Novel Biomolecular Carriers for Gene Delivery in Mammalian Cells

Investigation of DNA origami as a Potential Gene Delivery System

Author

Anne-Kathrine Kure Larsen

Supervisors

Leonid Gurevich

Evamaria Petersen

Master's Thesis

Nanobiotechnology

Aalborg University

June 2020

AALBORG UNIVERSITY

STUDENT REPORT

THE UNIVERSITY *of* EDINBURGH

AALBORG UNIVERSITY
STUDENT REPORT

Department of Materials and Production
Aalborg University
Skjernvej 4A
9220 Aalborg Øst
http://www.nano.aau.dk

Title:
Novel Biomolecular Carriers for Gene Delivery in Mammalian Cells. Investigation of DNA origami as a Potential Gene Delivery System

Project:
Master's Thesis

Project Period:
September 2019 - June 2020

Participants:
Anne-Kathrine Kure Larsen

Supervisors:
Leonid Gurevich
Evamaria Petersen

Total Page Number: 107
Appendix: 5
Finished: 3rd June 2020

Abstract:

The overall goal of this project was to investigate DNA origami as a novel method for efficient gene delivery and expression in mammalian cells. A plasmid encoding Cas9 and EGFP was used as a basis for DNA origami scaffold creation. PCR was utilized to amplify a target region in plasmid pCas9_GFP necessary for expression. PCRs revealed a template highly resistant to complete denaturation due to its GC-rich chicken β-actin promoter and chimeric intron region. A hairpin formation is suggested to prevent full-length amplification of the template as well. The double-stranded target region in plasmid lentiCas9-EGFP was successfully created instead. DNA origami modelling was performed, and two prospective designs are suggested. Simulations revealed apparent stability of the two structures in physiological conditions. Annealed oligo constructs and a model origami were used for initial studies of FRET detection of folded/unfolded states of DNA origami. The results hereof revealed a FRET signal for the annealed oligos, while further experimentation is needed to make a clear conclusion for denatured oligos and DNA origami.

Preface

This Master's thesis was conducted during the 3rd and 4th semester of the master programme in Nanobiotechnology at the Department of Materials and Production, Aalborg University. The project was conducted partly in the lab of Dr Paul De Sousa at the Centre for Clinical Brain Sciences, University of Edinburgh, and partly in the Nanobiolab at the Department of Materials and Production, Aalborg University.

The report will cover the general aspects of cellular delivery and uptake, followed by the materials and methods used in the project. Afterwards, a section describing DNA origami strategies and design and an experimental results section will follow. These results will further be discussed and the project concluded upon. Lastly, outlook and suggestions for further experimentation are presented.

Abbreviations used in the project is listed in "List of Abbreviations" or shown in brackets after first time use of the word. References are written using numbers, presenting the number in brackets relating to the referenced source in the bibliography. References in a sentence refer to that specific statement, while references after a full stop in the end of a subsection refers to the whole previous subsection. The bibliography lists the name and surname of the authors, title, journal with volume, issue (if applicable) and page numbers, or if a book, the publisher, and year of publication. In the case of three or more authors, the first author's name is written followed by *et al.* Figures without references are created by the author of this report. Both figures, equations, and tables are numbered in ascending order throughout the report.

I would like to thank Dr Paul De Sousa for giving me this unique opportunity to work in his lab, introducing me to the field of pluripotent and derivative stem cells, and for an extremely warm welcome to the lab.

Anne-Kathrine Kure Larsen

List of Abbreviations

AFM	Atomic force microscopy
A. G. LA polymerase	Advantage Genomic LA polymerase mix
aPCR	Asymmetric PCR
bp	Base pairs
CCMV	Cowpea chlorotic mottle virus
CME	Clathrin-mediated endocytosis
E. coli	*Escherichia coli*
DAPI	4,6-diamidino-2-phenylindole
dATP	Deoxyadenosine triphosphate
dCTP	Deoxyguanosine triphosphate
dGTP	Deoxycytidine triphosphate
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide triphosphate
dsDNA	double-stranded DNA
dTTP	Deoxythymidine triphosphate
EF-1α	Elongation factor 1α
EGFP	Enhanced green fluorescent protein
FBS	Fetal bovine serum
FRET	Förster resonance energy transfer
GC	Guanosine-cytosine

GFP	Green fluorescent protein
gRNA	Guide RNA
HEK293	Human embryonic kidney cells
HEK293T	Human embryonic kidney cells (stably expressing the SV40 large T antigen)
HEK293E	Human embryonic kidney cells (expressing Epstein-Barr Virus (EBV) nuclear antigen-1 (EBNA-1))
kb	Kilo base pairs
Nanovue	NanoVue^{TM} Plus Spectrophotometer (GE Healthcare)
nt	Nucleotides
PAM	Protospacer adjacent motif
PBS	Phosphate buffered saline
PCR	Polymerase Chain Reaction
Picodrop	Pico200 (Picodrop)
PEI	Polyethylenimine
RNA	Ribonucleic acid
ssDNA	Double-stranded DNA
$\mathbf{T_a}$	Annealing temperature
$\mathbf{T_e}$	Extension temperature
UoE	University of Edinburgh
WPRE	Woodchuck hepatitis virus post-transcriptional regulatory element
42hb	42 helix bundle DNA origami

Contents

1. Introduction

DNA origami has since introduction as a concept in 2006 gained much attention in a wide variety of fields as a method to construct nearly arbitrary 2D and 3D structures with high precision due to the unique base pairing of DNA. The principle is simple; using a single-stranded template of DNA as scaffold, a structure can be formed by designing oligos that are complimentary to different parts of the scaffold, acting as staples that holds the scaffold together at predefined places. When designed, scaffold DNA and staples are simply mixed, heated and cooled slowly down to self-assemble the structure. One of the interesting features of DNA origami that has been found is its ability to interface with cells, making it a promising candidate for drug- and gene delivery applications. [1]

While many studies focus on either drug delivery of DNA intercalator molecules loaded into the origami structures or orgiami as protective cages for delivery of other molecules or enzymes, plus for immunomodulation, only a few studies focus on gene delivery. Origami-mediated control of transcription by a transcription activator domain on the origami, and origami as a carrier for genes have been proposed, but not much focus has been put on using the origami itself for transcription. [1, 2]

This project was aimed on developing a novel gene delivery system for mammalian cells, using a GFP/Crispr-Cas9 encoding plasmid as a model. The idea was to fold the plasmid into a DNA origami, attach nuclear localization and cell penetrating signals to it and examine the delivery and subsequent transcription and protein expression in HEK293T-mCherry reporter cells with a genome integrated drug (doxycycline) inducible Tet-repressor/operator circuit as a model. In the presence of either doxycycline or functional delivery of Crispr-Cas9 and guide RNA to the Tet-repressor, the Tet-repressor would be inactivated, thus causing activation of the expression of mCherry fluorescent reporter. The transfection rates would then be compared to those of commonly used transfection agents Polyethylenimine and Lipofectamine.

As issues arose in the amplification of the selected plasmid, a thorough examination on this issue will be presented, together with proposed DNA origami designs, important considerations and initial experiments on the detection of folded/unfolded states of a model origami by FRET. Future experimentation necessary to reach the original goal will be outlined.

2. State of the Art

For many decades, gene delivery has been an important topic in biomedical sciences. It is a crucial first step when studying gene function and doing gene therapy, as it requires expression of an exogenous gene in the target cell. A variety of different methods have been evaluated and used during the years, including viral, nonviral and physical methods. These do all have different advantages and disadvantages and the field is constantly developing in the search for the most effective, simple and non-toxic ways to do cell transfection. [3]

Using replication-deficient viruses is one of the most common ways to transfer genes into cells and they were evaluated in the first gene therapy clinical trial in 1990, though the most studied DNA viral vector, a modified adenovirus, was discovered in 1953. The group of *Graham et al.* (1977) developed a cell line that was used for the first production of recombinant adenoviruses in 1977 and since that, adenoviral vectors have drawn much attention as gene delivery tools. These are highly efficient vectors compared to the less effective nonviral and physical methods. However, the nonviral solutions are often less toxic than viral vectors, entails higher flexibility in gene size and have lower immunogenicity. [3, 4]

Nonviral vectors are often build of DNA, proteins, polymers or lipids etc. utilizing the properties of each to create particles that can deliver genes into cells. This approach was investigated before the introduction of viral vectors, and the first method that gained acceptance was the calcium phosphate mediated gene transfer developed in the early 1970s. A few examples of nonviral gene delivery systems could be electroporation, gene guns and needle injection, together with cationic lipids and polymers. Lipofectamine is one of the most common cationic lipid based reagents used for transfection in a wide range of cell lines, together with polyethylenimine as a common cationic polymer transfection reagent. [5, 6]

Physical methods to deliver genes have gained increased attention during the last years. The principle is simple; genes are delivered using physical principles and therefore no reagents. As they are only relying on a physical force, any cytotoxicity or immunogenic reaction is avoided. This is indeed an advantage in order to be able to make safe transfection reagents for use in gene therapy but the methods need to be effective as well. While viral vectors are very effective, they have intrinsic properties that stimulate carcinogenesis

and immune response which is a fundamental issue when doing gene therapy. According to *Kamimura et al.* (2011) the solution might be to find a mechanism that can integrate transgenes into a selected site in a chromosome, thus avoiding several administrations and insertion mutagenesis. One the other hand, many non-viral methods are so ineffective that they are often considered unacceptable for clinical use. Furthermore, it is unlikely to find a single method that is optimal for gene delivery to all cell types. Optimally, the best delivery method would be by noninvasive routes that can target desired cells specifically and obviously present high delivery efficiency and appropriate action of the gene insert. [3]

Of natural compounds that have gained attention in drug and gene delivery over the past years is DNA. The unique base pairing properties of nucleotides gives DNA the ability to self assemble very precisely when having unique matching sequences. This is utilized in DNA origami where a scaffold and a mix of oligonucleotides are mixed and annealed to create specific shapes if designed properly. The first 3D origami structure to be created was a cube-like shape in 1991, assembled from DNA oligonucleotides, and since many other structures followed. A DNA tetrahedron first demonstrated by *Goodman et al.* (2004) with a simple step annealing and high yield have been a widely used DNA nanostructure in the research field of delivering diverse substances to cells. However, the concept of DNA origami itself with a scaffold and staples was demonstrated in 2006. It was presented as a method to design and construct nanostructures with the help of software to design the construct and relatively simple annealing steps to self-assemble the origami. While being proposed as a nanoscale material for many applications, it has also shown potential to interact with cells. Therefore, a growing interest in using such structures for cellular delivery has arisen. [1, 7]

An interesting study of a DNA origami triangle for gene delivery was performed by *Liu et al.* (2018) They created a DNA based nanocarrier for combined gene delivery and cancer therapy by linking the tumor therapeutic gene p53 to the DNA nanostructure loaded with a chemotherapeutic drug, doxorubicin, which is hydrophobic and intercalates into the base pairs of the DNA duplex. The nanocarrier was furthermore coupled to the MUC1 aptamer for targeted delivery to cancer cells, and the complex turned out to exhibit anticancer activity against a multidrug resistant tumor in both *in vitro* and *in vivo*. The gene delivery efficiency was furthermore compared to Lipofectamine by replacing the p53 gene with an EGFP gene and the percentages of positive EGFP cells were similar for both transfection methods. [2]

The question that now arises is, if one can obtain the same gene delivery efficiency, but by

the origami scaffold itself as the target genetic code to be delivered. A similar question was asked by *Lim et al.* (2015) as they point out that the large amount of nanocarriers needed for e.g. polymer carriers to deliver only a small copy number of genes gives some of the current nanocarriers a limitation. Their approach was to self-assemble plasmid DNA or siRNA into larger nanoclusters by the interaction of Zn^{2+} ions with the phosphate groups on the DNA backbone, and they obtained both efficient delivery of the nanoclusters into cells and expression of the target biomolecule. [8] The approach of using the cargo as the carrier at the same time eliminates both the use of a carrier and decrease the risk of toxicity associated with the carrier. Using DNA origami instead of the random packing of nanoclusters makes the process more complex, but at the same time it opens up to a wide range of opportunities with the ability to attach localization signals and fluorescent tracking dyes to the staples, or co-load compounds needed to be cotransfected with the plasmid DNA. It facilitate a higher degree of control over size and shape, potentially affecting the uptake fate as described in detail in the next chapter.

3. Gene Delivery by DNA origami

Since DNA origami as a method for constructing small, arbitrary shaped nanoparticles was first demonstrated by Paul Rothemund in 2006 [9], several potential applications of DNA origami have been proposed and many aspects of the field have been researched. One interesting feature of origami is its apparent ability to interface with cells, making it an attractive method in the field of cellular delivery. Although the potential for DNA origami in both materials science, drug- and gene delivery is large, it is still in an early stage and many hurdles have to be overcome before completely realizable. Some of the aspects that need to be considered and optimized for making DNA origami a fully implementable method is the stability of structures after administration, the targeting of specific cell types, how to improve uptake and target specific sub-cellular departments, and how to avoid that the structures are captured in endosomes. Lastly, the desired function of the structure should be both obtained and preserved until it reaches its final target domain in the cell. [1] All these aspects will be covered in the following sections, with the main emphasis on what is important when targeting the nucleus and making the DNA applicable for transcription.

3.1 DNA Origami Delivery

Delivering DNA nanostructures, and for that matter other nanostructures as well, a number of challenges have to be considered. These challenges depend both on the cell type and the characteristics of the particle in question for delivery.

3.1.1 Cell Targeting

For mammalian cells, different uptake pathways are possible, often varying among cell types. Often it is more simple and may not require any special targeting to deliver particles to cells, whose *in vivo* role is already uptake of foreign particles. That are for example monocytes and derivative macrophage and dendritic cells, all involved in the processing of foreign particles and presenting them to the immune system. For these it is often enough that the particle to be taken up is just in the preferred size range. [1] Other cells are, however, more selective and it will often require additional effort to make the particle suitable for uptake, although passive diffusion through the cell membrane may

be possible depending on their surface charge. An example of this is carbon nanotubes that have shown to be able to diffuse through the membrane and into the cytoplasm in a non endocytotic manner. [1]

DNA origami structures are negatively charged due to the phosphate groups on the backbone and therefore passive diffusion through the negatively charged cell membrane might not be the most probable outcome, unless charges can be concealed. An advantage is, however, that DNA origami is easy to modify by incorporation of ligands via attachment on staple strands for specific cell targeting. Also fluorescent tags can be applied by direct attachment or intercalation to track the pathway of origamis when entering a cell environment. As these functional groups are added to the staple strands, it is likewise possible to place them very precisely if needed. However, it has in fact been shown that origami structures can in some cases be able to effectively target specific cells without any modification. For example, in a study by *Zhang et al.* (2014) a triangle shaped origami was constructed and loaded with the anticancer drug doxorubicin and shown to passively target tumors, showing prominent effect *in vivo*, although this could probably be explained by enhanced permeability and retention effects in the tumor region. [10]

For targeted delivery as mentioned, a ligand able to recognize and bind to a surface molecule on the cell in question is often included in the DNA origami structure. This could be a DNA aptamer; a single-stranded DNA or RNA oligonucleotide that folds into specific 3D structures. They are typically in the size range from 20 to 80 bases and have the ability to recognize and bind target molecules with high affinity due to their compatibility with specific binding sites on target molecules such as nucleic acids, proteins and sugars. Furthermore, aptamers have several advantages over proteins in therapeutic applications like DNA origami because of their small size, easy synthesis and modification possibilities. [11]

If no specific receptor is to be targeted, it is instead possible to incorporate a cell penetration ability into the nanostructure. This can be obtained by adding cell penetrating peptides (CPPs) that can facilitate the entry into the cell. An example is a study where a tile based DNA nanostructure was decorated with cytosine-phosphate-guanine (CpG) which had to be delivered. The CpG oligodeoxynucleotides were assembled in loops on the DNA nanostructures to make it more resistant to nuclease digestion, while the attached CPP, a TAT peptide, had the effect of a greatly enhancing the uptake and accumulation in the endosomes in macrophages (which was the target destination for the study). [12]

3.1.2 Stability of DNA Origami Structures

Particles can also be delivered directly into the cytoplasm by electrotransfection where an electrical field is applied to disrupt the cell membrane and allow entry for the foreign particle. For DNA origami, this is however not always a safe route as a strong electric field can strip away the Mg^{2+} and other ions that normally stabilize the structure. However, a study by *Chopra et al.* (2016) suggests that folding origami in the presence of 1mM spermidine, a DNA condensating agent naturally occuring in cells, instead of $MgCl_2$ can produce structurally stable DNA origami structures that withstands an electric fields of 250 and 500 V/cm which were used for electrotransfection. Some of the structures were successfully internalized in Jurkat cells, mainly localized in the cytosol. [13]

Even when electroporation is not chosen to assist DNA origami in crossing the membrane, the ability of origami to retain stability in an environment with low Mg^{2+} concentration is still vital as this is the premise at physiological conditions. Many studies suggest that this ability to a large extend depends on the characteristics of the individual structure, among those size, shape and structure type, which in turn can affect the uptake efficiency described in section 3.2.5. DNA origami nanostructures are in general stable in $MgCl_2$ concentrations of 4-20 mM or NaCl concentrations of 1-3 M. In a study by *Ahmadi et al.* (2018) the stability of a 350 nm nanorod and a 50x25 nm nanobottle, both designed in CaDNAno, and a wireframe structure, designed in Daedalus, was evaluated in salt depleted media. The folding buffer was replaced with a Mg free buffer containing 5 mM Tris, 1mM EDTA and 30 mM NaCl and the structures were incubated for a day at 37 °C. For naked DNA origami, individual amorphous structures were observed, indicating denaturation of the origami under those conditions. On the other hand, DNA origami coated with linear polyethyleneimine (LPEI) at a N/P ratio (ratio of the amines in polycations to the phosphates in the DNA) ≥ 1 and DNA origami coated with chitosan at N/P ≥ 2 were stable under these conditions. Similarly, naked DNA origamis were highly susceptible to degradation in the presence of DNAse I, though the compact nanobottles showing the largest resistance, while LPEI and chitosan induced protection from this as well. Typical cell culture media contain as low as 0.6 mM $MgCl_2$ and fetal bovine serum (FBS), a typical supplement, contains several digesting enzymes. However, quite interesting the naked DNA origamis remained stable for up to a week in DMEM medium supplemented with 10 % FBS at 37 °C, though the concentration of them decreased slightly over time. The nanobottles did again show the largest resistance and degraded slower, but when the buffer exchange was performed only with DMEM medium, the nanorod and wireframe structure were still intact after one day at 37 °C, while the nanobottles were partially degraded. This could indicate that higher salt concentrations are needed to keep the

integrity of the more compact structure than the less compact nanorods and wireframe origami. The unexpected stability was hypothesised to stem from the growth media that contains a variety of free amino acids, inorganic salts and basic vitamins to possibly help sustain the nanoparticle integrity. [14]

During the recent 5 years, an alternative approach to the whole way of designing origami structures has gained ground. Increasing interest has risen in designing the wireframe, one-layer hollow structures due to an apparent larger stability and lower requirement of salt to keep their integrity. Also, more simple and easier rendering of the design were demanded. That resulted in new algorithms dedicated to the design of origamis with an open mesh structure, and so far these have in general shown to require fewer cations to stabilize the structures, making them more stable in buffers mimicking physiological conditions. Also, folding of the mesh structures has proven to be successful in both PBS buffer and DMEM. [1, 15, 16] Of programmes currently available, DAEDALUS and vHelix for Maya can be mentioned. Both programmes are based on the idea of folding the origami over a closed surface with a polygon shape as the skeleton. vHelix which is used in this project will be explained further in detail in section 5.1.2.

3.2 Uptake of DNA Nanostructures

In contrast to essential small molecules as amino acids and ions, larger molecules are transported over the cell membrane in a much more regulated way. These processes are all different types of endocytosis in which vesicles carrying the cargo into the cell are made by forming a cleft in the cell membrane with subsequent detachment from the membrane. Generally, there are two main types of endocytosis; phagocytosis and pinocytosis. Phagocytosis is the mechanism used when specialized cells as monocytes, macrophages and dendritic cells take up foreign and often large particles >250 nm for presentation to the immune system, as mentioned in section 3.1.1. Pinocytosis is on the other hand utilized in the uptake of fluids and smaller particles and can be divided into four subtypes: Clathrin-mediated endocytosis, caveolin-mediated endocytosis, clathrin and caveolin independent endocytosis and macropinocytosis. [11]

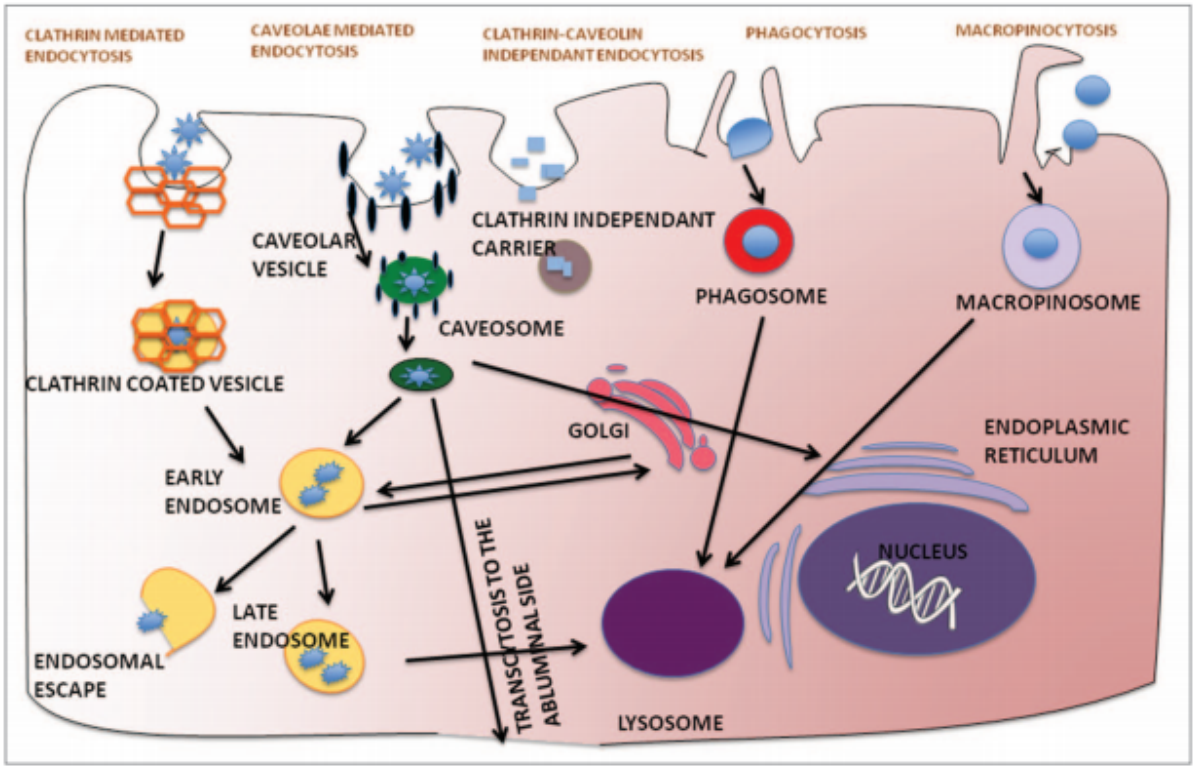


Figure 3.1: Taken from [17].

3.2.1 Clathrin-mediated Endocytosis

In clathrin-mediated endocytosis (CME), the most well studied of the uptake pathwways, the particles taken up are most often destined for the lysosomes, in which degradation of internalized particles takes place. It can be subdivided into receptor dependent and independent uptake. For molecules utilizing ligand-receptor pathways, receptor dependent CME is very common and therefore important for nanoparticles targeting such receptors. During CME, a formation of coated pits is triggered by the receptor-ligand bindings on the cytosolic side of the plasma membrane. The pits are formed by cytosolic proteins of which clathrin, a three-legged structure formed by three clathrin heavy chains, constitute the largest number. However, under physiological conditions, clathrin is dependent on adaptor proteins to self-assemble into the characteristic polygonal cages present in the accumulation of clathrin pits that further drives the deformation of the membrane to draw the cargo in question into the interior of the cell in the newly formed clathrin coated vesicle, developing into an early endosome. Next, a fusion is made between the early endosome and a pre-lysosomal vesicle that contains enzymes to produce a late endosome, maturing into a lysosome afterwards. For receptor independent CME, the internalization rate is often slower as the cargo will interact non-specifically with the membrane by charges and hydrophobic interactions. [11, 18]

3.2.2 Caveolin-mediated Endocytosis

This type of endocytosis is the most studied non-clathrin dependent mechanism and is seen in a wide variety of functions, e.g. signalling, lipid regulation and vesicular transport. Caveolae, meaning small caves in Latin, are typically 50-80 nm cavities in the plasma membrane, arising from caveolins, which are 21 kDa membrane proteins that attach to the cytosol side of the membrane via a hydrophobic sequence and binds cholesterol. Despite well studied, the exact link between caveolae-localized receptors and target molecules is still unknown, although a number of molecules such as cholera toxin and folic acid are known to utilize this pathway into the cells, and the pathway is known to have a role in transcytosis which several serum components and some viruses use to cross endothelium tissue. [18] In addition, a caveolar vesicle does not necessarily have the late endosome and lysosome as destination and can instead be transported to the caveosome. This organelle has, in contrast to others in the endocytic pathway, a neutral pH and not the usual endosomal markers. Here the loaded material can be sorted to enter the endoplasmic reticulum or trans-Golgi network.

3.2.3 Clathrin- and Caveolin Independent Endocytosis

Other pathways, so-called clathrin and caveolin independent mechanisms as a broad term, are not as well understood. However, it is known that other small cholesterol-rich microdomains on the plasma membrane do also play a role in endocytosis. Some of these structures are termed 'lipid rafts' which diffuse freely on the cell surface and are believed to be able to internalize within any endocytic vesicle, despite the exact mechanism not being well known.

3.2.4 Macropinocytosis

Macropinocytosis is in some aspects similar to phagocytosis as a mechanism to internalize larger molecules, however macropinocytosis can occur in a wider range of cell types. Both mechanisms are driven by actin assembly and engulfment of the particle. In macropinocytosis, a protuberance is formed in the membrane which collapse onto the particle to be taken up and fuse with the membrane. The produced macropinosomes have sizes of over 1 μm. The precise destiny of these is different within each cell type, but will in many cases fuse with lysosomes. Macropinocytosis is in general an unselective process, complementing the other mentioned pathways in nanoparticle uptake. [11]

3.2.5 Uptake Pathways for DNA origami

In line with the idea that many pathogens' entry into cells are affected by their size and shape, it has shown to a significant relation to the internalization of nanoparticles as well, still in a cell type specific manner. *Bastings et al.* (2018) studied the modulation of cellular uptake of DNA origami through control over mass and shape. They examined the uptake of 11 different DNA origamis, with the general shapes created in different size ranges, in human umbilical vein endothelial cells, bone-marrow-derived dendritic cells and HEK293 cells. They found that, in general, larger particles with greater compactness were preferentially internalized compared with elongated, high-aspect-ratio particles. Quite specific, they demonstrated a significant linear relationship between origami compactness and uptake efficiency for all three cell types. They observed that the compact structures with low-aspect-ratio with sizes form 50 to 80 nm resulted in the most efficient internalization. [19]

3.2.6 Endosomal Escape

After entering the cell by endosomal uptake, particles will, as described in section 3.2, most often be trapped in an endosome with lysosomal degradation as the end destination and therefore have to release a loaded cargo or escape itself if needed for further functionality. This needs to happen as fast as possible after entering the cell to preserve the stability of the particle taken up. A natural way for a particle to avoid ending up in a late endosome is by being taken up by the caveolae-mediated endocytosis which leads to the formation of a caveosome instead of an endosome that does not contain any of the enzymes related to degradation in lysosomes. This is also the way many bacteria and viruses avoid the enzymatic degradation. It is therefore rational if nanoparticles like drug carriers or DNA origami can be tuned to exploit this path. The drawback is, however, a slower internalization rate in most cells than for clathrin-mediated endocytosis. [11]

Other strategies than aiming for the caveolae-mediated pathway have been suggested as well, for example creating DNA nanostructures that could induce a change in lysosomal pH and thereby disrupt the lysosome to redistributed the carried drug to other sites of the cell. This has been observed in drug resistant tumor cells where the anti-cancer drug is often ending up in acidic organelles such as lysosomes, but if increasing the lysosomal pH, the drug could be redistributed. [20] An interesting study by *Lim et al.* (2015) investigated the self-assembly of nucleic acids in the presence of zinc and their subsequent ability to enter cells. The 200 nm nanoclusters were self-assembled from either plasmid DNA or siRNA through the interaction of Zn^{2+} ions with the phosphate groups of the

nucleid acids with the aim of transfection without a gene carrier, very much in line with the aim of this project. The nanoclusters were evaluated for internalization and transfection in HEK293 cells and the efficiencies were higher than for both PEI 1800 (100x) and poly-L-lysine (10x), but lower than for PEI 25 kDA (10x). The cellular uptake was, however, largest for the Zn/DNA nanoclusters. That means, PEI 25 kDA was still the most effective obtaining actual gene expression, but the nanoclusters were much more efficient entering the cell. That could indicate either less DNA protection or endosomal escape of the nanoclusters, however, their studies showed that they were capable of escaping to a lower degree. Furthermore, the study shows that the structures internalized by a mixture of macropinocytosis, clathrin-dependent endocytosis and clathrin independent endocytosis with no primary pathway defined. [8]

In another interesting study by *Mikkalä et al.* (2014) they used CCMV capsid proteins for the enhancement of cellular delivery. The capsid proteins can self-assemble on the origami surface and bind through electrostatic interactions with high yield via their positively charged N-terminus. Their study shows that delivery of coated origamis into HEK293 cells could be performed with a 13-fold improved delivery compared to naked DNA origamis, and the transfection efficiency far exceeded that of their positive control, Lipofectamine 2000 (L2K). At the same time, no significant cell toxicity was observed. When examining the origamis attached with capsid proteins in the target cells by confocal microscopy after transfection, colocalization of Cy3-labelled origamis with Hoechst-stained nuclei suggests that the origamis are within the cell nucleus or at least in very close proximity. [21] In another study by the same group, the two proteins, bovine serum albumin (BSA) and class II hydrophobin, were anchored to DNA origami via electrostatic interactions by a dendron polymer attached to the proteins via a cysteine-maleimide bonds, and accessed in terms of stability, transfection and immunocompatibility when delivered into HEK293 cells. Especially the BSA coating improved origami stability against endonucleases such as DNAse I as well as the transfection rate, and it facilitated endosomal escape to some extend, even though the structures did not reach the nucleus. [22]

A last example of a strategy to overcome endosomal escape is a nanoclew structure prepared by *Sun et al* (2015). They designed this structure to deliver Cas9 protein and sgRNA to the nucleus of mammalian cells for genome editing and found that coating with the cationic polymer PEI could help induce endosomal escape. Further fusion of nuclear localization peptides to Cas9 lead the complex into the nucleus. For a further note, they observed that editing was most efficient for the sgRNA whose sequence was partially complementary to the DNA nanoclew. [23]

These studies show that it is in fact possible to create strategies to make the nanoparticles escape the endosomal pathway and the following degradation in lysosomes. However, it is worth to mention that it is always difficult to draw general conclusions, as the situation will most often be different from cell type to cell type.

3.2.7 Targeting the Nucleus

Whereas aptamers are valuable for targeting a specific cell type, and CPPs are great at increasing the general cellular uptake, it is often necessary to use other strategies to make the nanoparticle arrive at the planned destiny if not the endosomes. First, as described in section 3.2.6, it often has to escape the endosomes before reaching the lysosomal stage. Thereafter, for the purpose of this project, it has to arrive at and enter the nucleus to facilitate transcription of the plasmid. One strategy for this is to include a nuclear localization signal (NLS) as demonstrated for a DNA tetrahedra in HeLa cells by *Liang et al.* (2014). By inhibiting the caveolin-mediated pathway and observing a large reduction in uptake of the bare DNA tetrahedra without a NLS, it was suggested that this was the main uptake pathway. Despite it should be the best pathway to avoid degradation, the ultimate destination was shown to be the lysosomes. However, when attaching a NLS lysosomal escape was possible, and the NLS-tetrahedra were now localized in the nuclei. [1, 24] Another example is the study by *Sun et al.* (2015) mentioned in section 3.2.6 where a NLS peptide was also fused to the cargo to facilitate entry into the nucleus. A last example of successful targeting of the nucleus is the virus encapsulation performed by *Mikkalä et al* (2014), described in section 3.2.6. [21]

3.3 DNA Transcription

Transcription is carried out by RNA polymerase together with general transcription factors which are additional factors necessary for recognition of promoter sequences, response to regulatory factors and conformational changes, all essential for the performance of RNA polymerase. In eukayotic cells, three different nuclear RNA polymerases are used to synthesize varying classes of RNA, RNA polymerase I, II and III. RNA polymerase II and its associated transcription factors are the most complex system, consisting of nearly 60 polypeptides. It is also the polymerase responsible for all messenger RNA and also several small RNA molecules, e.g. many precursors to small regulatory RNAs. [25, 26]

RNA polymerase II transcription can be divided into three stages: Initiation, elongation and termination. First, initiation sites are searched for, and gene specific regulatory factors bind near the site of initiation, that is at the core promoter. A core promoter is

the minimal DNA sequence required to specify nonregulated or basal transcription, and it helps the polymerase to position in a 'preinitiation complex'. Hereafter, a short piece of the DNA is unwound to be able to read the single stranded template strand. Often, short sequences of RNA are produced first, after which the polymerase starts synthesis of the full-length RNA. After approximately 30 bases of synthesized RNA, it is thought that the polymerase releases its contacts with the promoter region and the rest of the initiative machinery. Other factors that can support the elongating RNA polymerase can now be recruited. *In vitro*, many of the transcription factors used to initiate the transcription remain at the promoter site, probably marking already transcribed genes and making next round of recruiting faster. At the end of transcription, the RNA polymerase recognize termination signals and the transcription will end. Underway in the transcription process, activator and repressor proteins are continuously modulating the rate of transcription as a part of the control of gene expression. [25, 26]

It is important to note that for RNA transcription, the produced RNA transcript has the same sequence as the coding, or sense (+), DNA strand. It is thereby the antisense (-) strand which is the template strand which is read by the RNA polymerase in the transcription process. [26]

In this project, the plasmid pCas9-GFP was thought to constitute the template of the DNA origami, composed of a chicken β-actin promoter followed by a chimeric intron, a Cas9 sequence, EGFP sequence and a SV40 polyadenylation signal in the region determined as origami template. However, as the amplification of the template part of this plasmid involved some unseen difficulties, another plasmid was chosen in the end. This plasmid, lentiCas9-EGFP, will be focused on instead when analyzing the promoter and termination region in the following sections, as this will be the plasmid best applicable for origami production. This plasmid contains, in the origami template region, also both Cas9 and EGFP sequences, but has another promoter system which is an a EF-1α core promoter and a WPRE regulatory element.

3.3.1 EF-1α Core Promotor

The plasmid aimed to use for origami folding and transfection contains a 212 bp EF-1α core promoter for human elongation factor EF-1α, regulating the gene expression. Most eukaryotic cells abundantly express this elongation factor which is an enzyme catalyzing GTP-dependent binding of aminoacyl-tRNA to ribosomes. [27] A study by *Wakabayashi-Ito et al.* (1994) investigated the EF-1α promoter and introns by introducing a series of deletions and mutations in the promoter region and fusing it to bacterial chloramphenicol

acetyltransferease gene, determining the expression levels in HeLa cells. The promoter has a strong activity in various types of cells and is therefore useful for mammalian expression vectors. The expression is both transcriptionally and post-transcriptionally regulated, for instance is the expression of EF-1α enhanced in cell lines transformed with different oncogenes. In addition, other factors are known to affect the expression but its gene regulation in higher eukaryotes is still not completely understood. [27]

The study mentioned before showed that the 5' flanking region and a part of the first intron are essential for the promoter activity in HeLa cells. This was also reported to be the case for the 5' flanking region, the first exon and first intron when expressing in Chinese hamster ovary cells and human neuroblastoma IMR32 cell lines, compared to plasmids lacking the first intron. The intron region contains several SpI and ApI elements that seem to have an effect on the promoter activity. Progressive deletion of it caused a gradual decrease in promoter activity for most deletions, suggesting that multiple elements including SpI an ApI sites have additive effects on activity. [27] Their results further indicate that in the 5' flanking region from -129 to -42 contains regulatory cis-elements. Specifically they found two elements, EFP1 and EFP2, and several nuclear factors that bind to them. Creating internal deletion mutants in the elements ranging from -99 to -80 for EFP1 and -69 to -50 for EFP2 resulted in 25 % and 1 % activity compared to the wild-type promoter, indicating the importance of these elements for promoter activity. Furthermore, they tested if the two elements could work independently of each other and if repetitive sequences of the same element could preserve the promoter activity. However, the results showed an interdependent function of the elements as an enhancer, leaving one of them out causing a significant reduction in activity. Also, repetitive sequences of either one or both elements only increased the activity slightly. [27] While these results are a good indication on to which elements that are important for promoter activity it should be noted that the experiments were carried out in HeLa cells. Whether the same regulatory elements will also be essential in other cell types is not guaranteed.

Other elements in the promoter that can be valuable for the transcription initiation is the well-known TATA and CAAT box elements. Although eukaryotic promoter regulatory elements are not rigidly conserved and varies a lot, there are some consensus elements such as these boxes. TATA boxes have thus been found in around 25 % of human promoters, approximately 30 bp upstream of the transcription start site and they recruit the transcription pre-initiation complex and RNA polymerase. CAAT boxes are found in around 13 % of human promoters, binding the ubiquitous transcription factor NFY. [28]

3.3.2 Cas9 and EGFP

CRISPRs; clustered regularly interspaced palindromic repeats, have been described since 1987 in *Escherichia coli*, but in the mid-2000s elevated research into this area revealed it as a part of a quite impressing defense mechanism in several bacteria and archea. The CRISPR-Cas loci was found to consist of a CRISPR array of identical repeats with intercalated sequences that functions as memory signatures to recognize past invading DNA, encoding crRNA (crispr RNA), and an operon encoding the Cas protein components. The crRNA (trans-activating crispr RNA) was in 2011 found to be essential for crRNA maturation and to cause sequence specific immunity against parasite genomes comprising this. In 2012, it was shown that the *S. pyogenes* CRISPR-Cas9 protein is a dual-RNA-guided DNA endonuclease performing double-stranded breaks in DNA, guided by the tracrRNA/crRNA duplex. The dual tracrRNA/crRNA has since been engineered to be a single guide RNA, sgRNA, that has the two critical features originally performed by the two RNAs: a 20-nt sequence at the 5' end determining the target site for cleavage and a double-stranded structure at the 3' end required for binding to Cas9. This simple two component system opened up for a new way to simply and precisely editing DNA sequences, possible to use in an endless number of applications. Simply, changing the 20-nt guide sequences of the sgRNA will make it able to target the desired region of interest. However, the target recognition requires both base pairing with this guide sequence and presence of a short PAM (protospacer adjacent motif) sequence adjacent to the target sequence in the DNA, thereby limiting the direct targeting to DNA sequences with an adjacent PAM. [29] However, in a study by *Kleinstiver et al.* (2015) they engineered the *S. pyogenes* Cas9 protein to produce mutants with altered PAM specificites and thereby enabled editing of currently non-targetable sequences by the wild type protein. Furthermore, they identified a variant that exhibited improved specificity and better discrimination against off-target sites in human cells. [30]

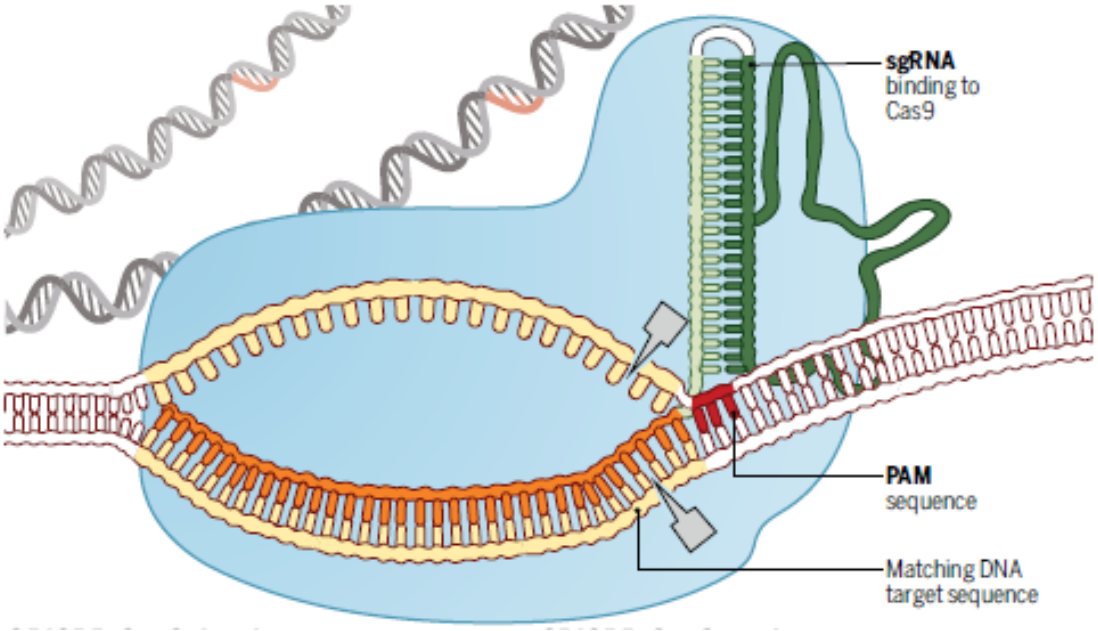


Figure 3.2: Illustration of the Cas9-gRNA complex during cleavage of a double-stranded DNA fragment. The illustration is taken from [31].

Enhanced green fluorescent protein is used as a reporter gene for immediate observation of expression.

3.3.3 WPRE Element

The woodchuck hepatitis virus post-transcriptional regulatory element (WPRE) is a post-transcriptional element that enhances gene expression by affecting nuclear mRNA processing, export an translation. Using a WPRE can therefore supplement an otherwise relatively low expression in transient expression in contrast to that of a stable cell line. [32] In a study by *Kim et al.* (2009) the influence of the WPRE on antibody production in HEK293E cells in combination with a CMV promoter. They found that WPRE increased the antibody production by 5.5-fold through an increase in mRNA levels and efficient export of the nuclear mRNA to the cytoplasm. [32]

Another study by *Schambach et al.* (2006) pointed out that different lengths of the WPRE are common, such as a 450 bp, 600 bp and 900 bp which they tested. They tested all three types for the influence on protein expression in murine fibroblast cells (SC-1), with the longest yielding the largest effect. However, the second largest element accounted for 74 % of the effects of the longest, meaning that the extra sequence on the 3' part was not absolutely necessary. [33] The previously mentioned study in HEK293E cells used a WPRE of 592 bp [32]. In this project, the WPRE is 589 bp which is corresponding

well with this size and also the second largest element in [33], possibly just being another variant of the regulatory element.

It is worth to mention that the actual effect on this element may not be the same in this setting, due to variations in cell lines and the utilized promoter. However, these studies indicates that it is possibly worth to include it in the scaffold for DNA origami design to optimize expression levels.

4. Materials and Methods

In the following sections, materials and methods used throughout the project will be presented.

4.1 Materials

Table 4.1: Materials used throughout the project, split into chemicals, biological material and buffers.

Chemicals	Description	Supplier
Agar	Lot: BCBP0217V	Sigma-Aldrich
Agarose	Lot: 117161	Fischer Scientific
Agarose	Lot: SLBN6401V	Sigma-Aldrich
Agarose I^{TM}	Lot: 19G0256361	VWR Life Science
Ampicillin sodium salt	Lot: 127K0609	Sigma-Aldrich
Ampicillin sodium salt	Lot: BCBB6153	Sigma-Aldrich
Betaine 5M	Lot: SLBZ7952	Sigma-Aldrich
DMSO	Lot: 25.0680407	TH. Geyer
DMSO (from Phusion^{TM} High-Fidelity DNA Polymerase)	Lot: 00853989	Thermo Scientific
DNAse free water	Lot: RNBC3293	Sigma-Aldrich
dNTP mix (10mM each)	Lot: 1361987	Invitrogen
dATP	Lot: 10044848	NEB
dCTP	Lot: 10044851	NEB
dGTP	Lot: 10044852	NEB
dTTP	Lot: 10044857	NEB
7-deaza-dGTP, Li-salt	Lot: 11422523	Roche Diagnostics
Ethidium bromide	Lot: SLBF7132V	Sigma-Aldrich
Extra staples	See Appendix C	
polyT staples	See Appendix C	
GelRed	Lot: 14PO731	Biotium
LB Agar powder	Lot: SLBD6887V	Sigma-Aldrich
LB Broth EZMix^{TM} powder	Lot: 107K8200	Sigma-Aldrich
Magnesium chloride sol. 50 mM	Lot: BRK25-111A	Bioline
Magnesium chloride	Lot: 00026651	PCRBIO
Mineral oil null (pcr reagent)	Lot: MKCF5776	Sigma-Aldrich
Nuclease free water	Lot: 00775580	Thermo Scientific
NuSieve®GTG®Agarose (LM agarose)	Lot: 0000306214	Lonza

Sodium Chloride	Lot: 17L184138	VWR Chemicals
Sodium hydroxide	Lot: SZBE2520V	Sigma-Aldich
SYBR Safe DNA gel stain 10,000x in DMSO	Lot: 2107413	Invitrogen
Trizma base	Lot: SLBZ1416	Sigma-Aldrich
Tryptone	Lot: L0159W	VWR Life Science
Yeast exstract	Lot: VM619326 346	EMD Milipore Coporation
5x Loading buffer	Lot: HLBB-108J	BioRad
Biological material	**Description**	**Supplier**
Advantage Genomic LA polymerase mix 5u/μL	Lot: 1904658A	TaKaRa
Biotaq™ Red DNA Polymerase	Lot: BRK25-111A	Bioline
Core staples	L-CP-11-1-2	Tilibit Nanosystems
NEB®5-alpha F'I^q Competent *E. coli*	-	NEB
EZ load™ Molecular Ruler 1 kb	Lot: 64250077	BioRad
*Bam*HI-HF	Lot: 10020678	NEB
*Bam*HI	Lot: 0000355155	Promega
*Eco*RV-HF	Lot: 10016059	NEB
*Eco*RI-HF	Lot: 0151803	NEB
*Mlu*I, GQ	Lot: 0000315619	Promega
*Not*I-HF	Lot: 10020162	NEB
*Pvu*I-HF	Lot: 10041276	NEB
*Sac*I-HF	Lot: 10043047	NEB
*Sap*I	Lot: 10036024	NEB
*Sma*I	Lot: 0841503	NEB
OneTaq 2xMaster Mix	Lot: 100050007	NEB
lentiCas9-EGFP #63592 in *E. coli* Stbl3	-	Addgene
pCas9_GFP #44719	-	P. De Sousa lab, UoE
p7560 scaffold	Lot: M-1-3-5	Tilibit Nanosystems
Taq DNA polymerase	Lot: 00815450	Thermo Scientific
Phusion™ High-Fidelity DNA Polymerase	Lot: 00853989	Thermo Scientific
Quick-Load Purple 1 kb Plus DNA ladder	-	NEB
1 kb DNA ladder	Lot: 1441709	NEB
100 bp DNA ladder	Lot: -	NEB

Buffers	Description	Supplier
Buffer D	Lot: 0000245201	Promega
Buffer E	Lot: 0000178048	Promega
TAE buffer	-	In house stock
5x Q5 High GC enhancer	Lot: 10009458	NEB
10x Advantage Genomic LA Buffer (Mg^{2+})	Lot: 1004656A	TaKaRa
10x CutSmart buffer	Lot: 10018445	NEB
10x Folding buffer	Lot: M-2-1-2	Tilibit Nanosystems
10x NH_4 Reaction buffer	Lot: BRK25-111A	Bioline
10x Taq buffer with KCl	Lot: 00837310	Thermo Scientific
10x Taq buffer with $(NH_4)_2SO_4$	Lot: 00825069	Thermo Scientific
Primers	**Description**	**Supplier**
F.pcas9gfp.5phos	5'-GGAATAAGGGCGACACGGAA-3' (5'-phos.)	Eurofins Genomics
R.pcas9gfp.5biot	5'-ATAGTCCTGTCGGGTTTCGC-3' (5'-biot.)	Eurofins Genomics
R.pcas9gfp	5'-ATAGTCCTGTCGGGTTTCGC-3'	Eurofins Genomics
Fpcas9gfp5p18m	5'-GGAATAAGGGCGACACGG-3' (5'-biot.)	Eurofins Genomics
Adj. forw. scaf.	5'-GGTCGCTGAGTAGTGCGC-3'	TAG Copenhagen
Rev. real creates com.	5'-GGAGCCTATGGAAAAACGCCA-3'	TAG Copenhagen
GCregionpCas9forw.	5'-CTTCACTCTCCCCATCTCC-3'	TAG Copenhagen
GCregionpCas9rev.	5'-TCATCCACCTTAGCCATCTC-3'	TAG Copenhagen
Forw-Lenti-outEF1a	5'-GGTTTATTACAGGGAC AGCAGAG-3'	TAG Copenhagen
Rev-Lenti-outWPRE	5'-GCTGCCTTGTAAGTCATTGG-3'	TAG Copenhagen
Oligos	**Description**	**Supplier**
TBIT-00108-Cy3	5'-ATTCGCCTTGGGAAG GGCGATAAGGC-3'	Eurofins Genomics
TBIT-00100	5'-TGCTTTGAAGTTTGAGATTAGA ACAATTAAATGTGTTTTTAG-3'	Eurofins Genomics
Oligo-sd-FRET	5'-GTTCTAATCTCAAACTT CAAAGCAGTACGCCTTAT CGCCCTTCC-3'	Eurofins Genomics
Oligo-ld-FRET	5'-GTTCTAATCTCAAACTTCAA AGCACCGACGCGCTGGTCCGC CTTATCGCCCTTCC-3'	Eurofins Genomics
Middle-ld-FRET	5'-GACCAGCGCGTCG-3'	Eurofins Genomics

4.2 Methods

4.2.1 Medium for Cell Growth

For bacterial growth, LB-medium was prepared by mixing 10 g/L tryptone, 5 g/L yeast extract and 5 g/L NaCl. When making plates for bacterial growth, 15 g/L agar was further added. The medium was autoclaved at 121 °C for 30 minutes. For the LB-agar plates, 200 mg/L ampicillin was added before pouring the plates. When using liquid LB-medium for plasmid isolation, 400 mg/L ampicillin was added.

4.2.2 Plasmid Purification

Plasmid pCas9_GFP [34] was obtained in solution after extraction by another student in the P. De Sousa lab (UoE). Later in the Nanobiolab (AAU), it was obtained from transformed *E. coli* 5-alpha F'Iq Competent cells (NEB). It was thereafter extracted from an overnight culture of a single colony, using a GeneJET Plasmid Miniprep Kit (Thermo Scientific), following the instructions from the supplier. lentiCas9-EGFP was extracted using the latter after receiving it as a bacterial stab culture from Addgene. [35] All extractions were performed using overnight cultures of their respective *E. coli* strain in LB medium, grown at 37 °C, 275 rpm.

Concentration measurements of extracted plasmid fractions were carried out using a NanoVue^{TM} Plus Spectrophotometer (GE Healthcare) or a Pico200 (Picodrop).

4.2.3 Restriction Enzyme Analysis

Restriction enzyme analyses were carried out on both plasmid pCas9_GFP and plentiCas9-EGFP, using 20 units of enzyme per 1 μL plasmid (stock from extraction) if at 100 % activity and 1x of the recommended buffer and for 1 hour at 37 °C if nothing else is mentioned. If activities did not match in double or triple digestions, equivalent extra enzyme was used compensating for reduced activity.

4.2.4 Sequencing

Sequencing of the target region in pCas9_GFP was performed by DNA Sequencing & Services (MRC | PPU, School of Life Sciences, University of Dundee, Scotland, www.dnaseq.co.uk) using Applied Biosystems Big-Dye Ver 3.1 chemistry on an Applied Biosystems model 3730 automated capillary DNA sequencer. The work was managed by the Cloning Group,

MRCPPU/Reagents and Services, School of Life Sciences, University of Dundee, who designed sequencing primers.

4.2.5 Polymerase Chain Reacion

Several PCRs were carried out during this project. In this section, main reaction solutions are presented, and any changes from these protocols are mentioned together with each result.

The sequences, tags and melting temperatures are shown in table 4.2.

Table 4.2: A list of all primers with their sequences, melting temperatures and tags. T_m values are the ones provided by the supplier of the primer.

Primer	**Sequence/T_m**	**Tag**
F.pcas9gfp.5phos	5'-GGAATAAGGGCGACACGGAA-3' (59.4 °C)	5'-phosphate
R.pcas9gfp.5biot	5'-ATAGTCCTGTCGGGTTTCGC-3' (59.5 °C)	5'-biotin
R.pcas9gfp	5'-ATAGTCCTGTCGGGTTTCGC-3' (59.4 °C)	-
Fpcas9gfp5p18m	5'-GGAATAAGGGCGACACGG-3' (58.2 °C)	5'-biotin
Adj. forw. scaf.	5'-GGTCGCTGAGTAGTGCGC-3' (54.9 °C)	-
Revrealcreatescom.	5'-GGAGCCTATGGAAAAACGCCA-3' (56.7 °C)	-
GCregionpCas9forw.	5'-CTTCACTCTCCCCATCTCC-3' (53.2 °C)	-
GCregionpCas9rev.	5'-TCATCCACCTTAGCCATCTC-3' (51.8 °C)	-
Forw-Lenti-outEF1a	5'-GGTTTATTACAGGGACAGCAGAG-3' (55.3 °C)	-
Rev-Lenti-outWPRE	5'-GCTGCCTTGTAAGTCATTGG-3' (51.8 °C)	-

For the many first reactions, the following standard protocol in table 4.3 when using A. G. LA polymerase was used with any deviations mentioned in the results section. All PCR runs were, if possible, performed by preparing mastermixes, ensuring a uniform solution of all reagents.

Table 4.3: PCR setup and program for PCR run (a1).

Composition	**Concentration**	**Amount (μL)**	**Final concentration**
Nuclease free H_2O	-	18.75	-
dNTP mix	10 mM (each)	1.00	0.4 mM
A.G. LA buffer	10x	2.50	1x
F.pcas9gfp.5phos	10 μM	1.00	0.4 μM
R.pcas9gfp.5biot	10 μM	1.00	0.4 μM
A.G. LA polymerase	5u/μL	0.25	0.05 u/μL
Template DNA	200 ng/μL	0.50	4 ng/μL
Total		**25**	
Temperature (°C)	**Time (min)**		**Repeats**
94	1		x1
98	10 (sec)		
52/55/57/60/68	0.5		x30
68	8		
72	10		x1
4	∞		

With Taq polymerase, the following protocol in table 4.4 were the standard if nothing else mentioned. Also here, mastermixes was prepared when possible and Taq polymerase always added last.

Table 4.4: PCR setup and protocol for the first amplification of pCas9_GFP with Taq polymerase and betaine/DMSO derivatives.

Composition	**Concentration**	**Amount (μL)**	**Final concentration**
Nuclease free H_2O	-	1	-
dNTP mix	5 mM (each)	1.25	0.25 mM (each)
Taq buffer with KCl	10x	2.5	1x
$MgCl_2$	50 mM	0.75	1.5 mM
Adj. forw. scaf.	10 μM	1.25	0.5 μM
Rev. real creates com.	10 μM	1.25	0.5 μM
Template DNA	700 ng/μL	1	28 ng/μL
DMSO	7.5/12.5/25%	5	1.5/2.5/5%
Betaine	2.5/3.75/5M	10	1/1.5/2 M
Taq DNA polymerase	1 u/μL	1	0.04 u/μL
Total		**25**	

Temperature (°C)	**Time (min)**	**Repeats**
95	3	x1
95	0.5	
51	0.5	x30
72	3	
72	15	x1
4	∞	

For reactions where the above mentioned conditions are not used, the conditions will be specified together with the result or in Appendix A which will be referred to when relevant.

4.2.6 Agarose Gel Electrophoresis

Agarose gel electrophoresis was used to confirm sizes of PCR products or confirm annealing of DNA oligos. Agarose gels were made by dissolving agarose in the desired buffer by microwaving until all agarose was completely dissolved. Afterwards it was stored at 60 °C to avoid polymerization if anything left. Preparation of a gel was done by adding 1 μL SYBR Safe DNA stain or GelRed per 10 mL gel, or 0.5 μL Ethidium bromide per 40 μL gel and pouring it into a casting frame with the required comb. If post-staining with SYBR Safe DNA stain or GelRed was used, no dye was added prior to casting. When set, the gel was placed in the electrophoresis tray and covered with buffer. The samples were prepared with 1/6 loading dye, and the desired DNA ladder was loaded as well. The electrophoresis was run using a Mini-Sub Cell GT from Bio-Rad, often at 90 V (70 V for preparative gels) and until it had ran the desired length. If not pre-stained, the gel was

post-stained by placing it in buffer with SYBR Safe/GelRed while shaking for around 40 minutes or until the bands were visible. The gel was imaged, and if preparative cut, under UV light to examine it.

For PCR products of expected 7,797 bp and 7,453 bp, 0,7 % agarose gel in 1xTAE buffer was used, and for products below this size, a 1 % gel was used instead. For origami samples and annealed oligos, a 2 % agarose gel in 0.5xTB buffer was used, while for purifying annealed oligos, a 1.5 % LMP agarose gel in 1xTAE buffer was used. PCR products were stained with SYBR Safe DNA stain or Ethidium bromide, while origami and oligos were stained with GelRed.

4.2.7 DNA Extraction from Gel

Extraction of PCR product from amplification of plentiCas9-EGFP was carried out by running a preparative agarose gel in 1xTAE buffer, afterwards cutting out the gel piece containing the product. It was packed in parafilm and frozen overnight at -20 °C. The DNA fragment was extracted from the gel using a Millipore Ultrafree DA Centrifugal Filter Device Kit, following the instructions from the supplier. Centrifugation steps were carried out at 5,000 g for 10 minutes throughout several rounds until no more of the gel piece was left. Precipitation of the DNA fragment was performed by adding 2x volume 96 % ethanol and 1/10 volume 3 M NaAc (pH 5.2), freezing it at -20 °C for around 2 hours. To obtain the DNA as a pellet, it was centrifuged at 20,000 g for 30 minutes, discarding the supernatant, adding 500 μL 70 % ethanol and centrifuging again at 20,000 g for 10 minutes. Last, the supernatant was again discarded while the pellet air dried until no ethanol was left in the tube. The DNA was resuspended in 20 μL DNAse free water.

4.2.8 Folding of Square Origami

The folding mixture for a 42hb origami sample is presented below in table 4.5. The origami folding was carried out in a Blue-Ray Biotech Turbocycler. The solution was heated to 65 °C for 10 minutes and then gradually cooled down from 60 to 40 °C with -1 °C/h. Last, it was cooled to 4 °C and stored at the same temperature.

Table 4.5: Folding solution for one 42hb origami batch of 20 μL.

Composition	**Concentration**	**Amount (μL)**	**Final concentration**
p7560 scaffold	100 nM	4	20 nM
10X folding buffer	450 mM	2	45 mM
$MgCl_2$	200 mM	2	20 mM
Core staple	746 nM	5.36	200 nM
polyT pass. strands	1539 nM	2.60	200 nM
Extra staples	12.5 μM (each)	0.72	450 nM (each)
TBIT-00100 (ATTO647N)	12 μM	0.36	216 nM
TBIT-00108-Cy3	10 μM	0.36	180 nM
Ultrapure water		2.60	
Total		20	

The extra staples solution was prepared by mixing 2 μL of TBIT-00101R, TBIT-00102R, TBIT-00103R, TBIT-00104R, TBIT-00105, TBIT-00106 and TBIT-00107 with 2 μL DNase free water.

4.2.9 Dialysis of 42hb Origami

To purify the origami after confirming the folding by gel electrophoresis, dialysis using 100 kDa dialysis membranes was performed. The membranes were cut and placed in 0.5xTB buffer with 11 mM $MgCl_2$ for at least an hour, and then the origami solution was transferred to the membrane together with 100 μL or less buffer. The dialysis membrane was clamped and placed in an electrophoresis chamber, running at 65 V for 15 minutes and last 2 minutes with reversed polarity. The origami solution was then pipetted carefully out of the membrane and stored at 4 °C. The concentration of the purified origami was estimated using Picodrop.

4.2.10 Oligo Annealing and Purification

Oligos were mixed at a concentration of 1 μM each. To simplify the presentation of methods and results, the oligos are named as follows:

Table 4.6: List of all oligos used for annealing of the constructs D-ST, A-ST, D-A-ST, D-LT, A-LT and D-A-M-LT. T_m values are the ones given from the supplier.

Oligo	Short name	**Sequence (length,T_m)**	**Tag**
TBIT-00108-Cy3	D	5'-ATTCGCCTTGGGAAG GGCGATAAGGC-3' (26, 68.0 °C)	3'-Cy3
TBIT-00100	A	5'-TGCTTTGAAGTTTGAGATTAGAAC AATTAAATGTGTTTTTAG-3' (42, - °C)	5'-ATTO647N
Oligo-sd-FRET	ST	5'-GTTCTAATCTCAAACTT CAAAGCAGTACGCCTTAT CGCCCTTCC-3' (44, 73.2 °C)	-
Oligo-ld-FRET	LT	5'-GTTCTAATCTCAAACTTCAA AGCACCGACGCGCTGGTCCGC CTTATCGCCCTTCC-3' (55, >75 °C)	-
Middle-ld-FRET	M	5'-GACCAGCGCGTCG-3' (13, 46.0 °C)	-

Solutions of D and ST (named D-ST), A and ST (named A-ST), D, A and ST (named D-A-ST), D and LT (named D-LT), A and LT (named A-LT) and D, A, M and LT (named D-A-M-LT) were prepared in 1xPBS buffer and annealed by heating to 95 °C for 5 minutes and cooling down from 95 to 25 °C, lowering the temperature 1 °C/min. When reaching 25 °C, the sample was cooled to 4 °C and stored at this temperature.

The annealing of the oligos was analyzed by running the samples in a 2 % agarose gel in 0.5xTB buffer.

Purification of the annealed oligos was performed to filtrate away loose strands in the solution to avoid background in fluorescence measurements later. The purification was carried out with a QIAEX II Gel Extraction Kit from Qiagen by first running a 1.5 % LMP agarose gel in 1xTAE buffer and cutting out the desired band under UV illumination. The annealed fragments were extracted from the gel piece by following the instructions provided by the kit. All centrifugation steps were carried out at 13,000 RCF. Last, the fragments were resuspended in at least 50 μL 1xPBS buffer to have enough for fluorescence measurements. Concentrations of the fragments were measured using Picodrop and recalculated in terms of mol/volume using the molar mass of the total constructs, as the solutions differ in composition. The solutions for fluorescence were all dissolved to a final concentration of 350 nM.

4.2.11 Fluorescence Measurements

Fluorescence measurements were carried out using a ISS chronos DFD fluorescence lifetime spectrometer and a quartz cuvette with dimensions of 5.25x5.25mm. The minimum requirement for sample volume was 50 μL and all measurements were performed using this volume. No optical filter was applied for any of the measurements performed. Baselines of solvent for the different solutions were subtracted from the data.

4.2.12 Denaturation of Origami and Oligos

To observe the effect on the FRET signal when the dyes separate, the origami and annealed oligos were denatured. Both oligos and origami were denatured adding NaOH to a concentration of 1 M, and origami denaturation by heating the sample to 65 °C for 5 minutes and quickly moving it onto ice was also tested. For NaOH denaturation, all results were normalized to the final concentration of 315 nM.

4.2.13 Structure Characterization with AFM

Atomic force microscopy was performed with a AFM NT-MDT Solver in semi contact mode on 42hb origami structures to characterize and determine folding and unfolding of the structures when denaturing them. Samples were prepared on a freshly cleaved mica plate by depositing 10 or 15 μL sample solution, letting settle for around 20 minutes and rinsing with 100 μL or 1 mL Milli-Q water. Scans were performed with varying sizes from 4x4 to 1x1 μm, 512 lines and a frequency of 0.4 to 1.0 Hz.

4.2.14 Origami Design

Two origami structures were designed using the plugin vHelix in Autodesk Maya 2019, and a BSCOR scaffold routing algorithm and spring relaxation software, as described in detail in section 5.1.2. For both designs, an octahedral platonic solid subdivided in triangles was chosen.

4.2.15 Simulations with oxDNA

Simulations were performed with the simulation code oxDNA as described in section 5.1.3. First, an energy minimization of 5000 steps with a stepsize of 0.00166 resulting in a 25 ps simulation was run for both structures. This was followed by a relaxation procedure performed by running an MD simulation for 10,000 steps with a stepsize of 0.005 giving a total runtime of 150 ps. A maximum backbone force of 486 pN was set for both minim-

ization and relaxation, together with a salt concentration of 0.1 M NaCl. For relaxation, a langevin thermostat was included to control the temperature, and average sequence parameters were used.

5. DNA Origami Design

Even though DNA origami has been shown to fold from a double-stranded DNA source [36, 37], in this project it was chosen to rely on creation of a single-stranded scaffold. When folding origami with a double-stranded scaffold source, it suffers from the possible re-annealing of the two complementary scaffold strands competitioning with the folding of the origami structure. Thermodynamically, the re-annealing of scaffold will often be favoured, and it therefore requires extra care to balance the assembly process in favour of the origami folding. This has shown to be possible by either having a molar ratio of staples to scaffold at at least 50:1 [37], or by additives in the folding process such as formamide to keep the scaffold strands separated [36].

If not balanced properly, origami folding might still be possible, but the yield will be low. Based on these considerations, it was chosen to rely on the creation of single-stranded scaffold as a high yield is required and excessive amounts of staples would be a significant cost.

5.1 Strategies

In the following sections, the general strategies for ssDNA creation and DNA origami design will be presented.

5.1.1 Creation of Single-stranded Origami Scaffold

In all matters, the part of the scaffold that should be used for the DNA origami scaffold would be created and amplified by PCR, as it is a straightforward method capable of producing large amounts of product. The initial idea for the creation of single-stranded DNA from the PCR product was inspired by *Murgha et al.* (2014) by the addition of a phosphate group at the 5' end of the forward primer that is used to extend the sense strand in PCR, and a biotin tag at the 5' end of the reverse primer that is used to extend the antisense strand. The antisense strand is the strand that is used for transcription by RNA polymerase and is therefore the wished product in the end. By using the enzyme lambda exonuclease that preferably hydrolyses phosphorylated double-stranded DNA, it should be possible to create single-stranded DNA in an easy way. The biotin moiety on the antisense strand is assumed to protect it, although lambda exonuclease hydrolyses

non-phosphorylated substrates at a greatly reduced rate. The biotin tag also serves as a way to purify the ssDNA if necessary, for example by using streptavidin on magnetic beads. Afterwards, the production of ssDNA can be confirmed by digestion of both untreated PCR product and lambda exonuclease treated product with the single-strand specific Exonuclease I. [38]

Another strategy to create ssDNA is by asymmetric PCR (aPCR) as demonstrated by *Veneziano et al.* (2018) who created ssDNA of up to 15 kb. In aPCR, one of the primers is present in molar excess over the second primer, in the mentioned study 50-molar excess, to amplify preferably the target strand. When favourably using one of the primers, the reaction will, after the non-target primer is used up, proceed linearly instead of exponentially as in standard PCR. Therefore, the yield is much lower than in standard PCR and it is important to reduce non-specific priming as the product from this can easily exceeed the target product by exponential reaction rate. [39]

5.1.2 DNA Origami Design

DNA origami design is in this project carried out using vHelix, a plugin for Autodesk Maya developed by Högberg Research Group at Karolinska Institutet. The program is developed to be used together with BSCOR software that routes the scaffold along the edges of a predefined polyhedral mesh and afterwards relaxes the structure using a spring relaxation algorithm. The product is an .rpoly format file which can be imported into Autodesk Maya for subsequent staple design. The scaffold routing algorithm was developed by Orponen group at Aalto University, while the spring relaxation software was developed by Högberg Research Group at Karolinska Institutet. By this, an open DNA structure with primarily one helix per edge, which is expected to be stable under physiological conditions as described in section 3.1.2, can be obtained. [15, 40]

A consideration on the design was how to make the scaffold amenable for transcription by RNA polymerase. As the transcription factors that recognize the transcription initiation site bind double-stranded DNA as described in section 3.3, this was deemed to be important. An origami structure is, after all, double-stranded as the scaffold base-pairs with the staples at all sites, but the single staple binds more than one place on the scaffold, and therefore the scaffold is not continuously double-stranded with the same staple running along. To create double-stranded regions so that the RNA polymerase transcription complex can hopefully recognize the right place, the most important sites in the promoter region were determined from literature studies. This is described in section 3.3.1. From this knowledge, two sequences in the promoter that is deemed to be most important for

recognition will be covered with one single oligo each to make it continuously double-stranded over these sections. The placements of oligo 1 and 2 are seen below in figure 5.1.

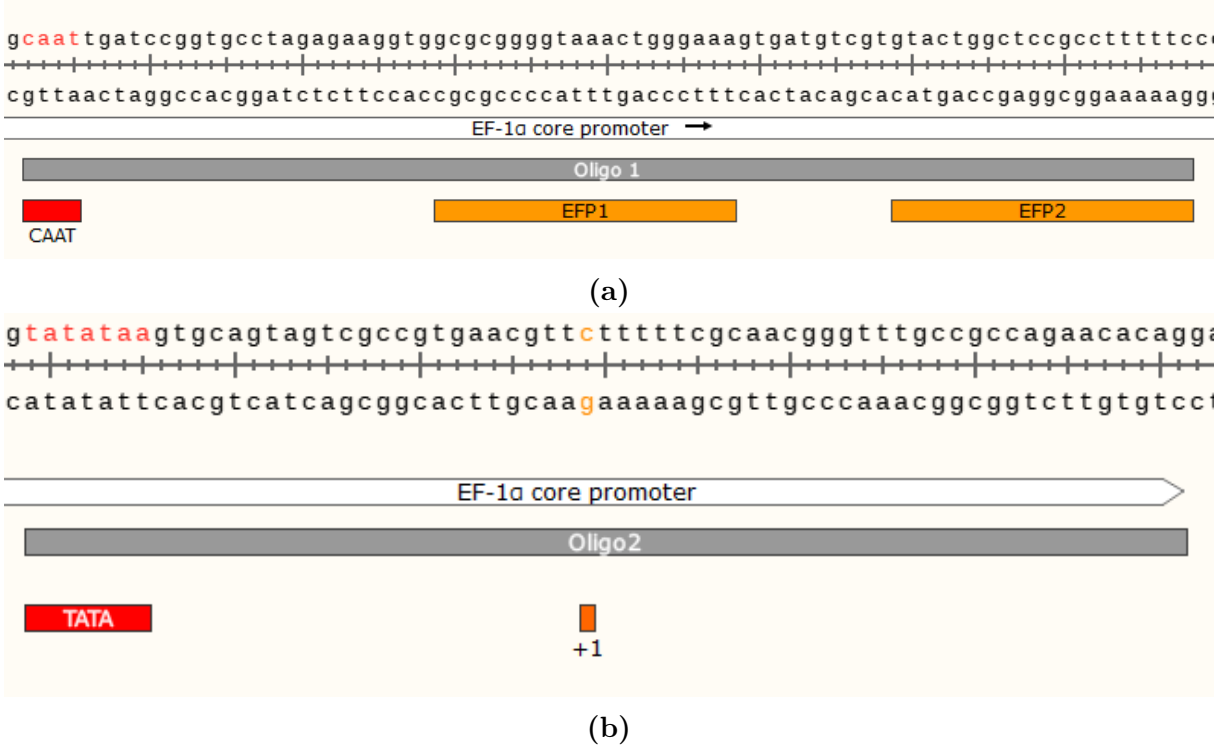


Figure 5.1: Illustration of the placement of the long oligos for double-stranded regions in the promoter region. (a) Oligo 1 is a total of 77 bases, while (b) Oligo 2 is a total of 63 bases. The boxes are indicating the placement of important promoter elements and the transcription start site. The illustration is obtained using SnapGene Viewer.

The design process is split into 4 steps: First, a 3D polygon mesh structure were drawn using Autodesk Maya 2019 as the structure was desired. Two designs were created in this project; one where the promoter regions (discussed further below) are packed inside the rest of the structure, and one where it constitutes a part of one of the outer polyhedral edges. It should be noted, that completely packing/hiding of those strands inside the structure was impossible, as the scaffold routing can only be done on a closed surface. Also, the space between the two sequences that had to be packed is only 20 nt, which means that only a short turn is possible between them. The immediate best way to pack it with these limitations was to create a 'pit' from the surface and into the inner part of the structure, as depicted in figure 5.2. This mesh constitutes the basis for the design of the first DNA origami structure, OH1.

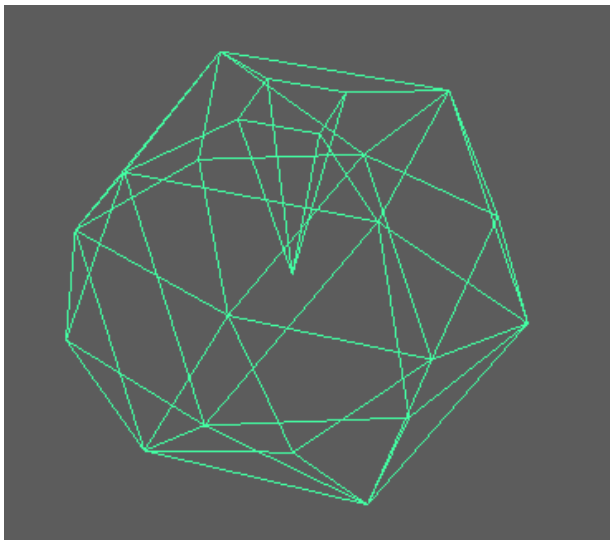


Figure 5.2: The final polyhedral mesh for OH1 used for scaffold routing. The mesh was created in Autodesk Maya 2019, from where this illustration is also printed.

Another structure, basically with the same polyhedral shape but with a flat side on the outside for the oligos, was designed as well to compare the two. This is named OF1, and the mesh can be seen in Appendix B, figure B1.

After the mesh creation, the scaffold routing and spring relaxation, to determine the least-strained DNA helix arrangement, were performed to generate a DNA origami with scaffold and staples routed along the edges of the polyhedral mesh. The step is performed automatically by the software. Afterwards, staple adjustment can be done manually by loading the structure in Maya with the vHelix plugin. First, any too large strand gaps were auto-filled such that the structure would not encounter strain at these sites when folded. Thereafter, manual staple adjustment was necessary a few places in the designed structure, to compensate for too long or too small staple sequences, that would either be difficult and expensive to synthesize or have too weak hybridization with the scaffold. The structure of the origami with the oligos packed is seen in figure 5.3.

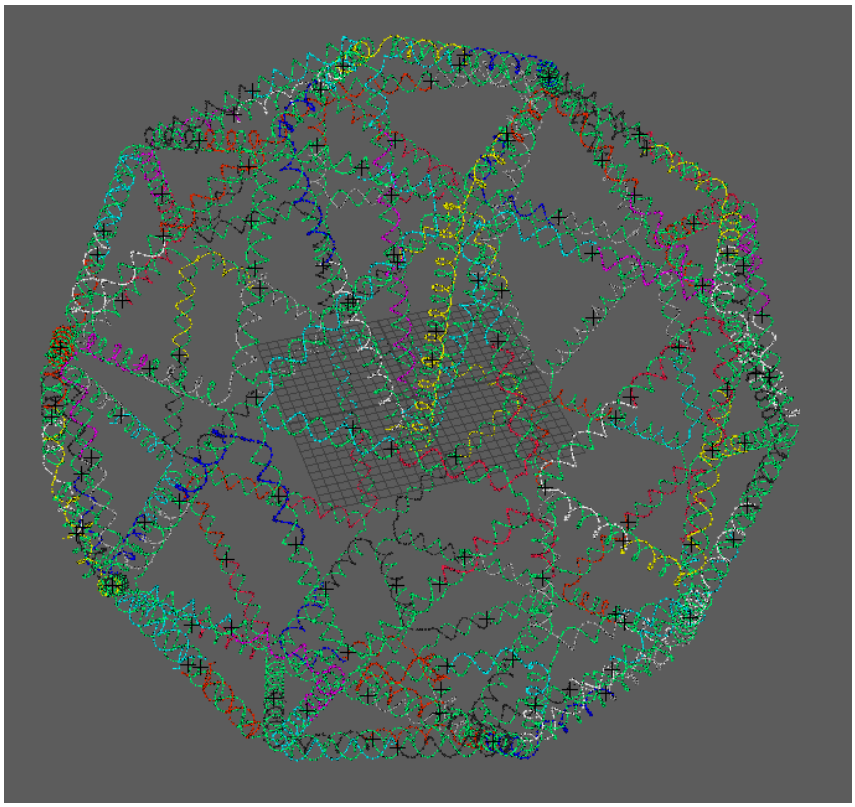


Figure 5.3: The polyhedral mesh of OH1 with the scaffold routed along all edges. Illustration printed from Autodesk Maya 2019.

Also, staple adjustment was performed to include the long oligos implemented for the promoter region and securing that the sequence of the plasmid would fit with the 20 nt space between them. Last, the sequence was added such that it fitted with the oligo regions. A zoom in on the region where the oligos meet is seen in figure 5.4.

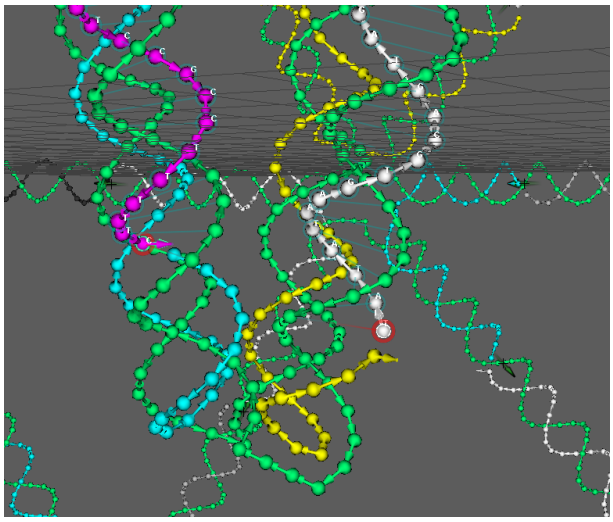


Figure 5.4: A closer look at the region where the 3' end of Oligo 1 (purple strand) can be seen in a turn together with the 5' end of Oligo 2 (white strand). A total of 20 bp are between the two oligos, covered by other staple end. Illustration printed from Autodesk Maya 2019.

The other structure with the oligos on the outside can be seen in figure 5.5 below, while a zoom in on the oligo region is seen in figure B2 in Appendix B.

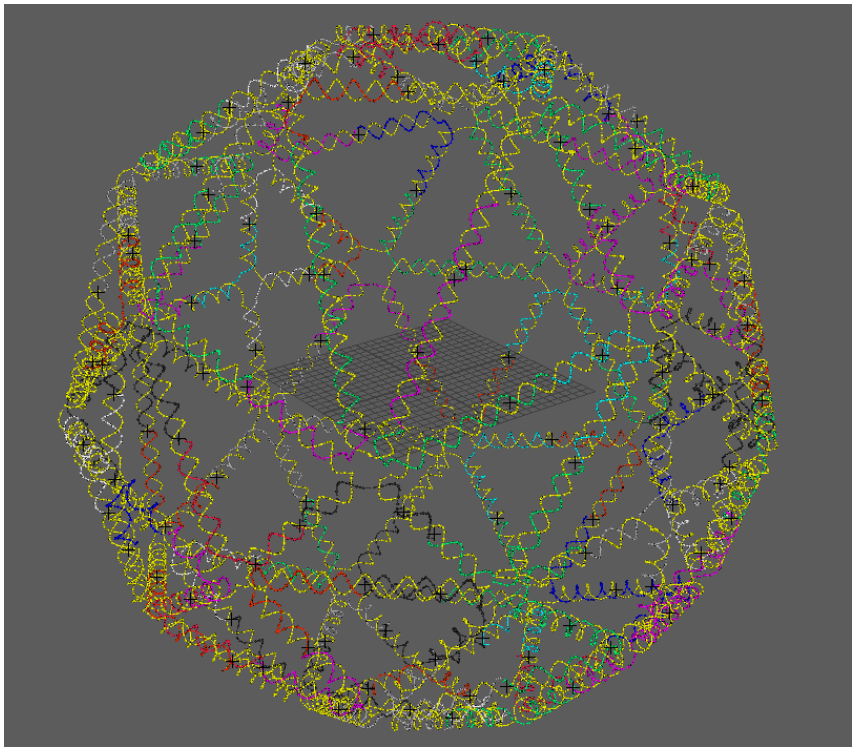


Figure 5.5: The polyhedral mesh of OF1 with the scaffold routed along all edges. Illustration printed from Autodesk Maya 2019.

The DNA origamis uses 6,054 bp and 5,885 bp of the total sequence of the 6,064 bp target

region in lentiCas9-EGPF for OH1 and OF1, respectively.

5.1.3 Simulations

The simulations of the two designs were carried out with the simulation code from oxDNA [41], using TacoxDNA to prepare the data from vHelix for an oxDNA simulation. [42] oxDNA was originally used to implement the coarse-grained DNA model introduced by T. E. Ouldridge, J. P. K. Doye and A. A. Louis and has since been developed by researchers in the groups of Doye and Louis at the University of Oxford. [41] In the simulations carried out in this project, molecular dynamics (MD) simulations are performed using the oxDNA2 code which is a newer version that introduced different widths for the major and minor grooves in the DNA helix. Furthermore, it includes an electrostatic interaction based on the Debye-Hückel model that enables the simulation of DNA structures in salt concentrations of 0.1 M NaCl and above which is relevant for the purpose of simulating the origami structures in this project in concentrations closer to physiological conditions than normally used for origami folding. However, the simulation is restricted to NaCl and not $MgCl_2$ as is usually the ion in question for origamis. [43]

5.2 Methods for Evaluation of Folding/Unfolding

To study folding and unfolding of DNA origami structures, measurement of fluorescence resonance energy transfer (FRET) is utilized. A previously designed DNA origami by *Sønderskov* (2016) is used as a model origami to test the method and its applicability. It is a square shape DNA origami consisting of 42 helix bundles, referred to as 42hb in this project. The structure is illustrated in figure 5.6 [44] FRET is a non-radiative transfer of energy between two fluorophores in close proximity. When a donor fluorphore is photoexcited, it can transfer energy to a ground-state acceptor fluorophore resulting in a typical relaxation of the acceptor by emitting light in its emission spectrum. The FRET efficiency can be given quite simple:

$$E = \frac{R_0^6}{R_0^6 + R^6} \tag{5.1}$$

The Förster radius, R_0, is the distance at which the energy transfer is 50 % and R is the distance between donor and acceptor. Förster radii are often around 3-6 nm. As seen in eq. 5.1, the efficiency falls off as $\frac{1}{R^6}$ and therefore the observation of FRET is only possible at small distances, typically of maximum 10 nm. This also means that FRET measurements are sensitive to even very small variations in donor-acceptor distances, making it

valuable when investigating small conformational changes. As R_0 is proportional to the spectral overlap between the two fluorophores, it is important to choose a suitable FRET pair. [45]

The fluorophore pair used in this project is the ATTO647N fluorophore, which was already incorporated into the DNA origami structure (depicted as a red circle in figure 5.6), and a Cyanine3 (Cy3) added to another staple in the configuration (depicted as a yellow circle in figure 5.6). A rough estimate of the distance between the two fluorophores is 5 nm.

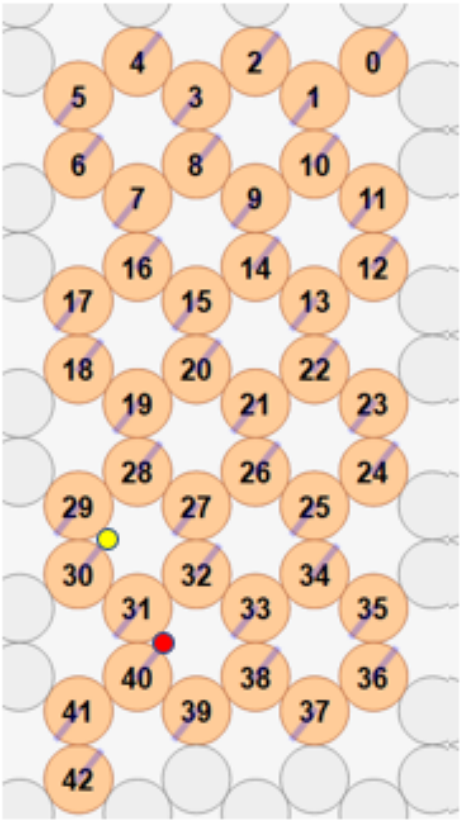


Figure 5.6: Illustration of the 42hb DNA origami. Fluorophores are depicted as circles; ATTO647N the red circle, and Cy3 the yellow circle. Illustration created using CaDNAno.

Cy3 is a cyanine fluorophore used as the donor with an absorbance maximum at 550 nm and an emission maximum at 568 nm. ATTO647N has an absorption maximum at 646 nm and an emission maximum at 664 nm, according to Eurofins Genomics. [46] The spectral overlap of the two dyes can be seen in figure 5.7 below.

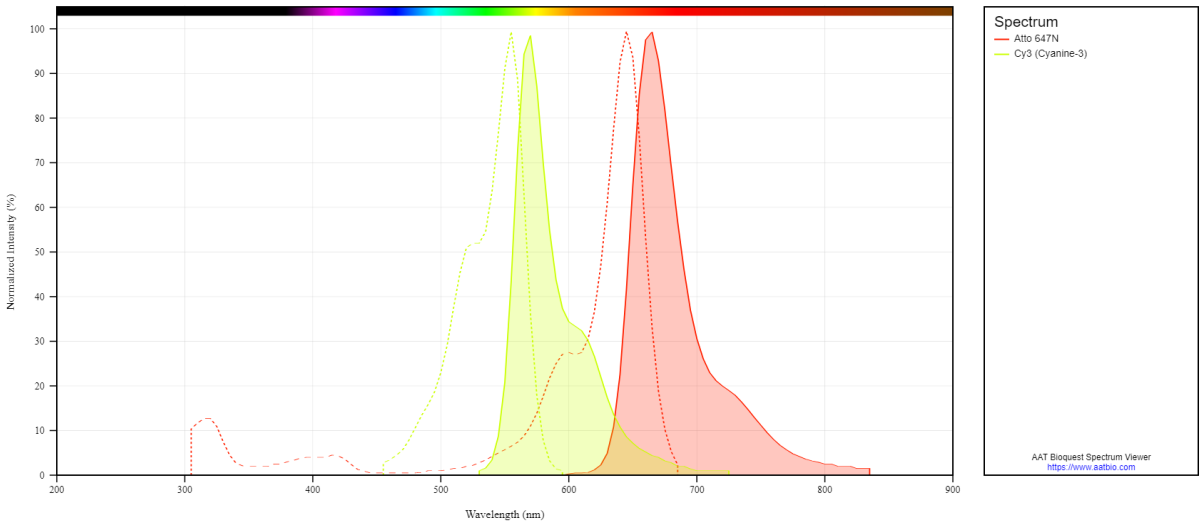


Figure 5.7: Spectral overlap between the donor dye Cy3 and the acceptor dye ATTO647N. Spectrum prepared with AAT Bioquest's interactive Spectrum Viewer.

To test the method and the distances between the fluorophores, two constructs of annealed oligos, utilizing the fluorophore staples, were designed. These are depicted in figures 5.8 and 5.9 below.

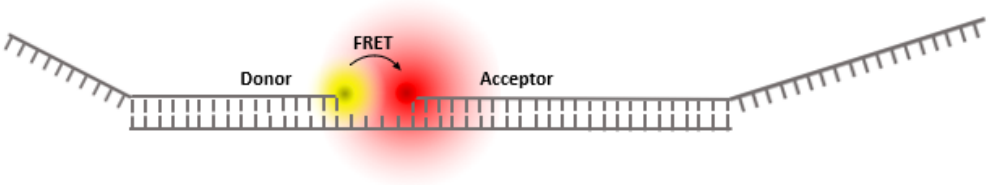


Figure 5.8: The D-A-ST oligo construct for a reference study of FRET detection in annealed and denatured state.

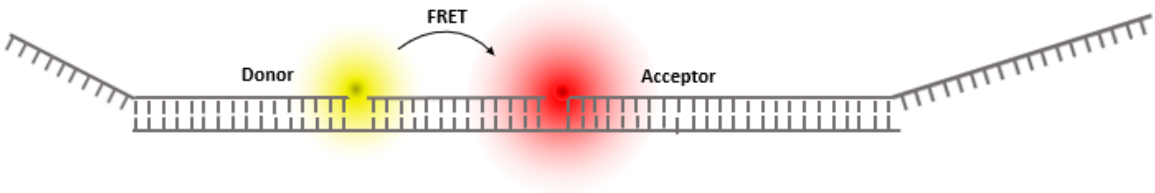


Figure 5.9: The D-A-M-LT oligo construct for a reference study of FRET detection in annealed and denatured state.

By the use of this technique it is assumed that it will be possible to detect if a DNA origami is intact or disintegrated. For the oligo constructs, FRET is expected to be detected when annealed. When denatured, FRET is expected to disappear. As well for DNA origami structures in their intact conformation, FRET is expected to occur, and for disintegrated structures, FRET is expected to disappear.

6. Experimental Results

In this chapter, results obtained during the project will be presented. This includes a section for each plasmid, the pCas9_GFP and lentiCas9-EGFP with subsections covering restriction enzyme analysis, sequencing of pCas9_GFP, amplification by PCR, and for lentiCas9-EGFP asymmetric PCR. Furthermore, results from simulations will be presented. A study of FRET detection for analysis of folded/unfolded states of a DNA origami structure will also be presented, carried out on a known model origami, 42hb.

6.1 Plasmid pCas9_GFP

In the following sections, the results from restriction enzyme digests, sequencing and amplification of plasmid pCas-GFP will be presented in that order.

6.1.1 Restriction Enzyme Digests

Restriction enzyme digestion was performed first time when starting the work with the first plasmid extract (p1). This digestion was performed with restriction enzymes *Mlu*I and *Bam*HI (Promega), performing both single and double digestions. The results can be seen in figure 6.1. Digesting with these enzymes should yield the full-length plasmid fragment of 9,302 bp for both single digests, and for the double digestion, fragments of 3,781 bp and 5,521 bp should be produced. It is seen that all fragments are fairly close to the expected size on the gel, however all slightly larger than the expected size; in lane 1 and 2, digestion with *Mlu*I and *Bam*HI, respectively, yields a fragment around 10-11 kb, while the double digestion shown in lane 3 yields fragments just below 4 and 6 kb. However, looking at the ladder, which is slanting, and considering the gel resolution, it was deemed sufficient. In lane 4, the plasmid is seen. Bands for both supercoiled and open circular conformations appear.

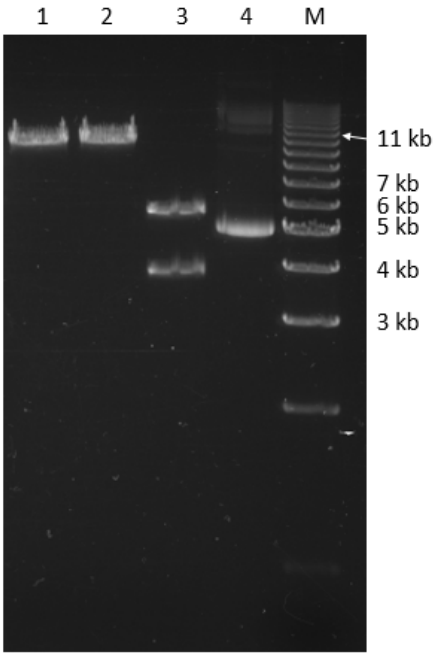


Figure 6.1: Restriction enzymes digests performed on pCas9_GFP. Lane 1: Single digestion with *Mlu*I, lane 2: Single digestion with *Bam*HI, lane 3: Double digestion with *Mlu*I and *Bam*HI, lane 4: Plasmid pCas9_GFP (p1). M denotes the M denotes the EZ load^{TM} Molecular Ruler 1kb.

When extracting the plasmid next time for other PCR sets, fraction (p2), digestions were performed again. This time with *Bam*HI-HF and *Sma*I, expecting to yield fragments at 8,665 bp and 637 bp, the results seen in figure 6.2. The small fragment in lane 3 is around 600 bp (white square), corresponding to the expected fragment of 637 bp. The larger fragments in both single and double digestions are, however, all slightly above the expected size, visible around approximately 9 kb in lane 3 for double digestion and around 10 kb in lane 1 and 2 for single digestion, where expected to be 8,665 bp and 9,302 bp, respectively.

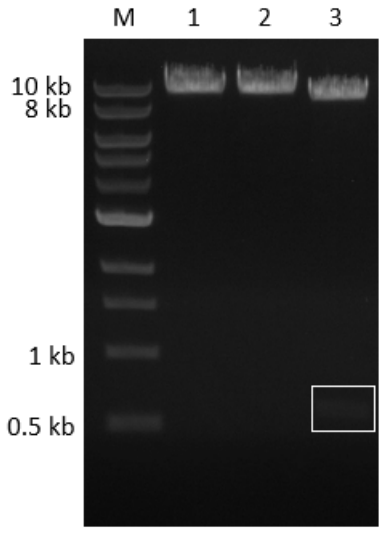


Figure 6.2: Restriction enzyme digestions with *Bam*HI-HF and *Sma*I. Lane 1: Digestion with *Sma*I, lane 2: Digestion with *Bam*HI-HF, lane 3: Double digestion with *Sma*I and *Bam*HI-HF. The white square indicates a fragment around 600 bp. M denotes the 1 kb DNA ladder (NEB).

Even though the gel resolution of large fragments was probably the reason for the too large fragments after the digestion, another digestion was performed to be certain. Here, *Bam*HI-HF and *Eco*RV-HF were used to digest the plasmid to fragments of 1,976 bp and 7,326 bp in the double digestion. The result can be seen below in figure 6.3. In lane 1 and 2 showing the result of single digestions where the whole plasmid length of 9,302 bp is expected to be seen, the fragment is just below 10 kb, and the resolution in this gel seems to be slightly better than in the last (figure 6.2). The fragment still looks slightly too large but it is also difficult to determine precisely. For the double digestion in lane 3, the smallest fragment is at 2 kb, corresponding well with the expected 1,976 bp. The largest fragment is just below 8 kb, nearly corresponding to the expected 7,326 bp but slightly too large for this fragment, however deemed sufficient for further procedure. Lane 4-6 shows the same digestion reactions with another plasmid fraction, (p2). Lane 7 and 8 shows the plasmid fractions (p1) and (p2), respectively, showing supercoiled (lower band) and open circular plasmid (upper band).

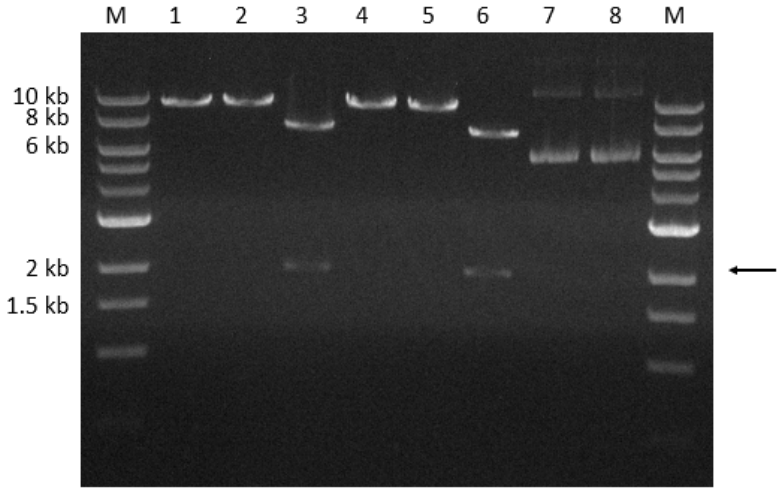


Figure 6.3: Restriction enzyme digests of pCas9_GFP with *Bam*HI-HF and *Eco*RV-HF. Lane 1-3: Digestion of plasmid extract (p1) with *Bam*HI-HF, *Eco*RV-HF and both, respectively, lane 4-6: Digestion of plasmid extract (p2) (not used in PCRs) with *Bam*HI-HF, *Eco*RV-HF and both, respectively, lane 7-8: Plasmid extract (p1) and (p2), respectively. M denotes the 1 kb DNA ladder (NEB). The arrow highlights fragments at low concentration in lane 3 and 6.

6.1.2 Sequencing of pCas9_GFP

Sequencing of the target region in pCas9_GFP was performed by MRC PPU DNA Sequencing and Services unit at University of Dundee. The plasmid sequence matched well with the given sequence in most regions, except a 2 bp insertion in the promoter and a few insertions, substitutions and a single deletion in the chimeric intron. However, also a 33 bp insertion followed by a 6 bp deletion were seen in the chimeric intron as visualized in figure 6.4.

Figure 6.4: Part of the sequencing result depicting a 33 bp insert and a 6 bp deletion in the actual plasmid. Illustration produced using Snapgene Viewer

6.1.3 PCR Amplification of pCas9_GFP

Several PCRs were performed in attempts to amplify the part of the plasmid that should be used as the DNA origami template. An illustration of the plasmid can be seen in figure 6.5 with sequence to be amplified marked in blue. A total of 7,797 bp were target region.

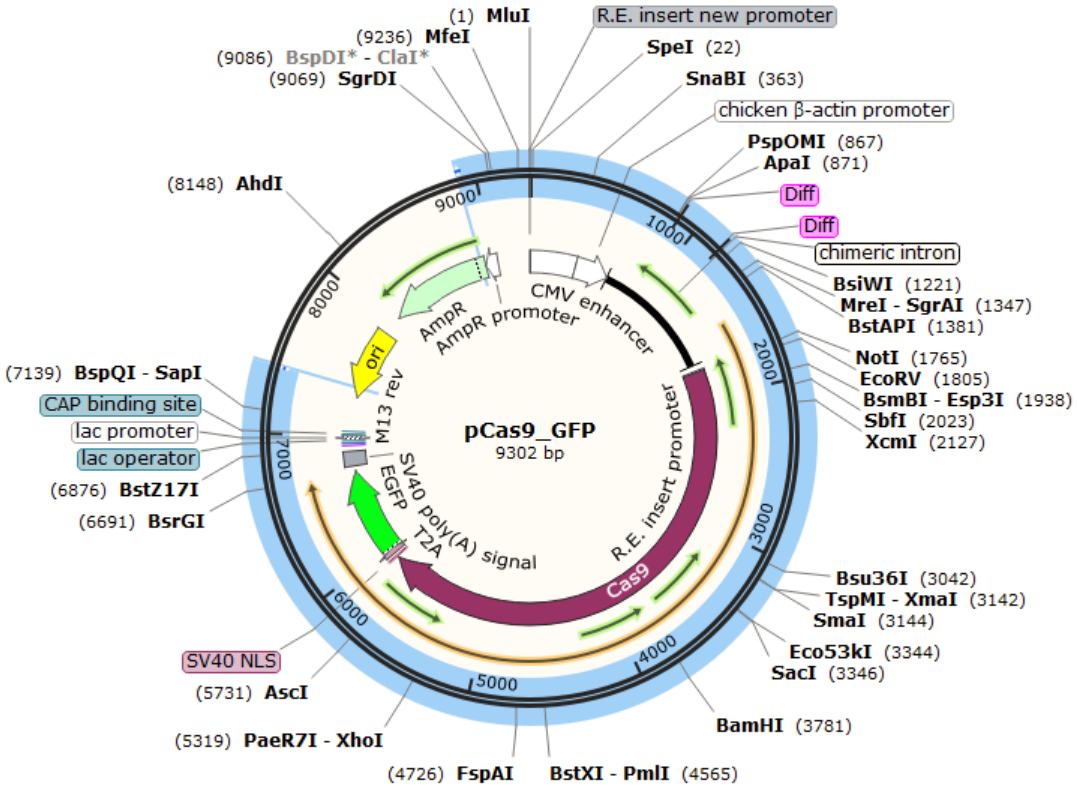


Figure 6.5: Illustration of plasmid pCas9_GFP. Amplified region marked blue. Illustration prepared in Snapgene Viewer.

Below in table 6.1 a list of all primer pairs is presented, showing the expected product size and the plasmid it was used for.

Table 6.1: A list of all primer pairs used in this project and their respective plasmids and target product size.

Primer pair/Plasmid	Forward primer	Reverse primer	Product size
1) pCas9_GFP	F.pcas9gfp.5phos	R.pcas9gfp_5biot	7,797 bp
2) pCas9_GFP	F.pcas9gfp.5phos	R.pcas9gfp	7,797 bp
3) pCas9_GFP	Fpcas9gfp5p18m	R.pcas9gfp_5biot	7,797 bp
4) pCas9_GFP	Adj. forw. scaf.	Rev real creates com.	7,453 bp
5) pCas9_GFP	GCregionpCas9forward	GCregionpCas9reverse	1,649 bp

For the many first PCRs until further mentioned, a plasmid extract (p1) was used as template. The concentration of was estimated using Nanovue to be around 200 ng/μL.

First amplification (a1) was produced following table 4.3 in section 4.2.5 with the results shown in figure 6.6. The annealing temperatures were varied from 52 °C to 60 °C (lane

1,2: 52 °C, lane 3,4: 55 °C, lane 5,6: 57 °C, lane 7,8: 60 °C), while lane 9 and 10 show the result from a two-step PCR reaction extending at 68 °C. With a target product of 7,797 bp (expected placement on the gel indicated by the arrow), it can be observed that most of the PCR products on the gel are too large; slightly above or around 8 kb, except for the lower fragment in lane 9 and 10. It is also seen that there are more than one product in all the lanes, however, in every reaction, one product is clearly most represented.

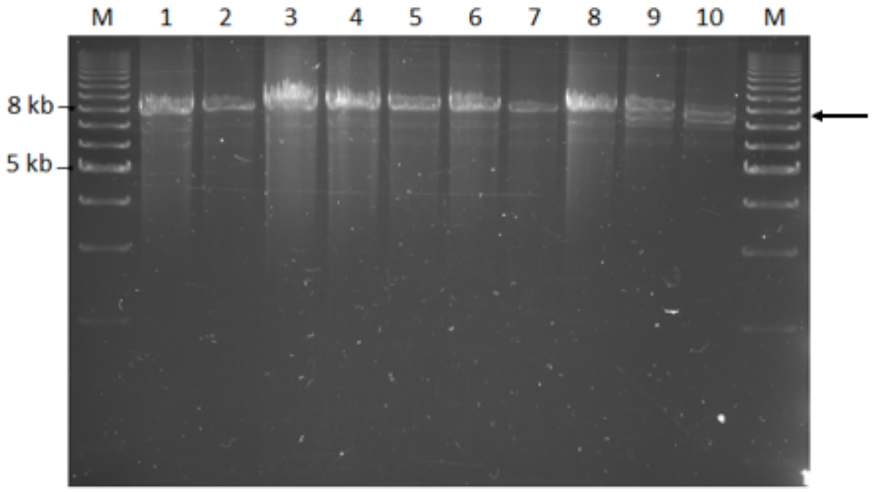


Figure 6.6: PCR (a1) on plasmid pCas9_GFP (p1) using A. G. LA polymerase and primer pair 1, extending 7,797 bp of the sequence. Lane 1,2: T_a at 52 °C, lane 3,4: T_a at 55 °C, lane 5,6: T_a at 57 °C, lane 7,8: T_a at 60 °C, lane 9,10: 2-step PCR with extension temperature at 68 °C. M denotes the EZ load^{TM} Molecular Ruler 1 kb used. The arrow indicates where the target product is expected on the gel.

In the second PCR run (a2) reproducing the result from lane 5, 6, 9 and 10 in figure 6.6 (T_a at 57 °C and two-step PCR) was attempted. Experiments with a reverse primer lacking the 5' biotin tag (R.pcas9gfp) were conducted as well to investigate eventual influence on the PCR reaction or the migration on the gel. Also, a Q5 High GC enhancer was added to some reactions to compensate for a high GC content in the amplified plasmid piece. The results can be seen in figure 6.7. The reaction setup was the same as in PCR (a1), except for the reactions where primer 'R.pcas9gfp' was used instead (lane 5, 6, 7) and the reactions where the Q5 High GC enhancer was added to 1x concentration (lane 8, 9, 10). Negative controls were carried out by adding no template DNA to the reaction (lane 2, 5, 8, 12).

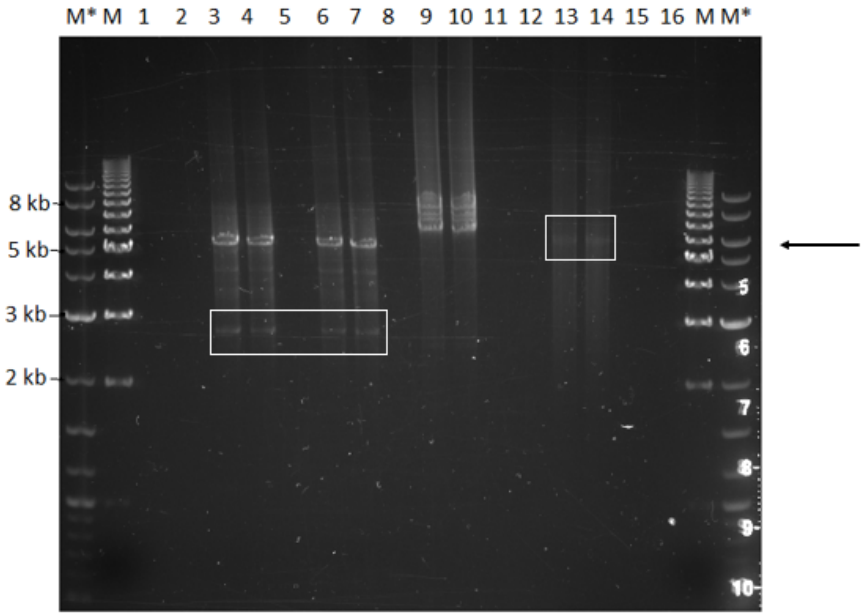


Figure 6.7: PCR (a2) on the plasmid pCas9_GFP (p1). Lane 2,3,4: Conditions as in (a1) with T_a at 57 °C (lane 2 as a negative control), lane 5, 6, 7: Conditions as in lane 2, 3, 4 but with reverse primer 'R.pcas9gfp' instead of 'R.pcas9gfp_5biot' (lane 5 as a negative control), lane 8, 9, 10: As in lane 2, 3, 4 but with 1x Q5 High GC enhancer added (lane 8 as a negative control), lane 12, 13, 14: As in (a1) two-step PCR (lane 12 as a negative control). Lane 1, 11, 15 and 16 are empty. M denotes the EZ load™ Molecular Ruler 1kb used, and M* a Quick-Load Purple 1kb Plus DNA Ladder. To be noted, the numbers in the left side of the image are from the gel tray in which the gel was placed.

As seen in figure 6.7, reproducing the results from first PCR was not successful. Lane 3 and 4 in figure 6.7 should correspond to lane 5 and 6 in figure 6.6, as it is performed with the same PCR conditions. However, the product sizes are very different, being around 8 kb in PCR (a1) and just above 5 kb in PCR (a2), indicated by an arrow. A product at low concentration can also be observed at around 2.8 kb (indicated by the white box) The same products are observed for the PCRs with the reverse primer without a 5' biotin tag. For the reactions including Q5 High GC enhancer (lane 9, 10), three or four products can be seen in the range between 6 and 8 kb. It can be suspected that one of these is the target product. The reactions made by a two-step PCR were also not reproducible, showing a product at low concentration just below 6 kb (lane 13, 14, indicated by a white box), opposed to products between 7 and 8 kb in the first PCR (a1). All negative controls showed no visible products on the gel.

To rule out any eventual influence of the dye on the DNA migration in the agarose gel, an identical gel was run and afterwards post-stained with the same dye. This did not affect the size of any of the products appearing on the gel, except yielding an additional

product around 1 kb, visible in many of the lanes, except for the controls. The small DNA fragment is probably visible on the post-stained gel, and not the pre-stained gel, as the dye migrates in the opposite direction of DNA, making small DNA fragments less visible if pre-staining is used. The result can be seen in Appendix A, figure A1.

As non-target products were observed on the previous gels, touch-down PCR was attempted to avoid mispriming. The reaction was let to anneal at first 72 °C and afterwards decreased 0.4 °C per cycle for 10 cycles an afterwards annealed at 68 °C for 20 cycles. Both 68 °C and 72 °C were tested as extension temperature and addition of Q5 High GC enhancer was also tested. The results of these reactions are seen in figure 6.8 in lanes 1 to 4, where no products with the correct size are obtained. Only products at 5 kb are seen, although in low concentration (marked by white square).

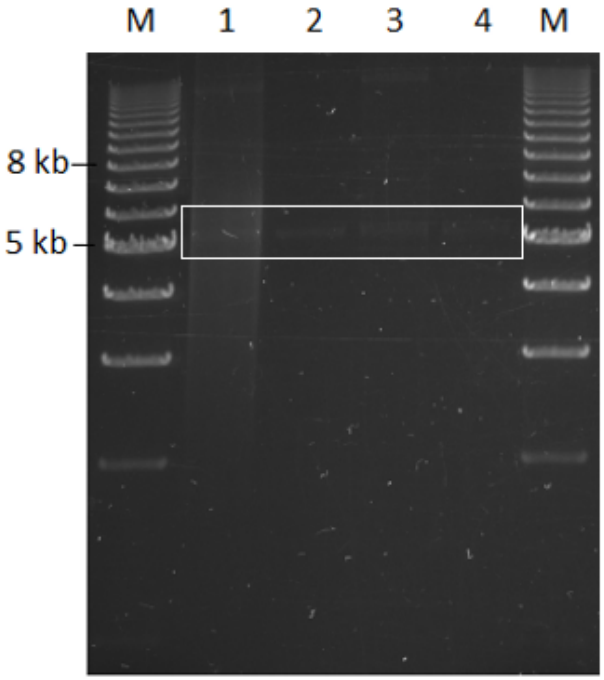


Figure 6.8: PCR products from touch-down PCR with A. G. LA polymerase on plasmid pCas9_GFP (p1). Lane 1: T_e at 68 °C, lane 2: T_e at 68 °C with Q5 High GC enhancer, lane 3: T_e at 72 °C, lane 4: T_e at 72 °C with Q5 High GC enhancer, lane 5: PCR (a1) product from lane 9 in figure 6.6, PCR (a1) product from lane 10 in figure 6.6, PCR (a1) product from lane 5 in figure 6.6, PCR (a1) product from lane 6 in figure 6.6. M: EZ load^{TM} Molecular Ruler 1kb.

As no target product was yet created, except from the possible products in PCR (a1), the forward primer 'F.pcas9gfp.5phos' was expected to have a too loose 3' end. A new primer lacking the two adenines at the 3' end was ordered and tested. The sequence and other properties can be seen in table 4.2. A new PCR run (a3) with the new primer and

otherwise same parameters as in (a1) was carried out, the results are seen in the gel image in figure 6.9. It is seen that non of the products represents the desired product at 7,797 bp.

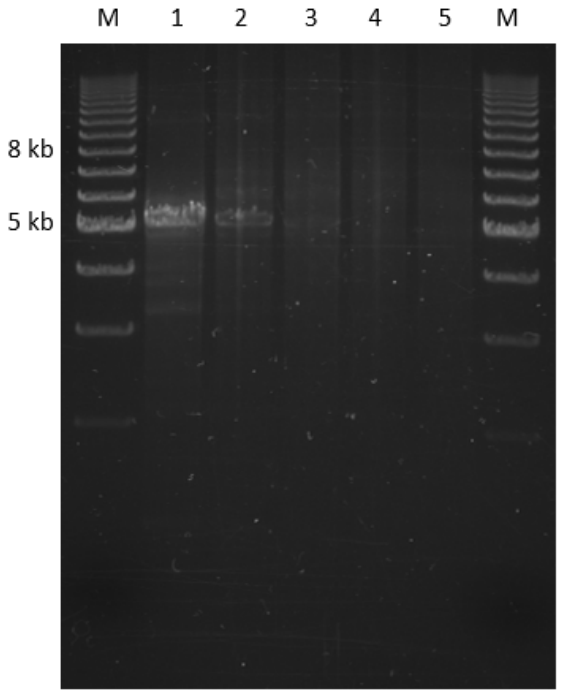


Figure 6.9: PCR (a3) with new primer 'Fpcas9gfp5p18m'. Lane 1: T_a at 52 °C, Lane 2: T_a at 55 °C, Lane 3: T_a at 57 °C, Lane 4: T_a at 60 °C, Lane 5: T_a at 68 °C (two-step PCR). All reactions carried out as in table 4.3.

To be noted, a OneTaq 2xmastermix was also tested as it is supposed to amplify robustly GC rich and long templates. It was tested at standard conditions and also with the Q5 High GC enhancer included in the kit, and the result can be seen in appendix A, figure A2. Predominantly, a 5 kb product was produced in this attempt.

A lower amount of template (30ng) was also tested in PCR (a4). Else, the PCR was carried out as in (a1), testing annealing temperatures at both 57 °C and 70 °C. Also , OneTaq 2x Master Mix was again tried at T_a at 57 °C (lane 3) but otherwise prepared as in table A2. The results can be seen in figure 6.10.

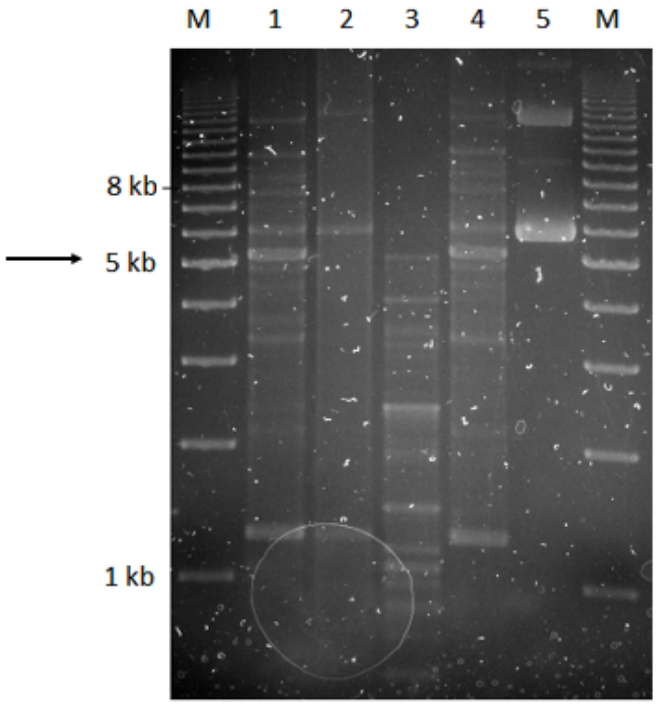


Figure 6.10: PCR (a4) with varied template amount of plasmid pCas9_GFP (p1). Lane 1: 30ng template and T_a at 57 °C, lane 2: 100ng template and T_a at 70 °C, lane 3: OneTaq 2x Master Mix, lane 4: 100ng template and T_a at 57 °C, lane 5: Template DNA.

Another variable in this PCR was the thermocycler being preheated to 94 °C and then placing the samples directly into the hot block. This was done in an attempt to reduce the possibility of nonspecific amplification. It is seen that in both lane 1, 3 and 4 many nonspecific products have been produced, however, a product around the targeted 7,797 bp is also visible in both lane 1 and 4. Comparing it to figure 6.7 and 6.9, the product just above 5 kb (marked by arrow) is visible here as well. In lane 2 with 70 °C annealing temperature this band is not seen and only three other products are visible around 1.3 kb, 6 kb and 13 kb. The two upper products are much in line with the size of the template product in lane 5. However, no target product is produced with T_a at 70 °C.

A PCR was carried out with a BioTaq Red DNA polymerase from Bioline as well to be able to test the influence of $MgCl_2$ concentration which was varied between 1.5 and 3 mM in 0.5 increments. Furthermore, different annealing temperatures were tested. The method and results can be seen in Appendix A in figure A3 and tableA2. Only products in the low kb range were observed. Also, it was tested if the PCR could be carried out with products from PCR (a1) as template, but this resulted in smeared lanes on the gel.

So far, no product with the correct size had been observed, except from a possible cor-

rect sized product in PCR (a1). The problems were assigned to the GC-rich template, especially around the promoter region. Therefore, additives like betaine and DMSO were considered and tested. Now, the plasmid extract solution was changed to a new, (p2), which was estimated to a concentration of around 700 ng/μL using Picodrop. In a first trial, 1, 1.5 and 2 M betaine together with 1.5, 2.5 and 5% DMSO were tested. The method can be seen in table 4.4 and the result in figure 6.11. Primer pair 4 was used. A dominant product in all lanes is seen between 5 and 6 kb, whereas products are also seen around 9 and between 3 and 4 kb. None of the products have the expected size around 7.5 kb. A slight change regarding the produced products can be observed over the range of betaine and DMSO concentrations, and the bands seem sharpest at 2 M betaine. A product just below 4 kb that is visible at 2M betaine and 2.5 % DMSO disappears at 5 % DMSO.

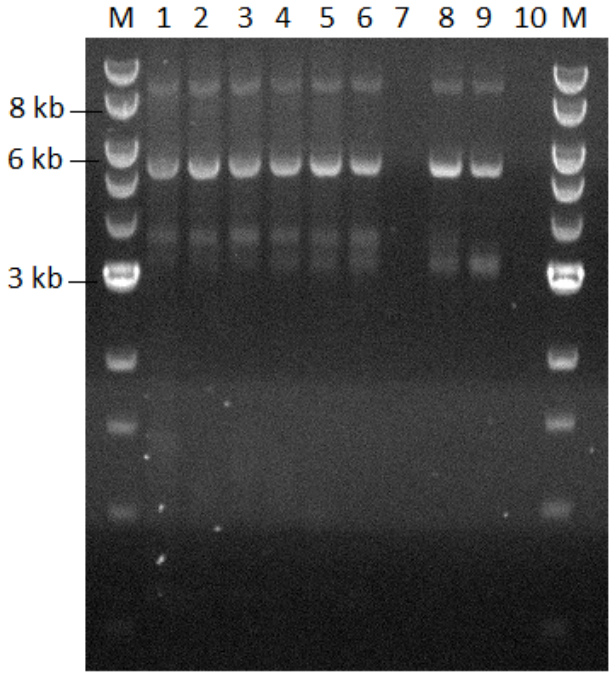


Figure 6.11: PCR (a6) with Taq polymerase, betaine and DMSO additives. Lane 1-3: 1 M betaine and 1.5, 2.5, 5.0 % DMSO, respectively, lane 4-6: 1.5 M betaine and 1.5, 2.5, 5.0 % DMSO, lane 7: Empty, lane 8-9: 2 M betaine and 2.5, 5 % DMSO, lane 10: Empty. M denotes the 1 kb DNA ladder (NEB).

A PCR with the primer set 3 was made using the same conditions as before with 3 % DMSO and 1 M betaine, and 5 % DMSO and 2 M betaine, both conditions tested with 1/10x, 1/50x and 1/100x template (p2). The thermal conditions were instead 96 °C, 5 min; (96 °C, 0.5 min; 50 °C, 0.5 min; 72 °C, 8 min)x30; 72 °C, 15 min. The same product around 6 kb is visible in lanes 1-5, with reduced intensity and smear when a

lower amount of template is used. The reduced smear is also obtained in the sample with highest template amount and 5% DMSO and 2M betaine, although the product still appears at 6 kb, not reflecting the desired product size.

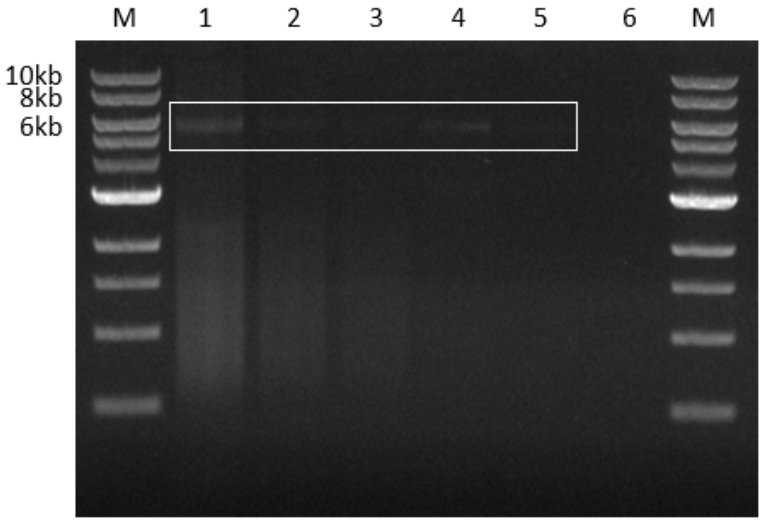


Figure 6.12: PCR (a7) with Taq polymerase and primer pair 3. Lane 1-3: 3% DMSO, 1M betaine and 1/10x, 1/50x and 1/100x template (p2), respectively, lane 4-6: 5% DMSO, 2M betaine and 1/10x, 1/50x and 1/100x template (p2), respectively. M denotes 1 kb DNA ladder (NEB).

It was decided to test the effect of betaine and DMSO additives together with A. G. LA polymerase. The same procedure as in table 4.3 was followed, except adding 0.5 μM primers instead of 0.4 μM, using a different template stock solution (p2), and adding 3 % DMSO and 1 M betaine. Two different annealing temperatures were used; 50 °C and 53 °C. In figure 6.13 it is seen that all samples yielded products in the correct size region, though the bands are smeared and it is difficult to assess whether there are one or two products. The latter was further confirmed running another gel (see figure A4 in Appendix A) with half the volume loaded where two distinct bands are visible with the lower product having the expected product size (lane 3 and 4). Besides this, the same reaction but with 10 % DMSO was conducted to see if the product could be narrowed down to be only the correct size but this yielded 3 products; one around 2.7 kb, one at 6 kb and a nearly invisible one around the expected product size (lane 1 in figure A4).

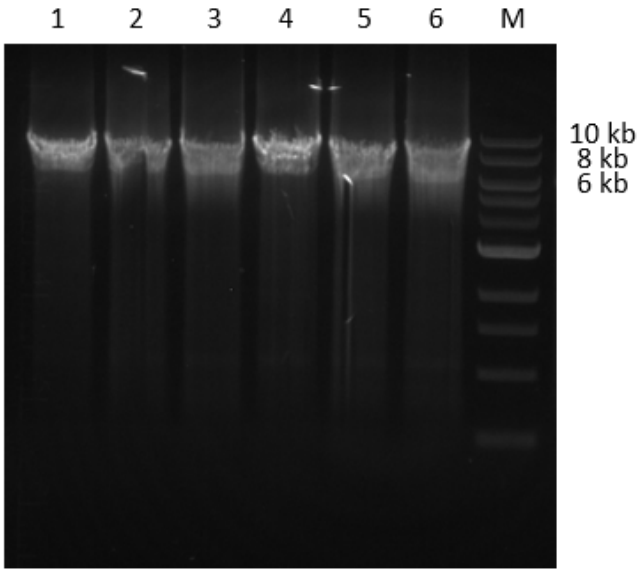


Figure 6.13: PCR testing the affect of betaine and DMSO additives in combination with A. G. LA polymerase. Prepared as in 4.3 but with 0.5 μM primers and 1M betaine, 3 % DMSO. Lane 1-3: T_a at 50 °C and 1/100x, 1/10x and 1x template (p2), respectively. Lane 4-6: T_a at 53 °C and 1/100x, 1/10x and 1x template (p2), respectively. M denotes a 1 kb DNA ladder (NEB).

A substitution of 3/4 of all dGTP in the dNTP solution with a 7-deaza-dGTP nucleotide was also tested together with the A. G. LA polymerase which had to date provided the most promising results using primer pair 3. It was tested both in combination with 5 % DMSO and 1 M betaine, but also with a 1xCES solution (0.54 M betaine, 1.34 mM DTT, 1.34 % DMSO, 11 μg/mL BSA). The result can be seen in figure 6.14, where the upper product in lane 3 with the substituted dGTP might be the right product with the expected size of 7,797 bp (marked with arrow), however it can not be fully verified.

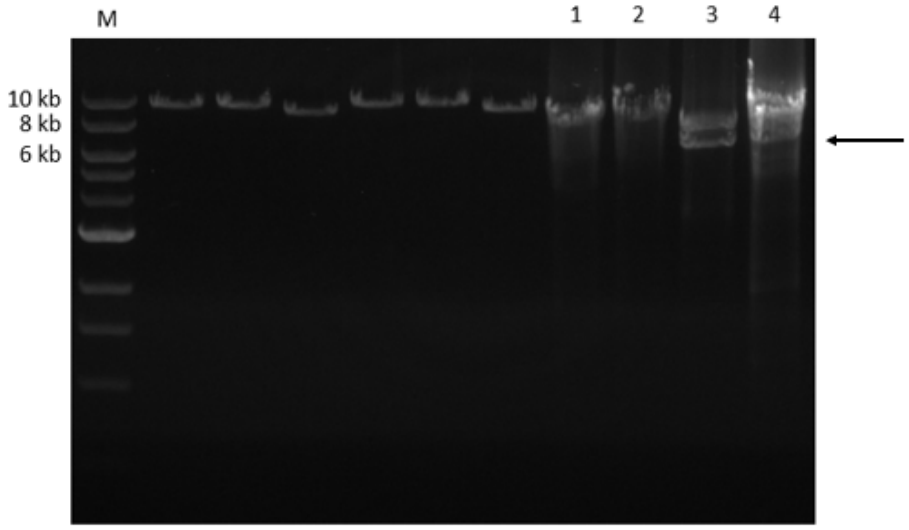
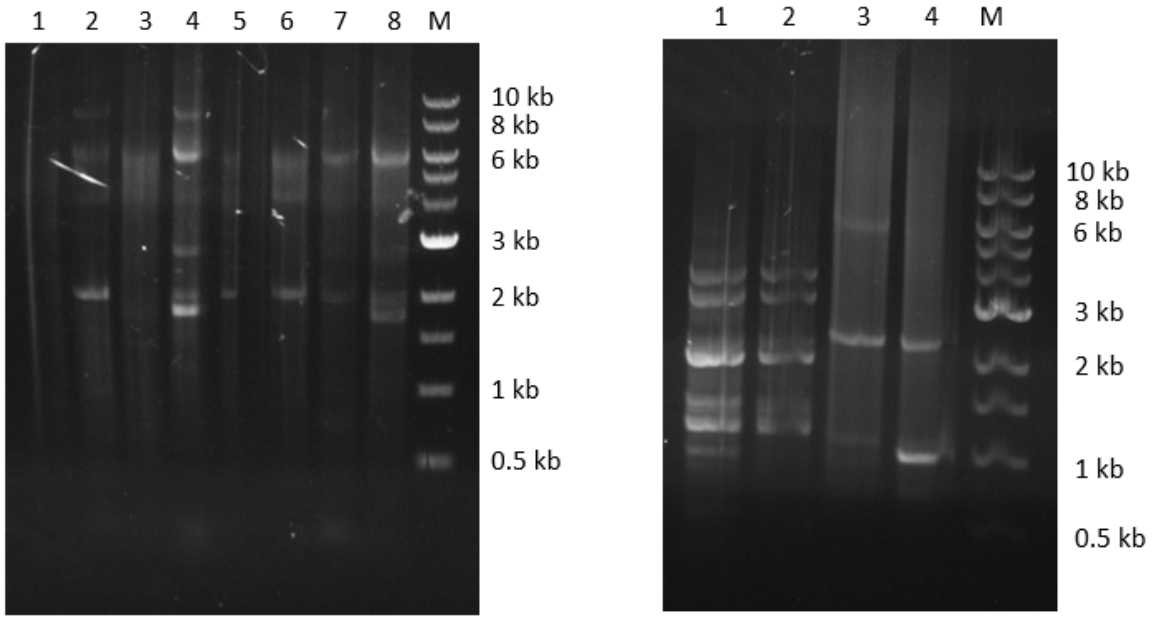


Figure 6.14: PCR evaluating substitution of 3/4 of the dGTP amount with a 7-deaza-dGTP. All 4 PCRs were set up with 0.4 mM dNTPs (each), 0.5 μM primers (primer pair 3), 1 μL 1/100 template and 1.25 U of A.G. LA polymerase with the matching buffer. Lane 1: Previous conditions including 5% DMSO and 1M betaine together with standard dNTP solution, lane 2: 1xCES solution together with the standard dNTP solution, lane 3: 5 % DMSO, 1M betaine and 7-deaza-dGTP, lane 4: 1xCES solution and 7-deaza-dGTP. M denotes a 1 kb DNA ladder (NEB). The arrow indicates product that could be the target product. All lanes not marked are not PCR products and can be ignored. The reactions were all performed by following cycles: 94 °C, 1 min; (98 °C, 10 sec; 50 °C, 30 sec; 68 °C, 8 min)x30; 72 °C, 10 min.

Phusion^{TM} High-Fidelity DNA Polymerase is a polymerase typically used for difficult templates, including GC-rich and for long-range PCRs up to 20 kb. As the plasmid has a high GC content, this polymerase was tested. The reactions were set up with the two enclosed buffers; a Phusion HF buffer and a Phusion GC buffer. Buffers were added in concentrations of 1x, providing a final $MgCl_2$ concentration of 1.5 mM, dNTPs at 200 μM each, primers at 0.25 μM and 1/100x template (p2). Both 0.05 and 0.02 U/μL polymerase were tested. As addition of DMSO at 3 % for GC rich sequences was suggested by the protocol, this was tested as well. As seen in figure 6.15a the tests were first run with primer pair 4, yielding several products but none with the correct size of 7,453 bp. It can be observed that no products are produced using the HF buffer, except for in lane 7. Primer set 3 was therefore only tested with GC buffer. The results of this are seen in figure 6.15b. This only yielded products at primarily smaller sizes.



(a) **(b)**

Figure 6.15: PCR amplification with Phusion^{TM} High-Fidelity DNA Polymerase and 1.5mM $MgCl_2$, 200 μM dNTP (each), primers at 0.25 μM and 1/100x template (p2) (general conditions). (a): All reactions with primer set 4. Lane 1: 1xHF buffer and 0.05 U/μL polymerase, lane 2: 1xGC buffer and 0.05 U/μL polymerase, lane 3: 1xHF buffer, 0.05 U/μL polymerase and 3 % DMSO, lane 4: 1xGC buffer, 0.05 U/μL polymerase and 3 % DMSO, lane 5-8: Same sequence of PCRs as lane 1-4 but with 0.02 U/μL polymerase. M denotes a 1 kb DNA ladder (NEB). (b): All reactions with primer set 3, 1xGC buffer and 3 % DMSO in addition to the general conditions. Lane 1: 0.05 U/μL polymerase, lane 2: 0.02 U/μL polymerase, lane 3: 0.05 U/μL polymerase and touchdown PCR, lane 4: 0.02 U/μL polymerase and touchdown PCR. M denotes a 1 kb DNA ladder (NEB). For PCR thermal conditions, see table A3 in Appendix A.

6.1.4 PCR Amplification of GC-rich Region

To simplify the amplification procedure and to deduce if the promoter region with a high GC content was in fact the problematic part of the plasmid, primers were selected to only amplify this part. Primer pair 5 in table 6.1 was used, supposed to yield a product of 1,649 bp with a GC content of 69 %. This is further illustrated in figure 6.16 below.

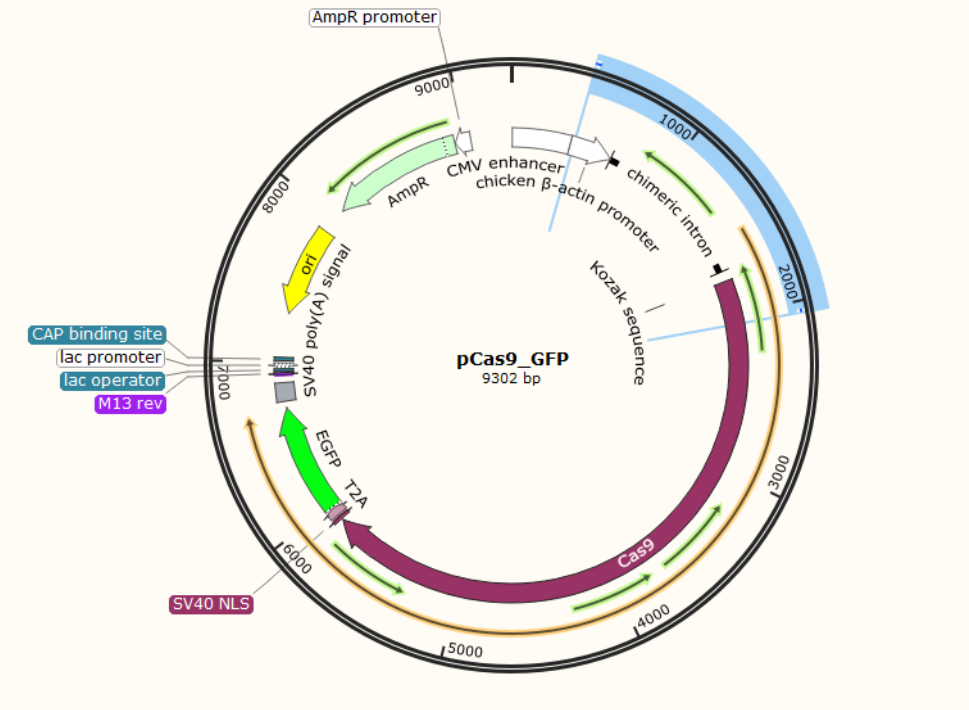
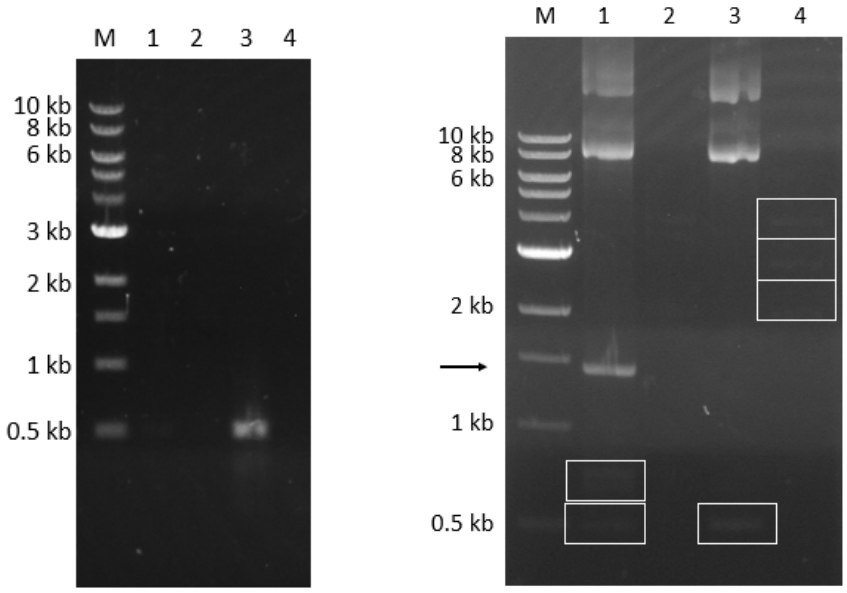


Figure 6.16: Illustration of the amplified GC-rich promoter region of pCas9_GFP (marked in blue). Illustration prepared in Snapgene Viewer.

Attempts were performed with both Taq polymerase and Phusion polymerase as seen below in figure 6.17. For Taq polymerase, two PCR buffers were tested, one with KCl and one with $(NH_4)_2SO_4$, both normal PCR and touch-down PCR were employed, and the addition of a preCESII solution (final concentrations of 0.56 M betaine, 1.4 mM DTT and 1.4 % DMSO) was utilized as suggested by *Ralser et al.* (2006) for GC-rich templates. [47] Only a product around 500 bp was obtained in these reactions, using the Taq KCl buffer and touch-down PCR. This product can be seen in lane 3, figure 6.17a.

For Phusion polymerase, two different concentrations of plasmid were used, and either 1x CG buffer together with 3 % DMSO or 1xHF buffer and 1xCES solution (0.54M betaine, 1.34 mM DTT, 1.34 % DMSO, 11 μg/mL BSA) were used, proposed by *Ralser et al.* (2006). [47] PCR with large template amount yielded most products. One product at around 1.4 kb was obtained with 1xGC buffer and 3 % DMSO (lane 1, figure 6.17b), otherwise only small products in low concentration were seen (marked in white box), together with products at 8 kb and above the DNA ladder for both reactions with high template amount (lane 1 and 3). PCR with low template concentration, HF buffer and 1xCES yielded products around 2, 3 and 4 kb at very low concentration (marked in white boxes).



(a) **(b)**

Figure 6.17: PCR of the GC-rich region of the promoter region. (a) Amplification using Taq polymerase at 0.04 U/μL, $MgCl_2$ at 1.5 mM, dNTPs at 0.2 mM (each), primer pair 5 at 0.5 μM each, 1/100x template (p2), preCESII solution giving final concentrations of 0.56 M betaine, 1.4 mM DTT and 1.4 % DMSO, and 1x of either Taq buffer with KCl or $(NH_4)_2SO_4$. Lane 1: Previous mentioned conditions, KCl buffer and standard thermocycling conditions, lane 2: $(NH_4)_2SO_4$ buffer and standard thermocycling conditions, lane 3: KCl buffer and touchdown PCR, lane 4: $(NH_4)_2SO_4$ buffer and touchdown PCR. For PCR thermocycling conditions, see table A4 in Appendix A. (b) Amplification of GC-rich region with Phusion polymerase, all reactions containing 0.2 mM dNTPs (each), 0.5 μM of each primer in primer pair 5, 0.05 U/μL polymerase. Lane 1: 1xGC buffer, 1x template and 3 % DMSO, lane 2: 1xGC buffer, 1/100x template (p2) and 3 % DMSO, lane 3: 1xHF buffer, 1x template and 1xCES solution (0.54M betaine, 1.34 mM DTT, 1.34 % DMSO, 11 μg/mL BSA), lane 4: 1xHF buffer, 1/100x template (p2) and 1xCES solution (0.54M betaine, 1.34 mM DTT, 1.34 % DMSO, 11 μg/mL BSA). M in both (a) and (b) denotes the 1 kb DNA ladder (NEB). Products at low concentration which are difficult to see are marked in white boxes.

Furthermore, reactions with Taq polymerase and an increased extension temperature to 80 °C were performed as shown in figure 6.18. In lane 3, a product around 1.6 kb is seen (marked with arrow), which corresponds to the expected size of the GC-region target. This was obtained with touchdown PCR, 1xCES solution (0.54M betaine, 1.34 mM DTT, 1.34 % DMSO, 11 μg/mL BSA), 1/10x template (p2) and the $(NH_4)_2SO_4$ buffer, running at the conditions seen in table A5 in Appendix A. This reaction was performed by Chris Stubberup, another MSc student working on the same plasmid.

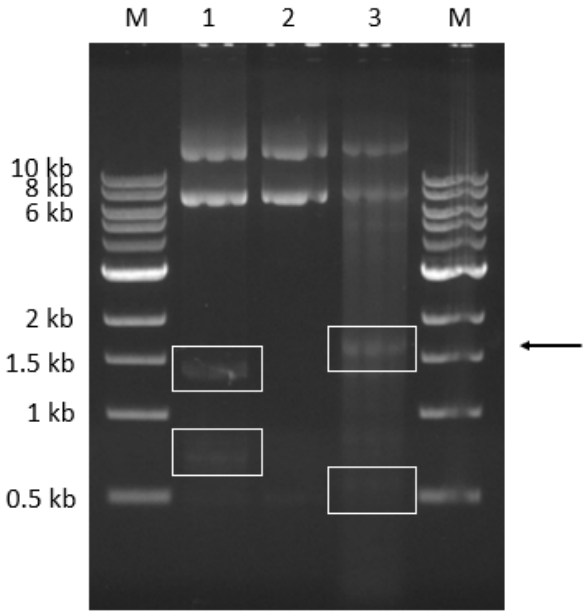


Figure 6.18: PCR of GC-rich region of plasmid pCas9_GFP with an extension temperature at 80 °C. Lane 1: As in lane 1 in figure 6.17b but with 5 % DMSO, lane 2: As in lane 3 in figure 6.17b but with 1xGC buffer instead, lane 3: Touchdown PCR performed with 0.02 U/μL Taq polymerase, 0.2 mM dNTPs (each), 1.5 mM $MgCl_2$, 1xCES solution (0.54M betaine, 1.34 mM DTT, 1.34 % DMSO, 11 μg/mL BSA), 1/10x template (p2) and $(NH_4)_2SO_4$ buffer, running at the conditions seen in table A5 in Appendix A (performed by Chris Stubberup, MSc student in the same course). Products at low concentrations difficult to observe are marked in white boxes. Arrow indicates product with the expected size.

Even though the expected product seems to appear, the product has a low concentration, and with the high extension temperature required for it to succeed, it was unreasonable that satisfactory results with the whole 7.8 kb target region could be obtained. Therefore, the plasmid was changed to a less GC-rich one.

6.1.5 Analysis of GC-region

To study the reason for repetitive products in the GC-rich region, the sequence was investigated for potential secondary structures. A potential hairpin could be pointed out as seen in figure 6.19 below.

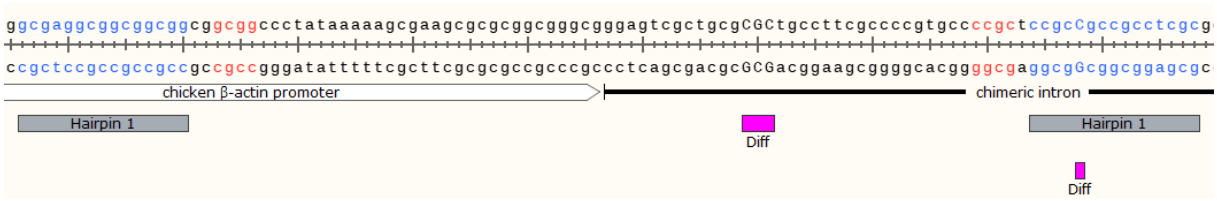


Figure 6.19: Illustration of the suspected hairpin stretching over the 3' end of the promoter and the 5' end of the chimeric intron. The illustration is made using SnapGene Viewer. The colours represent which regions that are expected to bind with each other to create the hairpin, and the longest fragment of the hairpin is also indicated by the grey features.

Further analysis of this sequence was performed using the RNA WebServer from Institute for Theoretical Chemistry, University of Vienna, yielding the prediction in the following graphical representation in figure 6.20. The colour representation is denoting base-pairing probabilities, but for unpaired regions, it will be the probability of bases being unpaired. [48, 49]

Figure 6.20: A minimum free energy prediction of the proposed hairpin in figure 6.19, produced using the RNA WebServer from Institute for Theoretical Chemistry, University of Vienna. The ruler indicates the probability of base-pairing. For unpaired bases, it is the probability of bases being unpaired. [49]

6.2 Plasmid lentiCas9-EGFP

Since attempts creating a single product of the expected size from plasmid pCas9_GFP with large enough yield and minimal non-specific amplification was not successful, it was decided to order another plasmid. The plasmid lentiCas9-EGFP has essentially the same genetic elements, including the desired sequence for Cas9 and EGFP. Instead of the GC-rich chicken β-actin promoter and chimeric intron, lentiCas9-EGFP has another promoter, the EF-1α core promoter with a GC content of only 59 % compared to 73 % for the chicken β-actin promoter and chimeric intron in pCas9_GFP.

6.2.1 Restriction Enzyme Digestion of lentiCas9-EGFP

lentiCas9-EGFP was first cut with *Bam*HI-HF and *Not*I-HF (NEB) in CutSmart buffer to yield segments of 7,640 and 5,537 bp. The result of digestions of three different plasmid extracts can be seen below in figure 6.21 in lane 2, 4 and 6. The products are in all lanes approximately around 7 and 9 kb which are both too large compared to the expected results.

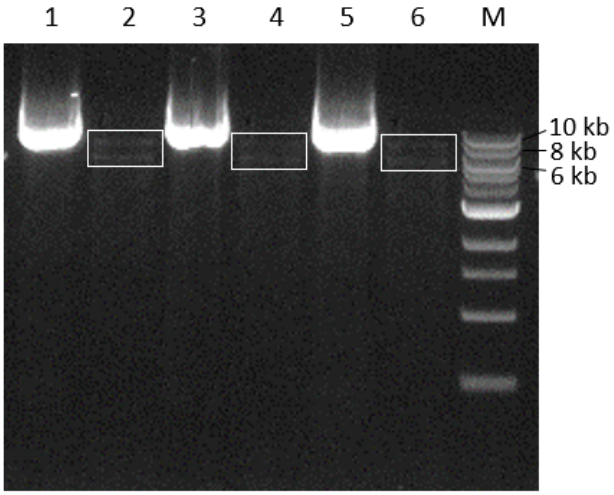


Figure 6.21: Plasmid lentiCas9-EGFP digested with BamHI-HF and NotI-HF. Lane 1, 3 and 5 showing the plasmid extracts 1, 2 and 3, respectively, while lane 2, 4 and 6 are showing the digests of plasmids from extract 1, 2 and 3, respectively. M denotes the 1 kb DNA ladder (NEB).

As the digestion shown in figure 6.21 yielded too large fragments, another digestion was performed. This was done with the result seen in 6.22. Lane 1 shows the result of the digestion performed with *Sap*I, *Pvu*I-HF and *Bam*HI-HF, expected to yield fragments of 1,379 bp, 2,033 bp, 4,186 bp and 5,579 bp. All fragments on the gel are corresponding fairly well to the expected sizes (given left to the corresponding bands in the gel). Lane 2 shows the digestion with *Sac*I-HF, *Pvu*I-HF and *Eco*RI-HF, expected to yield fragments at 191 bp, 1,174 bp, 1,422 bp, 2,720 bp, 3,608 bp and 4,062 bp. The fragments on the gel are corresponding fairly well with the expected sizes (indicated right to lane 2 (in lane 3)) for this digestion as well. It is, however, not possible to ensure the presence of a fragment at 1,174 bp due to the shadow caused by the gel doc system overlapping just at the spot. Also, the fragment at 191 bp is not visible, probably due to the resolution of the gel, but since the other fragments are matching in both digestion reactions, it was deemed sufficient to justify that the plasmid was in fact the expected plasmid.

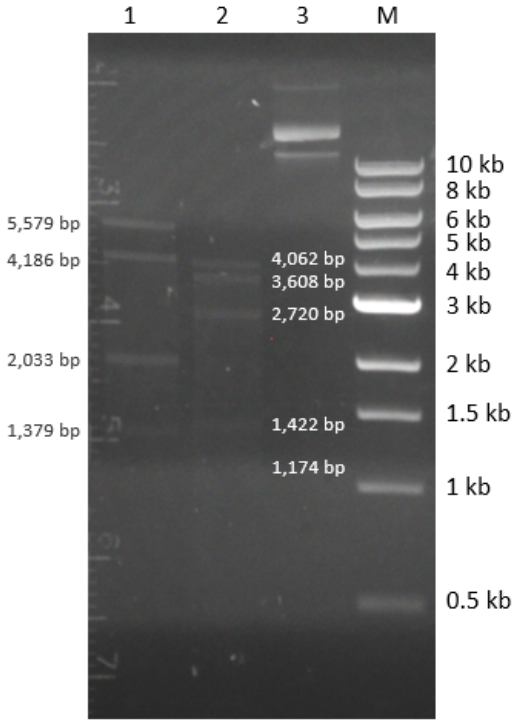


Figure 6.22: Plasmid lentiCas9-EGFP digested with; lane 1: SapI, PvuI-HF and BamHI-HF, lane 2: SacI-HF, PvuI-HF and EcoRI-HF. Lane 3 shows the plasmid fracion digested in lane 1 and 2. M denotes the 1 kb DNA ladder (NEB).

6.2.2 PCR Amplification of lentiCas9-EGFP

The primer pair (6) 'Forw-Lenti-outEF1a' and 'Rev-Lenti-outWPRE' with a product size of 6,064 bp was used in the following amplifications. An illustration of the amplified part (marked in blue) can be seen below in figure 6.23.

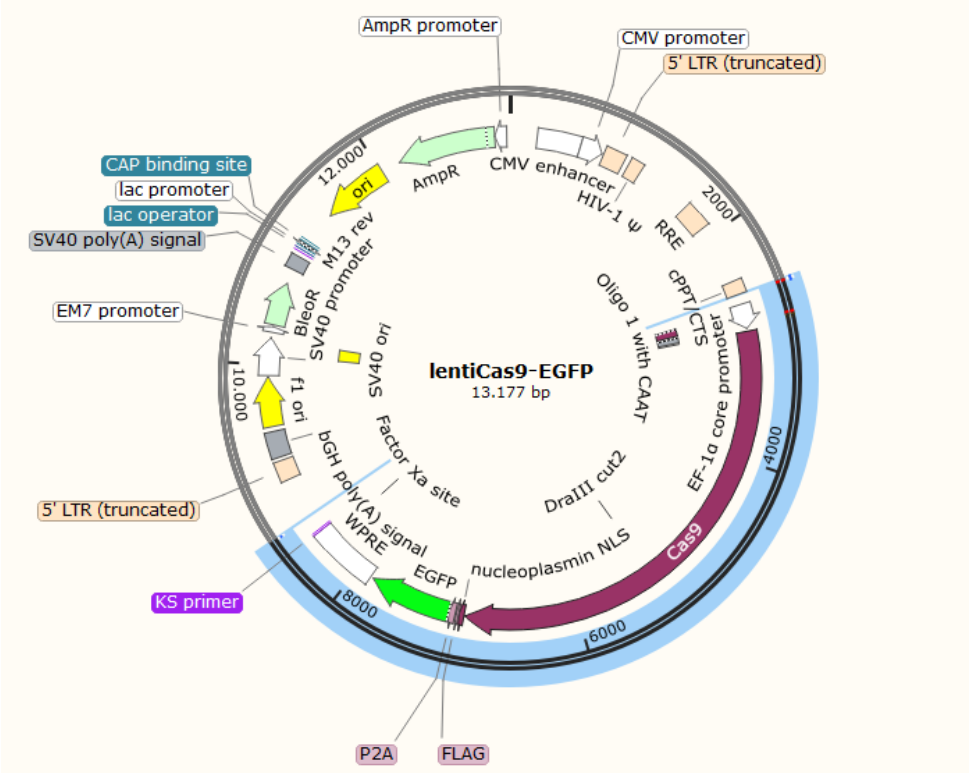


Figure 6.23: Illustration of plasmid pCas9_GFP. Amplified region marked blue. Illustration prepared in Snapgene Viewer.

The plasmid fraction (lp1) which was used for PCRs in the following results was estimated to have a concentration around 40 ng/μL using Picodrop. First PCR on lentiCas9-EGFP was run with 0.04 U/μL Taq polymerase and the standard 1xTaq buffer with KCl. 1.5 mM $MgCl_2$, 0.2 mM dNTPs (each), 0.4 μM primers of primer pair 6 (each) and 1x, 1/10x or 1/100x template (lp1) were used in the reaction. The cycle conditions were as in table 4.4 but with an annealing temperature at 47 °C and and extension time of 8 minutes. Also, a negative control without template was performed. The results can be seen in figure 6.24 below.

Lane 1 and 2 shows product at 6 kb, corresponding to the expected size, however only a smear appears in lane 3 where template from the plasmid extract stock (lp1) was used. The negative control shows, as expected, no products in lane 4.

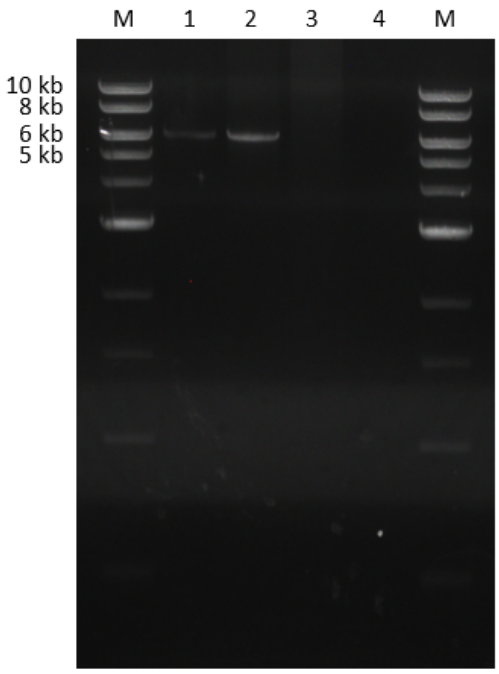


Figure 6.24: PCR on plasmid lentiCas9-EGFP with with 0.04 U/μL Taq polymerase, 1xTaq buffer with KCl, 1.5 mM $MgCl_2$, 0.2 mM dNTPs (each), 0.4 μM primers of primer pair 6 (each) and in lane 1; 1/100x template (lp1), lane 2; 1/10x template (lp1) and lane 3; 1/1 template, all of lentiCas9-EGFP; lane 4 shows a negative control without template, all with 1 μL products loaded. The cycle conditions were as in table 4.4 but with an annealing temperature at 47 °C and and extension time of 8 minutes. M is the 1 kb DNA ladder (NEB).

As the expected product was obtained, another PCR with same conditions but double volume was prepared with 1/10x template (lp1) to prepare for asymmetric PCR. The only difference was the extension temperature which was shortened to 6 min as it should be enough for this product size. The 6 kb PCR product was cut out of the gel and extracted as described in section 4.2.7, and the extracted product was run on a gel to confirm, see figure 6.25, lane 2, below. As only a smear was obtained when using 1xtemplate (lp1) which was the stock solution, while product occurred for the dilutions of template, these solutions were loaded on the gel as well. Lane 1 shows 1/10x template (lp1), clearly confirming plasmid with supercoiled and open-circle formations of it, while lane 3 with the stock solution (lp1) shows a smear and only a faint plasmid band, indicating degradation of the plasmid. The same is the case for other extracts (lp2) and (lp3) in lane 4 and 5, respectively.

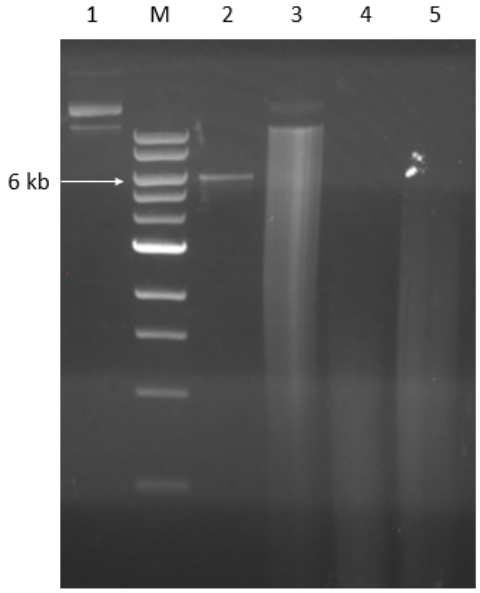


Figure 6.25: Lane 1: 1/10x lentiCas9-EGFP extract (lp1), lane 2: Product after gel extraction of PCR product on lentiCas9-EGFP, described in section 6.2.2, lane 3: lentiCas9-EGFP extract (lp1), lane 4: lentiCas9-EGFP extract (lp2), lane 5: lentiCas9-EGFP extract (lp3). M denotes the 1 kb DNA ladder (NEB).

With the extracted PCR product as template, an asymmetric PCR was carried out. It was prepared and run with the same conditions as the standard PCR, but with only the reverse primer, 'Rev-Lenti-outWPRE', as the antisense strand would be the desired strand for DNA origami folding. The primer was added at three different concentrations; 0.4, 0.6 and 0.8 μM. 1 μL of the extracted PCR product was added as template. The PCRs yielded no observable product.

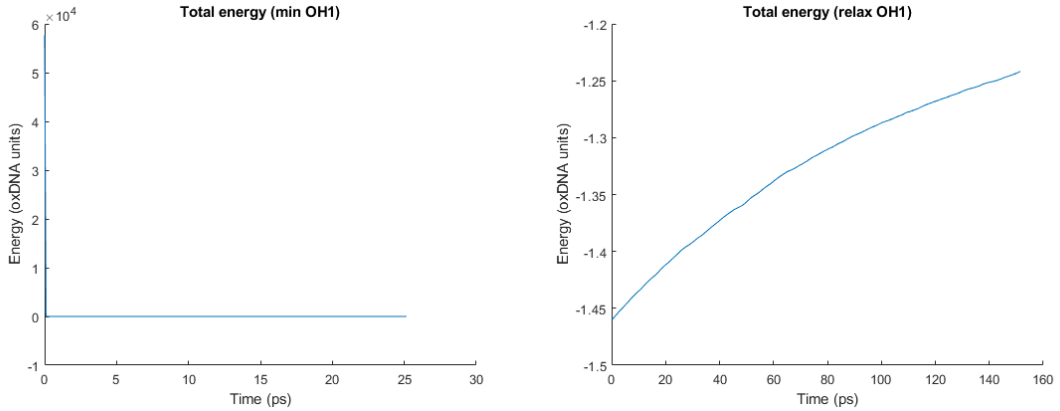
6.3 Simulations

Simulations were performed for both proposed designs for 12 ns using oxDNA. Minimizations and relaxations of the two structures can be seen in figure 6.26 for OH1 and in figure 6.27 for OF1. By mistake, minimization and relaxation was run at 300K for OH1 while at 310K for OF1, however if running OH1 at 310K, it yields very small differences when relaxing as seen in Appendix E, figure B2.

Looking at the total energies for both structures, a steep decrease is observed when the structure is minimized (figure 6.26a and 6.27a). This shows the importance of minimization before simulation, as large forces when the structure is falling into its first state can induce high stress and possible breakage of bonds. In both minimization and relaxation, a maximum backbone force is decided, above which the usual FENE potential is changed

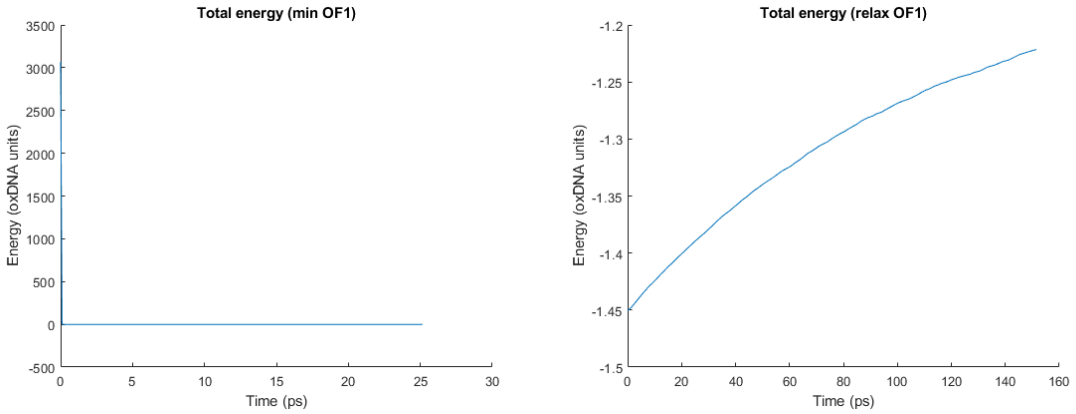
for a modified potential. This ensures that forces resulting from stretched bonds will not damage the origami, for instance by breaking base pairs. [50]

A relaxation of both structures, performed as a short MD simulation with the same maximum backbone force as for the minimization, was hereafter performed, see figure 6.26 and 6.27. The simulation is started from the energy minimized configuration and a nearly linear increase in total energy can be observed, however starting to increase less steeply in the end.



(a) (b)

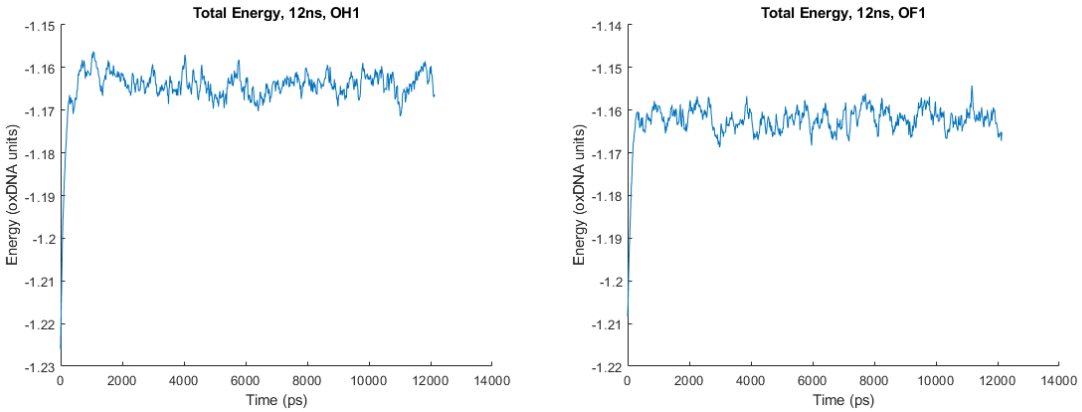
Figure 6.26: Minimization (min) and relaxation for OH1, showing the total energy over the simulation time.



(a) (b)

Figure 6.27: Minimization (min) and relaxation for OF1, showing the total energy over the simulation time.

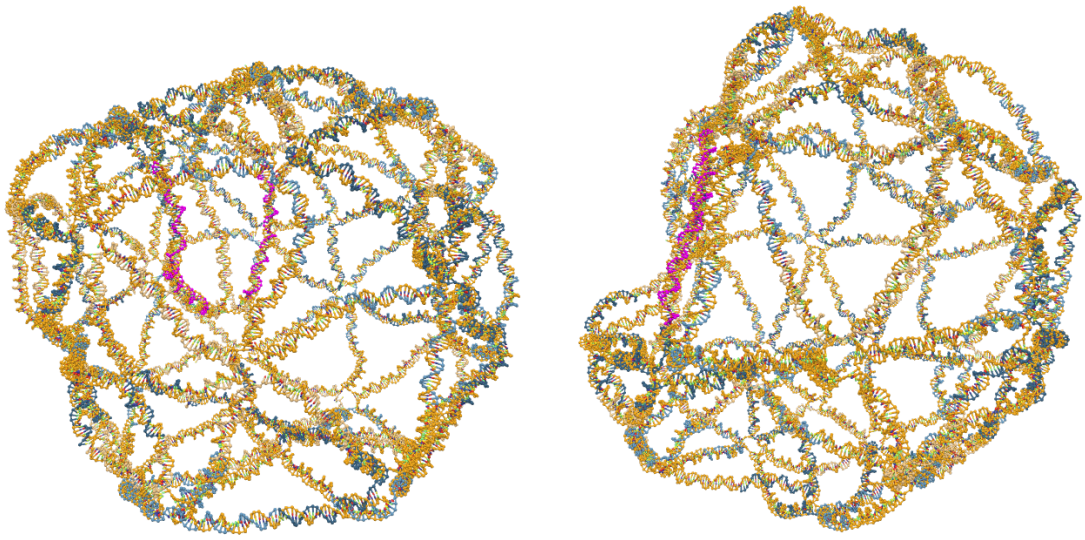
The total energy output after the simulation of the structures can be seen in figure 6.28.



(a) **(b)**

Figure 6.28: Total energy of the two origami designs. (a) Shows the OH1 structure with the double-stranded promoter regions packed inside, and (b) the OF1 structure where these regions are on the outside of the structure.

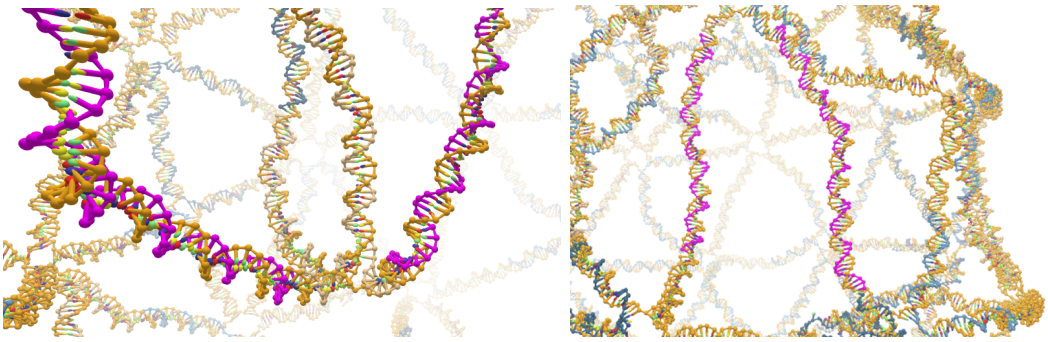
The final configurations of the two structures were altered after the simulation but still intact. Illustrations of both origamis in the last configuration can be seen in figures 6.29a and 6.29b below. The oligos ensuring double-stranded promoter regions as described in section 5.1.2 is marked in pink in both structures.



(a) **(b)**

Figure 6.29: Last conformations for (a) OH1 and (b) OF1, after 12 ns of simulation time. The oligos ensuring double-stranded promoter regions are marked in pink. The images are created using oxView.

A closer look on the doubles-stranded regions' configuration can be seen in figure 6.30 below.



(a) (b)

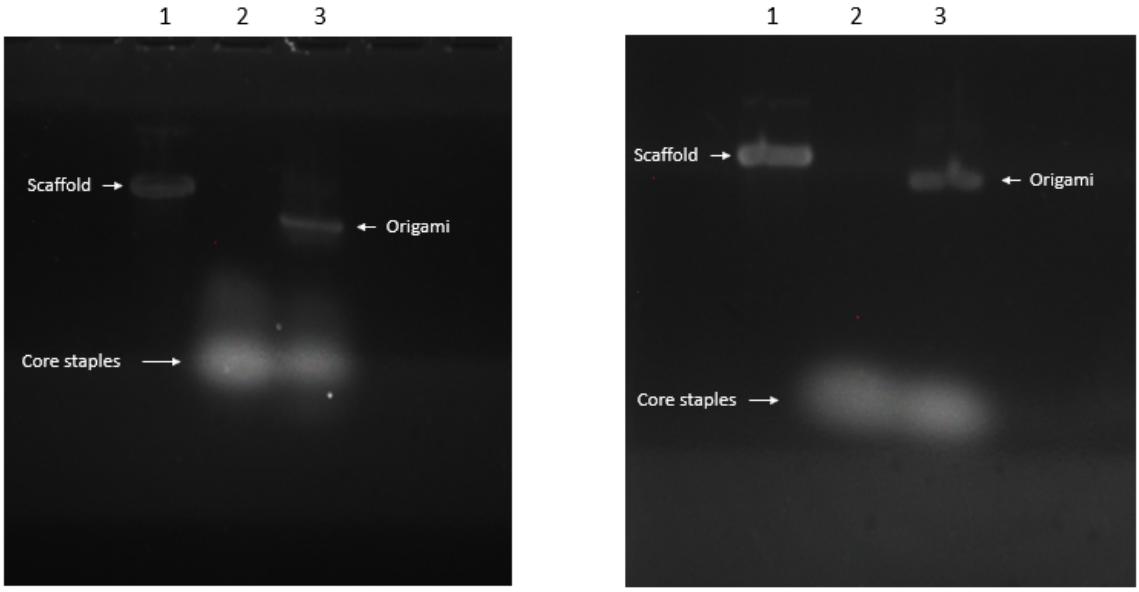
Figure 6.30: Close-up on the oligos for double-stranded promoter regions for (a) OH1 and (b) OF1. The images are created using oxView.

6.4 42hb Origami

The 42hb origami structure described in section 5.2 was used as a model origami to initiate the test of concept of using FRET to detect deformations in an origami structure. The results from these initial experiments will be relevant in further development of the original origami structure for use in transfection experiments by evaluating the folding/unfolding of origamis, both via basic studies of stability in cell culture media and directly in cells after transfection attempts using fluorescence microscopy.

6.4.1 Folding of 42hb Origami

To access whether the folding of the origami had occurred as expected, a gel electrophoresis was run as described in section 4.2.6. The results of the two origamis used in further experiements can be seen below in figure 6.31.



(a) **(b)**

Figure 6.31: Gel electrophoresis of folded 42hb origami. (a) Ori3 and (b) Ori4. In both, lane 1: p7560 scaffold, lane 2: Core staples, lane 3: 42hb origami. 2.5 μL of both origami solution, scaffold solution and core staple solution was loaded.

The origami in lane 3 in both images migrates faster than scaffold alone. When the DNA is packed, it is expected to affect the ability to run on a gel such that a difference between scaffold and origami is observed upon successful folding. The higher mobility is reasonable as closely packed DNA structures run better through the pores in the gel than both open circled and supercoiled DNA (open circular seen in lane 1 as the upper band, supercoiled seen as the lower and most intense band). In lane 3, excess staples after folding can be seen as a blurred band in the bottom of the gel in line with the band in lane 2 with only core staples. By observation from the gels, it seems like Ori4 has folded to a slightly better degree as no scaffold is left in lane 3, opposite to a shadow, probably from excess scaffold, in lane 3 in figure 6.31a.

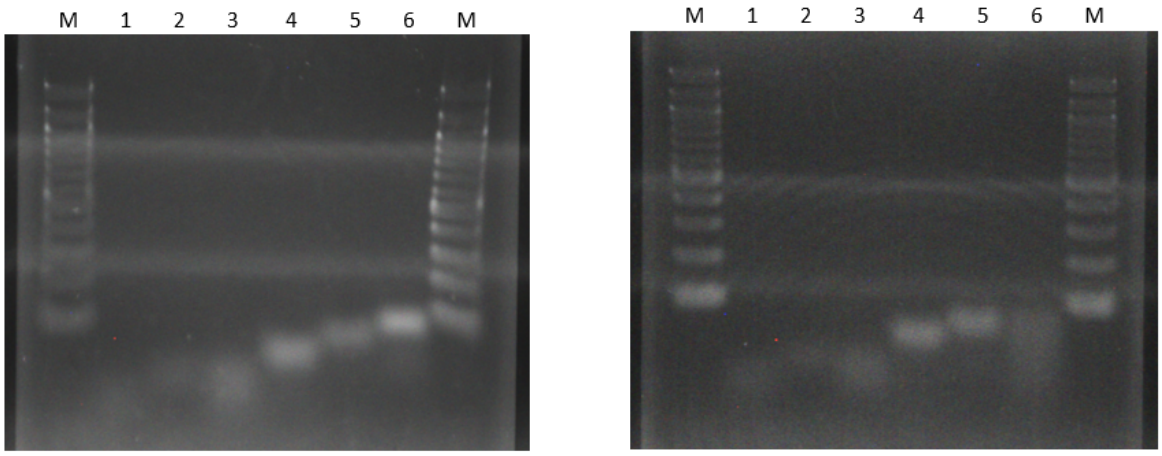
6.4.2 Annealing of Oligos for FRET Detection

As it is unsure if the distance between the fluorophores in the origami structure is too large for FRET detection, a reference study was made. Constructs as described in section 4.2.10 were designed utilizing the fluorophore staples and three separate oligos. The sequence, melting temperatures and tags are seen in table 4.6 in section 4.2.10. Short names are listed in the same table for convenience in the presentation of results.

The D-A-ST construct seen in figure 5.8 has a 4 bases between the fluorophores, giving an

approximate distance of 1.7 nm. This construct was expected to yield FRET if possible with the fluorophore pair. The D-A-M-LT construct seen in 5.9 was prepared with 17 bases between the fluorophore, 15 of the bases being paired with a short oligo to prevent bending of the strand. The distance here is similar to the one in the origami.

To obtain an idea whether the oligos were annealed together successfully, a 2 % agarose gel was run. Both single oligos with the donor fluorophore Cy3 attached (D) and single oligos with the acceptor fluorophore (A) were loaded, together with the annealed fragments, see figure 6.32.



(a) **(b)**

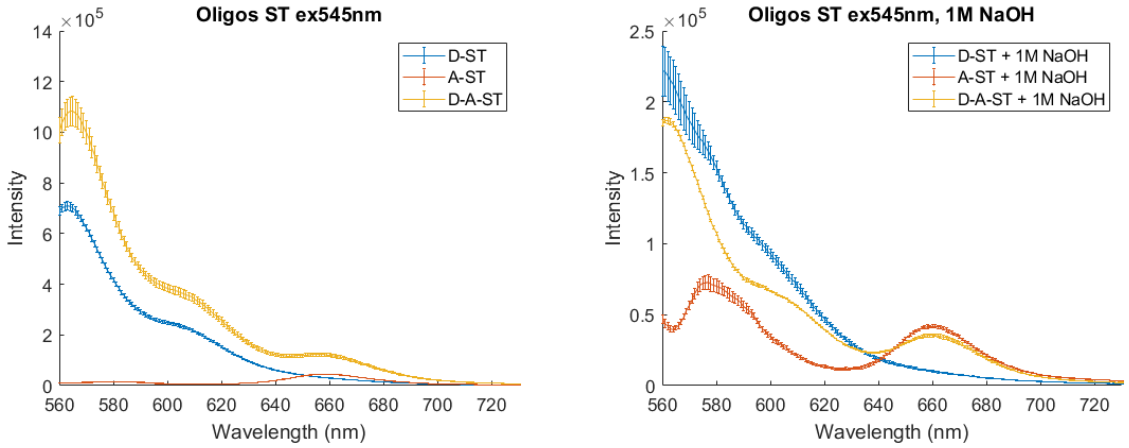
Figure 6.32: Gel electrophoresis of oligos solutions after annealing. (a) Lane 1: D, lane 2: A, lane 3: D-A, lane 4: D-ST, lane 5: A-ST, lane 6: D-A-ST. M denotes a 100 bp DNA ladder (NEB) (b) Lane 1: D, lane 2: A, lane 3: D-A, lane 4: D-LT, lane 5: A-LT, lane 6: D-A-M-LT. M denotes a 100 bp DNA ladder (NEB).

It can be noticed that in lanes with only a single oligo with no annealing partner is present, that is lane 1 and 2 in both images, the bands are visible but very blurred, so no specific placement and thereby size can be determined from it. In lane 3 in both images, D and A are loaded together but should not anneal together, and these bands are also blurred with an even larger band in length. However, in lane 4 and 5 in both images where annealing is expected, the bands are more clear and defined, however still blurred but to a smaller degree. In lane 6 in figure 6.32a where the whole configuration in figure 5.8 is supposed to be annealed together, the band is blurred but is even more clear than in any of the others, suggesting that at least some annealing has occurred. That is opposite to lane 6 in figure 6.32b, where the configuration in figure 5.9 is supposed to be created. Here, the band is blurred as for the single oligos and the length of the band is large, spanning the whole region of bands from lanes 1-5. This could suggest that the configuration is not

annealed properly.

6.4.3 Detection of Unfolding by FRET

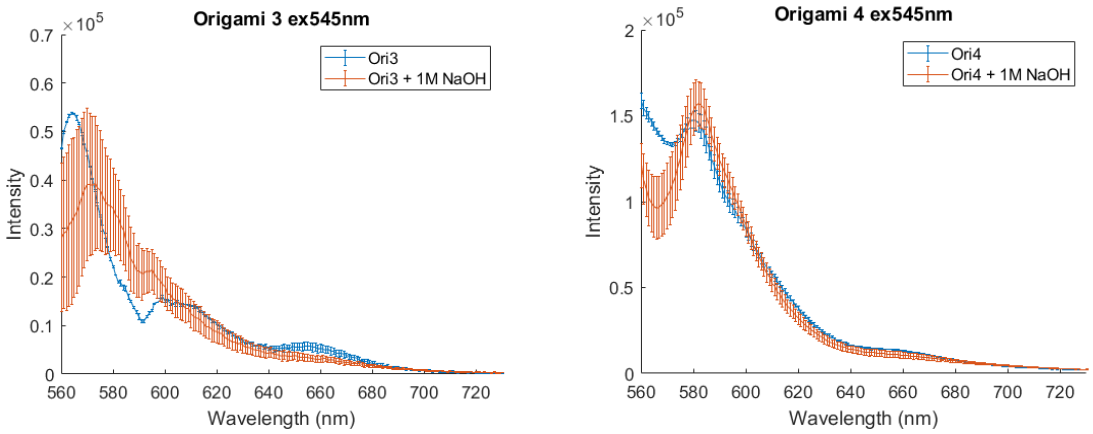
Fluorescence measurements were carried out on the short-distance oligo construct and origami samples to see if a FRET signal could be obtained. Different excitation wavelengths were tested, and 545 nm was found as the best to detect the peaks of both Cy3 and ATTO647N at the same time. The results of measurements of oligos can be seen in figure 6.33. The data obtained with annealed oligos was normalized to a concentration of 315 nM similar to the concentration after adding NaOH, assuming a linear relationship between counts and concentration.



(a) **(b)**

Figure 6.33: Fluorescence measurements of oligo constructs with an excitation wavelength of 545 nm. (a) Constructs D-ST, A-ST and D-A-ST measured while annealed, (b) constructs D-ST, A-ST and D-A-ST after denaturation with 1M NaOH. Notice the difference in scale bars.

Below in figure 6.34 measurements with the same conditions can be seen for both origamis.

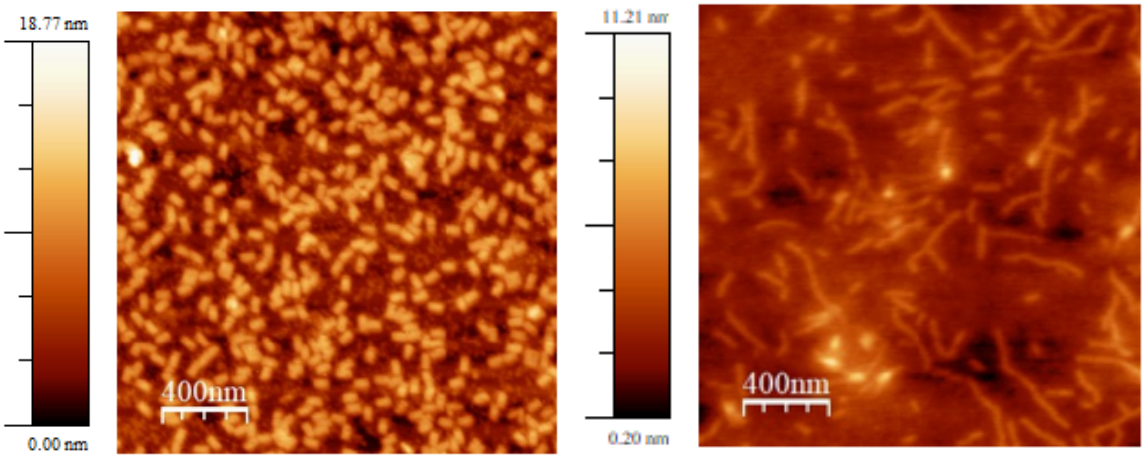


(a) **(b)**

Figure 6.34: Fluorescence measurements of Ori3 and Ori4 before and after treatment with 1M NaOH. (a) Shows Ori3, while (b) shows Ori4.

6.4.4 Confirmation of Folded and Unfolded States by AFM

To confirm folded or unfolded states of origami structures, AFM was utilized to image them. In figure 6.35 supposed folded states of Ori3 and Ori4 can be seen, before treated with sodium hydroxide. For Ori3 in 6.35a, uniform rectangular structures can be observed, indicating well folded origamis. For Ori4 in figure 6.35b, the same well defined structures cannot be observed and instead longer and thinner fibrillar structures are seen which indicate improper folding.



(a) **(b)**

Figure 6.35: AFM images representing (a) Ori3, and (b) Ori4 after purification with dialysis. The scan size is 2x2 μm.

In figure 6.36, the Ori3 solution after treatment with sodium hydroxide and a dialysis pro-

cedure to remove any excess sodium hydroxide can be seen. The uniform shaped objects on the surface indicates a worn AFM tip.

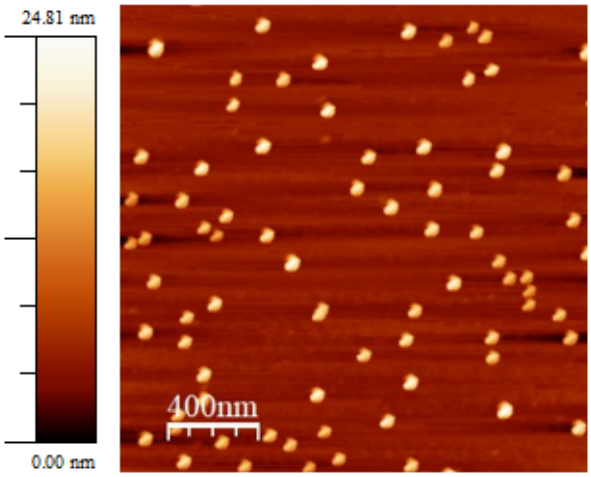
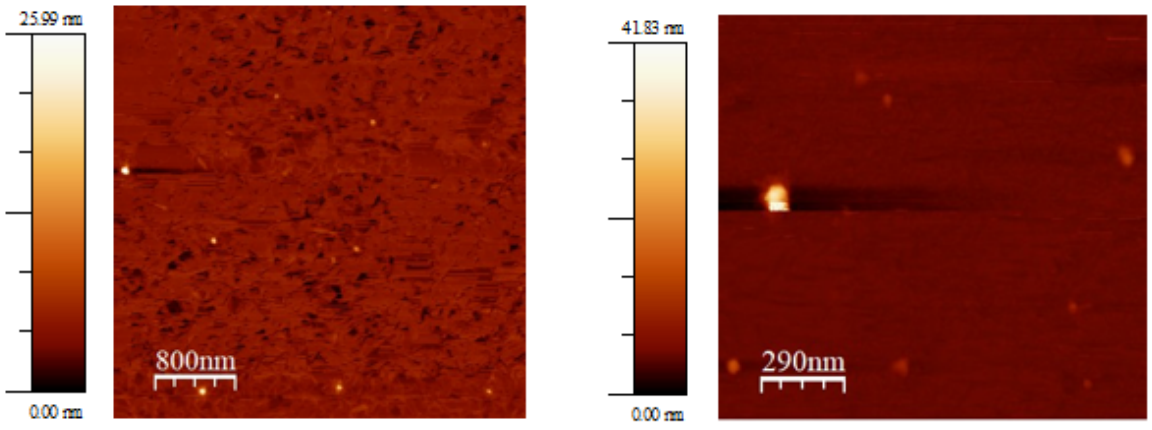


Figure 6.36: Deposition of Ori3 after treatment with 1M NaOH and following dialysis. The scan size is 2x2 μm.

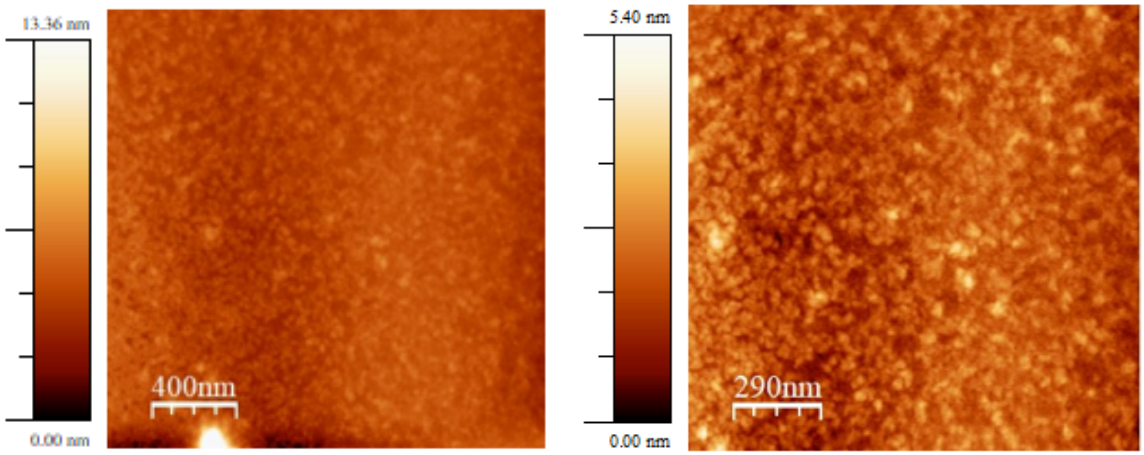
In figure 6.37 a sharp AFM tip was used to measure the sample from figure 6.36 again. Topologies of structures on the surface can be seen in figure B2, B2 and B2.



(a) **(b)**

Figure 6.37: The same sample as in figure 6.36 imaged with a new AFM tip.

Below in figure 6.38, origamis denatured by heating up and cooling quickly on ice can be seen instead.



(a) **(b)**

Figure 6.38: Both (a) and (b) shows Ori3 after heating to 65 °C for 5 min. and cooling on ice directly afterwards.

7. Discussion

7.1 Plasmid pCas9-GFP

7.1.1 Sequencing

For plasmid extract (p1), it was decided to sequence it to confirm the specific sequence of the target region for origami scaffold creation. This was considered to be important as the plasmid should be used for origami folding, and variations in sequence could influence the folding of the final origami. The sequencing result revealed variations from the sequence obtained from Addgene, with most changes present in the chimeric intron region and a single change in the chicken β-actin promoter. While simple to change the sequence used as scaffold in an origami design, it is questionable if the variations would have an effect on the expression of the gene products when transfected into cells. The 33 bp insert followed by a 6 bp deletion could especially be suspected to have an influence as it is large and therefore increases the chances of mutations at important sites. This is, however, difficult to answer as regulatory element systems in gene expression are often used in different variations, for instance as described for the WPRE element in section 3.3.3. While difficult to find indications on this issue regarding this specific promoter and intron construct, a study by *Callis et al.* (1987) showed that expression of a chimeric chloramphenicol acetyltransferase (CAT) gene could be increased by 100-fold by the addition of the *Adh1* intron 1 when located between the promoter and CAT coding region. The chimeric intron is in this plasmid located after the promoter as well. However, their study also revealed less importance of the last two introns of *Adh1*. [51] In a study by *Wakabayashi-Ito et al.* (1994) on the EF-1α promoter, both the 5' flanking region and the first intron were important for the promoter activity. However, the study also revealed that not all mutations that they carried out had a significant effect on the expression. [27] All sequencing results were obtained more than one time in the sequencing reactions and is therefore very certain. A T at 640, seen in figure 6.4 in section 6.1.2, has a low intensity though, and a small signal from A as well, which is not so certain. However because of the many reads, it is plausible. It should as well be noted that gene expression relies on many other factors, such as the cell type in question, and it is therefore difficult to determine the exact effect of these variations found by sequencing.

7.1.2 Amplification

To prepare the scaffold for DNA origami, amplification of a 7,797 bp part of the plasmid was attempted as illustrated in figure 6.5. As it is a large DNA fragment, the A. G. LA polymerase was chosen in the beginning as it is ideal for long templates. The first PCR reactions carried out with standard conditions and different annealing temperatures (figure 6.6) all produced products of which the most intense was a fragment around the expected 7.8 kb. It is difficult to assess whether it is actually the correct product because a large amount of DNA is present in the samples, making the bands broad and difficult to distinguish from each other.

The results from the first PCRs could not be reproduced, neither by addition of Q5 High GC enhancer, change of primer or by touch-down PCR. Interestingly, preheating of the thermocycler before placing the samples into it, which is suggested to reduce nonspecific amplification [52], yielded a large number of PCR products. Among these products, a fragment that could be the target product was observed. It is difficult to deduce why each of the products appears, but it could be hypothesized that the preheating had improved the melting of the template. This could have opened up the GC-rich region and made amplification possible at different sites, or at the same sites with varying elongation length.

The immediate reason for the repetitive products appearing at 5 kb would be that the primers are binding more specifically at other sites than intended. This cannot be disproved, but another suspicion was the GC-rich region, which, if not fully denatured, could stop the polymerase in the middle of the elongation, yielding products only from the non-GC-rich region. Looking closer at the chicken β-actin promoter and the chimeric intron, they have a GC-content of 68 and 74 %, respectively. Several studies, among them one by *Seifi et al.* (2012), suggest that the addition of betaine and DMSO can enhance amplification of such templates. In that study, amplification of a putative promoter region of mouse peroxisomal protein with a GC-content of 71.01 % was achieved by the addition of betaine and DMSO to the PCR reaction in a PCR buffer with ammonium sulfate. [53] Betaine reduces the melting temperature of DNA and decreases the dependency on base pairs, while DMSO disrupts base pairing. [54] Both substances should therefore be valuable, if the problem is a non-denatured template; if it does not denature fully from the beginning or in the case where it re-anneals again before primers anneal during lowering of the temperature to the annealing temperature.

Mixed concentrations of 1-2 M betaine and 1.5 to 5 % DMSO were therefore tested with both primer set 3 and 4, as seen in figures 6.11 and 6.12. These attempts did not yield the

expected target product. Again, a product between 5 and 6 kb was seen for both tests, while also a larger product around 9 kb, and smaller products around 3-4 kb, were seen in the one with primer pair 3. However, a mistake was made during these reactions by setting the extension time to 3 minutes instead of 8 minutes, thus raising the question if such short time is enough to extend the 7,453 bp fragment. The recommended extension time given by the supplier is 1 minute for fragments up to 2 kb; for longer fragments, it should be increased by 1 min/kb. [55] Even though a 15 min final extension was performed, it might have been too short time to complete extension for the large fragment.

Betaine and DMSO additives were tested as well using the A. G. LA polymerase which had so far showed the most promising results in the first PCR run. As seen in figure 6.13, all lanes have a broad band, later confirmed by loading less sample for at least lane 1 and 2, revealing that they consisted of two products, where one of the products could be the correct one. To test if this could be optimized to obtain only the target product, substitution of 3/4 of all dGTP in the dNTP solution with a 7-deaza-dGPT was tested, as suggested by *Musso et al.* (2006) who amplified GC-rich DNA sequences with GC-content from 67 % to 79 %. [54] The 7-deaza-dGPT reduces the number of hydrogen bonds between guanine and cytosine, and prevents formation of intramolecular G4 quadruplexes, purine-motif triplexes and G*G Hoogsteen base pairing [54], and might possibly make the double-stranded template anneal weaker. This would make the denaturation happen easier and thereby secure better primer binding. [54] However, as seen in figure 6.14, it did not yield a single target product as hoped for. The reason for this could be that the template still needs to denature fully in the first place as the effect of 7-deaza-dGTP will not be seen before the first round of amplification, where after it should lead to an easier denaturation of the templates.

It can be discussed if excision and purification of the possible target product from the gel should have been attempted. Digestion with a suitable restriction enzyme could then be performed to validate the product, or sequencing should be carried out to confirm it more precisely, as possible lacking ends of the product would not be visible by digestion.

Looking over the many amplification attempts performed, many non-specific products had been created. Especially, products around 5 kb, between 5 and 6 kb or at 6 kb recur often, and the sequence was investigated more in detail. Interestingly, the distance between the annealing site of the reverse primer 'R.pcas9gfp.5biot' and the 3' end of the chimeric intron is 5,686 bp (for the reverse primer 'rev. real. creates complementary', 5,620 bp). This 5.6 kb region has a total GC-content of 50 %, while it for the promoter and chimeric intron altogether is 73 %. If the denaturation in the latter region is a problem and the

template sticks together or re-anneals in these regions, it will naturally create products lacking the GC-region as the full amplification will be inhibited. This fits well with the fact that before additives like betaine and DMSO were used, primarily products around 5 kb occurs. When using additives and Phusion polymerase suited for high GC content afterwards, both smaller and larger products occur, probably due to the more opened up GC-region. However, if denaturation is still not complete, target product formation will still be inhibited and non-specific amplification will be dominant.

In the light of these observations it was decided to investigate the GC-rich region more deeply, and a primer set (5) was ordered to test amplification of just the promoter- and chimeric intron region, a total of 1,649 bp.

GC-region

The GC-rich region of the promoter and the chimeric intron region was tested to clarify if this was in fact the problem, or if it was because of the long template to amplify. As the primers for this GC-region only spanned 1,649 bp, this should at least be simple in terms of length. Both the addition of a preCESII solution suggested by *Ralser et al.* (2006), the use of a Taq polymerase buffer with $(NH_4)_2SO_4$ as suggested by *Seifi et al.* (2012) and touchdown PCR were tested together with Taq polymerase, and also Phusion polymerase with 1xCES or GC buffer and DMSO. Only products at around 500 bp, 1.4 kb, 8 kb and one above the ladder were appearing, shown in figure 6.17, indicating that the GC-region was in fact an issue.

The reason for the product around 1.4 kb was investigated a bit further, as it only lacked around 200 bp to be complete. It was speculated if possible to determine any secondary structures in one of the ends of the GC target sequence, for example hairpin structures. Looking at the sequence, it was in fact possible to find a potential placement for a hairpin, that could prevent the amplification of the last 205 bp, as seen in figure 6.19 in section 6.1.4. If such hairpin was created in the single stranded template when denaturing and reannealing again, it could easily stop the polymerase in the middle of the amplification process, promoting amplification of only shorter products without the hairpin sequence. [54] To support this argument, the suspected hairpin sequence was investigated using the RNA WebServer from Institute for Theoretical Chemistry, University of Vienna. This prediction seen in figure 6.20 could indicate that the possibility of a hairpin creation is present and could have yielded the 1.4 kb product instead of the expected 1.6 kb.

A last attempt creating the desired fragment was done by adjusting the extension temperature to 80 °C. This is not optimal for amplifications, as the activity of the polymerase

is compromised. The half life of Taq polymerase is around 1.6 h at 95 °C. [56] However, in combination with touchdown PCR, the 1.6 kb GC-region was amplified with Taq polymerase, seen in lane 3 in figure 6.18. This was most probably enough to prevent the hairpin formation, though not transferable to use in amplification of the large origami scaffold due to the compromised integrity of DNA polymerase at high temperatures and long amplification times.

After investigation into the GC-region, it could be confirmed that it was indeed an issue; not only because of the high GC-content but also possibly due to the hairpin structure. Further investigation could be done to try to optimize the PCR but the chances for a successful amplification were deemed low. It could be questioned if the possible target product obtained in some reactions together with other products could be purified and used in aPCR, however the low yield was thought to be a limiting factor. Another strategy considered was edition of the plasmid by cutting out the chimeric intron, however this would probably not solve the issue as the hairpin stretches over the 3' end of the promoter and the 5' end of the chimeric intron. The plasmid was therefore changed to the similar lentiCas9-EGFP.

7.2 Plasmid lentiCas9-EGFP

7.2.1 Amplification

PCR amplification of lentiCas9-EGFP was first tested with standard Taq polymerase conditions. The expected product size of 6 kb was obtained in first attempt in 2 out of 3 samples. The reaction containing the template from the plasmid extract stock solution (pl1) showed only a smear on the gel, indicating either an error in the preparation of the PCR reaction or that the template in this solution had been degraded. This could be due to the presence of nucleases in that solution. The latter is most probably the case, as restriction enzyme digest with this solution also only resulted in a smear (not shown). This stock solution was also used for the first digest that worked fine, but this was carried out immediately after the extraction, while the second digest and PCR were carried out approximately two months after. The dilutions prepared from the same extract two months prior remained fine though, indicating possible contamination in the stock solution (pl1) and hence degradation over time. [57]

7.2.2 Asymmetric PCR

For asymmetric PCR, two strategies can be used; either the aPCR is carried out directly on the plasmid using a molar excess of one of the primers, for example a 50 molar excess [39], or the double-stranded product from the standard PCR is extracted from a gel and used as template in an aPCR with only one primer.The second strategy was selected here, but a very low concentration of the PCR product was obtained after extraction, approximately 10 ng/μL measured with Picodrop. In the aPCR on this fragment, none of the reactions yielded a product. As only one primer is used, the increase in amplified product is only linear instead of exponential as in a PCR with two primers. This requires both a decent amount of starting template and equivalent matching or excessive concentration of the single primer to obtain the maximum rate of amplification. Probably the template and primer concentrations have been too low to provide a reasonable starting point for the reaction, or the primer has not been added excessively, however the maximum added primer concentration was corresponding to the total concentration of both primers in a normal PCR. Another explanation could be the number of cycles, which could have been increased for the aPCR to increase the yield.

7.3 General Considerations about PCR Amplification

It would have been interesting and more deductive to perform a more general investigation, changing one parameter at a time, especially with the A. G. LA polymerase that created the most promising results for the amplification of plasmid pCas9-GFP. However, with limited amounts of available polymerase, more parameters were often changed at a time. Also, considering the long extension times, it was a time consuming process. However, the varying parameters did also make it difficult to observe tendencies in the results and thereby decide about the next step. Further optimization could imply the use of both negative and positive controls. Negative controls were not always carried out and those are important to show any DNA contamination, especially valuable when having a lot of non-specific amplification as observed in many reactions. Positive controls would have been useful in the beginning, clarifying if every substance except primers and template worked as supposed in the PCR reactions.

Another strategy that could be considered, both for the normal PCR and for aPCR, is nested PCR where a fragment which is a little larger than desired is amplified first, and then inner primers are utilized to amplify that product, such that non-specific amplification can be easier controlled. [58] It is unsure whether it would optimize the normal PCR of pCas9-GFP as it would still require a full amplification of the large fragment, which

until now has proven to be difficult. Maybe the outer primers could be designed with a high T_m and the PCR run at correspondingly high T_a, thereby avoiding non-specific amplification. For aPCR, the single primer used could be designed as an inner primer, possibly ensuring a more specific binding.

7.4 Design of DNA Origami for Transfection

The designs of the DNA origamis were chosen out of a wide variety of possibilities. The open mesh structures were chosen because of the high promise for stability at physiological conditions [14, 16], which will be evaluated in the next section by discussion of the simulations. The designs were chosen to differ in the position of the double-stranded promoter regions to evaluate the influence on the expression potential and the stability. While difficult to assess what is important, when the uptake mechanism and transcription of single-stranded DNA fate is unknown, it was hypothesized that packing of these double-stranded regions could induce protection against nuclease digestion and environmental factors. A study by *Lim et al.* (2015) showed that packaging of plasmid DNA by the self-assembly of Zn^{2+}/DNA nanoclusters increased luciferase expression by around 1000-fold at a Zn^{2+} concentration of $\geq$ 549x10^{-3} M, compared to naked plasmid DNA. This indicates that structural DNA can increase both stability and internalization into cells. [8] It has also been shown by *Agarwal et al.* (2017) that another protection strategy by coating DNA origami with cationic poly(ethyleneglycol)-polylysine block copolymer induced resistance to nuclease digestion by DNAse I and FBS. [59] It should be mentioned though, that the DNA origami scaffold in this project consists of gene elements for expression of the candidate molecules, Cas9 and EGFP. Therefore, the outer part of the origami is also important, and protection strategies for the whole DNA origami could be of interest as well, for example by using the previous mentioned strategy. Regarding stability, a line can be drawn to the FRET studies in this project. If introducing one or more fluorophore pairs with a short enough distance between the donor and acceptor fluorophore into the origami structure, measurement of the DNA origami integrity during cell transfection could be performed. FRET has been used by *Heini Ijäs et al.* (2019) to investigate the reversible opening and closing of a dynamic DNA origami nanocapsule, controlled by pH, as a drug delivery vehicle. [60] This demonstrates the sensitivity, applicable for stability studies as well.

7.4.1 Simulations

The simulations performed on the two structures, OH1 and OF1, using oxDNA, showed an immediate integrity of both structures after simulating for 12 ns. The simulations were performed with inclusion of an electrostatic term serving as a method to mimic different salt concentration levels. The simulations were performed with a NaCl concentration set to 0.2 M, close to physiological conditions [43] which will be the environment for the origamis when used in transfection studies.

As the curve is not flattening for any of the relaxations, a complete relaxation with stable energy in the end was not obtained for any of the structures, and it should probably have run for a longer time. However, the final simulations without any maximum backbone force were run from the final configuration after these relaxations, and this did not result in any broken bonds. A steep increase in energy is seen for both structures in figure 6.28a and 6.28b, however the increase is largest for the OH1 structure, indicating that it had a longer way to reach its final structure after relaxation. Within a short time, they both obtain energy fluctuations in a constant small range, approximately after 2 ns for OH1 and 1 ns for OF1, indicating that the final structure is obtained quite fast and that they appear stable. It can be argued that longer simulation should be performed, as DNA helices display movement in both the ns, μs and ms range. [61] The simulations are, however, a good indicator of stability of the obtained structures, which as seen in figure 6.29 are deformed but retain the overall shape. Furthermore, the oligos are in OH1 still packed inside and for both structures intact as seen in figure 6.30. To obtain more quantitative data on the stability, RMSF values could be calculated to observe fluctuations in positions of strands and thereby get an indication whether some parts of the structure are more mobile than others and in risk of breakage. It could be interesting to perform "end-to-end" distance analyses of the different regions of the origamis, showing the movement of one region related to another in the structures, which could give an indication of stability as well.

Furthermore, it is worth to mention that oxDNA is a coarse-grained simulation program, hence limitations. All atom-simulations can be performed as well, in contrast to the nucleotide level of oxDNA, however for large systems as DNA origami it is often infeasible. Therefore, coarse-grained simulations are a good alternative. [62]

7.5 Detection of Folding and Unfolding of Origami Structures

7.5.1 Folding of Origami

Both Ori3 and Ori4 seemed to have folded based on the altered movement on the gel compared to the p7560 scaffold. However, looking at the AFM images in figure 6.35 it is observed that Ori3 is folded into the correct square structures, while Ori4 is only folded to some extend and possibly aggregated into more fibrillate structures. If considering the gel images, a small variation in the mobility of the origamis can be seen, though it is hard to determine, thereby questioning the gel validation of folding of this origami. The origamis were folded in the same way, and the only difference for Ori4 were a double reaction volume and a final concentration of 2.88 μM of the two fluorophore staples instead of 216 nM for the ATTO647N staple and 180 nM for the Cy3 staple. This was done to ensure that enough was present to bind on each origami, ensuring maximum fluorescence signals as low intensities had been observed for Ori3. This imbalance in staple concentrations would be the only immediate explanation for the changed folding pattern.

7.5.2 Annealing of Oligos

Two oligo constructs for FRET detection were designed; one with a distance of only 4 bases ($\sim$ 1.7 nm) between the fluorophores which was expected to yield FRET signal, and one construct with approximately the same distance between as in the 42hb origami, however it is difficult to assign precisely. In figure 6.32, the different strands are seen after the annealing procedure. It can be observed that single-stranded strands produce weak, diffuse bands in the gel, while strands with double-stranded segments produce a clearer band, however still diffuse. It is believed that the single-stranded DNA oligos will result in the smeared bands, as GelRed is an intecalating dye [63] and cannot intercalate in the same way as with double-stranded DNA. It can also be noticed that the annealed strands seem too large on the gel, as D-A-ST would constitute a total length of 72 bases (44 paired, 10 and 18 unpaired in each end) and D-A-M-LT would constitute a total length of 83 (53 paired, 10 and 18 unpaired in each end and 2 on each side of M). However, the DNA ladder only goes down to 100 bp, and a ladder with smaller fragments is needed to clarify this, although it is suspected to be poor gel resolution.

Deduced from the clearer gel bands from samples with expectedly annealed oligos, it was concluded, that D-ST, A-ST and D-A-ST all seemed to be annealed successfully. For the long construct, D-LT and A-LT seem annealed well, but D-A-M-LT result in a broad

smeared band, indicating an unsuccessful annealing. Yet it does instead support interpretation of the successful annealing of the short construct. The middle strand can be suspected to be the cause of the failed annealing, as the melting temperature is much lower for this strand than for the others in the construct and it will therefore anneal last.

7.5.3 FRET Measurements of Annealed Oligos and Origami

FRET was measured on the short oligo construct. Looking at figure 6.33a, the fluorescence peak for A-ST is at a wavelength of 658 nm, as expected for ATTO647N. Measuring the intensity for D-A-ST at this point, the intensity is increased 2.7 times when the Cy3 donor fluorophore is added to the strand, indicating that FRET is occurring. Looking at the D-ST construct it is seen that there is as well some background fluorescence at this point. Whether this stems from the ST strand or the D strand is unknown and it would therefore have been valuable to measure the D strand alone. If that is constituting to the background at 658 nm, the actual FRET signal is lower. However, from these results it is seen that FRET could be observed, although with a low signal. It was expected that that FRET would have be more intense, as the donor pair is often used. For example, *Santangelo et al.* (2004) used a similar approach to visualize real-time expression levels in living cells with Cy3 as the donor and Cy5 as the acceptor fluorophore. For fluorescence measurements with the same excitation wavelength, they obtained a clear FRET signal by an increased Cy5 intensity and correspondingly low Cy3 signal as the energy transfer was large. [64] ATTO647N and Cy5 are spectrally similar dyes, and with the same distance between the dyes in this project, a similar clear result was expected. Also, the quantum yield of Cy5 is reported to be 0.27 in PBS, while ATTO647N is reported to have one at 0.65 in PBS, so ATTO647N should actually have a higher fluorescence intensity. [65, 66] The reason for the low intensity in these results could probably be that it has degraded over time, since the fluorophore staple dates back to 2013.

Denaturation of the oligos was carried out with 1M NaOH and the fluorescence measurements hereafter are seen in figure 6.33b. The intensity at 658 nm for D-A-ST has decreased to below the intensity of A-ST, from the first view an indication of lost FRET while unfolding. However, this cannot be concluded as the intensity of the Cy3 has as well fallen drastically, as visualized in figure B2a. While the intensity of the ATTO647N fluorophore at 658 nm is now 92 % of the intensity before NaOH treatment, the intensity of Cy3 at its max, 563 nm, is only 30 % of the intensity before NaOH treatment. If only denaturing the strands, it would be expected that the Cy3 kept its fluorescence in a range similar to before, even though fluorophores can very well be affected by the surrounding environment and the orientation of them on the nucleic acid strands. [67] It

is however more reasonable that the high alkaline environment has influenced the dyes, apparently most for Cy3, as the fluorescence capability can be influenced by different environmental factors. [68] When Cy3 is not emitting light with an intensity comparable to before denaturation, the energy transfer cannot happen at the same level even if the fluorophores were close, and therefore no conclusion can be drawn about the denaturation on the background of FRET measurements.

For future experiments it would be more convenient to denature the oligos in a different way that cannot affect the fluorophore integrity. This could be done by thermal denaturation by heating the oligos above the highest melting temperature of the oligos and fast cooling on ice afterwards to prevent re-annealing. It should be ensured, potentially on an agarose gel, that re-annealing does not occur to ensure a correct basis for FRET measurements. To exclude any background signal, an extension of the ATTO647N strand with a quencher on the opposite end of the fluorophore could be considered to form a beacon that results in quenching of the signal when not bound to another strand. However, this would complicate the comparison to the DNA origami unfolding, as the staple with fluorophore will not necessarily dissociate from the scaffold and therefore not be able to form the beacon. As the oligo assay is made as a reference study for the origami, it will probably be best to stick with the current constructs.

As the FRET measurements on annealed oligos at denaturing conditions are deemed inconclusive in showing a folded/unfolded state, so is the measurements on origamis. Yet, it could look like there is a smaller FRET signal for Ori4 than for Ori3 at 658 nm in line with the less folded state of Ori4 confirmed from AFM images, though this is impossible to conclude definitively. The intensities are not normalized for the concentration of origamis, the counts are very low for a fluorescence measurements, and the standard errors in the measurements are very large. A conclusion would require these things sorted out. Therfore, the origami experiment should be repeated, preferably with a new ATTO647N staple to yield a higher intensity.

Experimental limitations affecting these results are as well the cuvette placement and concentration determination of the oligos. The cuvette used did not fit into the fluorescence spectrometer, and therefore a 3D printed cuvette holder was utilized. This holder made it possible to place the cuvette in the sample chamber, but it could not be placed with such accuracy that the exact same position was ensured when changing sample. This could possibly have influenced the measurements and should therefore be taken into consideration.

7.5.4 AFM

AFM imaging was carried out to confirm folded/unfolded states of origami. As already commented in section 7.5.1, the folded state of Ori3 could be confirmed, while Ori4 did not fold into the desired squared origami structures. Therefore, further AFM imaging after denaturing was only conducted on Ori3. For Ori3 denatured with 1M NaOH and washed after deposition on mica plate, a large clump could be seen on the image with height of around 25-30 nm (not shown). Therefore, the origami was dialysed to remove any excess sodium hydroxide. Small dots around 20 nm in height can be seen after dialysis in figure 6.36. These structures are too large to be origamis, but a clear similarity in shape can be observed, which could indicate a damaged tip, producing the same shapes when drawn over a deformation in the surface. It is not possible to conclude, whether it is folded/unfolded origami structures or another substance on the surface. To clarify the origami state after NaOH treatment and dialysis, another cantilever was used to measure the same sample as in figure 6.36, which is seen in figure 6.37. Small structures can be observed; if it is origami is hard to tell but looking at profiles of the structures it reveals that all has a height between 5 and 10 nm. Yet, the process of dialysing the origami solution will dilute the sample, and therefore only a few structures are seen. It is difficult to obtain a clear conclusion with such low number of structures, but it is indeed an indication of the unfolding. This could indicate that it is unfolded origami structures. The surface is not uniform, indicating damage of the mica plate, probably due to the sodium hydroxide. Also, if taking a closer look, the lines pointing in different directions could be hypothesized to be DNA strands.

A small amount that was left from the Ori3 solution without NaOH was denatured by heating to 65 °C for 5 min and quickly moved onto ice to prevent refolding. This sample was assessed using AFM as well, for a reference on how an origami would look in the denatured state without any possibility for denaturation agent to be seen on the surface. It was not used for fluorescence measurements, as a very little amount was left, and if diluted to a volume suitable for the cuvette size, the fluorescence count would be too low to consider acceptable. As the origamis were unfolded thermally without adding any chemical, the structures seen in figure 6.38a and 6.38b must be denatured origamis. Yet, the worn cantilever was used, and therefore the image is affected by a non-sharp cantilever. However, if the origamis were still in the folded conformation, they would look like small uniform shapes but with no characteristic square structure seen. That is however not the case, and instead an unstructured pattern is seen, indicating unfolding of the origami. The structures are also small, around 5 nm in height when largest, fitting well with unfolded or partly unfolded origamis.

8. Conclusion

The aim of this project was to investigate DNA origami as method to package plasmid DNA for subsequent transfection and expression in mammalian cells, using the HEK293T-mCherry reporter cell line. While a part of the process to obtain this goal has been studied, unseen challenges in creation of the single-stranded scaffold for origami folding occurred.

Investigation of the reasons for a difficult amplification of plasmid pCas9_GFP suggested that both the overall GC-rich chicken β-actin promoter and chimeric intron in the plasmid, together with the possible hairpin formation between the 3' end of the promoter and the 5' end of the chimeric intron could be the issue. It can be concluded that the use of diverse PCR additives such as betaine, DMSO and 7-deaza-dGTP and a polymerase suited for long amplicons all have an effect on the amplification, however, this did not result in the creation of a scaffold for DNA origami. This plasmid is therefore considered insufficient for this purpose.

In contrast, amplification of the target region in the other plasmid, lentiCas9-EGFP, was obtained successfully in first attempt by using Taq polymerase at standard conditions. Successful asymmetric PCR to create a ssDNA scaffold is still to be obtained.

Preliminary studies of FRET as a method to detect folding and unfolding of DNA origami were carried out on annealed oligo constructs and the 42hb model origami. The results for origamis were affected by a low fluorescence intensity of the fluorophore ATTO647N, however a FRET signal was obtained from the construct of oligos annealed with the shortest distance. Whether denaturation of both oligos and DNA origami yielded a decreased FRET signal due to the separation of fluorophores is questioned, as the integrity of the Cy3 fluorophore was affected by the highly alkaline environment induced by the denaturation with sodium hydroxide.

Lastly, two polyhedral mesh DNA origami constructs were computationally designed using a region of plentiCas9-EGFP as scaffold sequence. Possible important sequences of the promoter for expression were suggested and included in the DNA origami design. The two structures were simulated at physiological conditions and are from these observed to be stable after 12 ns.

9. Further Experimentation

First of all, the single-stranded scaffold DNA should be created. The target product from plentiCas9-EGFP has been obtained easily already as double-stranded by PCR and aPCR conditions should be optimized. That will include further test of conditions for the method based on the double-stranded product as template in amplification with only the reverse primer, by e.g. increasing the primer concentration to ensure excess and increasing the number of PCR cycles to compensate for the linear amplification. Other strategies could the lambda exonuclease treatment as mentioned in section 5.1.1 with the use of modified primers. Validation of the ssDNA creation can in both cases be validated by treatment with exonuclease I. [38] Second step would be folding of the DNA origami and optimization hereof, by using both gelelectrophoresis and AFM in the structure validation.

For cellular uptake studies, confocal fluorescence microscopy could be utilized to quantitatively assess the particle internalization. DNA origami should be added and incubated with the cells, and the nucleus, which is the target region, could be stained with either Hoechst 33324 or DAPI. The DNA origamis could be prepared with fluorophores on selected staple strand, and an overlay of measurements for the two fluorescent channels could be created to visualize the internalization. Further more, a Lysotracker could be utilized to stain lysosomes to assess whether the DNA origamis ended up in the lysosomes. This could also give an indication into the uptake pathway(s). [19, 22] For quantitatively measuring of particle uptake, flow cytometry could be utilized. DNaseI treatment should be performed to digest any non-internalized DNA origamis that would otherwise compromise the results.

The uptake could be assessed both for naked DNA origamis and for origamis with attached CPPs to target the specific cell type. Furthermore, the uptake could be evaluated for DNA origamis coated with the CCMV capsid proteins, described in section 3.2.6, that showed a 13-fold improved delivery compared to naked DNA origami in HEK293 cells and showed DNA origami colocalized with nuclei. [21] In addition, it could be tested whether the transfection efficiency could be increased by coating with the cationic polymer PEI, as suggested by *Sun et al.* to induce endosomal escape. Also, the use of NLS peptides, as suggested by the same study to direct the DNA origmami to the nucleus, should be investigated. [23] Transfection rates should further be compared to transfection rates of currently used agents, such as PEI or Lipofectamine.

Expression could in the first instance be assessed via the reporter gene, EGFP, and thereby indicate whether expression could be obtained. Quantitative measurements could hereafter be performed using flow cytometry. Subsequently, the expression of Cas9 could be evaluated by co-transfection with the gRNA plasmid directing Cas9 to the TET-repessor gene for cleavage. Efficient cleavage could be assessed by subsequent flow cytometry measurements of the mCherry reporter stably integrated into the HEK293T-mCherry cell line.

Further studies on the model origami, 42hb, should be conducted to investigate the potential of FRET for detection of folded and unfolded states of the origami. A new ATTO647N labelled staple should be used to ensure maximal fluorescence intensity, and the DNA origami could be investigated at different states of unfolding by using a denaturing agent at different concentration increments. Furthermore, unfolding by heat should be tested properly, as no denaturing chemical will be able to compromise these results. Before transfection experiments, it could be useful to transfer the method to use on the OH1 and OF1 DNA origami designs by incorporating donor and acceptor fluorophores in close proximity in the design by attachment to staple strands. This way, stability could be tested both in cell culture media before transfection, and stability (folded/unfolded states) could be assessed live during uptake, by using confocal fluorescence microscopy. This would be valuable as to see if the DNA origami retain integrity until reaching the nucleus, and if unfolding inside the nucleus, as would be required for transcription, takes place.

Bibliography

[1] Dhanasekaran Balakrishnan *et al*. "Delivering DNA origami to cells". In: *Nanomedicince (Lond.)* 14.7 (2019), pp. 911–925.

[2] Jianbing Liu *et al*. "A DNA-Based Nanocarrier for Efficient Gene Delivery and Combined Cancer Therapy". In: *Nano Lett.* 18 (2018), pp. 3328–3334.

[3] Kenya Kamimura *et al*. "Advances in Gene Delivery Systems". In: *Pharm Med* 25.5 (2011), pp. 293–306.

[4] F. L. Graham *et al*. "Characteristics of a human cell line transformed by DNA from human adenovirus type 5". In: *J Gen Virol* 36 (1977), pp. 59–74.

[5] Helena Sork *et al*. "Lipid-based Transfection Reagents Exhibit Cryo-induced Increase in Transfection Efficiency". In: *Molecular Therapy—Nucleic Acids* 5.290 (2016), pp. 3–11.

[6] Patti A Longo *et al*. *Transient Mammalian Cell Transfection with Polyethylenimine (PEI) (ch. 18 in Methods in Enzymology.* 1st ed. Elsevier Science Technology, 2013, pp. 227–240.

[7] R. P. Goodman *et al*. "The single-step synthesis of a DNA tetrahedron". In: *Chem. Commun.* 0.12 (2004), pp. 1372–1373.

[8] Kwang Suk Lim *et al*. "Nano-Self-Assembly of Nucleic Acids Capable of Transfection without a Gene Carrier". In: *Adv. Funct. Mater.* 25 (2015), pp. 5445–5451.

[9] Paul W. K. Rothemund. "Folding DNA to create nanoscale shapes and patterns". In: *Nature* 440 (2006), pp. 297–302.

[10] Qian Zhang *et al*. "DNA Origami as an In Vivo Drug Delivery Vehicle for Cancer Therapy". In: *AcsNano* 8.7 (2014), pp. 6633–6643.

[11] Di Sheng Lee *et al*. "Cellular processing and destinies of artificial DNA nanostructures". In: *Chem. Soc. Rev.* 45 (2016), pp. 4199–4225.

[12] Yijiao Qu *et al*. "Self-Assembled DNA Dendrimer Nanoparticle for Efficient Delivery of Immunostimulatory CpG Motifs". In: *ACS Appl. Mater. Interfaces* 9 (2017), pp. 20324–20329.

[13] Aradhana Chopra *et al*. "Electrotransfection of Polyamine Folded DNA Origami Structures". In: *Nano Lett.* 16 (2016), pp. 6683–6690.

[14] Yasaman Ahmadi *et al.* "(Poly)cation-induced protection of conventional and wireframe DNA origami nanostructures". In: *Nanoscale* 10 (2018), pp. 7494–7504.

[15] Erik Benson *et al.* "DNA rendering of polyhedral meshes at the nanoscale". In: *Nature* 523 (2015), pp. 441–444.

[16] Rémi Veneziano *et al.* "Designer nanoscale DNA assemblies programmed from the top down". In: *Science* 352.6293 (2016), aaf4388-1–8.

[17] Shravan Kumar Sriraman *et al.* "Barriers to drug delivery in solid tumors". In: *Tissue Barriers* 2 (2014), e29528.

[18] Irene Canton Giuseppe Battaglia. "Endocytosis at the nanoscale". In: *Chem. Soc. Rev.* 41 (2012), pp. 2718–2739.

[19] Maartje M. C. Bastings *et al.* "Modulation of the Cellular Uptake of DNA Origami through Control over Mass and Shape". In: *Nano Lett.* 18 (2018), pp. 3557–3564.

[20] Qiao Jiang *et al.* "DNA Origami as a Carrier for Circumvention of Drug Resistance". In: *J. Am. Chem. Soc.* 134 (2012), pp. 13396–13403.

[21] Joona Mikkalä *et al.* "Virus-Encapsulated DNA Origami Nanostructures for Cellular Delivery". In: *Nano Lett.* 14 (2014), pp. 2196–2200.

[22] Henni Auvinen *et al.* "Protein Coating of DNA Nanostructures for Enhanced Stability and Immunocompatibility". In: *Adv. Healthcare Mater.* 6 (2017), p. 1700692.

[23] Wujin Sun *et al.* "Self-Assembled DNA Nanoclews for the Efficient Delivery of CRISPR–Cas9 for Genome Editing". In: *Angew. Chem. Int. Ed.* 54 (2015), pp. 12029–12033.

[24] Le Liang *et al.* "Single-Particle Tracking and Modulation of Cell Entry Pathways of a Tetrahedral DNA Nanostructure in Live Cells". In: *Angew. Chem. Int. Ed.* 53 (2014), pp. 7745–7750.

[25] Steven Hahn. "Structure and mechanism of the RNA polymerase II transcription machinery". In: *Nature Structural Molecular Biology* 11.5 (2004), pp. 394–403.

[26] Jeremy M. Berg *et al. Biochemistry.* 7th ed. Palgrave Macmillan, 2012, pp. 883–900.

[27] Noriko Wakabayashi-Ito Shigekazu Nagata. "Characterization of the Regulatory Elements in the Promoter of the Human Elongation Factor-lα Gene". In: *The Journal of Biological Chemistry* 269.47 (1994), pp. 29831–29837.

[28] Joshua P. Ferreira *et al.* "Modulating ectopic gene expression levels by using retroviral vectors equipped with synthetic promoters". In: *Syst. Synth. Biol.* 5 (2011), pp. 131–138.

[29] Jennifer A. Doudna Emmanuelle Charpentier. "The new frontier of genome engineering with CRISPR-Cas9". In: *Science* 346.6213 (2014), pp. 1258096-1–9.

[30] Benjamin P. Kleinstiver *et al.* "Engineered CRISPR-Cas9 nucleases with altered PAM specificities". In: *Nature* 523 (2015), pp. 481–485.

[31] Jennifer A. Doudna Emmanuelle Charpentier. "The new frontier of genome engineering with CRISPR-Cas9 (rev. summary)". In: *Science* 346.6213 (2014), p. 1077.

[32] Keun-Soo Kim *et al.* "Enhancement of Recombinant Antibody Production in HEK 293E Cells by WPRE". In: *Biotechnology and Bioprocess Engineering* 14 (2009), pp. 633–638.

[33] A. Schambach *et al.* "Woodchuck hepatitis virus post-transcriptional regulatory element deleted from X protein and promoter sequences enhances retroviral vector titer and expression". In: *Gene Therapy* 13 (2006), pp. 641–645.

[34] Q. Ding *et al.* "Enhanced efficiency of human pluripotent stem cell genome editing through replacing TALENs with CRISPRs". In: *Cell Stem Cell.* 12.4 (2013), pp. 393–394.

[35] S. Chen *et al.* "Genome-wide CRISPR screen in a mouse model of tumor growth and metastasis". In: *Cell.* 160.6 (2015), pp. 1246–1260.

[36] Björn Högberg *et al.* "Folding DNA Origami from a Double-Stranded Source of Scaffold". In: *J. AM. CHEM. SOC.* 131 (2009), pp. 9154–9155.

[37] Yang Yang *et al.* "DNA Origami with Double-Stranded DNA As a Unified Scaffold". In: *ACS Nano* 6.9 (2012), pp. 8209–8215.

[38] Yusuf E. Murgha *et al.* "Methods for the Preparation of Large Quantities of Complex Single-Stranded Oligonucleotide Libraries". In: *Plos One* 9.4 (2014), e94752.

[39] Rémi Veneziano *et al.* "In vitro synthesis of gene-length single-stranded DNA". In: *Scientific Reports* 8.6548 (2018), pp. 1–7.

[40] *vHelix - Free-form DNA nanostructure Design.* 2020. URL: `http://vhelix.net/` (visited on 01/06/2020).

[41] Petr Šulc et al. "Sequence-dependent thermodynamics of a coarse-grained DNA model". In: *The Journal of Chemical Physics* 137.13 (2012), p. 135101.

[42] A. Suma *et al.* "TacoxDNA: A user-friendly web server for simulations of complex DNA structures, from single strands to origami". In: *J. Comput. Chem* 40 (2019), p. 2586.

[43] Benedict E. K. Snodin *et al.* "Introducing Improved Structural Properties and Salt Dependence into a Coarse-Grained Model of DNA". In: *The Journal of Chemical Physics* 142.23 (2015).

[44] Steffan Møller Sønderskov. *Nanophotonics circuits formed by DNA self-assembly (Master thesis in Nanobiotechnology)*. Institute of Physics and Nanotechnology Aalborg University, Denmark, 2016, pp. 1–113.

[45] Clare E. Rowland *et al.* "Intracellular FRET-based probes: a review". In: *Methods Appl. Fluoresc.* 3 (2015), p. 042006.

[46] Eurofins Genomics (webpage). *Fluorescent Dyes - Select The Best Label For Your Assay!* 2020. URL: `https://www.eurofinsgenomics.eu/en/dna-rna-oligonucleotides/custom-dna-rna-oligos/custom-dna-oligos/dna-modifications/fluorescent-dyes/` (visited on 25/05/2020).

[47] Markus Ralser *et al.* "An efficient and economic enhancer mix for PCR". In: *Biochemical and Biophysical Research Communications* 347 (2006), pp. 747–751.

[48] RNAfold WebServer. URL: `http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi` (visited on 27/05/2020).

[49] Gruber AR *et al.* "The Vienna RNA Websuite". In: *Nucleic Acids Research* 36.2 (2008), W70–W74.

[50] Jonathan P. K. Doye *et al. The oxDNA coarse-grained model as a tool to simulate DNA origami*. URL: `https://arxiv.org/ftp/arxiv/papers/2004/2004.05052.pdf` (visited on 30/05/2020).

[51] Judy Callis *et al.* "Introns increase gene expression in cultured maize cells". In: *GENES DEVELOPMENT* 1 (1987), pp. 1183–1200.

[52] Ayaz Najafov Gerta Hoxhaj. *PCR Guru. An Ultimate Benchtop Reference for Molecular Biologists*. 1st ed. Academic Press, Elsevier, 2017, p. 23.

[53] Tahere Seifi *et al.* "Amplification of GC-rich Putative Mouse PeP Promoter using Betaine and DMSO in Ammonium Sulfate Polymerase Chain Reaction Buffer". In: *Avicenna Journal of Medical Biotechnology* 4.4 (2012), pp. 206–209.

[54] Marco Musso *et al.* "Betaine, Dimethyl Sulfoxide, and 7-Deaza-dGTP, a Powerful Mixture for Amplification of GC-Rich DNA Sequences". In: *Journal of Molecular Diagnostics* 8.5 (2006), pp. 544–550.

[55] Thermo Scientific. *PRODUCT INFORMATION. Taq DNA Polymerase (recombinant)*. 2015.

[56] Huseyin Avni Oktem *et al.* "Single-Step Purification of Recombinant Thermus aquaticus DNA Polymerase Using DNA-Aptamer Immobilized Novel Affinity Magnetic Beads". In: *Biotechnol. Prog.* 23 (2007), pp. 146–154.

[57] Henry C. Chiou *et al.* "Enhanced resistance to nuclease degradation of nucleic acids complexed to asialoglycoprotein-polylysine carriers ". In: *Nucleic Acids Research* 22.24 (1994), pp. 5439–5446.

[58] Yao-Guang Liu Robert F. Whittier. "Thermal Asymmetric Interlaced PCR: Automatable Amplification and Sequencing of Insert End Fragments from P1 and YAC Clones for Chromosome Walking". In: *Genomics* 25 (1995), pp. 674–681.

[59] Nayan P. Agarwal *et al.* "Block Copolymer Micellization as a Protection Strategy for DNA Origami". In: *Angew. Chem. Int. Ed.* 56 (2017), pp. 5460–5464.

[60] Heini Ijäs *et al.* "Reconfigurable DNA Origami Nanocapsule for pH-Controlled Encapsulation and Display of Cargo". In: *ACS Nano* 13 (2019), pp. 5959–5967.

[61] Rodrigo Galindo-Murillo *et al.* "On the absence of intrahelical DNA dynamics on the ms to ms timescale". In: *NATURE COMMUNICATIONS* 5 (2014), p. 5152.

[62] Erik Poppleton *et al.* "Design, optimization, and analysis of large DNA and RNA nanostructures through interactive visualization, editing, and molecular simulation". In: *Nucleic Acids Research* (2020), pp. 1–12.

[63] F. A. P. Crisafuli *et al.* "Characterizing the interaction between DNA and GelRed fluorescent stain". In: *European Biophysics Journal* 44 (2015), pp. 1–7.

[64] Philip J. Santangelo *et al.* "Dual FRET molecular beacons for mRNA detection in living cells". In: *Nucleic Acids Research* 32.6 (2004), e57.

[65] AAT Bioquest. *Spectrum [Cy5].* URL: `https://www.aatbio.com/fluorescence-excitation-emission-spectrum-graph-viewer/Cy5_cyanine_5` (visited on 30/05/2020).

[66] ATTO-TEC GmbH. *Product Information. ATTO 647N.* 2019. URL: `https://www.atto-tec.com/fileadmin/user_upload/Katalog_Flyer_Support/ATTO_647N.pdf` (visited on 30/05/2020).

[67] Asif Iqbal *et al.* "Orientation dependence in fluorescent energy transfer between Cy3 and Cy5 terminally attached to double-stranded nucleic acids". In: *PNAS* 105.32 (2008), pp. 11176–11181.

[68] Nicole Kretschy *et al.* "Sequence-Dependent Fluorescence of Cy3- and Cy5-Labeled DoubleStranded DNA". In: *Bioconjugate Chem.* 27 (2016), pp. 840–848.

Appendix A

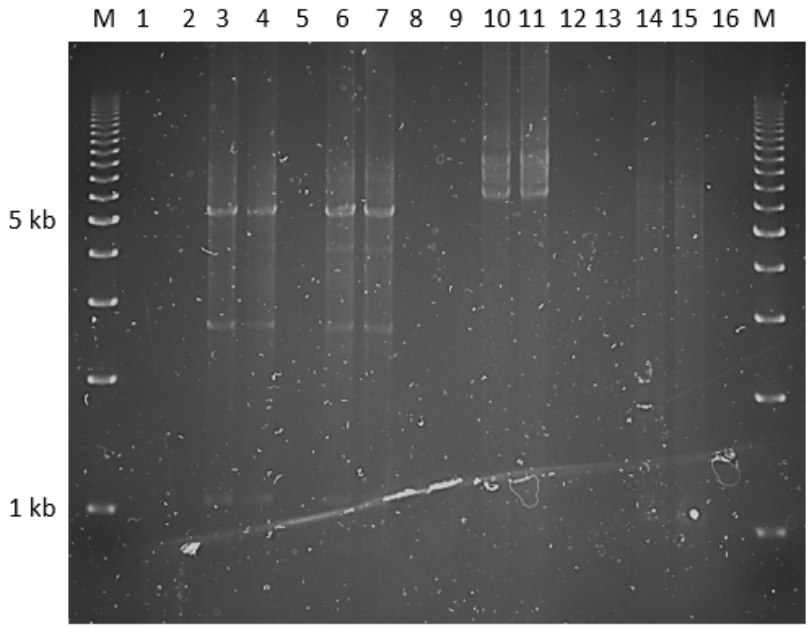


Figure A1: PCR (a2) post stained with SYBR Safe DNA stain. Lanes correspond to lanes in figure 6.7.

Figure A2: PCR setup for amplification with OneTaq 2x mastermix polymerase.

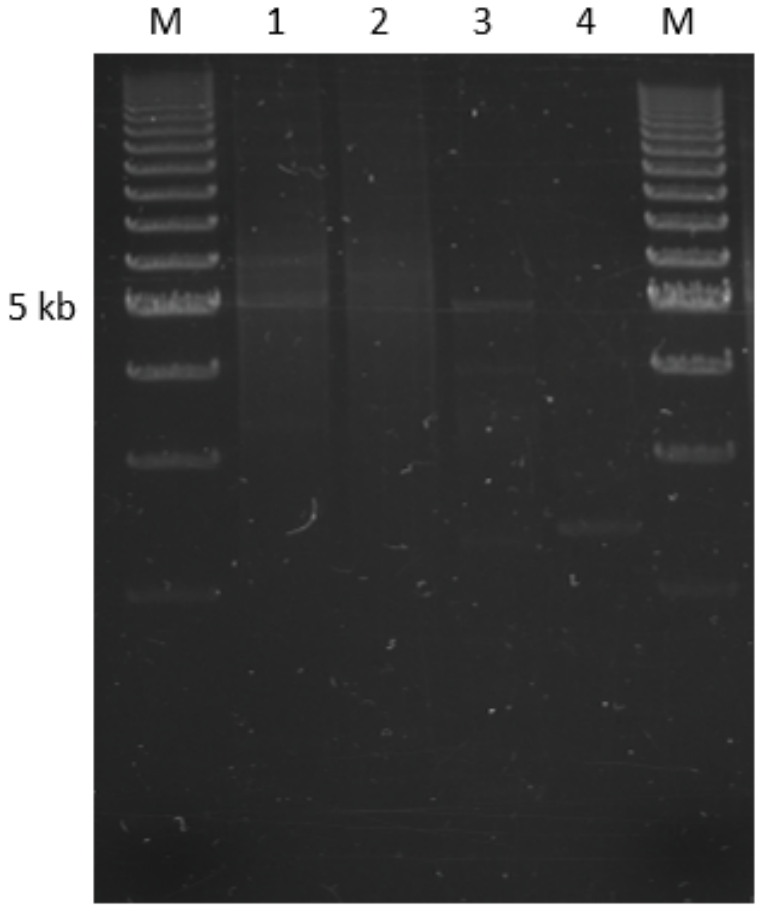


Figure A2: Lane 1: A.G. LA polymerase with T_a at 57 °C prepared as in table 4.3 but with 100 ng template, lane 2: A.G. LA polymerase with T_a at 68 °C prepared as in table 4.3 but with 100 ng template, lane 3: OneTaq mastermix prepared as in table **??**, lane 4: OneTaq mastermix GC prepared as in table **??**

Composition	**Concentration**	**Amount (μL)**	**Final concentration /amount**
Nuclease free H_2O	-	11/6	-
OneTaq 2xmastermix	2x	12.5	1x
Fpcas9gfp5p18m	10 μM	0.5	0.2 μM
R.pcas9gfp.5biot	10 μM	0.5	0.2 μM
Template DNA	20 ng/μL	0.5	10 ng
GC enhancer	5x	-/5	1x
Total		**25**	

Temperature (°C)	**Time (min)**	**Repeats**
94	3	x1
94	0.5	
54	0.5	x30
68	8	
68	10	x1
4	∞	x1

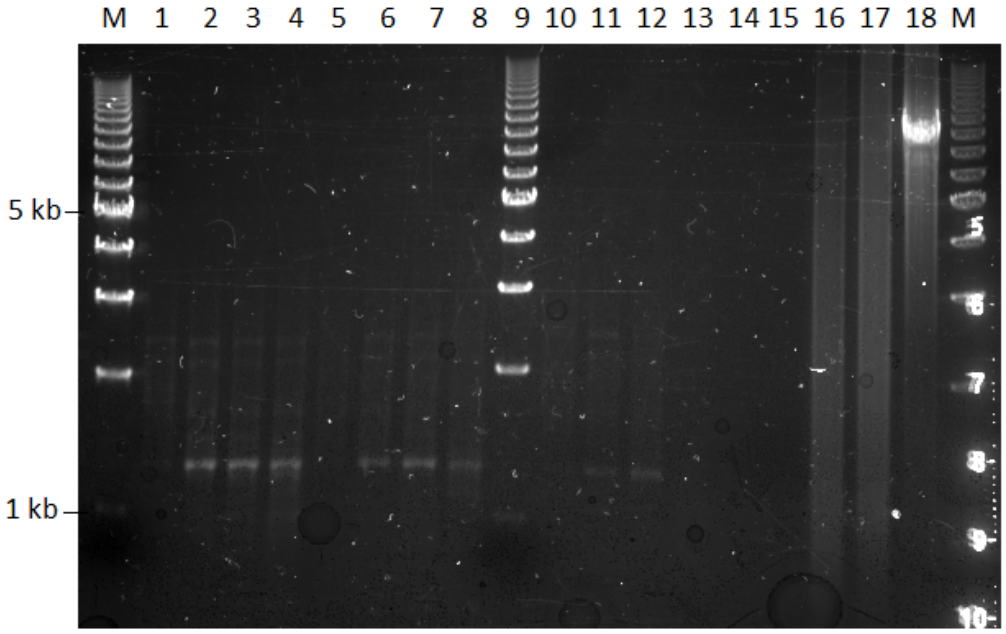


Figure A3: PCR (a5). Lane 1-4: 54 °C and 1.5, 2, 2.5 or 3 mM $MgCl_2$, respectively, lane 5-8: 56 °C and 1.5, 2, 2.5 or 3 mM $MgCl_2$, respectively, lane 10-13: 58 °C and 1.5, 2, 2.5 or 3 mM $MgCl_2$, lane 14-15: Neg. controls with 1.5 and 2.5 mM $MgCl_2$ at T_a at 56 °C, lane 16-17: PCR (a1) product from lane 3 in 6.6 as template in reaction, lane 18: PCR (a1) product from lane 3. M is the marker where EZ load^{TM} Molecular Ruler 1kb is loaded.

Table A2: PCR setup for amplification with Biotaq Red DNA polymerase.

Composition	**Concentration**	**Amount (μL)**	**Final concentration /amount**
Nuclease free H_2O	-	29/28.5/28/27.5	-
dNTP mix	10 mM (each)	1	0.2 mM (each)
NH_4 buffer	10x	5	1x
$MgCl_2$	50 mM	1.5/2/2.5/3	1.5/2/2.5/3 mM
F.pcas9gfp.5phos	10 μM	2.5	0.5 μM
R.pcas9gfp.5biot	10 μM	2.5	0.5 μM
BioTaq Red polymerase	1u/μL	2.5	2.5 u
Template DNA	4 ng/μL	1	4 ng
Total		**50**	

Temperature (°C)	**Time (min)**	**Repeats**
95	3	x1
95	0.5	x35
54/56/58	0.5	
72	3	
72	15	x1
4	∞	x1

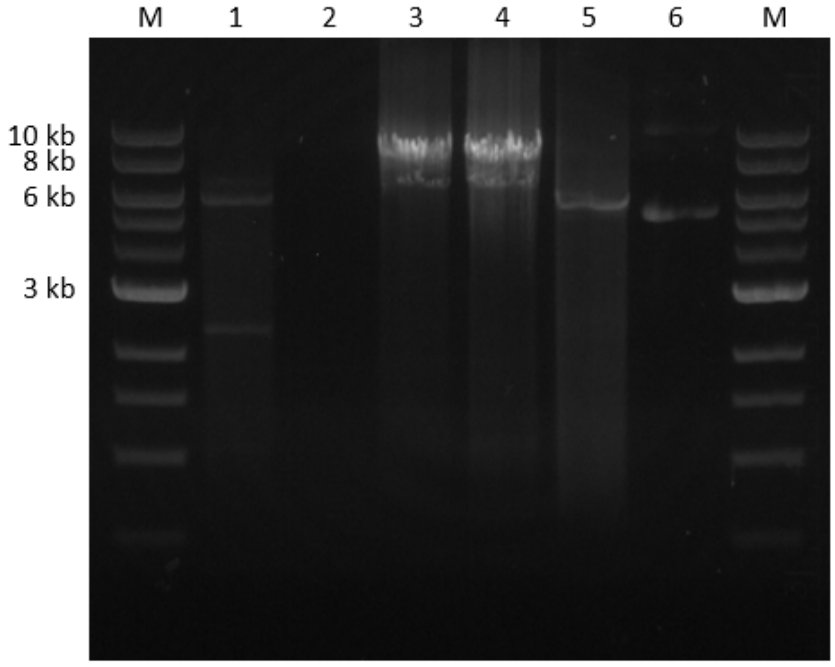


Figure A4: Lane 3 and 4: Half volume loaded sample of the products in lanes 1 and 2 in figure 6.13.

Table A3: Phusion PCR conditions. Upper table shows standard PCR and lower table shows touchdown PCR.

Temperature (°C)	**Time (min)**	**Repeats**
98	0.5	x1
98	10sec	
65.7	0.5	x30
72	2	
72	10	x1
4	∞	x1

Temperature (°C)	**Time (min)**	**Repeats**
98	5	x1
96	0.25	
75-65 (-0.5 °C/cycle)	0.5	x20
72	2	
96	0.25	
65.7	0.5	x30
72	2	
72	10	x1
4	∞	x1

Table A4: PCR conditions for amplification of GC-rich region with Taq polymerase. Upper table shows standard PCR and lower table shows touchdown PCR.

Temperature (°C)	**Time (min)**	**Repeats**
95	3	x1
95	0.5	
47	0.5	x30
72	1.7	
72	15	x1
4	∞	x1

Temperature (°C)	**Time (min)**	**Repeats**
95	3	x1
95	0.5	
57-47 (-0.5 °C/cycle)	0.5	x20
72	1.7	
95	0.5	
47	0.5	x30
72	1.7	
72	15	x1
4	∞	x1

Table A5: Touchdown PCR conditions for amplification of GC-rich region with Taq polymerase, performed by Chris Stubberup, MSc student in the same course.

Temperature (°C)	**Time (min)**	**Repeats**
98	5	x1
96	0.25	
62-47 (-0.5 °C/cycle)	0.5	x30
80	1.7	
96	0.5	
47	0.5	x30
80	1.7	
80	10	x1
4	∞	x1

Appendix B

Additional images describing the design of OF1:

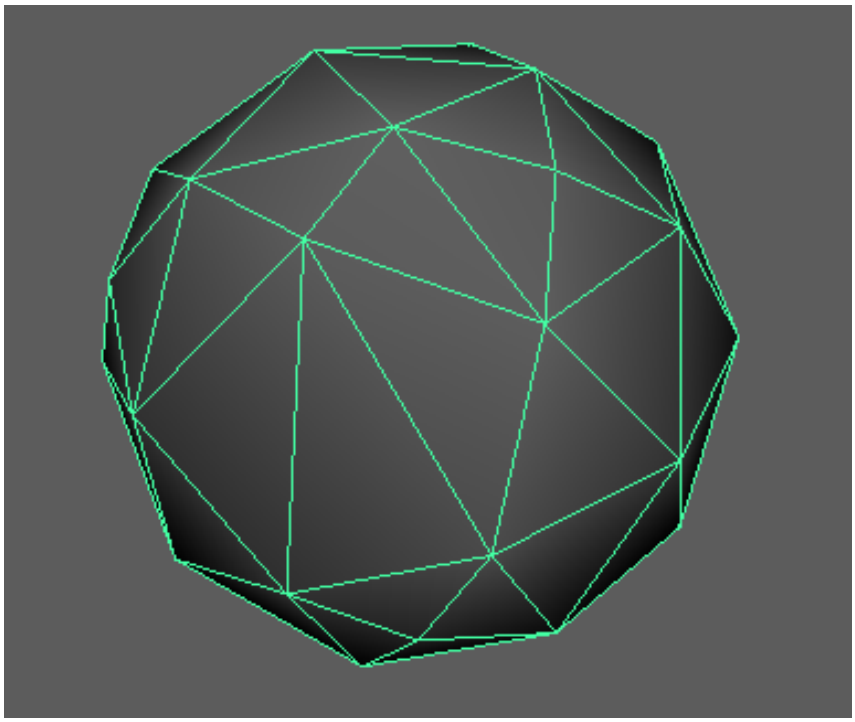


Figure B1: The OF1 DNA origami mesh.

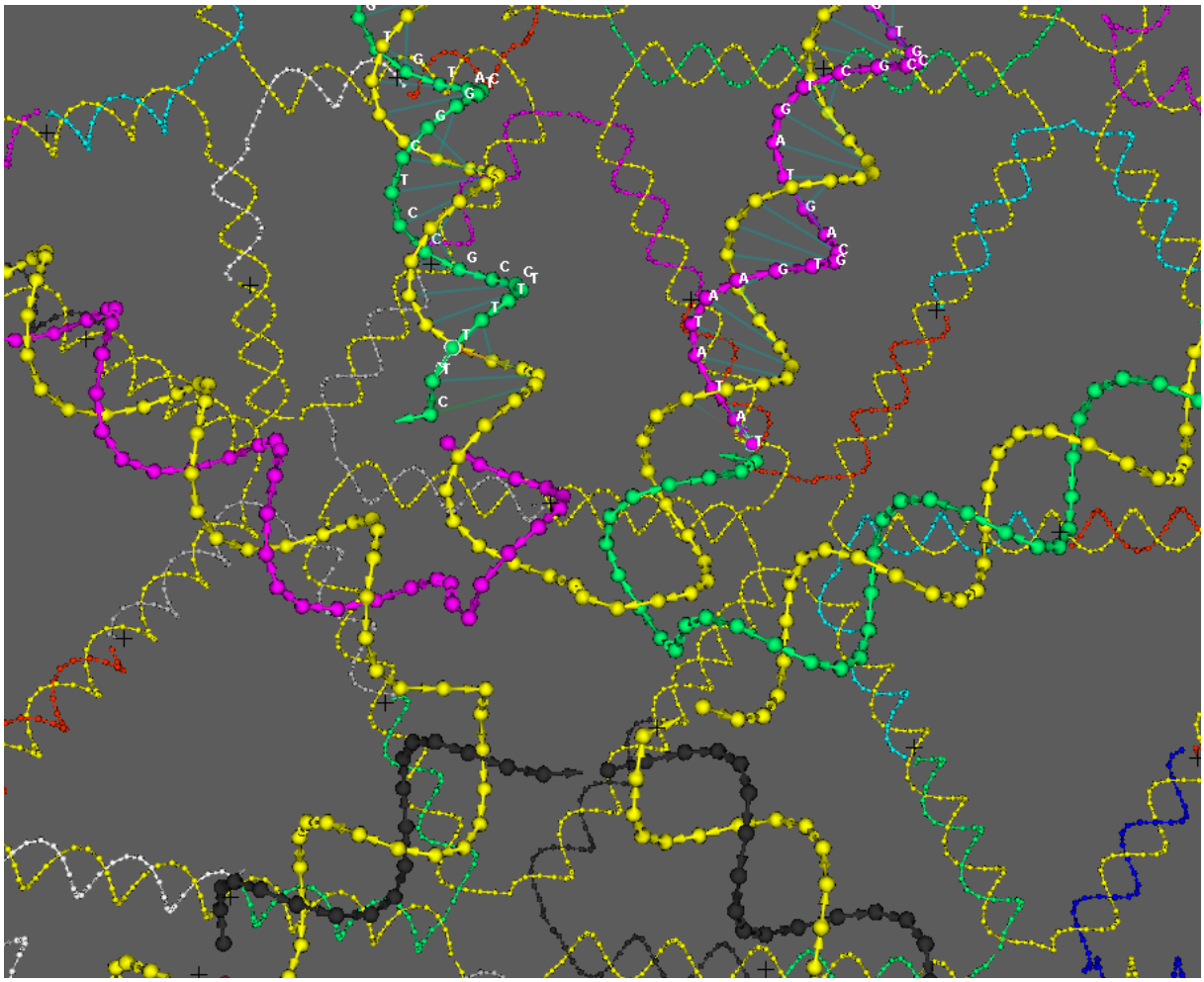


Figure B2: Oligos for double-stranded promoter regions in the OF1 origami.

Appendix C

Core staples in 42hb:

AGTTAGCGTAACGATCTAAAGTGATACCCCCACGCGTATAGC
AGGTGAACGGTCGCCCGTCGATAAACAGTACCAAGCGCGAAA
AATTCGAGAGCTAAACAGGAAACTATTAAAGAACGTGGACTC
TCCTTTGTCCTCGTGAACGGTACGCCAGGAGAGAGCGGCCCT
ATTTTAAACGAGCAAGAAGTGTTTTTATTGCCCTTACTCGTC
TTATCATGCGCGTAGAGGCCACCGAGTAGTGAGACTAGGGGC
TTATCATGCGTAACCGTTGTAGCAATACGGCCTTGCTTGCTG
CATATTCTCAACAGATAACCTCATAGCGTTAGGCAATATTTATAAGAAA
CTGATTAATTGTTTGGATTATATATCAAAACGTCATAGGAAT
GGAAGGTTAAATCAAATTCGCTGTGTAGGTAAAACTGTCCAG
GAATTGAAGCGGAAAATAATGGACAGCTGGCGAAAAAACCAC
GCTCGCCAATAGTGCTGAATTCTGCCAGGGGATAGAAAACAT
TTTCTTATTAATAAGACAAAACAGTATGTTAGCAA
TGATACCTTTAGAATATTAAAAAAACGACGGCCAG
CTTGTTATATTTTGATATTTTGATAAATAAGGCGT
CGTGGTGGCCATCTAATAGATGTTATTATTTTCCCAGTCACG
CAGGGCTAAATTTTCGCGAGAGAATAAACACCGGA
GAGTAAAGGTGGGCAACTAATAGTAACATTAAGTTGGGTAAC
CTGGTTGATTTTTTATCCAATTTACTAGAAAAAGC
TGACCTCCTTGAATAATATCTAACAAAGGGGGGATGTGCTGC
AAAAATCACAGAGGAATGGATTATTTACGCAGAACAATATTACCGCCAG
GCAAATGAGAGTCAGCTTAGGTTCTTACCAGTATAAATCGCC
AAGAATACGTGGATACGAGCCTAATTGCCCCAGCA
GATAGTTCGTTGGGCGATTGATGATTAAGACTCCTGCTTTAA
GATTCACCAGTCACGAGGCGGTTCACCAAAAGAGT
ATTGGCACAGCAGAAGATAAATAAAATCATTTCTCCCATCAA
CGTCTGATGAGGCGGTCAGTAGCCAGCAAACCTCATGGCAAA
GCTCAATCCATTGCAACTATCTTCTTTGAGCGGTCAGGGAAG
CTCACATGGAAGCACACAATTCCACACAACCGAGC
AAGCGGTTAGACAGTAGAATCCAACGTCAAAGGGCGAAAAAC
TTGCAGCCCAGTCGGGAAACCACCCTTCTAGTCTTTAATGCG
TGGCCCTAATCCTGCGTATAACGTCTATCAACTAAATCGGAA
CACCGCCCCAGCTGCATTAATATTCTGGATAGCCCTAAAACA
AGCTGATAATCAGTCTATGGTCCCTAAAGGGAGCCCCCGATT
TCACCGTTAAGTTTAAGCCAGAATGGAAACAAATAACCAGAG
GGGCAACCCAACGCGCGGGGAACGACCATAAAAATACCGAAC
TTATAAAGGCGAAATATTTACTCGAATTCAGGAAGGTCAATCTGAACGGTAATCGT

TCAGGCCGATTAAAGGGATTTCCACGCTGACTCTAGAGGATCTATTTAA
AATTAACCACCACAGGCGAACGTGGCGAGAAAGGAACGCTGC
AGAGCGGCAACTCGGTATTAGGATGATGAAACGGCACTTTTG
CGTGCTTCCCGAACAATACATAAAACAACGTCGGAATTTCAA
CTACAGGTTTGCGGTTAGGAGTTTTTTATAGCCAGTTAAATG
TGCGCCGTAGAGCTTGACGGGGAAAGCCCCCGCCGCAGAAGG
CGCTTAACTGTCCATCACGCAGGGTGGTCGCATTTCACATAA
ACTCAGGCCGTAATATCACCATAGCTATCTTACGA
AGAGTCCAAGAATACTTAAGTGTCCTACCTTTCTCAGGAGAAAGCTTTC
TGAAAGTATTAAGAGGCTGCACAACCTAAAACGAATGACCAA
GGAACCTATTATTCTAGCGGGTTATTAGTAGTAAAAAGATTA
TTAATGCCCCCTGCCAGGCGGGCATTTTGTTTAATAGCAAAG
TAACAGTGCCCGTAGAGGGTTGACTGTATTGTGAAACCCTGA
CCTTGAGCAACGGAGATTTGTATCATTATAACGGGGTCAGTGCCGGAAT
CATTCCACAGACAGCCCTCATCAAGCAAGCCCAAT
TCAATCAACCGGAAAATCCAAGAGAGAAACAAGAAAGCCTTT
GCCGGAAGCCGCCAAACAGCCGGAAGCGTCAACAAATAAAAA
CGACCTGACCCTCATAATTTGATTAACTAATGCAGATATATT
AATCCTCATCGCCTGATAAATAGAGCCGAACGAGC
ATCACCGTTTGTTTAACGTCAATGGCTTAGAGCTTGGGCGCG
CAGCATTCTTGATATTCACAAAGCGCAGGAGTGTACTGGTAA
TCAAGTTTTTAAGATAAATCAACAGTTCTGAGAAACTGTTTA
GAACGAGCGTTTGCAGCCACCTAAGGGAACCGAACAGAGGCA
TTGGGCTGGTAGAAATTGTATGGGATTTTGCTAAACAACTTT
TGAGATGCGGTCATCCTCAGACGAGGCGCAGACGGAACACTC
CCTGTAGCATGTACGAGCCACGTACCGCGATACAGTCTCTGA
TTCAACTTAACGGAGCTTTCGTAGTAAATGAATTTTCTGTAT
TTAATCAGCGCGTTAACCGCCCTCCATGTTACTTAGCGATTA
TTACCTTATCTACGAACAGCTTTTTGTCGTCTTTCCAGACGT
ATGCGATTGCCTTTCACCCTCTGTGTCGAAATCCGCAAAGTA
CACATTCTTGCCCTCATGACAAGAACCGAGACTTTCATCGGA
GAGCCAGATATTGAGAATACCCAAAAGAGACGAAGCCCTTTT
ACTGGCAGGGAGGGAAGGTAACATACCAGTCAGGAGCGCCGA
AATAACGCGGAAATTATTCATCCATTTGCATTACCCATCGAT
ACGCAATAGTAAGCCAATGAAATAACCCCCAGCTAGGGAGGT
TTAGACTAGAGGAACTTTACAATAAGAAGAATATAATGCTGTGTTTAGC
AATCTCCAAGACAGTTAGACTCCTCAAGTCATAATGACGAGA
GGATAGCTTAATGCACCGTGTAGTTAATTTCATCTATTGTAA
ACTGCGGCGGATTGAAAACAGATATTATCATGTTTTAAATATTTCGCAA
TTCATTGCTATTATCGGGAGACCAGTTAAAGTACGGTGTCTGTAGTTTG
GCGTTTTATAAGGCAACTAATCGTTGAAAATAGAA
AAATATCGTCAAAAATGAAAACCTAATTCTGTAATGGATTGA
AGACTTCTAATAGTATTTTTGAATTAATGGTTTGAAATACCG
GCCCGAAAACACCAAGATTTAAGGCTCCCAACAGTTTCAGCGGAGTGAG
AAAATCACTGAACAGTCTTTCCATTCCATATAACAGTTGATT
AAAAAAAGGAATACTCCACGGAATAAGTTTATTTTTCAGTTG
AAGTCAGCAATAAATGAGTAAGTCTGGCAGGAACGCGAACTCGAACCAC
AGGGTAAATGACCAACTGGCTCATTAAACAGTAGCAACCACC
TTGAGCGATCTTACTTAGCGAACCTCCCCTAAGAATTCATCG

GGATTAGCAAGAGTAATCTAACTTTGAAAGAGGACAAGGCACTGAGGAA
CAACGCTCCACCAGGACAGAAAGGTGTAGACAACAACCATCG
ATCCTGACTAATATGTAATTCATGTAATATAGCTTAGATTAT
CAATTTTCCGCCGCAGCAGCAAGGTTTACACCCTCATTTTCAGGGATAT
AATTGCTACGATTTGAACCAGCATCTTTAGAAGGAACCCTCAGCAGCGA
AGCTCAATTATCCCCCGCCTCAGCCCCCGTTTTGCGGCCGCTTTTGCGG
AAAAGGAGATCGTCTTAGGATTGAAACAAAAGAATACACTAA
AGCATGTATTTCAATTACCTGCGGCACCGCTTCTGTGCCAAG
AATCGGCTCGCGCACATCAAGTTGAGGAGACAGTG
ATTCCAATTTGAATTACATTTAGCCGTCGTAAGCA
CCAAGTAATAACGGAATTACCCACTAACACGAATA
TTTTATTCGCGAGGCGTTCCCAATTCTGAACAAGCCCTTTTA
AGGTAAACAGAGAGATAGCAAGTAGCACGGAATTAAGGAACC
GAGGCATCGACAAACATTACCGCTTATCCGGTATTGACTTGC
TACAAAGTTTGCCAACGACGACACTATCATAACCC
AGAGATCTATGACCTAAGAATTAGCAAACCAACAG
GCCTTTAAGATTCAGTCACAAAACATATAAAAGAA
GGTAGCTAAAATGTACGCAAAGACACGTTTACCAGGAGGGGG
CGGAGAGCGGGAGAAAATAATAATCAATTATATTTTCATTTG
TGATAAAGTCCAATAGGTGGCTCAATAGAAAATTCTATTACA
TTCTAGCCGCAAGGTAGATAACCTTATCATGGTCAATAACCT
ATGATATATAAATAACGTAGATTACCAGCGCCAAAAACGAAC
CATCAATAACCCTCAACGCGCTATTAAAACCATTAGATACAT
GACAGTCCTCAAATTATTACGGGGCGACATTCAACAAGAAAA
GGCCGGACAATGCCCAACATGCATCGAGCGAACGAGTAGATT
ACAACGCATACAAATTGGGTTATATAACAAAAGACGCTGAGA
TCTGACCTAAATAACTAGCATATTGTATCCGTAAT
CGGTTTAAGTTAAATCAGTACCTATTTCATCTTTGACCCCCA
AAATGCTGATGCAAAACCAATCTTCCTGATGGAAACATCGGG
TATATGTGTATCATATGCGTTCAGATTCAAAAGGGAGAAAAC
ACCTCCGATAGTGATTGAAAATGCTTCTCAGGTTTAATTATT
AAGCAAACCCGGGTACCACAGACAATATTTTTGAAAAAGACG
TGGGAACAAACAAAGAGGCGAAAACCAGGCAAAGCACGTTGT
TTCTCCGACGTTAACCTCGATTGGCTATTGACCTGAAAGCGTCCGCTTT
AGCGAGTTCGCATTTAAGCTACGAACTGCCAACAGAGATAGATGTCGTG
CTTTCATTCAGCTCGTGTAATTCGCCATGTAATAAAAGGGACGAATCGG
TCAGCTTACAACATATATGGTAAATACATACATAA
TTTGAGCAAAAGAAACTTTACCTGGAGTGGTTTGCGTTGCGCTCACTGC
CAGTACATATCTAACGGCTGATTTTCTTTTTGCGTATTGGGC
ATATATGATACAGTGGTTAGAACCTACCACTTCTG
AGGGGACCAGGCAAGGCAACGAGCTGAAAAGGTGGTTTGCGG
GATGAATTGAGTGATTGAAAGGCATCACCTGGTAATATCCCA
GTGCCGGATTATTCAGAAACCATCCCATTAGCAGC
GCCATTCTACAAAATGTCTTTGTCCTGATAACATACATCAAA
GCAACTGTGATTGCGAACGGGCTGTTTACATTAGAAGTCAGA
CGGTGCGAGAAACACCGCACTTTCAGCTGAACACCGGTCTTT
GGCCTCTTCGCTATTACGCAGAACAGTAAAGCCGTACGACGA

Table A5: Core staples in 42hb.

PolyT staples in 42hb:

AGGAACAACTAAAGGAATTTTTTT
ATTTACCGATTGGCGACAGGATTTGCACACAAGAAAAAGTACTTTCGAG
TTTTTAGGTCAGACGTTCCAGTAAGCTTTTT
TTTTTCGCATAGGCAAATCAACGTTTTT
ATGAAACATTAGCACAAGAAAAGATAGC
TTTTTACGTCACCAAGAACCGCCACCTTTTT
TTTTTTAACAAAGCCCAAAAGGATTTTT
TGCTCATATTACCCTGGCTGACGGTGTACAGACCAGGTTTTT
ACTTGAGTAAAGGTCGAGGAA
TTTTTCCGTCACCGCTGAGTTTCGTCTTTTT
TTTTTATTACGAGGCATAGTACATAACG
TTTTTACCAGTACAAACTACAACG
AGAGCAATAAAAACCAAAATAGCGAGTTTTT
TTTTTCCAGAAGGAAACGAATTATCATTTTT
TTTTTAGGCTTTTGCAAGCTTCAAAGCGATTTTT
TTTTTCAATAATAAGAGAGGCCGGAATTTTT
TTTTTACCAGACCGGAACTTTAATTGCTCTTTTT
AGAGTACGCAAACTATTAAGCTGTACCA
TTTTTATTAGTTGCTATGGTTGAGGCTTTTT
TTTTTCTTTTGATAAGAGGTCATTCATCAATAATCATA
TTTGAAGATAGAAGGCGCCCAAGATTTTGTAAATCGGTCAGTAATATCA
TTTTTAAATCAGATCCTTAAATCAAGTTTTT
TTTTTGCGAATAATAATTAGCAACGGCTATTTTT
TTTTTGCATTAACACCTCAGAGCTTTTT
TTTTTAGAGAATATTTGAGTTAAGCCTTTTT
TTTTTATAAAGCTAGCCTGAGAGTTTTT
AATCGGTAATAAAGTCCAATATCTACTAATAGTAGTATTTTT
AATTGAGAAGCCAACTTTTTA
TTTTTGTAGGGCTTCGAACAAAGTTATTTTT
TTTTTTCTGGAGCAAACAAGAGGTCATT
GAATCGAATATGTACCCCGGTTGATATTTTT
TTTTTCTGAGAGACTACCGCTCAACATTTTT
TTTTTATCAGAAAAGCCTGGTGTAGATGGTTTTT
ACGAGGGTTTTTCAGCAGATA
GTCACGTCCAAAAACGTAATC
AGAATCCATTTATCAAAATCATAGGTTTTTT
TTTTTTAATTTTCCCTTCCAGTAATATTTTT
TTTTTGCGCATCGTAACGTATCGGCCTCATTTTT
GACGACACGTGCATGTCAACCTCTAATCATCCTGTGAGCTAA
TTTTTAAGAAATTGCGTATAGCAAGCTTTTT
TTTTTGGAAGATCGCACTCCAGCCGCCAGGG
TGCACGTAATCCTGTCAGATG
TTTTTCATCAATATAAAACAGAAATATTTTT
TTTTTTGGTTGTGAAATGTCCCGTTTTT

TTTTTCTCAGAACCGCCACCCTCACGTAACA
CGCACGAGCCCGAGAAATCCC
TTTTTTCAATATCTGTCGCTATTAATTTTTT
TTTTTCCAAAATAATAGCTGTTTTTTTT
CCCCGCTTTATGACATTCATGTGGATGTTCTTCTAAGTTTTT
TTTTTCAGTGCCACGCTGAGATTAACACTTTTGAC
TTTTTCCTGTGTGAAATTGTTATGGTCA
ATCCGCTTAAAGTGTAAAGCCTGGGGTTTTT
TTTTTGAAATACCTACACGCCTGCAATTTTT
TTTTTTGCCTAATGAGTTTGATGGTGGTTTTTTT
GAACTCAAACAGGAAAAACGCTCATGTTTTT
TTTTTCAGAGGCTTTGATTAAACGGGTAATTTTT
TTTTTGCCTGAGTAGAAAACCCTCAATTTTT
TTTTTCCGAAATCGGCAATAGGGTTGAGTTTTTT
CAAGTGTATTAGTAATAACATCACTTTTTTT
TTTTTCTAGGGCGCTGGATGGCAATTTTTTT
TTTTTGTTGTTCCAGTTTGGAACA
AAAGCGAAAGGAGCGGGCGTTTTT
GTTTCCAGGACTAAGATATTCTCAGTGAAATTCGAAAGAAGTGCTATCA
TTTTTGTCATACATGGCTTTTGATCACCCTC
TTTTTAATACGTAATGCCACTACGAGATGAACCTTCAT

Table A5: PolyT staples in 42hb.

Extra staples in 42hb:

Name	Sequence
TBIT-00101R	GCAAAGACAATAGAACGCAAGACAAAGAATG TTAAACA ACATTTCATTTG
TIBT-00102R	GCAAAGACAATAGAAAAACTTTTTCAAATTT AAAATAAC AACCAATTAAT
TBIT-00103R	GCAAAGACAATAGAAGCAACTACAAAATACCCT CAGTTC ATCGATAAGTGTGAGGCTTGCAGGG
TBIT-00104R	GCAAAGACAATAGAAGAAGTTTCAGAGCCGAG CCACAG CGTCAGATATAAATAACCGATATATT
TBIT-00105	AATCCCCAAATCACATCATAA
TBIT-00106	ACCAAGTGCCATTCAGGCTGCGCCAGGG
TBIT-00107	AATCGTCTCAACCGTAAATAA

Table A5: Extra staples in 42hb.

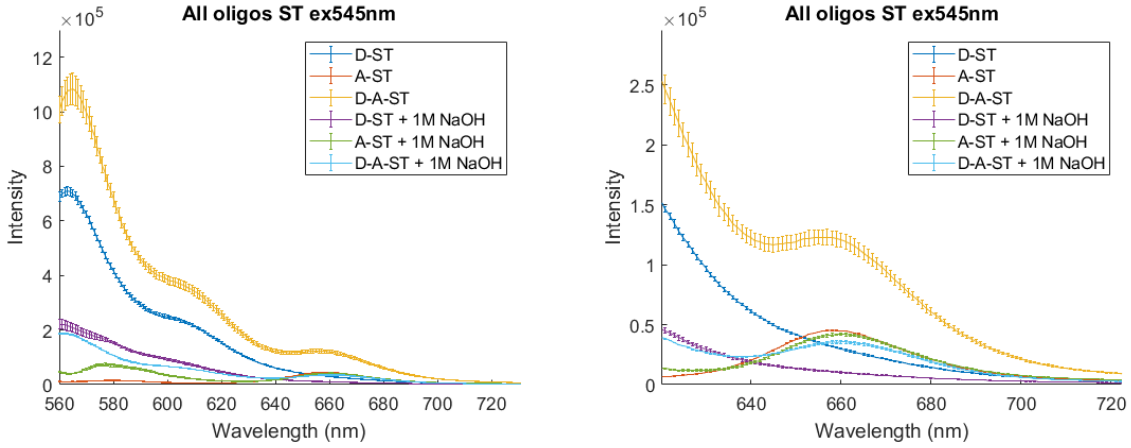
Staples with attached fluorophores in 42hb:

Name	Sequence
TBIT-00100	ATTO647N-TGCTTTGAAGTTTGAGATTAGAACAATTAAA TGTGTTTTTAG
TBIT-00108-Cy3	ATTCGCCTTGGGAAGGGCGATAAGGC-Cy3

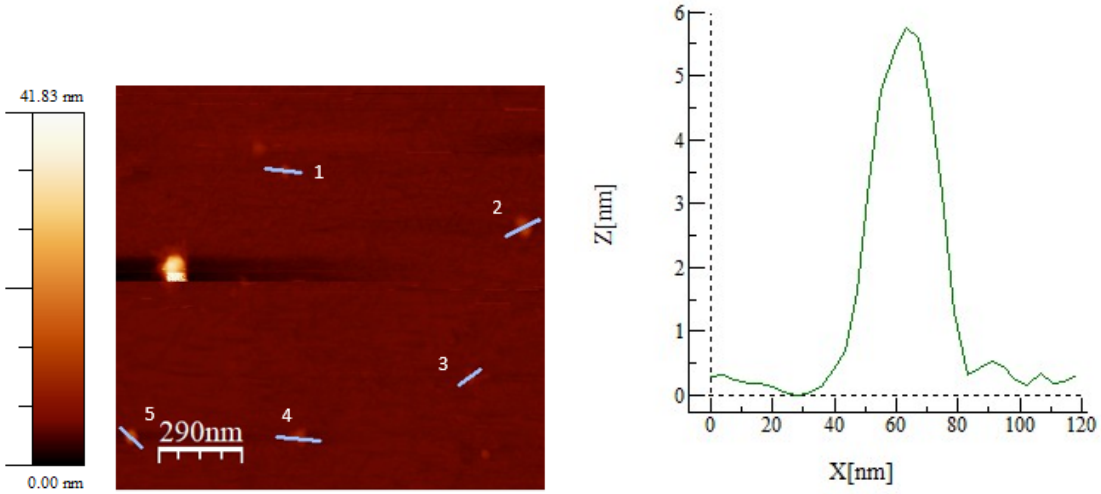
Table A5: Staples with fluorophores in 42hb.

Appendix D

Additional plots of FRET detection of oligo constructs, here all constructs before and after denaturation showed together:



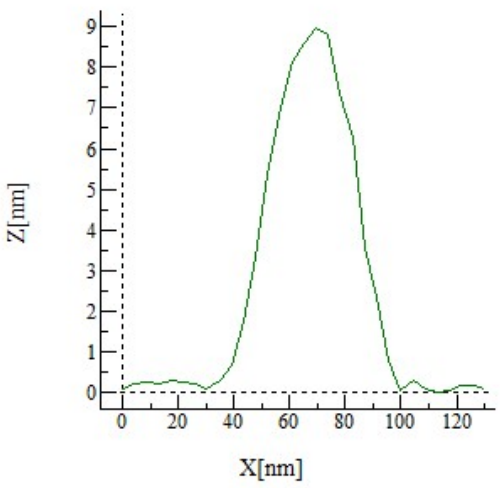
(a) **(b)**



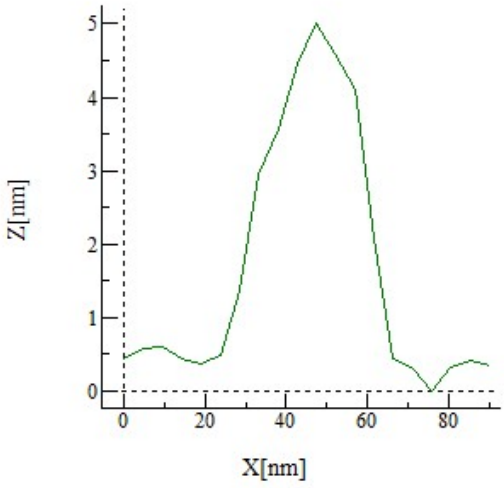
(a) Measuring profiles of structures in figure 6.37

(b) Profile of 1 from (a).

Figure B2

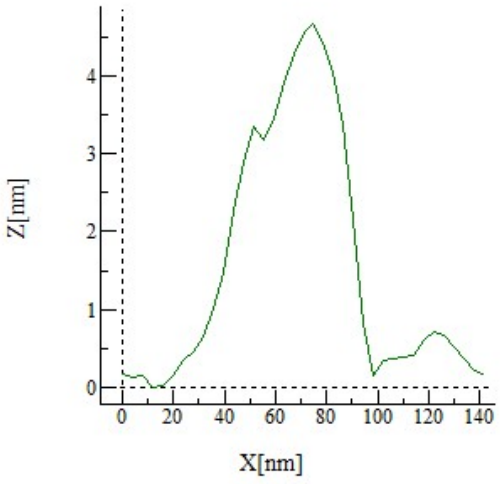


(a) Profile of 2 from (a).

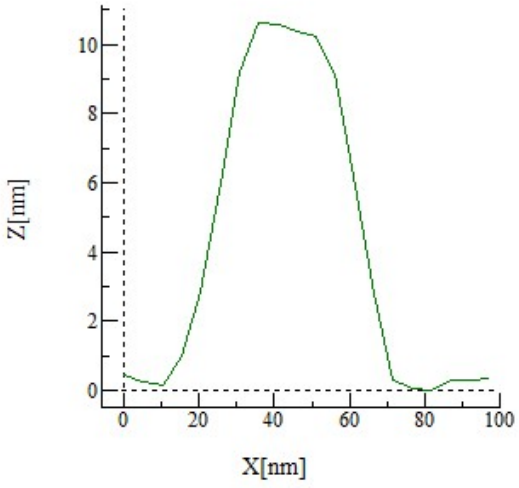


(b) Profile of 3 from (a).

Figure B2



(a) Profile of 4 from (a).



(b) Profile of 5 from (a).

Figure B2

Appendix E

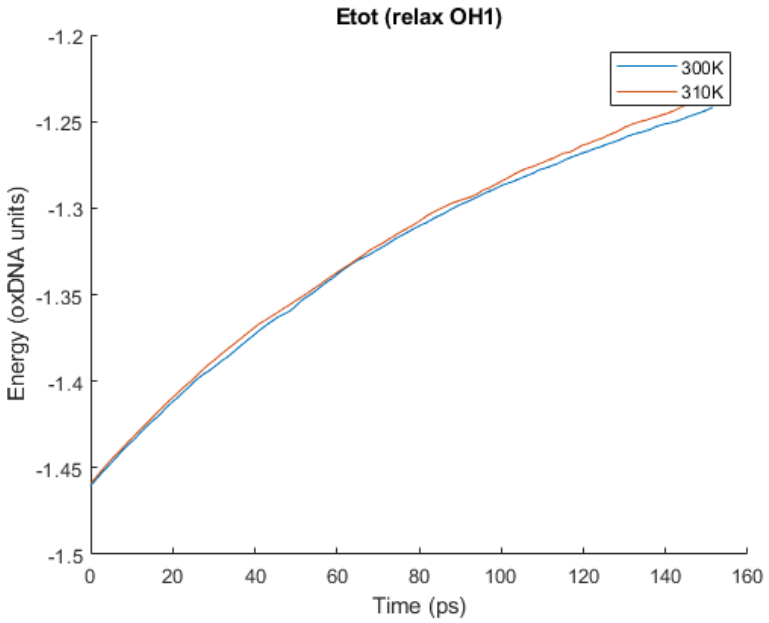


Figure B2: Total energies after relaxation of OH1 at 300 K and 310 K.