a cross-correlation of a continuous or discrete function with itself. A cross-correlation being defined as:

$$(f \star g)(\tau) \equiv \int_{-\infty}^{\infty} f(t)g(t+\tau)dt \tag{6}$$

With f and g two continuous functions, f denoting the complex conjugate of f and τ the time lag. Substituting f for g gives us the autocorrelation of f. When working with discrete data the cross-correlation is given by[24]:

$$(f \star g)[n] \equiv \sum_{m=-\infty}^{\infty} f[m]g[m+n] \tag{7}$$

Instead of correlating in the time-domain one can also do correlations in the spatial domain. When working with images this is often a discrete 2-dimensional spatial correlation, which we shall call $c_{i,j}$. This is given by

$$c_{i,j} = \frac{\sum_m \sum_n [f(m+i, n+j) - \overline{f}][g(m,n) - \overline{g}]}{\sqrt{\sum_m \sum_n [f(m,n) - \overline{f}]^2 \sum_m \sum_n [g(m,n) - \overline{g}]^2}} \tag{8}$$

Where $f(m,n)$ and $g(m,n)$ are the intensity of the pixel at point (m,n) in the original and translated image respectively. $\overline{f}$ and $\overline{g}$ are the mean values of the intensity matrices f and g.

Correlation is often done in Fourier space, as a convolution is simply a multiplication in Fourier space with a complex conjugate and a fast Fourier transform is a computationally less taxing operation. If we take $\mathbf{F} = \mathcal{F}(f)$ and $\mathbf{G} = \mathcal{F}(g)$ as the Fourier transforms of our intensity matrices we can compute

$$C = \mathbf{F} \circ \mathbf{G}^* \tag{9}$$

where $\mathbf{G}^*$ denotes the complex conjugate of $\mathbf{G}$. After this we can obtain the cross correlation by applying an inverse Fourier transform.

$$c = \mathcal{F}^{-1}(R) \tag{10}$$

What we call a semi-autocorrelation in this thesis is a spatial cross-correlation of an image with a copy of itself rotated by 180 degrees. The peak of this correlation can now be used to obtain the x and y position of the bead relative to the centre of the region of interest. How this works is that the correlation will peak when the beads completely overlap. As the two images being correlated are mirrored with regards to the centre the location of this peak tells us how far the bead is from the centre and in which direction. With these two values we can find the x and y position of the bead. All our data acquisition is done in LabView. In figure 4 an example of a semi-autocorrelation of a bead as seen in LabView is shown. With this method the sub-pixel accuracy we can obtain is limited only by the level of noise in our signal, as the correlation can be done in steps smaller than our pixel size.

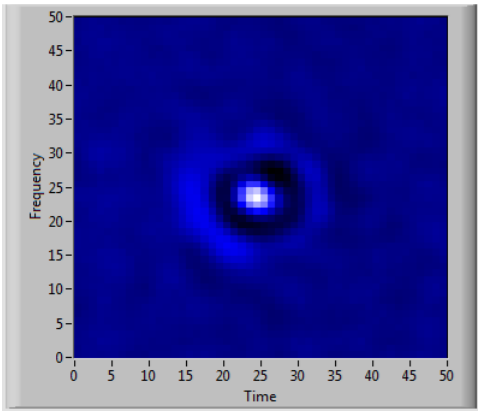

Figure 4: Semi-autocorrelation of a bead

For the z-position a different technique was used. A region of interested around a stuck bead was selected after which a z stack kernel was obtained by imaging the bead for a set amount of frames for a chosen set of z values, these z values were obtained through use of a piezo driver. A mean was obtained of the image for every z-position. The regions of interest of actual beads can then be correlated with this kernel and the z position obtained through a maximum of this correlation with one of the images from the kernel for each frame. The precise z position value can lie between the values of the images in the kernel, the correlation allows for interpolation. In figure 5 a fit of an ROI frame to a z-stack kernel is shown.

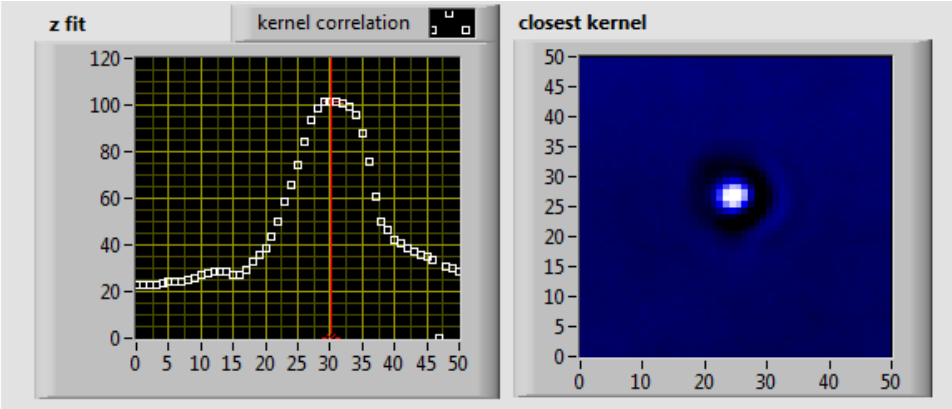

Figure 5: Fit of bead image to a kernel. The left figure shows the maximum of the correlation plotted against the kernel number. The right figure shows the kernel which has the highest correlation maximum

This fit does not give us the exact z position, as the kernel values are based on the values of the piezo. However, as we want to see changes in position and not necessarily the absolute position, this is not a problem.

Aside from the three positions of the bead we also use our LabView program to acquire other parameters per frame. These are the maximum of the correlation and dynamic range of the region of interest. The former can be used for a goodness of fit test and the latter is a good indicator for general quality of the image. We also acquire the motor speed and piezo position for all ROI's in a single measurement per frame.

2.4.3 Bead Positional Signal

Now that we have a method to obtain positional data of our bead it is useful to look at the various movements we expect to be present in this signal. The best way to do this is to look at the power spectrum of the signal, or the absolute value of the Fourier spectrum of the signal. Provided that this Fourier spectrum is properly normalised. This gives us the frequencies present in the signal and their amplitudes, so long as these frequencies are temporarily stable and are either sinusoidal or a sum of sinusoidal signals. We expect that we will find four separate major contributions to the signal. The movement of the tethered bead due to the Brownian forces acting upon it, the movement of the bead due to the folding and unfolding of the G-quadruplex, drift and background noise. As we are going to be working with digital images we can assume our background noise to be Gaussian noise [25]. We expect the Brownian motion to be roughly several tens of nanometers for a tethered bead. The folding and unfolding should present itself as a jump in the signal of around 20 nm. Drift can be larger than this but is much slower and should show itself in the signal as a slow increase or decrease in the mean of the position. For a size estimate of the noise I refer to section 2.4.4. It should be easy to discern between movement of the bead, drift or otherwise, and noise in the Fourier spectrum as Gaussian noise contains very high frequency relative to any physical movement of our tethers. In figure 6 some positional data of a bead with visible drift is shown and its power spectrum.

We can see drift, movement and noise in the signal in figure 6 A. The power spectrum shows a strong zero peak followed by a quickly falling curve. Note that generally the power spectrum is symmetrical, minus the zero peak. This is due to the negative frequencies which are a general property of the Fourier transform as discussed in section 2.3. I have removed these negative frequencies in figure 6 **B**. For any calculations done in this thesis on the power spectrum of this data we remove the mirrored part. The quickly falling curve is the low frequency movement which contains the drift and other movement of the bead. Considering that this is a stuck bead we should see neither folding nor unfolding and Brownian

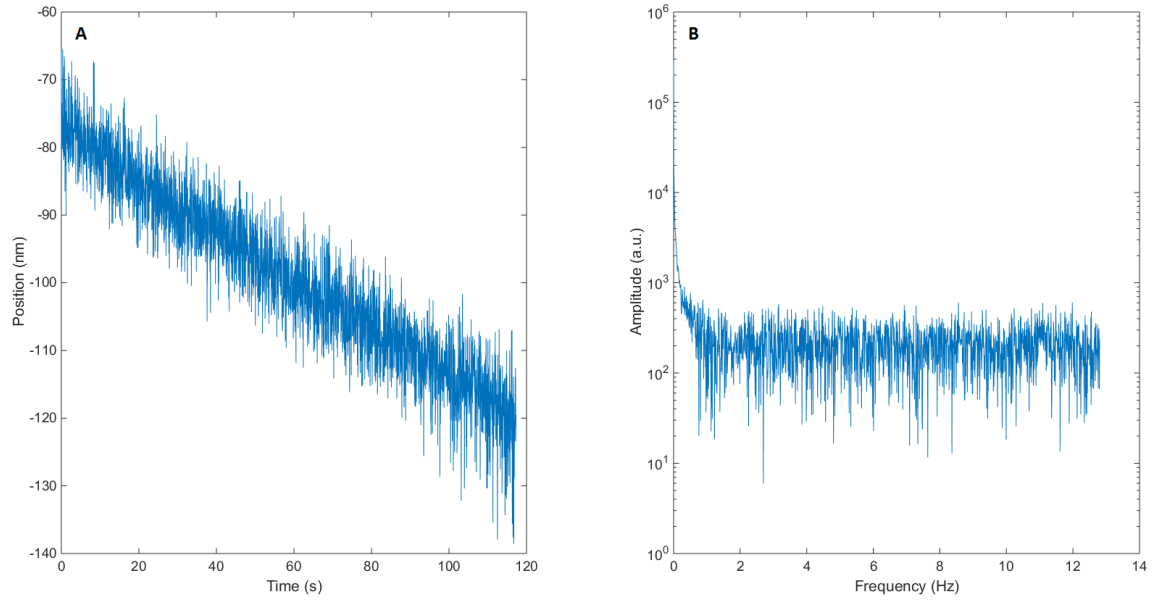

Figure 6: **A**: Positional signal of a bead in nanometers showing signs of drift and noise, **B**: logarithmic plot of the power spectrum of the signal plotted as amplitude against frequency.

motion. After this the power spectrum remains relatively constant, these are the frequencies caused by the Gaussian noise.

Although it is easy to distinguish between the contributors to the noise and movement, discerning between different types of low frequency movement is harder in Fourier space. Fortunately this is not a problem. Subtracting a wide band moving average from the signal, if drift is present at all, will remove the drift without altering the slightly higher frequency bead movement caused by forces such as Brownian motion. After this the power spectrum can be used to discern noise from the movement we want to analyse.

2.4.4 Noise Estimation & Parseval's Theorem

As was stated in section 2.4.2 the sub-pixel accuracy we can obtain is dependent on the level of noise present in our signal. Although our noise also contains lower frequencies we assume that the movement we want to analyse contains no very high frequencies whatsoever, based on earlier research done towards DNA looping. To estimate the noise of our positional data we can use a mathematical analysis through Parseval's theorem. This theorem, in its discrete form, states that the total energy of a dataset summed over all its elements is equal to the total energy of the dataset's discrete Fourier transform summed over all of its frequencies components. Or in mathematical form:

$$\sum_{n=0}^{N-1} |x[n]|^2 = \frac{1}{N} \sum_{k=0}^{N-1} |X[k]|^2 \tag{11}$$

With $x[n]$ the nth element of dataset x and $X[k]$ the kth frequency component of its Fourier transform. Note that for Parseval's theorem we have to normalise by the length of our data, hence the $1/N$ prefactor on the right hand of the equation. With the assumption that our noise is of a relatively high frequency compared to the actual movement of the bead we can sum over the high frequency region of our power spectrum to obtain a rough estimation of our noise level and therefore our sub-pixel accuracy.

For the data shown in figure 6 we subtract the local mean obtained with a moving average. If we then apply Parseval's theorem by summing over the final third of the power spectrum, after removing the mirrored half, and multiply this by three we obtain a noise estimate of 4.7 nanometers. Looking at the data this looks like a good estimate, we can consider our assumptions to hold.

3 Set-up

3.1 Types of Samples

For these experiments we created three different types of samples. Each of these was prepared inside a channel made by cutting a rectangle out of parafilm of 5 mm × 25 mm. The parafilm was put between a Thermo Scientific microscope slide of 76 mm × 26 mm and a 24 × 50 mm coverslip with a thickness of 170±5 μm. This was then heated so as to melt the parafilm and seal the channel. Two holes are drilled through which fluids can be flushed. Before the samples are made the coverslips are treated with 8% hydrofluoric acid for 1 minute before being put in demineralised water. The HF etches the glass of the coverslip which cleans it thoroughly and minimises opportunities for unwanted non-specific binding. A schematic representation of the sample is shown in figure 7.

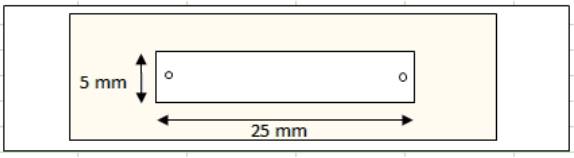

Figure 7: Schematic representation of an empty sample.

The three types of samples created were as follows:

- 'Stuck' samples consisting of silica beads bound straight to the glass
- Tethered samples of silica beads attached to the glass with standard DNA strings
- Tethered samples of silica beads attached to the glass with DNA containing one or multiple G-quadruplexes

For all the samples the beads used had a diameter of 1 μm. The first of these, the samples containing stuck silica beads, were used to determine the accuracy of the set-up. These beads should in theory not move at all, unless they have bound unspecifically to the glass or other objects, so all apparent movement is noise or artifacts from the set-up. These samples are easily created by filling the sample with buffer (20 mM PBS at pH 7.4), then adding silica bead solution and flushing out loose beads, after three minutes of waiting time. The sample is sealed after preparation.

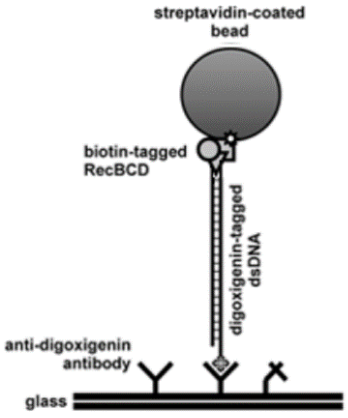

Figure 8: Schematic representation of the tethers used in the experiments [26]

The second and third sample types require more elaborate preparation methods. An example of what the tethers in these samples look like can be seen in figure 8. After the sample is filled with buffer a solution of 0.01 μ g/ml antidigoxigenin is added. After this a modified version of the PBS is added which contains 20 μg/ml BSA. This is done so that anywhere in the sample where there is no antidigoxigenin is covered in BSA, which should hinder beads from attaching directly to the glass. In an eppendorf filled with the same modified buffer the DNA and beads are mixed. Depending on the type of sample this DNA is either standard double helix DNA or DNA containing quadruplexes. The beads are covered in streptavidin which allows the DNA to attach to the bead. This is due to the fact that our DNA tethers have digoxigenin on one end, to bind to the antidigoxigenin on the glass and biotin on the other, which can bind to streptavidin. This solution is flushed through the sample so that the digoxigenin on the other end of the DNA can attach itself to the antidigoxigenin. Finally the sample is washed with more buffer and sealed. For the exact quantities in appendix A the lab protocol used for preparing the tethered samples is shown.

3.1.1 DNA Amplification Using PCR

For the creation of the tethered samples appropriate strands of DNA had to be isolated and created. This was done using PCR. The base DNA used for our PCR was pUC18. For the regular double-helix dna this pUC18 is put through the PCR procedure as described in section 2.2. A G-quadruplex can be inserted by using the enzymes HindIII and BanII to cut a piece of DNA from the pUC18 creating two handles into which a quadruplex can be spliced. The entire sequence of the PCR product made including the location where these enzymes cut is shown in figure 9.

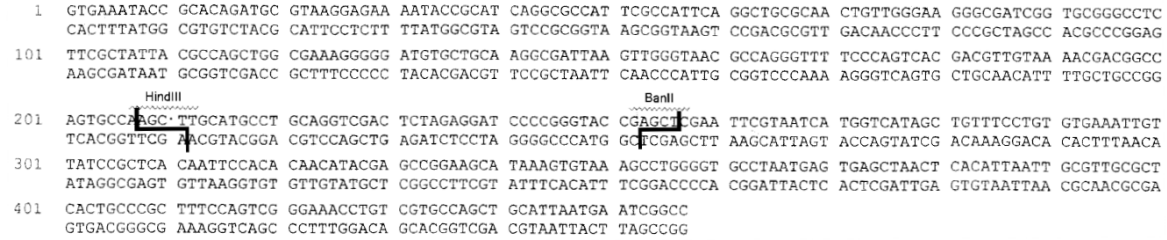

Figure 9: Base pair sequence of the DNA strand used. The locations where the enzymes remove the normal base pairing to be replaced with a single-stranded G-quadruplex are shown [27].

The G-quadruplexes used in our experiments were made by a previous student [27]. For the standard double helix DNA tethers used in one of our sample types we ran a new PCR on four batches. Thus creating simply the DNA shown in figure 9. In figure 10 the result of the PCR process is shown.

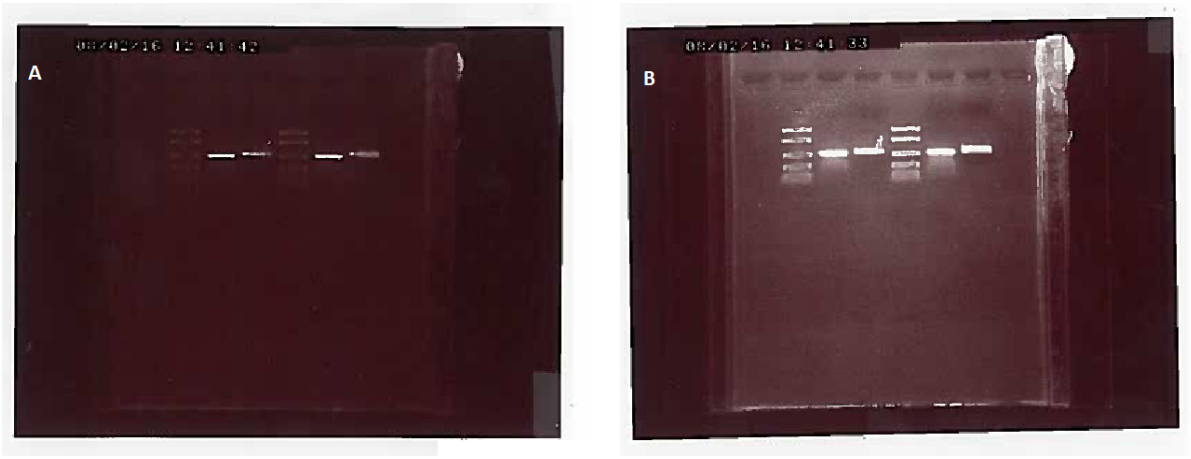

Figure 10: **A**: Normal image of our PCR product for standard double helix DNA, **B**: Overexposed image of our PCR product for standard double helix DNA. The ladder visible in both images is a reference mixture of DNA of known lengths showing the 50, 200, 400, 850 and 1500 base pair marks, increasing from bottom to top

In figure 10 14 bands can be seen. Each vertical lane is a sample, in each lane DNA is separated by electrophoresis and made visible with ethidium bromide, an intercalating dye. These batches were run at 120 V for 15 minutes. The gel was a 2 % Agarose slab. The two ladder tracks of 5 bands were markers to calibrate how far a certain length of DNA would move through the buffer. Next to each ladder are two bands for each batch. These are the unwashed and washed versions of the respective batches. With washed we mean that a PCR clean-up kit was used on the final product. Without washing one can not make sure that the DNA is fully rid of enzymes, primers, salts, dyes and so forth. The batches made measured to be 124 ng/μL and 141.5 ng/μL respectively with a nanodrop spectroscope. Furthermore from the image it can be seen that the length is, as expected, around 400 base pairs.

Both the G-quadruplex DNA and standard DNA were diluted to

3.2 Centrifugal Force Microscope

To study the mechanics of G-quadruplex DNA in real time we have used a centrifugal force microscope, or CFM, inspired by Halvorsen and Wong [1]. The general idea is to mount a normal light microscope on a rotating stage. Powered by an electromotor this stage can rotate fast enough to apply a centrifugal force field to an ensemble of particles on a sample whilst visualising them at the same time. A cross section and top down view of the CFM used in our experiments are shown in figure 11

The CFM is built on an Aerotech ADRT 150-135 rotary stage. Furthermore the driver for the stage is an Aerotech Soloist CP20. A ThorLabs LED528E is used for the light source (1.5 mW at 525 nm with a bandwidth of around 20 nm) and filtered with a bandpass filter in order to exclude ambient light. The camera and objective are an Allied Vision GC2450 CCD camera and Olympus PLN20 20x objective respectively. For controlling the z-position of the sample we use a Thorlabs TPZ001 150 V piezo driver.

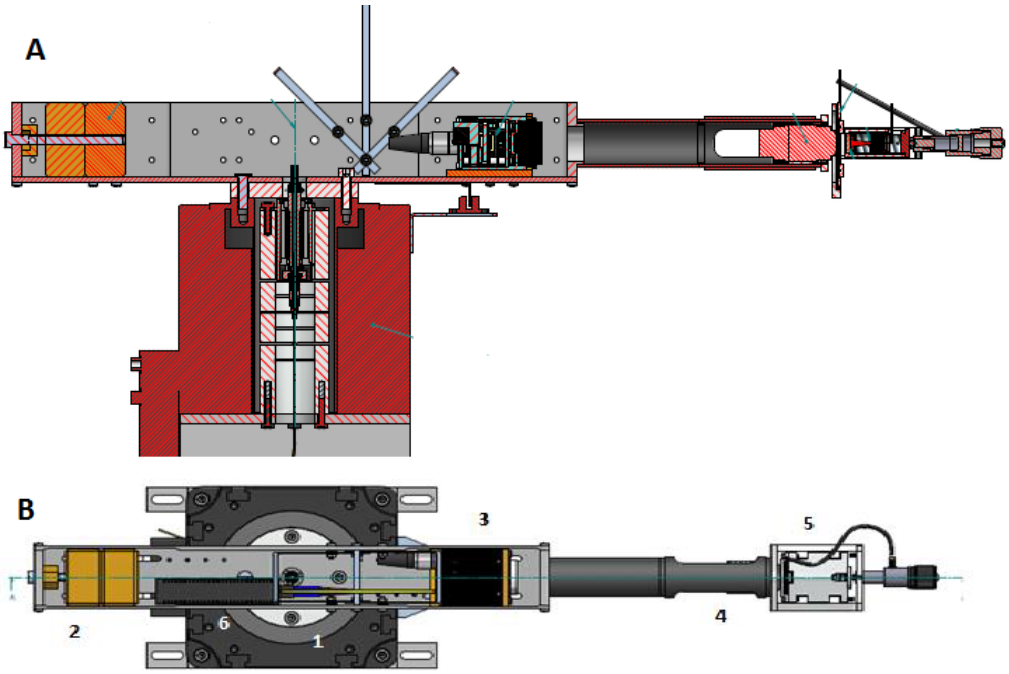

Figure 11: **A**: Cross section of the CFM, **B**: The CFM as viewed from above. The labels are as follows **1**. Motor, **2**. counterweight, **3**. camera, **4**. objective, **5**. sample holder and **6** box for transferring data to a computer. [28]

These components allow for the CFM to move with a maximum rotation speed of 600 revolutions per minute. At 600 rounds per minute the 1 micron diameter silica beads with which the samples are created experience a force of 1.19 pN. This is more than adequate as the DNA looping and folding we want to see occurs at forces in the femtonewton range [29]. The 150 V of the piezo driver allow us to move the sample 1.5 μm in the z-direction.

4 Results & Discussion

4.1 Forces on Tethered Beads

Although we can apply force on our tethers by using the CFM to put our silica beads in a force field, there are obviously more forces involved, even when the CFM is stationary. It would be ideal for the accuracy of our force measurements that said other forces are negligible, or can in other ways be accounted for. In reality however there will be other forces which will add to the noise of the signal or be visible in other ways. Before analysing any data obtained with the CFM we therefore first take a look at the forces on the bead tether system.

First of all it is easy to establish that when looking at a tethered bead in our sample that the gravitational and Brownian forces on the DNA tether are negligible compared to those on the bead. A silica bead with a diameter of 1 μm with a density 2.643 g/cm^3 weighs 1.39 fg. If we do a back-of-the-envelope calculation for the tether we can say that several hundreds of base pairs weigh less than a megadalton. If we take an overestimate and say the maximum weight of a tether is one megadalton, this would still only correspond to less than 2 attograms [30]. It is clear that, for these forces, the tether is negligible compared to the bead

So we can focus on the forces on the bead. These can be split into four different forces. Brownian motion in the buffer, gravitational pull on the bead, the pull from the CFM and the friction between buffer and bead caused by the latter two forces.

$$\vec{F}_{\text{tether}} = \vec{F}_{\text{grav}} + \vec{F}_{\text{brown}} + \vec{F}_{\text{CFM}} + \vec{F}_{\text{fric}} \tag{12}$$

Note that we assume here that the fluid inside the sample is incompressible. For the friction term we assume that the movement of our system is governed by the Langevin equation for Brownian motion. The equation for this motion as described by Langevin is [31]

$$m\frac{d^2\vec{x}}{dt^2} = -\lambda\frac{d\vec{x}}{dt} + \vec{\eta}(t) \tag{13}$$

with $\vec{x}$ one degree of freedom of the particle, m the mass of the particle, λ a damping coefficient and $\vec{\eta}(t)$ a noise term which represents the effect of collisions between the bead and the molecules of the fluid. This is given by a Gaussian probability distribution with the correlation function

$$\langle \eta_i(t)\eta_j(t')\rangle = 2\lambda k_B T\delta_{i,j}\delta(t-t') \tag{14}$$

Where k_B is Boltzmann's constant, T is the temperature in Kelvin and δ is the Dirac-delta function. For three-dimensional movement we have to add up the three possible degrees of freedom.

This is essentially a simplified version of the Langevin dynamics for a multi-particle system which reads

$$M\frac{d^2X}{dt^2} = -\nabla U(X) - \lambda\frac{dX}{dt} + \sqrt{2\lambda k_B T}\delta_{i,j}\delta(t-t') \tag{15}$$

with M the masses of the particles and $U(X)$ the interaction potential potential between particles. We see that if we assume that there is no interaction between our beads, and we drop the interaction potential, we arrive back at equation 13.

If we assume now that our beads are overdamped. That is to say always in a state of terminal velocity, which for short time-scales holds we can find the contribution from both frictional and Brownian forces. We first then use that for every degree of freedom we have $\frac{1}{2}k_BT$ of thermal energy which gives us $\frac{3}{2}k_BT = 6.213 * 10^{-21}$ J if we take $T = 300$ K. This is equal to $\frac{1}{2}mv^2$ so we can insert the mass of our bead (1.39 pg) to get the terminal velocity. This gives us $v = 9.7 * 10^{-3}$ m/s. If we were to work with normal Brownian dynamics this velocity would give us a force of $F = -91.4$ pN. However to arrive at this value we assumed that thermal motion and friction always cancel each other out. The actual force on the bead from frictional and Brownian sources should therefore always be 0. This is obviously also not the case so for a more detailed look we look at a paper from Chen et al from 2009 where a more quantitative analysis of the forces resulting from thermal fluctuations on DNA and ended up with a small result: a residual entropic non-directional force in the range of around 40 fN for similar tethers [32].

As for the gravitational forces our samples are inserted vertically into the set-up meaning that directionally the gravitational force pull the bead parallel to the glass whereas the CFM pulls it perpendicular. The Brownian motion is obviously random and the frictional forces act in the opposite direction of the forces causing them. As for a size estimate we can easily establish the gravitational force. As we are

dealing with beads in a fluid we have to replace the mass m in the equation $F_{\text{grav}} = m * g$ with the relative mass

$$m = (\rho_{\text{bead}} - \rho_{\text{buffer}})\frac{4}{3}\pi R^3 \tag{16}$$

If we insert the bead density and a buffer density of 1.004 g/cm^3 into this equation we obtain a gravitational force of $F_{\text{grav}} = 8.4$ fN. This is approximately a factor 5 smaller than the thermal forces and as such does not play a large part in bead movement. We could however expect that movement in the direction of the gravitational force will be slightly larger than its perpendicular counterparts.

Lastly we look at the force field caused by the microscope. This is dependent on the rotation speed. This is simply the centripetal force

$$F_{\text{cent}} = mr\omega^2 \tag{17}$$

where m is the mass of the object, r the radius of curvature and ω the angular velocity. In our case the radius r is the distance between our sample and the axis of rotation of the microscope which is 0.35 m. Once again we substitute the relative mass using Archimedes' principle. This gives us an equation for the force on our beads

$$F_{\text{CFM}} = (\rho_{\text{bead}} - \rho_{\text{buffer}})\frac{4}{3}\pi R^3\omega^2 \tag{18}$$

we assume that the force is uniform over the sample. If we insert the densities we can plot the force as a function of revolutions per minute. The result of this is shown in figure 12

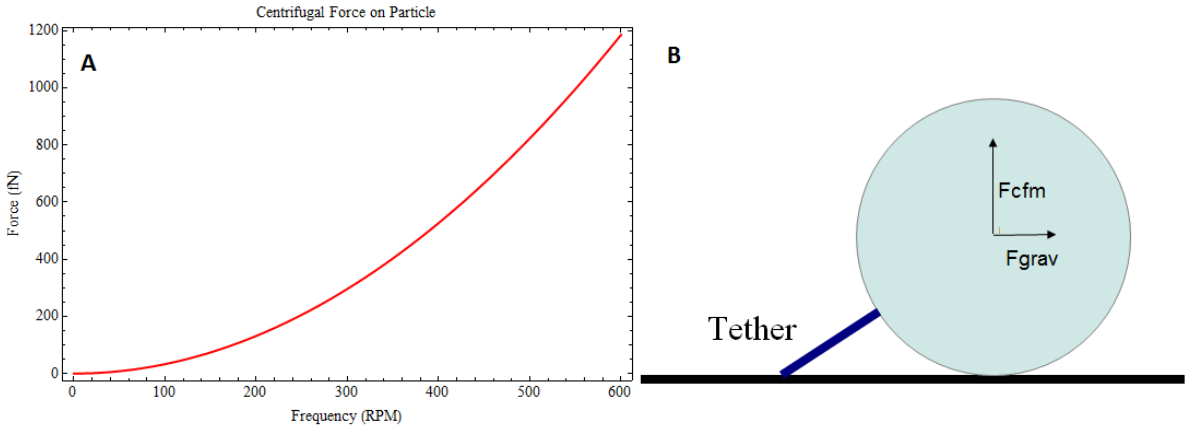

Figure 12: **A**:Centripetal force on 1 micron silica beads as a function of the rotational speed of the CFM, **B**: Schematic diagram of relevant forces on bead

4.2 Sample Preparation Protocol

I spent a significant amount of time creating samples to be used in the CFM. After a period of coming to grips with the basics of biochemical sample preparation, I made attempts to tweak the protocol in consultation with the group technician.

The first samples created for this thesis used a protocol made by a previous student [33]. To arrive at the final protocol in appendix A several modifications had to be made after trial and error attempts during my research. Firstly the old protocol used two versions of a buffer dubbed PTC. This was a buffer containing Tris, KCl, $MgCl_2$ + 6 H_2O, EDTA, BSA and Heparin. The difference between the two versions depended on the quantity of BSA. As the research was done several years ago new buffer had to be prepared. Instead of recreating the PTC a simple phosphate-buffered saline (PB) solution was attempted first, due to its easy production and a lack of solid arguments for including the other ingredients of PTC. The quality of samples was not significantly better or worse than the ones created with PTC, so the PTC was replaced by two versions of the PB. One without BSA and one with 1 mg/μL BSA dubbed PB2.

The previous protocol also called for sonicating the bead buffer mixture in an eppendorf before adding DNA. This was done to to remove possible clusters of beads from the solution. With proper shaking and mixing it was found that clustering was not a problem. Furthermore the streptavidin attached to the beads could be affected by sonication and hence lead to a lower chance of acquiring properly tethered beads. For these reasons the sonication was removed from the protocol.

Regardless of the tweaks, sample creation success is very low. Both of the tethered sample types are

often unsatisfactory even when the protocol is followed exactly and succesfully. Results with no tethers, only stuck beads, no beads whatsoever and so forth are very common. Aside from this mechanical mishaps, such as air bubbles in the trough or a leak in the parafilm, are hard to avoid consistently. Air bubbles are not necessarily a problem when the CFM is stationary, causing anomalies only at the interface between the bubble and the sample, but when rotated could be very disruptive. As the CFMs major advantages lie with an increase in efficiency a better sample creation set-up or protocol is necessary. One option is to replace the two tubes of silicon, through which the solutions are flushed manually with a pipette, with a drip system. That is to say connect one tube with a small container, such as a pipette tip or eppendorff, and fasten this at a great enough height that gravity will drag any liquid placed in the container through the trough. With this system air bubbles should no longer be a problem.

4.3 Characterisation of Positional Signal

Although already constructed, calibration of the microscope took up a sizable amount of time. A variety of problems on various fronts were encountered and numerous ways were devised to circumvent or solve them. Early on during my research, computer and LabView problems obstructed attempts to get acquianted with the CFM. Once this was solved to a workable degree I began work on calibrating the set-up.

4.3.1 Engine Artifacts

From work done by a previous student it was established that the motor used to drive the CFM created artifacts in the data [34]. As such the first priority during this research was to use the stuck silica bead samples to accurately characterise the set-up and remove or circumvent these artifacts. The first artifact found during this research was that the motor, without necessarily driving the microscope to rotate could cause a periodic signal to appear in the data. A test run was done in which several stuck beads were tracked for around 5000 frames. Every 1000 frames another step in the process of acquiring data was done.

- 0-1000 No rotation, engine off
- 1000-2000 No rotation, engine on
- 2000-3000 100 RPM rotation, engine on
- 3000-4000 No rotation, engine on
- 4000-5000 No rotation, engine off

The resulting data of this measurement for one bead is shown in figure 13. The change in position seen between frames 2000 and 3000 can be explained because the Z-position of the particle changes as force is applied. This causes a shift in the x and y position as well.

It is worth noting that this is a worst case example. These artifacts were not always present in the data.

It is clear that once the engine is turned on, in this case, a periodic signal is visible in the data. After rotation this is no longer visible, however it is clearly more noisy than the final 1000 frames where the engine is simply turned off. This suggests that the engine simply being on introduces artifacts into our data. This could be due to its power draw. A drop in voltage could cause the LED, which is connected to the same power source, to lose intensity. This in turn could decrease the dynamic range of the ROI's and hence the resolution. In figure 14 the dynamic range and goodness of fit of the same dataset are shown. We can see a drop in dynamic range as soon as the engine is turned on.

Furthermore we see that the maximum of the correlation increases strongly when the engine starts rotating, while the dynamic range decreases. This could possibly be explained by the change in z position as the engine starts moving, this would suggest however that our goodness of fit is correlated to our z-position

Another possibility, which could explain the periodicity of the artifact is that the engine oscillates around a position it considers to be its starting position when turned on. The artifact turned out to be unpredictable and when the microscope was carefully initialised it could be kept absent from data. For further research on the same set-up I have included an initialisation protocol in appendix B.

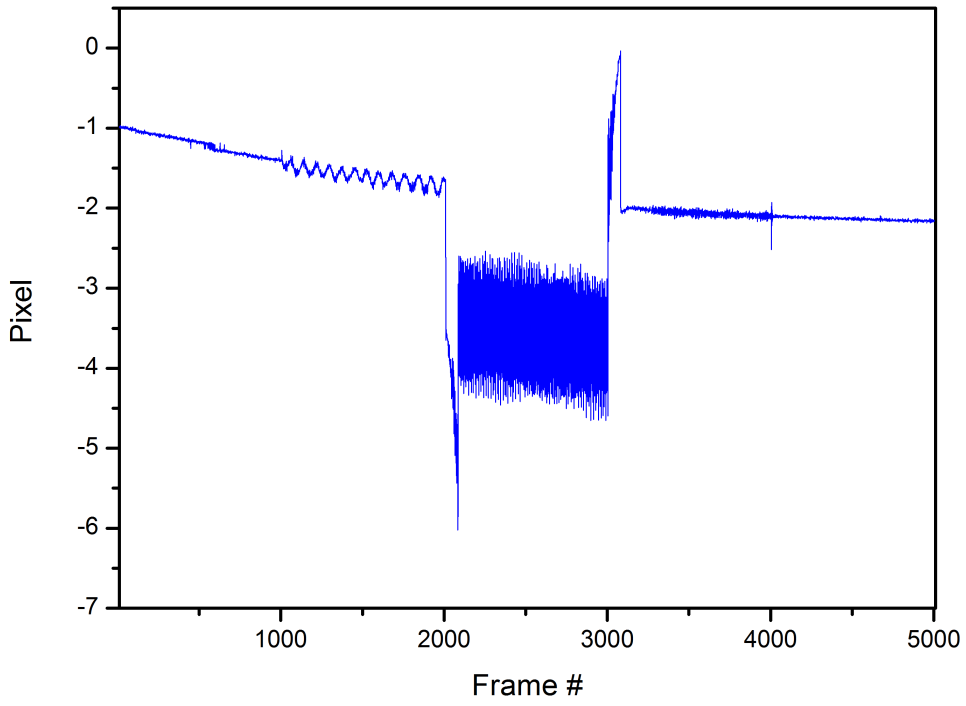

Figure 13: X position of bead in pixels

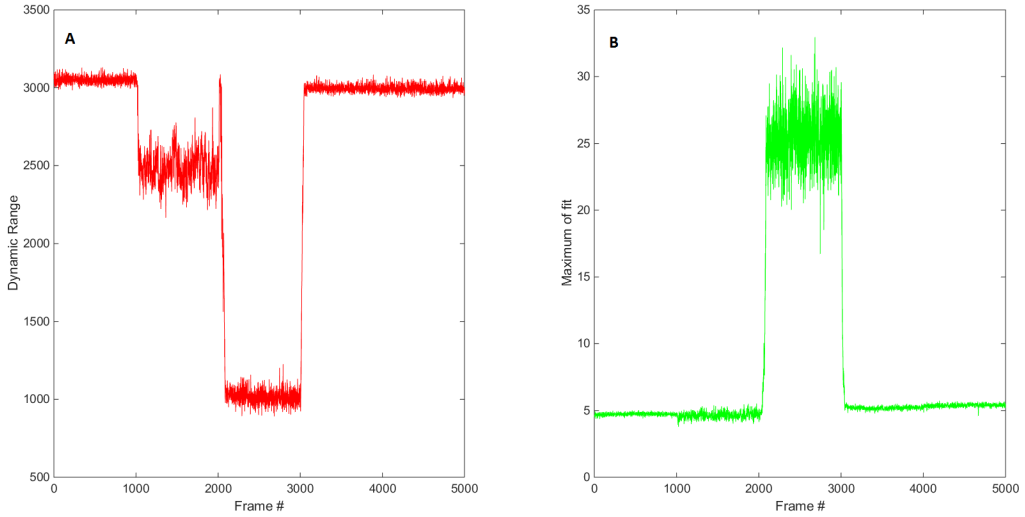

Figure 14: **A**: Dynamic range of image per frame, **B**: Maximum of the fit of the bead position per frame

4.3.2 9-Fold Symmetry

A second problem we encountered was that the engine appeared to have a 9-fold symmetry [34]. The data obtained from the microscope therefore contains a periodic movement of 9 times the rotational speed of the microscope. To be able to obtain accurate data with the microscope rotating this artifact had to be isolated and removed. The first solution we tried was to find the peak in in the power spectrum of the data and average it out. We expect the interesting movement of the bead to have a considerably lower frequency than the 9-fold peak and hence we would not be removing any information on the movement when removing the peak. That is to say Brownian motion we expect to be high frequency,

whereas the kinetics of the G-quadruplex should be much slower. In figure 15 an example is shown of data of a stuck bead rotating at a 100 revolutions per minute and an attempt to remove the artifact peak. Although the result could have been improved by replacing the averaging with a more complex substitution, such as a linear approximation, and further improved by proper zero padding, it became apparent that this method was coarse and complicated to accurately automate. Furthermore aside from the large 9-fold peak, there seem to be smaller but non-negligible peaks at other frequencies too high to correspond to the movement of the bead as well.

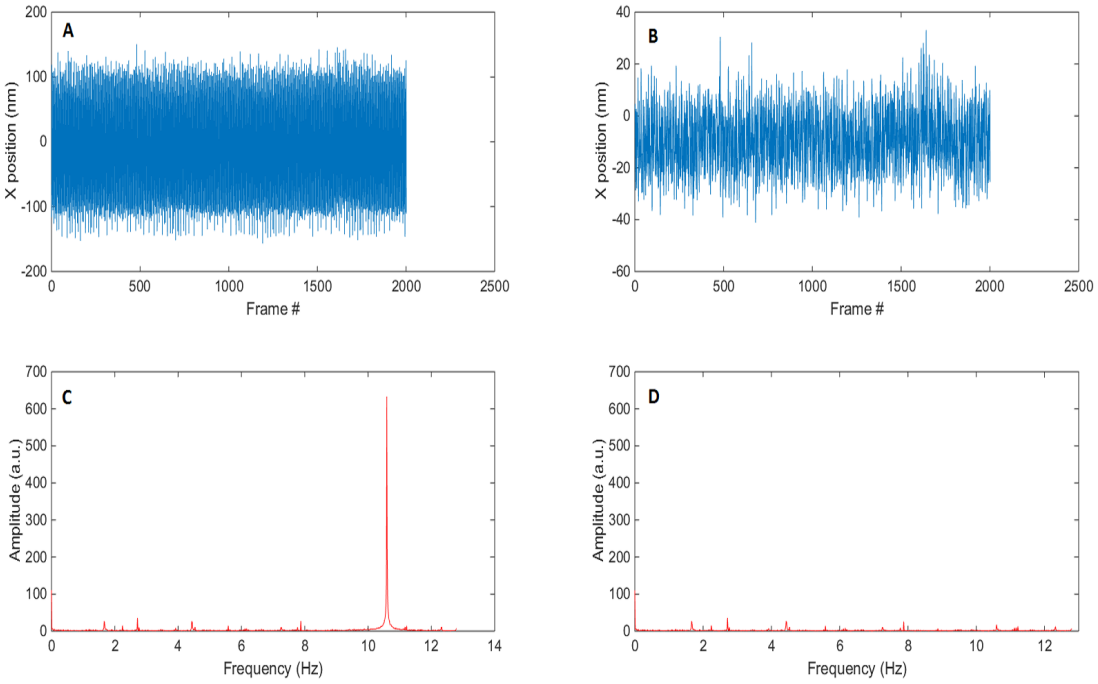

Figure 15: Example of peak removal for a stuck bead rotating at 100 RPM. **A**: X position of bead in nanometers unmodified, **B**: X position of bead in pixels restored from modified Fourier transform, **C**: Unmodified power spectrum of X position, **D**: Power spectrum of X position with peak averaged out. I have chosen to plot the power spectrum in a linear plot here to emphasise the peaks present.

In order to circumvent having to manually remove the peaks a second solution was tried by measuring at revolution speeds for which the acquisition frame rate of our camera would be a multiple of the 9-fold peak. The GC2450 camera used in these experiments has a frame rate of 14.8 frames per second at 2448 x 2050 pixels. For most measurements we used a fourth of the total pixels at 1280 x 1024. This gave us a frame rate of 25.6 frames per second. At this frame rate we should be able to rotate at 170.6 and 341.2 RPM without seeing the artifact in our data. Although this appears limiting, by tuning the height of the frame any frame rate between 14.8 and 50 frames per second is achievable and therefore many rotation speeds can be used while utilising this technique.

In figure 16 the X and Y positions of a stuck bead measured at 170 revolutions per minute are shown.

It is clear that the 9-fold symmetry in the X direction is not removed immediately through this method. What appears instead is an apparent sine wave with a relatively long period. This could be explained by the fact that the engine will never rotate exactly at the required speed perfectly. So instead of avoiding the 9-fold symmetry what we see is temporal aliasing.

Temporal aliasing occurs when one does not fulfil the Nyquist criterion. The Nyquist criterion, sometimes called the Nyquist sampling theorem, states that the sampling frequency used when analysing a signal should be at least twice the highest frequency contained in that signal [35]. So $f_s \geq f_{sig}$ has to hold. When this is not the case we are undersampling and aliasing occurs. This can lead to artifacts or misrepresentation of the signal. In figure 17 this can be seen.

The undersampled sine wave appears to have a lower frequency than it actually has. This appears to be happening with our data when we rotate the CFM in sync with the 9-fold symmetry.

As an analogy we can view this as being similar to a beat frequency, where the beat is caused by

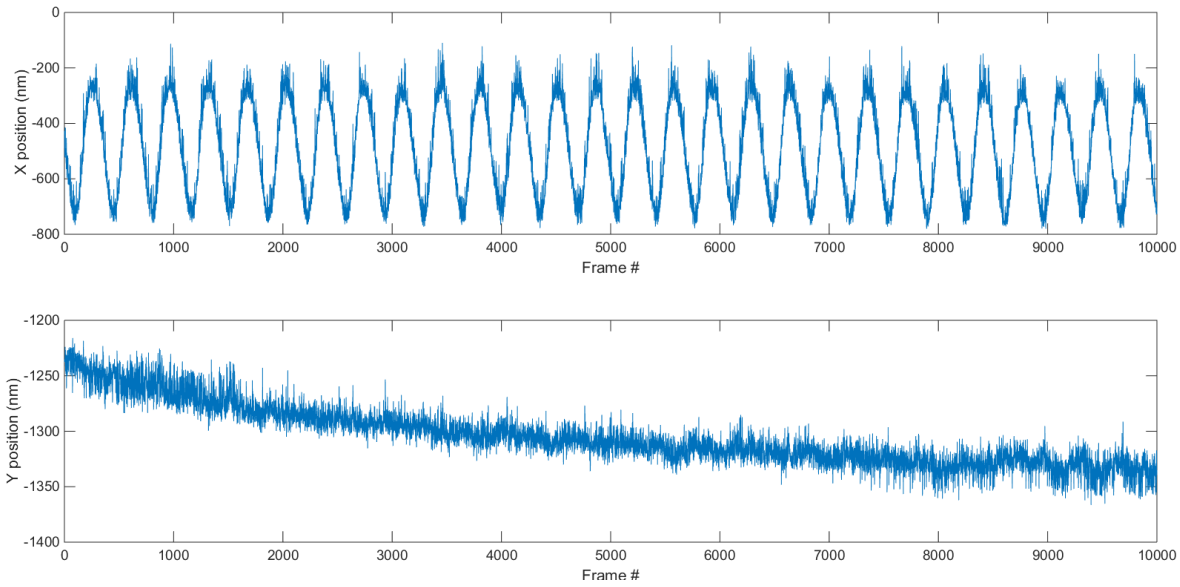

Figure 16: X and Y positions of a stuck silica bead measured using the CFM at 170 revolutions per minute

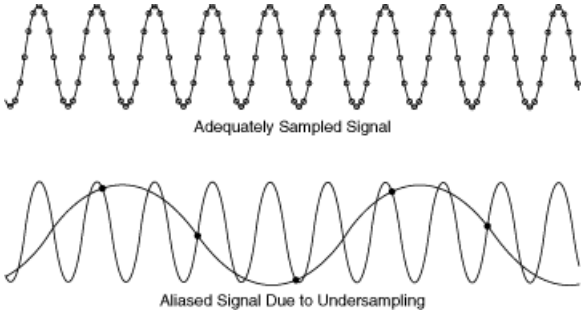

Figure 17: Representations of a correctly sampled and undersampled sine wave [36].

the slight difference in frequencies between the frame acquisition and rotation speed. Trigonometrically this looks like

$$\cos(2\pi f_1 t) + \cos(2\pi f_2 t) = 2\cos\left(2\pi\frac{f_1+f_2}{2}t\right)\cos\left(2\pi\frac{f_1-f_2}{2}t\right) \tag{19}$$

We see that if the frequencies f_1 and f_2 are close to each other then the second term on the right hand side of the equation which depends on $\frac{f_1-f_2}{2}$ will cause a periodic variation in the amplitude of the first term on the right hand side. Although not fully removed the larger period and larger relative amplitude of the periodic signal compared to the movement data makes it a better candidate for removal in Fourier space.

A first attempt at doing this involved manually estimating the period and fitting a series of sine waves with periods in the vicinity of the estimate. From the fit with the smallest residuals the amplitude, phase and frequency were obtained after which this sine function could be subtracted from the data. An example of this can be seen in figure 18, where the same data shown in figure 16 was used.

We can see that there is still a visible periodic component in the data with the sine subtracted. This suggests that our fit was inadequate to fully remove the 9-fold symmetry from the data. This may lie in our assumption that the beat frequency we see is a well defined sine function. If the engine rotates at a slightly varying frequency due to a natural instability a more complex artifact could be present then we would have expected. In an attempt to more accurately model the actual artifact a fit programme was written which would include higher harmonics of the beat frequency found. In figure 19 we again fit and subtract a periodic signal from out data but this time have included the 2nd and 3rd harmonics.

Now we see less periodicity in the end result. Looking at the autocorrelation function of the residual data however a small periodicity still appears to be present. We have tried subtracting a series of higher harmonic signals. This lead to little or no improvement of the residual data. It would appear that more engine stability is the best way of solving this problem.

As the beat frequency is a lot lower than the normal 9-fold periodicity it allows us to gather some

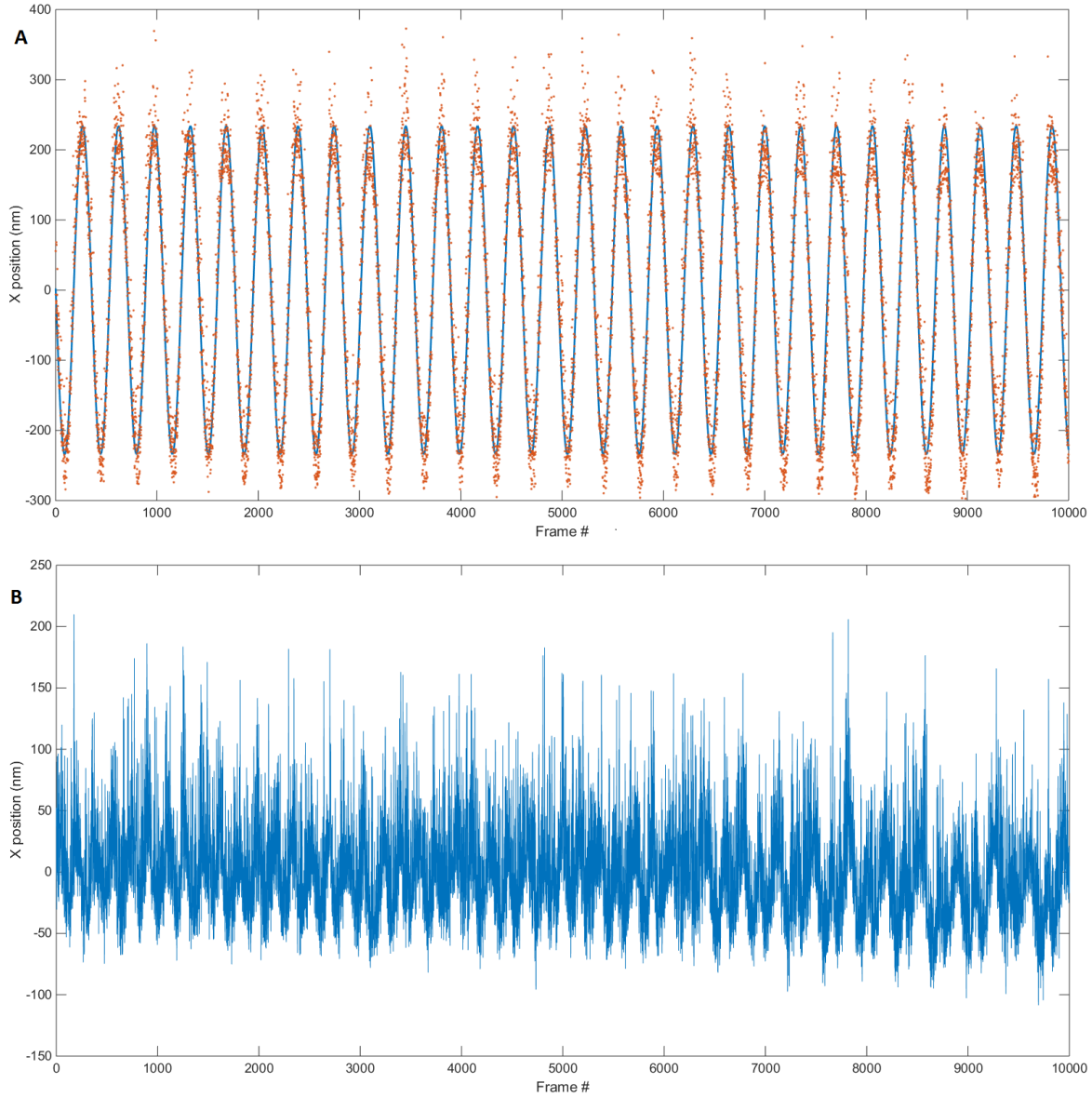

Figure 18: **A**: X position of bead with mean subtracted and best sine fit, **B**: fitted sine subtracted from data

information from the data regardless. We can smooth out the periodic signal by using a simple running average filter. Set at the correct width we can subtract the local mean of the data in order to remove the periodic signal without sacrificing information on the bead movement. This is however a rough and unrefined method and other possibilities of removing or preventing the periodicity in the data would be preferable.

To summarise and show the complexity of the artifact I have simulated some positional data by combining a clean sine function with randomly distributed numbers, to represent movement with a periodic artifact in it and compared its power spectrum to that of our actual data at both 100 rounds per minute and 170 rounds per minute. The result can be seen in figure 20.

We see that in the case of a noise signal added to a clean sine function we can still see one single strong peak. When the CFM is rotating at 100 rounds per minute we see a very strong peak which should be at 9 times the rotational speed around 15 Hz. Due to the Nyquist criterion, which in our case is $\frac{25.59}{2} = 12.795$, which is below the expected frequency the peak is shifted to a lower frequency, as was already determined by a previous student [34]. Aside from this we also see strong peaks at various points in the power spectrum at too high frequencies to be noise. When we rotate with the frame rate and rotation speed in sync we see that the 9-fold peak is almost invisible but we still see groups of smaller peaks throughout the power spectrum in the high and medium frequency range. We also see stronger peaks in the low frequency range which is the beat frequency.

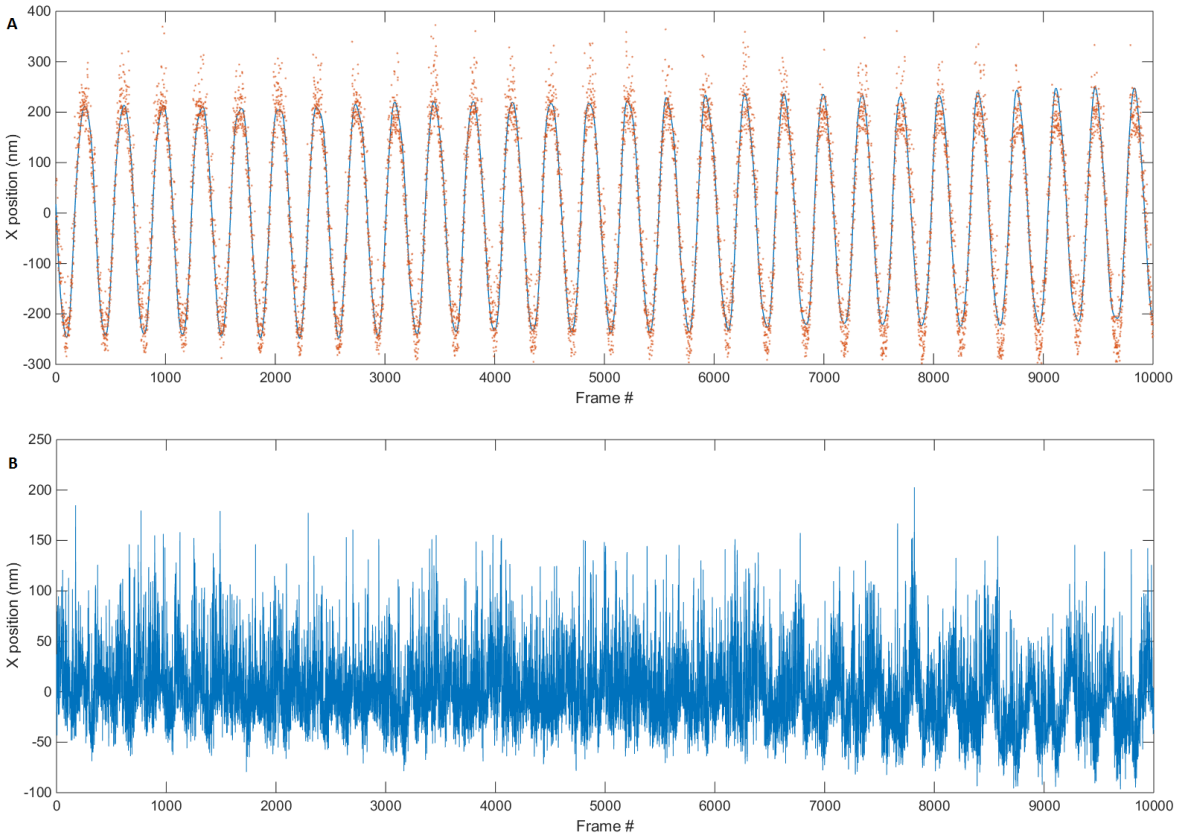

Figure 19: **A**: X position of bead with mean subtracted and best periodic fit including 2nd and 3rd harmonics, **B**: fitted sine subtracted from data

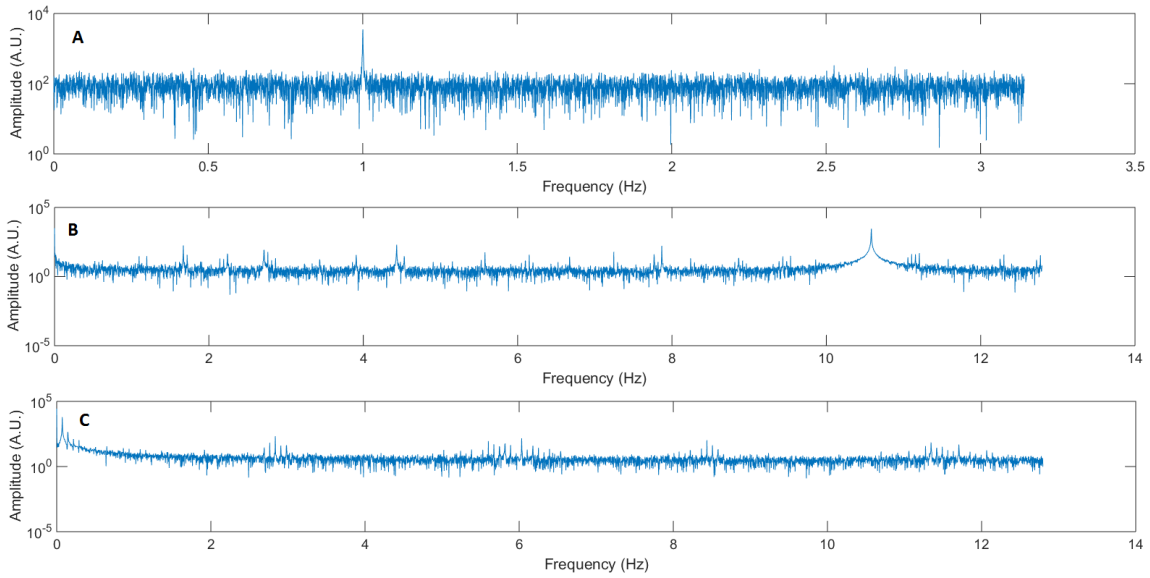

Figure 20: Logarithmic power spectra of **A**: simulated data consisting of a clean sine function of frequency 1 with added noise, **B**: a stuck bead rotating at 100 rounds per minute, **C**: a stuck bead rotating at 170 rounds per minute.

All of this confirms that the engine artifacts present in our set-up are more complicated than was postulated by previous students.

4.3.3 Set-up Accuracy & Noise Estimation

After identifying the inherent problems of our set-up and setting up the modified sample creation protocol, I tried obtaining analysable results. The first step was to determine the accuracy of the set-up, now that we know its inherent problems, and obtaining data on stuck beads without rotation. This was followed by acquisition of data on the two types of tethered samples at zero rounds per minute.

The engine artifacts we discussed are of course not present when the engine is not in use and our microscope is stationary. We can safely derive an estimate of the sub-pixel accuracy we can achieve using data of stuck silica beads. Samples with stuck beads were inserted into the microscope and imaged at 25.59 frames per second. Several tens of regions of interest were selected and positional data obtained for around a hundred thousand frames.

If present, drift was subtracted from the data and the r position calculated from the x and y positions through $r = \sqrt{x^2 + y^2}$. Extreme peaks, caused by a passing loose particle for example, were filtered out and a histogram made of the r positions. A Gaussian curve was then fit to the histogram to obtain the mean movement and standard deviation in r. One would expect a stuck particle or tethered particle to be identifiable through this mean position as a tethered particle would have a much higher mean movement in r. The variance in the x and y positions were also obtained to verify the resulting curve. Also if the bead or tether is stuck in an unusual way or on one side this should be visible by a large difference between the respective variances. A plot of a Gaussian fit to the histogram and the filtered r position of one such stuck stationary bead is shown in figure 21.

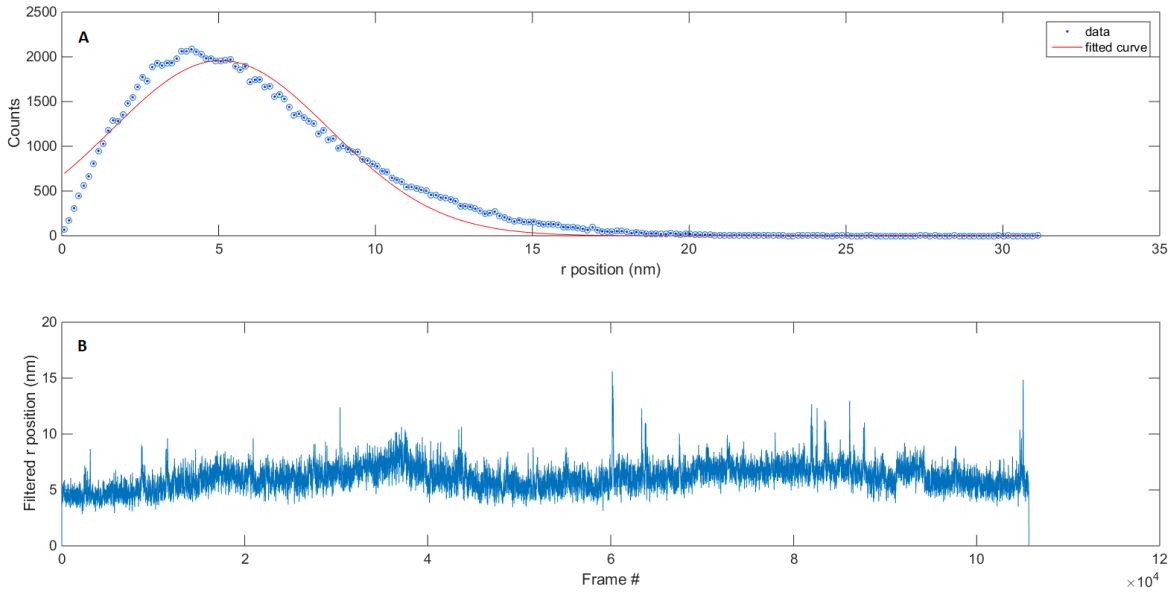

Figure 21: **A**:Gaussian fit of r positional values of a stationary 1 micron silica bead,**B**: filtered r position of a stationary 1 micron silica bead.

The first thing we notice is that the histogram of r values is not symmetrical. I encountered this skew for almost all beads analysed. This makes sense as we are using a Gaussian fit as an approximation. As we are calculating r through the square of the movements in the x and y directions we will never have a Gaussian curve as the mean will never be zero, unless of course there is no movement at all.

To take a more detailed look at the signal in the table below the values obtained from the fit as well as the variances obtained from the raw x and y positional data are shown.

Mean r (nm)	Variance in r (nm)	Variance in X (nm)	Variance in Y (nm)
5.1	6.25	12.8	37.5

We see that the variance for the y position is almost three times higher than that of the x position. A possible explanation is that gravity acts in the y direction and hence influences the position. However for a truly stuck bead the gravitational force, which is 13.64 fN, should not be noticeable in the positional data.

In figure 22 histograms of the separate x and y positions of the bead are shown.

We can see that the y position appears skewed in the negative direction. It also appears to have a non-zero mean. The negative y direction is the direction in which the gravitational force pulls the

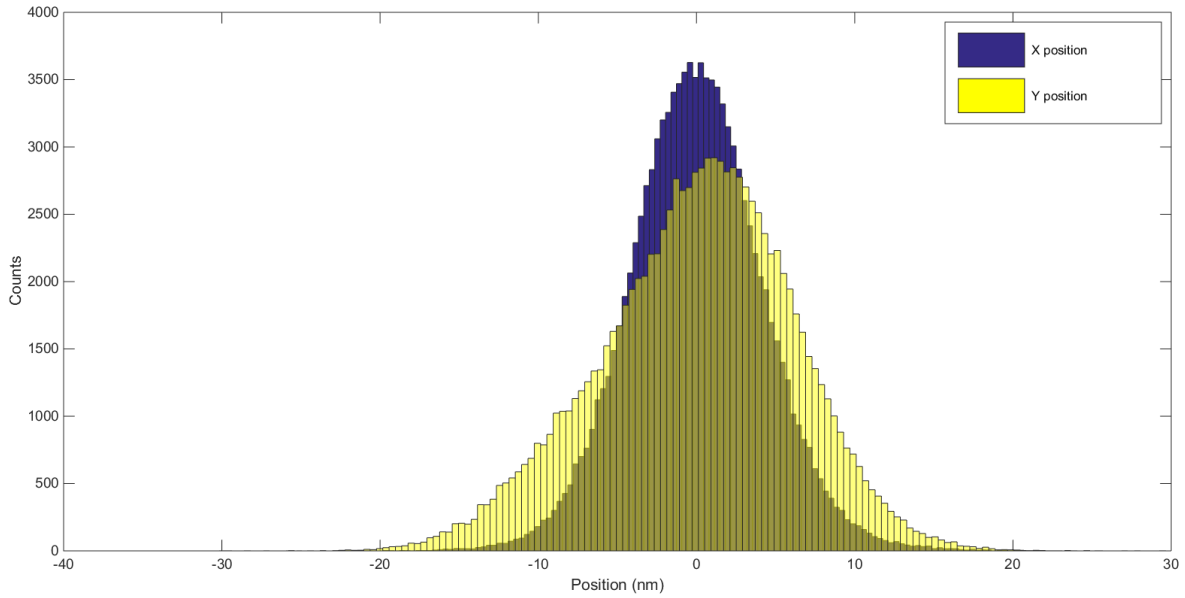

Figure 22: Histograms of x and y positions of stuck stationary silica bead

bead. This could suggest that the bead isn't as stuck as one would expect and the effect of gravity is noticeable in the y position.

Despite this unexpected difference between x and y directional movement an estimate of the sub-pixel accuracy can be made using Parseval's theorem. I summed over the final third of the power spectrum of the x and y positions of 10 stuck beads and multiplied this by three to obtain values for the accuracy of the set-up. We assume that regardless of how the bead is stuck its movement should not be fast enough to be in or close to this region of the power spectrum. The mean of these values gives us a second estimate of how accurate our set-up is. Through this method I obtained values of 7.1 ± 2.3 nm for the x position and 9.5 ± 3.3 nm for the y position. Note that this gives us variances of 5.3 nm and 11.2 nm for the x and y positions respectively, which is lower than what was found from the raw data. Regardless both variances should be adequate as the folding or unfolding of a G-quadruplex we want to visualise decreases or increases the tether length, and hence the movement, by roughly 20 nm.

4.4 Identification of Beads

To determine whether or not a tether was identifiable through its mean movement I obtained the mean r position for several tens of regions of interest of all three samples, stuck, standard DNA tether and G-quadruplex tether. This was done firstly without rotation so as to avoid the problems with the 9-fold periodicity. The mean r position and its standard deviation were obtained with the same Gaussian fit as used in figure 21, the variance in the x and y positions were also obtained in order to check to what degree a Gaussian fit was actually reliable for each ROI. If one or more large jumps in the r position of the region of interest could be identified by hand the mean r position was obtained for every range visible. The results are shown in figure 23. Each ROI was imaged for at least 60000 frames.

For the stuck beads we see that most beads have a mean r position of around 10 nm. This is comparable to the noise level we expect and we can assume these beads are truly stuck. However there are also several anomalies. First of all, jumps were seen in the data of several ROI's. When this was the case multiple mean r positions can be seen in figure 23 for the same ROI. The noise or background did not change so these jumps appear to be movement from the bead. There are also several beads which have a much higher mean r position of around 60-80 nm.

A possible explanation for both is that not all beads are truly stuck to the glass. Our silica beads are coated in the protein streptavidin which could non-specifically bind to a contamination on the glass, or simply the glass itself. Unlike both the tethered samples, the stuck samples are not washed with BSA. The BSA is used in the tethered samples in order to prevent this non-specific binding, from streptavidin or the other proteins, from occurring. This would explain the larger mean movement. In both the tethered samples beads of similar mean r movement can be seen. These samples were washed with BSA making non-specific binding a less likely event, on the other hand the presence of antidigoxigenin and

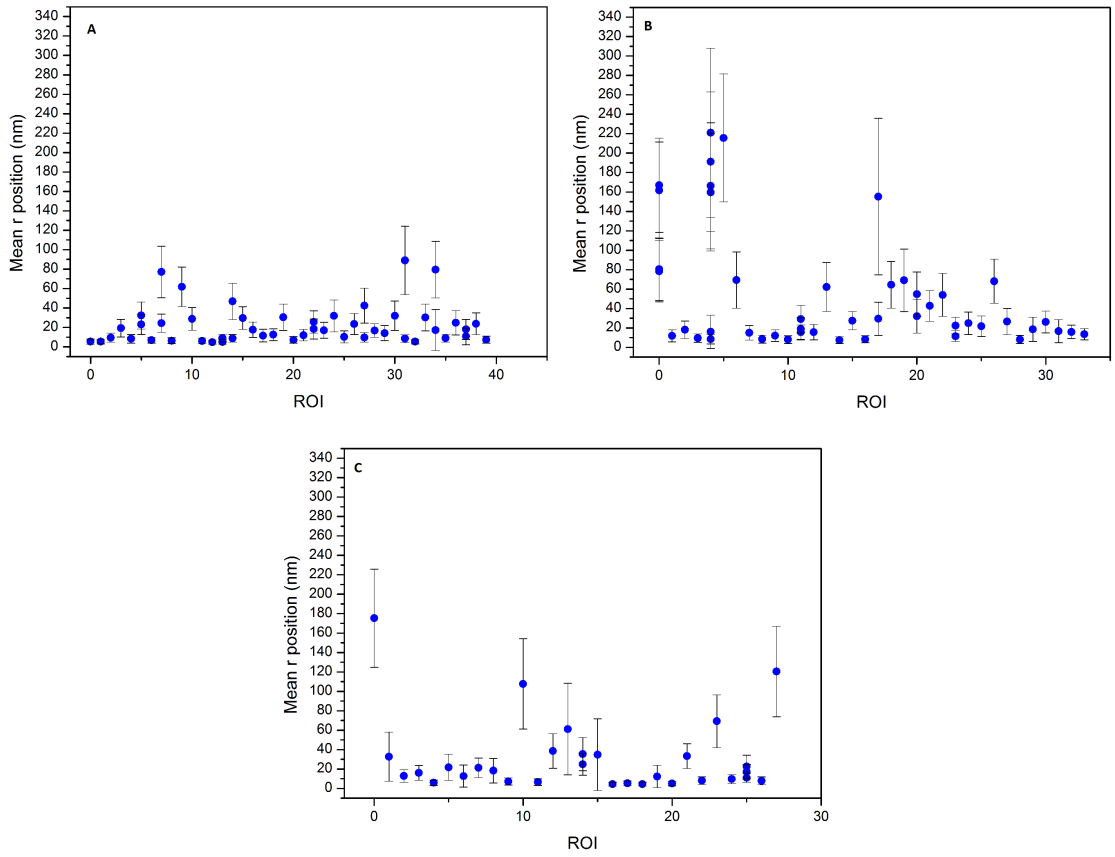

Figure 23: Mean r position of beads **A**: a stuck bead sample,**B**: a standard double helix DNA sample,**C**: a G-quadruplex tethered sample all measured without rotation. Note that in plot **B** one value of 554 ± 203 is also present but left out of range in order to keep the plot readable.

DNA allows for more unwanted binding to occur.

In figure 24 the mean r positions of the stuck beads are shown with the variances of the x and y position for each fitted region. This was done to see whether or not the anomalous beads could be detected through a large difference in variances. If a particle is non-specifically bound it could be in a situation where its movement is strongly restricted in one direction but not so in another.

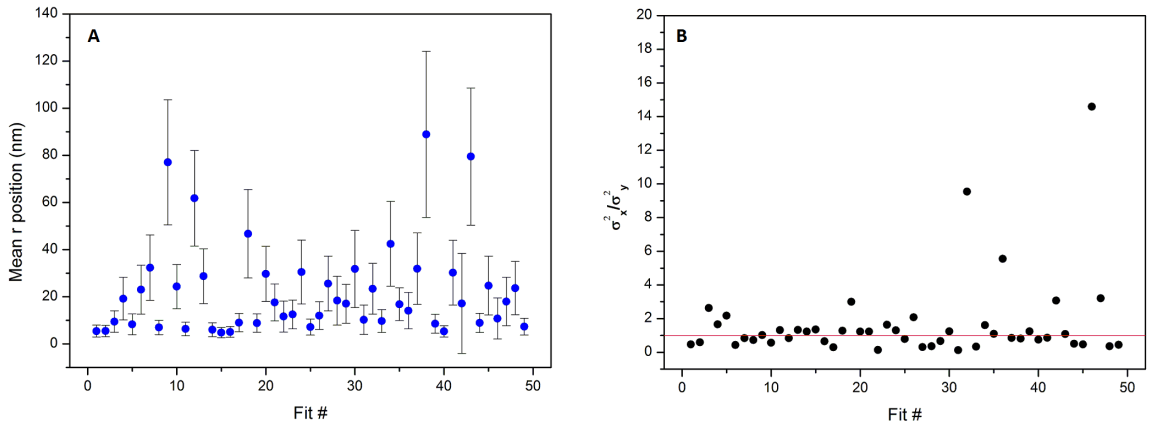

Figure 24: **A**: Mean r position of stuck silica beads without rotation, **B**: variance in x position divided by variance in y position of the regions fitted in **A**, the red line shows unity.

We can see from the figure that the relative variances of x and y provide little guidance. Some higher

mean r values have relative variances close to one whereas some apparently stuck beads show a large difference between the two positions. Any fit which has a relative variance which is far from one should be considered to be at least partly misguiding.

Regardless it would appear that both tethered samples contain a lot of stuck beads, be it non-specifically bound or fully stuck. With a tether length of around 200 nm any mean movement below a 100 nm can not be a properly tethered bead. This is problematic as the samples measured and shown in figure 23 B and C were already the most successful in a long series of attempts. Only around one in three samples is usable and even those appear to have far less tethers than one would like to see. We can see from the images that the tethered samples do have several beads with a significantly higher mean movement, almost reaching the 200 nm mark. These are likely to be tethered particles. These particles also however show more jumps. They were imaged for over 60000 frames (over 40 minutes) though so for shorter measurements less variation can be expected.

This brings us to another problem, that is the lifetime of the tethered samples. Once made using the protocol from appendix A the beads in the sample show movement for only a few hours. After this the overlarge part of the beads is stuck. After a day the samples are unusable. This could be due to Brownian motion or gravity eventually causing the tether to bind to its surroundings thus sticking the bead. No matter the explanation it makes remeasuring a sample over the course of several days impossible. Extensively studying one sample to obtain data on more than 30 or 40 beads is also made difficult by this. Considering the success rate of created samples is low improving the sample creation protocol in order to increase the success rate or lifetime would lead to a substantial improvement in obtaining data on DNA tethers.

4.5 Folding & Unfolding of Tethers

Using the CFM, rotating in resonance with the frame rate, I tried to obtain data on beads, both tethered and untethered, with the CFM rotating. This was done at a frame rate of 25.6 frames per second.

A similar analysis to what was done in section 4.4 was done on all three samples rotating at 170 rounds per minute. At this speed the CFM exerts approximately 95 fN of force on the beads which should be enough to see folding and unfolding of a G-quadruplex occur. Moreover at this frequency we minimise the 9-fold periodicity caused by the engine and only have to deal with the slower beat frequency. Before obtaining the mean r positions of the beads a short time-scale moving average filter was used to remove the beat signal from the data. The results are shown in figure 25.

We look first at the stuck beads. Similar to the stationary beads we see that most beads have a mean r of around 10 nm comparable to the noise level with a few outliers around 50 nm. As this is similar to the stationary beads we can assume that our running average filter is adequate for removing the beat frequency and determining the mean r position of beads at rotation speeds whose 9-fold periodic artifact has the same periodicity as the frame acquisition rate.

Unfortunately both the tethered samples show even less mean r positions of large enough size to suggest the beads in question are tethered than without rotation. For the G-quadruplex sample there is even just one region of interest which has a bead with a mean movement in r of over 100 nm. It is impossible to say anything on the movement of the tethered beads with so few tethers. What we can say is that the uncertainties in the mean movement for the higher values are relatively large. Although we can distinguish between two different mean movements in one region of interest visually, for small jumps of several tens of nanometers the uncertainties of the fit will most certainly overlap. The Gaussian fit, whether or not it holds, might not be the best way to determine when a G-quadruplex folds or unfolds.

At the same frame rate we can also obtain data showing the beat frequency instead of the 9-fold peak when we obtain data at 340 rounds per minute. At this speed the CFM exerts 381 fN of force on the beads in the sample which should be more than enough to see folding and unfolding of a G-qaudruplex. I attempted this first for a sample of beads tethered with DNA without a G-quadruplex. At this speed the CFM shook the table and its protective plexiglas case wildly though. Further obtaining of data at speeds this high was deemed unnecessary until more accurate data was gathered at lower speeds. Note that by tuning the frame rate other speeds in between 170 and 340 rounds per minute can be used whilst still circumventing the 9-fold peak artifact.

For completeness the mean r position of the tethers as obtained at this speed are shown in figure 26. The beat frequency was removed by a short bandwidth running average filter, like with the data obtained at 170 rounds per minute.

The smaller amount of ROI's is due to the fact that a significant portion of beads came unstuck from the glass at this speed. 381 fN of force should not be enough to do this to a tether or a properly stuck bead. But the mechanical stability of the set-up as a whole, as seen by the shaking of the table it

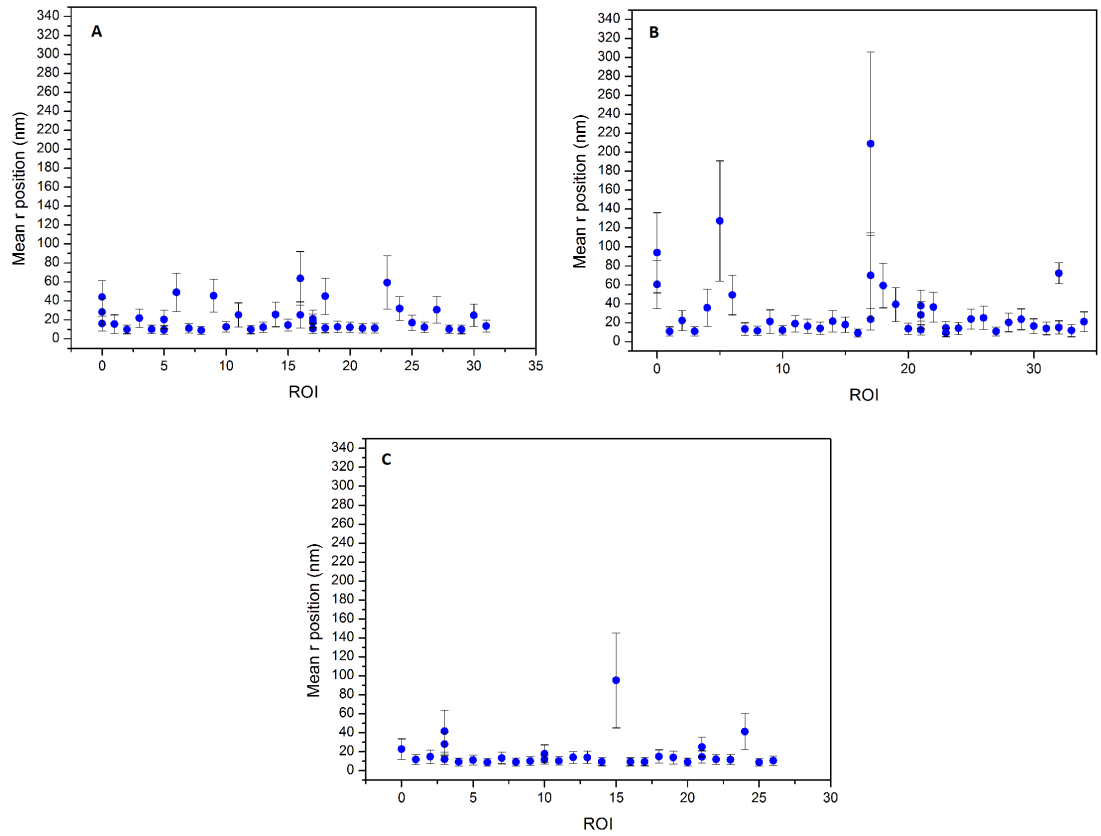

Figure 25: Mean r position of beads **A**: a stuck bead sample,**B**: a standard double helix DNA sample,**C**: a G-quadruplex tethered sample all measured with the CFM rotating at 170 rounds per minute

is on, could have played a part here. Another thing we notice is that unlike the stationary samples or those rotated at 170 RPM, all mean r positions are close to each other.

One explanation for this is that the only beads that were left to image were properly stuck but the high speed increased the noise level of the data thus leading to a higher mean r position, which in the case of stuck beads depends solely on the noise. Another explanation is that at these speeds a simple moving average filter is not enough to accurately remove artifacts thus leading them to have a larger effect on the signal. Regardless no ROI shows a mean r position of larger than a 100 nm, which we would expect for a tethered particle.

Let it be clear that there is by no means enough data at this speed to say anything concrete about the sample. However it is apparent that if we want to measure at these speeds, or higher ones, the set-up should be reinforced to make sure the increased stress on the sample and microscope does not influence the data.

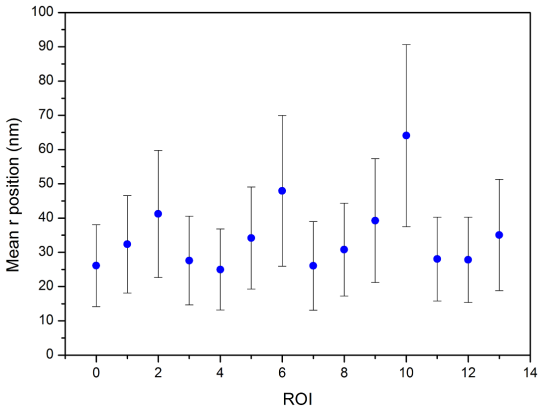

Figure 26: Mean r position of beads in a standard double helix DNA sample, measured with the CFM rotating at 340 rounds per minute.

4.6 Rayleigh Distribution Fits

To confirm whether or not this apparent lack of tethers is a feature actually present in the sample and not a fault in the analysis I first took a look at the fits we used for obtaining the mean movement

and standard deviation. We know that our movement is not described by a Gaussian curve, yet as a reasonable approximation such a fit was used. Our movement, being a dataset of positive-valued positions is much better described by the Rayleigh distribution given by:

$$f(x,\sigma) = \frac{x}{\sigma^2} e^{-x^2/(2\sigma^2)} \tag{20}$$

Unfortunately the realisation that this would be a more accurate fit to obtain information on the movement of our beads came too late to reanalyse all our data. In order to determine whether or not the discrepancies found could be explained wholly through using the wrong fit I have compared the results of a Rayleigh fit to the results obtained with a Gaussian fit. First of all in figure 27 a comparison between a Rayleigh probability distribution and a Gaussian distribution fitted to the r values of a stuck bead is shown. Note that unlike the Gaussian fit in figure 21 **A** the y axis is now normalised and does not represent absolute counts but probability.

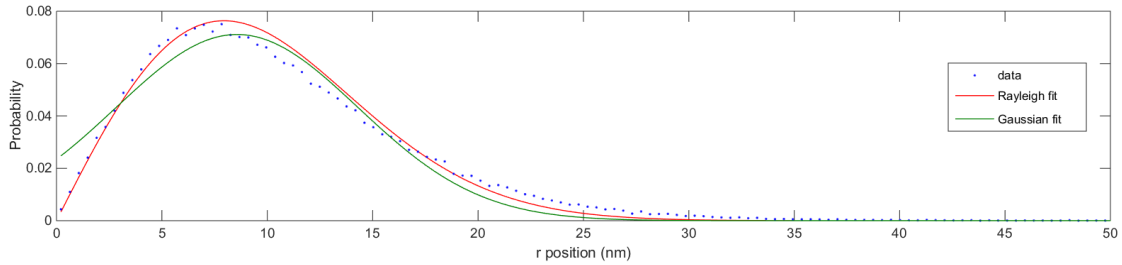

Figure 27: r position of a stuck bead measured with the CFM without rotation with a Rayleigh distribution fit and a Gauss distribution fit.

We can see that the Rayleigh distribution describes our data more accurately. Nonetheless the difference between the peak found by the Gaussian fit and the actual peak in data appears to be of the order of size of several nanometers, hence why we used it as an approximation. The discrepancies found, on the other hand, were often tens of nanometers in size. In figure 28 the difference between the mean movement and standard deviation found by the Rayleigh distribution and Gaussian distribution is plotted for 20 fits of data from a sample with standard double-helix DNA tethers measured by the CFM without rotation.

As the Gaussian fit appeared to misfit the peak slightly higher than its actual location I have plotted the difference in mean position as the maximum of the Gaussian fit minus the maximum as found by the Rayleigh fit. For the standard deviation the same was done to see whether the uncertainty would improve or become worse with the correct fit applied.

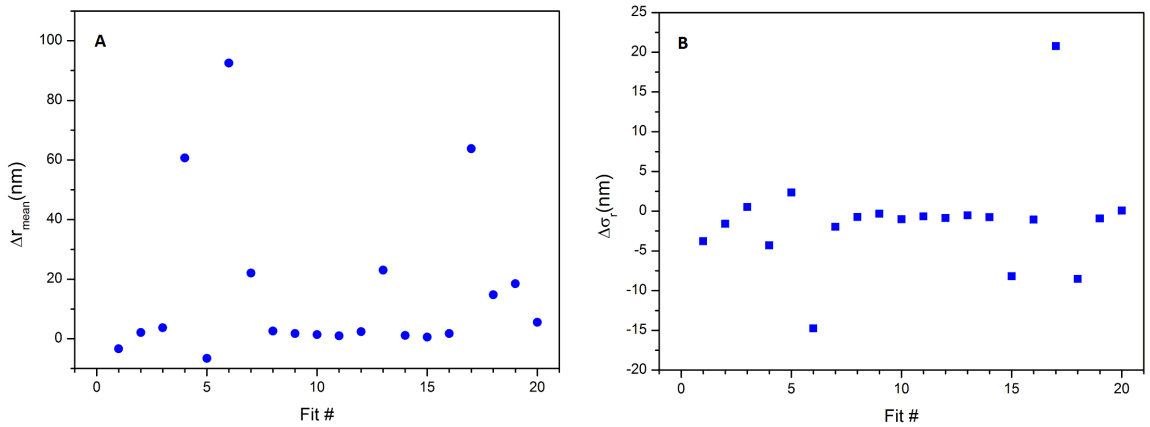

Figure 28: Comparison of fit results on beads in a standard double-helix DNA sample. **A**: Difference in calculated mean r position of bead, **B**: difference in calculated standard deviation of the mean r position of bead. In both cases the result from the Rayleigh fit was subtracted from the result found by the Gaussian fit.

We can see from the figure that the Gaussian fit overestimates the mean r position of the bead compared to the Rayleigh fit. Sometimes as much as several tens of nanometers. The uncertainty in the position, if we look at the difference in standard deviation, mostly increases when using the Rayleigh fit. Most values lie between 0 and 10 nanometer. The larger differences were determined from fits that calculated a higher mean r position. The difference between the estimate of a Rayleigh fit and Gaussian fit therefore seems to increase as the mean r value increases.

This is problematic. With the Gaussian fit we found very few mean r position values of a large enough size to be a possible tether. Now with the correct fit we find that these values should be even smaller, thus the choice of fit does not explain our apparent lack of tethers. Furthermore the increase in standard deviation means that the uncertainties determined in the previous sections should be larger for most datapoints. Nonetheless it is clear that the Rayleigh fit should be used in future research.

4.7 Variance Analysis of Tethers

Another discrepancy worth a closer look is the differences in variances between the x and y positions. A good way to do this is to compare the relative variances of the x and y positions with the apparent tether length. This apparent tether length needs to be defined though. From simple geometry we can calculated that the maximum r positions we can observe in the x,y-plane is $\sqrt{l^2 + 2Rl}$ where l is the length of the tether and R the radius of the bead.

Due to the forces acting on the bead and the fact that DNA strands are not rigid and will therefore never be completely stretched this is not a good estimate for our expected r position. In figure 29 a plot is shown from a paper by Nelson et al. from 2006 where the mean r position of a DNA tethered silica bead is calculated for various tether lengths.

Note that although the figure says that the x axis is in units of base pairs, judging from the text of the paper and other plots we expect this to be a mistake and the x axis values represent tether length in nanometers.

The measurements in figure 29 were done with beads with a diameter of 480 nm. Hence we can not directly use the values from the plot to calculate our apparent tether length. This is not a great problem as, when the bead is of sufficiently larger size than the tether, which in our case it is, the mean movement increases linearly with bead size [38]. With our beads having a diameter of 1 micron we can scale the curve in figure 29 with a factor 2 to estimate our apparent tether length from the mean position obtained.

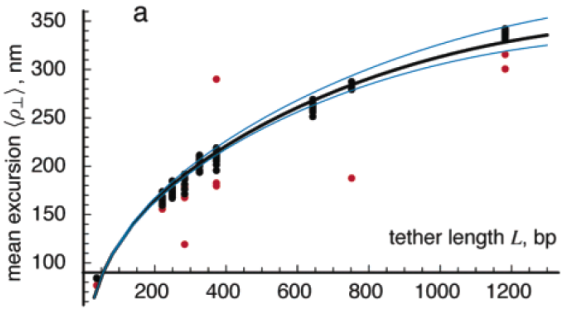

Figure 29: Expected mean excursion of a 240 nm radius tethered bead plotted as a function of tether length. Dots represent measurements and the line a fit curve [37].

We can see if we scale the plot that we can disregard any mean movement below 80 nm as this would correspond to a tether length of smaller than around 40 nm. As can be seen in figure 10 our standard double-helix DNA consists of around 450 base pairs. 1 nm is approximately 3 base pairs so this gives us a tether length of 150 nm. For the G-quadruplex DNA we have two double stranded handles of 412 base pairs combined and a single stranded quadruplex of 86 base pairs. For the single stranded DNA 1 base pair corresponds to 1 nm so this gives us a tether length of around 220 nm. From figure 29 scaled by a factor 2 we can see that we should expect a mean r position of 250 and 350 nm respectively for the two tether types.

To make sure this is a correct assumption we use the following equation [38]

$$\frac{\langle \mathbf{r}^2 \rangle}{L\xi/3} = 2 + \frac{4N_r}{\sqrt{\pi}\mathrm{erf}(N_r)} \tag{21}$$

Where $\langle \mathbf{r}^2 \rangle$ is the mean of the square of the r position, L is the length of the tether in nanometers, ξ is the persistence length in nanometers and N_r is a dimensionless number given by $R/\sqrt{L\xi/3}$ with R the radius of the bead. If we insert our tether lengths and a persistence length of 50 nm, which we take to be a reasonable estimate based on the same paper where we obtained the equation, we can obtain another estimate, this time the root mean square of the r position. This gives us a root mean squared r position of 248 nm for the standard double-helix DNA and 275 nm for the tethers with a G-quadruplex. It would

appear the scaling of figure 29 is not entirely correct. A difference of 75 nm can not be explained by the small difference between the RMS value of the r position and mean r position. For this reason, and the uncertainty over whether the x axis has the correct units, I have opted to calculate the apparent tether lengths using equation 21 instead.

Looking at the data plotted in figures 23 and 25, very few of the beads measured in the G-quadruplex samples reach a mean r position of over 100 nm and the same applies to the beads in the standard double-helix DNA tethered samples. Although these positions were obtained with Gaussian fits, which were incorrect, we saw in section 4.6 that using the correct Rayleigh fit almost always decreases the mean position. Taking the previously discovered oddities in the difference between the variance of the x and y position it would be interesting to look at the apparent length of our possible tethers plotted against the relative variances of their x and y positions. Ideally we would like to see that tethers with apparent lengths close to what we would expect to see of a good DNA tether have little or no difference between the x and y variances whereas tethers which appear to short or too long have a noticeable relative difference. In figure 30 the apparent tether length of all mean r positions higher than 80 nm has been extrapolated and the relative variances of the ROIs plotted, as well as the calculated curve of equation 21.

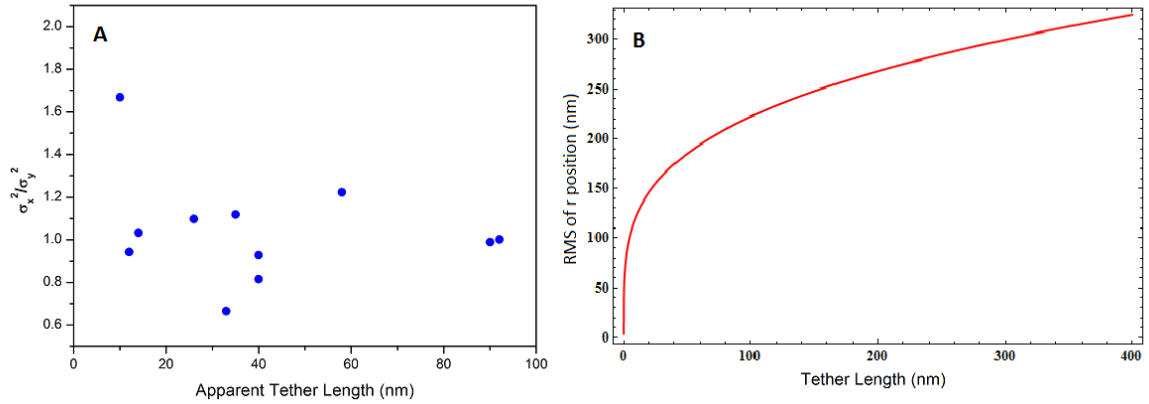

Figure 30: **A**: Relative variance plotted against the apparent tether length of all beads with a mean r position higher than 80 nm, **B**: Calculated curve showing the relationship between tether length and the root mean squared of the r position.

Unfortunately even the beads found with higher mean r movement are nowhere near the tether length we would expect. Although we could be off due to the difference between the RMS of r and the mean of r, this could never cause a difference of 150 nm or more. We see no beads around the 250 to 275 nm range. There also appears to be no strong correlation with the relative variance of the beads. This suggests that not a single one of the beads analysed was a good DNA tether. This leads us back to the problems with sample creation which appear to be even larger than was already suspected. Note that the values used here are the ones fitted with a Gaussian curve. We should expect the actual apparent tether lengths of these beads to be even lower than they are.

5 Conclusion

In this thesis an attempt was made to use a Centrifugal Force Microscope, as a cheaper more efficient alternative to current single particle measuring methods, to analyse the dynamics of G-quadruplex DNA. Work done by a previous student found that the set-up experienced several problems creating artifacts in the data [34]. During this research these problems were found to be more severe than was initially thought. Because of this the acquisition of data on G-quadruplex tethers was greatly delayed. Attempts were made to remove the artifact by finding and removing the frequency in the power spectrum of the data, this proved laborious and inefficient. A workaround for the artifacts was found by synchronising the rotation speed of the microscope with frame acquisition rate of the camera, in order to replace the artifact with a beat frequency. This proved more succesful. At higher rotation speeds exceeding 300 rounds per minute, general instability of the set-up caused strong vibrations which made this method unusable however.

Although eventually data could be obtained on tethered samples, these samples proved to have been created unsuccessfully, showing very few actual tethers. Furthermore oddities were found in the distribution of x and y positions of the beads studied. The most likely explanation for this is the occurence of various unwanted non-specific bindings in the created samples, emphasising that the sample creation process needs to be vastly improved in order for the set-up to be efficient. Nonetheless standard double-helix DNA was made in order to be used as a control sample, which is still useful for future research.

In spite of this I managed to properly characterise the set-up finding the sub-pixel accuracy to be good enough to gather information on kinetics of DNA molecules, being in the 10 nm range. Furthermore even with artifacts present, with an arm length of around 30 cm, we still obtain positional information accurate to 400 nm. Hence the microscope is still a viable machine for future studies on single particles.

6 Outlook

For future research the first solution one would think of is to replace the rotary stage engine with one which does not produce artifacts. In this thesis the same rotary stage was used as in the original experiment by Halvorsen & Wong [39], yet they make no mention of any problems with artifacts. This can probably be explained by the fact that they used much larger beads in their experiments, with 2.8 μm diameters, thus minimising the effects of such artifacts. A replacement engine could however still be an option. Overall mechanical stability is still a major issue. The table and the set-up is on, the case it is in and other tertiary materials needs to be further stabilised in order to allow for high speed measurements. The engine can technically achieve 600 rounds per minute but the vibrations the CFM causes at these speeds make any data acuired hard to use. As one of the major advantages of the CFM is the range of forces it can apply in one instrument.

Secondly the sample creation is key for experiments using the CFM and has proven to be unreliable with the current protocol. In contact with other research groups with a similar set-up, it was found they only had a one in three success rate creating samples. Our success rate was even lower with dozens of created samples producing not a single proper tether. It would be crucial to further research to either create a new considerably more succesful protocol, or find a different way of making samples altogether. The CFM could be modified, to accommodate a different sample, less prone to leaks, breaking and air bubbles. Another option, as was mentioned in section 4.2, to minimise air bubbles and other mechanical defects, could be to use a drip system. In this case inconsistencies caused by the manual factor of using a pipette would be removed.

These options would only increase the success rate slightly. The amount of samples created which are unusable due to mechanical defects is relatively small compared to samples which are structurally fine, but show no tethers or other discrepancies regardless. The sample creation protocol would have to be improved biochemically as well. One way to do this would be to return to the protocol used by previous students, only keeping the modifications which improved the mechanic aspect of the samples [33]. There is unfortunately no strong evidence to suggest one particular chemical or aspect of the protocol causes the lack of tethers. Throughout this research I made sure that the chemicals used were fresh, ordering new batches through the group when necessary and creating new buffer on a regular basis. The G-quadruplex DNA used was a few years old but kept frozen properly until use. Also the DNA made during this research showed a lack of tethers as well, suggesting that the age was not an issue.

Finally, until the CFM is made more stable, for gathering information on G-quadruplex folding and unfolding, existent methods such as optical traps and atomic force microscopes are alternatives.

These methods are still laborious and expensive, which was the motivation for the creation of the CFM, however.

7 Acknowledgements

First and foremost I would like to thank my direct supervisor Dr. Gerhard Blab. Through nearly two years of research he provided a great amount of time and effort to make this thesis possible. He was crucial in keeping spirits up through less succesful periods and aside from teaching me many things such as programming, chemical work and other research skills which I did not yet possess, was always up for coffee and conversation on more leisurely subjects. I would also like to thank Prof. Dr. Hans Gerritsen for allowing me to work in his research group and giving key advice at times. I would furthermore like to express my gratitude to ing. Dave van den Heuvel, who provided the necessary chemicals, helped work on the sample creation protocol and did all the HF etching to order. Dr. Arnout Imhof also deserves a mention for taking the much appreciated time to be a second reviewer of this thesis.

Lastly I would like to thank all of the members of the biophysics group for their good company during coffee breaks, tolerating my occasional lack of professionalism and teaching me more about life than one would expect to learn during a physics master's thesis. With a special mention for Johannes and our ongoing table tennis rivalry.

8 Bibliography

[1] Ken Halvorsen and Wesley P Wong. Massively parallel single-molecule manipulation using centrifugal force. *Biophysical Journal*, 98(11):L53–L55, 2010.

[2] James D Watson, Francis HC Crick, et al. Molecular structure of nucleic acids. *Nature*, 171(4356):737–738, 1953.

[3] Rosalind E Franklin and Raymond G Gosling. Evidence for 2-chain helix in crystalline structure of sodium deoxyribonucleate. *Nature*, 172:156–157, 1953.

[4] Robert D Wells and Stephen C Harvey. *Unusual DNA structures*. Springer New York etc, 1988.

[5] Andrew N Lane, J Brad Chaires, Robert D Gray, and John O Trent. Stability and kinetics of G-quadruplex structures. *Nucleic acids research*, 36(17):5482–5515, 2008.

[6] Stephen Neidle. Human telomeric G-quadruplex: The current status of telomeric G-quadruplexes as therapeutic targets in human cancer. *FEBS Journal*, 277(5):1118–1125, 2010.

[7] Shankar Balasubramanian, Laurence H Hurley, and Stephen Neidle. Targeting G-quadruplexes in gene promoters: a novel anticancer strategy? *Nature reviews Drug discovery*, 10(4):261–275, 2011.

[8] Giulia Biffi, Marco Di Antonio, David Tannahill, and Shankar Balasubramanian. Visualization and selective chemical targeting of RNA G-quadruplex structures in the cytoplasm of human cells. *Nature chemistry*, 6(1):75–80, 2014.

[9] Giulia Biffi, David Tannahill, John McCafferty, and Shankar Balasubramanian. Quantitative visualization of DNA G-quadruplex structures in human cells. *Nature chemistry*, 5(3):182–186, 2013.

[10] Susanna Lynch, Heather Baker, Sarah G Byker, Dejian Zhou, and Kumar Sinniah. Single molecule force spectroscopy on G-quadruplex DNA. *Chemistry–A European Journal*, 15(33):8113–8116, 2009.

[11] Scott L Forman, James C Fettinger, Silvia Pieraccini, Giovanni Gottarelli, and Jeffery T Davis. Toward artificial ion channels: a lipophilic G-quadruplex. *Journal of the American Chemical Society*, 122(17):4060–4067, 2000.

[12] Sarah Burge, Gary N Parkinson, Pascale Hazel, Alan K Todd, and Stephen Neidle. Quadruplex DNA: sequence, topology and structure. *Nucleic acids research*, 34(19):5402–5415, 2006.

[13] James D Watson and Francis HC Crick. The structure of DNA. In *Cold Spring Harbor symposia on quantitative biology*, volume 18, pages 123–131. Cold Spring Harbor Laboratory Press, 1953.

[14] Julian Leon Huppert. Four-stranded nucleic acids: structure, function and targeting of G-quadruplexes. *Chemical Society Reviews*, 37(7):1375–1384, 2008.

[15] Shozeb M Haider, Gary N Parkinson, and Stephen Neidle. Structure of a G-quadruplex–ligand complex. *Journal of molecular biology*, 326(1):117–125, 2003.

[16] Harvey Lodish, Arnold Berk, S Lawrence Zipursky, Paul Matsudaira, David Baltimore, James Darnell, et al. *Molecular cell biology*, volume 4. WH Freeman New York, 2000.

[17] Liana Oganesian and Jan Karlseder. Telomeric armor: the layers of end protection. *Journal of cell science*, 122(22):4013–4025, 2009.

[18] Christiane Schaffitzel, Imre Berger, Jan Postberg, Jozef Hanes, Hans J Lipps, and Andreas Plückthun. In vitro generated antibodies specific for telomeric guanine-quadruplex DNA react with stylonychia lemnae macronuclei. *Proceedings of the National Academy of Sciences*, 98(15):8572–8577, 2001.

[19] Kary B Mullis, Henry A Erlich, Norman Arnheim, Glenn T Horn, Randall K Saiki, and Stephen J Scharf. One of the first polymerase chain reaction (pcr) patents, July 28 1987. US Patent 4,683,195.

[20] Gerald Schochetman, Chin-Yih Ou, and Wanda K Jones. Polymerase chain reaction. *The Journal of infectious diseases*, 158(6):1154–1157, 1988.

[21] Elias M Stein and Guido L Weiss. *Introduction to Fourier analysis on Euclidean spaces*, volume 1. Princeton university press, 1971.

[22] Richard G Lyons. *Understanding digital signal processing*. Pearson Education, 2010.

[23] Lord Rayleigh. Xv. on the theory of optical images, with special reference to the microscope. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 42(255):167–195, 1896.

[24] Richard D Keane and Ronald J Adrian. Theory of cross-correlation analysis of piv images. *Applied scientific research*, 49(3):191–215, 1992.

[25] Mark R Banham and Aggelos K Katsaggelos. Digital image restoration. *Signal Processing Magazine, IEEE*, 14(2):24–41, 1997.

[26] Maria Spies, Mark S Dillingham, and Stephen C Kowalczykowski. DNA helicases, 2002.

[27] Jolien Marsman. Sample preparation to study the effects of tension on the dynamical behavior of the G-quadruplex DNA. Bachelor's thesis, Utrecht University, 2012.

[28] S. Deelen. Technical drawings of centrifugal force microscope. *figures 2,3*, 2014.

[29] Yih-Fan Chen, JN Milstein, and Jens-Christian Meiners. Femtonewton entropic forces can control the formation of protein-mediated DNA loops. *Physical review letters*, 104(4):048301, 2010.

[30] William M Haynes. *CRC handbook of chemistry and physics*. CRC press, 2014.

[31] Paul Langevin. *On the theory of Brownian motion*. 1908.

[32] Yih-Fan Chen, Gerhard A Blab, and Jens-Christian Meiners. Stretching submicron biomolecules with constant-force axial optical tweezers. *Biophysical Journal*, 96(11):4701–4708, 2009.

[33] Rosalinde Pots. Optimization of sample preparation and first measurements in a study on the dynamics of G-quadruplex DNA. Bachelor's thesis, Utrecht University, 2013.

[34] Alexandru Pavel Nechifor. Optimization of the mechanical stability of a centrifugal force microscope. Bachelor's thesis, Utrecht University, 2014.

[35] Richard C Dorf and Robert H Bishop. Modern control systems. 1998.

[36] National Instruments. *NI LabWindows/CVI manual*, 2010.

[37] Philip C Nelson, Chiara Zurla, Doriano Brogioli, John F Beausang, Laura Finzi, and David Dunlap. Tethered particle motion as a diagnostic of DNA tether length. *The Journal of Physical Chemistry B*, 110(34):17260–17267, 2006.

[38] Darren E Segall, Philip C Nelson, and Rob Phillips. Volume-exclusion effects in tethered-particle experiments: bead size matters. *Physical review letters*, 96(8):088306, 2006.

[39] Supplemental material, massively parallel single-molecule manipulation using centrifugal force. `http://www.cell.com/cms/attachment/2024918574/2044649422/mmc1.pdf`. Accessed: 31-05-2016.

A Sample preparation protocol

Time	Operation	Comment
0:00	Wash microscope glass with EtOH followed by H_2O and blow dry with N_2	
0:05	Cut parafilm with template, stick between slide and etched coverslip and heat until sealed	
0:15	Fill sample with phosphate buffer	
0:20	Mix 23 μL phosphate buffer with 7 μL antidigoxigenin (0.01 μg/ml) in eppy	
0:25	Add 20 μL antidig solution, and wait for 20 minutes	
0:47	Wash with PB2 (1x75 μL) then add 75 μL PB2	
0:49	Wait for 30 minutes in meantime mix 23 μL PB2 with 5 μL beads in eppy	
0:52	Add 2 μL DNA and homogenise before using	
1:19	Wash sample with PB2 (2x75 μL) then add 30 μL DNA/bead solution	
1:25	Wait 3 min in meantime grab 2-component acrylic glue	
1:28	Wash with PB2 (2x75 μL) close endings with 2-component acrylic glue	

B Microscope initialisation protocol

- Start PC
- Turn set-up power on
- Start Vimba Viewer to check if camera connection works
- If not, restart PC, leaving set-up on
- Start Vimba Viewer, camera should work now
- Close Vimba Viewer, open LabView AcquireData
- Turn on continuous update
- Check whether piezo control works, if not restart LabView
- Put an ROI over a bead
- Turn engine on, if dynamic range changes, restore engine to starting position and/or restart LabView until this no longer happens
- Begin measurements