Master's Thesis

Nanostructures on a Vector
– Enzymatic Oligo Production for DNA Nanotechnology

Camilla Sandén
LITH-IFM-A-EX--12/2686--SE
Nanostructures on a Vector
– Enzymatic Oligo Production for DNA Nanotechnology

Camilla Sandén

Master's Thesis carried out at the Swedish Medical Nanoscience Center,
Department of Neuroscience at the Karolinska Institute

Supervisors:  
Björn Högberg
SMNC, the Karolinska Institute  
Johan Hurtig
IFM, Linköpings University

Examiner:  
Daniel Aili
IFM, Linköpings University

Linköping, June 15th, 2012
Nanostructures on a Vector, Enzymatic Oligo Production for DNA Nanotechnology.

Camilla Sandén

Sammanfattning
Abstract

The technique of DNA origami utilizes the specific and limited bonding properties of DNA to fold single stranded DNA sequences of various lengths to form a predesigned structure. One longer sequence is used as a scaffold and numerous shorter sequences called staples, which are all complementary to the scaffold sequence, are used to fold the scaffold into intricate shapes. The most commonly used scaffold is derived by extracting the genome of the M13 phage and the staples are usually chemically synthesized oligonucleotides. Longer single stranded sequences are difficult to synthesize with high specificity, which limits the choices of scaffold sequences available. In this project two main methods of single stranded amplification, Rolling Circle Amplification (RCA) and the usage of helper phages, were explored with the goal to produce both a 378 nt scaffold and staple sequences needed for folding a DNA origami structure. To facilitate imaging by Transmission Electron Microscopy (TEM) of this small structure, the DNA origami structure was created to form a polymer structure. Production of the scaffold sequence in high yield was unsuccessful and no well-defined polymers were found in the folded samples, though a few results showed promise for further studies and optimizations. Due to time constraints of this project, only production of the scaffold sequence was tested. Unfortunately the scaffold produced by the helper phages was of the complementary strand to that used to design the DNA origami structure, and could therefore not be used for folding. The correct strand was produced by the RCA where the yield was too low when using Phi29 DNA polymerase for proper folding to take place, though small scale RCA by Bst DNA polymerase on the other hand showed promising results. These results indicate that the scaffold production may not be far off but still more experience in producing intermediate size oligonucleotides may be necessary before succeeding in high yield production of this 378 nt long sequence. The promise given by this production is to enable high yield, high purity, low cost and also an easily scalable process set-up. This would be an important step in future DNA nanotechnology research when moving from small scale laboratory research to large scale applications such as targeted drug delivery systems.

Nyckelord
Keyword

DNA origami, Rolling Circle Amplification, Phage, Scaffold, Oligonucleotides
ABSTRACT

The technique of DNA origami utilizes the specific and limited bonding properties of DNA to fold single stranded DNA sequences of various lengths to form a predesigned structure. One longer sequence is used as a scaffold and numerous shorter sequences called staples, which are all complementary to the scaffold sequence, are used to fold the scaffold into intricate shapes. The most commonly used scaffold is derived by extracting the genome of the M13 phage and the staples are usually chemically synthesized oligonucleotides. Longer single stranded sequences are difficult to synthesize with high specificity, which limits the choices of scaffold sequences available. In this project two main methods of single stranded amplification, Rolling Circle Amplification (RCA) and the usage of helper phages, were explored with the goal to produce both a 378 nt scaffold and staple sequences needed for folding a DNA origami structure. To facilitate imaging by Transmission Electron Microscopy (TEM) of this small structure, the DNA origami structure was created to form a polymer structure. Production of the scaffold sequence in high yield was unsuccessful and no well-defined polymers were found in the folded samples, though a few results showed promise for further studies and optimizations. Due to time constraints of this project, only production of the scaffold sequence was tested. Unfortunately the scaffold produced by the helper phages was of the complementary strand to that used to design the DNA origami structure, and could therefore not be used for folding. The correct strand was produced by the RCA where the yield was too low when using Phi29 DNA polymerase for proper folding to take place, though small scale RCA by Bst DNA polymerase on the other hand showed promising results. These results indicate that the scaffold production may not be far off but still more experience in producing intermediate size oligonucleotides may be necessary before succeeding in high yield production of this 378 nt long sequence. The promise given by this production is to enable high yield, high purity, low cost and also an easily scalable process set-up. This would be an important step in future DNA nanotechnology research when moving from small scale laboratory research to large scale applications such as targeted drug delivery systems.
SAMMANFATTNING

ACKNOWLEDGEMENT

I would like to give thanks to:

Björn Högberg for the opportunity, and supervision in completing my master’s thesis.

Cosimo Ducani, Corinna Kaul, Alan Shaw and other members of the Högberg for their support within the lab.

My examiner Danial Aili and supervisor Johan Hurtig of LITH for their guidance.

Daniel Wickman for joining me in the fight against computer breakdowns.

Ferdinand Choong for advice and use of computer.

My excellent coworkers at the Swedish Medical Nanoscience Center for their comradery.

Swedish Medical Nanoscience Center and the Department of Neuroscience at the Karolinska Institute for use of their core facilities.
**ABBREVIATIONS**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Adenine</td>
</tr>
<tr>
<td>AFM</td>
<td>Atomic force microscopy</td>
</tr>
<tr>
<td>bp</td>
<td>Base pair</td>
</tr>
<tr>
<td>BSA</td>
<td>Bovine serum albumin</td>
</tr>
<tr>
<td>C</td>
<td>Cytosine</td>
</tr>
<tr>
<td>cfu</td>
<td>Colony forming units</td>
</tr>
<tr>
<td>Da</td>
<td>Dalton</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>dNTP</td>
<td>Deoxyribonucleotide triphosphate</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>EM</td>
<td>Electron microscopy</td>
</tr>
<tr>
<td>g</td>
<td>Gram or Gravitational acceleration</td>
</tr>
<tr>
<td>G</td>
<td>Guanine</td>
</tr>
<tr>
<td>HC</td>
<td>Honeycomb</td>
</tr>
<tr>
<td>kb</td>
<td>Kilo base pairs</td>
</tr>
<tr>
<td>LB</td>
<td>Luria broth</td>
</tr>
<tr>
<td>M</td>
<td>Molar</td>
</tr>
<tr>
<td>nt</td>
<td>Nucleotide</td>
</tr>
<tr>
<td>NZY</td>
<td>Casein hydrolysate, Yeast</td>
</tr>
<tr>
<td>OD</td>
<td>Optical density</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>PEG</td>
<td>Polyethylene glycol</td>
</tr>
<tr>
<td>pfu</td>
<td>Plaque forming units</td>
</tr>
<tr>
<td>PPB</td>
<td>Phage prep buffer</td>
</tr>
<tr>
<td>RCA</td>
<td>Rolling circle amplification</td>
</tr>
<tr>
<td>rpm</td>
<td>Revolutions per minute</td>
</tr>
<tr>
<td>RT</td>
<td>Room temperature</td>
</tr>
<tr>
<td>SDS</td>
<td>Sodium dodecyl sulfate</td>
</tr>
<tr>
<td>SEM</td>
<td>Scanning electron microscopy</td>
</tr>
<tr>
<td>SQ</td>
<td>Square</td>
</tr>
<tr>
<td>T</td>
<td>Thymine</td>
</tr>
<tr>
<td>TBE</td>
<td>Tris, Borate, EDTA</td>
</tr>
<tr>
<td>TEM</td>
<td>Transmission electron microscopy</td>
</tr>
<tr>
<td>UV</td>
<td>Ultra violet</td>
</tr>
<tr>
<td>YT</td>
<td>Yeast, Tryptone</td>
</tr>
</tbody>
</table>
1 INTRODUCTION

1.1 BACKGROUND

The field of DNA origami research is young and originated in 2006 [1]. It uses DNA to fold structures in the same manner as paper is used in the Japanese art of origami to fold different shapes, this is how DNA origami got its name [2]. It utilizes the well-defined and specific properties of DNA to construct nanoscale structures. Due to the predictable binding in base pairing [3] DNA can theoretically be programmed into any desired structure [4]. When folding a DNA origami structure one long single stranded DNA sequence is used as a scaffold and shorter single stranded oligonucleotides are used as staples to get the desired shape [1].

Since DNA origami is a relatively new science there are no commercial applications on the market yet. Though the research and usage of synthetic biology and genes are expanding, the needs for oligonucleotides which are cheaper and with low error rate are growing [5]. The goal is to be able to produce DNA sequences of any sequence or size. Today the most frequently used scaffold strand is purified genome from the M13mp18 phage and the staples are chemically synthesized [2]. The disadvantage of using this scaffold is that the length of the scaffold cannot be chosen in respect to the designed structure, it creates an upper size limit to structures. In the cases where the scaffold sequence is too long for the designed structure the excess bases are commonly placed in a loop at one edge of the structure [1]. It might also be preferable to control the exact sequence of the scaffold to enhance binding strength in crucial positions [6].

1.2 PURPOSE AND AIDS

The main goal of this project was to produce the scaffold and the staples required for a designed structure by taking advantage of a strain of Escherichia coli. The cells internal machinery such as DNA proofreading, DNA replication and cell growth were utilized to produce copies of the sequences with a low error rate. The other component in the main goal was the fact that this method can be considered an all-in-one or one-pot reaction since all the building blocks of the design is present from start. This strategy of a more straight forward approach may help in the scale up process. To reach the main goal minor aims were specified:

- To create a cell line with transformed cells containing the phagmid with the designed sequences.
- To amplify the sequences by rolling circle amplification (RCA) in vitro.
- To amplify the sequences by inoculation with helper phage in vivo.
- To use the amplified sequence to fold three different polymer DNA origami structures, each with two variations on the lengths of the protruding staple strands.
1.3 **Scope**

The focus of this project was towards the functionality of the production of scaffold and staple sequences. The scaffold sequence as well as the DNA origami designs used for verifying the process were defined in the project specification.

The project was divided into two sequential phases:

1. **Scaffold production**

   This first phase is a simplified version of the complete process. Only the scaffold sequence was produced and synthetic staples ordered to fold the DNA origami structures. In this first phase the protocol for the production was set up which then was used in the second phase. This phase was considered complete when successful scaffold production was confirmed by proper folding of at least one of the designed DNA origami structure into well-defined polymers. If more than one set of staples form polymers, when folded with the scaffold, the staples generating the highest yield and the most well-structured polymers were chosen for the next phase of the project.

2. **Scaffold and staples production**

   The second phase was the complete aim of the project which was to create the all-in-one production of the scaffold and staple sequences. The production was limited to just one set of staples determined by the first phase.

The production pathways that were investigated are outlined below.

Two different methods of amplification were used to produce the single stranded scaffold/scaffold-staples in both phases. The initial steps for both methods were identical. The pseudogene was constructed by using PCR and extended primers including the restriction sites for HindIII and Nb.BsrDI to elongate the scaffold/scaffold-staples sequence. The pseudogene was inserted into the vector pDrive with ligation of blunt ends creating a phagemid containing genes for antibiotic resistance, the f1 origin for phage genome replication and the designed scaffold/scaffold-staples sequence. Ultracompetent XL10-Gold cells were transformed with the phagemid by heat shock and resulting colonies subsequently screened in respect to the pseudogene by HindIII digestion.

One method was based on the principles of the RCA using the Phi29 or the Bst DNA polymerase. First the phagemid was extracted and then the pseudogene was cut out by HindIII digestion. The pseudogene was circularized by ligation of the blunt ends by the T4 ligase. To create a suitable template for the RCA reaction the phosphate backbone was cut (nicked) by Nb.BsrDI. The RCA was evaluated with respect to a few factors; large (complete phagemid) or small (circularized pseudogene) template, nicked or un-nicked, added primer or not and finally Phi29 or Bst DNA polymerase. The product from the RCA was a single stranded continuous repeat of the strand on which the primer or nicked sequence was located.

The other method utilizes the ability of helper phages to replicate and incorporate genome produced from a phagemid by generating a single stranded circular copy and placing it into new phage particle. A culture of transformed XL10-Gold cells in the exponential growth phase was inoculated with the VCSM13 helper phage and after incubation at 37 °C the phage genome was extracted.

The final step of both methods were digestion with BseGI to cut out the scaffold and staples from the single stranded products.
2 Theory

2.1 The Properties of DNA

The deoxyribonucleic acid (DNA) molecule is most known for its capability to store information. Because of its predictable structure the use of the molecule has expanded into the nanoscience research [1]. This section focuses on the properties of DNA that are exploited in the DNA origami research.

2.1.1 Watson-Crick Base Pairing

The building blocks of DNA are four different nucleotides that form a polymer with a repeating sugar-phosphate backbone that the bases are attached to. The four different bases are; adenine (A), cytosine (C), guanine (G) and thymine (T). The bases bind to each other by hydrogen bonds which are only possible between A-T and C-G. The DNA structure is compatible with any sequence of bases and because of the specific bonding of base pairs the sequence of one strand completely determines the sequence along the other strand (also known as the complementary strand) [3].

2.1.2 Secondary Structure

The primary structure of DNA refers to the exact sequence of nucleotides from the 5’ to the 3’ end. When two strands bind together they form a double helix where the strands are antiparallel, where the 5’ end of one strand is the 3’ end on the complementary strand. This is known as the secondary structure of DNA [7]. When DNA is circular the double helix can be twisted into a superhelix by twisting around its own axis, the DNA molecule is then considered to be supercoiled. In comparison to the open circular conformation, supercoiled DNA is more compact in its shape [8], [9], [10].

In cases where the DNA sequence has inverted repeats separated by a few bases these sequences can bind to each other when the DNA is single stranded. These structures looks like hairpins and are known by that name [7].

2.1.3 Restriction Site

Restriction endonucleases, or more commonly called restriction enzymes, are enzymes that can modify DNA. A restriction enzyme can only modify DNA at a restriction site which is a short, often double stranded, sequence unique to each restriction enzyme. The cleavage done by the restriction enzymes can be divided into different categories defined by the product, blunt ends or sticky ends. In the case of blunt ends the restriction enzyme cleaves both strands at the same position. Sticky ends are the result when the restriction enzyme cleaves the two strands at different positions and then an overhang of a couple of bases is created which, on the two fragments, are complementary. Other restriction enzymes can cut out just one base from one of the strands within the restriction site, this is called nicking [11].
2.1.4 Holliday Junction

A Holliday junction is essentially a crossover between strands. It is as travelling on a road, sooner or later a junction will come where another road crosses, an ordinary four way intersect is exactly what a Holliday junction looks like (Figure 1). In DNA origami Holliday junctions are placed evenly spaced throughout the design to give stability to the structure [12]. Though the junction with four branches is the most frequently used one, junctions have been known to be stable from three [13] up to six branches [14] separated with a minimum angle of 60° [15].

![Holliday Junction Image](image.png)

**FIGURE 1:** The Holliday junction is made up of four oligonucleotides illustrated by the colored arrows. They form a cross by base pair bonding of complementary strand which in the figure are the black lines.

In nature, branched DNA such as Holliday junctions play an important role in recombination. It is formed as an intermediate when two helices breaks and form crossover linkage with one another [16]. This takes place during meiosis, the forming of sex cells, and generates a genetic diversity from limited genetic material [17], [18].

2.2 How It Started

It all began in 1960 when a challenge was given to write the entire 24 volumes of the Encyclopedia Britannica on the head of a pin. The thought of miniaturization was already popular at this time but to complete this challenge a new way of approaching the problem was necessary. To fit all the pages on the given area it had to be 25 000 times smaller than the actual size of the Encyclopedia [19].

One of the first DNA nanostructures was a two dimensional crystal made in 1998 [20]. In 2006 Rothemund suggested another way of using DNA to build nanostructures. He used one long single stranded DNA sequence as a scaffold and mixed this with shorter single stranded oligonucleotides also known as staples. When hybridized the staples folded the scaffold into the designed shape [1]. This method is called DNA origami which refers to the Japanese art of folding a flat sheet of paper into different shapes [2].

There are two different overall strategies to produce nanostructures;

- Top-down: Manipulations to a nanoscale template are done by using externally controlled devices.
- Bottom-up: By using the properties of some biomolecules, such as DNA, self-assembly in parallel steps can be made without the need for manipulations.

The higher predictability in bottom-up constructions as well as the possibility to exact control things on an even smaller scale has made this strategy the more desired one [18]. In DNA nanotechnology Rothemund was the first to use the self-assembly technique of DNA origami [1].

Choosing DNA as a raw material to construct nanoscale objects instead of other candidates such as proteins has several advantages [21]. There are only four different chemical building blocks that follow the very predictable Watson-Crick base pairing [3]. Another aspect of the DNA helix is that for each complete turn there are 10.5 base pairs, this can be used to calculate crossovers between strands and thus facilitates the different construction. Thanks to these characteristics DNA has proven to be a programmable structure for building DNA nanostructures of any given design [4].
2.3 STEPS IN DNA ORIGAMI DESIGN

The design process of a DNA origami structure goes from using the designers imagination and creating a rough sketch to computer software where the scaffold sequence is inserted and the staple sequences are generated and adjusted to give a stable nanostructure. This is a step by step description of the process in designing a DNA origami structure.

1. The design process starts with deciding on the desired shape of the structure and drawing a geometric model of it. Fill the shape with an integer number of helices and mark out where crossovers are needed to hold the parallel helices in place. The length of the x-axis (parallel to the helices) is easy to estimate because the length of one turn is 3.6 nm. Along the y-axis the helices interact with each other and the distance between two helices appears to depend on the spacing of crossovers. The helices are not close-packed due to electrostatic repulsion [1].

2. Insert the single stranded scaffold into the pattern so that it takes up one of the two strands in every helix. Scaffold crossovers can only be placed where there is a tangent point between helices [4].

3. Data that reflects how the scaffold sequence run through the model is used as an input to the computer aided part of the design process. One turn of the helix corresponds to 10.5 bp but the program uses an integer number of bases between crossovers, 21 bp for 2 turns. The program then helps in the process of generating staple sequences, which are complementary to the scaffold [1]. Staple crossovers must be located at least 5 base pairs away from a scaffold crossover to gain maximum stability in the design. The design leaves no bases unpaired [4].

4. Since the program approximates 21 base pairs for 2 turns, this causes strain in some crossover places depending on the design of the structure. To balance this staples are designed so they alternate directions, if one faces the minor groove the other one faces the major. The twist of the scaffold crossovers are calculated and moved if necessary [1].

5. Finally when folding the scaffold a “seam” usually runs through the design due to the winding way the scaffold is laid out. To stabilize this bridges are created which are staples running across the seem to strengthen the design [1].

2.4 CLASSICAL SEQUENCE SYNTHESIS

2.4.1 OLGONUCLEOTIDES

The standard chemical DNA synthesis was developed in the 1980's and is usually named the phosphoramidite four-step process. This is a cyclical process that elongates a chain of nucleotides by adding nucleotides at the 5' end of the oligonucleotide. In the first cycle of the synthesis an initial protected nucleoside tethered bound to a CPG (controlled pore glass) or a PS (polystyrene) bead with its 3' hydroxyl acts as a template on which the oligonucleotide can grow. The flow of reagents over the template is done stepwise to induce a stepwise addition of bases in the construction of the oligonucleotide and when the sequence is complete the oligonucleotide is cleaved from the solid support.

The most common errors in the synthesis of oligonucleotides are insertions and deletions. Deletions can be as frequent as 0.5 % and insertions 0.4 % per position which results in less than 30-40 % correct sequences when synthesizing a 100 nt oligonucleotide. Purifications with high performance liquid chromatography (HPLC) or polyacrylamide gel electrophoresis (PAGE) can reduce 90 % of these impurities but because these methods use the difference in size when purifying they have no effect against the other types of mutations [5]. The vast majority of the DNA nanostructures that have been characterized depend on synthetic DNA to fold the finished product. This classical chemical synthesis is not able to produce the high-quality and longer sequences that are being requested from the researchers around the world [22].
2.4.2 Genes

The classical DNA synthesis method is typically used when synthesizing oligonucleotides with a length up to 150-200 nt. Longer oligomers up to 300-600 nt can be synthesized but the yield is very low. For synthesis of genes other strategies than the ones used for synthesis of oligonucleotides must be used. Though the two most commonly used gene assembly processes, ligation based and PCR-driven, both use oligonucleotides as a template to produce a gene [5].

2.5 Folding DNA Origami Structures

The two major principles used when folding are the Watson-Crick base pairing and the formation of Holliday junctions [12]. Nowadays the most common scaffold used is derived from the bacteriophage M13mp18 whose genome is single stranded DNA of 7249 nt [2]. While the length of M13mp18 is enough for folding a variety of smaller structures, one future goal is to costume-design origami structures with a mass up to 1 GDa which is 100 times larger than the M13mp18 genome [6]. When folding smaller structures the excess bases are either left at one end of the structure as a tail or staple strands are made to form it into a loop tucked away at one edge to minimize interaction with other structures [1], [23].

The advantage of using a single stranded scaffold is that rehybridization with the complementary strand is impossible and this would improve the yield of DNA origami structures because the only possibility for the scaffold is binding to the staple strands [1]. If a double stranded DNA sequence is used as the scaffold it will rather hybridize with the staple strands as long as there are staples in the folding reaction for both of the scaffold strands. This opens up for the possibility to use longer sequences as scaffolds since it might enable the use of entire genomes of prokaryote or eukaryote cells [24].

The number of base pairs for one complete helix turn can be varied by adding or removing base pairs in the design between crossover positions. The strain given by these modifications gives the structure an overall twist, which when designing a nanotube structure with 11 bp/turn surprisingly gave better folding than the straight design [25].

When folding a structure made from one scaffold strand all the staple strands are designed to completely cover the entire sequence. Folding a polymer structure demands a different use of staple strands. One strategy is to design staples that stick out from the monomer structure and at the opposite end the complementary strand will be left unpaired [4]. A similar way of using this strategy is to first fold the monomers where they all have unpaired scaffold bases at the edges and subsequently add staples that match the two sites desired to bind together [23]. A third way is to separately fold monomers and a scaffold frame both with protruding staple strands that are complementary in such a way that the monomers fit into the frame [26]. The purpose of these strategies is to bind a monomer to another monomer in a controlled fashion. In this way larger structures than can be built without depending on the production of longer scaffold sequences.

Folding is done by first mixing the scaffold with staples. The reactions mix is then rapidly heated to denature the double stranded components. This is followed by a slow cooling to allow hybridization and minimize errors in the folding process. A high GC base pair content in the design can give a higher rate of mispairing during the folding [4]. To further lowering the error rate an excess of staples is always used [1].
2.6 Imaging

Atomic force microscopy (AFM) was the most common method for analyzing DNA origami structures but lately Electron microscopy (EM) has become more popular as well as super-resolution fluorescence microscopy and high-speed AFM.

The problems with creating three-dimensional nanostructures are the lack of structural integrity and strong interactions with surfaces which make them difficult to confirm. Because of these problems mostly compact structures with geometric shape has been characterized. Imaging with AFM uses a charged surface which because of the interaction flattens the structure, this can be avoided by using EM [27].

2.6.1 Atomic Force Microscopy

In AFM a sharp probe is used to scan the surface and the force between the probe and the surface is measured which then gives information about the topography. The concept is much like Scanning tunneling microscopy (STM) but here the sample does not have to be conductive nor is there a requirement of vacuum and low temperatures [28].

2.6.2 Electron Microscopy

EM uses a beam of electrons that are excited in high vacuum from a special cathode using high voltage or heat which scatters or passes through the sample giving a black and white picture of the sample. Magnetic lenses are used to focus the electron beam [29]. There are two major methods in EM that are frequently used. One is Scanning electron microscopy (SEM) where the electron beam is scanned across a solid sample and the back-scattered electrons are detected. This method requires vacuum and conductive samples so the samples are metallized before imaging. In the case with DNA origami the metallization masks the finer details which make SEM not a very good imaging method for intricate nanostructures. The other method is Transmission electron microscopy (TEM) which also requires a special sample preparation and high vacuum but the strength is that it has a high resolution, at best 0.1 nm depending on different factors such as acceleration voltage. In TEM only the electrons transmitted through the sample are detected. The sample has to be thin and for better contrast a sample preparation called staining is used where elements with high atomic number such as uranium salts (uranyl formate) bind electrostatically with the anions like DNA [27].

2.7 Expanding on the DNA Origami Structure

By using the programmable structure the ability to attach other nanoparticles, nanostructures or proteins on the DNA origami design makes it attractive in a multiple of applications. By knowing the exact location and direction of each staple, proteins can be assembled in any desired pattern [30]. One unique possible binding position can be regarded as a pixel. The size of a pixel is the determined by the size of the staples, with shorter staples there are more unique binding sites in the design for e.g. proteins to bind. By marking some staples with biotin or fluorophores these can be considered binary 1's and the unmarked binary 0's [1]. One problem is that a protein attached to the origami structure is bigger than the size of one pixel which is about 6 nm. When placing proteins in a predetermined design they stoichiometricly interact and the resulting pattern may shift from the desired one if their binding sites are close [27].
2.8 FUTURE APPLICATIONS

At this time no applications of DNA origami has made its way out to the market but the research is intensive in many different fields. A few of the possible areas of interest are listed below.

2.8.1 MOLECULAR COMPUTING

Complex calculations in computers are made possible by a large set of binary digital circuits incorporating logic gates. By using DNA origami structures with markers these can be used in these circuits because of their ability to act as binary 1's and 0's [21].

2.8.2 DRUG DELIVERY SYSTEMS

DNA origami structures are also good candidates for target drug delivery because the structure is biocompatible, they can be designed to contain particles, fully encapsulate them and release them when presented with a trigger [31]. The outside of the structure can be coated with proteins and other target signals for transporting to the desired location before releasing the cargo. The disadvantage with this method is that DNA origami may be susceptible to nuclease and due to this risk releasing its cargo too early [6]. A new successful study shows the possibility of using DNA origami structures for delivery and time release of the intercalating drug doxorubicin which induces apoptosis in breast cancer cells [32].

2.8.3 NANO ELECTRONICS

Another possibility is using DNA origami in molecular electronics to build circuit boards by attaching nanowires, carbon nanotubes or metal nanoparticles [1].
3 MATERIAL AND METHOD

3.1 AGAROSE GEL ELECTROPHORESIS

Separation by gel electrophoresis utilizes the differences in molecular size, conformation and net charge which essentially results in molecules with a large cross section area migrate less efficiently through the porous gel than molecules with a smaller cross section area when affected by an applied voltage across the gel [33]. A low voltage for a longer time was preferable when separating longer fragments and for shorter fragments higher voltages was used to minimize the diffusion of the bands. When separation of shorter fragments a higher concentration of agarose (VWR) in the gel was used, which gives a higher resolution due to the decrease in migration rate when the pores were smaller. The ethidium ions moves in the opposite direction of DNA due to its positive charge and intercalates with the DNA. The migration rate of the sample DNA should not exceed that of the ethidium bromide dye (Sigma-Aldrich), if this happened no ethidium bromide was left in the gel to visualize the sample DNA. This was used to image DNA on the gel by exposure to UV light [34]. The choice of voltage, running time and agarose content is displayed in each gel picture under the result section.

When separating folded DNA origami nanostructures with agarose gel electrophoresis additions of 10 mM MgCl$_2$ in both the gel and running buffer were made. The divalent Mg ion causes a change in the conformation of the Holliday junction from an open and loose to a well stacked cross. This promoted a tighter folding of the nanostructures and ensured that minimal changes in the structure conformation took place during the separation [35]. The same conditions were used when separating single stranded DNA >200 nt, shorter fragment were run on a denaturing polyacrylamid gel, using the divalent ions to inhibit hybridization between strands.

The sample volume was mixed with 1x of the Gel loading dye blue 6x (New Enland Biolabs) that contained EDTA to bind excess divalent Mg ions to stop any further enzymatic reactions, SDS to release enzymes bound to DNA, Bromophenol as a tracking dye of the migration rate and the high-density Ficoll 400 which enables the total loading volume to sink into a well. The running buffer used was 0.5x TBE buffer and the final concentration of ethidium bromide in the agarose gel was 0.7 µg/mL [34].

3.2 CREATING THE PSEUDOGENE

3.2.1 DESIGN OF SCAFFOLD AND STAPLES

The scaffold sequence was 378 nt long and was part of a longer sequence, the vector pBluescriptII(+). The vector was screened for an appropriate 378 nt sequence in respect to restriction sites used later on and GC-content, preferably 37-43 %. The exact sequence of the scaffold was not important only that it was known and stable, the choice of using the pBluescriptII(+) as a template for the scaffold was only a matter of convenience. The scaffold used in this project had a GC-content of 43 % and had a size of 117 kDa.

Scaffold: 5’ CCGGGTCAATACGGGATAATACCGGCACATAGCAGAATTTTTAAAAAGTGCTCATCATT
GAAACCGTCTCTGCGGGGGCCAAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACC
CCTCGTGCACCCAAATCTGACATCTCATCGTCTTTACTTACCGACGTTCTCCTGAGCTGGGAACAAAAACAGG
AAGGCAATCTGCACAAAAAAAGGGAAATAAGGGCGACACGAGAATGTGAATACACTCATCTCTCTT
TTCAATATTATTGAGCATTATTACAGGTTATGTCTCTCTATGAGCAGCAGATACATATTTGGATGATTAGA
AAAATAAAACAAATAGGGTTTCCGCACATTTCCCCGAAAAAGTG 3’

Staples were designed to match the chosen scaffold sequence to form a polymer DNA origami structure, described in detail in 3.9.1.
3.2.2 PSEUDogene with ScaFFoLD

To create the pseudogene, which was used for ligation into a vector, the scaffold sequence was elongated with restrictions sites and hairpin structures. The designed primers had a 20 bp recognition for the pBluescriptII(+) and added sites for restriction enzymes HindIII (AAGCTT), Nb.BsrDI (GCAATG) and hairpin structures with restriction site for BseGI (GGATGNN). Creation and amplification of the double stranded pseudogene was done by using PCR as illustrated in Figure 2a. In total the pseudogene sequence was 440 bp (Figure 2b).

Forward primer: 5’ AAGCTTAACATCCGCGGAACGCGGATGTTCCGGCGTCAATACGGGATAA 3’
Reverse primer: 5’ AAGCTTGCAATGTCATCCGCGTGCACTTTTCGGGGAAATGTGC 3’

The PCR was run with a first denaturing step 5 min at 95 °C and then 36 cycles with denaturing 30 s at 95 °C, annealing 30 s and elongation 3 min at 72 °C and at last a final elongation step 5 min at 72 °C. Three PCR with different annealing temperatures at 54 °C, 56 °C and 58 °C to evaluate which temperature give the highest yield. Each reaction volume was 50 µL where final concentrations of the components were 0.3 nM dNTP mix (Thermo Scientific), 1x of the 10x PCR Buffer (New England Biolabs), 1.5 mM MgCl2 (New England Biolabs), 0.5 uM of each primer (Sigma-Aldrich) and 50 ng of the template pBluscriptII(+) (Fermentas). Due to the use of Taq Polymerase (New England Biolabs) in the PCR the product will have an extra adenine in the 3’ end [36].

3.3 LIGATION OF THE PSEUDogene INTO THE VECTOR

The ligation reaction was done with the Cloning kit from Qiagen and the attached protocol was followed. The vector used was the pDrive cloning vector provided in the Cloning kit, which was delivered as a linear double stranded DNA with a single uracil overhang in the 3’ end. This sticky ends were perfect matches for the PCR product which has an extra adenine, this ligation can be done without any pretreatments of the pseudogene.

The vector pDrive is 3.85 kb long and the sequence include genes for resistance against the antibiotics Ampicillin and Kanamycin as well as the f1 origin used in replication of phage genome.

The total volume of the reaction-ligation mixture were 10 µL and the components used were 5 µL 2x Ligation master mix, 1 µL 50 ng/µL pDrive cloing vector and 2 µL PCR product. The ingredients were mixed gently and incubated 30 min at 4 °C.
3.4 TRANSFORMATION

The transformation was done by heat shock using XL-10 Gold ultracompetent cells (Stratagene) and the ligation-reaction mix from previous step, each transformation was done with 30 µL of cells which were treated with 1.2 µL of β-mercaptoethanol (Stratagene). In the positive control the cells were transformed with 1 µL of the vector pUC18 (Stratagene) delivered with the XL10-Gold cells. The negative control was treated with the same conditions as the other cells but no plasmid was added to transform the cells. Two transformations were done with the ligation-reaction mix in different amounts, one with 3 µL specified by the Stratagene protocol and the other with 10 µL.

The heat shock was performed by incubating the cells on ice 30 min then heat to 42 °C for 30 s and then back on ice for 2 min. To each tube with cells 270 µL of preheated 42 °C NZY+ media (Fluka Analytical) was added and the cells were incubated 1 h at 37 °C on shaker at 225-250 rpm.

200 µL of the transformed cells were plated and dilutions 1:1 and 1:20 were made. The cells were plated on freshly made LB-agar (Sigma-Aldrich) plates with 50 µg/mL ampicillin (Sigma-Aldrich) and incubated at 37 °C overnight.

3.5 FRESH OVERNIGHT CULTURE

The fresh overnight culture was an intermediate step with the purpose of amplifying one chosen colony. A single isolated colony was picked from the LB-agar plates with the transformed cells and added into a tube with 3 mL LB media (Sigma-Aldrich) and a final concentration of ampicillin at 100 µg/mL. The culture was incubated over night at 37 °C on shaker at 225-250 rpm.

3.6 SCREENING COLONIES

The colonies from transformation were screened in respect to the inserted pseudogene. Because of the chance of self-ligation of the pDrive the pseudogene was not expected to be found in all of the resulting colonies from the transformation. A couple of colonies were chosen and separate fresh overnight cultures were made as described in 3.5. The plasmid DNA was extracted with the method described in 3.7.1.1.

3.6.1 HindIII SCREENING

The extracted plasmid DNA was then cut with the restriction enzyme HindIII described in 3.7.1.2. Colonies with the insert showed two bands, one longer (3851 bp) for the pDrive vector and one shorter (440 bp) for the pseudogene, when the DNA was separated with gel electrophoresis.

3.6.2 PCR SCREENING

The alternative screening method used the extended primers used when creating the pseudogene. The product from this PCR, executed with the same conditions as described in 3.2.2, was the pseudogene because of the choice of the primers that would only amplify that sequence.


3.7 Amplification of the Pseudo gene

Two different methods for amplifying the pseudo gene were tested as described in 3.7.1 and 3.7.2. In both cases the product was a single stranded DNA sequence.

3.7.1 Amplification Method 1 – In vitro with RCA

The method used a series of manual additions of restriction enzymes, ligase and polymerase in stepwise reactions. In the end RCA is used to amplify the pseudo gene. Before the RCA was used transformation of the pseudo gene to the right format to act as an efficient template was done.

3.7.1.1 Plasmid Extraction

The fresh overnight culture was centrifuged 10 min at 3000 g to collect all the cells in the pellet. The Plasmid Mini Kit 1 (Omega Bio-Tek) was used to extract the plasmid from the cells. Lysis of the cells took place and by precipitation most of the cell material was removed. The plasmid DNA was then separated from the rest by a filter and after a two of washing steps finally eluated.

The concentration of the extracted DNA was measured with a Nanodrop 1000 Spectrophotometer (Thermo Scientific), the eluation buffer was used as blank.

3.7.1.2 HindIII Digestion – Cutting out the pseudo gene from the vector

The extracted plasmid was digested with FastDigest HindIII which cuts out the pseudo gene from the plasmid since the restriction sites for this enzyme were located at both ends of the pseudo gene. The result from this digestion was sticky ends with complementary overhang sequences because the restriction site is a palindromic sequence.

The reaction was carried out in a total volume of 25 µL were 10 µL of Plasmid DNA, 2.5 µL of FastDigest Green 10x buffer (Fermentas), 0.5 µL FastDigest HindIII (Fermentas) and incubated 1 h at 37 °C and 10 min at 80 °C to inactivate the enzyme after the digestion.

3.7.1.3 Ligation into Circular Pseudo gene

Ligation of the sticky ends resulted in circular DNA where the pseudo gene made up the entire circle.

To the 25 µL of digested DNA from previous step an additional 8 µL of Rapid Ligation buffer (Fermentas) and 2 µL T4 Ligase (Fermentas) was added and the total volume was corrected to 40 µL. The mixture was incubated 10 min at 22 °C and 10 min at 65 °C to inactivate the enzyme.

3.7.1.4 Nb.BsrDI Digestion - Nicking the Pseudo gene

To create a template for RCA the complementary strand to the scaffold was kept circular and the other strand was nicked, one base was cut out, with the restriction enzyme Nb.BsrDI.

The entire preparation, 40 µL, from previous step was used and the final volume after additions in this step was 70 µL. To the Ligated DNA an extra 7 µL of NEB2 10x buffer (New England Biolabs), 3.5 µL Nb BsrDI (New Enland Biolabs) was added and incubated 2 h at 65 °C and then inactivated 25 min at 80 °C.
3.7.1.5 RCA

The product from the RCA was a single stranded continuous repeat of the complementary strand of the circular template. The Phi29 and Bst DNA polymerase used the nicked strand as a template and displaced this strand as it moved around the closed circle (Figure 3).

An alternative was instead of having the nicked circular DNA, is to use the circular DNA before nicking as a template. To enable amplification by RCA a primer was designed that annealed with the complementary strand of the scaffold which then acted as a template for the DNA polymerase (Figure 3).

RCA primer: 5' ACCGCACATAGCAGAAC 3'

Both the nicked templates as well as the closed circular template with the primer were used in two different sizes, the full phagemid and the smaller circularized pseudogene.

The total reaction volume when using the nicked template was 100 µL including 40 µL of nicked circular DNA, 10 µL Phi29 10x buffer (Fermentas) or ThermoPol Reaction buffer (New England Biolabs), 10 µL of 10 mM each dNTP mix (Thermo Scientific), 2 µL of 10 mg/mL BSA (New England Biolabs).

When using the primer an annealing step was done before the RCA. To the 40 µL of circular DNA an equal molarity of the primer was added as well as 10 µL Phi29 10x buffer or ThermoPol Reaction buffer and finally the volume was corrected to 83 µL. Annealing was done by stepwise cooling 5 min at 94 °C, 10 min at 60 °C, 10 min at 45 °C and finally 10 min at 37 °C [37]. After the annealing 10 µL of 10 mM each dNTP mix and 2 µL of 10 mg/mL BSA was added.

All amplification was done by adding 50 U Phi29 or 50 U Bst DNA polymerase incubation for 16 h at 30 °C or 65 °C.
3.7.2 Amplification Method 2 – In vivo with amplification by Helper Phage

This method utilized the phases own ability to do RCA inside the infected cell. This is the natural way that phages with single stranded DNA genome amplify [38], [39]. A helper phage is a phage that lacks the f1 origin and is unable to replicate by itself. Though its genome contains all the vital information to create new phage particles, the one thing that’s missing is the single stranded DNA to be packaged in it.

The plasmid used to transform the cells contains the f1 origin, it was therefore considered to be a phagemid, and thus could be amplified and packed into phages when the cells were infected with a helper phage [40].

3.7.2.1 Culture

To amplify the cells further a larger culture was made. 600 µL of the fresh overnight culture was added into a 1000 mL flask containing 100 mL 2x YT media (MP Biomedicals), 500 µL 1 M filtered MgCl$_2$ (VWR) and 100 µg/mL ampicillin (Sigma-Aldrich). The culture was incubated at 37 °C on shaker 200 rpm until OD$_{600}$ = 0.45. The culture was inoculated with a multiplicity of infection of 20, VCSM13 helper phages (Stratagene) per cell, and left over night on the shaker at 37 °C.

3.7.2.2 Helper Phage Extraction

The cells were separated from the helper phages by centrifugation 20 min 2200 g at 4 °C, done twice for higher purity. The cells form the pellet and the phages stayed in the supernatant. Precipitation of the phages was done by adding 4 g PEG8000 and 3 g NaCl to separate them from the media. The solution was incubated at most 5 min at RT and then placed on ice for at least 30 min. Centrifugation 30 min 10000 g at 4 °C resulted in the phages forming the pellet. The pellet corresponding to 100 mL of culture was resuspended in 2 mL of 10 mM Tris pH 8.5 and a last centrifugation 5 min 3000 g at 4 °C was done to pellet down any remaining cells. The supernatant was transferred to a fresh tube [41].

3.7.2.3 DNA Extraction

The capsid proteins were stripped by using 4 mL PPB2 (0.2 M NaOH, 1 % SDS) which unravelled the proteins at which point they were unable to keep the capsid formation. After this 3 mL PPB3 (3 M KOAc, pH 5.5) was used as a neutralization buffer. After incubated 10 min on ice the solution was centrifuged for 30 min 16500 g to collect the proteins in the pellet. The supernatant was mixed with at least 2 volumes 99.5 % ethanol (Kemetyl) and incubated 50 min on ice and the sample was centrifuged for 30 min at 16500 g. The DNA forms the pellet and two washing steps with 75 % ethanol (Kemetyl) were done. The pellet was resuspended in 340 µL 10 mM Tris pH 8.5 [41].

3.7.2.4 Phenol-Chloroform Extraction and Ethanol Precipitation

The phenol-chloroform extraction was done to reduce the amount of residue proteins from the previous DNA extraction. To this solution 1 volume of phenol-chloroform (AppliChem) was added and subsequently vortexed and centrifuged 3 min at maximum velocity to separate the two phases, the lower organic phase which dissolved the proteins and the upper aqueous phase which dissolved the DNA. The aqueous phase was transferred to a new tube and the steps were repeated an additional two times.

In the ethanol precipitation firstly 1/10 volume 3 M sodium acetate was added to the retrieved volume from the previous step. After this 3 volumes of 100 % ethanol was added and incubated 30 min at -20 °C before centrifuged 15 min at maximum velocity. The precipitated DNA forms the pellet which was washed twice by addition of 1 mL 75 % ethanol, mixed gently and centrifuged in the same manner as before. The final pellet was air dried and resuspended in 10 mM Tris pH 8.5 buffer [42].
3.8 **BseGI Digestion – Cutting out the Scaffold**

In this step the template was single stranded with exception only to the hairpin structures where the double stranded restriction site for BseGI was placed. The execution of the digestion of the product from the RCA and the helper phage amplification were identical.

To 50 µL of the RCA product 10 µL Tango 10x buffer (Fermentas) and 5 µL of the restriction enzyme BseGI (Fermentas) was added and the total reaction volume adjusted to 100 µL. The digestion was done 24 h at 55 °C.

3.9 **DNA Origami Structures**

3.9.1 **Designs**

The DNA origami structure was designed to form polymers, because of the small size of the monomer, to facilitate the confirmation of proper folding by TEM. In the designs of the structures two different ways of stacking the helices were used, the square (SQ) and the honeycomb (HC) lattice (Figure 4a-c). The three structures used were the 9 helices square bundle (9H-SQ), the 18 helices honeycomb bundle (18H-HC) and the 17 helices twisted honeycomb bundle (17H-HCTw). The overall twist in the last structure was made by inserting an extra base pair between crossovers for each two full turns of the helix. This shifts the amount of base pairs from 10.5/turn to 11/turn which creates a strain that decreases by the polymer changing from straight to twisting around its own axis.

To form polymers each monomer had a number of protruding strands, and at the opposite edge of each protruding strand the complementary strand was found on the scaffold with no internal staple binding to it. This way the monomers behave similar to blocks of LEGO fitting together perfectly (Figure 4d). For each structure two lengths of the protruding strands were designed; 8 and 4 bp for the 9H-SQ and 4 and 2 bp for the structures 18H-HC and 17H-HCTw.

![Figure 4](image-url)
3.9.2 FOLDING
Scaffold and staples were mixed and folded during an elaborate stepwise cooling repeated in cycles for a total time of 16 h.

The folding was done with a total volume of 20 µL with a final concentration of 10 mM MgCl₂, 10 mM Tris, 5 mM EDTA. The scaffold and each staple for each specific design were added at a molar ratio of 1:10.

3.10 IMAGING THE STRUCTURES

3.10.1 SAMPLE PREPARATION
The sample solutions were individually immobilized on a carbon film coated mesh copper grid FCF400-CU (Electron Microscopy Sciences). Preparations of the grids were done by glow discharge which essentially uses low pressure and high voltage to create a positively charged and hydrophilic layer on the grid [43].

To each grid 5 µL of sample solution was added for 3 min which was then removed by drying it off on a filter paper. The surface of the grid was stained for 20 s with 40 µL of 20 mg/mL uranyl formate (Electron Microscopy Sciences), 0.02 M NaOH and finally the grid was dried in the same manner as the previous step.

3.10.2 TEM
Pictures were taken with magnifications between 22000x and 44000x with the TEM FEI Morgani 268 using a voltage of 80 kV.
4 RESULTS

4.1 PSEUDOGENE

Extension of the scaffold by the extended primers was successful at all tested annealing temperatures (54, 56 and 58 °C). The agarose gel in Figure 5 showed a single band positioned between the 400 and 500 bp bands of the DNA ladder which was the appropriate migration of the linear double stranded 440 bp pseudogene. A slight increase in yield when 56 °C was used can be suggested from the brightness of the band compared to the ones of 54 °C and 58 °C.

![Figure 5](image)

**FIGURE 5:** The PCR products were separated on a 1 % agarose gel run at 120 V for 1 h. All three reactions produced bands at the same length which corresponds to the expected 440 bp. The only difference in the PCR set-up was the annealing temperature which was used to name the lanes. The middle lane shows a slight increase of intensity indicating a higher concentration of DNA.

4.2 PHAGEMID

To check if the ligation of the phagemid was successful a separation of a 1 % agarose gel at 120 V for 1.5 h was done. The size of the pDrive vector and the phagemid are quite similar due to the fact that the pseudogene was small in size. The migration rates were the same and no conclusion about the insert efficiency could be drawn because of the resolution not being sufficient.

4.3 TRANSFORMATION

Transformation with XL10-Gold ultracompetent cells by heat chock resulted in approximately 2300 cfu/mL when transformed with the lower amount of phagemid and 3800 cfu/mL when transformed with excess amount. The cfu-counts were done on the 1:20 dilutions and no duplicates were made which makes the calculations a rough estimate. Though it was enough for this experiment to confirm successful transformation.

Before the transformation was accomplished a number of factors were investigated. The most crucial factor seemed to be the use of cells with high viability and competence. The choice of antibiotic and its concentration was also important as well as the usage of fresh LB-agar plates. Another factor was the human error, methods such as heat shock needed practice to be completed successfully.

4.4 SCREENING OF COLONIES

Both screening with HindIII digestion to cut out the pseudogene from the phagemid and screening by PCR using the extended primers to amplify just the pseudogene were completed. Only the HindIII digestion was reliable since it turned out that due to a possible contamination of the extended primers the PCR screening gave a lot of false positive results.
4.4.1 **HindIII Digestion**

By utilizing the fact that restriction sites for HindIII was located at both ends of the pseudogene, the HindIII digestion was used as a screening method. The digested DNA from a colony with the insert showed two bands when separated on an agarose gel, 3851 bp and 440 bp which corresponded to the lengths of the pDrive vector and the pseudogene.

15 colonies were screened and only one colony, number 5, showed positive results for the inserted pseudogene. The PCR product, which was the pseudogene, was run as a control for the migration of the pseudogene. All digested samples were placed in the well on the right of the undigested sample (phagemid) from the same colony. In samples from all colonies there was a difference in the migration of the undigested phagemid compared to the digested phagemid which were slower. Even though no insert was found, these results were consistent. This change in migration rate reflected the change in conformation from circular to linear, which influenced the cross section area, due to the fact that the pDrive vector itself had one restriction site for HindIII.

There was a distinct difference between samples in how the pairs of sample DNA ran in the gel (Figure 6). They can be divided into two subgroups where colony sample 1, 2, 3, 4, 7, 8, 10, 11, 12, 13 was called group 1 and sample 6, 9, 14, 15 was called group 2 for future references. In group 1 the samples ran as discussed above where the pDrive self-ligated and the pseudogene is most likely not present in these colonies. In group 2 both the samples in a pair ran slower than group 1 thus indicating a longer sequence.

Sequencing over the insert area was done with colony 5 and 14 were only the results from colony 5 was positive for the pseudogene. This showed the correct pseudogene sequence except a point mutation C>T at position 387, which was compensated by changing the corresponding nucleotide in the relevant staple for each design.

![FIGURE 6: HindIII screening of 15 colonies (number 1 to 15) on a 1 % agarose gel run at 120 V for 1 h. The sample loading was normalized to 50 ng DNA in each well. For each colony two samples were loaded, in the well to the left the extracted phagemid and to the right the HindIII digested phagemid was run. The PCR, annealing temperature 56 °C, created pseudogene was run as a reference to identify the migration rate of the pseudogene.](image)
4.4.2 PCR
The alternative to screening by HindIII digestion was to use the primers designed for elongating the scaffold. These extended primers would then amplify only the pseudogene and one band at this length would show on the agarose gel. Though unfortunately even the negative control showed a band at this length (Figure 7) which made it impossible to distinguish the true positive from the false positive results.

4.5 Amplification by RCA

4.5.1 Phi29 DNA polymerase
The results from the RCA-tests with the Phi29 DNA polymerase all showed the same trend in different extents, the yield was much lower than expected, by comparing rough calculations based on gel images. The band at barely 400 bp was the double stranded scaffold and the band migrating further at about 250 bp was probably the single stranded scaffold (Figure 10c). In the cases where only the upper band was visible indicates poor amplification of the template. Though in all tests, no matter of the factors tested, the RCA product stuck to the well which indicated a very long DNA fragment and good amplification. The digested RCA product gave bands as described indicating a large amount of double stranded DNA, which can only originate from the template itself. The result was then contradictive when considering that the single stranded RCA product was a long fragment and the digestion of this product indicated a high amount of double stranded scaffold.

There was one instance when the RCA primer was used in combination with the annealing program run before the RCA where the amplification was much higher. These settings were run with the template primer concentrations 1:1 and 1:10, where the 1:1 was the most successful (Figure 8a). The digested RCA product did not give the bands expected. When using the large template the expected band on the gel would correspond to the 378 nt scaffold, 23 nt hairpin and 3.85 knt vector and finally in small amounts the residue double stranded template. When the small template was used the expected bands were the 378 nt scaffold, 23 nt hairpin and a small amount the residue double stranded template. The template giving the best amplification determined by the gel in was the small closed circle. There was no clear band for the single stranded scaffold obtained after digestion but rather a smear reaching down to the expected migration length. One probable explanation to the unexpected longer bands and the smear in all lanes could be incomplete digestion of the RCA product.
4.5.2 **Bst DNA Polymerase**

The bands given by the digested RCA product when the DNA polymerase was switched from the Phi29 to Bst is shown in the Figure 8b and c. These gel images are of the same gel but with different exposure times to enhance the very strong band and then the very weak ones. The reactions were all done with the same concentrations of the compounds with the only change being the total reaction increasing from 20 µL to 100 µL. In smaller scale all templates except the large closed circle with added RCA primer shows an overloaded fast migrating band. In the larger scale reactions multiple band are showing corresponding to the single strand residue vector or incomplete digestion. The only sample with a band at the length of the single stranded scaffold was given by the large nicked circle template.

![Figure 8](image)

**FIGURE 8:** Abbreviations used in all three gel pictures: S=Small circle, L=Large circle, N=Nicked, P=Primer 1:1 and P+=Primer 1:10. Samples were separated on a 1.5 % agarose gel with 10 mM MgCl₂ run at 70 V for 4 h. a) Each set of conditions were displayed in two lanes, the one to the left was the RCA product when using the Phi29 DNA polymerase and to the left its BseGI digest. Amplification using the template and primers in 1:1 and running the annealing program before RCA was started gave the highest amount of DNA. The digestion may not have been complete and no clear band corresponding to the single stranded scaffold was found, though it might be present in the smear at shorter band lengths. b) and c) were taken on the same gel but with different exposure times in trying to visualize both the weak and the strong bands. The pictures show only the BseGI digestion of RCA run with the Bst DNA polymerase. The first four sample lanes were from reactions with a total volume of 20 µL and the last four from reactions with a total volume of 100 µL, the ratio of the components were the same.

4.6 **Amplification by Helper Phages**

Results from only one of the five amplifications by the helper phage method had a slight positive result. After phenol-chloroform extraction two bands at about 1 and 2 kb became visible, before the extraction this DNA was probably located amongst the aggregates stuck in the well. The migration of the lower band suggested that it could possibly correspond to the approximately 4300 nt long single stranded DNA which was the phage genome derived from the phagemid. The upper band could possibly contain the about 7000 nt long single stranded genome of the helper phage itself. Though the expected yield from the amplification was much higher
than actually produced. The broad vague bands in all lanes that were located below 500 bp were most likely RNA residues from the DNA extraction.

After digestion with BseGI was done, the single stranded pseudogene was expected to be visible in the gel (Figure 9), but no such band could be found. The digestion was done with two different volumes of the Crude P-C, 50 % template as described in 3.8.1 and 25 %, though the total reaction volume was constant. This was done to investigate if the lower salt concentrations in the 25 % template reaction would give a higher degree of digestion. Though in all cases no digestion product could be identified.

Since the experiment was done with negative results a couple of times a few factors were varied with the purpose of increasing the rate of success. Incubation time of the culture after inoculation with the helper phages were tested at 3 h, 5 h and 17 h (overnight). The best result was given with 5 h incubation. At 3 h the production of phages with genome from the phagemid might be too low because of the lack of time to produce phage particles but when increasing the incubation time to 17 h there might be too much cells in the culture which then contributed to more protein and DNA contamination of the end product.

Another factor was the ratio of PPB2 and PPB3 which in the first two rounds of the amplification was too low. This error originated from unoptimized protocols which was corrected to the specifications described in 3.7.2.3.

Though the most important factor of them all seemed to be the usage of viable helper phages with a correct titer. If using a helper phage batch which have been thawed and refrozen, calculated titer would differ from the viable titer and this would effectively lower the multiplicity of infection.

4.7 FOLDING

Folding of the three DNA origami structures, with the two different lengths of protruding strands each, was done firstly focusing on using the scaffold produced by RCA using the Phi29 DNA polymerase and the nicked small circularized pseudogene as the template. In each occasion the result showed undefined irregular short fragments with a dimension of 2-10 nm wide and 50-200 nm long. The expected polymers should have been straight and of considerable length which indicated that the scaffold was poorly folded, if folded at all. When comparing the folded scaffold with the scaffold without any added staples no significant differences was found (see Figure 10a). The shapes in the picture could be double stranded scaffold or pseudogene which was expected to have a size of about 125x2 and 145x2 nm. Other probable objects that might be proteins, e.g. the
66 kDa BSA added in the RCA compared to the 117 kDa scaffold, and enzymes from all the reaction steps the create the pseudogene.

The folded samples were also compared with the scaffold by separation on an agarose gel. The only band visible was located at barely 400 bp which corresponded to the double stranded scaffold. The lane for the scaffold showed two bands, one at barely 400 bp and one at about 250 bp (Figure 10c). The upper band was probably double stranded scaffold and the much weaker lower band single stranded scaffold. This lower band was not present in the lanes with folded structures but if some structures were folded from this the concentration would have been too low to confirm (Figure 10b).

The scaffold produced from the small and the large closed circular templates, with the scaffold primer concentration 1:1 and annealing of the primers before the RCA with the Phi29 DNA polymerase, were both used to attempt to fold the polymer structure without success. The TEM pictures were similar to those of the folded scaffold produced from the small nicked circle with Phi29 DNA polymerase.

**FIGURE 10:** a) TEM pictures of each folded structure and the scaffold without additions of staple strands as a control, scale bars are 100 nm long. There was no polymers found in either of the samples and the samples all looked like the scaffold control which indicates no folding. b) The nanostructures separated on a 2 % agarose gel with 10 mM MgCl₂ at 70 V for 3 h. Number 0 was the scaffold and 1=9H-SQ-8bp, 2=9H-SQ-4bp, 3=18H-HC-4bp, 4=18H-HC-2bp, 5=17H-HCTw-4bp, 6=17H-HCTw-2bp. All samples had the same migration rate of that expected from the double stranded scaffold. Even though barely visible in the lane of the scaffold there was a faster band at about 250 bp which after folding disappeared. c) The BseGI was the digested RCA product (Phi29, nicked small circular template) which then was used for the folding. 0=BseGI= scaffold.
5 Discussion

The main goal of the project to successfully construct an all-in-one production of the DNA building blocks to fold a DNA origami structure was not completed. Due to unforeseen hinders the second phase, the scaffold and staples production, was never started. The production of enzymatic scaffold and staples in the way described in this project is highly likely to be successful taken in consideration a few changes and some more time spent to optimize the protocols further.

The aim of creating a cell line with transformed cells containing the pseudogene could then be considered partially successful. Cells were transformed with the phagemid containing the pseudogene, however, the direction in which the pseudogene was inserted was not direction specific. When the designs were made for the pseudogene and the staples for the DNA origami structures, a mistake was made. Not considering the non-directional cloning vector used the pseudogene could be inserted in both directions. Unfortunately in the one colony found with the insert it was placed in the wrong way, in consideration to the direction of the f1 origin. When staples for the DNA origami structures were designed only the position of the nicking site was considered. This resulted in only the scaffold produced from the amplification by RCA was the one matching the staples. The scaffold produced from the amplification by helper phages was the complementary scaffold and not usable for folding. Further screening would most likely reveal a colony with the pseudogene inserted in the correct direction. Though in this case the aim of producing the scaffold sequence with the helper phage strategy became impossible due to the direction of the f1 origin.

Timesaving improvements to the transformation step can be done by making sure to use highly viable and competent cells. One property of the XL10-Gold ultracompetent cells and the pDrive vector that was not exploited was the possibility to perform blue-white colony screening. In comparison to the screening methods used in this project, the HindIII digestion or the PCR amplification where in both cases the products needs time consuming separation with gel electrophoresis, morphological techniques such as blue-white screening offers a quicker detection of colonies with the pseudogene insert.

RCA was applied to amplify the sequence with small degree of success, however the technique requires further optimization to increase yield and purity. Initially there was a persistent problem with the RCA where after amplification the DNA content was too low to give visible bands on any agarose gel. A decision was made to increase the volume of sample DNA into each reaction. After this change RCA products were detected in agarose gels, which previously was not the case. Though this change increased salt concentrations in each reaction step, which might inhibit enzymatic activity. A test run with desalting samples before the RCA was done only to lose the DNA content in the last step. This approach could possibly increase the amplification if preformed successfully which could not be performed due to time constraints within the limits for this thesis.

The scaffold production from the RCA had low success rate and very low yield. Confirmation by gel electrophoresis indicated the presence of a RCA product in all tests, but only after the DNA sample volume had been adjusted. Though after digestion with BseGI only a very low amount of possible single stranded scaffold could be detected, and then only in a few instances. This resulted in poor folding, if no folding at all. The samples seemed to contain a high amount of double stranded and a very low amount of single stranded scaffold. This is counterintuitive to the confirmed long RCA product which is always single stranded.

Increasing the RCA reaction volume surprisingly had a strong negative effect on the amplification. In small scale reactions scaffold production seemed promising, except when using the large unnicked circle which may be due to strong secondary structure [44]. Despite increasing the reaction volume, keeping the ratio of the components constant, the same result could not be obtained. Though these larger reaction volumes have been successfully used in other preparations. Considering the small reaction volume especially the Bst DNA polymerase generated possible high yield single stranded scaffold which would be interesting for further testing since this was just discovered just before project closing.
The aim to fold 3 different polymer DNA origami structures was not achieved due to the problems encountered in earlier aims. While pilot experiments were performed they did not produce positive results.

A few thoughts come to mind, regarding possible future improvements to the project. It might have been overambitious to try to produce such a long sequence without experience generating intermediate lengths of enzymatic oligonucleotides.

Alternatively, it could be productive adopting the strategy of assembling large DNA nanotubes by using tiles constructed of only 5 oligonucleotides with lengths ranging from 26 to 42 nt described by Rothemund et al., 2004 [45] and Ekani-Nkodo et al., 2004 [46].

It would also be preferable to construct an experiment where the enzymatically produced oligonucleotides could be compared to synthetic ones. This both enables test folding to check if the sequences gives a well-defined polymer structure and the sequences could also be used as a reference of migration rate when using gel electrophoresis. Though one of the challenges in focus when designing this project was to explore if longer sequences could be produced with the methods specified.
BIBLIOGRAPHY


