CHALMERS
UNIVERSITY OF TECHNOLOGY

A Rolling Circle Amplification-Based Methodology for Making Long, Sequence-Repeating, DNA Duplexes

Master's thesis in Biology and Biological Engineering

NORA DEKONING

Department of Biology and Biological Engineering
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2018

Master's Thesis 2018

A Rolling Circle Amplification-Based Methodology for Making Long, Sequence-Repeating, DNA Duplexes

Nora Dekoning

CHALMERS
UNIVERSITY OF TECHNOLOGY

Department of Biology and Biological Engineering
Division of Chemical Biology and Division of Chemistry and Biochemistry
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2018

A Rolling Circle Amplification-Based Methodology for Making Long, Sequence-Repeating, DNA Duplexes
NORA DEKONING

Examiner: Fredrik Westerlund
Division of Chemical Biology
Department of Biology and Biological Engineering

Co-Promotor: Marcus Wilhelmsson
Division of Chemistry and Biochemistry
Department of Chemical and Chemical Engineering

Supervisors: Robin Öz and Jesper Nilsson

Department of Biology and Biological Engineering
Division of Chemical Biology
Chalmers University of Technology
SE-412 96 Gothenburg
Telephone: +46 31 772 1000

'One of the beautiful things about science
is that it allows us to bumble along,
getting it wrong time after time,
and feel perfectly fine
as long as we learn something each time'

- Martin A. Schwartz

Acknowledgements

This chapter is dedicated to the people that have helped me directly or indirectly during the five months that I was working on this project.

First, I want to sincerely thank my promotors **Fredrik Westerlund** and **Marcus Wilhelmsson**. Thank you for your blind leap of faith in this random Belgian girl, because of it I have had the opportunity to work on such an interesting and challenging project. It has been my absolute pleasure to work with the both of you.

My supervisors, **Jesper** and **Robin**, for teaching me so many new things, for putting my questions before their own work. Both of your never ending optimism kept on surprising me and always made me feel better in the lab days that were less successful. I could not have wished for any better supervisors! Also, thank you so much for the proofreading and honest feedback!

To **Mats Nilsson and his group**, who provided the RCA protocol and follow-up information.

The **Westerlund Group** for making me feel welcome from day one. Spending my days in the microscopy lab, while you accepted my music choices, made life great! **Vandana, My, Rick, Ville, Robin, Elina, Kai, Sriram, Eka and Sune**, thank you for all the laughs, the help, the advice and the lovely group activity.

The **Wilhelmsson Group** for inviting me to all discussions, meetings, out-door activities and more. **Moa, Anders, Jesper, Tom, Pauline, Sangamesh and Mattias**, the ice-skating, lunches and the division conference are a couple of the activities that made my Erasmus memorable!

In particular: To **Pauline**, you invited me to activities from day one. The first week was absolutely amazing! Not a single moment of homesickness, because you kept me so busy. Thank you for being such a perfect addition to my Erasmus experience! You will always be welcome in my home, wherever I am. To **My**, for letting me invite myself to your activities like the GBG beerweek. It was an absolute honour to get to know your amazing music taste, but not only that, it was an honour to get to know you. Don't change, please! To **Gerard**, for being the cutest Catalan around. Okay, if I am honest: I only like you because of your cats. No, just kidding. I am so happy that I decided to be social that one day in the lab, everything got better and better after that moment. PS: I am holding you to that tour in Barcelona! To mama **Dasha**, the lively and loving way that you move through life is so catchy and beautiful to behold. Thank you for being my mama abroad and feeding me on multiple occasions!

To my friends back home, you all have been a great support. I have missed you! To **Marie**, for always checking up on me and making me feel missed. To **Hannah**, for answering all my questions about thesis writing, presentations and just for being your chaotic self. To **Kemp**, for distracting me when I needed it and for being stimulated to go on Erasmus yourself. I am so proud! To **Rani**, through all the ups and downs I know that you will be there for me and I will be there for you. To **Robin**, you have the power to always make me laugh. Our conversations can go from serious to absolute nonsense to serious again. The simplicity of it all is something that I will never take for granted.

Last, but most definitely not least, I want to thank **my mom** for always being there for me. You have raised me to be independent, to land on my two feet, and I think that it is working out just fine. PS: Yes, mom, I ate a lot of bananas during this last month!

Abstract

In vitro studies of sequence-specific DNA-protein interactions using techniques where long DNA molecules is needed are currently limited by the size of synthesized DNA molecules, or is restricted to the commercially available DNA. For single molecule DNA imaging techniques using for instance nanochannels confinement, the small sized molecules make the interactions difficult to detect. The commercially available DNA, on the other hand, does not allow any freedom in choosing or changing the DNA sequence. Therefore, a method for producing long DNA molecules containing a sequence of choice would alleviate these limitations and greatly improve the possibility to study DNA-protein interactions.

The general concept of this paper was to create long, double-stranded DNA molecules with a sequence that is specifically designed to interact with the protein internal host factor. To make such molecules, three experimental procedures were developed, building on the single-stranded DNA products from a rolling circle amplification (RCA) reaction. The first experimental procedure was based on the presumed perfect annealing of smaller complementary strands to the single-stranded RCA product. The second experiment assumed that single-stranded gaps would be present in the duplex after annealing, hence the addition of a polymerisation reaction. In the third experiment, the RCA was run for 24 hours, to allow double-stranded product to be formed in the RCA.

The DNA strands were visualized using fluorescence microscopy, with the goal of studying them and their protein interactions in nanochannels. To be able to use this detection method, an external fluorescent dye, YOYO, is used in the main aim of the project. However, as these types of dyes change the native structure of DNA, an extra aim was to use the fluorescent base analogue tC incorporated into one of the duplex strands, leaving the native structure intact.

All three experimental procedures were shown to be capable of producing apparently double-stranded DNA molecules, that were larger than 100 kilo-base pairs albeit with a broad size distribution. This shows that the main aim in terms of procuing DNA molecules has been completed. The tC-containing DNA molecules were not visible under the microscope with the settings used, but appears to be promising.

Table of Content

List of Symbols, Glossary and Abbreviations

NT: NUCLEOTIDES

KBP: KILO BASE PAIR

DSDNA: DOUBLE-STRANDED DNA

IHF: INTEGRATION HOST FACTOR

RCA: ROLLING CIRCLE AMPLIFICATION

PCR: POLYMERASE CHAIN REACTION

TC: TRICYCLIC CYTOSINE

DNA: DEOXYRIBONUCLEIC ACID

A: ADENINE

G: GUANINE

C: CYTOSINE

T: THYMINE

SSDNA: SINGLE-STRANDED DNA

DNTP'S: DEOXYNUCLEOSIDE TRIPHOSPHATES

Φ29 POL: Φ29 DNA POLYMERASE

RCPS: ROLLING CIRCLE PRODUCTS

NT: NUCLEOTIDES

PLP: PADLOCK PROBE

W: DA OR DT

R: DA OR DG

N: ANY NUCLEOTIDE

I_0: INCOMING LIGHT INTENSITY

I: LIGHT INTENSITY AFTER PASSING THE SAMPLE

A: ABSORBANCE

ε_λ: MOLAR ABSORPTIVITY AT A CERTAIN WAVELENGTH Λ

S_0: GROUND STATE

S_1: FIRST ELECTRONIC STATE

S_2: SECOND ELECTRONIC STATE

VR: VIBRATIONAL RELAXATION

IC: INTERNAL CONVERSION

FRET: FÖRSTER RESONANCE ENERGY TRANSFER

DUTP: DEOXYURACIL TRIPHOSPHATE

YO: OXAZOLE YELLOW

YOYO: 1,1'-(4,4,8,8-TETRAMETHYL4,8-DIAZAUNDECAMETHYLENE)-BIS[4-[(3-METHYLBENZO-1,3-OXAZOL-2-YL)-METHYLIDENE]-1,4-DOHYDROQUINOLINIUM] TETRAIODIDE

EMCCD: ELECTRON MULTIPLYING CHARGED COUPLED DEVICE

TIRF MICROSCOPY: TOTAL INTERNAL REFLECTION FLUORESCENCE MICROSCOPY

P3: PRIMER 3

P4: PRIMER 4

P31: PRIMER 31

P32: PRIMER 32

TRIS-HCL: TRIZMA HYDROCHLORIDE

DTT: 1,4-DITHIOTHREITOL

BSA: BOVINE SERUM ALBUMIN
ATP: ADENOSINE TRIPHOSPHATE
FITC: FLUORESCEIN ISOTHIOCYANATE
CY5: CYANINE 5
ROS: REACTIVE OXYGEN SPECIES
APTES: (3-AMINOPROPYL)TRIETHOXYSILANE
ATMS: ALLYLTRIMETHOXYSILANE
BME: 2-MERCAPTOETHANOL
BP: BASE PAIRS

Figures and Tables in the Text

1. Introduction

Protein-protein and protein-DNA interactions play a central role in virtually all biological functions. These interactions are therefore also critical in the mechanisms causing diseases, as it typically is the mutations in the binding interfaces of proteins that are responsible for perturbed function[1]. By trying to understand these protein interactions *in vitro*, it is possible to develop methods for treatment, diagnosis and even the prevention of those kind of diseases[2,3].

Solid-phase oligonucleotide synthesis is typically used for synthesizing specific oligonucleotide sequences to study DNA-protein interactions. This methodology has the advantage of being a rapid, fully automated and inexpensive way to make custom-made sequences[4]. The disadvantage is that the practical length of these oligonucleotides is limited to *ca*. 200 nucleotides (nt), but typically even fewer, as overall yields and purification becomes challenging. Although these smaller strands can be joined together to form longer products[4], the studied DNA-protein interactions will be limited to the length and the purity of the initially synthesized DNA strands. Many techniques used to study oligonucleotides and their interactions, *e.g.* nanochannels, require long, uniform, double-stranded DNA molecules to be able to detect the interactions. A fast and efficient way of producing such DNA, containing user-defined sequences, would benefit the field significantly.

In the first part of this project it was tried to make long (kbp) double-stranded DNA (dsDNA) molecules with a specific sequence known to interact with the integration host factor (IHF), a histone-like protein[5]. The method of choice was rolling circle amplification (RCA). RCA is an enzymatically driven synthetic technique powered by RNA or DNA polymerases. This amplification has a couple of benefits compared to the well-known polymerase chain reaction (PCR) such as the capability to be carried out at lower temperatures without the need of thermo-cycling, which is something that does not happen in living cells[6,7]. It also needs only one primer for each circular substrate to generate hundreds of repeats[8]. Currently the RCA methodology has been applied frequently in biomedical research[9,10] because of its simplicity; more specifically in diagnostics[11], biodetection[12], nanotechnology[13,14] and immunoassays[15,16].

The second part of the project involved the incorporation of fluorescent base analogues in DNA. Unmodified or native DNA is difficult to study with fluorescence as it is an intrinsically non-fluorescent molecule. This has led to the development of numerous fluorescent base analogues[17] and dyes for imaging and detecting DNA. One of the base analogues that has proven very useful as a probe is 1,3-diaza-2-oxophenothiazine, also known as the tricyclic cytosine (tC) analogue, because it has properties that closely resembles the naturally occurring nucleobase cytosine[18], with the added value of exhibiting a high and stable fluorescence when incorporated in dsDNA, without affecting the overall DNA[19]. This makes it possible to incorporate a tC molecule in close proximity to the sites of interest without altering the DNA structure to a significant extent.

By applying the RCA methodology in combination with short single-stranded DNA oligomers containing the tC base analogues, it could be possible to obtain dsDNA molecules with incorporated tC bases on sequence specific sites in the duplex. In this way it would be possible to make large and intrinsically fluorescent DNA molecules and study their interactions with the IHF protein by fluorescence means.

If a robust RCA-based protocol can be designed to obtain dsDNA molecules that interact with the IHF protein, the principle could easily be expanded to study numerous DNA-protein interactions to a more detailed extent using *e.g.* nanochannels. This would help to pave a way towards novel treatments and diagnosis. If the introduction of fluorescent base analogues is successful, it would minimize the perturbation effects that fluorescent dyes can have on the dsDNA molecules. This would provide a means for producing results *in vitro* that more closely resemble the situation *in vivo*.

2. Theoretical background

2.1. DNA

Deoxyribonucleic acid (DNA) is the carrier of the genetic information of all organisms and was discovered in 1944 by Oswald Avery and co-workers[20]. The structure of this important molecule was further examined by Watson and Crick in 1953[21] and will be explained briefly in the following section. The discovery of the structure paved the way for further research, making DNA one of the most researched molecules. Because of this, it is now known that all living organisms have a genome, the genetic material of a cell, that can construct a replica of itself to reproduce and pass on the genetic information. Some of the key biological processes involved in this DNA replication will also be explained.

2.1.1. Structure

In DNA, two different types of heterocyclic bases are found, namely purines and pyrimidines. In total there are four different bases: adenine (A), guanine (G), cytosine (C) and thymine (T)[22] (Figure 1). The bases combine with a deoxyribose sugar to form nucleosides.

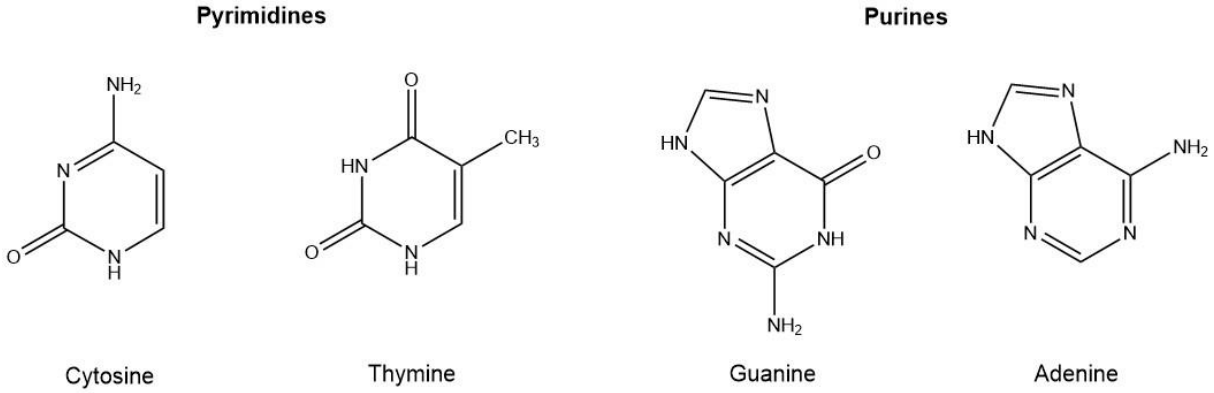

Figure 1. The four bases of DNA.

When nucleosides get phosphorylated they are called nucleotides. Nucleotides can have mono-, di- or triphosphate group(s) attached to the nucleoside sugar moiety.

The total structure of a single-stranded DNA (ssDNA) molecule is composed of nucleotides linked by 3'-5' phosphodiester bonds and this chain has directionality, meaning that there is a difference between the two ends of the chain (Figure 2). The sequence of DNA gives the primary structure of a nucleic acid. By convention, a strand is typically read from the 5' to the 3' end. The 5' end is the carbon atom that has no residue attached and the 3' end is an hydroxyl group that has not formed a phosphodiester bond (Figure 2).

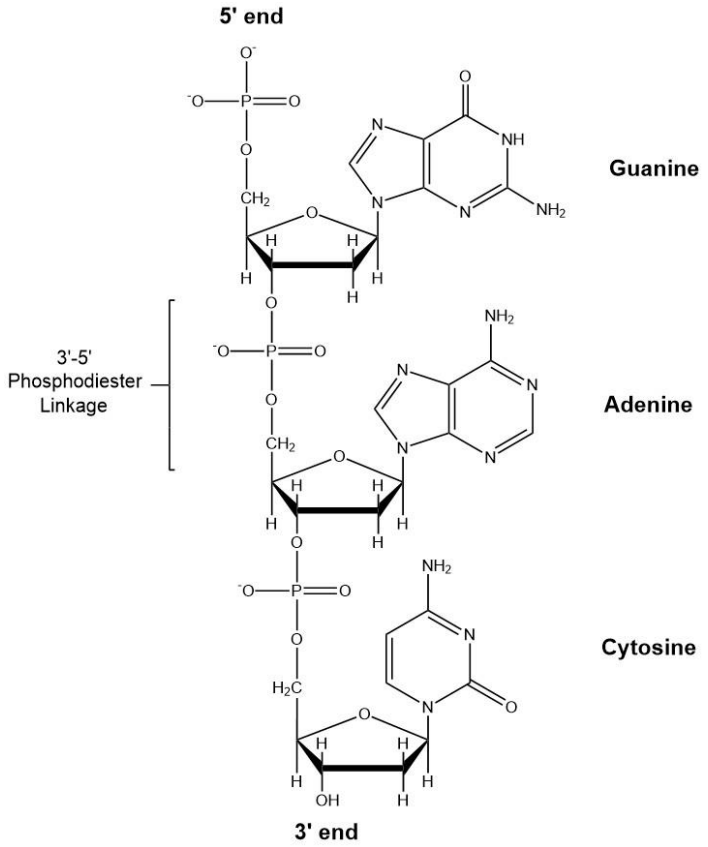

Figure 2. The chemical structure of the nucleotide residues joined together by 3'-5' phosphodiester linkages. The 5' end is the nucleotide with the free 5'-phosphoryl group and the 3' end is the nucleotide with the free 3'-hydroxyl group. The backbone is negatively charged since at least one oxygen on each phosphorous atom is deprotonated.

Two ssDNA molecules can form a duplex. This double-stranded DNA (dsDNA) has two antiparallel strands bounded to each other by hydrogen bonds formed between the complementary bases and was first described in 1953[21]. The pairing of the bases is specific; adenine pairs with thymine and cytosine pairs with guanine (Figure 3). The A/T pair has two hydrogen bonds and the G/C pair has three hydrogen bonds. This specific pairing makes the strands complementary to each other and are therefore capable of serving as templates for each other.

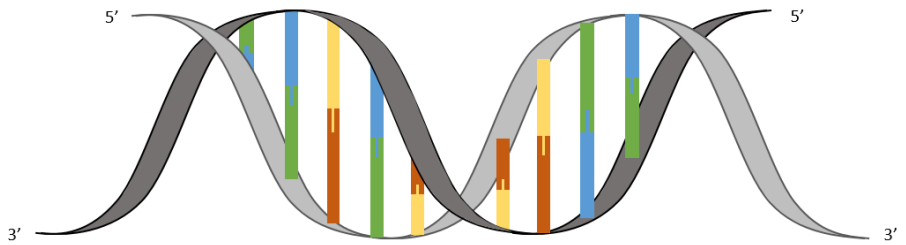

Figure 3. Schematic representation of the double helix structure of dsDNA. The green blocks stand for G, the blue ones for C, the yellow ones for A and the red ones for T. The light and dark grey structures represent the backbone of the two strands.

The duplex consist of an interior that is hydrophobic and a hydrophilic, negatively charged backbone. The double helix is stabilized by several weak forces and interactions, that make up the secondary structure. These forces are strong enough to enforce the conformations but also weak enough to give it a certain flexibility. The different interactions present in dsDNA are hydrogen bonds between the

bases, stacking interactions (van der Waals forces), hydrophobic effects and charge-charge interactions, with the latter having a destabilizing effect.

2.1.2. Biological Processes Involving DNA

Replication of genetic information proceeds by duplicating the DNA of the parent cell to make two daughter cells. The DNA replication mechanism must be rapid and accurate to ensure efficient genetic information transfer and it involves a polymerisation and ligation reaction. Both reactions are possible *in vitro* under favourable conditions and these *in vitro* processes will be explained in the sections below.

2.1.2.1. Polymerisation

The DNA polymerases are a group of enzymes responsible for synthesizing complementary strands in DNA synthesis. This strand synthesis is made by adding deoxynucleoside triphosphates (dNTP's) to the 3' end of the growing chain (See Figure 4a). Since the dNTP concentration is usually equal in cells, the DNA polymerase spends most of its time avoiding the wrong nucleotides[22], because only one of the four options is the right choice. The incorporation of a wrong dNTP delays the synthesis[23]. The polymerase enzymes are involved in DNA replication, -synthesis and -repair and most organisms have more than one DNA polymerase present to fulfil these duties[24].

Under favourable conditions, such as the presence of dNTP's and an initiating oligonucleotide (primer), polymerases have the ability to work *in vitro*. Two important properties when choosing a polymerase for the synthesis *in vitro* are processivity and fidelity[24]. Processivity is the number of nucleotides the polymerase can add, each time it binds to the template, before falling off. Fidelity involves the accuracy of this nucleotide addition. It is preferred that the polymerase has an adequate proofreading system, recognizing the incorrect base pair, and a 3'→5' exonuclease activity, capable of detaching the wrong nucleotide from the growing strand, to ensure a minimum amount of erroneous base pair-mutations.

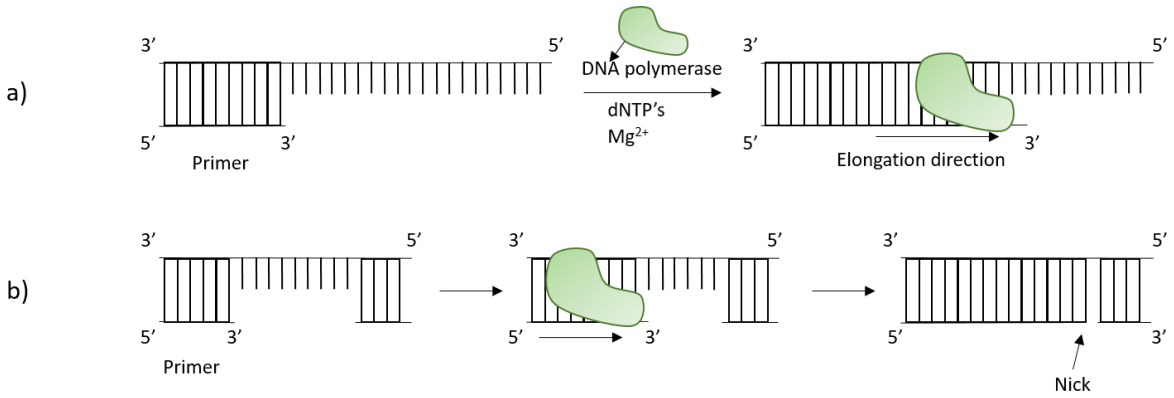

Figure 4. Schematic representation of the elongation of a complementary strand by polymerisation: (a) The synthesis of the complementary strand starting from a primer sequence. The DNA polymerase (green) incorporates the dNTP's in the 5'→3' direction. The elongation stops when the strand is complete. (b) The addition of nucleotides to fill in a gap of an incomplete dsDNA molecule. The DNA polymerase also incorporates the dNTP's in the 5'→3' direction and falls off when it reaches the next dsDNA region, leaving a nick between the two double-stranded regions.

The different polymerases used in the *in vitro* polymerisation reactions in this project were DNA polymerase I and ϕ29 DNA polymerase, the properties of which are listed in Table 1.

DNA polymerase I has the ability to fill up gaps in dsDNA, because of its bifunctional activity[24] (Figure 4b). Bifunctional activity means that it has, next to 3'→5' exonuclease activity, 5'→3' exonuclease activity as well allowing for nick-translation. The optimal pH for this polymerase is 7.4 and the optimal temperature is 37°C, and importantly Mg^{2+} or Mn^{2+} in addition to Zn^{2+} is required for the polymerase to function. The error rate of this polymerase is < 9 x 10^{-6} bases (supplier product sheet) and the processivity is low.

The ϕ29 DNA polymerase (ϕ29 pol) is highly processive (>70 kb before falling off the DNA template strand)[25] and is popular in DNA amplification procedures where displacement of the growing DNA strands is needed, because the polymerase is able to do the displacement together with the polymerisation process[25–27]. The replication rate is 1400-1500 bases/min[28]. In the presence of Mg^{2+} it has high fidelity, because it extends the growing strand 2 x 10^6 times slower with a mismatched nucleotide[27,29], and a high insertion discrimination with an error frequency of 2.2 - 4 × 10^{-5} ensuring high accuracy[29]. Unlike most DNA polymerases, ϕ29 pol does not require accessory proteins to have an optimal polymerisation performance[27].

Table 1. The properties of the DNA polymerases used.

DNA polymerase	3'→5' exonuclease	5'→3' exonuclease	Strand displacement	Processivity
DNA polymerase I	Yes	Yes	No	Low[30]
ϕ29 DNA polymerase	Yes	No	Yes	High

2.1.2.2. Ligation

Nicks in DNA strands can be sealed with a ligation reaction. A nick is a missing phosphodiester linkage between the 3'-hydroxyl group of one nucleotide and the 5'-phosphate group of the adjacent one (Figure 5). The DNA ligase used in this project, is T4 DNA ligase. This enzyme has the capability to seal nicks between two ssDNA fragments annealed to another ssDNA strand[31]. T4 DNA ligase needs ATP to be able to join the nicks and is also dependent on the presence of Mg^{2+} [31,32].

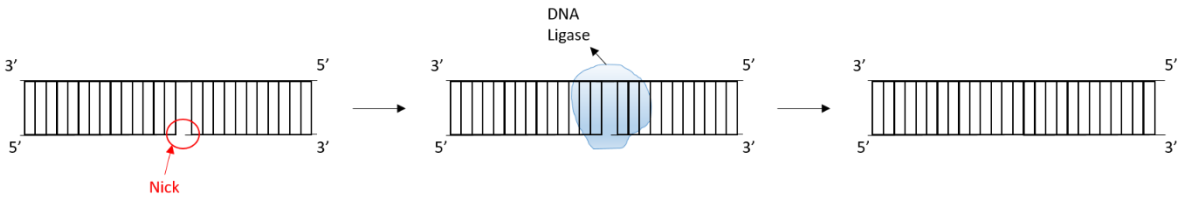

Figure 5. Schematic representation of the ligation mechanism.

2.2. Rolling Circle Amplification (RCA)

The rolling circle reaction scheme is frequently used in nature for the replication of circular plasmids and viral genomes. This inspired the development of a protocol *in vitro* that was first described in 1995[33]. A paper discussing variations on this biochemical mechanism with different circles and polymerases was published in 1996[8]. At that time, rolling circle products (RCPs) exceeding 2000 nt

were reported. The protocol involves a circular DNA molecule that is used as a template for the production of long ssDNA or ribonucleic acid (RNA) molecules by certain DNA or RNA polymerases. The amplification is isothermal and has been reported to have fewer errors compared to a PCR reaction together with good reproducibility[34].

In the RCA reaction a circular DNA molecule is required as the starting molecule; the initial reports only used circular DNA molecules as substrate. In 1994 however, it was reported that certain oligonucleotides, with the use of a ligation reaction, could be made into a circle by adding a shorter complementary strand[35]. The starting linear DNA molecule used to form the circular template is called a padlock probe (PLP) and has typically a size of 25-100 nt[7]. The combination of the PLP molecules and RCA was developed in 1998[36]. The shorter complementary strand (called the target) allows the PLP arms (Figure 6) complementary to the target, to come together for ligation and becomes the primer for the elongation reaction (Figure 7). In this project the sequence of interest is in between the PLP arms.

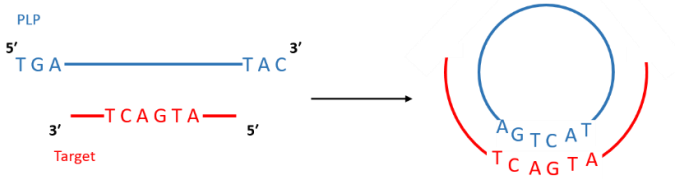

Figure 6. A representation of the PLP (blue), PLP arms (bases displayed at the 5' and 3' in blue) and the target (red) with the complementary bases (red) to these arms. The strands will rearrange themselves to form the hydrogen bonds between the complementary bases. This way the PLP gets the shape of a circle with a nick between the 5' thymine and the 3' cytosine. The nick is sealed in a ligation reaction.

During the RCA reaction the ϕ29 DNA polymerase adds nucleotides starting from the 3'-end of the target and, due to its strand displacement abilities, produces in this way a long continuous strand consisting of multiple repeating copies of the complementary sequence of PLP[25] (Figure 7). The procedure can elongate the ssDNA strand by hundreds to thousands of copies made by the DNA polymerase[13]. The RCP length in nucleotides is dependent on the reaction time of the RCA.

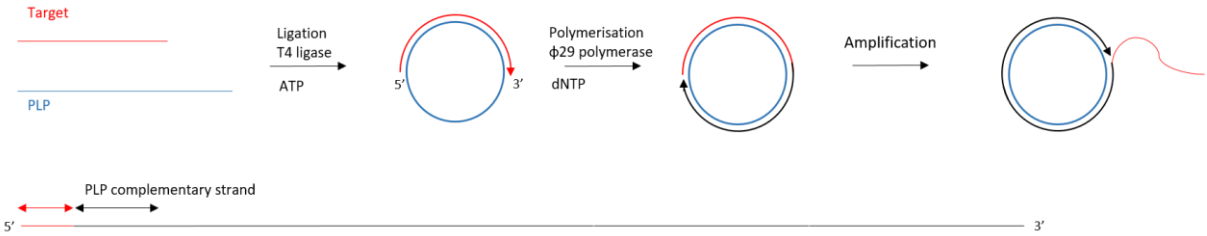

Figure 7. Schematic illustration of the RCA procedure.

2.3. DNA-protein interaction with Integration Host Factor (IHF)

The growing use of single molecule DNA experiments allows for a better understanding of how DNA and certain proteins interact with each other. One of the proteins that is of interest in those studies is the integration host factor (IHF). IHF can wrap DNA in such a way that it forms higher-order structures[5] and is involved in multiple processes carried out in bacteria, such as gene expression, DNA replication

and chromosome packing etc. IHF has two long 'arms' and they bind with, or wrap around, the minor groove. The DNA-protein interactions are made with the minor groove and the phosphate backbone (Figure 8)[37,38].

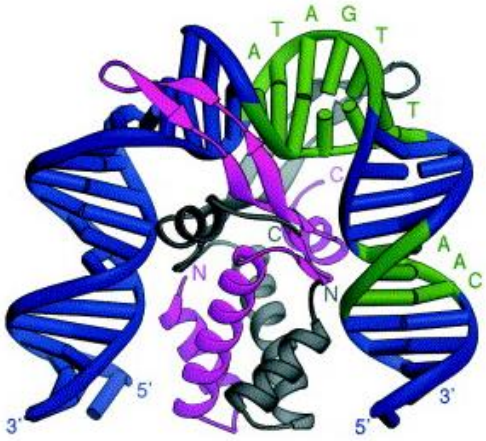

Figure 8. Structure of the DNA-protein complex with IHF[38]. (Reprinted with permission)

The RCPs in this project were specifically designed for the interaction and binding with the IHF protein. The sequence specific binding of IHF is characterized by the core sequence consensus (underlined) WATCAANNNNTTR, which was defined by Friedman[39]. The 'W' in the sequence can stand for dA or dT, the 'R' for dA or dG and 'N' can be any nucleotide. When the core consensus is TATCAANNNNTTG, a dA/dT-rich region is required at the 5'-end of this core consensus. In some other variants of this sequence the core consensus alone is enough to ensure binding, but the presence of the dA/dT-rich region increases the binding[40]. In this project the following core sequence, TATCAAGCCGTTG, combined with a dT region at the 5'-end, was present in the oligonucleotides used.

2.4. UV-vis Spectroscopy

Spectroscopy studies the interactions between light and matter, by measuring the absorption or emission of light by a sample. The spectral region of UV-vis spectroscopy lies between 180 and 700 nm and is divided into two sub-domains: near UV (180-400 nm) and visible (400-700 nm). The absorption in the UV-vis domain occurs when photons from a light source, lamp, interact with molecules or ions in the sample that is examined[41]. Spectrophotometry can be used in quantitative measurements, such as the determination of concentration via the Lambert-Beer law (Equation 1),

$$\log \frac{I_0}{I} = A = \varepsilon_\lambda \, l \, c \qquad (1)$$

I_0 stands for the incoming light intensity, I is the intensity of light after passing through the sample, A is the absorbance (i.e. the amount of light absorbed by the sample), ε_λ is the molar absorptivity [L mol^{-1} cm^{-1}] at a certain wavelength λ [nm], l the pathlength of the light through the sample [cm] and c is the molar concentration of the sample [mol L^{-1}]. The molar absorptivity is specific for the compound being analysed.

In the case of DNA in aqueous solutions, nucleobases absorb around 180-305 nm, with a maximum at around 260 nm. The absorption measured at 260 nm wavelength is typically used with the Lambert-Beer law to calculate the DNA concentration[41].

2.5. Fluorescence Spectroscopy

Some compounds emit light after being electronically excited by photons. The emission of light from excited singlet states to the ground state is called fluorescence and occurs very fast, typically on a nanosecond time scale[42]. Aromatic molecules often have, to some extent, fluorescent properties and such compounds, with significant emission, are called fluorophores. The quantum yield is an important characteristic of fluorophores. The quantum yield of fluorescence is the number of photons emitted per absorbed photon. A high quantum yield is an indication for a bright fluorophore, as brightness is proportional to the multiplication of the quantum yield and the molar absorptivity. A fundamental characteristic of fluorescence is that the emitted photon is lower in energy than the absorbed photon resulting in a Stokes shift. Good fluorophores show a distinctive difference. Also, since emission can be collected in a way where excitation light is geometrically excluded from the detector, a very high signal-to-noise ratio can be achieved[41]. The fluorescence properties of molecules are often used for detection methods because it is one of the easiest and most sensitive techniques. It is so sensitive that it can detect emission at single molecule level[43,44].

The photophysical processes involved in optical spectroscopy can be visually described in a Jablonski diagram (Figure 9). It shows the transition of electrons between a molecule's different electronic states. The different electronic states are visualized as thick black lines and include the ground state (S_0), the first electronic state (S_1) and the second electronic state (S_2). Each electronic state has several vibrational levels (grey lines). Following the absorption of light (photons), several processes can occur. A fluorophore is usually excited to a higher vibrational level of S_1 or a higher electronic state. In this case vibrational relaxation (VR) occurs when this photon rapidly relaxes to the lowest vibrational state of the electronic state. Internal conversion (IC) occurs when the photon undergoes a horizontal transition between electronic states. Fluorescence typically occurs when the photon returns to the ground state from the lowest vibrational state of S_1.

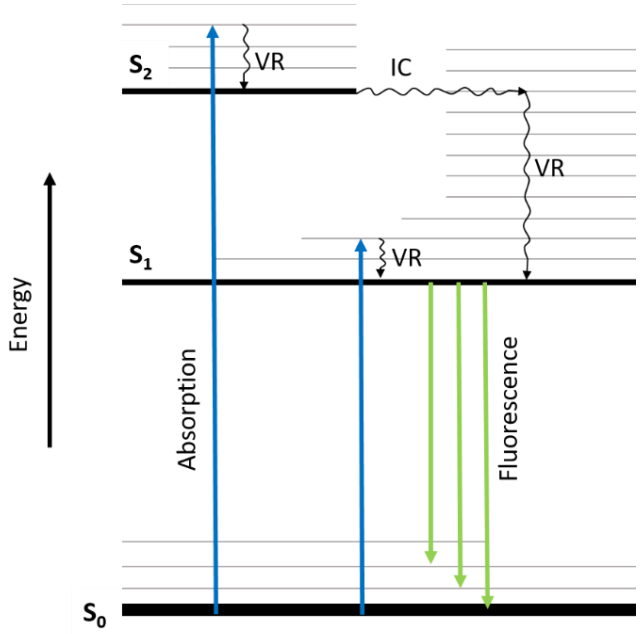

Figure 9. A simplified Jablonski diagram showing the different electronic states (thick black lines, S_0, S_1, S_2) and vibrational levels (thin grey lines). Absorption (blue arrows), fluorescence (green arrows), vibrational relaxation (VR) and internal conversion (IC) are indicated.

2.5.1. Fluorescent Base Analogues and Dyes

A challenge with using fluorescence spectroscopy in nucleic acid research is that DNA and RNA nucleotides are virtually non-emissive[45]. A solution to this problem is to label DNA with internal or external fluorophores. There has been extensive research into making different types of fluorophores for DNA and RNA.

One of the options to make DNA detectable by fluorescence is *via* external binding of a dye. External dyes are typically very bright, with a high quantum yield, and are thus useful as probes in applications where the specific place of binding is less important, for example when using gel electrophoresis. However, they are less useful when a high level of sequence specific detail is needed, such as in Förster resonance energy transfer (FRET) studies[46]. One way of external binding is the addition of an intercalator like YOYO, see section 2.5.1.3. A potential drawback is that an intercalator affects the native DNA structure[47].

An alternative way of fluorescent labelling is the incorporation of modified nucleobase analogues, which are typically designed to mimic the overall function of the native base, but with the added value of being fluorescent. Such synthetic labelling are sometimes denoted internal modifications[48]. Several fluorescent base analogues have been developed for all of the five nucleobases (the fifth being uracil, naturally occurring in RNA)[17]. These modified bases often closely resemble their natural counterpart in size and binding properties towards the other natural bases in the same DNA strand. Also they possess the possibility to form a stable base pair with at least one of the natural bases on the opposing DNA strand in a dsDNA molecule. Because of these traits, their incorporation has a minimally disruptive effect on the DNA structure. A general drawback with fluorescent base analogues is that they are significantly less bright compared to external dyes[17].

During this project, the fluorescent base analogue tC, the ATTO 647N dye labelled deoxyuracil triphosphate (dUTP) and the dye YOYO were used.

2.5.1.1. 1,3-diaza-2-oxophenothiazine (tC)

1,3-Diaza-2-oxophenothiazine (tC) is a fluorescent cytosine base analogue where the native structure has been extended with two additional aromatic rings (Figure 10). tC discriminates well between A and G targets[49] and thus favourably forms a Watson-Crick base pair with guanine bases[50]. It also shows very robust fluorescent properties with different neighbouring bases and chemical environments, and the addition of the base analogue is essentially non-disruptive to the native DNA structure[18]. tC has a redshifted emission compared to many other fluorescent base analogues, is reasonably photostable and has an excitation wavelength well separated from the natural bases (Figure 11). These characteristics made it the base analogue of choice in this study[18].

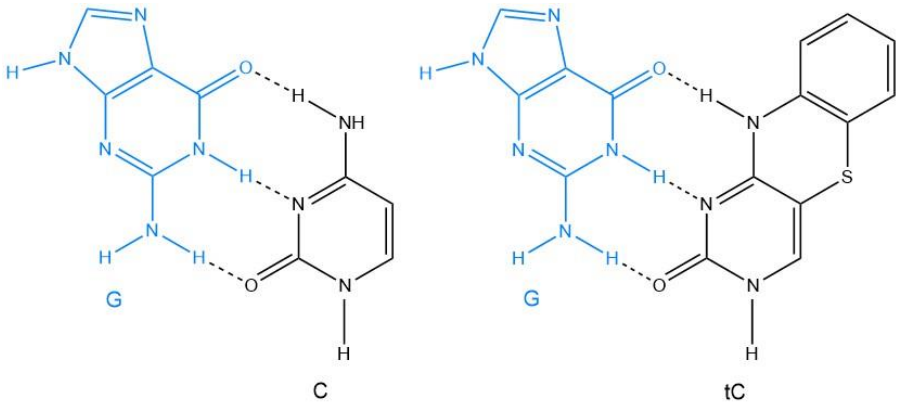

Figure 10. Chemical structures of cytosine (C) and tricyclic cytosine (tC) with their respective hydrogen bonds (dotted lines) to guanine (G).

The absorption and emission spectra of tC incorporated in double-stranded DNA is seen in Figure 11. The lowest energy absorption band has a maximum at 392 nm and the emission maximum is at 503 nm[18]. The average quantum yield is 0.19 and the extinction coefficient is 13500 mol $L^{-1}cm^{-1}$ [18].

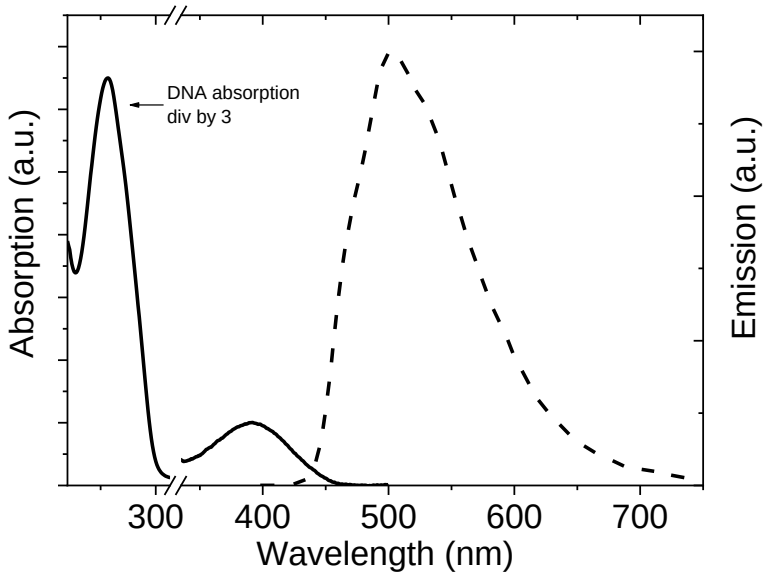

Figure 11. Normalized absorption (solid line) and emission (dashed line) spectra of the fluorescent base analogue tC incorporated in dsDNA.

2.5.1.2. Aminoallyl-dUTP-ATTO 647N

Aminoallyl-dUTP-ATTO 647N is a deoxyuracil triphosphate (dUTP) molecule labelled with ATTO 647N dye (Figure 12). This labelled nucleotide was used as a control for the RCA protocol. The molecule has an excitation maximum at 644 nm and an emission maximum at 669 nm. The spectra can be seen in Figure 14.

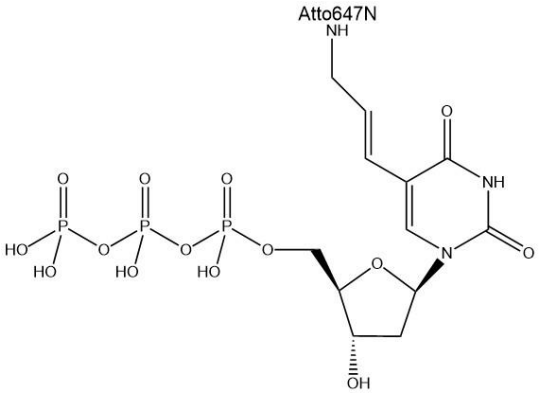

Figure 12. Structural formula of Aminoallyl-dUTP-ATTO 647N.

2.5.1.3. YOYO dye

The DNA binding dye YO, also known as oxazole yellow, and its homodimer 1,1'-(4,4,8,8-tetramethyl4,8-diazaundecamethylene)-bis[4-[(3-methylbenzo-1,3-oxazol-2-yl)-methylidene]-1,4-dohydroquinolinium] tetraiodide (YOYO; Figure 13, left), are highly fluorescent when intercalated into dsDNA but poorly fluorescent when free in aqueous solution, corresponding for YOYO to a quantum yield that is 3200 times lower, and have no visible interaction with monomeric monophosphate nucleotides[51,52]. This makes the dyes popular for research concerning DNA because the background noise is limited. When comparing YO and YOYO with each other it is concluded that YOYO has a higher binding affinity to DNA than YO but both dyes interact in the same way[53]. Though the fluorescence of the dye is brighter when bound to dsDNA, YOYO is also capable of binding to ssDNA[54].

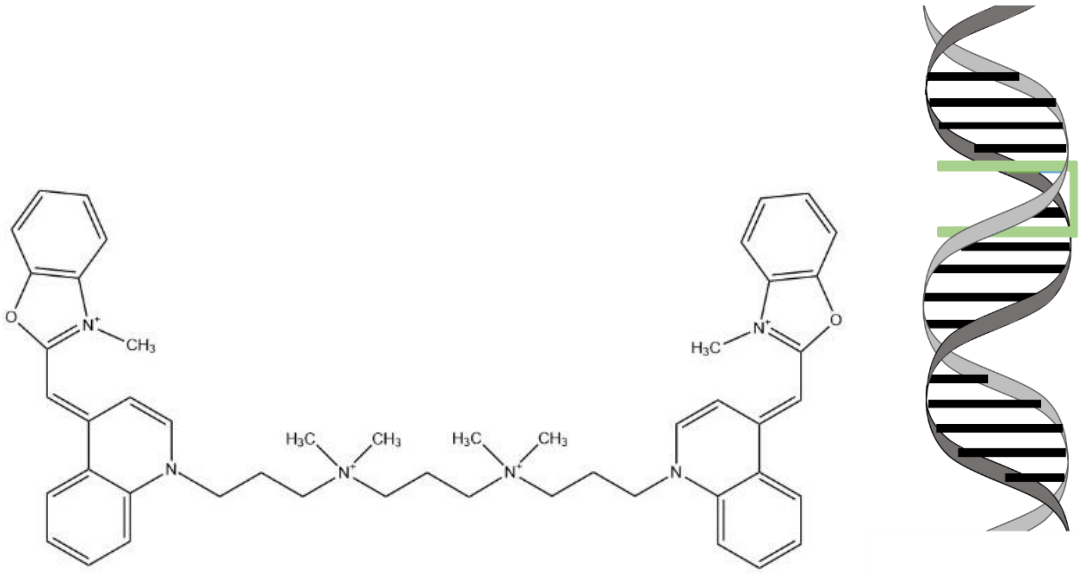

Figure 13. The chemical structure of YOYO (left). A schematic representation of bis-intercalation of YOYO (green) binding with DNA (right).

The mechanism of binding of YOYO to DNA is through bis-intercalation[53,55] (Figure 13, right). The dye intercalates between the bases in the DNA structure by inserting the two aromatic ring structures. This insertion elongates the DNA[47]. When intercalated into DNA, YOYO has an excitation maximum of 489 nm and an emission maximum of 509 nm[55] (Figure 14). The quantum yield of YOYO, when incorporated in dsDNA, is dependent on the mixing ratio. For a 0.01 chromophore:bases mixing ratio it is measured to be 0.35 and the extinction coefficient of YOYO at 457 nm is 96100 mol $L^{-1}cm^{-1}$ [56].

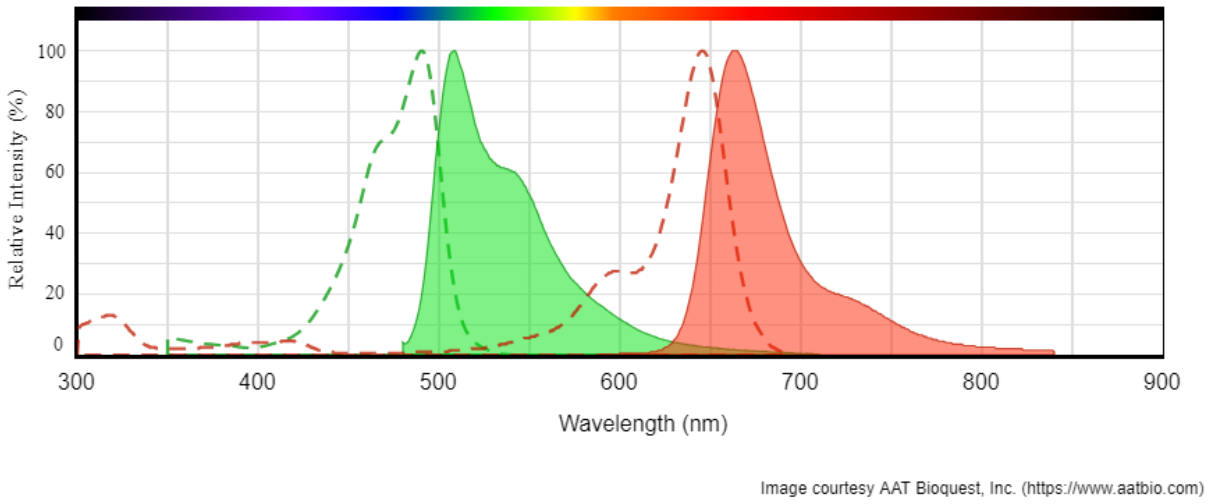

Figure 14. The absorption (dotted line) and emission (full line) spectra of YOYO are shown in green. The absorption(dotted) and emission (full) of ATTO 647N are shown in red.

2.5.2. Fluorescence Microscopy

Because of advancing technology on filters, cameras and mirrors used in fluorescence microscopy, it has gained the ability to be used for detection down to the level of single molecules[57]. Fluorescence microscopy was the detection method used throughout this work.

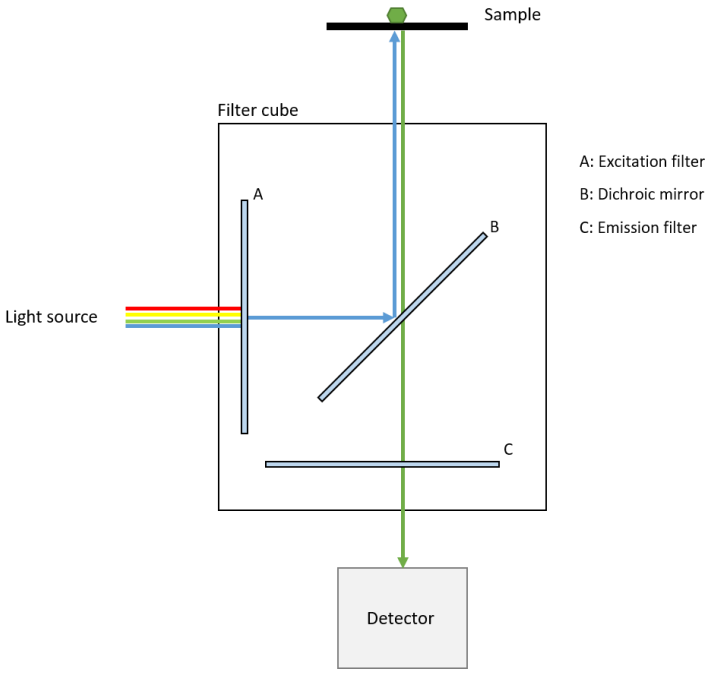

Figure 15. Schematic representation of an inverted fluorescence microscope with its main components.

This type of microscopy uses emission from added fluorophores or inherently fluorescent compounds to detect and visualize for instance small molecules. It consists of a light source for excitation, an excitation filter and an emission filter, an objective to focus the light and a detector to detect the emitted photons (Figure 15).

The types of light sources that used to be the most common for fluorescence microscopy were mercury and xenon lamps, but these lamps do not always provide an equal intensity across the spectrum. Lasers and LED sources are increasing in use, because they are smaller in size and achieve near-monochromatic light[58].

The filters usually have a dichroic mirror that is capable of reflecting the excitation light coming from the light source (Figure 15, blue arrows) to the sample and transmitting the emitted light, because it consists of longer wavelengths (Figure 15, green arrow). By adding this dichroic mirror the weak emitted light can be separated from the much stronger excitation light. The combination of the dichroic mirror, the excitation filter and the emission filter forms the filter cube, which can be changed depending on the fluorophore used.

The excitation and emission light pass through the same objective, which must be in perfect alignment. This objective gathers the emitted light in such way that it can form an image. To detect the small amount of emitted photons a sensitive detector is needed. An example of such a detector is an electron multiplying charged coupled device (EMCCD).

2.6. DNA Stretching

DNA molecules can be viewed as polymers made up of several hundreds of thousands of nucleotides. In solution, DNA is coiled up, which makes estimating characteristics such as size and mass difficult, unless the DNA is stretched. Different methods have been developed to be able to get the DNA to

stretch, such as the use of magnetic[59] and optical[60] tweezers, where one or both ends of the DNA molecule is attached to a bead, stretching DNA on positively charged glass, attaching DNA to a surface with the flow of the buffer being the force that causes stretching[61] and through nanochannels[62]. These methods typically use fluorescence microscopy for detection and imaging, and often total internal reflection fluorescence (TIRF) microscopy.

The two methods used in this study are stretching on glass and confinement of DNA in nanochannels. Both will be more explained in the following paragraphs.

2.6.1. Stretching on Glass

To be able to stretch DNA on glass, the surface of the glass has to be modified in a chemical way, for example making it positively charged, or by adding a polymer coating to make it hydrophobic[63]. This charge or hydrophobicity allows one end of the DNA backbone to anchor to the surface and the flow of the solution will induce the stretching. The glass surface has to be impeccably clean to get the best possible results, otherwise the stretching of DNA will not occur.

2.6.2. Nanofluidics and Nanochannels

Stretching of DNA in nanochannels employs nanofluidics, which is the study and manipulation of fluids confined within nanostructures. Together with microfluidics, nanofluidics has had a rapid increase over the last 10-20 years for use in research of small-scale systems and in single-molecule experiments. When scaling down to the nanometer level, surface-to-volume ratios and surface effects become increasingly important[64]. Because of the high surface-to-volume ratio, all molecules present will be close to the surface and this can induce non-specific binding to the walls of the channel. Even molecules with a small affinity to the surface can give problems, such as clogging the channel entrances.

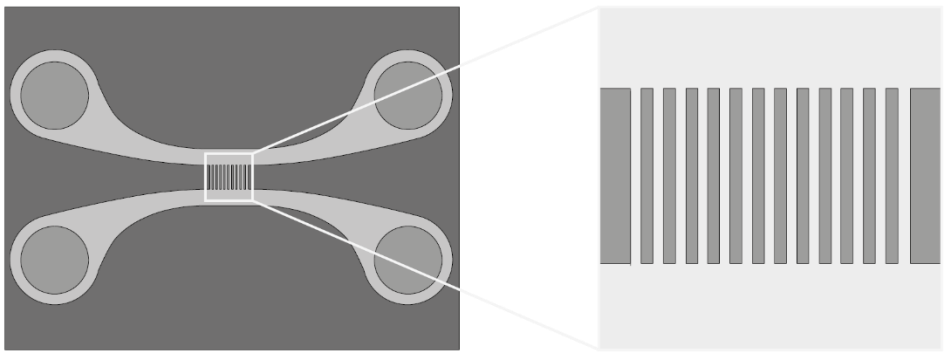

Figure 16. Nanofluidic chip with four loading wells (dark grey circles) and microchannels (left). These two microchannels are connected in the middle by rows of nanochannels (150x100nm²; right)[65].

Compared to other methods of stretching DNA, nanochannels have the advantage that every molecule can be stretched because of the confinement of the channel and thus they are not attached to any surface. An example of such a nanofluidic device is shown in Figure 16. The extension in the nanochannels scales linearly with the contour length. An extra remark that has to be made is that the

stretching of DNA is dependent on the solvent that they are in, mainly on the presence of ions in that solution. The degree of extension is higher when the ionic concentration in the buffer is low.

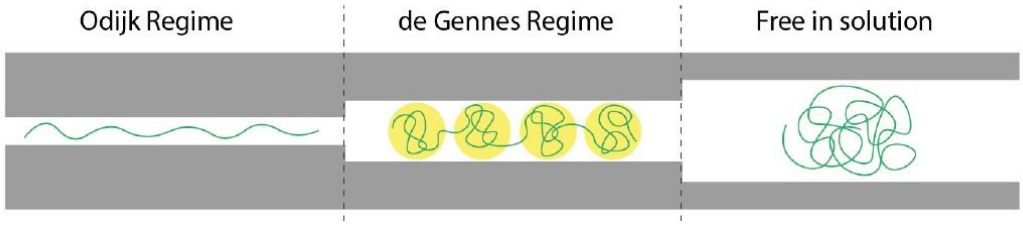

Figure 17. Different physical regimes in different types of confinement in the nanochannels[65].

Depending on the dimensions of the channels, different physical regimes can exist. As seen on Figure 17, DNA coils up when it is free in solution. When the space gets limited due to a decreasing channel size, DNA will start to fold differently. The two best studied and characterized regimes are the Odijk regime and the de Gennes regime. The Odijk regime is observed when the space is too limited for DNA to coil up, leading the DNA to be stretched along the length of the channel. During the de Gennes regime the channel is wide enough to allow for DNA to make a continuous series of smaller coils, but too narrow for it to coil up completely. In practice, most channel dimensions usually lie between these two regimes[66].

3. Methods

This chapter aims to provide a detailed description of the experimental methods and instruments used throughout this project.

3.1. Oligonucleotides

Designed oligonucleotides were used to create complementary strands to the single-stranded DNA produced in the RCA procedure described in section 3.4 below. The complementary oligonucleotide sequences consisting of all natural nucleobases are referred to as primer3 (P3) and primer4 (P4). The ones having incorporated tC base analogues are referred to as primer31 (P31), with only one tC base analogue, and primer32 (P32), with four.

All oligonucleotides were manufactured by ATDBio Ltd. More detailed information about these oligonucleotide sequences and their parameters can be found in Appendix 1. The oligomers were received as lyophilised and desalted solids. All the oligonucleotides were dissolved in 100 µL Tris hydrochloride (Tris-HCl) (10 mM, pH=7.5) and aliquoted. All oligomer samples were stored in -20 °C.

3.2. Stock Concentrations

The concentrations of the stock solutions were determined spectrophotometrically using a Varian Cary 4000 Spectrophotometer. A cuvette was used with a path length of 0.3 cm and a volume of 60 µL.

The absorption at 260 nm together with the molar absorptivity and the path length was used to determine the concentration, *via* Lambert-Beer's law (equation 1, p.8). The calculated stock concentrations are given in Table 2. All oligomer concentrations in this thesis refer to the molar of the entire strand, *i.e.* not base pair, nor base concentration. From these stock concentrations all further dilutions were prepared using the Tris-HCl buffer.

Table 2. The measured absorbance at 260 nm and the calculated stock concentrations.

	PLP	**Target**	**P3**	**P4**	**P31**	**P32**
A_{260}	0.224	0.126	2.15	1.95	1.77	2.57
Concentration [µM]	1.11	1.09	21.5	19.0	18.2	27.9

3.3. Different Experiments

Three different types of experiments have been performed in this study. Figure 18 provides an overview of the different reaction schemes used in each experiment and its sub-experiments. The next sections will describe these intermediate steps in a more detailed manner, always using the mixture from the previous step as starting solution.

In Experiment 1, the RCA protocol is followed by an annealing process and a ligation reaction. The terms 1X and 10X refer to the starting concentration of PLP and Target used. For the RCA (10X 1h) and

the RCA (10X 0.5h) the annealing can differ between a process with P3 & P4 added or P32 & P4 added to the mixture.

Experiment 2 has an additional polymerisation step before the ligation reaction and Experiment 3 has only one step, namely RCA, where the amplification time is prolonged to 24 hours.

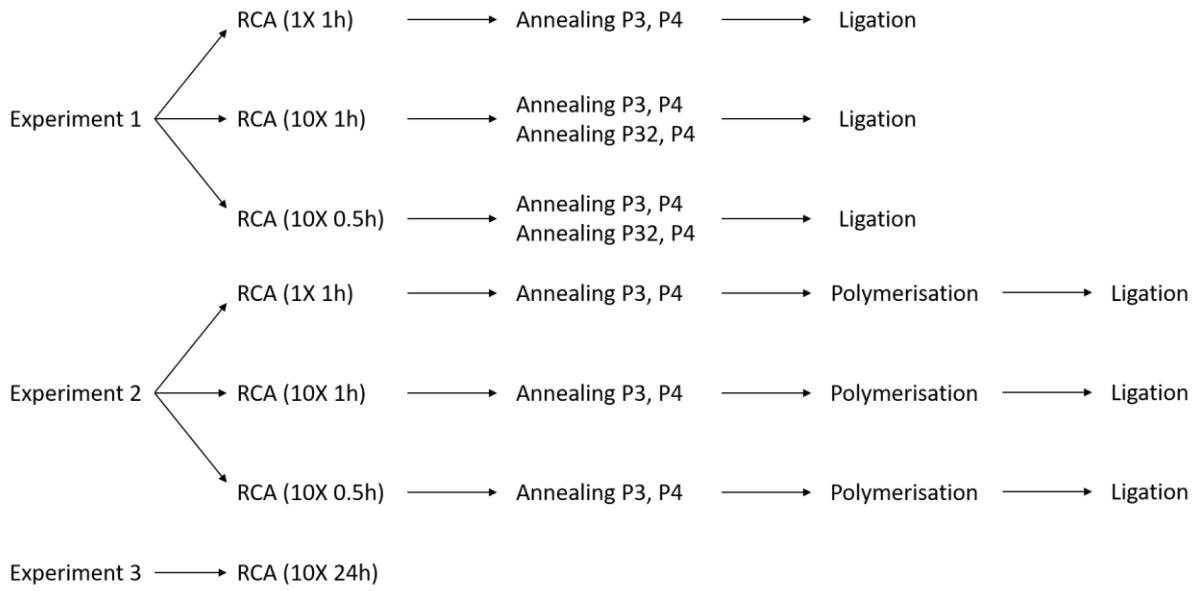

Figure 18. An overview of the three different experiments performed during this study with their intermediate, follow-up reaction steps. More detailed information of these steps will be provide in the text. Rolling circle amplification (RCA) is used in every single experiment as the first step, varying in concentration (1X, 10X) and amplification time (0.5h, 1h, 24h). The second step is the annealing of primer 3 (P3) & primer 4 (P4) or primer 32 (P3), containing tC molecules, & P4.

3.4. Rolling Circle Amplification (RCA)

A starting point of the protocol for RCA was provided by Prof. Mats Nilsson at Stockholm University.

1X 1h

The first step in the RCA procedure is a ligation reaction, in which the circular DNA templates for the polymerase reaction are made. The ligation was performed with 100 pM PLP (ATDBio Ltd), 10 pM Target (ATDBio Ltd), 0.02 U/µL T4 DNA ligase (NEBiolabs) in 1× ϕ29 DNA polymerase buffer (33 mM Tris-acetate, pH 7.9 at 37 °C, 10 mM magnesium-acetate, 66 mM potassium-acetate, 0.1% (v/v) Tween 20, 1 mM DTT; Thermo Fisher Scientific), 0.2 µg/µL bovine serum albumin (BSA; NEBiolabs) and 0.68 mM ATP (Thermo Fisher Scientific). The final volume was adjusted to 20 µL with milliQ water. This mixture was incubated for 15 min at 37 °C. Immediately after the incubation, the reaction was inactivated for 1 min at 65 °C.

After the ligation reaction, the amplification step was performed by 160 mU/µL ϕ29 DNA polymerase (Thermo Fisher Scientific) in 1× ϕ29 DNA polymerase buffer, 125 µM dNTP's (NEBiolabs) and 0.2 µg/µL BSA. The final volume was adjusted with milliQ water to 30 µL. The sample was incubated for 1 hour at 37 °C and inactivated for 1 min at 65 °C.

10X 1h

The ligation was in this case performed with 1 nM PLP, 1 nM Target, 0.04 U/µL T4 DNA ligase in 1× ϕ29 DNA polymerase buffer, 0.2 µg/µL BSA and 6.8 mM ATP. The volume was also adjusted to 20 µL by addition of milliQ water. The sample was incubated at 37 °C for 15 min and after this inactivated at 65 °C for 1 min.

The amplification step was performed after the circular DNA templates were ligated. To the previous mixture, 160 mU/µL ϕ29 DNA polymerase was added together with 1× ϕ29 DNA polymerase buffer, 1.25 mM dNTP's and an extra addition of 0.2 µg/µL BSA.. After adjusting the volume to 30 µL with milliQ water, the sample was incubate at 37 °C for 1 hour. The inactivation of the reaction took place immediately after at 65 °C for 1 min.

For samples named '10X 0.5h' the incubation of the amplification step was decreased to 30 minutes.

10X 24h

The ligation reaction forming the circular DNA templates followed the same protocol as described in the 10X 1h method.

The amplification step was performed when adding 0.5 U/µL ϕ29 DNA polymerase together with 1× ϕ29 DNA polymerase buffer, 1.8 mM dNTP's and 0.2 µg/µL BSA. The volume was adjusted to 30 µL with milliQ water. The mixture was incubated at 37 °C for 24 hours and inactivated at 65 °C for 1 min.

3.4.1. RCA with Fluorescently Labelled dUTP

For the reactions involving fluorescently labelled dUTP, the making of the circular templates follows the procedure described in 1X 1h. Further, the amplification step is performed by mixing 160 mU/µL ϕ29 DNA polymerase in 1× ϕ29 DNA polymerase buffer, 123 µM dNTP's (NEBiolabs), 2 µM Aminoallyl-dUTP-ATTO 647N (Jena Bioscience) and 0.2 µg/µL BSA.

3.5. Annealing Process

A 50 mol% excess of P3 and P4 was used in the annealing reaction, in which the fully complementary P3 & P4 binds to the single-stranded rolling circle products. The RCA process runs approximately 1000 cycles in 60 minutes, meaning that the total amount of DNA is amplified a 1000 times. Based on this estimated reaction rate, the final concentrations were 83 nM for both P3 and P4 in the 10X 1h samples and 41.5 nM in the 10X 0.5h samples. In the annealing reactions where P31 & P4, and P32 & P4 were added to the RCPs, the final concentrations were the same.

After preparing the solutions containing the RCPs and complementary strands, the annealing reaction was carried out in a temperature-controlled water bath. Table 3 lists four of the annealing processes used in this study.

Table 3. An overview of some of the annealing protocols used.

Annealing experiment	Warm-up	Highest temperature	Cool-down
1	22 °C → 95 °C	20 min at 95 °C	95 °C → 22 °C

	Rate: 3.65 °C/min		Rate: 1 °C/min
2	22 °C -> 95 °C Rate: 3.65 °C/min	20 min at 95 °C	95 °C → 22 °C Rate: 0,5 °C/min
3	22°C → 80 °C Rate: 2.90 °C/min	20 min at 80 °C	80°C → 22 °C Rate: 1 °C/min
4	22°C → 80 °C Rate: 2.90 °C/min	20 min at 80 °C	80°C → 22 °C Rate: 0.5 °C/min

A different annealing process, based on the protocol that has been used for the specific labelling of RCPs with fluorophore containing complementary strands[67], was also attempted to have a comparison and to strive for a maximal amount of annealed complementary strands. This annealing protocol is described in Table 4.

Table 4. Annealing process based on the specific labelling of RCPs.

Annealing experiment	**Warm-up**	**Heat step**	**First cool-down**	**Hybridisation**	**Second cool-down**
5	22 °C → 70 °C Rate: 3.20 °C/min	5 min at 70 °C	70 °C → 55 °C Rate: 0.5 °C/min	15 min at 55 °C	55 °C → 22 °C Rate: 0.5 °C/min

3.6. Polymerisation Reaction

The polymerisation reaction that fills in the gaps between the annealed strands was carried out with 0.4 U/µL DNA polymerase I (NEBiolabs) in 1× NEBuffer (50 mM NaCl, 10 mM Tris-HCl, 10 mM $MgCl_2$ and 1 mM DTT; NEBiolabs) and an additional amount of 125 µM dNTP's (NEBiolabs). The sample was incubated at 37°C for 60 min and was stopped at 65°C for 10 min.

The 1X 1h, 10X 1h and 10X 0.5h samples of Experiment 2 used the same polymerisation reaction, as every component was added in high excess.

3.7. Ligation Reaction

The ligation reaction that seals the nicks in the dsDNA was carried out by mixing 160 mU/µL T4 DNA ligase (NEBiolabs) in 1× of the ϕ29 DNA polymerase buffer (Thermo Fisher Scientific) with an extra addition of 1.70 mM ATP (Thermo Fisher Scientific). The reaction was incubated at room temperature for 60 min and deactivated at 65°C for 10 min.

The temperature used here was lower than the ligation temperature used in the RCA protocol, because the optimal temperature for T4 DNA ligase lies between 16 - 22 °C. 37 °C was used to speed up the process in the RCA protocol, but here is was more important that the ligase was capable of sealing every nick present in the DNA molecules.

The 1X 1h, 10X 1h and 10X 0.5h samples in Experiment 1 and 2 used the same protocol for the ligation reaction, since all components were added in high excess.

3.8. Fluorescence Microscopy

To visualize the RCA-based product used in this thesis an inverted fluorescence microscope was used. See section 2.5.2 for more information.

3.8.1. Fluorescent Staining of DNA With YOYO

The samples containing the P3 and P4 (*i.e.* without tC) were stained with the fluorescent dye YOYO. Typically the theoretical molar concentrations of DNA-YOYO ratio is 5:1. Due to the uncertain concentration of DNA produced in the RCA, the relative amount of YOYO was increased. 0.2 µL of 40 µM YOYO (Thermo Fisher Scientific) was added to 10 µL of the sample, thus it was present in high excess. Following the addition of the dye, the solution was incubated at 50°C for 60 min.

3.8.2. Microscope Settings

The microscope used was an Axio Observer Z1 from ZEISS equipped with an EMCCD camera (Photometrics Evolve) and a Lumencore SpextraX light source (Nikon instruments). Light of 475 nm was used to excite YOYO. The emission filter used was a fluorescein isothiocyanate (FITC) filter. The excitation and emission spectra of YOYO, as well as the light passage through the FITC filter is showed in Figure 19. For all subsequent images a 100× TIRF oil immersion objective was used. The exposure time was set to 100 ms and the light intensity for excitation was 20% of the maximal LED light.

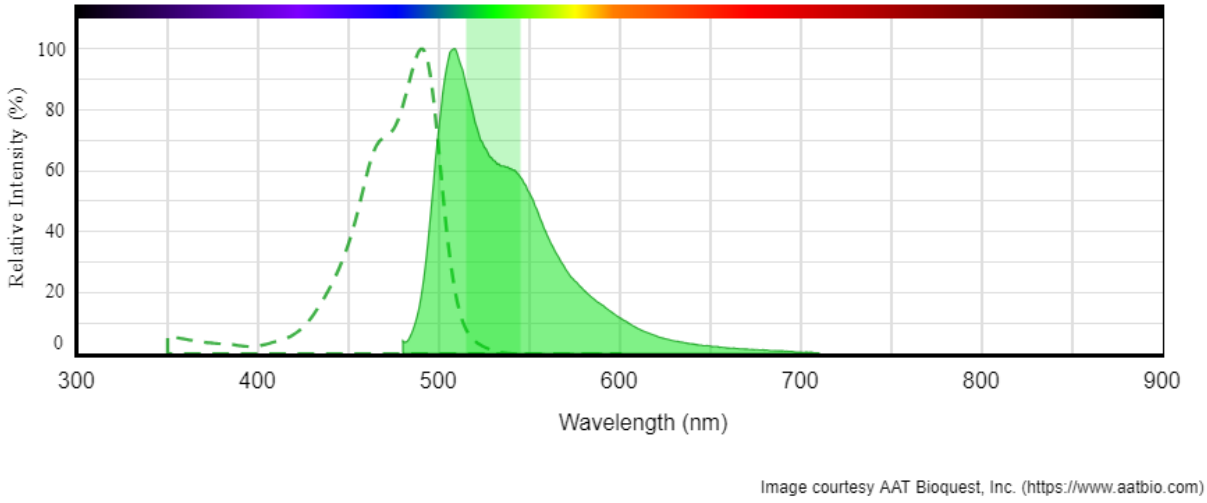

Figure 19. The absorption (dotted line) and emission spectra (full line) of YOYO. The transmission of the FITC filter is displayed as a rectangle in light green.

For visualizing aminoallyl-dUTP-ATTO 647N, light at 630 nm was used to excite the dye. The emission filter was a Cyanine 5 (Cy5) filter (Figure 20). The same objective was used as with YOYO, the exposure time was set to 300 ms and the light intensity for excitation was 100%.

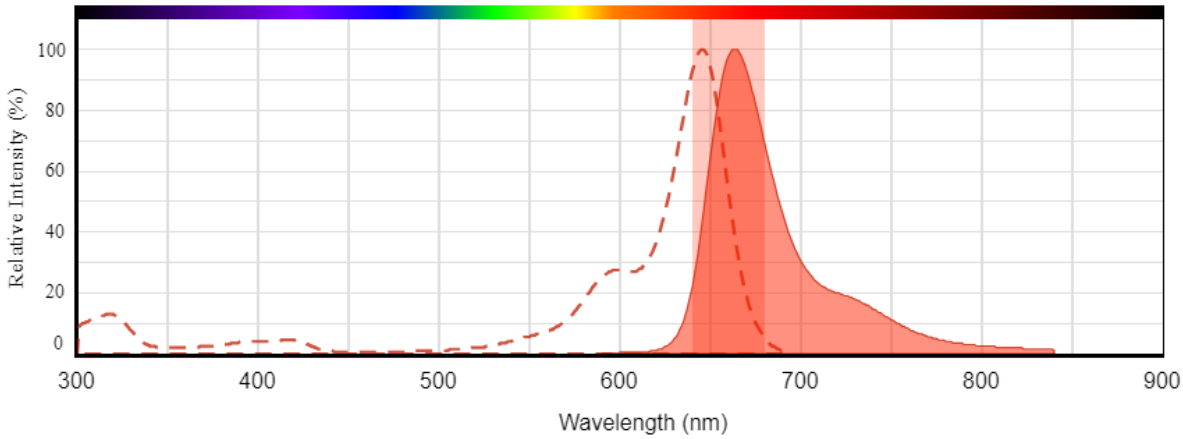

Figure 20. The absorption(dotted line) and emission (full line) of ATTO 647N. The transmission of the Cy5 filter is displayed as a rectangle in pink.

For the visualisation of the samples containing tC base analogues, the filter set 90 HE (Zeiss) was used as emission filter (Figure 21). Light at 385 nm was used to excite the tC base analogues in the sample. The exposure time was set to 100 ms and the light intensity for excitation was 20% from the maximal LED inlight.

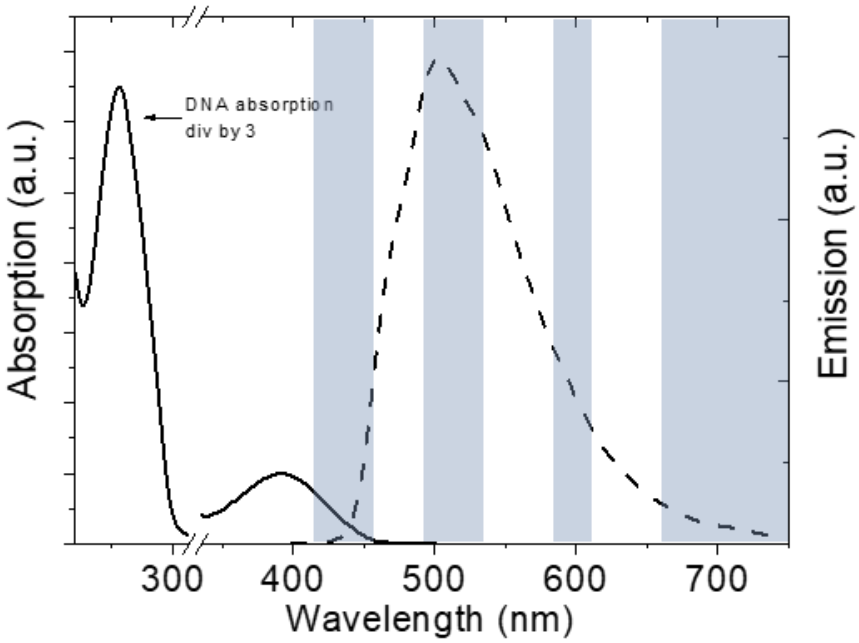

Figure 21. The absorption (full line, with peak at 390) and emission spectra (dotted line) of tC when incorporated in dsDNA. The transmission of the emission filter set 90 HE is portrayed by the grey rectangles.

3.9. DNA stretching

Prior to visualizing the DNA in the microscope, the samples were stretched. Three different methods, for stretching the DNA, are used: positively charged slides, activated coverslips, or nanofluidics in nanochannels.

To all samples was, prior to stretching, a reducing agent added. This reducing agent reacts with the reactive oxygen species (ROS) that get released when YOYO emits light and thus makes sure that the ROS cannot react with or damage the DNA present.

3.9.1. Positively Charged Slides

The DNA was stretched on VWR Microscope Slides (26x76x1 mm; positively charged; white). On top of this slide a Marienfeld Microscope Cover Glass was placed. To 10 µL of the sample 1 µL DTT was added as reducing agent and 3-10 µL of the resulting mixture was pipetted in between the positively charged slide and the cover glass. The slide was placed on the microscope when the solution had spread under the entire cover glass.

3.9.2. Activated Coverslips

Uncharged Marienfeld Microscope Cover Glasses were added in a solution of (3-aminopropyl)triethoxysilane (APTES; ≥98%; Sigma Aldrich) – allyltrimethoxysilane (ATMS; 95%; Sigma Aldrich) – acetone (Sigma Aldrich) in a 1:1:100 ratio. The cover glasses were kept in the solution at room temperature for at least 30 min before use, where after they were rinsed with acetone and water at least three times each and dried with air. One coverslip was placed on top of a Menzel-Gläser (Bevelled edges, frosted end, Thermo Fisher Scientific). 3 µL of the solution, containing stained DNA and DTT, was pipetted between the Menzel-Gläser and the charged cover glass. As soon as the solution had spread under the activated coverslip, the edges were sealed with nail polish. After this the slide was placed on the microscope.

3.9.3. Nanochannels

The nanofluidic device pictured in Figure 16 was used for stretching the DNA, with nanochannels that had a dimension of 150x100 nm². A lipid coating was applied to lower the affinity to the surfaces and prevent the clogging of the channels (see section 2.6.2)[68]. To make the coating visible, the lipids can be fluorescently labelled.

First the channels were cleaned by adding 10 µL milliQ water in every loading well. The water was flushed through the channels and removed afterwards. 10 µL Lipid buffer (0.11× TBE + 100 mM NaCl) was loaded. While flushing, the channels were studied under the microscope using white light to make sure that all the air was replaced by buffer. The excess of buffer was taken out after flushing and replaced by 7 µL of a 1% fluorescein labelled lipid coating in each well. The coating was studied under the microscope with the same settings as described for YOYO in section 3.8.2. After full coating of the micro- and nanochannels in the device, the excess lipid was taken out. The channel was once again pre-washed with a Tris-EDTA buffer (TE; 10 mM TRIS and 1 mM ethylenediaminetetraacetic acid, EDTA,) and excess of buffer was taken out. After these steps, the stained DNA containing sample could be loaded.

Once loaded, the N_2 pressure-driven flow was started by applying pressure in the loaded well. In this way, DNA molecules were flushed inside the microchannel. To be able to force DNA molecules in the nanochannels, pressure had to be applied on both sides of the microchannel. Once in the nanochannels, they were imaged as described in 3.8.2.

10X 24h

The 10X 24h sample was made into a 1:5 dilution with milliQ water. 10 µL of the dilution was stained with YOYO in the same manner as described in section 3.8.1. After staining, 3 v/v% of 2-mercaptoethanol (BME) was added as a reducing agent. 5 µL of this mixture was loaded in one of the wells.

4. Results and Discussion

The main aim of the project was to make long double-stranded DNA molecules with a specific repetitive sequence. An additional aim was to incorporate fluorescent base analogues at specific places in the sequence to make the DNA inherently fluorescent. This Results and Discussion part is an overview of the most relevant experiments done to pursue these goals.

4.1. RCA experiments

The RCA protocol is the starting point of this project, and it was therefore important to make sure that this amplification was working well for these specific oligonucleotide sequences. Aminoallyl-dUTP-ATTO 647N can be incorporated into the DNA molecule by a DNA polymerase as a substitute for dTTP. Adding labelled dUTP thus gives a method to check if the amplification is working: if labelled dUTP is incorporated in the ssDNA molecules, an overlap will be visible between the YOYO emission and the emission of ATTO 647N. The RCPs collapse in blobs of DNA and are therefore visible as dots, typically at a size of around 1 µm[67].

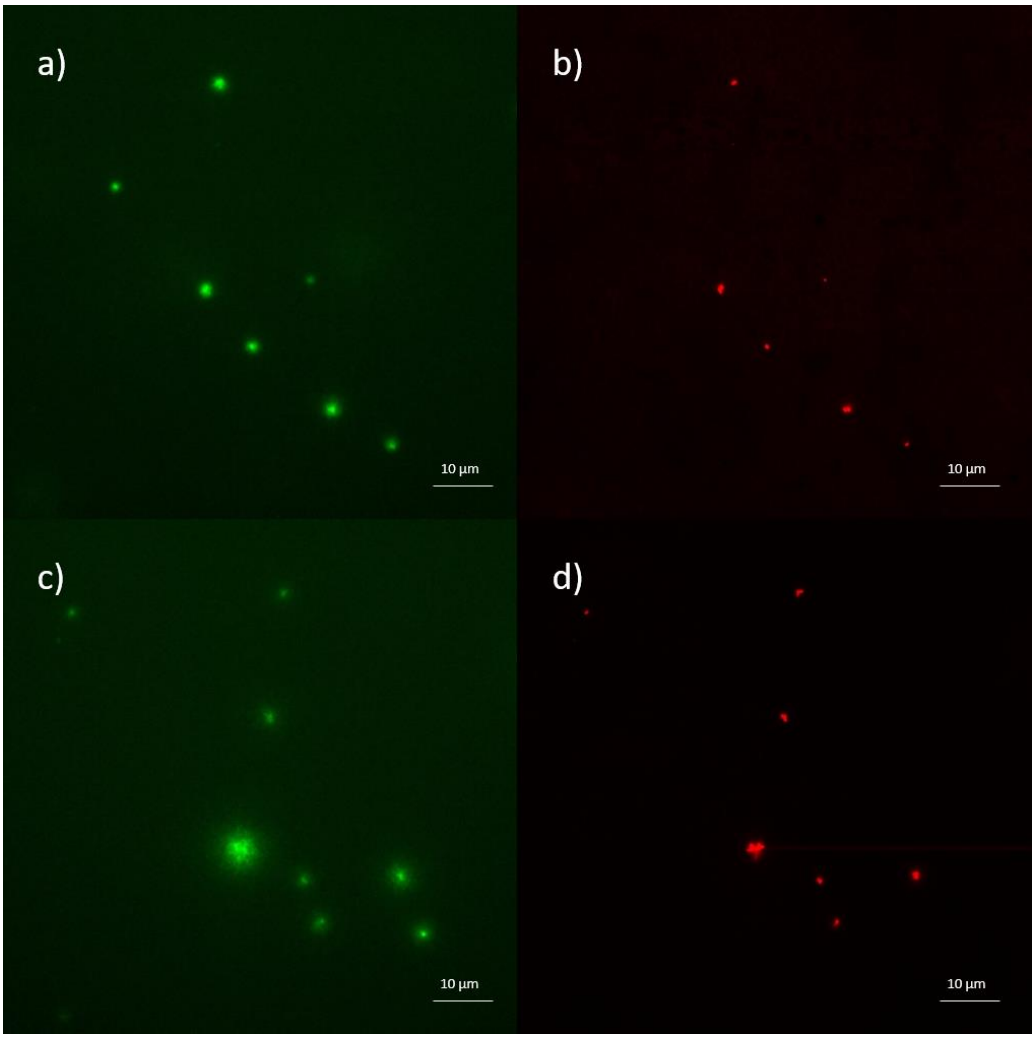

Figure 22. Fluorescence microscope image of RCA products on a positively charged slide. a) The emission of YOYO from the RCPs is visible. b) The emission of ATTO 647N dye is visible from the same RCPs as in figure a. c) Shows the emission of YOYO from several RCPs. d) Shows the emission of ATTO 647N dye from the same RCPs as in figure c.

In Figure 22 an overlap is visible between the emission of YOYO and the labelled dUTP, which means that the dUTP analogues get incorporated into the ssDNA molecules during the RCA procedure. Almost all RCPs show both emission of YOYO and emission of ATTO 647N. The sample was investigated on a positively charged slide.

Although the RCA was based on an already existing and often used protocol, not all the results were as expected. An observation was that the general yield of the RCA procedure was lower. Even though the overall amount of RCPs in the 10X samples was higher than in the 1X samples, it was still low when comparing it to other RCPs starting from different oligonucleotides. The RCPs were present in the 1X samples, but the amount was very low, which made it more difficult to detect. The most logical explanation would be that the ligation reaction to form the circular DNA molecules was not working optimally. This could be due to the design, length and/or sequence, of the arm of the PLP, because the reference sample used had a PLP length of 80 nt, instead of the 68 nt used here.

4.2. Stretching on Positively Charged Slides

The long ssDNA molecules obtained by the RCA protocol were used as the starting point to create dsDNA molecules. The main question was: is it possible to obtain dsDNA molecules and how many steps are necessary to obtain them? This was evaluated by a series of experiments listed below, which will be discussed in the following sections.

- Annealing of complementary strands with a ligation reaction to seal the nicks (Experiment 1).
- Annealing followed by an extra polymerisation step and ending with a ligation reaction to seal the nicks (Experiment 2).
- Running the amplification step of the RCA protocol for 24 hours (Experiment 3).

All three experiments were tested on positively charged slides first.

4.2.1. Experiment 1

In the ideal situation the long RCPs should be able to form dsDNA when annealed with two short complementary strands (34 nt), that fill up the whole repetitive sequence, followed by a ligation reaction to seal the nick between the different shorter strands bound to the RCPs. This way of obtaining the dsDNA was therefore attempted first, as it was also thought that this method would be the most useful for achieving the native-like fluorescent dsDNA molecules with tC.

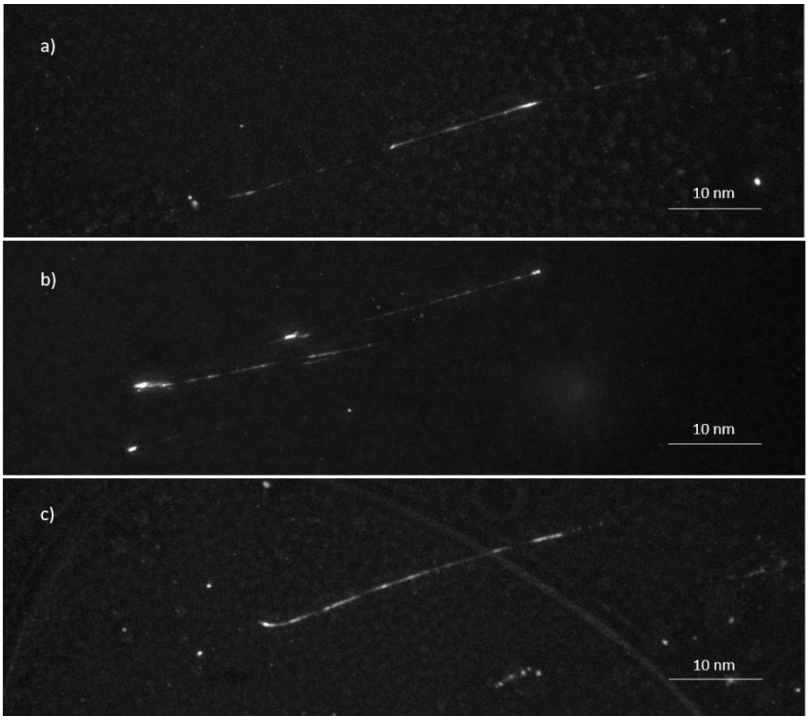

Figure 23. Fluorescence microscope images of DNA strands, stained with YOYO, from Experiment 1 stretched on positively charged slides. The samples showed here are all 10X 1h samples. a) Annealing procedure: 95°C and 1 °C/min cool-down. b) Annealing procedure: 80°C and 1 °C/min cool-down. c) Annealing procedure: 65°C and 1 °C/min cool-down.

The stretched molecules from the samples seen in Figure 23 all look similar; they show a heterogeneous staining and an overall weak fluorescence signal. Different types of annealing processes were attempted varying the temperature and the cool-down rate. Annealing procedures going up to 95°C showed precipitation in some cases, and were thus avoided in the following samples. Different concentrations of PLP and target were also investigated (1X, 2X, 5X and 10X). The 1X had only a small amount of stretched molecules, as was expected from the results in section 4.1. However, none of the other variations resulted in a notable improvement when viewed on the positively charged slides.

A stretched dsDNA molecule stained with YOYO should show a homogeneously bright staining over the length of the molecule. From the results shown in Figure 23, and mainly the heterogeneous staining, it was concluded that the annealing procedure fills only parts of the strands and that these parts are more fluorescent than the ssDNA parts in between (Figure 24). Therefore a polymerisation step was added in Experiment 2 to try to fill up these gaps.

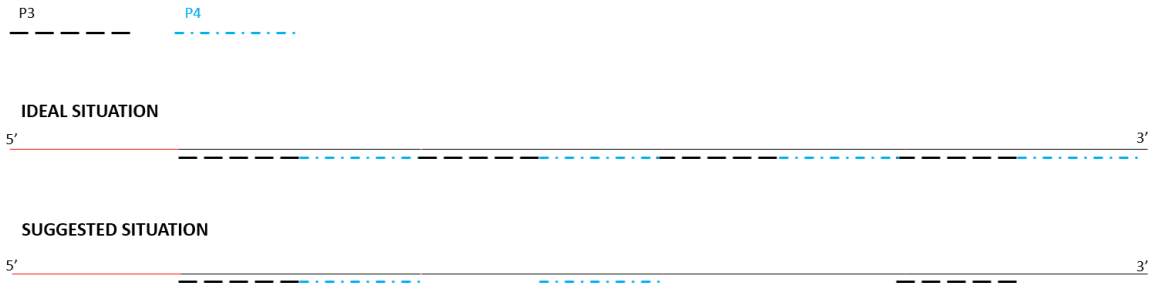

Figure 24. Schematic representation of what would have been the ideal situation following Experiment 1, namely the complete filling of the RCPs with small complementary strands primer 3 (P3) & primer 4 (P4) and the situation suggested by the stretching results shown in Figure 23.

The number of base pairs of the DNA molecules was not accurately determined, as that would require the addition of a reference sample containing dsDNA with a known length that can be distinguished from the original sample. However, a rough estimation of the number of base pairs of the DNA molecules suggested that the RCA process was quite successful in creating long (kbp) molecules.

It is important to note that the number of bases visible in one pixel is much higher than the 34 nt long complementary sequences used. An estimation of the amount of base pairs in one pixel is roughly 300 bp. If the assumption, that the brighter spots indicate dsDNA, is correct, it would mean that the brighter regions contain multiple complementary strands in close proximity to each other.

4.2.2. Experiment 2

The results from Experiment 1 deemed it necessary to have an additional polymerisation step. In Experiment 2 this polymerisation reaction was added and the annealing procedure was kept at a maximum of 80°C, to avoid precipitation, with a slow cool-down of 0.5 °C/min, to maximize the number of complementary strands that can be annealed. Unexpectedly, the results looked very similar to the results from Experiment 1 (Figure 25). From the picture it is obvious that there is a distribution of different sized strands, but none of them shows a homogeneously bright staining.

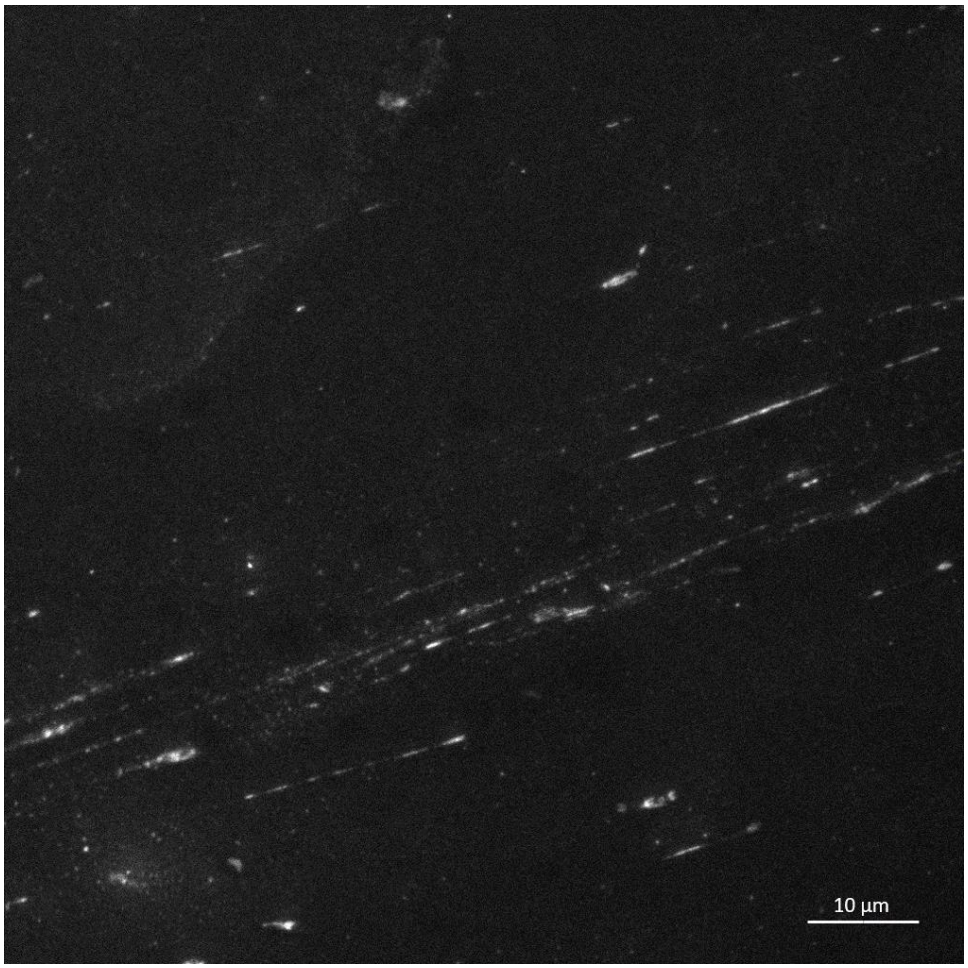

Figure 25. Fluorescence microscope images showing the result of Experiment 2 stretched on positively charged slides. A 10X 1h sample is shown here, fluorescently labelled with YOYO. The strands appear heterogeneously stained and a wide variety of different sized molecules is visible.

In theory, a sufficient amount of DNA polymerase I and high excess of dNTP's had been added, so complete formation of double-stranded products was expected. A potential reason for not stretching and bad staining could be the high salt concentration in the sample. However, no noticeable improvement was observed when the samples were diluted with milliQ to lower the salt concentration, indicating that this was not the cause.

Even though the annealing temperature was kept at a maximum of 80°C, some precipitation was still observed as milky threads in the solution (Figure 26b). Sometimes it was already visible somewhere along the experiment, for example after the polymerisation reaction, and sometimes it was only visible at the very end. It is possible that this precipitate is dsDNA, but it is unclear what brings about this precipitation. This hypothesis is strengthened by microscope images showing fluorescent structures that look like they are surrounded by small fluorescent strands, which could indicate dsDNA (Figure 26a).

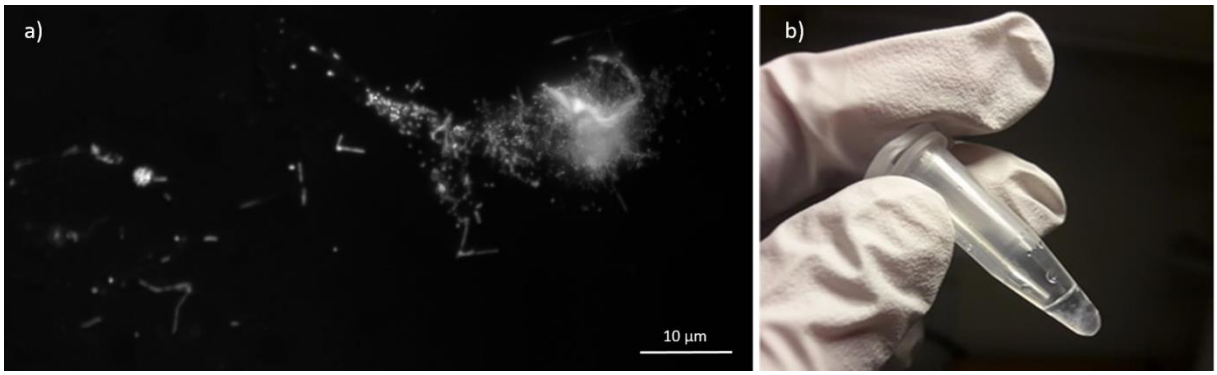

Figure 26. a) Fluorescence microscope image of small fluorescent strands in close proximity to a bigger fluorescent molecule, showing small fluorescent lines that could indicate dsDNA strands. b) An example of the precipitation observed in the reaction tubes after the completion of Experiment 2.

4.2.3. Experiment 3

For Experiment 3 the amplification step was prolonged to 24 hours. A paper in 2014 by Ducani et al. discussed that the RCA procedure only obtains purely ssDNA rolling circle products, it was necessary to add a single-stranded DNA binding protein[69]. They claimed that otherwise ϕ29 DNA polymerase started to produce double-stranded DNA. After one hour it was reported that the RCPs would consist out of 95% ssDNA and 5% dsDNA and after increasing the incubation time to 72 hours it would almost completely consist of dsDNA.

Though this way of achieving dsDNA was not useful for the incorporation of the tC base analogue, it was carried out to see if these finding were also applicable on the oligonucleotides used in this project.

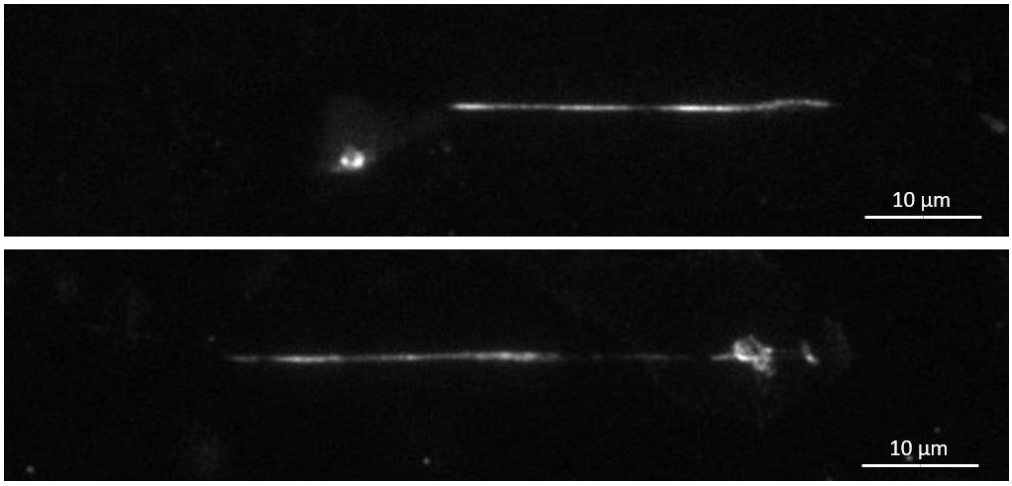

Figure 27. Fluorescence microscope images from Experiment 3 portraying stretched DNA molecules of a 10X 24h sample. The sample was stained with YOYO and shows a heterogeneous staining.

When running the amplification reaction for 24 hours, the sample gets highly viscous indicating a high concentration of dsDNA. Notable here is that precipitation was, however, never seen. Figure 27 shows fluorescence microscope images of the results of Experiment 3. The fluorescent staining is a little better than that of Experiment 1, but it is still too heterogeneous. With these images, it was not possible to conclude if these heterogeneous stained structures were dsDNA, even though, the fact that they were stretched out suggests that they are. This suggested to stop the use of the positively charged slides, since it was previously reported that dsDNA would be produced when the amplification period was extended[69].

4.3. Stretching on Activated Coverslips

Every experiment that was investigated on the positively charged slides gave approximately the same unexpected result. It was therefore decided to test some of the samples on regular, uncharged glass slides with activated coverslips. Through this stretching method, it was possible to show that many of the strands obtained from Experiment 2 and 3 were in fact mostly double-stranded. This raised a few questions: Were all the different steps in Experiment 2 necessary? Is Experiment 1 enough to form dsDNA molecules?

On these slides the different experiments were tested backwards. First old samples from Experiment 2 and Experiment 3, that had been stored at -20°C, were tested again. When they appeared to contain dsDNA, Experiment 1 was carried out again to answer if those conditions were sufficient as well. It was important to try Experiment 1 again, because the annealing procedure would be the most suitable for the incorporation of tC base analogues, and thereby addressing the second aim of this study.

4.3.1. Experiment 2

When loading the stored samples of Experiment 2 (10X 1h), they appeared to have a high concentration of dsDNA with a broad size distribution. As is seen in Figure 28, the stretched out DNA strands are now homogeneously stained. In the right top corner of the middle picture (b) a strand is visible that is consisting of a bright part and a less bright part. A possible explanation could be that the strand is not fully double-stranded.

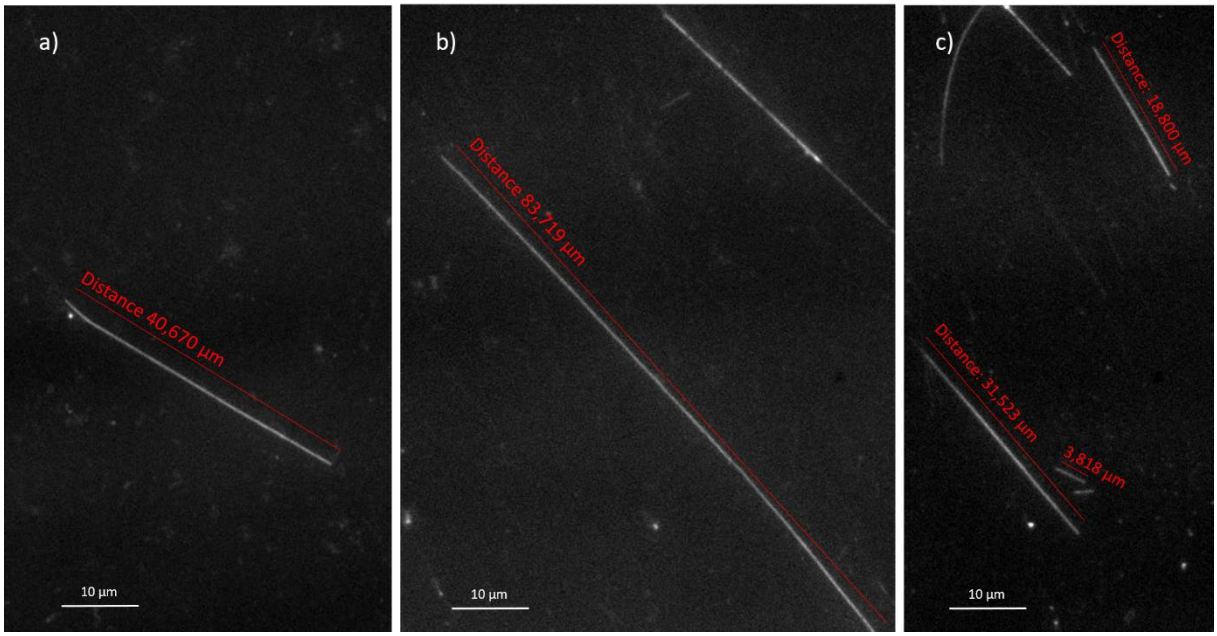

Figure 28 Fluorescence microscope images of one 10X 1h sample, from Experiment 2, stretched on activated coverslips and fluorescently labelled with YOYO. The red lines indicate the length of the displayed part of the molecules in µm. In image b this is a underestimation of the total length of the strand, as a part of the molecule is not shown in the image.

The sizes indicated here in red are ranging from 3.8 µm to at least 83.7 µm, which shows the broad size distribution. The 83.7 µm strand indicates a large dsDNA molecule and the length displayed is an underestimation of the actual length as only part of the strand is visible in Figure 28b. Although the exact size in base pairs is unknown, a rough estimation in length can be made when using λ-DNA. The DNA molecules from λ-DNA are of known length, 48,502 base pairs. The strands were not imaged and measured in the exact same buffer as was used during this experiment, but in a 0.5× TBE (Tris base, boric acid and EDTA) buffer. In the 0.5× TBE buffer, the λ-DNA molecules had an average length of 17 µm, which meant that approximately 2850 bp were visible per µm (V. Singh, personal communication). This value gives a rough estimate for the length of the molecules, and indicates that the size distribution shown in Figure 28, is ranging from 11 to more than 230 kbp. To make a better estimation, it is necessary to image and measure the λ-DNA under the same conditions as the sample itself.

When the incubation time of the amplification step was lowered to a half hour, the overall stretched strands appeared to be shorter. Figure 29 shows two of the longest strands found after that experiment, which are less than 20% of the length of the longest strands found in the one hour amplification samples.

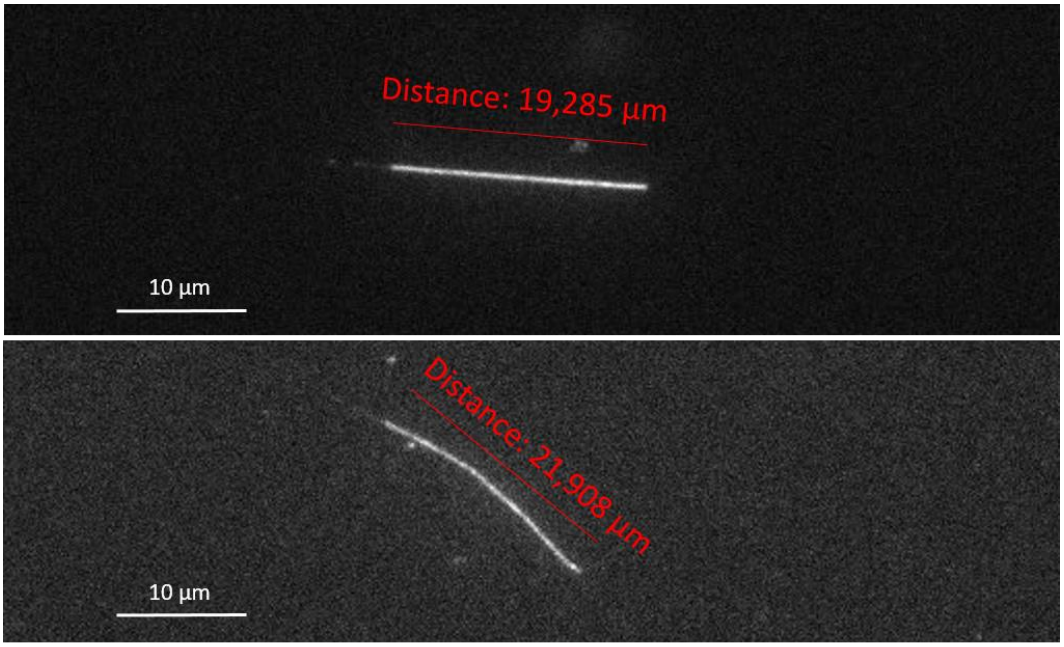

Figure 29. Fluorescence microscope images showing two of the longest stretched DNA strands on activated coverslips found in a 10X 0.5h sample when stained with YOYO. The lengths of the molecules are indicated by the red lines.

The results show that Experiment 2 is working. These dsDNA molecules are all sequence specific repetitions of the IHF protein binding site, except for the first 40 base pairs coming from the target and perhaps some mutations, but the frequency of erroneous base pairs is low. Out of these results, however, it was unclear how many bases in the complementary strand there were after the annealing procedure and how many there were due to the dNTP incorporation by DNA polymerase I. To resolve this unanswered question, Experiment 1 had to be repeated.

When stretching the samples of Experiment 2 on the activated coverslips, it gave a better result compared to the positively charged slides, but it came with a challenge as well. Not all slides gave the same result of stretching when testing the exact same sample. It could happen that most DNA strands on the slide were still coiled up. It is unclear why coverslips from the same batch can differ significantly in performance. It could be because they require a careful handling and need to be rinsed thoroughly to avoid any particles sticking on the surface. One small dust particle, for example, could be a large disturbance for the stretching process. Unfortunately, these kind of contaminations are only visible when the sample has already been applied, showing no, or only a few, stretched strands together with most DNA molecules still coiled up in solution.

4.3.2. Experiment 3

When stretching the refrigerated samples of Experiment 3, it was shown that a similar result can be achieved with this method. The high viscosity reported in section 4.2.3 had already indicated that the concentration of DNA could be significant in this sample, but this was not proved with the experiments on positively charged slides. When applying the undiluted sample on these activated coverslips, however, it showed a 'web' of stretched out strands (Figure 30), which corresponds to a high concentration of DNA.

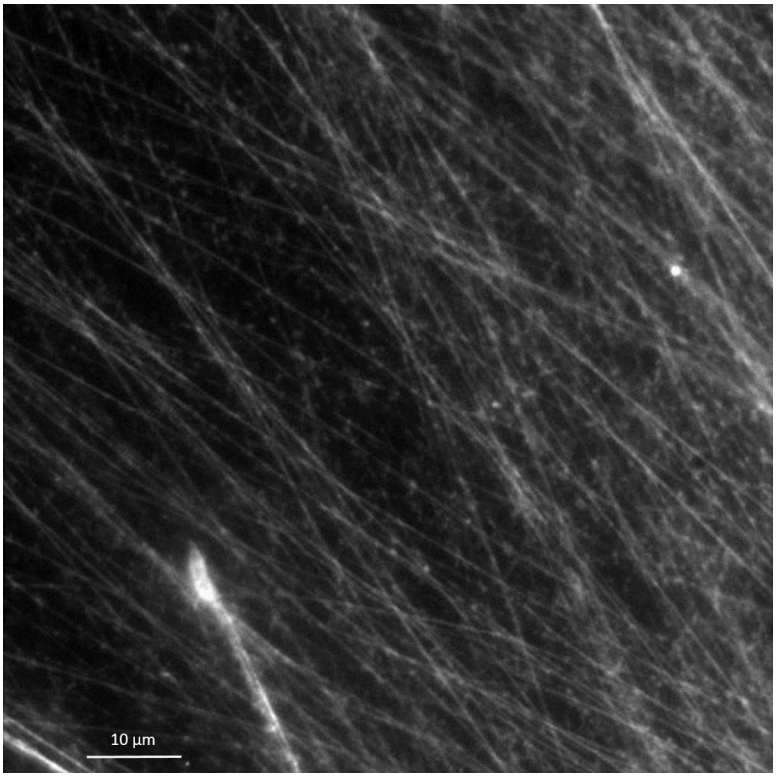

Figure 30. Fluorescence microscope image of a 10X sample that was amplified for 24 hours. A high concentration of long dsDNA molecules is visible, when stained with YOYO, and they form a 'web'-like structure on the slides, due to the different strands stretching in an overlapping manner.

To obtain better images and a better indication of the length of the strands, the sample had to be diluted with milliQ water. The dilution series performed were 1:5 and 1:10. Some of the strands found in those dilutions appeared to be almost four times larger on the slides compared to the molecules found in Experiment 2.

The fact that Experiment 3 works, means that not as many reaction steps are needed to arrive at dsDNA molecules starting from the RCA protocol. The downsides of using Experiment 3 is that there is barely any control of the size distribution of the strands and that the mixture at the end of the experiment could still be a combination of ssDNA and dsDNA, as was reported by Ducani et al[69]. if the strands are as big as they appear to be, this is also not preferred when studying DNA-protein interaction in nanochannels, as they might be too long to be stretched in one channel. In Experiment 2, this size distribution is presumed to be less wide and the amount of ssDNA, if still present, is expected to be lower, although it is not known how big the percentage of ssDNA is in either one of the experiments.

An extra disadvantage to Experiment 3 is that it cannot be used to form dsDNA with tC base analogues incorporated at a certain position. Some polymerases are capable of incorporating base analogues, but these incorporations will not be sequence specific. This means that all cytosine nucleobases could be replaced by a tC base analogue in the dsDNA molecules. When the ϕ29 DNA polymerase starts to produce the complementary strand, with tC base analogues in the template, it will probably be that the polymerase favours incorporating of dGTP over the incorporation of dATP, but the error-rate will be higher than during the incorporation of naturally occurring nucleobases[70].

4.3.3. Experiment 1

The first experiment was re-done by adding complementary strands P3 & P4 to 10X 1h RCA samples, using two of the different annealing protocols, namely protocol 4 and 5 explained in section 3.5. After the annealing, a ligation reaction was added. In contrast to what was concluded in section 4.2.1, it looks like the strands are homogeneously stained and they appear to be fully double-stranded (Figure 31).

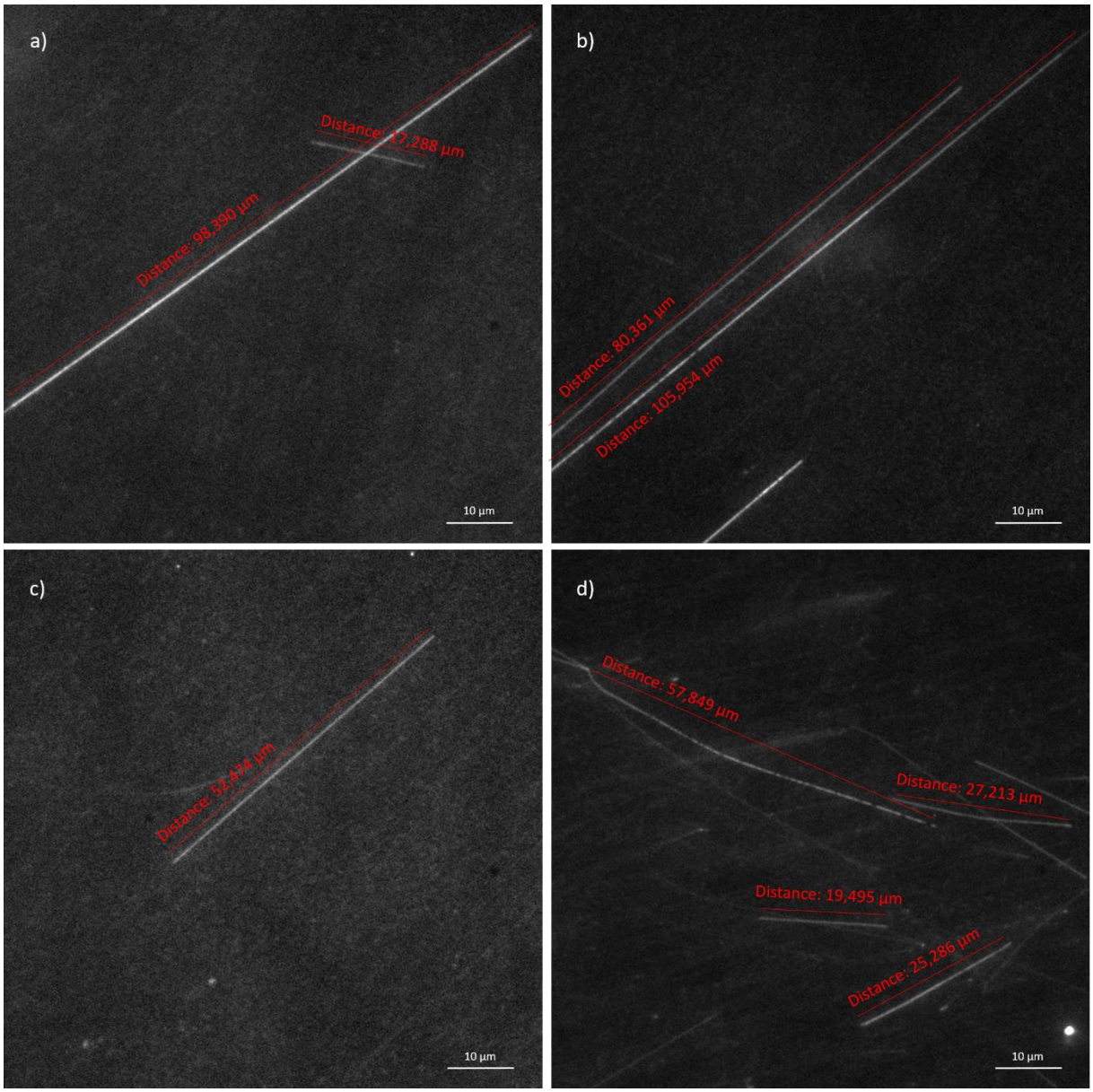

Figure 31. Fluorescence microscope images of stretched DNA molecules obtained by Experiment 1, after being stained with YOYO. a) and b) show DNA molecules obtained from the annealing protocol with a heat step at 70 °C for 5 min followed by a hybridisation step at 55 °C for 15 min and a cool-down of 0.5 °C/min. c) and d) show DNA molecules obtained from the annealing protocol with a heat step at 80 °C and a cool-down of 0.5 °C/min. The red lines indicate the length in µm of the parts of the molecules shown.

Figure 31 shows the two different annealing protocols used; a & b show the results obtained with the annealing protocol 5 and c & d show the results from annealing protocol 4. On the basis of the microscope images obtained from this experiment, no noticeable differences could be observed

between the two different annealing procedures. Annealing protocol 5 involves a heat step at 70 °C and a hybridisation at 55 °C and has been used in previous research involving RCPs to specifically label the rolling circle products to determine their presence and quantity. This specific labelling is achieved with a shorter, externally labelled strand, complementary to a part of the repetitive sequence. It is reported that approximately 500 to 1000 fluorophores are annealed to the RCPs obtained from one hour amplification (± 1000 repeats), but that saturation is more often reached at 700 annealed complementary strands[67]. This indicates that it is not likely that the full length of the RCPs gets annealed with their smaller complementary counterparts. Images b and d both show a strand with parts of the molecule being less fluorescent, these parts could indicate areas where there is still ssDNA present. As previously reported, it is important to keep in mind that the resolution of the images only allows to see approximately 300 base pairs in each pixel, meaning that it is not possible to tell whether these homogeneously stained strands in images a and b (Figure 31) are completely dsDNA. It is possible that gaps remain present, due to the annealing process not being perfect. Two different options can be tried in the future to address this, namely damage repair with the ATTO 647N labelled dUTP and loading them in the channels, where single-stranded parts would appear more flexible.

The sizes indicated by the red lines in Figure 31 are in agreement with the other experiments reporting a size distribution. In the figure, the sizes roughly range from 49 to more than 300 kbp.

It is concluded from these results that Experiment 1 looks promising and that the annealing process has a significant part in the making of the dsDNA molecules, but that it is still uncertain whether these strands are fully double-stranded. In Experiment 2 this can be concluded with more certainty due to the added polymerisation step.

4.4. Activated Coverslips Compared To Positively Charged Slides

The positively charged slides are charged in a chemical way, by an adhesive formula from VWR. The charge is applied over the surface of the glass slides.

The mixture of acetone, APTES and ATMS activates the glass surface of the coverslip by covalently binding positively charged amino groups. These positive charged groups attract and bind the negatively charged backbone of dsDNA.

The most significant difference between the two types of charging is that the activation of the coverslips was always done right before the actual stretching of the DNA samples, because it had often already lost its stretching ability a day after being activated. The difference between the method of charging or the chemical composition of the surface layer cannot be addressed, due to the adhesive formula in the commercial slides that is not public information.

4.5. Nanochannels

Due to the long dsDNA molecules found in the samples from Experiment 2 and 3, it was attempted to load a sample of Experiment 3 in the nanochannels. This resulted, in a non-coated nanofluidic device, in most of the DNA molecules sticking to walls of the loading well and clogging up the channel, meaning that applying a lipid coating was necessary[71]. When the coating was applied, as reported in 3.9.3, the DNA molecules could move free in solution.

It was observed in the microchannel, in agreement with what was reported on the activated coverslips, that the sample consisted of a broad range of different sized DNA molecules. Figure 32 displays some of the microscopic images taken from different dsDNA strands trapped in a nanochannel. The molecules range in Figure 32 from 3.2 to 14 µm. To make a rough estimation of the length of these DNA strands a reference can be used in the same manner as the reference for the activated coverslips, namely stretching λ-DNA in a 0.5× TBE buffer. In similar sized channels the average length of the λ-DNA was reported to be around 6 µm under these buffer conditions[72]. This results in 8100 base pairs present in one µm, suggesting that the observed strands range from 26 to 113 kbp. This was shorter than was expected when viewing the 10X 24h sample on the activated coverslips. The coverslips might overstretch the DNA, but also the ionic concentration of the sample has an, not to be ignored, influence on the extension of the DNA strands in nanochannels. A higher dilution of the sample will lower the ionic concentration. At lower ionic strengths the DNA will be more extended[73]. Thus it is important that a reference sample is obtained in the exact same conditions to get a more reliable estimation in length.

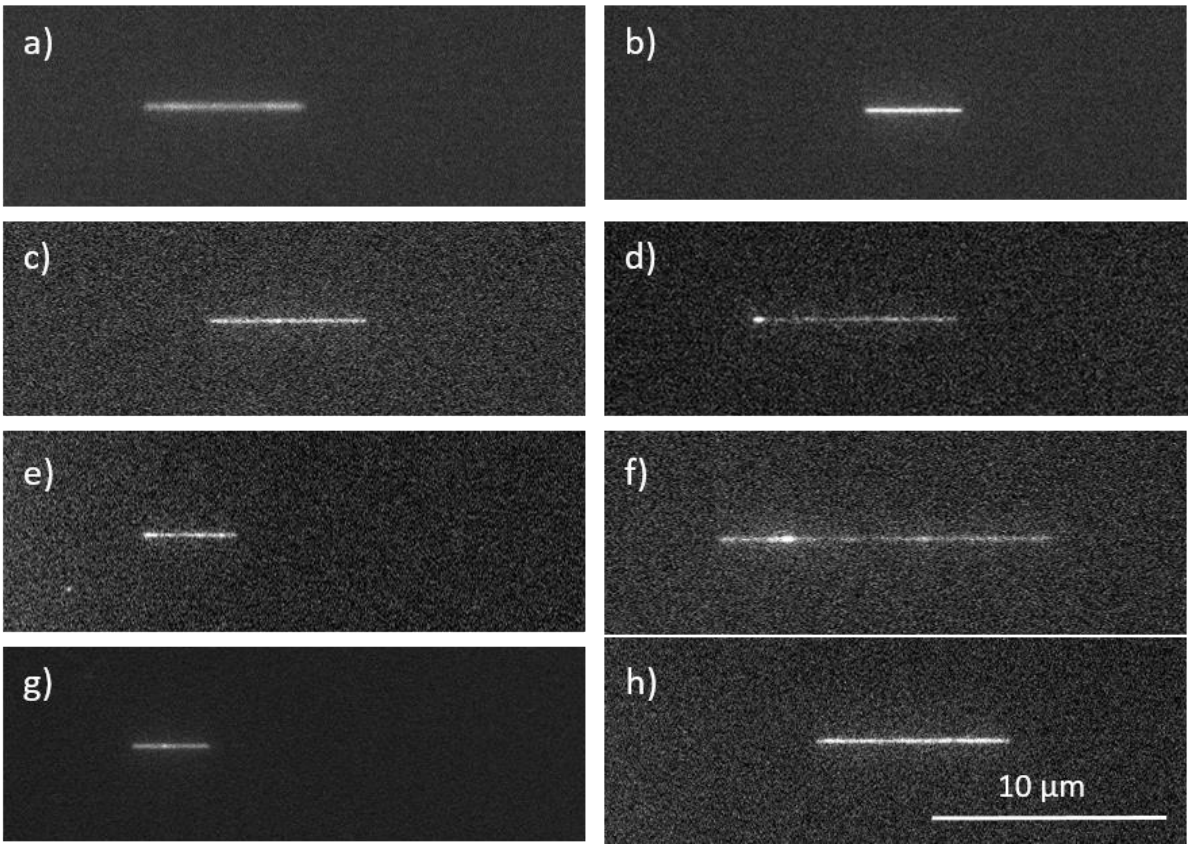

Figure 32. Fluorescence microscope images of stretched dsDNA molecules in the 150x100 nm² nanochannels. All molecules portrayed here come from the same 10X 1h sample of Experiment 3, after being fluorescently labelled with YOYO. a, b, c, e, g and h show a homogeneously stained molecule, but d and f show a heterogeneously stained DNA strand.

The manner of staining (heterogeneous/homogeneous, bright/subdued) was varying for different molecules in the same sample. As is seen in Figure 32, heterogeneous staining was present in the nanochannels as well (d and f). It is unsure why this remains. The movement of the molecules and the capability to be forced in the nanochannels, strongly indicates that the molecules were dsDNA. It has been reported that, even after a 24 hour incubation at 50 °C, the DNA molecules will still display a distribution in fluorescence emission between different molecules[73], making the difference between

images b and c as expected. However, the heterogeneous staining in one molecule (seen in images d and f) has not been reported. A possible explanation could be that the proteins present in the samples were still attached to some parts of the molecule and this is interfering with the bis-intercalation of YOYO. Purification of the sample might indicate whether this explanation is correct. The addition of this step would also be advised before studying the DNA-protein interactions with IHF. The downside of a purification step, however, is that there is a risk of cutting up the DNA in smaller pieces.

4.6. dsDNA Molecules with tC Base Analogues

The promising results from Experiment 1, stretched on activated coverslips, increased the probability of obtaining kbp sized, internally labelled dsDNA molecules by using this method. When repeating the experiment with tC base analogues, it appeared that the tC molecules present were not fluorescent enough to make the strands visible in the settings used. If the annealing process is perfect, there should be four tC base pairs present in every repetitive sequence of 68 nt. Taking into account that every pixel consists of approximately 300 bp, it would mean that 17 tC base pairs should be present in every pixel. This was hoped to give a sufficient amount of emission, but the quantum yield of tC (0.19) and the extinction coefficient (13500 mol $L^{-1}cm^{-1}$) are lower than those of YOYO (0.35 and 96100 mol $L^{-1}cm^{-1}$ respectively), making the total brightness of YOYO at least 13 times stronger compared to that of tC. Important to mention is that the emission filter used could not capture the full emission spectrum of tC. In the case of a weaker probe, this causes a significant loss in captured emission. The use of a different filter, would probably increase the brightness of the captured emission and is recommended in further research.

When staining these samples with YOYO, the DNA strands became visible, but the staining was more heterogeneous and less bright than other samples tested on the activated coverslips (Figure 33). It is unknown why this is the case, but it shows that it is highly possible that the tC containing strands are annealed to the RCPs.

It is reasonable to think that the tC base analogues are present in the stretched molecules, but that the internal labelling is not strong enough to make the strands visible with the detection method used in this study.

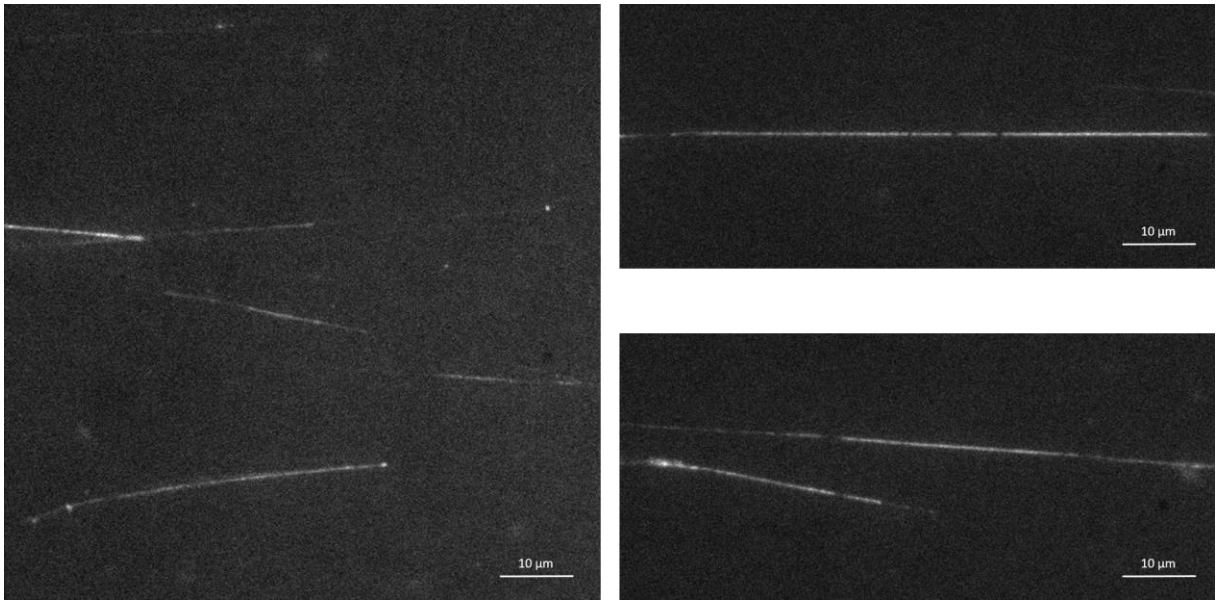

Figure 33. Fluorescence microscopy images of YOYO stained DNA strands containing tC molecules. The sample portrayed follows Experiment 1.

5. Conclusion

Several conclusions can be drawn based on the results and discussion described in the previous chapter.

The RCA protocol used in this study is adoptable for the designed oligomer sequences, but the yield of RCPs was lower when comparing them to the RCPs of different starting PLP and target sequences.

The outcome of this project demonstrates that the manually activated coverslips provides a better stretching and therefore improved analysability compared to the commercially available positively charged slides. However, the reason is not fully understood.

Experiment 1, in which only the annealing and ligation reaction was present, showed promising results, but more research will be needed to find out if those alone are enough to obtain fully double-stranded DNA molecules. However, because the molecules showed parts of bright and homogeneous staining, it can be concluded that the annealing protocol plays an important role in the making of the dsDNA molecules.

When adding an extra polymerisation step before the final ligation, as was done in Experiment 2, it was seen that the resulting molecules look similar to those of Experiment 1, in which no polymerisation step was used. However, there is a higher probability that these DNA molecules are double-stranded, due to the use of DNA polymerase I, which should be capable of filling in the potential gaps left behind in the annealing process. The roughly estimated lengths varied from 11 to more than 240 kbp. It was shown that the length of the DNA molecules can be varied by adjusting the reaction time in the amplification step of the RCA procedure, giving a control over the end result.

If the amplification step of the RCA is prolonged to 24 hours, it was shown that the ϕ29 DNA polymerase starts to produce dsDNA molecules, as previously reported by Ducani et al[69]. However, it is still unsure what drives this formation. When stretched on the charged coverslips, some strands appeared to be multiple times larger than reported for Experiment 2, but the molecules stretched in the confinement of nanochannels indicated strands up to 110 kbp. The possible control, reported in Experiment 2, is completely lost due to this prolonged amplification step and this is expected to give the sample a broader size distribution.

The methodology described in Experiment 2 is preferred over that of Experiment 1 and 3 with the initial aim of this project in mind; It gives a higher probability of dsDNA molecules than in Experiment 1 and when compared to Experiment 3, it retains the feature to control the length of the obtained molecules. Also, the size distribution is assumed to be narrower, because it is unsure when the ϕ29 DNA polymerase stops producing ssDNA and starts producing dsDNA molecules, and the percentage of ssDNA molecules is most likely smaller.

It can be concluded from the results in this thesis that a rolling circle amplification-based methodology can be used to obtain kbp sized, double-stranded DNA molecules, which consist of numerous repetitions of the specific binding sequence of IHF. It was also shown that more than one RCA-based method was capable of achieving this. These findings provide new possibilities to study DNA-protein interactions, and are expected to be of particular importance for techniques requiring long DNA molecules.

The internal labelling of the dsDNA molecules with tC base analogues did not yield products that could be imaged using the applied detection techniques. This was thought to be due to, at least in part, the emission filter used that did not display a perfect overlap with the emission of tC, possibly in combination with a shortage of tC molecules present in the strands to make the whole DNA molecule fluorescent. More research is needed to draw more relevant conclusions concerning this, but staining the same sample with YOYO and imaging this, confirmed the presence of DNA molecules, which indicate that the method has potential for the second aim of the project.

6. Future Outlook

This chapter presents possible further steps that might be able to help to achieve the end goal of this thesis, namely DNA molecules for improving the study of DNA-protein reactions.

The limited time gave no room for a complete optimisation of the protocol. The starting sequences gave a lower yield than was expected and to increase this, it is advised to increase the size of the PLP arms to improve annealing. Every component added, in the different reaction steps was in high excess to ensure an adequate amount of product strands and a reasonably fast reaction. All the reactions were incubated for an amount of time that would ensure product, but it might well be that these periods of time in the ligation and polymerisation steps were unnecessarily long or maybe even too short to seal all the nicks or convert all the ssDNA to dsDNA respectively. The stretching of the samples on activated coverslips gave a significant improvement compared to the positively charged slides, but this method also has to be optimized to make it more robust and reliable. It is advised to move on from stretching on glass to stretching in nanochannels.

The samples are assumed to still have both ssDNA and dsDNA in solution for every attempted experiment. The addition of S1 nuclease could be a way to get rid of the ssDNA molecules that are not wanted in the end. S1 nuclease is a known single-strand-specific endonuclease, meaning that it hydrolyses the still single-stranded regions in dsDNA molecules and thus removes them. The downsides of this is that it adds another step to the protocol and can cut up the DNA into smaller pieces if they are not completely double-stranded.

The method of estimating the length of the DNA duplexes in the Results and Discussion section is imprecise. A reference sample of dsDNA molecules with a known length, such as λ-DNA, should be added to make a reliable estimation of the length of the dsDNA molecules in the samples. This would be a valuable asset to show that the protocol is in fact capable of producing such long lengths and it would give the possibility to make a more accurate indication of the size distributions present in the experiments.

Nanochannels are already presented as an alternative way to make the stretching of the dsDNA molecules more reproducible. The intention was to detect and measure the molecules in nanochannels during this project, but due to time limitations this was only made possible for one sample of one experiment, namely Experiment 3. To be able to study the molecules in the type of nanofluidic devices presented in the Nanofluidics and Nanochannels section, it might be necessary to purify the mixtures to avoid clogging of the channels. A method has to be found to do this purification at the same time as minimizing the breaking of the dsDNA molecules. To avoid the strands from sticking to the channel walls, coating of the channels with a lipid layer is recommended. The use of nanochannels will also be favourable when studying the interactions between the dsDNA molecules and the IHF protein.

After the optimization and the commissioning of the nanochannels, the molecules should be tested how and if they interact with the IHF protein. If these interactions are in fact readily observable, it opens the possibility to apply the protocol to different sequence specific proteins. If the protocol shows to be robust for different DNA-protein interactions, more studies can follow to obtain new insights in, for example, several diseases induced by these interactions.

More research is most definitely needed to conclude whether the protocol can be applied for the internal labelling of the molecules with base analogues, to avoid using bis-intercalator dyes such as YOYO. When working further with fluorescence microscopy as a detection method, it would be advised to work with a long pass 450 nm filter. If this does not prove helpful for the detection of tC labelled strands, it might be that the fluorescence microscope and/or settings, used in this project, are not sensitive enough to pick up the weak emission of the tC base analogues. Further optimisation will have to be performed.

7. Bibliography

[1] M. W. Gonzalez, M. G. Kann, *PLoS Comput. Biol.* **2012**, *8*, DOI 10.1371/journal.pcbi.1002819.

[2] T. Ideker, R. Sharan, *Genome Res.* **2008**, *1*, 644–652.

[3] M. G. Kann, *Brief. Bioinform.* **2007**, *8*, 333–346.

[4] C. B. Reese, *Org. Biomol. Chem.* **2005**, *3*, 3851.

[5] K. Drlica, J. Rouviere-Yaniv, *Microbiol. Rev.* **1987**, *51*, 301–319.

[6] M. Fakruddin, K. S. Bin Mannan, A. Chowdhury, R. M. Mazumdar, M. N. Hossain, S. Islam, M. A. Chowdhury, *J. Pharm. BioAllied Sci.* **2013**, 245–252.

[7] M. G. Mohsen, E. T. Kool, **2016**, *12*, DOI 10.1007/s11897-014-0247-z.Pathophysiology.

[8] D. Liu, S. L. Daubendiek, M. A. Zillman, K. Ryan, E. T. Kool, *J. Am. Chem. Soc.* **1996**, *118*, 1587–1594.

[9] J. H. Kim, M. Jang, Y.-J. Kim, H. J. Ahn, *J. Med. Chem.* **2015**, *58*, 7863–7873.

[10] M. M. Ali, F. Li, Z. Zhang, K. Zhang, D.-K. Kang, J. A. Ankrum, X. C. Le, W. Zhao, *Chem. Soc. Rev.* **2014**, *43*, 3324.

[11] V. V Demidov, *Expert Rev. Mol. Diagn.* **2002**, *2*, 542–548.

[12] X. Chen, B. Wang, W. Yang, F. Kong, C. Li, Z. Sun, P. Jelfs, G. L. Gilbert, *J. Clin. Microbiol.* **2014**, *52*, 1540–1548.

[13] W. Zhao, M. M. Ali, M. A. Brook, Y. Li, *Angew. Chemie - Int. Ed.* **2008**, *47*, 6330–6337.

[14] M. Guo, I. Hernández-Neuta, N. Madaboosi, M. Nilsson, W. van der Wijngaart, *Microsystems Nanoeng.* **2018**, *4*, 17084.

[15] L. Zhou, L.-J. Ou, X. Chu, G.-L. Shen, R.-Q. Yu, *Anal. Chem.* **2007**, *79*, 7492–7500.

[16] B. Schweitzer, S. Roberts, B. Grimwade, W. Shao, M. Wang, Q. Fu, Q. Shu, I. Laroche, Z. Zhou, V. T. Tchernev, et al., **2010**, *20*, 359–365.

[17] M. Wilhelmsson, Y. Tor, *Fluorescent Analogs of Biomolecular Building Blocks: Design and Applications*, John Wiley & Sons, **2016**.

[18] P. Sandin, L. M. Wilhelmsson, P. Lincoln, V. E. C. Powers, T. Brown, B. Albinsson, *Nucleic Acids Res.* **2005**, *33*, 5019–5025.

[19] K. C. Engman, P. Sandin, S. Osborne, T. Brown, M. Billeter, P. Lincoln, B. Nordén, B. Albinsson, L. M. Wilhelmsson, *Nucleic Acids Res.* **2004**, *32*, 5087–5095.

[20] M. Cobb, *Curr. Biol.* **2014**, *24*, R55–R60.

[21] J. D. Watson, F. H. C. Crick, *Nature* **1953**.

[22] L. A. Moran, R. A. Horton, G. Scrimgeour, M. Perry, *Principles of Biochemistry*, Pearson (New International Edition), **2014**.

[23] T. A. Steitz, **1999**, 17395–17399.

[24] H.-M. Eun, in *Enzymol. Prim. Recomb. DNA Technol.*, **1996**.

[25] L. Blanco, A. Bernad, J. M. Lázaro, G. Martín, C. Garmendia, M. Salas, A. Bernads, J. M. Lharo, G.

Martins, *J. Biol. Chem.* **1989**, *264*, 8935–8940.

[26] X. Y. Li, Y. C. Du, Y. P. Zhang, D. M. Kong, *Sci. Rep.* **2017**, *7*, 1–10.

[27] C. Garmendia, A. Bernad, J. A. Esteban, L. Blanco, M. Salas, *J. Biol. Chem.* **1992**, *267*, 2594–2599.

[28] I. Gascón, J. M. Lázaro, M. Salas, *Nucleic Acids Res.* **2000**, *28*, 2034–2042.

[29] J. a. Esteban, M. Salas, L. Blanco, *J. Biol. Chem.* **1993**, *268*, 2719–2726.

[30] K. Struhl, **2011**, 1–15.

[31] H.-M. Eun, in *Enzymol. Prim. Recomb. DNA Technol.*, **1996**.

[32] A. V. Cherepanov, S. De Vries, *Eur. J. Biochem.* **2003**, *270*, 4315–4325.

[33] A. Fire, S. Q. Xu, *Proc. Natl. Acad. Sci.* **1995**, *92*, 4641–4645.

[34] L. Linck, U. Resch-Genger, *Eur. J. Med. Chem.* **2010**, *45*, 5561–5566.

[35] M. Nilsson, H. Malmgren, M. Samiotaki, M. Kwiatkowski, B. Chowdhary, U. Landegren, *Science (80-.).* **1994**, *265*, 2085–2088.

[36] J. Banér, M. Nilsson, M. Mendel-Hartvig, U. Landegren, *Nucleic Acids Res.* **1998**, *26*, 5073–5078.

[37] P. A. Rice, S. Yang, K. Mizuuchi, H. A. Nash, *Cell* **1996**, *87*, 1295–1306.

[38] T. W. Lynch, E. K. Read, A. N. Mattis, J. F. Gardner, P. A. Rice, *J. Mol. Biol.* **2003**, *330*, 493–502.

[39] D. I. Friedman, *Cell* **1988**, *55*, 545–554.

[40] L. M. Hales, R. I. Gumport, J. F. Gardner, *J.Bacteriol.* **1994**, *176*, 2999–3006.

[41] F. Rouessac, A. Rouessac, *Chemical Analysis : Modern Instrumentation Methods and Techniques*, **2004**.

[42] J. R. Lakowicz, *Principles of Fluorescence Spectroscopy*, Springer US, **2006**.

[43] S. Weiss, *Science* **1999**, *283*, 1676–1683.

[44] S. Gurrieri, K. S. Wells, I. D. Johnson, C. Bustamante, *Anal. Biochem.* **1997**, *249*, 44–53.

[45] W. Hauswirth, M. Daniels, *Chem. Phys. Lett.* **1971**, *10*, 140–142.

[46] L. M. Wilhelmsson, *Q. Rev. Biophys.* **2010**, *43*, 159–183.

[47] Y. Wang, H. Schellenberg, V. Walhorn, K. Toensing, D. Anselmetti, *Mater. Today Proc.* **2017**, *4*, S218–S225.

[48] B. Juskowiak, *Anal. Bioanal. Chem.* **2011**, *399*, 3157–3176.

[49] K. Y. Lin, R. J. Jones, M. Matteucci, *J. Am. Chem. Soc.* **1995**, *117*, 3873–3874.

[50] L. M. Wilhelmsson, A. Holme, P. Lincoln, P. E. Nielsen, B. Norde, **2001**, 2434–2435.

[51] L. D. Simon, K. H. Abramo, J. K. Sell, L. B. McGown, *Biospectroscopy* **1998**, *4*, 17–25.

[52] H. S. Rye, J. M. Dabora, M. A. Quesada, R. A. Mathies, A. N. Glazer, *Anal. Biochem.* **1993**, 144–150.

[53] A. Larsson, C. Carlsson, M. Jonsson, B. Albinsson, *J. Am. Chem. Soc.* **1994**, *116*, 8459–8465.

[54] K. H. Abramo, J. B. Pitner, L. B. McGown, *Biospectroscopy* **1998**, *4*, 27–35.

[55] H. S. Ryel, S. Yue, D. E. Wemmer, M. A. Quesada, R. P. Haugland, R. A. Mathies, A. N. Glazer, **1992**, *20*, 2803–2812.

[56] A. Larsson, C. Carlsson, M. Jonsson, *Biopolymers* **1995**, *36*, 153–167.

[57] E. Lang, J. Baier, J. Köhler, *J. Microsc.* **2006**, *222*, 118–123.

[58] K. R. Spring, M. W. Davidson, "Microscopy U: The Source of Microscopy Education," can be found under https://www.microscopyu.com/techniques/fluorescence/introduction-to-fluorescence-microscopy, **n.d.**

[59] I. De Vlaminck, C. Dekker, *Annu. Rev. Biophys.* **2012**, *41*, 453–472.

[60] K. R. Chaurasiya, T. Paramanathan, M. J. Mccauley, M. C. Williams, **2010**, *7*, 299–341.

[61] R. Roy, S. Hohng, T. Ha, *Nat. Methods* **2008**, *5*, 507–516.

[62] F. Persson, J. O. Tegenfeldt, *Chem. Soc. Rev.* **2010**, *39*, 985.

[63] J. F. Allemand, D. Bensimon, L. Jullien, A. Bensimon, V. Croquette, *Biophys. J.* **1997**, *73*, 2064–2070.

[64] J. O. Tegenfeldt, C. Prinz, H. Cao, R. L. Huang, R. H. Austin, S. Y. Chou, E. C. Cox, J. C. Sturm, *Anal. Bioanal. Chem.* **2004**, *378*, 1678–1692.

[65] V. Müller, **2017**.

[66] K. Frykholm, L. K. Nyberg, F. Westerlund, *Integr. Biol.* **2017**, *9*, 650–661.

[67] J. Melin, J. Jarvius, J. Göransson, M. Nilsson, *Anal. Biochem.* **2007**, *368*, 230–238.

[68] F. Persson, J. Fritzsche, K. U. Mir, M. Modesti, F. Westerlund, J. O. Tegenfeldt, *Nano Lett.* **2012**, *12*, 2260–5.

[69] C. Ducani, G. Bernardinelli, B. Högberg, *Nucleic Acids Res.* **2014**, *42*, 10596–10604.

[70] G. Stengel, B. W. Purse, L. M. Wilhelmsson, M. Urban, R. D. Kuchta, *Biochemistry* **2009**, *48*, 7547–7555.

[71] F. Persson, J. Fritzsche, K. U. Mir, M. Modesti, F. Westerlund, J. O. Tegenfeldt, *Nano Lett.* **2012**, *12*, 2260–2265.

[72] E. Abad, A. Juarros, A. Retolaza, S. Merino, R. Marie, A. Kristensen, *Microelectron. Eng.* **2011**, *88*, 300–304.

[73] L. Nyberg, F. Persson, B. Åkerman, F. Westerlund, *Nucleic Acids Res.* **2013**, *41*, DOI 10.1093/nar/gkt755.

[74] M. J. Cavaluzzi, *Nucleic Acids Res.* **2004**, *32*, 13e–13.

8. Appendix

8.1. Appendix 1

Oligonucleotide name	Sequence 5'-->3'	5' modification	Length [bp]	GC% [%]	Tm [°C]	MM [g/mole]	ε [L/mole.cm]
PLP*	TGAGGAGCTGCATACACCGCGCAACGGCTTGATACGCAACGCAAAAAACCGATGTAAGACACTATTAC	Phosphate	68	48.5	71.2	20944.6	670300
Target*	GGTGGTGGATCAGCTCCTCAGTAATAGTGTATCTCTCACA	/	40	47.5	64.8	12302	386500
Primer3 (P3)*	CGCGCAACGGCTTGATACGCAACGCAAAAAACCG	Phosphate	34	55.9	68.6	10423.8	329400
Primer4 (P4)*	ATGTAAGACACTATTACTGAGGAGCTGCATACAC	Phosphate	34	41.2	60.3	10458.9	341900
Primer31 (P31)°	CGCGCAACGGCTTGATACGCAACG(tC)AAAAAACCG	Phosphate	34	52.9	66.8	10608	323910
Primer32 (P32)°	CGCG(tC)AACGG(tC)TTGATACG(tC)AACG(tC)AAAAAACCG	Phosphate	34	44.1	62.9	10926	307440

* All values were calculated by IDT oligo analyzer. The site uses the ε values for the dNTP's from the paper that is in reference [74].

° The ε value is made up of the value from P3 given by IDT oligo analyzer and substract the difference in ε value between C and tC (ε(C) = 7400 M^{-1}[74] and ε(tC) = 13500 M^{-1}[18]. The difference has to be multiplied by 0.9, because it is in a duplex: ε(tC-C) = 5490 M^{-1}) for every tC present. Other parameters were taken from the atdbio information sheet.