Assembly and intracellular delivery methods for nucleic acid nanotechnology

Catherine Hua Fan
St Peter's College
University of Oxford



Thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy at the University of Oxford

Michaelmas Term, 2023

ABSTRACT

The field of nucleic acid nanotechnology has mainly taken DNA and RNA out of their biological context and uses their properties to revolutionise the control of molecular self-assembly; but can we bring this technology back into the cell? Since their inception, nucleic acid nanostructures have developed into almost any shape imaginable, but one application area for these constructs that has not been explored as much — despite the advantages in size, biocompatibility and the programmability of these constructs— is at the interface with biology. Challenges preventing the breakthrough of this field into more intracellular applications involve improving the cellular uptake of the assembled constructs and, once inside the cell, protecting these nanostructures from nuclease degradation or unfolding. Overcoming these issues with new intracellular delivery methods or self-assembly designs that allow for *de novo* construction within the cell would open up a plethora of applications in diagnostics, therapeutics, drug delivery, as well as providing programmable tools for studying molecular biology within the cell itself.

This thesis investigates three approaches for designing intracellular nucleic acid nanostructures. The first strategy presented is a new isothermal single-stranded assembly method, inspired by knitting, for designing nucleic acid nanostructures inside the cell. I show the characterisation of both a single-knit and multi-knit structure, through both computational modelling as well as experimental analysis. The resulting knitted constructs were primarily focused on *in vitro* replicable DNA systems to allow for ease in characterisation, although I also delve into the concept as a co-transcriptionally folded RNA method. Using atomic force microscopy, I was able to visualise the rolling circle amplified multi-knitted constructs. Taking inspiration from knitting, this assembly technique has the potential to be used for further complex designs and patterns. I then explore alternative approaches towards bringing nucleic acid nanostructures into the cell. This includes the use of confinement as a new method for delivering asymmetrical DNA origami cryoelectron tomography tags into mammalian cells through the creation of transient pores in the membrane to allow for the cellular uptake and targeting of the DNA nanostructures. Finally, I demonstrate three variations of bistable DNA hairpins to show reconfiguration kinetics that could be used in the design of future intracellular nucleic acid nanostructures. With these three projects, I was able to demonstrate a variety of *in vivo* implementation of DNA nanostructures, which can inspire future exploration of this promising field of intracellular nucleic acid nanostructures.

Acknowledgements

First and foremost, I would like to thank my supervisor, Professor Andrew Turberfield, whose guidance and scientific knowledge were paramount to the development of the projects in this thesis. From helping me develop the novel concept of DNA knitting, to supporting me in establishing a new tissue culture set up for the lab, Andrew has allowed me a good measure of independence throughout the past four years, while providing the necessary input when needed. Equally, I would like to thank Dr. Jonathan Bath, whose reassuring presence in the lab and expertise in a vast realm of fields was immensely valuable to my growth as a Ph.D student.

I would also like to thank the many other members of the Turberfield group, past and present. In particular, Dr. Erik Benson and Dr. Joakim Bohlin were essential to guiding me through the usage of oxDNA to model my nucleic acid nanostructures. Dr. Antonio Garcia Guerra and Dr. Rafael Carrascosa Marzo also offered a plethora of advice in optimising my experimentations. Dr. Seham Helmi, Alma, Qian and the rest of the group members are owed many thanks for making the lab space an enjoyable environment. A special thanks goes out to Sanna Piippo-Henderson for keeping the lab clean and well-stocked.

I would like to thank Dr. Emma Silvester and Dr. Lindsay Baker, who I collaborated with for my intracellular delivery experiments. Emma provided me with her DNA origami signposts and walked-me through the process of assembling the structures. Dr. Zaw Win from the Baum group was also helpful in these experiments for showing me the 4Dcell dynamic cell confiners as well as the homemade versions, and working with me to obtain preliminary results.

As for the others who have trained me on various equipment, I would like to thank Dr. Casey Adam from the Contera group for teaching me how to use the Cypher AFM and Dr. Tom Catley from the Pyne group for guiding me through the use of the Bruker Fastscan-XR AFM and TopoStats AFM processing software. I would like to thank the Royce Partner Equipment Access Scheme for allowing me to travel and work in Sheffield with Dr. Alice Pyne and her group.

Special thanks goes out to my amazing friends, Piyush, Sanja, Surya, the Hotmamis (Anna, Luna, Elisa, Selma, Brooke and Sophie), my fellow Peterites (especially Stephanos for showing me how to submit my thesis) and many others who have made my time at Oxford one of the best experiences of my life! My heart goes out to them all.

I would also like to acknowledge the remarkable fact that this thesis has travelled through six different countries (United Kingdom, Italy, Thailand, Malaysia, Singapore and Turkey) throughout its writing process.

Finally, I would like to thank my parents, Ying and Yuan, and my brother Andrew, who have always supported me in all my endeavours, no matter how outlandish or over-the-top. In particular, I would like to thank my mum for the constant reminders to upload my thesis before the deadline. Mum, I did it!

Table of Contents

1 Introduction 8

1.1 Scope of Thesis 8

1.2 Nucleic Acids and Nucleic Acid Nanotechnology 9

1.2.1 The structure and biophysical characteristics of nucleic acids 9

1.2.2 Programming the shape of DNA nanostructures 11

1.2.3 Single-stranded DNA and RNA assembly 13

1.2.4 Assembling dynamic nucleic acid nanostructures 16

1.3 Bridging the Gap between Cell-Free and Intracellular Nucleic Acid Nanostructures 18

1.3.1 Nucleic acid nanotechnology inside the cell and in whole organisms 18

1.3.2 Intracellular and in vivo delivery of nucleic acid nanostructures 21

1.3.3 Challenges of current intracellular delivery methods 23

1.3.4 Nucleic acid nanofabrication inside the cell 24

1.4 Summary and Thesis Structure 25

2 Materials and Methods 27

2.1 NUPACK 27

2.2 Kinefold and coFold 27

2.3 oxDNA and oxView 28

2.4 Nucleic Acid Nanostructure Preparation 29

2.5 Gel Electrophoresis 30

2.6 Microscopy 31

2.7 Tissue Culture and Cell Confinement 31

3 Designing Nucleic Acids to Assemble in a Knitted Fashion 33

3.1 Introduction 33

3.2 Author Contributions 33

3.3 Designing a Single-Stranded Nucleic Acid Nanostructure 33

3.4 Knitting 35

3.4.1 Stitches and knit patterns 35

3.4.2 Mathematics of knitting 35

3.4.3 Translating knitting into nucleic acid assembly 36
3.5 Designing a Single DNA Knit Stitch 37
3.5.1 Initial design 37
3.5.2 Optimising looping with oxDNA simulations 39
3.6 Designing a Multi-knit DNA Nanostructure 42
3.6.1 Stacked and conjoining hairpins 42
3.6.2 Unfolding dynamics of a simplified Plus-Minus multi-knit structure 42
3.6.3 Designing a multi-knit DNA structure to self-assembly through RCA 45
3.7 Designing a Single RNA Knit Stitch 52
3.7.1 Converting the DNA single-knit structure to RNA 52
3.7.2 Converting the DNA multi-knit structure to RNA 54
3.8 Chapter Summary and Outlook 54
4 Analysing the DNA Knitted Structures 56
4.1 Introduction 56
4.2 Author Contributions 56
4.3 Analysing the Single DNA Knit Stitch Structure 56
4.3.1 Controls for the HTL structure 56
4.3.2 Relieving the well aggregation 59
4.3.2 PAGE analysis of knitted and non-knitted reactions 60
4.4 Analysing the Multi-knit DNA Structure 62
4.4.1 Analysing the multi-strand multi-knit DNA structure by PAGE 62
4.4.2 Analysing the single-stranded multi-knit DNA structure by PAGE 64
4.4.3 Analysing the single-stranded multi-knit DNA structure by AFM 65
4.5 Improving the AFM Resolution of the Multi-knit Nanostructures 66
4.5.1 Optimising the adsorption of the nanostructures to the mica 66
4.5.2 Imaging the multi-knit structure with specialised equipment and software 67
4.5.3 Analysis of the template strands 68
4.5.4 Analysis of the single knit hairpin 70
4.5.5 Analysis of the multi-knit structures 71
4.5.6 Potential discussions about looping analysis 72
4.6 Chapter Summary and Outlook 73

5 **Intracellular Delivery of DNA Origami CryoEM Tags** **75**
5.1 Introduction 75
5.2 Author Contributions 75
5.3 Confinement Chambers 75
5.3.1 Using confinement as an intracellular delivery method 75
5.3.2 Commercial 4Dcell confiners 76
5.3.3 Homemade cell confiners 78
5.4 Utilising DNA Origami cryoEM Tags for Intracellular Delivery 78
5.4.1 DNA origami signposts as cryoelectron tomography tags 78
5.4.2 Tagging intracellular proteins with DNA signposts 79
5.5 Optimising Confinement for Intracellular Delivery 79
5.5.1 Experimental design 79
5.5.2 Finding the optimal confinement height and pressure 82
5.6 Intracellular delivery and tagging of DNA signposts 83
5.6.1 Tagging an intracellular membrane protein 83
5.6.2 Issues encountered and improvements required 84
5.7 Chapter Summary and Outlook 85
6 **Bistable DNA Hairpins as a Reconfigurable Mechanism** **86**
6.1 Introduction 86
6.2 Author Contributions 86
6.3 Using NUPACK to design my DNA hairpins 86
6.3.1 Introduction to NUPACK 86
6.3.3 Designing my DNA hairpins in NUPACK 87
6.3.3 Comparing NUPACK 3.0 and NUPACK 4.0 90
6.4 Analysing the Bistable Hairpin Designs 92
6.4.1 Reaction parameters with each hairpin design 92
6.4.2 Short stem hairpins 92
6.4.3 Long stem hairpins 93
6.4.4 Extended long stem hairpins 94
6.4.5 Mirrored extended long stem hairpins 96
6.5 Chapter Summary and Outlook 98

7 **Conclusion & Future Research Directions** **100**
7.1 Summary and Discussion 100
7.2 Outlook and Directions for Future Research 102
7.2.1 Larger multi-knit nanostructures 102
7.2.2 Modifications to the sequences 104
7.2.3 RNA Knitting 105
7.2.4 Simulations of these nucleic acid knitted structures 106
8 **References** **107**
9 **Appendix** **118**
A Sequences 118
A.1 Chapter 3 and 4 118
A.2 Chapter 6 123
A.3 Primer sequences 126
B Supporting Figures 127
B.1 Chapter 3 127
B.2 Chapter 4 130
B.3 Chapter 5 137
B.3 Chapter 6 138

Chapter 1

Introduction

In this Chapter, I will explain the scope of this thesis and discuss the essential background literature. I will introduce the idea of using nucleic acids as programmable, self-assembled nanostructures, and then explore the relevant advancements made with this technology, briefly discussing the various assembly techniques, and focusing on single-stranded assembly which is the subject of my work presented in Chapters 3 and 4. Finally I will discuss methods towards bringing nucleic acid nanostructures into the cell, focusing on intracellular delivery and *in vivo* folding.

1.1 Scope of Thesis

In this thesis, I present three approaches towards bringing nucleic acid nanostructures into the cell. Most DNA nanostructures are created using *in vitro* systems as it is easier to control the assembly parameters as well to characterise the structures. A downside to *in vitro* assembly is that it is then difficult to introduce these complex structures into the cell, limiting their biological and medical applications. Therefore, new strategies for *in vivo* assembly or intracellular delivery methods are necessary to advance the field.

DNA and RNA have remarkable molecular recognition properties and structural features that make them promising templates to pattern materials with nanoscale precision. Their binding specificity and thermodynamic stability, nearly infinite choice of sequences that bind reliably to their complementary partners, rapid synthesis through automated methods and capabilities to be acted upon by a large variety of enzymes allow for an almost unlimited creation of addressable nanostructures in one, two and three dimensions. As there already exists a vast library of nucleic acid nanostructures assembled and characterised *in vitro*, creation of a reliable and effective intracellular delivery method would allow effective exploration of the application of these constructs in a biological context. Alternatively, a whole new assembly approach can be designed to allow for *de novo* construction, which carries the advantages of having fewer components, as well as using the cell's own machinery for continuous amplification. These concepts are a treasure chest of untapped potential as the study of nucleic acid nanotechnology, particularly within cells, is still at a very early stage of discovery and development.

This thesis is divided into three projects. For Project 1, I will show the design process for creating a single-looped and multi-looped DNA knitted nanostructure as well as go through the experimental analysis of these structures. For Project 2, I will show a new tool for delivering DNA origami into mammalian cells. I will explain the use of a commercial and homemade cell confiner, and go through the experimental design. I will then go through the preliminary experiments showing promising intracellular delivery. For Project 3, I will show a trio of bistable hairpins and discuss their reconfigurable dynamics and discuss how the concept can be used to bring nucleic acid nanostructures into the cell.

1.2 Nucleic Acids and Nucleic Acid Nanotechnology

1.2.1 The structure and biophysical characteristics of nucleic acids

Nucleic acids are central biomolecules that form complex polymers responsible for the development, functioning and reproduction of all known living organisms. These nucleic acids are classified as either deoxyribonucleic acids (DNA) or ribonucleic acids (RNA), and although they share many similar properties, they have chemical differences that result in their distinct biological roles, as well as provide unique characteristics when used as a biomaterial. As shown in Figure 1.1A, nucleic acids are made up of smaller subunits called nucleotides. Each constituent nucleotide is made up of a pentose sugar, either a ribose for RNA or a deoxyribose that is lacking a hydroxyl group for DNA. The sugar molecule is bound to a hydrophobic nitrogenous base called a nucleobase on the 1′ carbon, forming a nucleoside. Each nucleoside is attached to a phosphate group on the 5′ carbon, through covalent phosphodiester bonds. This creates a repeating sugar-phosphate backbone, where consecutive monomers are held together by two ester bonds between phosphate groups. This specificity of binding locations creates the 5′-3′ asymmetric directionality of nucleic acids.

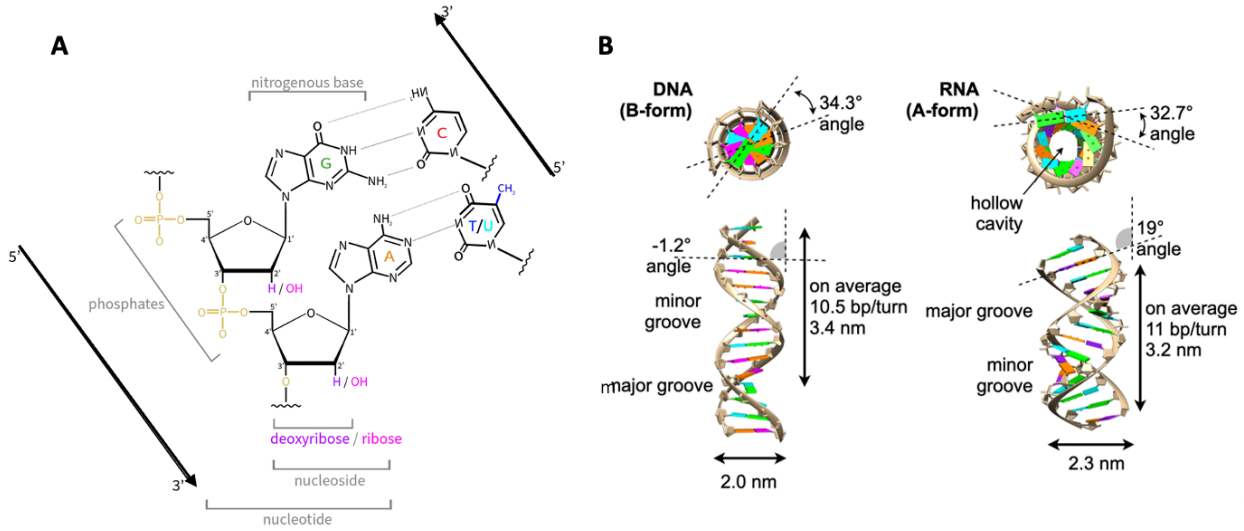


Figure 1.1: The structure of DNA and RNA. A) Chemical structure of the alternating sugar-phosphate backbone in nucleic acids, with the Watson-Crick base-pairing between relevant nitrogenous bases. Phosphate groups are coloured in brown and are connected to wavy lines indicating continuation of the backbone pattern. The difference in hydroxyl groups between DNA and RNA is highlighted in purple and pink, respectively. The 5' to 3' directionality is indicated with the arrows. B) Structural differences between B-form DNA (left) and A-form RNA (right) double helices, as taken from [1].

Alongside a distinct sugar component in the backbone, DNA and RNA also have different sets of nitrogenous bases essential to their structure and functions. They share three nucleobases, namely adenine (A), cytosine (C) and guanine (G), and differ as thymine (T) in DNA is replaced by uracil (U) in RNA, which lacks the methyl group on the 5-carbon. Nucleobases are further categorised

as pyrimidines (C, T/U) which form a single six-membered ring, and purines (A, G) which are composed of a six-membered ring fused to a five-membered ring. Due to their multiple hydrogen donors and acceptors, pyrimidines and purines can interact via hydrogen bonds in specific arrangements, known as Watson-Crick base pairing [2] as shown in Figure 1.1A, where G pairs with C, forming three hydrogen bonds and A pairs with T/U, forming two hydrogen bonds. A unique feature to RNA is that G and U can form a wobble-pair, which has comparable thermodynamic stability to Watson-Crick base-pairs and is nearly isomorphic [3, 4].

The ability to form complementary base pairs serves as the foundation of nucleic acid hybridisation, resulting in unique structural conformations. Whereas RNA features highly varied and often complex three-dimensional structures to serves its many biological functions, DNA is most commonly seen in a double-helical structure. First described in 1953 by Watson, Crick, Wilkins, Franklin and Gosling [2, 5, 6], the DNA double helical structure contains nucleobases hydrogen bonded in an anti-parallel orientation with one strand reading from 5′ to 3′ and the other from 3′ to 5′. This produces a double stranded structure with the negatively charged sugar-phosphate backbone on the periphery of the helix and the hydrophobic nitrogenous bases stacking on each other in the interior channel. Alongside specific base pairing interactions, the nitrogenous bases within a DNA or RNA strand are oriented in a way to allow for optimal stacking interactions. This results in a π-π stacking of the hydrophobic aromatic rings within consecutive nucleobases along the helical axis and contributes to the stability of the overall structure [7]. In order to maximise the π-π overlap and the Van der Waals interactions, nucleobases adopt rotated and tilted arrangements. This asymmetrical disposition of the backbone creates spaces referred to as ‘grooves’, categorised as the wider and deeper “major groove” and the narrower and shallower “minor groove”.

Base-pairing and base-stacking interactions are slightly stronger in RNA compared to DNA [8, 9], and alongside sugar puckering [10] result in structural differences in the helices as shown in Figure 1.1B. DNA prefers the B-form conformation while RNA strictly adopts the ribbon-like A-form, with both conformations forming right-handed double helices. In B-form DNA, the nucleobases are mostly centred around the helical axis, whereas in A-form RNA they are considerably displaced towards the major groove of the helix leaving a more open cylindrical core. The additional hydroxyl group in the ribose sugar of RNA also induces a change in its endo conformation, where the 2′ or 3′ carbon of the sugar is displaced upward out of the plane, resulting in a change from C2′-endo seen in B-form DNA to C3′-endo in A-form RNA [10]. This difference in twisting of the sugar molecule results in narrower and taller helices in B-form DNA than in A-form RNA, with a helical diameter of 2 nm compared to 2.3 nm, respectively, and an average distance between consecutive base pairs of 0.34 nm compared to 0.28-0.29 nm, respectively. The helical pitches (height per turn) are also different as B-form DNA has a pitch of 3.4 nm equivalent to about 10.5 base pairs, while A-form RNA has a helical pitch of only 2.8 nm resulting in a turn every 11 bp [11].

These structural changes caused by the additional hydroxyl group in the ribose sugar of the RNA backbone increases the number of possible geometries classified based on the observation that hydrogen bonding interactions between RNA bases exist between ‘edges’ of the two contributing bases: the Watson-Crick edge, the Hoogsteen edge or the sugar edge [12]. Bases can also interact in one of two orientations relative to the glycosidic bonds giving rise to 12 possible interaction geometries which result in at least two hydrogen bonds. Only six of these orientations are even

theoretically possible with DNA. Non-Watson-Crick base pair arrangements are frequently found in alternative secondary RNA structures and non-helical RNA motifs. These non-canonical motifs include isomorphous A-T and G-C^+ Hoogsteen base pairs, as well as homo-purine G-G and A-A pairs that are constituents of base triplets in triplexes, the C-C^+ base pair comprising three hydrogen bonds in the intercalated tetrastranded i-motif, and the G-U wobble pair is found in t-RNA. With RNA strands adopting such complex and varied three-dimensional architectures, predicting the 3D structure of a RNA strand from its sequence poses a significant challenge.

Nucleic acids are flexible molecules that can undergo structural changes and adopt different conformations based on environmental conditions and interactions with other molecules. This flexibility is essential for various biological processes, but also makes nucleic acids an ideal biomaterial for creating complex nanostructures. Specific environmental conditions such as pH, ionic strength, type of cations, molecular crowding and binding partners like proteins or low-molecular-weight metabolites contribute to the pathway of nucleic acid folding within the cell, but can also be used to manipulate the folding *in vitro* to create rational designs on the nanoscale. Additionally, the rigidity of these nanostructures can be programmed in based on the amount of single-stranded and or double-stranded regions. The persistence length (l_p), which is the decorrelation length of the direction of the polymer chain and determines the bending stiffness of a polymer, in single-stranded DNA and RNA are low with values of l_p around 1-2 nm allowing for regions of flexibility, whereas doubled-stranded nucleic acids are considerably stronger with l_p values of 50 nm for dsDNA and 62-64 nm for dsRNA. The folding and unfolding of these structures can also be tuned as nucleic acids can denature at melting temperature and anneal together when cooled. With the wide range of stable base pairs beyond the typical Watson-Crick base pairing contributing to numerous structural motifs found for DNA and RNA, nucleic acids can be programmed into a variety of different secondary and tertiary geometries. Furthermore, nucleic acids can be synthesised cheaply and with a wide range of modifications and can be acted on by various enzymes. With this vast tool-kit of both biological and synthetic origins, DNA nanofabrication has great potential for tailored-made medical and technological applications.

1.2.2 Programming the shape of DNA nanostructures

Nadrian Seeman's initial idea of DNA as a programmable material came to life in 1980 as he was sitting in his campus pub, reflecting on Dutch artist M. C. Escher's woodcut art-piece "Depth" (Figure 1.2A), depicting topologically isomorphous fish-like flying objects organised periodically in three-dimensions. When envisioning DNA patterned in a similar fashion (Figure 1.2B), Seeman realized that the branched junctions of the DNA strands could be connected together by the predictable base pairing interactions of single-stranded overhangs, called sticky ends, resulting in deliberately assembled 3D crystalline materials. A two-dimensional crystal was achieved in 1998 [13] inspiring a new generation of DNA nanotechnology. Many of these early DNA nanostructures were flat arrays of repeated units of naturally occurring DNA motifs as shown in Figure 1.2C [13, 14]. A common motif used was the Holliday junction, a four-way junction comprising four strands that is naturally formed with the cell during homologous recombination [15]. In particular, these synthetic Holliday junctions were rendered migrationally immobile by making arms with unique sequences of oligomeric nucleic acids allowing for growth into larger quadrilateral shapes [16]. In the following years, methods to self-assemble discrete geometrical objects using sequential

multi-step strategies were developed to create a three-dimensional DNA cube [17], a DNA tetrahedron assembled from four oligonucleotides [18], a DNA octahedron [19] and endless other complex architectures.

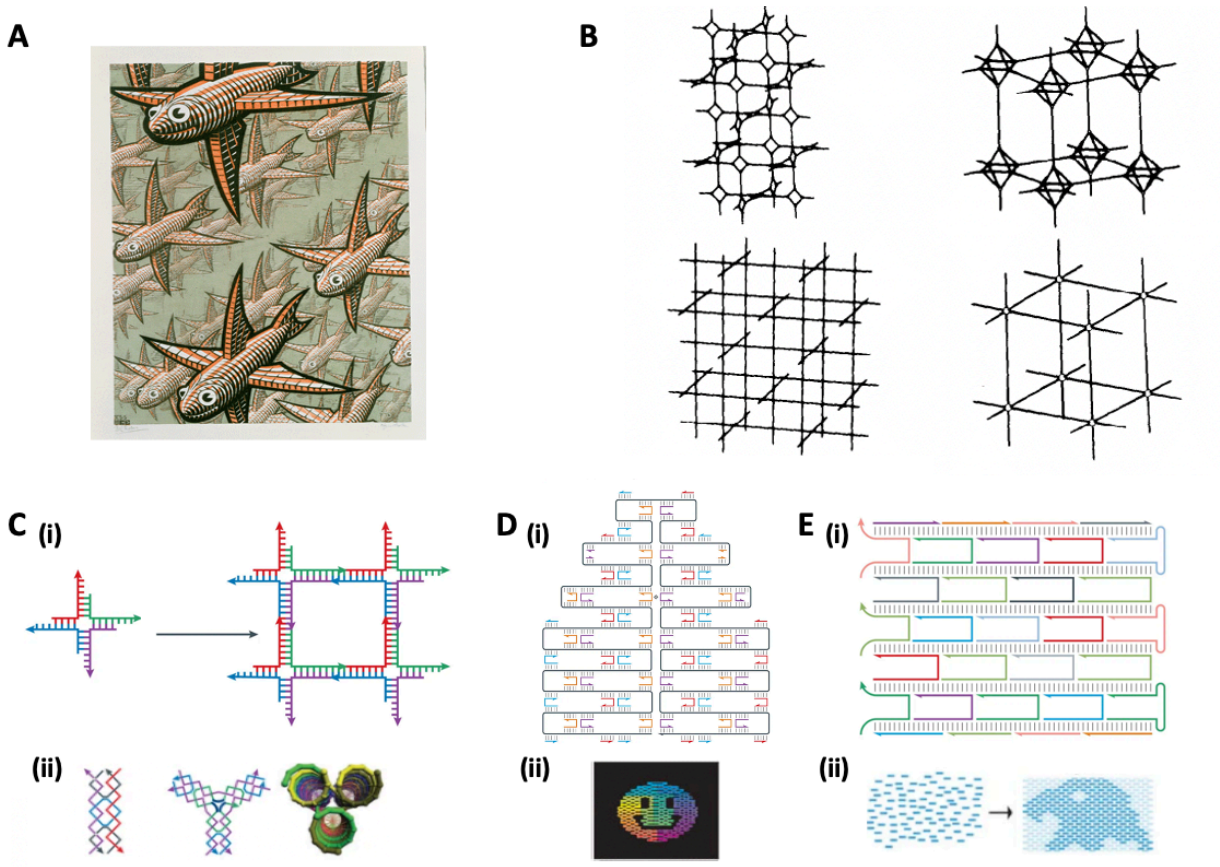


Figure 1.2: The initial concept of DNA nanotechnology and an overview of DNA assembly techniques. A) "Depth" a woodcut art piece by M.C. Escher, 1955 (c) The M.C. Escher Company - The Netherlands. All rights reserved. https://mcescher.com/. B) N. Seeman's initial depictions of N-connected DNA networks, as adapted from [16]*. In these stereoscopic projections, the dark lines represent double helical stretches, the circles represent junction regions, and the short lines on the periphery of each figure represent unsatisfied valences. C-E show three different DNA assembly techniques: C) DNA tile assembly* [20]*, D) DNA origami* [21]*, and E) DNA brick assembly* [20]*. In these three schematics, the coloured arrows represent unique oligonucleotides that bind together to form a motif in C), bind a large scaffold strand (black) in D), and anneal together in E) to form a folded DNA nanostructure. Example schematics of the structures built with each technique are depicted in part (ii).*

Soon after the use of the Holliday junction motifs came the double-crossover (DX) molecules comprising two DNA double helices linked together by two strand exchanges [22] resulting in more geometrically rigid and stable DNA nanostructures. The rational design of these DNA motifs to form a bigger lattice is the basis of DNA tile assembly (Figure 1.2C). Soon a toolbox of DNA tiles was created with motifs including the discussed Holliday junction (HJ) tiles [23] and double-

crossover (DX) tiles [13, 24], as well as paranemic-crossover (PX) tiles, multi-arm junction tiles, and DNA tetrahedron tiles [25] were used to create increasingly complex arbitrary shapes and patterns.

In 2006, the field of DNA nanotechnology was transformed by Paul Rothemund as he proposed a whole new DNA construction method termed DNA origami [21]. This new assembly technique, shown in Figure 1.2D uses a long scaffold strand, normally derived from single-stranded DNA from a viral genome, that is raster filled into a predesigned pattern and held together by short 'staple' oligonucleotide strands. DNA origami has a relatively simplified protocol as the staple strands normally do not need to be purified and are combined with the scaffold strand all together in a 'one-pot' reaction, relying on a thermal ramp to assemble the desired structure. Initially planned out by hand, the DNA origami assembly method has been combined with open source computer interfaces, such as caDNAno [26], to streamline the design process. This revolutionary assembly method has been used to make a broad range of structures ranging from the two-dimensional, such as the DNA smiley face [21], as well as a plethora of complex three-dimensional structures [21], and has improved the complexity of DNA nanostructure design as one could now construct up to 200 addressable points in an area of 8,000-10,000 nm^2.

With DNA origami, each distinct shape requires a new scaffold routing design and synthesis of new staple strands. Therefore, a multi-stranded DNA 'brick 'assembly was developed without the need for a long scaffolding strand as seen in Figure 1.2E [27]. Building off of the modular assembly of two-dimensional DNA tiles, DNA brick assembly works by using standardised small components mediated strictly by local binding interactions. Each canonical DNA brick is a 32 nucleotide single strand with four 8 nucleotide binding domains that can form interconnected staggered duplexes with one another in a stepwise assembly. Additionally, as this method doesn't require a biologically derived strand, each component can be synthetically modified allowing for greater chemical diversity. This assembly method is useful in making discrete asymmetrical structures, where the DNA bricks act like attachable children's blocks to make complex three-dimensional designs [28, 27].

1.2.3 Single-stranded DNA and RNA assembly

A distinct limitation to the DNA assembly approaches discussed so far, is their need for hundreds of DNA strands of different sequences to produce a single nanostructure. The costs to produce such large libraries of strands is not necessary a restriction as the cost of oligonucleotide synthesis continues to decrease [29], but the limiting factor is the environment these multi-strand assemblies are formed in. Many of these complexes require a multi-step thermal ramp for denaturing and annealing the strands, specific salt concentrations and accurate stoichiometries, therefore limiting them to precise *in vitro* reactions. Additionally, the use of multi-components increases the chance of defects through errors in the synthesis process or during experimentation. These defects lower the local addressability and resolution of the structure when used as a material with an expected nanometer precision. Attempts have been made to reduce the number of strands involved in the assembly method, such as a DNA octahedron formed by a single 1.7 kb strand in the presence of only five 40-mer oligonucleotides [30]. To reduce the components in nucleic acid assemblies to a

single long strand is a challenge that requires the balance of complex functionality and topological simplicity to avoid kinetic traps imposed by knots.

In the pursuit of unimolecular assembles, relatively simple DNA structures were assembled to create single-stranded 79 nucleotide (nt) four-arm junctions [31], 160 nt paranemic crossovers [32], and a 286 nt DNA tetrahedron [33]. Although simple in design, a unique feature to these constructs is that they can be replicated through rolling circle amplification (RCA). RCA is an isothermal nucleic acid replication process that can rapidly synthesise multiple copies of circular DNA or RNA as shown in Figure 1.3. This amplification process utilises high processivity polymerases such as Phi29, Bst3 and Vent *exo*-DNA polymerases and T7 RNA polymerase that bind to a primer strand annealed to the circular template (Figure 1.3A) and continuously adds nucleotides to create periodic daughter strands, displacing the daughter strand ahead of itself as it moves around the circle (Figure 1.3B) [34]. RCA reactions can be run at constant temperature — around 37 °C for most of the polymerases mentioned — in solution or in complex biological environments. Therefore RCA allows for the isothermal replication of long single-stranded DNA or RNA that can then self-assemble into their designed unimolecular nanostructures.

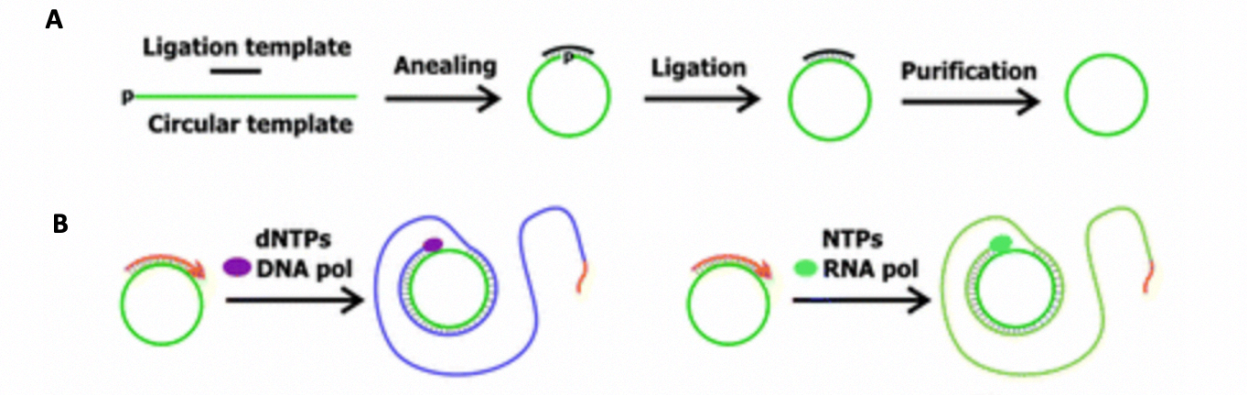


Figure 1.3: Overview of rolling circle amplification. This figure is adapted from [34]. *A) Circularisation of a RCA template (green) through the annealing and ligating of a primer strand (black) to the linear template strand. B) With the addition of dNTPs or NTPs and a high processivity polymerase, the continuous amplification of the circularised RCA template can generate long ssDNA (left) or RNA (right), respectively.*

As DNA origami assembly progressed, it inspired the 2014 conception of single-stranded RNA origami [35]. To form the RNA lattice, a combination of programmable motifs, such as hairpin and kissing loop interactions (KLs) as shown in Figure 1.4A, hold the single stranded RNA in its desired shape. Through the utilisation of the unique tertiary interactions of RNA, a new crossover pattern called a 'dovetail seam' was created where 2 to 3 base pairs cross the seam at each helix to create arrays of coaxially stacked helices that are both sequence-specific and non-topologically linked. These initial RNA origami structures could be assembled either by heat-annealing over mica or by *in vitro* co-transcriptional folding, and could be up to 660 nt in length. Further research into the conformational landscape of several RNA origami [36] has led to kilobase-sized single-stranded RNA origami nanoscaffolds [37].

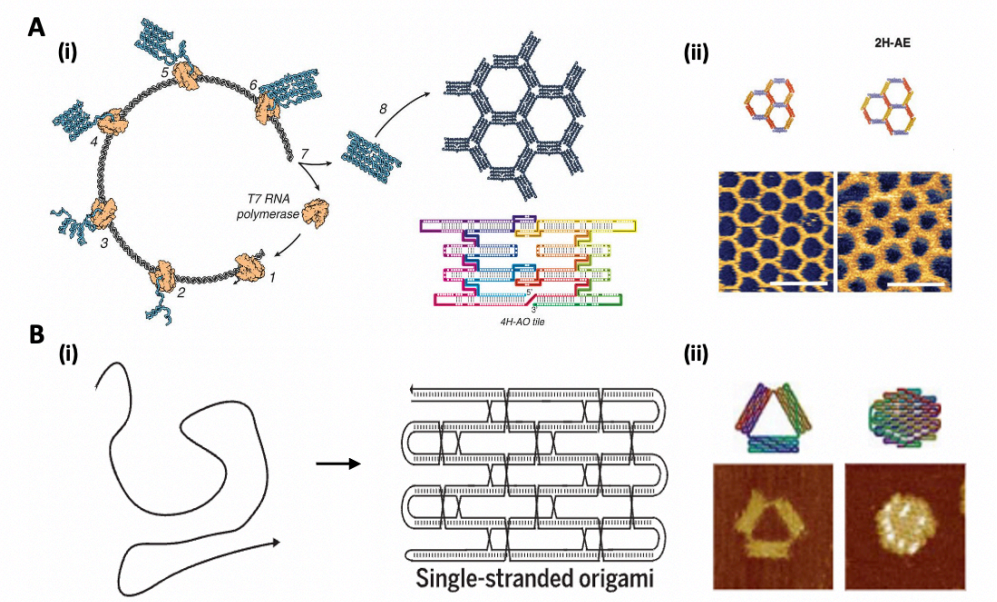


Figure 1.4: Design pathways for single-stranded DNA or RNA assembly. A) Schematic of RNA origami as adapted from [35]. *(i) Shows the co-transcriptional folding process of a 4H-AO tile and with a zoomed in representation of the kissing loop interactions. (ii) AFM images of the mica-annealed 2H-AO and 2H-AE tiles. B) Schematic of ssDNA origami as adapted from* [38]. *(i) Shows the folding process of the DNA strand and (ii) shows the AFM images of a ssDNA triangle and ssDNA smiley face.*

With the success of ssRNA origami, the development of ssDNA origami arose in 2017 [38]. In this technique as depicted in Figure 1.4B, partially complemented double-stranded DNA and parallel crossover cohesions are used to construct a structurally complex yet knot-free structure that can be folded smoothly from a single strand of DNA up to 10 kb. These structures were densely compact and could be programed to form protruding hairpins or loops to attach other moieties. The unimolecular nature of these ssDNA origami structures allowed for the *in vitro* or *in vivo* amplification of the DNA template through rolling circle amplification or intracellular transcription of a transfected DNA template. The replicated template would still require a purification step and also requires the use of an external thermocycler to run a program for annealing the final structure.

The ability to fold *de novo* designed nucleic acid nanostructures is more suited to unimolecular folding instead of multistrand assembly. Single-stranded folding is independent of reactant concentrations therefore offering higher assembly yields and more robust folding kinetics than multi-stranded structures, which rely on concentration-dependent intermolecular self-assembly. Additionally, with less components, this process can be scaled up with high replication fidelity as cells have various machineries and pathways for replication. In particular, the mutation rate of *E. coli* DNA replication machinery is approximately four orders of magnitude lower than the highest-fidelity PCR polymerase [39]. This exploration of co-transcriptionally folded RNA assemblies, particularly within the cell, will continue in Section 1.3.4 discussing "Nucleic acid nanofabrication inside the cell".

1.2.4 Assembling dynamic nucleic acid nanostructures

As such an interdisciplinary field, nucleic acid nanotechnology benefited from a coalition of techniques. A shift towards a more automated design process made it easier to move from constructing static two-dimensional constructs to complex three-dimensional and dynamic designs [40]. The development of new microscopes, particularly with atomic force microscopy and transmission electron microscopy, also allowed for accurate confirmation of the folded structures. These are just a few of ways the growing interest in molecular self-assembly has pushed for higher complexity and functionality of nucleic acid nanostructures.

Whereas structural nanotechnology can serve as templates to arrange inorganic or biomolecules for electronics and photonic application of biomimetic functions, the development of dynamic nanostructures allowed for the creation of autonomous molecular motors. The field of reconfigurable DNA structures was jumpstarted in 2000 with the concept of using isothermal DNA strand displacement to open and close a DNA 'tweezer' [41]. Figure 1.5 depicts various nanostructures utilising a toe-hold mediated strand displacement (TMSD) reaction. TMSD allows for the exchange of one nucleic acid strand with another, as an invader strand binds to an overhang of a strand in a double-stranded complex and displaces an incumbent strand via a branch migration process. In addition to the sequence of the strand, the strand length, pH and ionic conditions can affect the kinetics of displacement [42]. The kinetics of TMSD are investigated in Chapter 6 with the design and analysis of various bistable hairpins.

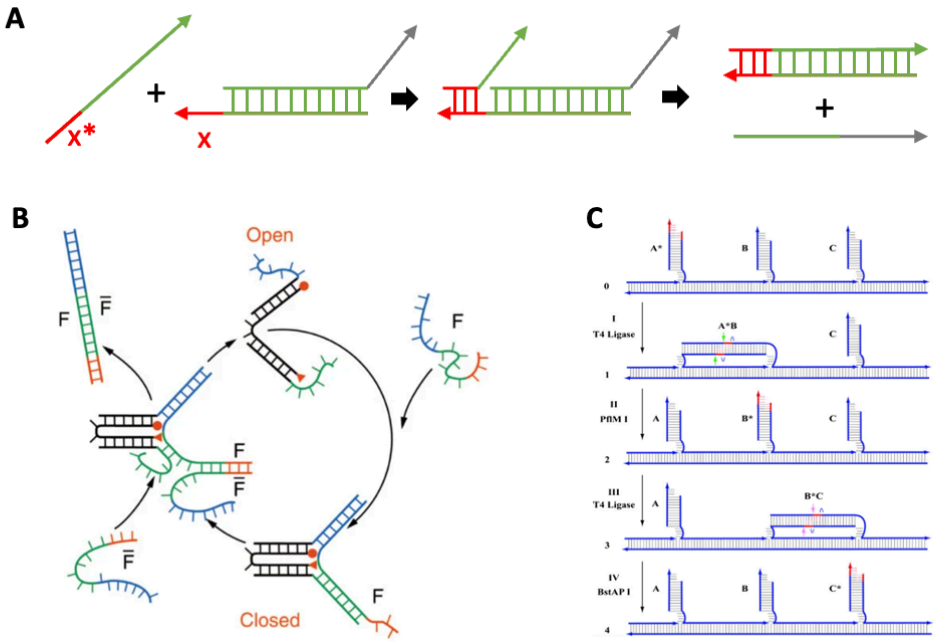


Figure 1.5: Toe-hold mediated strand displacement (TMSD) and its use in DNA nanotechnology. A) A schematic of DNA toehold exchange: a green duplex is shown with a toehold sequence (x) in red, which can be bound by an invader strand with the complementary sequence (x). As the invader strand is bound to the original duplex, it will start to displace the top strand through branch migration, resulting in a new duplex and displaced strand. B-C show examples of TMSD in different reconfigurable structures: a DNA tweezer* [41] *and a DNA walker* [42].

Figure 1.5B in particular shows the use of an unpaired 'toehold' extension on the nanostructure that can bind to the complete complement on the free 'fuel' strand and displace the existing motif on the nanotweezers to trigger a reconfigurable change. TMSD has led to other dynamic DNA nanostructures such as DNA walkers (Figure 1.5C) that can autonomously move along a molecular track [43, 44], as well as structurally switchable 3D DNA assemblies [45]. The usage of fluorophores, particularly, Förster resonance energy transfer (FRET) pairs [46, 47] also allows for more accurate analysis of these displacement reactions.

41Alongside strand displacement, a plethora of techniques and design add-ons have been used to create dynamic DNA nanodevices as highlighted in Figure 1.6. Aptamers are single-stranded oligonucleotides that fold into defined architectures to bind target proteins, making them an obvious choice in motif to add to a design to increase its complexity. Aptamers can be used in various ways, such a through a split aptamer system that enables fluorescence in the presence of a complementary sequence [48], or they can be combined with their appropriate ligand on the same structure to create a lock-and-key system as seen in Figure 1.6A on the DNA origami box to create a controllable lid [49]. Another approach shown in Figure 1.6B is the ion-mediated hybridization of multiple short single-stranded DNA strands. Here, a cation-responsive DNA origami device demonstrates millisecond-scale responsiveness due the change in ion levels allowing for the quick reconfiguration of the system [50]. Figure 1.6C shows the use of the photoisomerization molecule azobenzene, on a DNA origami nanoscissor to photoactivate the opening and closing of the structure in response to irradiation with either ultraviolet or visible light [51]. Furthermore, the DNA origami-based printer in Figure 1.6D is another example of an autonomous molecular robot, which can pattern incoming DNA strands in two dimensions [52]. The final device shown in Figure 1.6E is a DNA rotary ratchet motor, which uses the directional motion achieved within the framework of Brownian ratchets [53].

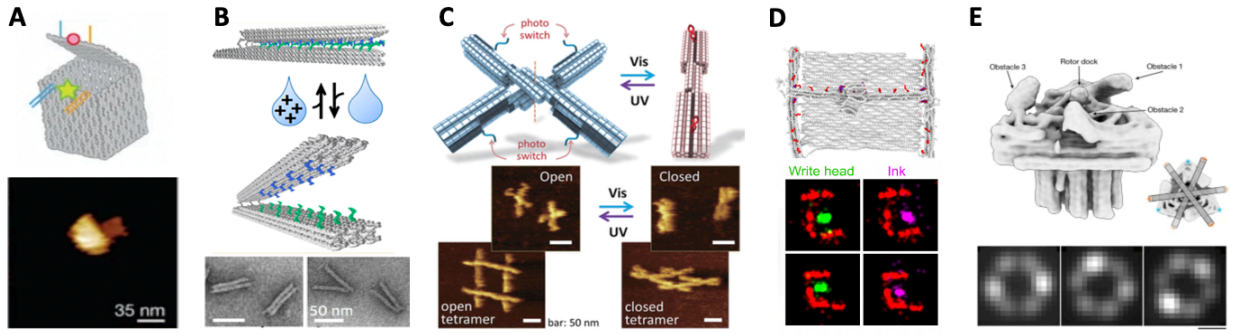


Figure 1.6: Overview of complex and reconfigurable nucleic acid devices. 2140*A) DNA origami box with controllable lid* [49]*, B) ion-activated reconfigurable DNA nanostructure* [50]*, C) photoactivated reconfigurable DNA scissors* [51]*, D) DNA origami printer* [52] *and E) DNA rotary ratchet motor* [53]*. Each design has a corresponding microscopy image with AFM images provided for A and C, TEM images for B, DNA-PAINT images provided for D, and single particle fluorescence images given for E.*

So far, I have mainly discussed the development of DNA nanostructures, as the field is far more advanced than its RNA counterpart. With the folding pathway of RNA harder to predict than with DNA, the design of RNA nanostructures has been quite limited. This increased library of potential

motifs can provide complex secondary and tertiary geometries with RNA [54]. The first synthetically programmed RNA based nanostructures were introduced in 2004 as akin to programmable jigsaw pieces [55], then came the organisation of intracellular reactions with rationally designed RNA assemblies [56], and ultimately the previously brought up method of RNA origami [35]. Although the use of RNA nanostructures so far has been scarce, these nanostructures hold the great potential to be genetically expressed and folded entirely within a cell, opening up a plethora of new biological applications.

1.3 Bridging the Gap between Cell-Free and Intracellular Nucleic Acid Nanostructures

1.3.1 Nucleic acid nanotechnology inside the cell and in whole organisms

Based primarily in cell-free systems, the field of nucleic acid nanotechnology has been able to scale up design complexity and allow for quantitative characterisation of reaction mechanisms to an extent that is infeasible for engineered gene circuits or other cell-based technologies. A caveat to this *in vitro* approach is the limitations to adopting biological applications for many of these nanodevices. These nanostructures have been carefully designed to fold and exist in a specific environment, involving fixed ion concentrations and temperatures, that might not match the physiological conditions within the cell. Additionally, they were designed in reaction pathways without the interference from DNA and RNA processing enzymes and other confounding factors that might be encountered in cells. In test tube you can precisely control the concentration of all components and reaction kinetics can be monitored with high time resolution. Molecular crowding and cellular compartmentalisation can adversely affect the performance of multi-stage, multi-input logic circuits and other systems with many interacting components. The diffusion coefficient of synthetic DNA molecules in the cytoplasm is 5-100 times smaller than in water [57] and the rates of hybridisation between complementary single-stranded nucleic acids are different in the cellular environment than in an aqueous buffer [58]. Therefore, much work needs to be done to bridge the gap between folding these nucleic acid nanostructures in well-mixed reaction buffers to assembling them in spatially structured, densely packed cellular environments.

By transitioning nucleic acid nanotechnology from the test tube to the cell, this field has the potential to impart important biophysical techniques that facilitates single-molecule observations in biology and programmable dynamics in response to specific biological cues. The ability to program DNA and RNA to any size, shape and ligand patterning can be advantageous to create novel intracellular devices as represented in Figure 1.7. In particular, nucleic acids serve as an ideal drug delivery vehicle as it is biocompatible and can be modified with addressable molecules, such as aptamers and ligands, capable of targeting a specific diseased environment. Additionally, you can fine tune the size and shape of these nanostructures as well as the placement of the molecules on the carrier to increase their stability and effectiveness as drug carriers [59]. Various DNA cages have been developed as drug delivery tools, such as a nucleolin-targeted and telomerase responsive DNA nanotube depicted in Figure 1.7A [60] as well as DNA origami with intercalating antitumour drugs showing circumvention of drug resistant cells [61]. In particular, the DNA tetrahedron is one of the most widely used DNA nanostructures for biomedicine, due to

their superior cellular uptake and biodistribution to targeted cells. DNA tetrahedra have been shown to penetrate skin [62] as well as pass the blood brain barrier [63–65]. An example tetrahedron drug carrier is shown in Figure 1.7B. As carriers for siRNA, DNA tetrahedron can improve circulation time in the blood compared to the lone parent siRNA [66]. Similar to the DNA tetrahedron, triangle-shaped DNA origami also exhibit optimal tumor passive targeting accumulation, showing that discrete geometries are important in the delivery of these nanostructures [67]. In addition to these singular constructs, multi-component nanostructures can also be designed therapeutics, such as the aptamer-tethered DNA 'nanotrain' for targeted transport of molecular drugs in cells [68].

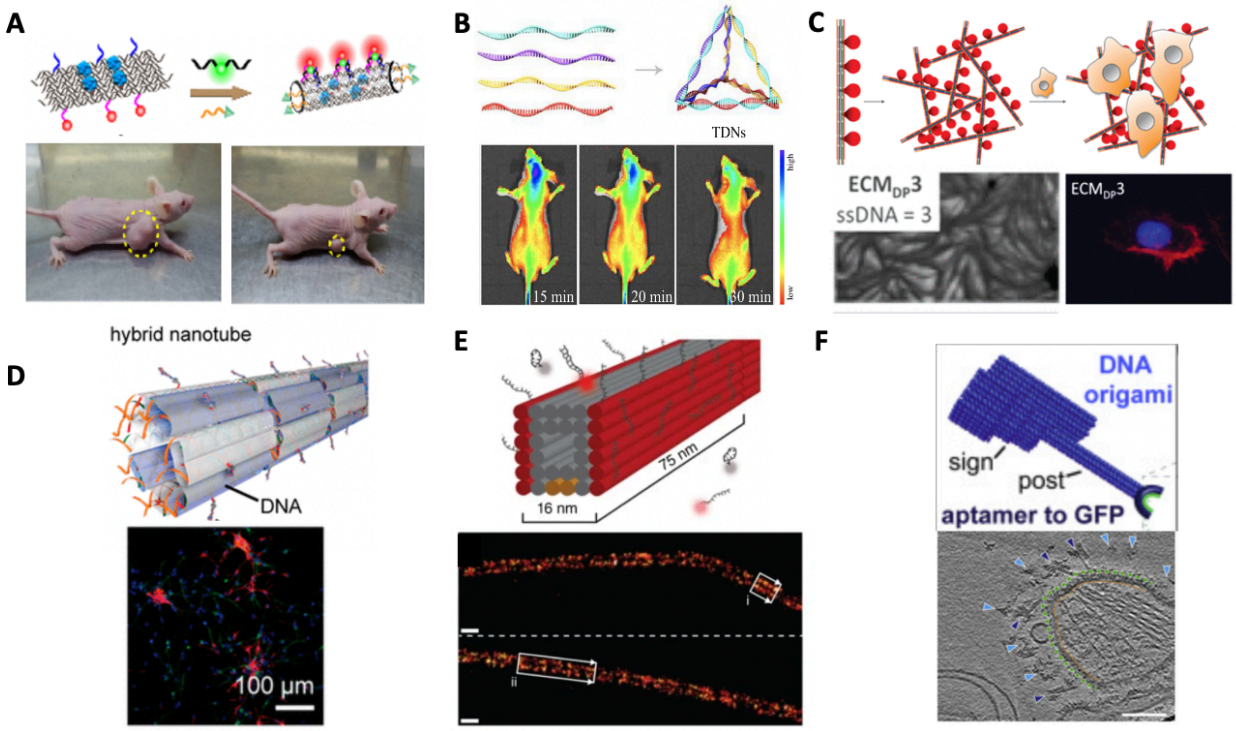


Figure 1.7: Overview of intracellular nucleic acid nanotechnology. A) nucleolin-targeted DNA nanotube for inhibition of tumour growth in xenograft-bearing mice [60]. *Tumour growth is shown to be reduced in treated mice (right) compared to untreated mice (left) after 21 days since injection. B) DNA tetrahedron relieving depression like behaviour in LPS-induced neuroinflammation in mice* [65]*. The biodistribution of Cy5-labeled DN tetrahedron in BALB/c nude male mice. C) DNA-based extracellular matrix to structurally tune the cytoskeletal arrangement of human cervical cancer cells* [69]. *Shown is an AFM image (left) of the DNA-protein matrix with 3 base single stranded regions, and an fluorescence microscopy image (right) of the resulting cell morphology, with actin labeled in red and the nucleus labeled in blue. D) DNA nanotube to stimulate the differentiation of neural stem cells* [70]. *A confocal micrograph is shown of neurons after 7 days of culturing on the DNA nanotubes. E) DNA-PAINT as a multiplexing super-resolution bio-imaging tool* [71]. *A DNA-PAINT image is given for Cy3-labeled strands along a DNA origami polymer. F) DNA origami signposts as cryoEM tags* [72]. *A computed tomograph slice is shown of the SPOTs binding to the outer membrane of sfGFP-gB vesicles.*

Alongside the delivery of cancer drugs, another biomedical application for nucleic acid nanostructures is within tissue engineering and regenerative medicine. DNA materials offer the possibility of finely positioning growth and differentiation factors into a well-defined structure and can program the cellular environment for a predesigned tissue growth outcome. As an adaptive material, DNA has been combined with protein to build a structurally tunable extracellular matrix for the growth of human cervical cancer cells [69] as shown in Figure 1.7C. More distinct geometries depicted in Figure 1.7D include DNA nanotubes that have been functionalised with RGD peptides to enhance the differentiation of neural stem cells into neurons [70].

Not only are nucleic acid nanostructures useful as biomedical devices within cells, but their properties can translate into powerful molecular tools for bio-imaging as they are capable of single molecule observations. A particular technology developed is DNA point accumulation for imaging in nanoscale topology (DNA-PAINT), which is a super resolution fluorescence technique that relies on the transient binding and unbinding of oligonucleotides to specific positions on an origami scaffold resulting in 'blinking' events [73]. Multiplexing images for a large number of distinct cellular species is difficult within cells due their dense and sticky environment, but as evident in Figure 1.7E, DNA-PAINT has been used in four-colour two-dimensional and three-colour three-dimensional imaging of proteins in fixed cells [71]. Additionally, this technique has been combined with fluorescence lifetime imaging microscopy (FILM) to create FL-PAINT that uses lifetime information as a multiplexing parameter for target identification without the requirement of any fluid exchange, leaving the sample undisturbed for imaging up to three targets in a cell [74]. Alongside fluorescence-based imaging, DNA has also been designed as asymmetrical signposts to identify proteins on the outer cellular membrane by electron cryotomography [72]. As shown in Figure 1.7F, these DNA origami nanostructures have a dense, high contrast "sign" head and a "post" tail incorporating an aptamer sequence, increasing the multiplexing capabilities of cryoelectron microscopy compared to the canonically used gold-nanoparticle. These DNA nanostructures will be discussed in more detail in Chapter 5 of this thesis.

Regarding their intracellular applications, one aspect to note is the stability of nucleic acids, as short unmodified nucleic acids have half-lives on the order of minutes with the cell [75]. There are a number of chemical modifications [76, 77] to the sugar, base and backbone of nucleic acids to enhance their intracellular stability. DNA is generally more stable and less prone to mutations as it serves as a long-term storage of genetic information, whereas RNA is more prone to damage and has a shorter lifespan it is constantly synthesised and degraded within the cell. Forming distinct geometries has shown to improve intracellular lifespan of nucleic acid nanostructures compared to simple linear DNA as various 2D and 3D shapes were tested to be more stable up to 12 hours in various cell lysates [78]. DNA origami structures have also shown to be more nuclease resistant compared to linear duplexes or plasmid DNA [79]. As many DNA origami structures unfold in low-concentration magnesium environments, their utility in biological fluids can be compromised through denaturation induced by physiological salt concentrations, but they can be optimised for low salt conditions by increasing the thermal denaturation temperature of their component parts. DNA nanostructures can also be coated with oligolysine polymers to increase their cellular uptake, biodistribution as well as stability with the cell [80]. Additionally, stabilised survival uptake into endosomal compartments and in mouse models, exhibit modest increase in pharmacokinetic bioavailability [81].

Work still needs to be carried out to improve issues of instability and low payload of common DNA assemblies to expand their intracellular applications. Additionally, many established nanodevices require systemic injection to achieve targeted delivery within the cell. New supramolecular modifications as well as a focus on *de novo* assembly of these nanostructures are therefore needed to make this technology a part of next-generation patient-specific nanomedicines.

1.3.2 Intracellular and in vivo delivery of nucleic acid nanostructures

As discussed, nucleic acid nanostructures have great potential in a wide spread of biomedical and biophysical applications. As there already exists a large collection of well characterised *in vitro* assembled structures, the reasonable next step would be to use an effective delivery method to bring these nanostructures into the cell as well as into whole organisms. When entering the cell, there are a few considerations including the pathway of cellular uptake, avoiding or escaping the endosomes, and achieving specific subcellular localisation. In regards to entering immune cells whose roles involve uptake of foreign particles, delivery may not require any specific targeting or penetration design considerations as long as these nanostructures are within the preferred size range between 50-200 nm [82]. However, many other cell types are more selective in the particles that are taken up, therefore distinctive mechanisms to stimulate entry and internalisation need to be considered.

Most of the current intracellular delivery techniques are more suited for permeation into bacterial cells compared to the efficiency in mammalian cells. For example, bacterial cells can handle a greater electrical field strength for electroporation than mammalian cells [83]. Overall, bacterial cells are more robust and contain less cellular components that might interfere with delivery. Therefore, characterisation of intracellular delivery techniques to bring nucleic acid nanostructures into mammalian cells is limited.

The shape of particles affects their cellular uptake, hence there is great advantage in the programmability of the size and shape of nucleic acid nanostructures. It has been shown that larger DNA origami structures with greater compactness were preferentially internalised compared to elongated, high-aspect ratio structures [84]. This phenomena has also been shown with other nanoparticles, as mammalian cells preferentially internalised compact hydrogel nanodiscs compared to longer nanorods [85]. Additionally, these compact nanostructures, such as DNA cages and tetrahedra, were found to improve the cellular uptake and gene silencing abilities of antisense oligonucleotides [86] and siRNA [66], compared to if these strands were delivered unprotected.

There already exists various commonly used methods to stimulate cellular uptake of naked linear and plasmid DNA and RNA that can potentially be used for the intracellular delivery of larger and more complex nucleic acid nanostructures, a few of which are depicted in Figure 1.8. Electroporation is one commonly used gene transfer technique that employs an external electric field to stabilise the permeation sites on the cellular membrane to create temporary pores to allow nucleic acid transport [87]. Additionally, electroporation of naked DNA delivered into adherent cells showed significant binding, through simple electrostatic interactions, of various intracellular proteins to productively traffic the DNA through the cell [88]. Electrotransfections have been used to deliver polyamine folded DNA origami structures as shown in Figure 1.8A [89]. This method

has also been used to deliver gene-encoding multimeric DNA origami bricks [90] into mammalian cells.

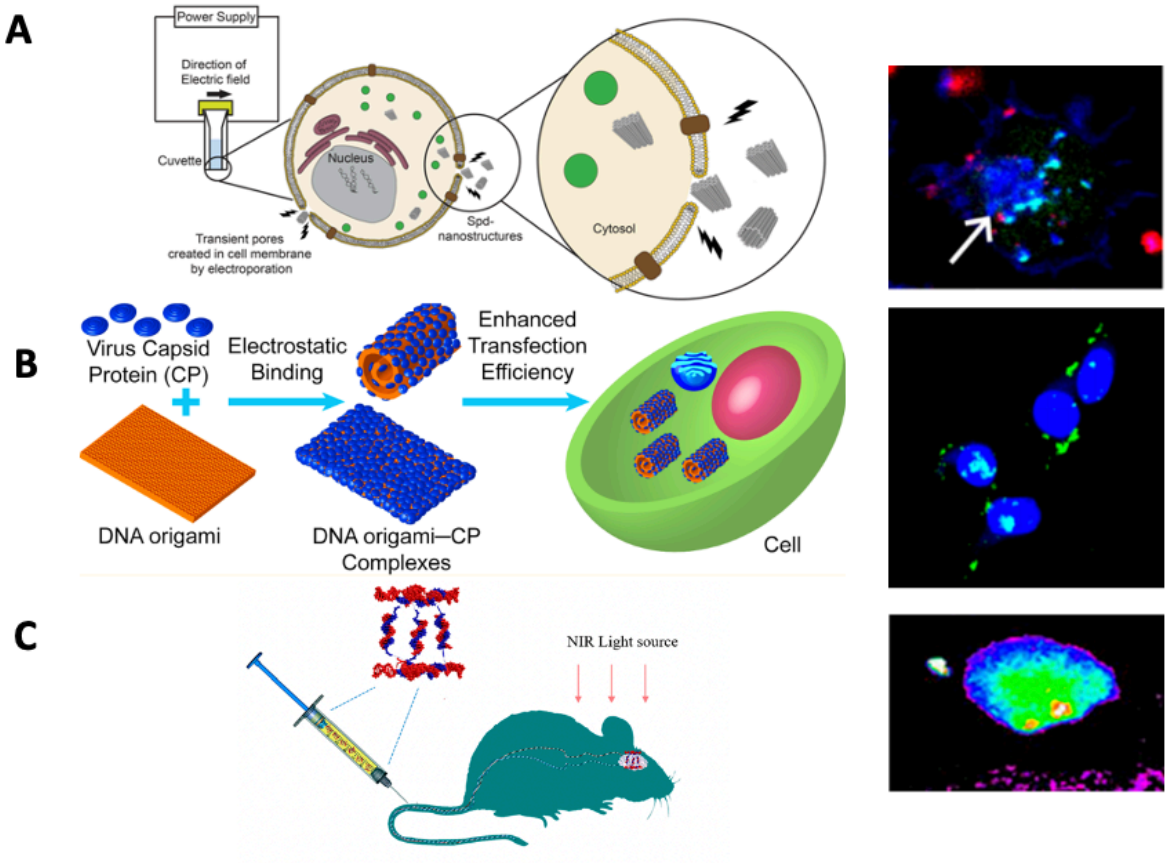


Figure 1.8: Overview of intracellular delivery methods for nucleic acid nanostructures. A) Electrotransfection of DNA origami into Jurkat cells [89]. *In the fluorescence micrograph, the nanostructures are labelled red, the nuclei blue and the lysosomes in green. B)* 64*DNA arrays coated with CCMV CPs to aid delivery into HEK293 cells* [91]. *The DNA nanostructures are shown in green and nuclei in blue in the fluorescence micrograph. C) Injection of DNA nanocages into a U-87 MG xenograft mouse model to inhibit tumour growth* [64]. *An ex vivo fluorescence image is given of the brain 90 min after inject to show biodistribution of the nanocages.*

Another canonical gene delivery method is through liposome-based transfection. Lipofections utilise lipid complexes to form a liposome around the DNA structure, which having a similar composition as the cell membrane, allows for easier transfer by endocytosis or direct membrane fusion [98]. Lipid-based transfection of larger DNA nanostructures have been touched upon briefly with the use of a lipid-coating to improve cell membrane penetration of DNA origami [99].

Similarily to applying lipids to the outer surface of a nanostructure, they can also be decorated with cell-penetrating molecules which may allow entry into cells even in the absence of the ability to bind to specific receptor. This technique has been used on DNA six-helix bundles functionalised with cell-penetrating peptides for delivery into Ramos cells [99] as well as with DNA tetrahedron appended with tumour-penetrating peptides to increase the rate of uptake into glioblastoma cells

[100]. Viral capsid proteins have also been used to coat DNA origami two-dimensional arrays to compact the surface and improve the cellular attachment and delivery into cells as shown in Figure 1.8B [91]. In many of these cases, specialised ligands were used to act as receptors to proteins found on the surface of the specific cells of interest. Therefore, this technique of attaching cell-penetrating particles to nucleic acid nanostructures would require a new coating for delivery into different cell types.

Microinjections straight into the cytoplasm [92] or nucleus [93, 94] is also a technique to deliver nucleic acids into single cells. This method uses a needle or a film of silicon nanoneedles [95, 96] to pierce the cell membrane for successful delivery into the cell. Direct injection of DNA nanostructures carrying anticancer drugs has been used to deliver these nanodevices into the targeted tumour-related areas in whole organism mouse models as depicted in Figure 1.8C [67] as well used to inject transcription-modulating DNA origami rods straight into the yolks of single-celled zebrafish embryos [97]. Overall this technique is useful for site-specific delivery of nucleic acid nanostructures.

1.3.3 Challenges of current intracellular delivery methods

Although there are various methods to potentially deliver nucleic acid nanostructures into the cell, there is no universal method that will work for every structure shape and cell type. Although shape can improve membrane transport, it has been shown that cellular uptake kinetics are more cell-type-dependent, as even compact DNA nanostructures failed to saturate specialised endocytosing dendritic cells [84]. Therefore, careful consideration needs to be used depending on the cell types used in each experiment.

There are also specific challenges to each of the delivery methods mentioned. For electrotransfections in a traditional cuvette, DNA origami has shown to form aggregated clumps [89]. The strong electric field applied can strip away the ions that stabilise the nucleic acid nanostructures. Therefore, careful consideration needs to take place in the folding design of these nanostructures in buffers with little to no Mg^{2+}. On the other hand, although microinjections can deliver intact nucleic acids directly into the cell, this technique is limited to a relatively small number of cells. Lipid-based transfections which have shown overall better delivery of a large number of probes through the membrane, are also limited in the fact that a significant fraction of the transported nanoparticles are enclosed in endosomes and do not reach the cytoplasm [101, 102].

Once administered to the cell or body, nucleic acid nanostructures have to cope with various encountered *in vivo* conditions such as clearance from endosomes inside the cell or by the liver, spleen, kidney or mononuclear phagocytes within whole systems, as well as execute their function at their specific target. As mentioned previously, more compact DNA nanostructures can enter the cell unaided, but when tracking the cellular entry mechanism using fluorescent live cell imaging showed that DNA tetrahedra enter HeLa cells by a caveolin-mediated endocytic pathway [103]. Therefore, most of these nanostructures are localised in the lysosomes, needing further endosomal escape mechanisms. Additionally, different nucleic acid delivery methods can result in vast differences in cellular uptake timing, amount and subcellular distribution.

1.3.4 Nucleic acid nanofabrication inside the cell

Intracellular production and folding of nucleic acid nanostructures is a possible way to circumvent many of the challenges involved in intracellular delivery. *De novo* assembly corresponds well with the replicable nature of nucleic acids and does not require many external components as it can utilise the cell's own machinery and replication pathways. The main consideration needed is the folding pathway of what is most likely going to be a single-strand of RNA, therefore, one will need to avoid kinetic trapping and misfolding within the cell.

For DNA origami nanostructures, there have been a few cases of replicating the long scaffolding strand inside cells, then harvesting the DNA to fold *in vitro* [38, 104]. This process can be streamlined to replicate and fold the strand entirely within the cell. Although, instead of DNA, RNA is a better candidate due to the simultaneous transcription and folding of this molecule within the cell. This co-transcriptional process of folding intracellular RNA nanostructures is shown as two examples in Figure 1.9.

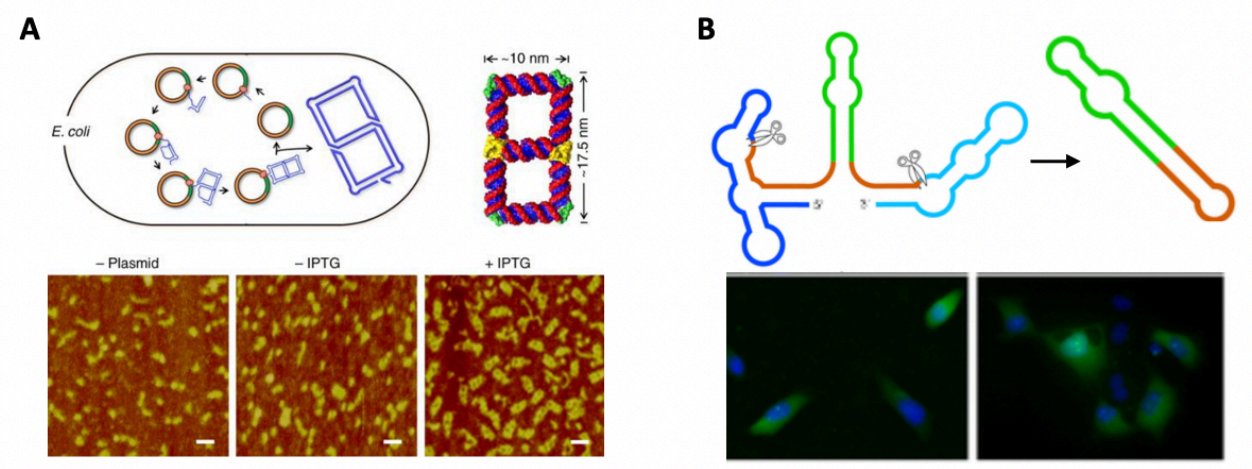


Figure 1.9: Co-transcriptionally folding RNA nanostructures inside the cell. A) RNA double nanosquares are folded inside E. coli cells and the lysed out to examine by AFM [105]. *B) RNA Broccoli aptamers are expressed and circularised in mammalian cells through ribozyme activity* [106]. *Fluorescent microscopy images show the distribution of aptamers (green) in HeLa and COS-7 cells (blue nuclei).*

The exploration of *in vivo* assembled nucleic acid nanostructures was conducted in Chengde Mao's group in 2018 [105]. Here, the researchers were able to assemble a two-dimensional double square structures directly in *E. coli* cells from transformed plasmids. The transformed cells were lysed and sonicated to analyse the folded structures by polyacrylamide gel electrophoresis (PAGE) and atomic force microscopy (AFM). Through microscopy, they were able to confirm correct folding of the nanosquares, particularly of the integral 3-way loop motif. Additionally, based on the band intensities in the PAGE graph, it can be estimated that 11.2% of all RNAs produced in the cells were the programmed double-square structure.

With regards to having a more complicated replication process, there has not been as much progress in the *in vivo* fabrication of nucleic acid nanostructures within mammalian cells. Unlike bacteria that only have a single type of RNA polymerase, eukaryotes utilise up to three types of RNA polymerases. Each eukaryotic RNA polymerase has its own unique role in the synthesis of specific RNAs. Therefore, when considering using this machinery for the production of a synthetic RNA nanostructure, specific promoter and terminator sequences need to be included to allow for successful transcription. These lengthy sequences will likely be incorporated into the final RNA nanostructure, which can interfere with the folding pathway. One way to bypass this is through autocatalytic sequences, which has been used in the *in vivo* transcription of circular RNA aptamers inside mammalian cells [106]. Here, they used Twister ribozyme sequences, after the promoter sequence and before the terminator sequence, that rapidly undergo autocatalytic cleavage by endogenous RNA ligases. Therefore the promoter and terminator sequences are removed from the final replicated RNA strand that is circularised in the process.

More work needs to be done to fully utilise the biological nature of nucleic acids. This is particularly important with *de novo* assembly inside mammalian cells, to fully capture the full capabilities of nucleic acid nanodevices in the biomedical and biophysical fields.

1.4 Summary and Thesis Structure

There is currently no other molecular engineering technology that enables the fully *de novo* design of a similarly complex and diverse set of biomolecular systems with such molecular precision and addressability, as seen with DNA and RNA. Therefore progression needs to be carried out towards more *in vivo* applications for nucleic acid nanotechnology that require a minimal fabrication method balanced with design simplicity and complex function. In this chapter, I have showcased the background literature that highlights the potential for nucleic acid nanostructures inside the cell. As the field has already characterised a vast catalogue complex structures from autonomous molecular robots to drug targeting devices, by either delivering these already assembled constructs into the cell or by folding them directly inside, we can create a whole new intracellular platform for nucleic acid nanotechnology applications. In this thesis, I propose three methods with the prospective capacities to allow for the intracellular usage of various nucleic acid nanostructures.

In Chapter 2, I will summarise the design software and experimental techniques I used throughout my work in this thesis.

In Chapters 3 and 4, I will discuss my first project: the design and analysis of a new single-stranded DNA and RNA assembly process inspired by knitting. In Chapter 3, I will revisit the concept of single-stranded nucleic acid assembly and provide relevant information on how that process can translate into knitting and different patterned designs. I will then go over how I used NUPACK and oxDNA to model the design parameters for a single and multi-knit DNA structure, as well as KineFold and CoFold to predict the folding pathways of these structures. In Chapter 4, I will show the analysis of the single and multi-knit DNA designs by gel electrophoresis and through atomic force microscopy. I will go over the optimisation steps for immobilising the flexible DNA knitted nanostructures onto the mica for AFM. I will then briefly explain how the concept of nucleic acid knitting can be used for co-transcriptionally folding of RNA nanostructures directly inside the cell.

I finished off these chapters by reviewing how the multi-knit DNA design can be improved and discuss the potential future experimental directions.

Chapter 5 will be covering my second project, which uses a confinement tool to deliver DNA origami cryoelectron tomography tags into mammalian cells. I will discuss both commercial and homemade versions of the confinement chambers, as well as go over the experimental design. I will then show preliminary data with fluorescent confocal microscopy to optimise the confinement parameters. Additionally, I will show intracellular membrane tagging by the DNA origami signposts after intracellular delivery through compression. Finally, I will discuss the issues that arose during my setup and explain how future experimentations can be improved.

Chapter 6 introduces my third project, which goes over the concept of a reconfigurable DNA nanodevice. In particular, I will review three different bistable hairpins with their own distinct folding pathways. These hairpins are designed to have two energetically favourable states to allow for confirmational switching through toehold mediated strand displacement. With these hairpins, I will discuss how modelling these systems can unravel the kinetics behind dynamic nucleic acid nanostructures as well as discuss their use for compacting structures for cellular entry.

Finally, in Chapter 7, I will summarise the key findings of this thesis as well as discuss the future research directions directly building on the work presented in the thesis.

Chapter 2

Materials and Methods

This Chapter summarises the materials, experimental procedures and the data analysis methods used and/or developed for the work in this thesis. First, I will discuss the opensource software used to design and characterise the nucleic acid nanostructures as well as go through the sample preparation protocols. I will present the characterisation techniques, followed by the data analysis methods. Finally, I will briefly discuss the tissue culture setup and subsequent cell confinement techniques used.

2.1 NUPACK

NUPACK 3.0 [107] and the current cloud version NUPACK 4.0 [108, 109] are online web interfaces developed within the Beckman Institute at Caltech for the analysis and design of nucleic acid structures. This software can create ordered complex ensembles as well as analyse strand interactions in a test tube. In particular, I used NUPACK to design the sequences for my single-knit DNA designs, multi-knit DNA designs and reconfigurable hairpin designs. All reactions were run at 37 °C and standard salt conditions. NUPACK is limited by the exclusion of pseudoknots in a structural ensemble. Therefore for my complex knitting structures, which form pseudoknot-like loops, I had to combine the secondary structural predictions from NUPACK with other software. In addition to the design of sequences, I used NUPACK to analyse the melting temperatures and stability of the interaction domains within my rolling circle amplified multi-knit DNA nanostructure. Further discussion of the features and usage of NUPACK will be explored in more detail in Chapters 3 and 6, when designing the knitted nanostructures and bistable hairpins, respectively. The temperature-dependent parameter sets are based on John SantaLucia's work on nucleic acid thermodynamics [110, 111]. A comparison of NUPACK versions 3.0 and 4.0, particularly focusing on their free energy parameters was investigated in Chapter 6. NUPACK 4.0 can also be used as a C++ library Python package. This enables more customisable and complex sequence design and analysis.

2.2 Kinefold and coFold

Kinefold [112] is a web server developed by the Isambert Lab within the Institut Curie to predict the lowest free-energy folding pathway of RNA and DNA, including pseudoknot and entangled helices using clustered stochastic simulations [113]. KineFold simulates nucleic acid folding pathways at the level of the formation and removal of helix regions, which are known to be the time limiting steps of nucleic acid folding kinetics. I used the program to generate the hairpins of my multi-knit DNA and RNA nanostructures. KineFold has a sequence input limit of 400 nt, therefore I had to split up my multi-knit DNA sequence into two parts to analyse the interactions. For my RCA reactions, I used a co-transcriptional rate of a new base added every 3 ms, as

representative of the T7 polymerase [114]. Alongside predicting the folding pathway of the structure, KineFold also generates helix tracing graphs for each sequence. These graphs follow the lowest free-energy during the folding process plotted against time and monitors the five most stable helices. I used these graphs to analyse the predicted kinetics of the helices within my structures.

CoFold [115] is a web server developed by the Meyer Lab within MDC Berlin, as a thermodynamic-based tool for predicting RNA secondary structure that takes co-transcriptional folding into account. This software utilises two different thermodynamic energy parameters [116, 117]. I used CoFold to produce arc plots of the folding pathways of my RNA single-knit structure. Compared to KineFold, CoFold has a larger size limit for the input sequence, at 50,000 bases, which is useful to modeling long sequences such as my DNA multi-knit structure. A caveat to CoFold, is that the arc plots are less straightforward to interpret than the schematics created in KineFold. Additionally, the CoFold arc plots may be limited in predicting structures with multiple interactions in a single helix.

2.3 oxDNA and oxView

oxDNA is a simulation code developed by the Ouldridge, Doye and Louis groups for the coarse-grained modeling of DNA [118]. oxDNA can perform both molecular dynamics and Monte Carlo simulations of oxDNA and oxRNA models, run on single CPUs or single CUDA-enable GPUs. The package also features a Forward-Flux Sampling interface to study the kinetics of rare events. Additionally, users can alter the behaviour of their systems by adding external forces, through harmonic force mutual traps, that can be used to pull on or apply torques to strands or confine nucleotides within semi-planes or spheres. The package also includes oxpy, a Python library allowing the control of the oxDNA simulations using Python scripts. Additional tests can then be generated to study nucleic acid-related quantities, such as analysing binding and unbinding events during the trajectory of a reaction. I used oxDNA to optimise the sequence designs of my single-knit and multi-knit DNA and RNA models. The simulations of these designs utilised strategically placed external forces, as harmonic force mutual traps with stiffness of 0.1. Simulations were run from a range of 1e5 to 1e9 steps at 25 °C and a salt concentration of 0.5. Additionally, with the help of Dr. Jonathan Bath, we created a Plus-Minus multi-knit DNA model to simulate the dynamics and structural characteristics of lattices of different looping patterns. Root mean square fluctuation calculations and force-extension curves were generated from the processed data.

To aid the design of oxDNA models, the browser-based visualisation tool oxView was developed by the Šulc, Yan, Doye, Turberfield and Louis groups [119]. OxView utilises a Three.js JavaScript library to create a smooth and seamless oxDNA configuration viewing and editing experience. I used oxView to design and relax my DNA and RNA nanostructures. For many of these structures, I manually manipulated the structures into their desired shapes and relaxed them by connecting to the Ox-Serve NANOBASE.org server and running a short 10,000 step simulation using their GPUs.

Additionally, I used the online service oxDNA.org, as developed by the Šulc lab at Arizona State University, to run a few of my oxDNA designs [120]. The server uses oxDNA and oxRNA models and equilibrium sampling of assembled designs to simulate DNA and RNA nanostructures.

2.4 Nucleic Acid Nanostructure Preparation

Single-stranded DNA oligonucleotides were ordered from Integrated DNA Technologies (IDT). Shorter sequences were ordered at 25 nmole scale with standard desalting purifications. For the longer (>100 nt) single-stranded sequences in my HTL model, these strands were ordered at 4 nmole scale Ultramers with standard desalting purifications. Double-stranded DNA templates for my multi-knit DNA structure were ordered as gene fragments from Twist Biosciences. DNA oligonucleotides and templates were resuspended in 1X Tris-EDTA (TE) buffer (Thermofisher).

For the strands in my single-knit HTL model, the hairpin and thread strands are quenched at 95 °C for 30 seconds then on ice for 30 seconds in 10 mM $MgCl_2$, while the lock strand is quenched in MilliQ water. To create a "knitted" reaction, 5 µL of 10 µM quenched hairpin is incubated with 5 µL of 10 µM quenched thread for 15 minutes. Then 8 µL of this hairpin-thread mixture is incubated with 8 µL of 10 µM quenched lock strand for 15 minutes. Then 12 µL of this hairpin-thread-lock strand is incubated with 6 µL of 10 µM toehold displacement strand for 15 minutes. Finally, 12 µL of this hairpin-thread-lock-toehold displacement strand mixture is incubated with 8 µL of 5 µM invader strand for 15 minutes. For the "non-knitted" reaction, the sample is similar, except that 5 µL of 10 µM thread is incubated with 5 µL of 20 µM lock strand. Then 8 µL of this thread-lock mixture is incubated with 8 µL of 4 µM hairpin. 12 µL of this thread-lock-hairpin mixture is incubated with 6 µL of 10 µM invader strand. Finally, 12 µL of this thread-lock-hairpin-invader strand mixture is incubated with 8 µL of 5 µM toehold displacement strand.

To assemble the multi-knit DNA template, I used a plethora of molecular cloning steps. Firstly, the template was amplified by Q5 polymerase (NEB). 12.5 µL of the Q5 2X MasterMix (NEB) was combined with 1.25 µL of a 10 µM forward primer and 1.25 µL of a 10 µM reverse primer, as well as up to 1 µg of the DNA strand, and topped up to 25 µL with water. The following PCR protocol is used: an initial denaturation step at 98 °C for 30 seconds, then 40x cycles of 98 °C for 20 seconds, 72 °C for 30 seconds and 72 °C for 30 seconds, with a final extension at 72 °C for 2 minutes. The template halves are then circularised using Golden Gate Assembly. Here, 3 µL of template 1 and 3 µL of template 2 are combined with 1 µL BsaI GG Enzyme Mix (NEB) and µL BsmBI GG Enzyme Mix (NEB) with 2 µL of 10x ligase buffer, and topped up with 20 µL water. This mixture is incubated at 37 °C for an hour, then 42 °C for an hour, with a final heat inactivation step at 60 °C for 5 minutes. To nick the top strand of the circular template, 10 µL of the sample is mixed with 1 µL Nt.BbvCI (NEB) and 2 µL of 10x NEB CutSmart buffer, and topped up with 20 µL of water. This mixture is incubated at 37 °C for an hour with a heat inactivation step at 80 °C for 20 minutes. Finally, to start the rolling circular amplification of my template, 10 µL of the nicked template is mixed with 3.75 µL of a dNTPs mixture (10 mM each dATP, dCTP, dGTP and dTTP) (NEB), 5 µL Phi29 buffer (NEB), 1 µL Phi29 polymerase (NEB) and topped up with 30.25 µL of water. This mixture is incubated at 30 °C for 2 hours with a heat inactivation step at 65 °C for 10 minutes. The polymerised product is then checked on a NanoDrop for the concentration, which should range between 1000-2000 ng/µL. The final sample can then be analysed by gel electrophoresis or through atomic force microscopy.

For the DNA origami signposts, nanostructures were gifted by Emma Silvester or assembled using the given protocol [72]. 10 nM of the M13mp18 scaffold is combined with 100 nM of each staple strand, 2.4 µM Cy5-bound fluorescent strands and 6 µM aptamer strands. A 100 µL mixture is then

annealed using the program: 65 °C for 15 minutes, 7 cycles of 65 °C for 5 minutes followed by 24 cycles at 59 °C for 35 minutes, 5 cycles at 35 °C for 10 minute and finally, 15 cycles at 30 °C for 5 minutes. The samples are immediately purified through PEG precipitation, where the 100 µL annealed sample is mixed with 100 µL PEG precipitation buffer (1X TE, 12.5 mM $MgCl_2$, 500 mM NaCl, 15% w/v PEG 8000) and spun for 30 minutes at 2000 rcf at 20 °C. The supernatant is then removed and the pellet is resuspended in 10 µL 1X folding buffer (1X TE and 12.5 mM $MgCl_2$). This spin and wash step is repeated two more times. The final resuspension is in a volume of 10 µL 1X folding buffer.

For the bistable DNA hairpins, the hairpins are quenched in 10 mM $MgCl_2$ and incubated with the toehold displacement strands at a ratio of 1:2 (hairpin : toehold displacement strands). The mixed sample is then annealed from 95 °C to 15 °C at a rate of -1 °C per min.

2.5 Gel Electrophoresis

Gel electrophoresis is used in this thesis to analyse the size and approximate shape of the nucleic acid structures. In particular, polyacrylamide gel electrophoresis (PAGE) is used to resolve small nucleic acid strands. For the polyacrylamide gels, a 30% 29:1 Acrylamide:bis-acrylamide stock solution (Sigma-Aldrich) was mixed with 50X TAE (Tris, acetic acid and EDTA) buffer (Thermofisher) diluted with water to give a final 1X buffer concentration, and the desired acrylamide concentration, which ranged from 10% (for the HTL model) to 15% (for the bistable hairpins) for experiments. Approximately 5 mL of the mixture is prepared per gel, combined with 1/100 volume of 10% ammonium persulphate (Sigma-Aldrich) and 1/1000 volume of tetramethylenediamine (Sigma-Aldrich). This mixture is poured between two glass plates with a 0.75 mm spacing, and an appropriate gel comb is added. The gel is left to polymerise for at least 10 minutes. Gels are then transferred to a Mini-PROTEAN tetra cell vertical gel electrophoresis system (Bio-Rad) and topped up with 1X TAE running buffer. For gel electrophoresis analysis of my single-knit HTL model, 10 µL of a 1 µM sample mixed with NEB loading dye to a final concentration of 1X. For my multi-knit DNA design, as much concentration for each sample is loaded into the wells. A 100 bp DNA ladder (NEB #B7025) can also be added as a reference point for the approximate strand length. These polyacrylamide gels are then run for 45 minutes to 90 minutes at 90 V.

The nucleic acid samples in each gel are stained with SYBRGold, which is an intercalating nucleic acid dye that can be used to stain double-stranded DNA, single-stranded DNA and RNA [121]. If screening for FAM signal, SYBRGold is added afterwards, as the signals overlap. Gels are scanned on the Amersham™ Typhoon™ biomolecular imager. With the gel scanner, appropriate filters are used for SYBRGold, FAM and Cy5 signals. For SYBRGold or FAM signal a laser wavelength at 488 nm is used with a filter bandpass of 525 nm. For Cy5 signal
A laser wavelength at 532 nm is used with a filter bandpass of 670nm.

Scanned polyacrylamide gels are then analysed in ImageStudioLite (LICOR) and ImageJ [122]. These are image processing software that can be used for the quantification of gels. I used these programs to adjust the colour balance as well as quantify the signal of each band intensity.

2.6 Microscopy

My multi-knit DNA nanostructures were imaged by atomic force microscopy on the Asylum Cypher ES as well as the Bruker FastScan-XR AFMs. AC40 biolever mini cantilevers were used with the Cypher, while FastScan-D cantilevers were used with the Bruker AFM. For either AFM, samples were imaged in liquid tapping mode. Samples were adsorbed onto mica through one of three different protocols as more thoroughly discussed in Chapter 4.5.1. The first adsorption protocol, mainly used with the Cypher AFM, utilised an imaging buffer of 12 mM $MgCl_2$, 4mM $NiCl_2$ and 1X TE. The sample is diluted to 100 ng in the imaging buffer, then 10 µL is deposited onto cleaved mica and incubated for approximately 2 minutes. The sample is then topped up with 50 µL more imaging buffer. The second absorption protocol was a modified version of the first as inspired by an AFM method for imaging G-quadruplex wires [123]. This protocol uses the same conditions as the first, except that 20 mM of $NiCl_2$ is used. The third adsorption protocol was based off the Pyne group's protocol [124] and mainly used with the Bruker AFM. This protocol uses an imaging buffer containing 3 mM $NiCl_2$ and 20 mM HEPES pH 7.4. First, 20 µL of the imaging buffer is deposited onto the cleaved mica, then 4 µL of 10 ng of sample is combined with this solution. The sample is incubated on the mica for 30 minutes, followed by 4 wash steps where the imaging buffer is replaced. The sample is topped with 40 µL of imaging buffer.

The typical scan size for an AFM image ranges from 120-250 nm, with a pixel density of 256-512 pixels/line and a line rate of 3-5 Hz. The imaging PeakForce (PF) setpoint is set to around 70, 000 pN, with a PF acquisition frequency of 8 and PF amplitude ranging from 5-10,000 nm.

To analyse the AFM images, structures can be flattened directly on the software, in the scanning probe microscopy tool Gwyddion [125] or through the AFM image analysis program TopoStats [126]. TopoStats in particular, was used to trace the structures and measure their respective sizes.

Fluorescence microscopy was performed using a Zeiss Axio Observer with filters for Cy5 and GFP visualisation. Bright-field was used in cases where a membrane dye was not available. Unless otherwise specified, confocal images were taken at 100x magnification with oil immersion.

2.7 Tissue Culture and Cell Confinement

MCF10A and Hek293T cells were cultured as described in Anchal Singh's principles for animal tissue culture [127]. These conditions include maintaining cells at an adequate temperature around 37 °C, 5% CO_2, and growing them in Dulbecco's Modified Eagle Medium, with high glucose GlutaMAX supplement (Thermofisher) combined with 10% fetal bovine serum (Thermofisher) and 1% antibiotic-antimycotic solution (Sigma-Aldrich).

For dynamic cell confinement experiments, cells were adhered to 35 mm petri dishes (ibidi) 2 hours before experimentation. For the static cell confinement experiments, approximately 30,000 cells were seeded into 35 mm petri dishes the day before imaging. If the Hek293T cells expressing eGFP-cd63-BMP7 were being used, 500 ng/mL doxycycline was also added at this point. Before confinement, cells were imaged with and without the DNA origami signposts as a control. The cells were then confined at an appropriate height for an hour in media containing 10 nM DNA

signposts. With the dynamic 4Dcell chambers, cells can be imaged throughout confinement. The cells were then released from confinement and were allowed to recover for another hour. The cells were then washed and resuspended in new media free of DNA signposts and images were collected here. The cell confinement protocols are discussed more thoroughly in Chapter 5.

In preliminary experiments, a phalloidin dye (Thermofisher) is used to stain the dead cells. CellTracker Green CMFDA dye (Thermofisher) was also used to stain the cell membrane. In later experiments, a HEK293T cell line was engineered to express GFP on the cd63-BMP7 proteins to aid visualisation of the cell membrane for fluorescence microscopy.

Chapter 3

Designing Nucleic Acids to Assemble in a Knitted Fashion

3.1 Introduction

In this Chapter, I will explain a novel single stranded nucleic acid assembly method inspired by knitting. Expanding on the discussion in Chapter 1, I will further explore the process of single stranded DNA and RNA assembly as well as go over the concepts and mathematics behind knitting. Using this background information, I will explain how I went from designing a single knitted DNA structure to a multi-knitted construct. In conjunction, I will show how I used various software to predict and model the folding pathways of these single-stranded structures.

3.2 Author Contributions

All sequences in this Chapter were designed in NUPACK and KineFold by Catherine Fan. The DNA models in oxDNA were created by both Catherine Fan and Dr. Erik Benson. Both members worked on creating the structures in oxView, with Catherine Fan particularly focused on the single-knit structure and Dr. Erik Benson helping with the multi-knit model. The Plus-Minus knitted models in Section 3.6.2, and subsequent dynamic simulations, were run by Dr. Jonathan Bath.

3.3 Designing a Single-Stranded Nucleic Acid Nanostructure

Single-stranded nucleic acid assembly was introduced in this thesis back in Sections 1.2.3 and 1.3.4, focusing on ssDNA and RNA assemblies *in vitro* as well as *in vivo* co-transcriptionally folded RNA. The main idea behind these assemblies is to use strategically placed motifs to fold a long strand into a desired architecture. The use of non-canonical nucleic acid motifs, which form easier within a single strand due to effective increased local concentration and improved stoichiometric control, further aids the complexity of the shapes these structures can form. The biggest challenge of single stranded assembly is within this folding pathway as kinetic traps can form as the main mode of assembly shifts from intermolecular interactions to intramolecular interactions.

In the case of self-assembly without an external energy input, the structure is expected to come together at the global free-energy minimum. With single stranded self-assembly, this process rarely reaches the expected ordered equilibrium, and instead, is arrested at the intermediate local minima in the free-energy landscape, resulting in metastable “kinetically trapped” states [128, 129]. Therefore, directing the system to the ordered ground state by escaping local free-energy minimum is crucial in the development of sophisticated single-stranded nucleic acid nanostructures.

In unimolecular assembly, not only do interactions within the particle need to be chosen precisely so that the ordered state minimises the system's free energy, consideration is also required to avoid forming molecular knots. A molecular knot is a mechanically interlocked molecular architecture that is analogous to a macroscopic knot, and as a kilo-bases long single strand of DNA or RNA is programmed to self-fold, it can find itself trapped topologically. While stronger interparticle bonds stabilise the ordered equilibrium state, it is also the formation of transient bonds and frequent bond-breaking events that help the self-folding nucleic acid strand escape these kinetic traps.

In addition to the use of strategic bond forming and breaking events, a method to improve single-stranded assembly is through the use of co-transcriptional folding. During co-transcriptional folding, a nascent RNA molecule exits from the polymerase and transitions through intermediate structural states that ultimately determine the final RNA structure [130]. The 5′-3′ polarity of RNA synthesis is key to this assembly process by directing an order of folding. In addition to the kinetic traps that occur early on during transcription, nucleic acids have also been shown to enter a late maturation process where the strand enters a kinetically trapped state after co-transcriptional folding that takes hours to escape [36].

During this RNA replication process, the folding pathway can also be affected by the polymerase speed. The speed of replication of polymerases can range from a new base added every 3 milliseconds for T7 polymerase and one base every 200 milliseconds for most eukaryotic RNA polymerases [131, 132]. Folding of the strand may not be immediately complete after transcription and further structural changes may occur due to a slower or faster folding process. Therefore, careful consideration is required in regards to which transcriptional setup is used as the time it takes for a system to reach its ordered equilibrium state can result in more or less kinetic trapping. In general, high processivity and high replication speed polymerases aid co-transcriptional folding as they can pass through kinetic barriers imposed by the secondary structure, allowing for longer range interactions to form [133].

Self-complementarity is an intrinsic property of single-stranded assembly, but this strong secondary structure may present challenges to synthesising the DNA template. This issue was encountered in my work in this thesis, as the template sequences designed for my single-stranded assemblies could not be synthesised by the well-known nucleic acid synthesis company, Integrated DNA Technologies (IDT). This was particularly due to the repetitive nature of my template, as a few complexities identified by IDT that prevented the manufacturing of my sequence include: one or more of my repeated sequences being greater than 8 bases comprising 70% of the overall sequence, the length of my repeated sequences needing to be less than 11 bases, and the my hairpin stem sequences requiring reduction in the length of the stem or complement to less than 10 bases. Luckily, by splitting my template in half, which would later require reattachment through molecular cloning, I was able to synthesis my templates with the company Twist BioSciences.

Self-assembly of nucleic acids is just bringing simple components together to form an ordered product. New methods for single-stranded nucleic acid assembly can be created by looking towards existing structures composed of a single-stranded particle. In particular, one can imagine nucleic acids like a string of yarn that can be knitted into all new assembly patterns. In this thesis, I will take these considerations for single-stranded assembly and combine it with the well-known art of knitting.

3.4 Knitting

3.4.1 Stitches and knit patterns

Knitting is an ancient skill still used today to create intricate works of art or produce items fulfilling practical needs. With the earliest knitted artifacts, a pair of socks, dating back to the 11th century CE in Egypt [134], knitting as a technology has developed throughout the ages from hand-knitting to using ivory or metal needles to high-capacity modern day knitting machines. Alongside the use of conventional fabrics, knitting can also be conducted using a plethora of material from spinning yarn made from wood [135] to even knitting with nucleic acids as to be shown in this thesis.

At the most basic level, knitting is just pulling a loop of yarn through another loop to form a stitch (Figure 3.1A). There are two fundamental stitch patterns in knitting: the knit stitch (Figure 3.1B (i)) which loops from back to front, and the purl stitch (Figure 3.1B (ii)) which loops from front to back. All other patterns, a few of which are shown in Figure 3.1C, are just a combination of these two. You can form the moss stitch by alternating knit and purl stitches on the same row, the stockinette stitch by creating a row of knits then turning and making a row of purl stitches, the garter stitch by creating knit stitches on each side, and ribbing patterns by creating combinations of alternating knit and purl stitches within the same row. The canonical form of knitting is flat-knitting, which creates rows of loops using two knitting needles, where the piece is tuned from front to back on each row. An alternative form to knitting, particularly for creating continuous pieces, is knitting "in the round" which uses multiple needles to form a seamless tube.

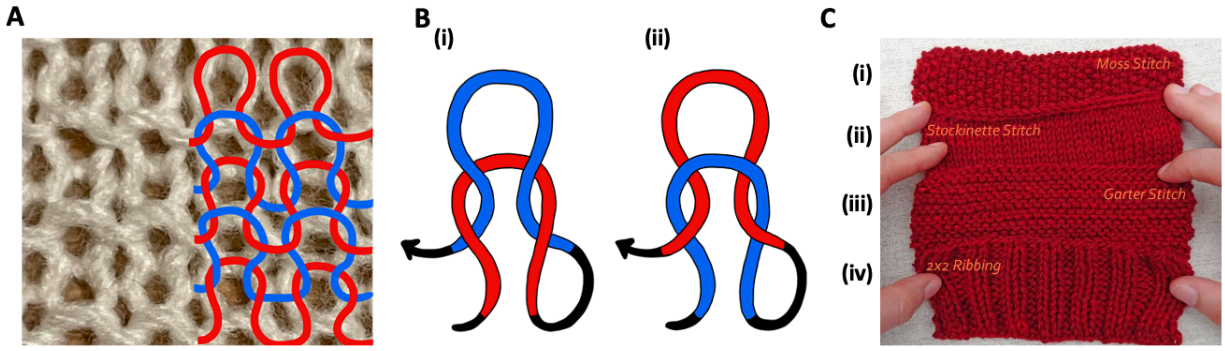


Figure 3.1: Stitches and patterns in knitting. A) Knitted pattern with alternating looping in each row. The red rows loop from the front to back and the blue rows loop from the back to front. B) Simplified schematic of the (i) knit stitch and (ii) purl stitch. C) Various knitting patterns are shown, including the (i) moss stitch, (ii) stockinette stich, (iii) garter stitch and (iv) 2x2 ribbing. This picture is taken from Jo Creates, 2017, https://jo-creates.com.

3.4.2 Mathematics of knitting

Knitting is not just a handicraft hobby, it is also a form of complex architecture. This rudimentary technology can be used to engineer complex three-dimensional objects with tunable stretchiness.

Physicist Elisabetta Matsumoto shows that a knitted fabric is a type of artificially engineered material that acquires its properties from the designed structure and not from the base material [136]. By varying the stitch combinations, you can alter the elasticity, mechanical strength and three-dimensional structure of the textile. The resulting knitted fabrics show completely different properties, mainly increased elasticity, compared to their individual strands of yarn.

Knitting can also be associated with knot theory, a set of mathematical principles that define how knots form [137]. As a knot is considered a tangled circle, enclosed with crossings that are impossible to untangle, a knitted stitch is just an interlocking networking of slipknots, repeating themselves back to back. Knot theory has been used to explain how DNA folds and unfolds [138], showcasing the close association between nucleic acid assembly and knitting.

As a knitted textile is comprised of slipknots, pulling a single loose thread can unravel the entire garment. But if all ends are closed, each individual knit cannot be untangled. The overall knitted fabric can then be considered held together by a square lattice of stitches connected to one another horizontally. This pairwise linkage between stitches is the basis for ribbon knots, which forms a self-intersecting disk [136].With this system, knitting becomes a computationally intensive assembly method which comprises topology and entanglement.

3.4.3 Translating knitting into nucleic acid assembly

Looking back at ssRNA origami assembly discussed in Chapter 1, these RNA structures can contain stable hairpins that fold up in the order they are produced by the polymerase, and then each hairpin can condense in a sequential fashion dictated by long-range interactions to create a global fold. A caveat to these hairpin-hairpin interactions, is that they do not have strong thermodynamic stability, therefore they are not good structural components for a DNA or RNA nanostructure as bringing them out of there ideal thermal conditions can lead to unfolding. It is here, where knitting can provide a solution. If nucleic acids could assemble in a similar fashion to knitting, where each individual loop is not topologically knotted, but the addition of a new row of loops holds the previous row in place, the resulting structure would have improved overall structural stability.

Looking back at the knitted garment in Figure 3.1A, this structure is just a combination of loops, which potentially can be created with nucleic acids. In particular, the knit stitch (Figure 3.1B(i)) shows the folding pathway a nucleic acid strand would follow to form each loop. As the strand continues on the right side, it can loop into an existing loop through complementary interactions between both structures. This interaction falls in line with being a pseudoknot-like interaction [139], which normally consists of a double-hairpin structure with an extended quasi-continuous double helical stem region as shown in Figure 3.2. These interactions are formed when bases outside a hairpin structure pair with bases within the hairpin or internal loop, which perfectly describes the interaction needed for knitting. To form a knitted loop, this interaction is best represented by the Type H pseudoknot seen in Figure 3.2A. The incoming strand interacts with the hairpin loop, and in the knitting process this strand would continue pushing through and form a loop within the hairpin structure. If done repeatedly, rows of knitted nucleic acids can form similarly to a knitted fabric.

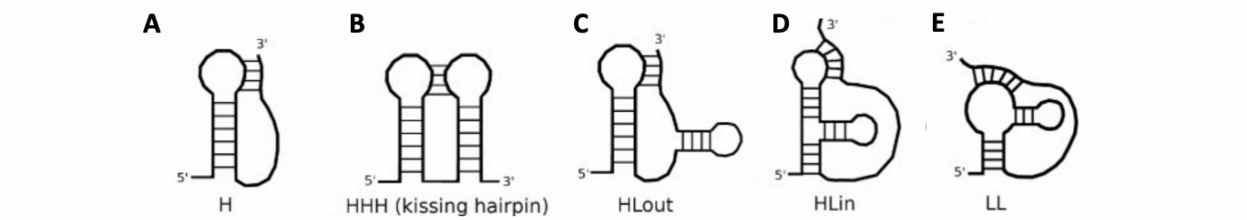


Figure 3.2: RNA pseudoknot types. This figure is adapted from [139]. *A) Type H, B) Type HHH, C) Type HLout, D) Type HLin, and E) Type LL pseudoknots. These pseudoknot types are from Pseuodobase++* [140] *classification.*

One of the main challenges of including pseudoknot-like interactions in a nucleic acid nanostructure design is that these secondary structures are difficult to predict computationally, as they are NP-hard for various energy models [141, 142]. There are a few programs that can be used to simulate the folding of these structures, such as KineFold [112] and coFold [115], but careful consideration is still needed in the design of the nucleic acid sequences. Additionally, topological constraints can be a consequence of looping as right-handed turns generated by the double-helix can produce left-handed turns elsewhere on the loop. There needs to be an interplay between this torsional strain and the thermodynamic stability of the interactions. Generally, a singular strand threading back through a loop of itself is entropically disfavoured. Sequence design can minimise these energetic barriers and enable the design of spontaneous knot formation [143].

Although there is not a physical needle knitting the nucleic acid nanostructures, the DNA or RNA strand can rely on folding in the direction and pattern of complementary sequences and related molecular interactions. When imaging folding nucleic acids in this fashion, you can start to see the similarities to knitting. With nucleic acids and their complex secondary and tertiary interactions, new stitch patterns not possible with just a needle and yarn can theoreticaly be achieved. Knitting as a single-stranded assembly for nucleic acid nanostructures has the potential to create the next generation of nucleic acid nanostructures.

3.5 Designing a Single DNA Knit Stitch

3.5.1 Initial design

To achieve a knitted structure with DNA, I start with constructing a single looped structure. This single knit design, termed the hairpin-thread (HT) model, consists of a single hairpin strand with a stem of 10 base pairs (bp), and a separate thread strand as shown in Figure 3.3A. The thread is designed to have dumbbell ends, made of parallel hairpins with stems of 8 bp and loops of 4 nucleotides (nts) each, to prevent the thread from completely pulling through the hairpin when looped in. The thread will then attach to a complementary sequence on the hairpin strand in its extended region, in what is termed the E domain (green), to bring both components closer. This domain is 20 bp long to ensure the thread does not spontaneously unbind from the hairpin. Finally, a complementary region between the thread and a part of the hairpin loop, called the C domain (orange), promotes the thread strand to loop into the hairpin through a pseudoknot-like interaction.

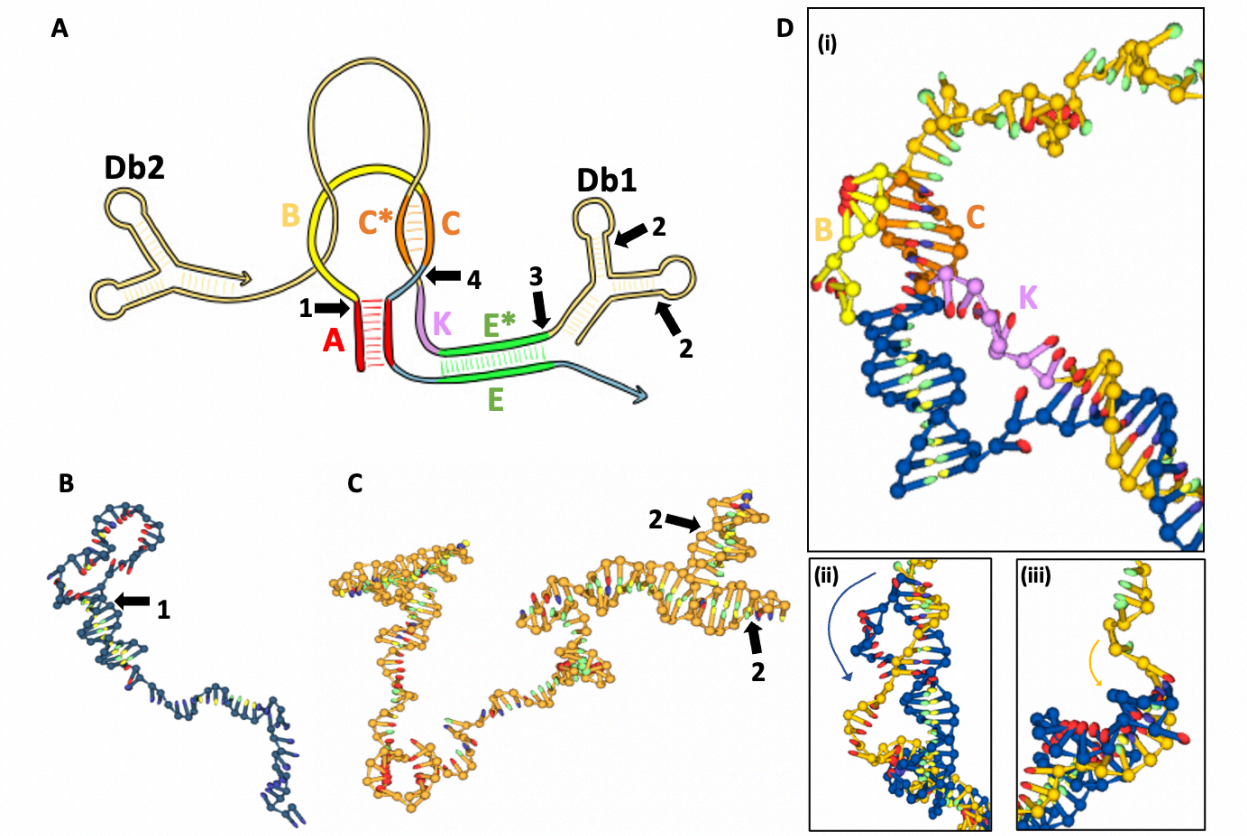


Figure 3.3: Designing a single knit DNA structure. A) schematic of the hairpin-thread (HT) model. In oxDNA these components are made separately, starting with the hairpin in B) where a mutual trap is placed at arrow 1 to promote folding, then the thread strand in C) is made where mutual traps are placed at both arrows 2 to promote folding of the dumbbell ends (Db1 and Db2). Next the thread is attached to the hairpin strand through the E domain by placing a mutual trap at arrow 3. Finally, looping is promoted through binding at the C domain by placing a mutual trap at arrow 4. D) Non-knitted structures in oxDNA. (i) The highlighted parameters, B, C and K, require optimisation to promote looping. (ii) The B domain needs to be longer to allow the thread strand to loop in and (iii) the C domain needs to twist the thread strand towards the hairpin more.

I used oxView to design the strands for the HT model and assembled then in oxDNA. To construct the hairpin strand (Figure 3.3B) from a linear sequence, a mutual trap — which is a harmonic force that at every moment pulls a particle towards a reference particle [107] — is added between the base pair that would be at the top of the hairpin stem. All the mutual traps here are indicated by black arrows in Figure 3.3A in the order they were applied. When the simulation is run with a trap stiffness of 0.1, the complementary sequences in domain A (red) are pulled together to form a hairpin stem of 10 base pairs. As for the thread strand (Figure 3.3C), the dumbbell ends are formed first in a separate simulation. Here the mutual traps are applied to the top of hairpin stems in the dumbbells. Once formed, the thread strand is attached to the hairpin strand by applying a mutual trap at the rightmost end of domain E, allowing for a 20 base pair complementary region to form. This interaction brings the strands closer together, so that adding a mutual trap at the base of domain C positions the thread strand ideally to loop into the hairpin.

To optimise this looping potential, a few domains need to be modified as shown in Figure 3.3D. The B domain (yellow) in the hairpin needs to be increased to provide more space for the thread to loop in (Figure 3.3D (i)). The K domain (purple) would also benefit from an increase in length, as a longer region here would allow for more flexibility of both the thread and hairpin strands. Most importantly, the C-C* domain needs to be optimised to better position the thread strand towards the hairpin (Figure 3.3D (ii)). The initial C-C* domain was created with 6 bp on the basis that a full turn in a DNA double helix is about 10.5 bp. Therefore, half a turn would in theory position the thread strand 180° from its original position around the hairpin strand allowing for it to enter the hairpin loop. This was not the case, as the oxDNA model in Figure 3.23(ii) shows that the thread strand (yellow) twists up and away from the hairpin when the C-C* domain is only 6 base pairs. Therefore this domain needs to be increased to bring the thread towards the hairpin through overtwisting.

3.5.2 Optimising looping with oxDNA simulations

My first optimisation steps were to increase domains B and K. In the initial HT model, the B and K domains were both 8 nts each. These lengths were arbitrarily chosen as they work approximately with the hairpin stem, which was designed to be 10 bp for stability. To optimise the B domain, various lengths were tested out ranging from 12-16 nts, as overviewed in Appendix Figure B.1. The 16 nt domain was the first to show spontaneous wrapping of the thread strand onto the hairpin, therefore this length was chosen for the B domain of subsequent designs. It is important to keep the B domain length at a minimum, as an increased hairpin loop can result in unwanted secondary structures, preventing knitting. To increase the space between the E-E* and C-C* domains, two more nucleotides were added to domain K to improve the flexibility of the thread strand, increasing this domain to 10 nts. Similar to the B domain, there is an interplay in balancing the length of the K domain and how flexible the overall structure is. For my single-knit design, I want the thread strand to loop into the hairpin without having to pull the overall hairpin closer towards the attachment site at the E-E* domain. This increase in the K domain should allow for more movement of the thread strand, and therefore provide more opportunity for the strand to bind to the hairpin loop. All subsequent models have these domain lengths unless otherwise specified.

My next step to improve the HT model was to increase the C-C* domain. As the initial 6 bp length was substantially undertwisted, I doubled the domain length to 12 bp, as well as tested out longer regions from 13-17 bp as shown in Appendix Figure B.2. Looping events occurred when the C-C* domain was 13, 15 or 17 nts as depicted by the pink backets in Figure 3.4A. A C-C* domain length of 15 nts was the most successful as five out of ten simulations were able to loop the thread into the hairpin strands within a fixed period of time (Appendix Figure B.3). Next, I analysed the number of steps in a simulation it took for the thread to loop and unloop from the hairpin when the C-C* domain was 15 nts. As shown in Figure 3.4B as a representative simulation, the thread strand is still not looped into the hairpin by 1.0e6 steps. The thread finally loops into the hairpin at 9.6e7 steps until it unloops again at 5.03e8 steps. This should provide enough time — as it takes more than 4 times the amount of steps to unloop than it took to loop — to trap the looped thread inside the hairpin using a lock strand as shown in Figure 3.5A. This lock strand strongly binds the thread strand by 20 base pairs at the M domain. The other end of the lock forms a similar dumbbell as the thread to prevent the thread strand from unlooping. An alternative lock system was designed using

two separate top and bottom lock strands as depicted in Appendix Figure B.4. Although this two lock system appeared to work, the single lock strand system was chosen for subsequent experiments simply for stoichiometry purposes as it added less strands into the overall reaction.

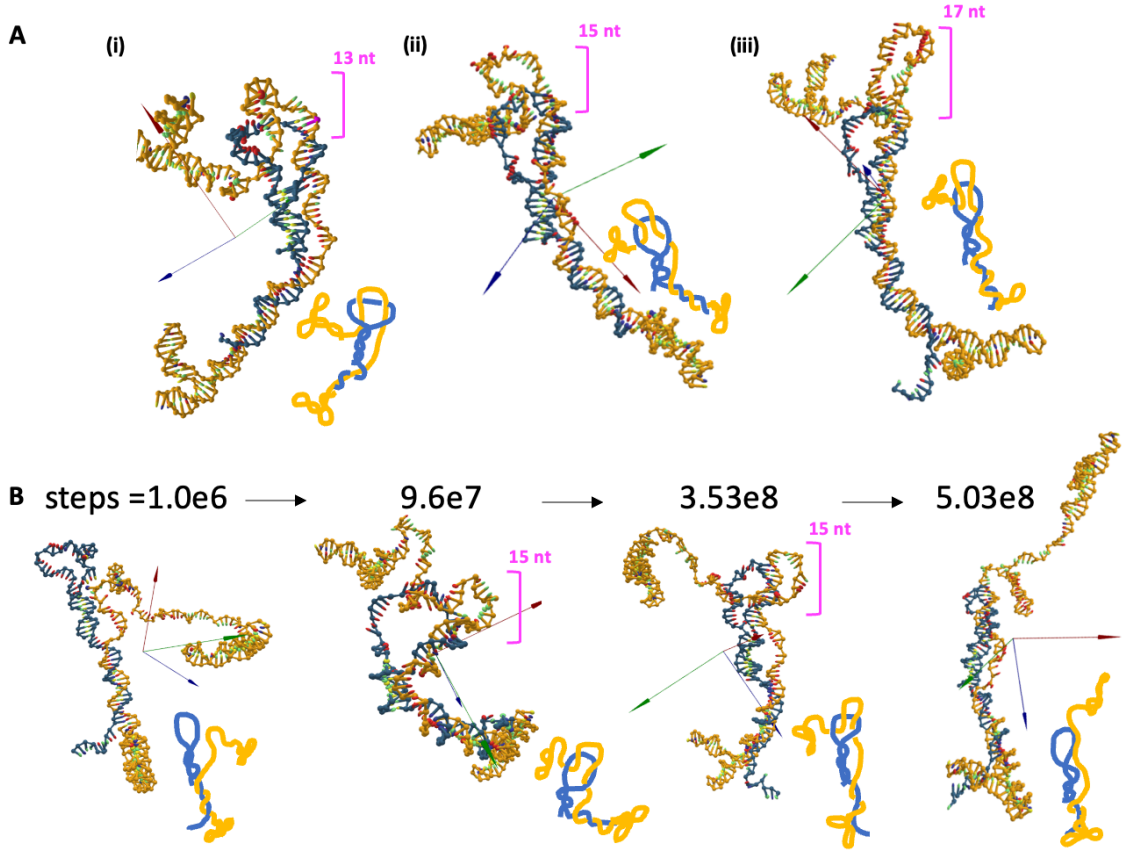


Figure 3.4: Optimising the C-C domain. A) Increasing the C-C* domain to (i) 13 nt, (ii) 15 nt and (iii) 17 nt. B) Timeline for looping and unlooping for a 15 nt C-C* domain model, starting from a time step count of 1.0e6, where it is unlooped, to it looping at 9.6e7 steps until it unloops at 5.03e8 steps. All models are made in oxDNA. Simplified schematics of each model are displayed on the righthand side. Regions of looping are indicated by the pink bracket.*

Also notable in Figure 3.5B is the appearance of the toehold displacement and invader strands. These strands are useful to determine whether the thread is correctly locked in the looped state, as the lock strand can bind to the thread strand, whether or not it is correctly looped. The toehold displacement strand binds to the 10 nt toehold, F domain, on the hairpin strand displacing the 20 base pair E-E* interaction between the hairpin and thread. The invader strand binds to a 10 nt region, L domain, on the thread strand and can displace the C-C* interaction between the hairpin and thread. With the introduction of these displacement strands, the thread is no longer bound to the hairpin strand in any regions. Therefore, the only way the hairpin and thread can still be associated together is if the thread strand is looped and lock into the hairpin, in a "knitted" state as depicted in Figure 3.6A. Alternatively, if the lock and invader strands are added to the thread first, before combining the thread and hairpin strands together, the thread can no longer loop into the hairpin. Therefore, when the toehold displacement strand is added in this "unknitted" reaction, as depicted in Figure 3.6B, then the thread and hairpin strands will fall apart. These reactions can be

distinguished by gel electrophoresis, which will be analysed in Chapter 4. The use of fluorophores will also aid the identification of each strand on a polyacrylamide gel. As shown in Figure 3.5B, a Cy5 fluorophore is added to the 5′ end of the hairpin strand, and a FAM fluorophore can be added to the 5′ end of the lock strand as well as the 3′ of the toehold displacement strand. These fluorophores were chosen as they do not overlap in wavelength with each, allowing for multiplexing when run on a polyacrylamide gel.

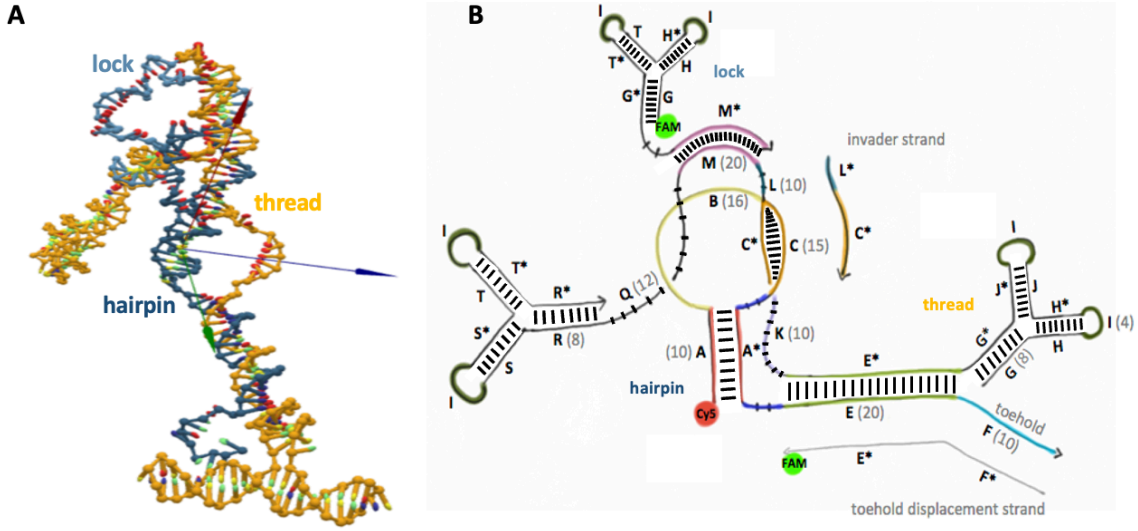


Figure 3.5: Optimised single knit DNA design. A) Hairpin-thread-lock (HTL) model in oxDNA and corresponding schematic in B) with toehold displacement strand, invader strand and associated fluorophores. The domains are indicated by letter and colour. The number of base pairs in each section is noted in grey.

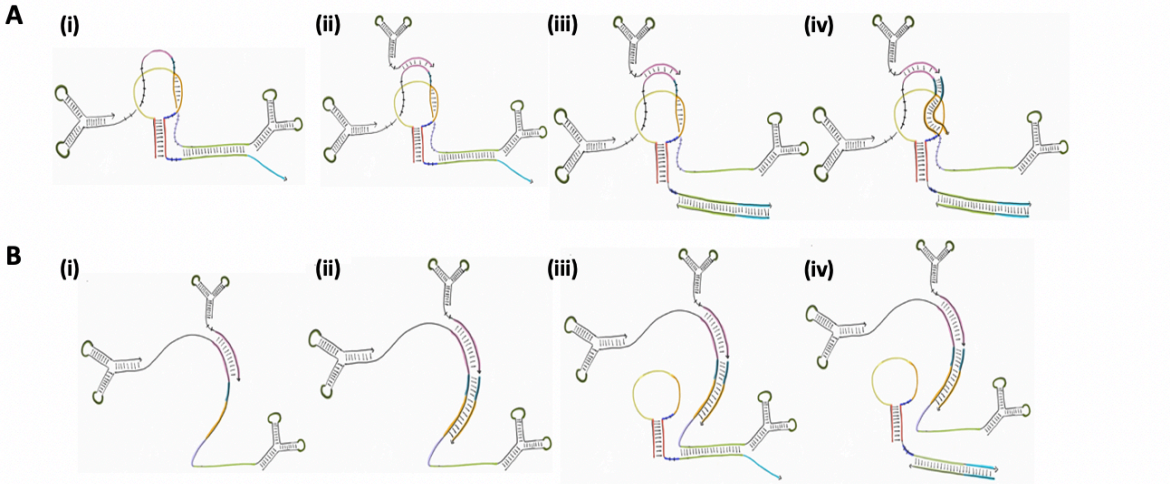


Figure 3.6: Schematic of the A) "knitted" and B) "unknitted" reactions. In A) (i) the thread is looped into the hairpin, then locked in looped state in (ii). The displacement of the E-E domain in (iii) by the toehold displacement strand, as well as the displacement of the C-C* domain in (iv) by the invader strand removes any binding regions between the hairpin and thread, therefore the only way these three components are associated together is if the thread is locked in place. In B) (i) the lock is added to the thread first, then in (ii) the invader strand is bound to the thread as well. This prevent the thread from looping into the hairpin in (iii), therefore when the toehold displacement strand is added in (iv), then the thread is no longer associated with the hairpin.*

3.6 Designing a Multi-knit DNA Nanostructure

3.6.1 Stacked and conjoining hairpins

With the creation of a single-knit DNA structure, the next step is to design a multi-knit system. There are various methods to creating a multi-knit DNA system as shown in Figure 3.7. By building off of the single knit design, you can get a stacked hairpin system (Figure 3.7A). The hairpin structure is a reasonable design to take up for each "knitted" strand as it is already in a looped structure. Additionally, you can add dumbbell ends to each "knitting" hairpin, to prevent them from pulling through the previous hairpin loop (Figure 3.7B). Alternatively, you can consecutively join individual hairpins, threads and lock strands (Figure 3.7C), which loop and lock into place in a similar fashion to the single knit system.

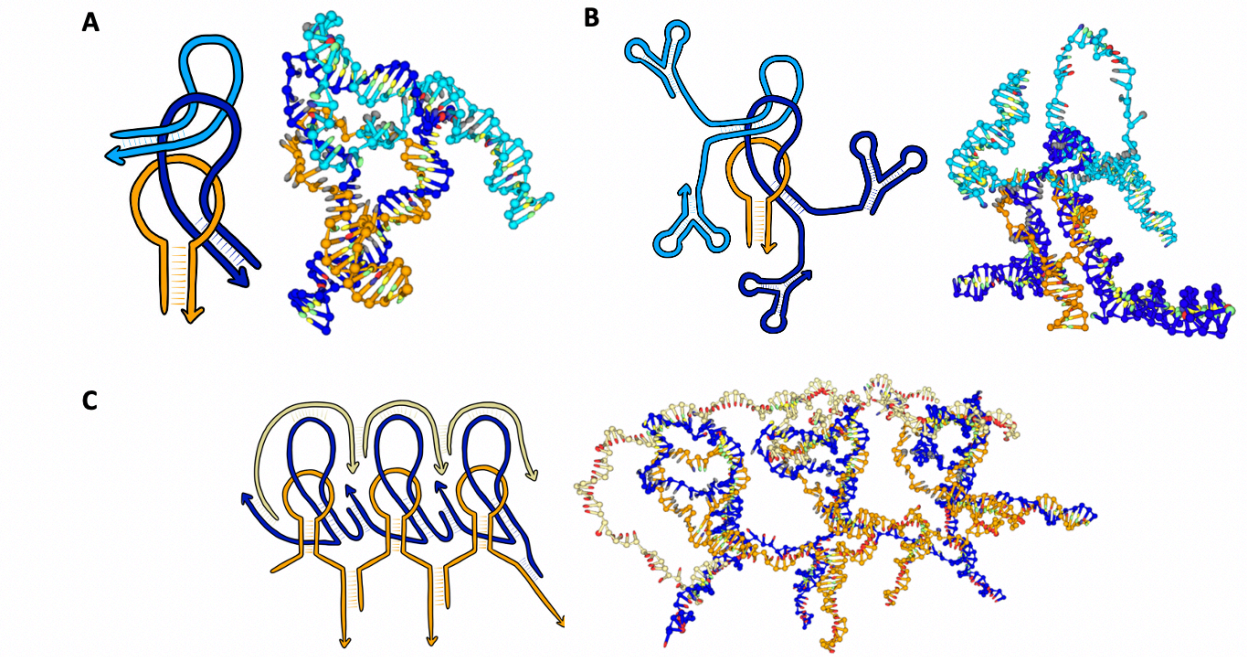


Figure 3.7: Various designs for a multi-knit DNA nanostructure. A) Three stacked hairpins, B) three stacked hairpins with dumbbell ends and C) 3x3 conjoining hairpins, threads and lock strands. The models are created in oxDNA, with a drawn schematic on the left for representation.

3.6.2 Unfolding dynamics of a simplified Plus-Minus multi-knit structure

To further investigate the characteristics and dynamics of a multi-knit structure, a simplified Plus-Minus (PM) system for a DNA knitted structure was created. The PM model is built off of a two loop motif (Figure 3.8A), where threaded loops are made by linking helical domains with minus (-) and plus (+) handedness so that the net crossing number is zero. A + helix was constructed by multiplying columns 1 and 4 of the oxDNA configuration (.dat) file of a - helix by -1. Single-stranded domains of 5 nt were used to introduce a half turn between the position where loops are closed and the position where adjacent loops are linked, to allow for easier tessellation of the multi-looped structure. In Figure 3.8B, the two-loop motif was duplicated to create eight rows of eight

loops, totaling 3320 nts. This tessellated structure was then relaxed using oxServe CPU settings as shown in Figure 3.8C.

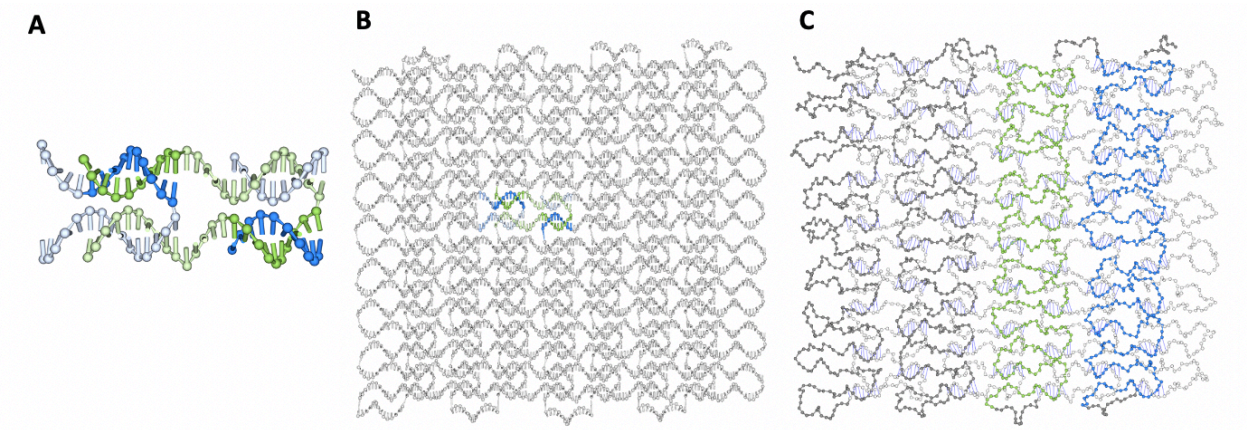


Figure 3.8: Constructing the PM (+/-) model in oxDNA. A) Two loop motif. One loop is indicated in green and the other loop is indicated in blue. The minus helix is shaded dark and the plus helix and other unpaired segments are shaded in light. B) Tessellated 8x8 two-loop motif. The two loop motif was duplicated to give 8 rows of 8 loops, totalling 3320 nts. An individual two-loop motif is highlighted in the centre. C) Relaxed model of the tessellated structure. The base pairs are omitted and harmonic traps used during relaxation are shown as solid lines. These models were constructed and relaxed in oxView, using oxServe CPU settings.

Knitting patterns can be described by the sequence of + and – helices at the base of the loops along the length of the template reading from 3′ to 5′. Whether a loop is threaded from above or below depends on the knitting pattern. Arbitrary patterns can then be created, with three particular patterns of interest to investigate as shown in Figure 3.9. One of the patterns created was the PMPM motif (Figure 3.9A(i)), which alternates + and – helices throughout, resulting in rows threaded over and under each other. Another pattern is the PMMP motif (Figure 3.9B(i)), which repeats a pattern of + - - + helices throughout, resulting in loops threading alternatively under and over in a chessboard pattern. The final pattern investigated is the PMPM_MPMP motif (Figure 3.9C(i)), which alternates rows of + - and - + helices, resulting in all loops threading from the same side. The mean structures for each pattern was calculated from the root mean square fluctuation (RMSF), which looks at individual residue flexibility during a simulation, as shown in part (ii) of Figure 3.9. It is interesting to note that all three patterns resulted in twisting of the overall structure around the center point. This is most likely due to torsional strain produced by the helices, which is evident as the motif that mirrored itself, PMPM_MPMP, was the most flat (Figure 3.9C(ii)) and all the loops threaded from the same side. This is closest to the garter stitch pattern seen in knitting (Figure 3.1C(iii)), which has all knit stitches in each row resulting in a flatter structure.

In addition to the overall structure for each pattern made, the unfolding dynamics were also investigated by applying an external force through harmonic mutual traps on opposite ends of the structure. Opposing mutual traps were applied to the 5′ and 3′ ends of the structure, with a force stiffness of 0.1 and rate of +/- 1e-6 (0.57 pN/nm and 08 1e-6 nm). The unfolding pathways are shown in part (iii) of Figure 3.9, where paired monomers are represented by regions of blue and unpaired monomers represented in green. The trajectories are read from left to right, and zoomed

in regions of interest are shown in the orange boxes. From these unfolding pathways, it can be seen that both PMPM and PMPM_MPMP unfold from both ends, whereas PMMP only unravels from one end. It is most likely arbitrary which end the structure unfolds from, but more trajectories are required to make a conclusive statement.

Force-extension curves for unfolding can also be obtained from the processed data as shown in part (iv) of Figure 3.9. These models look at the force measured as a function of the molecular extension when disrupting the stacking interactions during unfolding. Bigger steps in force are seen in the PMMP patterned model compared to the profiles for the PMPM or PMPM_MPMP models. This is likely due to the former pattern having two adjacent base paired - domains to require unfolding while the latter two patterns alternate +/- domains.

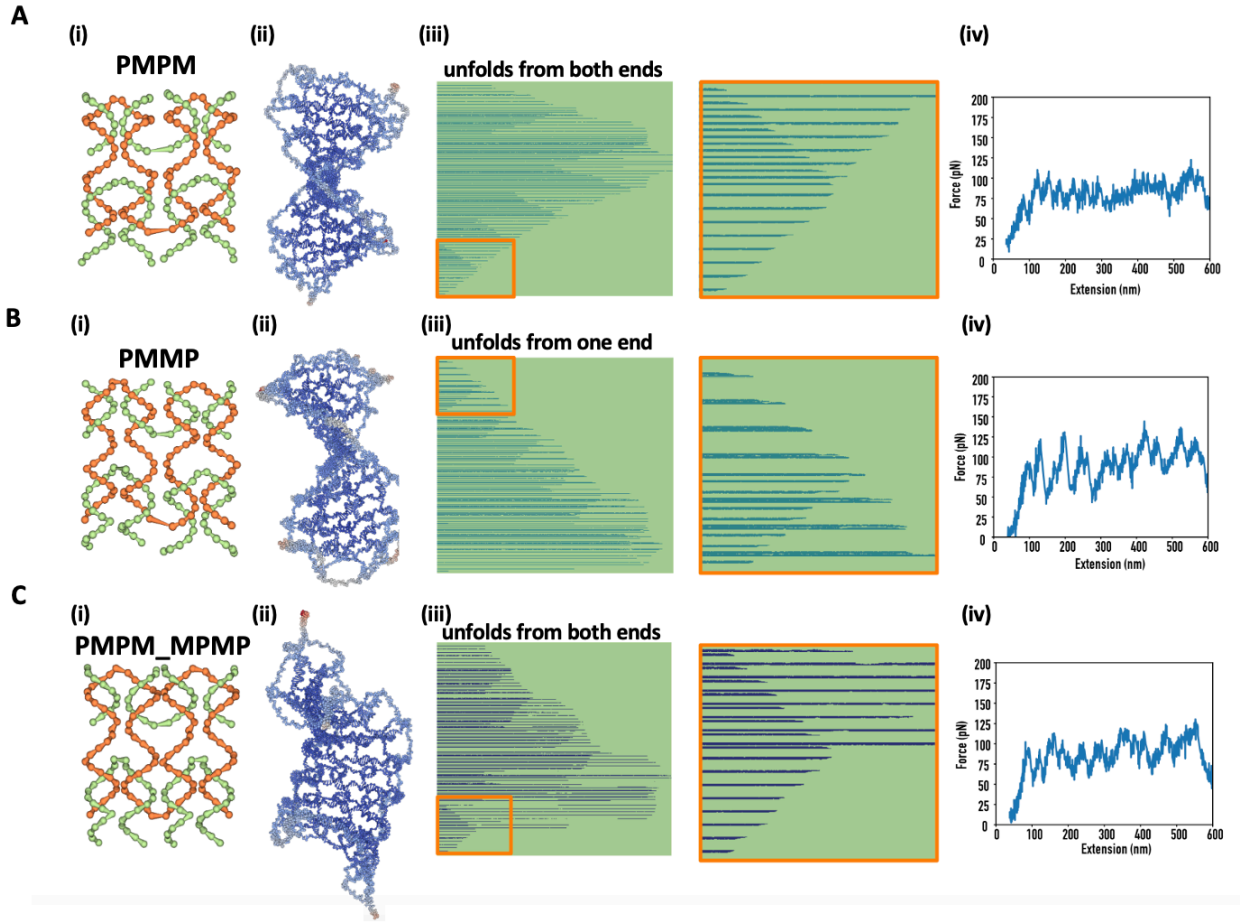


Figure 3.9: Simulating the unfolding dynamics of different PM motifs. A) PMPM model, B) PMMP model, and C) PMPM_MPMP model. For each model, (i) shows the motif, (ii) is the mean structure as calculated by RMSF, (iii) shows the order the bases become unpaired as the structure is unfolded with a zoomed in view of a section of interest as indicated by the orange box, and (iv) shows the force-extension curves for unfolding. In the unfolding pathways in (iii) the paired monomers are represented in blue and the unpaired monomers are represented in green, with the trajectory reading from left to right. The simulations were run on oxDNA and processed using relevant oxDNA packages.

The Plus-Minus system is a simplified model of a multi-knit DNA nanostructure that can output useful data on helical looping. As there is no assembly mechanism to exert control over threading direction, the exact knit patterns described for this system are likely impossible to recreate experimentally. By adding more sequence-specific interactions through hairpin structures and other motifs, can we create similar patterns with self-assembled DNA.

3.6.3 Designing a multi-knit DNA structure to self-assembly through RCA

The Plus-Minus multi-knit models were useful in studying the general trends of thread looping, but as they lack any sequence specificity nor incorporate any special motifs, a new design is required for the synthesis of DNA multi-knitted nanostructures. This new model builds off of the optimised parameters from the single-knit design, but also incorporates single-stranded assembly to more resemble its knitted namesake. The unimolecular nature of this structure also works in advantage of *de novo* assembly and has the potential to be folded intracellularly. To aid the assembly of the nanostructures, individual hairpins are programmed in to promote looping and the DNA template is designed for rolling circle amplification.

Before choosing specific sequences, I first checked if the exact parameters used in my single-knit model would work with a single-stranded multi-knit design. In oxView, as shown in Figure 3.10, I designed rows of three hairpins to loop into each other through mutual traps at the interaction domain, which is the equivalent to the C-C* domain of the single-knit model. To allow looping of consecutive rows, each hairpin loop contains two interaction domains, one to allow looping into an existing row and another to allowing looping of the next row of hairpins. With two interactions sites per hairpin loop, each domain was increased from 15 to 17 nts to improve stability, Additionally, two regions of 10 nt were added between the interaction domains to space out the looping. Linker regions (highlighted in red) were also added between each hairpin to allow flexibility of the incoming hairpin to loop into the previous one. As shown in Figure 3.10E, these parameters can form a relaxed 3 by 4 lattice of knitted hairpins.

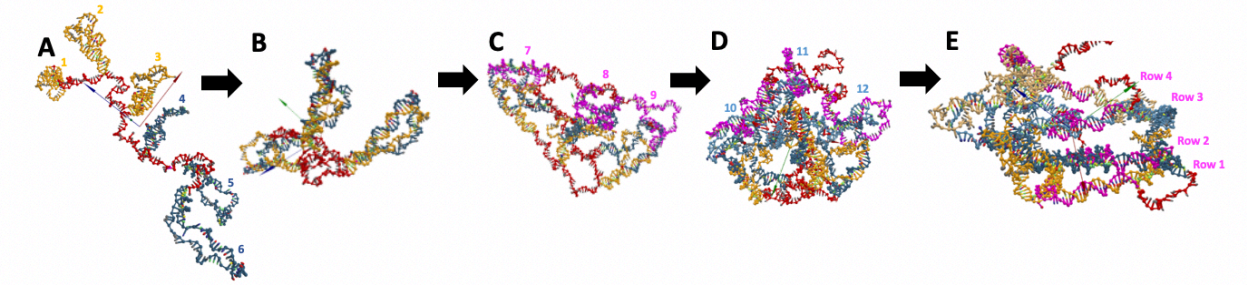


Figure 3.10: Looping the multi-knit structure in oxView. A) Six individual hairpins depicted in yellow and dark blue as a continuous strand. B) Mutual traps are added to the hairpins to result in two rows of three looped hairpins. Hairpin 4 loops into hairpin 3, hairpin 5 loops into hairpin 2 and hairpin 6 loops into hairpin 1. C) Three new hairpins in pink are added to create a third row. Hairpin 7 loops into hairpin 6, hairpin 8 loops in hairpin 5 and hairpin 8 loops into hairpin 4. D) Three more hairpins in light blue are added to create a fourth row. E) The four rows are relaxed in oxDNA, stretching out the overall structure. In the final structure, the rows are highlighted in pink. The linker regions between the hairpins are highlighted in red.

To incorporate rolling circle amplification (RCA) into my multi-knit system (Figure 3.11A), I aimed to make the template as minimal in length as possible as *in vitro* RCA normally utilises a short template strand around 80-200 nt, to allow for the continuous replication of a repeated sequence [34]. Due to the right-handed nature of DNA, the minimal model requires at least four rows to allow the repeated structure to continue knitting. Each row must alternate knit and purl stitches due to the right-handed nature of DNA, therefore two rows would not suffice as there would not be enough unique interaction domains. I included only two hairpins in each row to reduce the length of the template. This minimal repeating motif of eight hairpins is shown in Figure 3.11B, with four rows of two hairpins alternating looping from behind or from the front in each row. The beginning two hairpins form the base of the structure, then hairpins 3 and 4 loop from the back into hairpins 2 and 1, respectively. Due to the right-handed nature of the DNA duplex, the subsequent hairpins 5 and 6 will loop in from the front into hairpins 4 and 3, respectively. Finally, hairpins 7 and 8 continue to loop from the back into hairpins 6 and 5, respectively. If the template is repeated through rolling circle amplification, then hairpins 1 and 2 can continue the structure by looping from the front into hairpins 8 and 7, respectively. This multi-knit motif is rendered in oxView (Figure 3.11C) where the side profile (ii) displays the front and back weaving of the structure. These oxDNA models of the multi-knit design were formed by manual insertion and relaxation of the hairpin loops, unlike the oxDNA models of the single-knit design, which utilised mutual traps to form the knitted loops. Therefore, these models are only used to investigate the final shape characteristics, and do not represent the folding pathway.

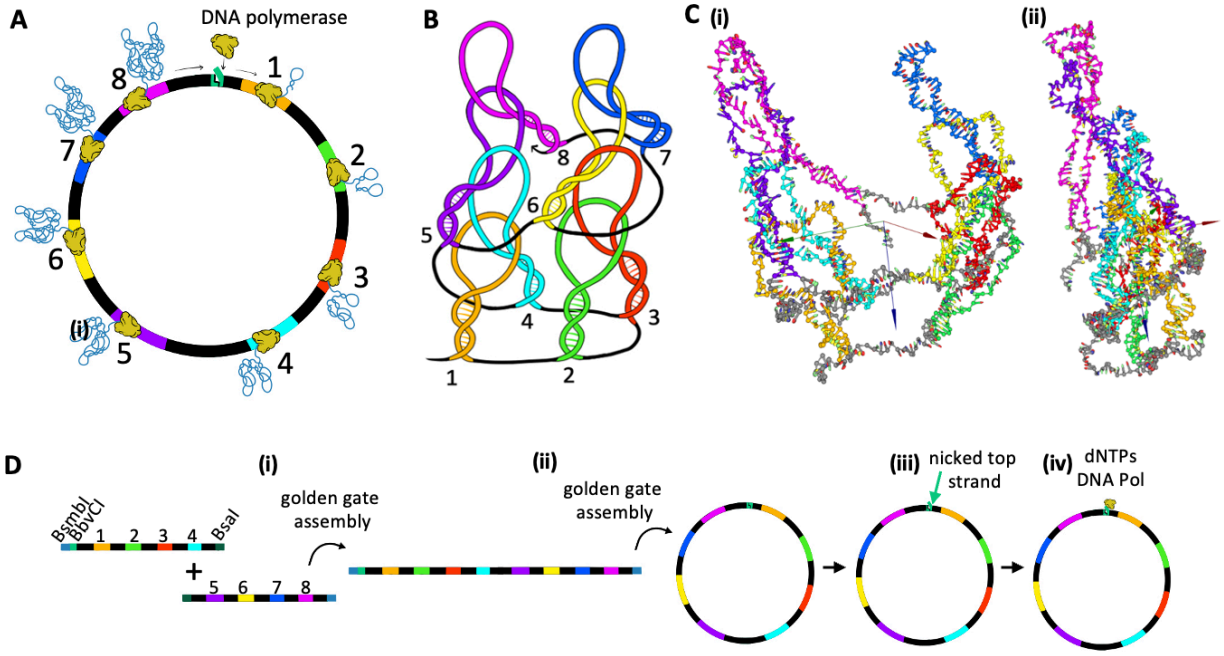


Figure 3.11: Overview of the multi-knit DNA nanostructure design. A) Schematic of rolling circle amplification of the multi-knit structure. B) Schematic of a single round of RCA of the multi-knit template, producing a motif consisting of 4 rows of 2 hairpins. C) The front (i) and side (ii) views of the multi-knit motif in oxDNA. D) Schematic of the molecular cloning process to assemble the two half templates to form a circular template for RCA. (i) The half templates are connected through the golden gate enzyme BsaI and then (ii) circularised using the golden gate enzyme BsmBI. (iii) The top strand of the circular template is nicked using the enzyme Nt.BbvCI and finally (iv) dNTPs and a high processivity DNA polymerase are added into the reaction to start RCA.

As mentioned in Section 3.3 of this chapter, there are challenges to synthesising DNA templates with repetitive and complex sequences. I encountered this issue in a few ways when attempting to order my template. Canonically, RCA templates are made from a short single strand of DNA circularised by a splint strand. Due to the length of my multi-knit structure, even when reduced to its minimal eight hairpin design, it was still over 800 nt, which is too long for most nucleic acid synthesis companies to make as a single strand. Therefore, I had to order my structure as a double-stranded DNA template. My second crisis came when attempting to order my template as a gBlock from IDT. Due to the repetitive sequences, IDT found my template too complex, therefore I had to split my template in half and incorporated recognition sites to reassemble the template in one piece as well as circularise it. The first half template is 415 bp, while the second half is 408 bp and together they form a 768 bp circle.

The molecular cloning process to create the circular RCA template is shown in Figure 3.11D. First the two half templates are joined together by the golden gate assembly enzyme, BsaI, then the linear strand is circularised by another golden gate assembly enzyme, BsmBI. These golden gate assembly enzymes are designed to cut out their recognition sites, therefore this assembly method is functionally scarless [144] and results in less sequences for the "knitting" structures to interact with. As the canonical single-stranded RCA template uses the splint strand as a primer for the polymerase, I can create an equivalent attachment point by using a nicking enzyme, which in this case was Nt. BbvCI. This nicking endonuclease only cuts the top strand of the double-stranded template. Finally, it is with the addition of dNTPs and a high processivity DNA polymerase that amplification can occur.

In addition to specific enzyme recognition sites, my RCA template utilises specific sequence domains generated from NUPACK to form the multi-knitted structure, many of which follow the parameters from my multi-knit model in Figure 3.10. As my minimal model contains eight hairpins, there are then sixteen domains — a sequence and its complement — to make the stems of each hairpin. These domains are 10 nt long to ensure stability of the hairpin, and are denoted in red in Figure 3.12. Within each hairpin is two 17 nt long interactive domains that allow for the pseudoknot-like interactions that promote knitting. These domains are colour-coded in Figure 3.12. Hairpin 1 contains domain A (dark purple) with its complement domain A* in hairpin 8, as well as domain B (yellow) with its complement in hairpin 4. Hairpin 2 includes domain C (pink) with its complement C* in hairpin 7, as well as domain D (orange) with its complement in hairpin 3. Additionally hairpin 3 incorporates domain E (turquoise) with its complement in hairpin 6. The final hairpin in template 1, hairpin 4, contains the aforementioned complement strand to domain B as well as domain F (blue) with its complement in hairpin 5. In template 2, hairpin 5 also has domain G (light purple) with its complement in hairpin 8. Next, hairpin 6 includes the previously mentioned complement to domain E as well as domain H (light green) with its complement in the incoming hairpin 7, which also the complement to domain C. And as mentioned, hairpin 8 has the complement sequences to domains G and A. Sandwiched between the two interaction domains as well as after the second domain in each hairpin are two 10 nt non-inactive domains (underlined), which provide separation between the binding domains as well as promote flexibility of the hairpin loop. Each hairpin loop also starts with a double-thymine spacer region to separate the first binding domain from the hairpin stem. Finally, between each hairpin is a 20 nt linker region (green) to provide flexibility to the structure to allow for the knitting to occur.

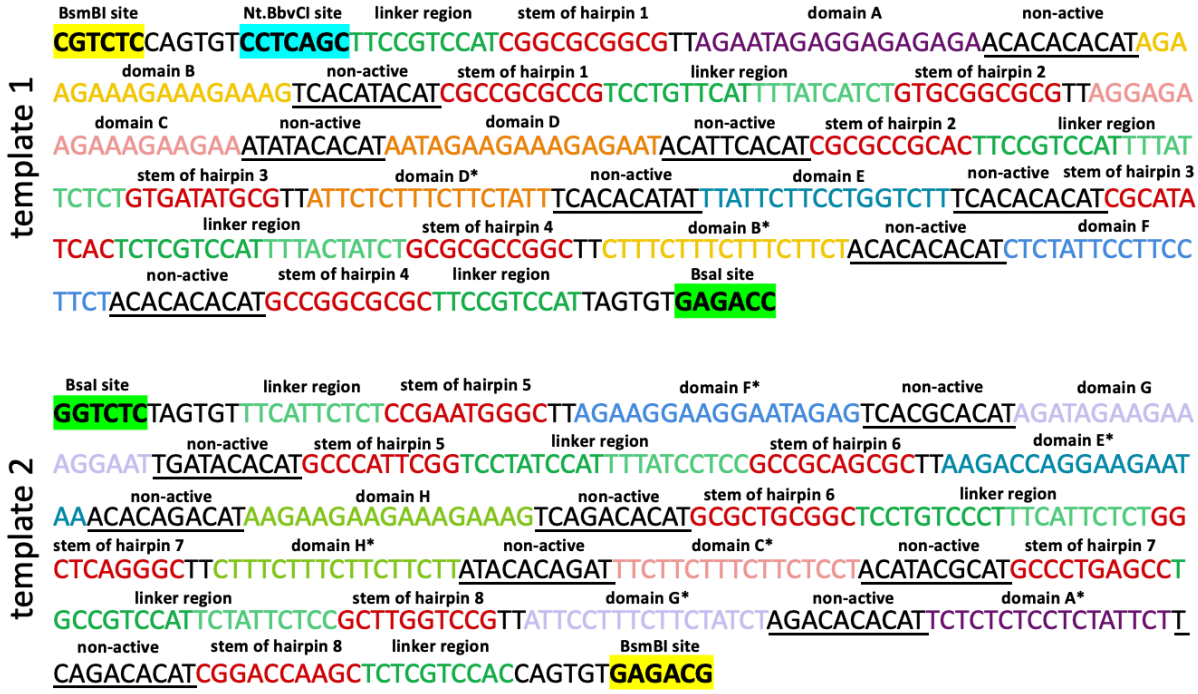


Figure 3.12: Annotated versions of the two half templates for multi-knit RCA. Depicted is just the top strand, in the 5'-3' direction, of the double-stranded template. The enzyme recognition sites are highlighted in bold. The two templates assemble together at the BsaI site bolded in green, which is removed in the cloning process. The longer template then circularises at the BsmBI site bolded in yellow, which is also removed in the cloning process. The Nt.BbvCI site bolded in blue remains and acts as the attachment point for the polymerase. Domains of 10 bases for the hairpin stems are coloured in red. 20 bases long linker regions between hairpins are coloured green. 10 bases long non-interacting regions are underlined. Finally, various binding domains of 17 bases are noted in specific colours. Matching complements utilise the same colours. A 2 thymine spacer is incorporated after the first stem domain of each hairpin.

To ensure the stability of each hairpin stem and the interaction domains, I analysed the melting temperatures of the sequences in NUPACK as shown in Figure 3.13. While the 10 bp hairpin stems had 50% of their bases unbound between the temperatures 45-70 °C, the longer 17 nt interaction domains were stronger, holding together until 85-90 °C. Although these are very high melting temperatures for the interaction domains, they might not reflect the actual stability of the structures. There may be torsional strain caused by overtwisting at the interaction domain, therefore the final structure might not have the full 17 base pairs bound for each domain.

The strong thermostability of these domains signifies that these interactions will not be transient, which ensures the maintenance of the entire knitted lattice. For future multi-knitted designs, shorter domains may allow for more transient binding that can be utilised for complex knitting patterns. In particular, it would be interesting to manipulate the looping patterns by modulating the annealing temperatures for each hairpin so that each row had different thermostabilities.

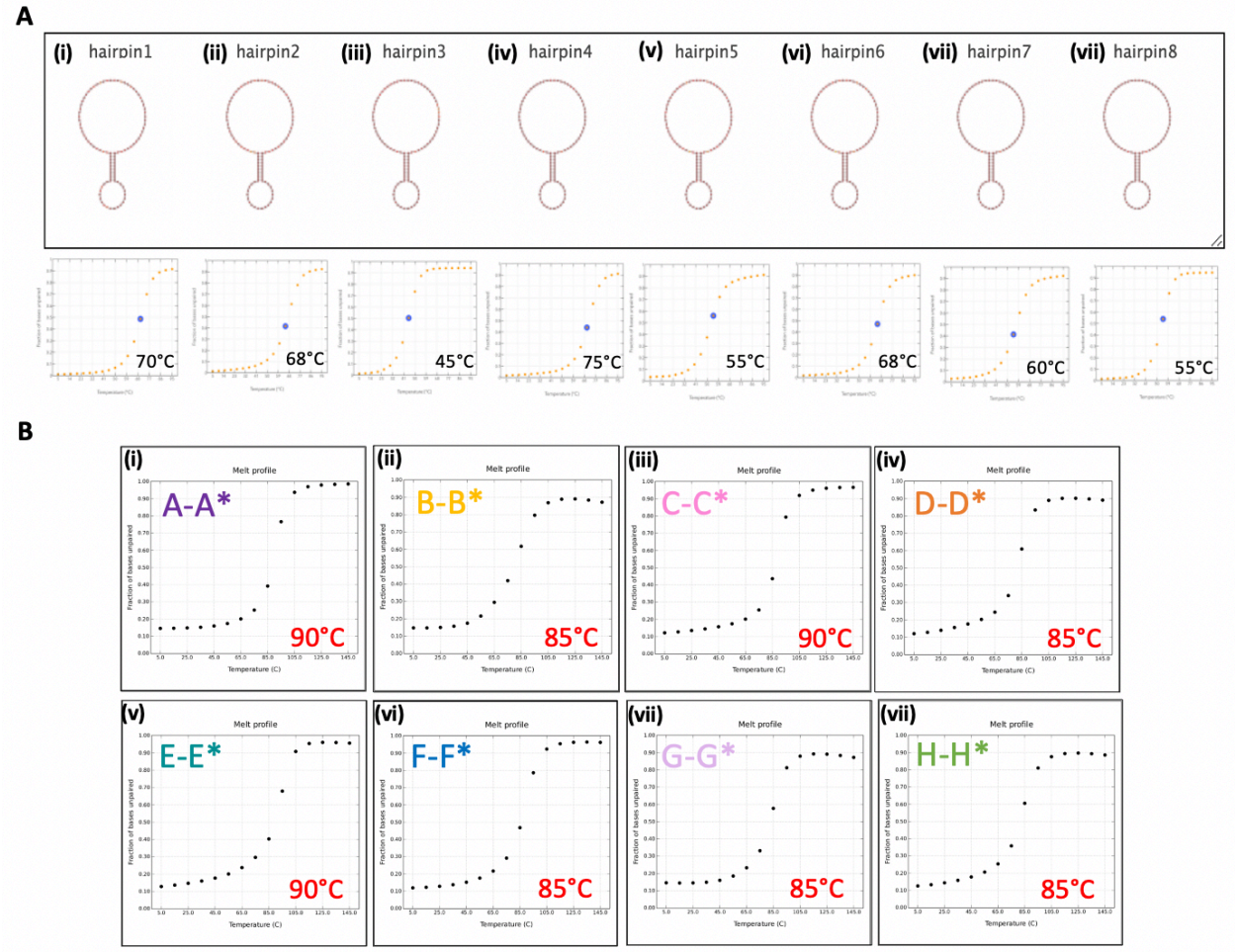


Figure 3.13: Melting curves for the multi-knit A) hairpins and their B) 17 bp interactions domains. In A) the representation of the whole hairpins are given on top. The melting curves are generated for their stem domains only for (i) hairpin 1, (ii) hairpin 2, (iii) hairpin 3, (iv) hairpin 4, (v) hairpin 5, (vi) hairpin 6, (vii) hairpin 7, and (viii) hairpin 8. In B) melting curves are generated for domains (i) A-A, (ii) B-B*, (iii) C-C*, (iv) D-D*, (v) E-E*, (vi) F-F*, (vii) G-G* and (viii) H-H*. For all graphs, the melting temperature where 50% of the base pairs are unbound, are written in C°. Hairpin schematics and melting plots are generated in NUPACK.*

In addition to analysing the melting curves of the sequences, I used KineFold to predict the folding pathway of the hairpins and their interaction domains at their lowest possible free-energy states. Due to the length of the sequences, I could only examine four hairpins at a time as KineFold has a limit of up to 400 nts per simulation. I split up the simulations to look at the folding and interactions between hairpin 1, 2, 3 and 4, between hairpins 5, 6, 7 and 8, between hairpins 3, 4, 5, and 6 and finally between hairpins 7, 8, 1, and 2 as shown in Figure 3.14. Each hairpin was formed correctly and the pseudoknot-like interactions —as represented by the coloured lines between the hairpin loops — were successful in all domains, except for between hairpin 1 and 8 which were shown to possess two additional interaction sites. This extra binding might prevent the looping of the next hairpin, which would be hairpin 4 into hairpin 1, if hairpin 1 is overtwisted. As long as the other

hairpin in that row is looped in correctly, the overall stability of the multi-knot structure should still be there, albeit more floppy.

Alongside the folding pathway, a helix tracing graph is plotted for each structure as depicted in part (ii) of each section in Figure 3.14. These graphs follow the lowest free-energy during the folding process plotted against time and monitors the five most stable helices. Additional helices may exist during the folding pathway. Although all the hairpins appear to follow similar folding pathways, the first hairpins in simulations C and D of Figure 3.14 which represent the helices for hairpins 3 and 7, respectively, form at lower free energy levels than the first hairpins in simulations A and B, therefore these structures might have more stable initial folding steps. It is interesting to note that the final helix in simulation B (depicted in dark blue), hairpin 8, varies in form until reaching its final structure. There could be various transient interactions occurring during this period. These interaction do not seem to affect looping of hairpin 1 into hairpin 8 as simulation D does not have any perturbations in the tracing of those helices. A similar transient binding and unbinding pattern occurs with the third helix (depicted in light blue) in simulation C. As hairpin 8 loops into hairpin 5, these transient interactions in hairpin 5 may be the cause of the additional interactions seen in hairpin 8.

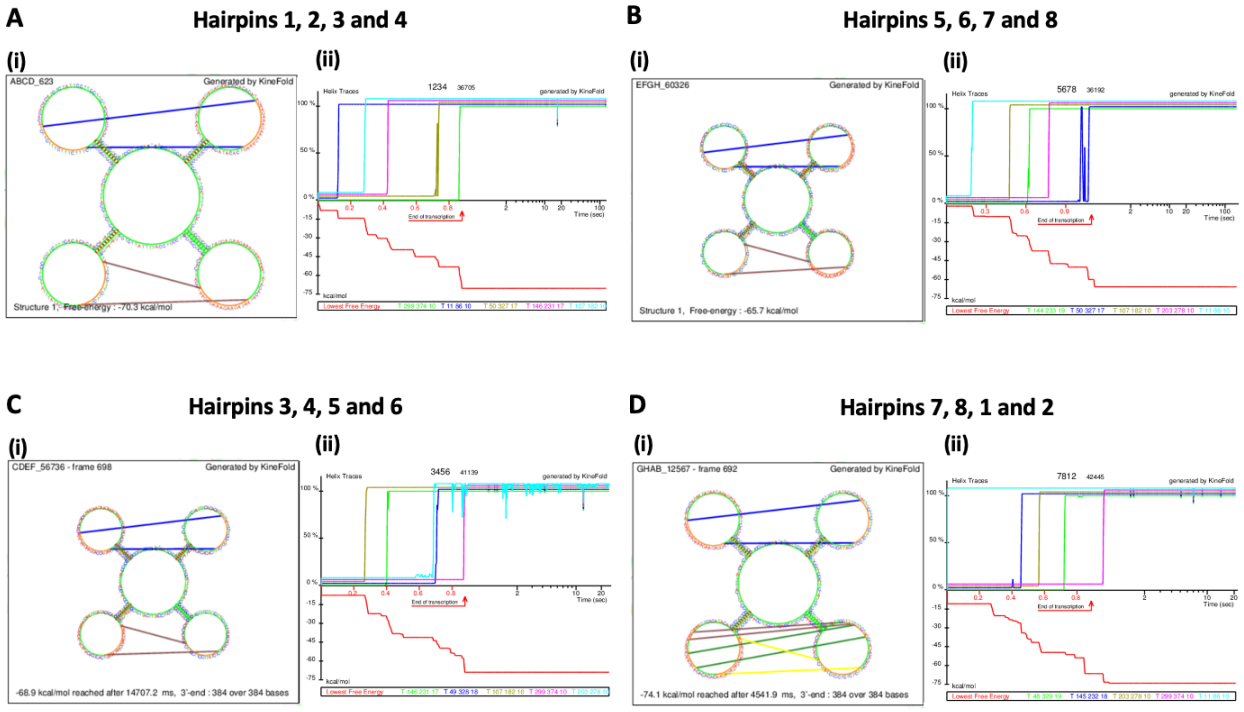


Figure 3.14: Folding pathways and helical tracing graphs of the hairpins in the multi-knit structure as predicted by KineFold. (i) Represents the lowest free-energy structures and (ii) shows the helix tracing graphs. Section A) depicts hairpins 1, 2, 3 and 4. Section B) depicts hairpins 5, 6, 7 and 8. Section C) depicts hairpins 3, 4, 5 and 6. Section D) depicts hairpins 7, 8, 1 and 2. The coloured lines in the folding pathway of part (i) indicate a pseudoknot-like interaction. The helix graphs in part (ii) are plotted to monitor which helices are formed during the lowest free-energy folding pathway. The helix presence is averaged over time and is displayed graphically between 0% (not formed) and 100% (always formed). The end of transcription is depicted by the red arrow. Note that the time-axis is first in linear scale during transcription and in log scale thereafter.

As depicted by the red arrows in Figure 3.14, the predicted speeds of transcription are similar between all hairpins. These KineFold simulations were all conducted with the polymerase speed of a new base added every 3 milliseconds, which closely represents the rate of the T7 RNA polymerase [145]. It would be interesting to see if a slower polymerase speed would result in similar folding pathways of the hairpins. Due to limitations in KineFold, preventing the analysis of more than two of my hairpin structures when using a slower polymerase parameter, I continued to use the faster polymerase speed for my initial designs.

Alongside KineFold, I used the CoFold to corroborate the folding pathway predictions. Canonically, CoFold is used to predict the secondary structures of RNA sequences, but I was able to utilise the program to depict the folding pathways of my multi-knit DNA sequences. Shown in Figure 3.15, an arc plot is generated using Turner thermodynamic parameters [116, 117] depicting eight hairpins interact in the same fashion as suggested by KineFold. A few interactions are missing, potentially due to limitations in overcrossing in the generated arc plot. In the plot, hairpin 1 interacts with hairpins 2 and 8, hairpin 2 interacts with hairpins 3 and 7, hairpin 3 interacts with hairpins 2 and 6, hairpin 4 interacts with hairpin 5, hairpin 5 interacts with hairpin 4, hairpin 6 interacts with hairpins 3 and 7, hairpin 7 interacts with hairpins 2 and 6, and hairpin 8 interacts with hairpin 1. One of the interactions not predicted in CoFold, is the one between hairpins 1 and 4. This was also suggested by KineFold, as hairpin 1 might have additional interactions with hairpin 8, prevent looping of hairpin 4 into hairpin 1. This is also seen with hairpins 5 and 8. Very few additional interactions are shown, suggesting that there are not many kinetic traps in the folding pathway.

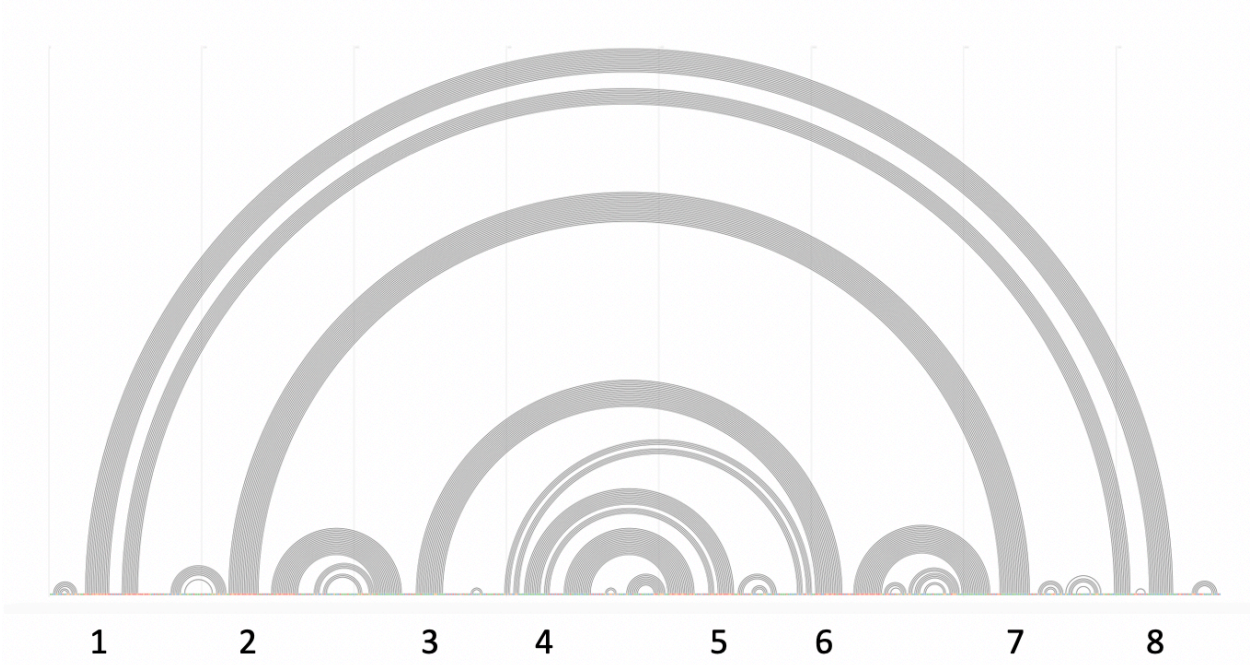


Figure 3.15: Arc plot predictions using Turner [116, 117] *thermodynamic parameters are used to predict the folding pathway of the multi-knitted DNA structure as generated by CoFold. Hairpin 1 interacts with hairpins 2 and 8. Hairpin 2 interacts with hairpins 3 and 7. Hairpin 3 interacts with hairpins 2 and 6. Hairpin 4 interacts with hairpin 5. Hairpin 5 interacts with hairpin 4. Hairpin 6 interacts with hairpins 3 and 7. Hairpin 7 interacts with hairpins 2 and 6. Hairpin 8 interacts with hairpin 1.*

In addition to the Turner thermodynamic parameters, CoFold also utilises Andronescu thermodynamic parameters [116, 117]. The difference between these parameters is that the Turner method only uses predictions made by free energy minimisation, while the Andronescu method uses both the available experimental free energy data in addition to structural data. An arc plot using the Andronescu parameters is depicted in Figure 3.16. In this plot, hairpin 1 interacts with hairpin 8, hairpin 2 interacts with hairpin 3, hairpin 4 interacts with hairpin 5, and hairpin 6 interacts with hairpin 7. These interactions correspond to the domains in the multi-knit model, albeit missing a few binding domains. In contrast to the Turner model generated arc plots, which did not have a lot of unwanted secondary structures, the Andronescu model generated arc plots appear to have additional interactions in between hairpin 2 and 3, as well as between hairpins 6 and 7. These structural motifs are likely to be within the hairpin loop regions, but it is difficult to interpret where else in the model these interactions may be associated with.

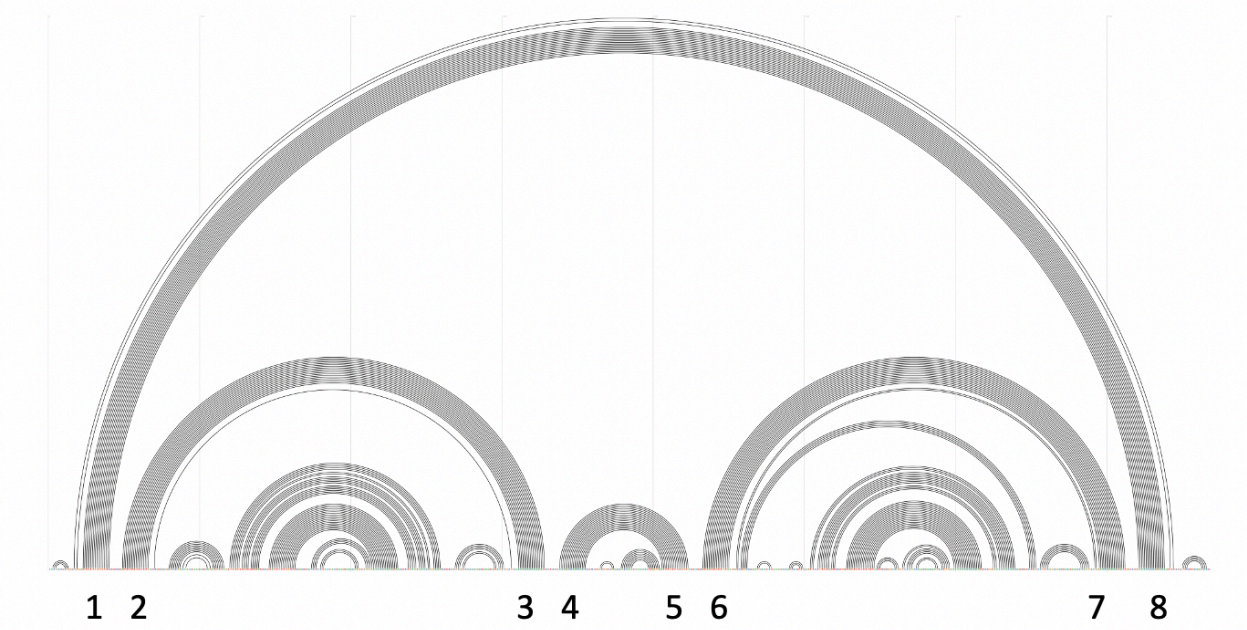


Figure 3.16: Arc plot predictions using Andronescu [116, 117] *thermodynamic parameters are used to predict the folding pathway of the multi-knitted DNA structure as generated by CoFold. Hairpin 1 interacts with hairpin 8. Hairpin 2 interacts with hairpin 3. Hairpin 3 interacts with hairpin 2. Hairpin 4 interacts with hairpin 5. Hairpin 5 interacts with hairpin 4. Hairpin 6 interacts with hairpin 7. Hairpin 7 interacts with hairpin 6. Hairpin 8 interacts with hairpin 1. Additional interactions exist between hairpins 2 and 3, as well as between 6 and 7.*

3.7 Designing a Single RNA Knit Stitch

3.7.1 Converting the DNA single-knit structure to RNA

Alongside designing knitted structures with DNA, I also explored RNA knitted designs since nucleic acid knitting as an assembly method is optimal for co-transcriptional folding inside the

cell. To see if the RNA strand would knit using the same parameters as my DNA knitting model, I converted my single-knit DNA design to a single-knit RNA one. To begin with, I combined the three separate hairpin, thread and lock strands into a single strand as shown in Figure 3.17A. With the help of Dr. Erik Benson, we modeled the structure in oxRNA (Figure 3.17B), which did not require any additional changes to the domain lengths. Finally I checked the folding pathways as predicted by KineFold (Figure 3.17C) and CoFold (Figure 3.17D), which both showed correct folding of the single-stranded structure with minimal to no unwanted kinetic traps. For the KineFold model, a single pseudo-knot like interaction is predicted exactly at the location of the knitting interaction that promotes looping. In the CoFold model, this model is also predicted as shown by the arc interaction highlighted by the *s. These characterisations show that the RNA single-knot model can likely be made using the DNA single-knit model parameters.

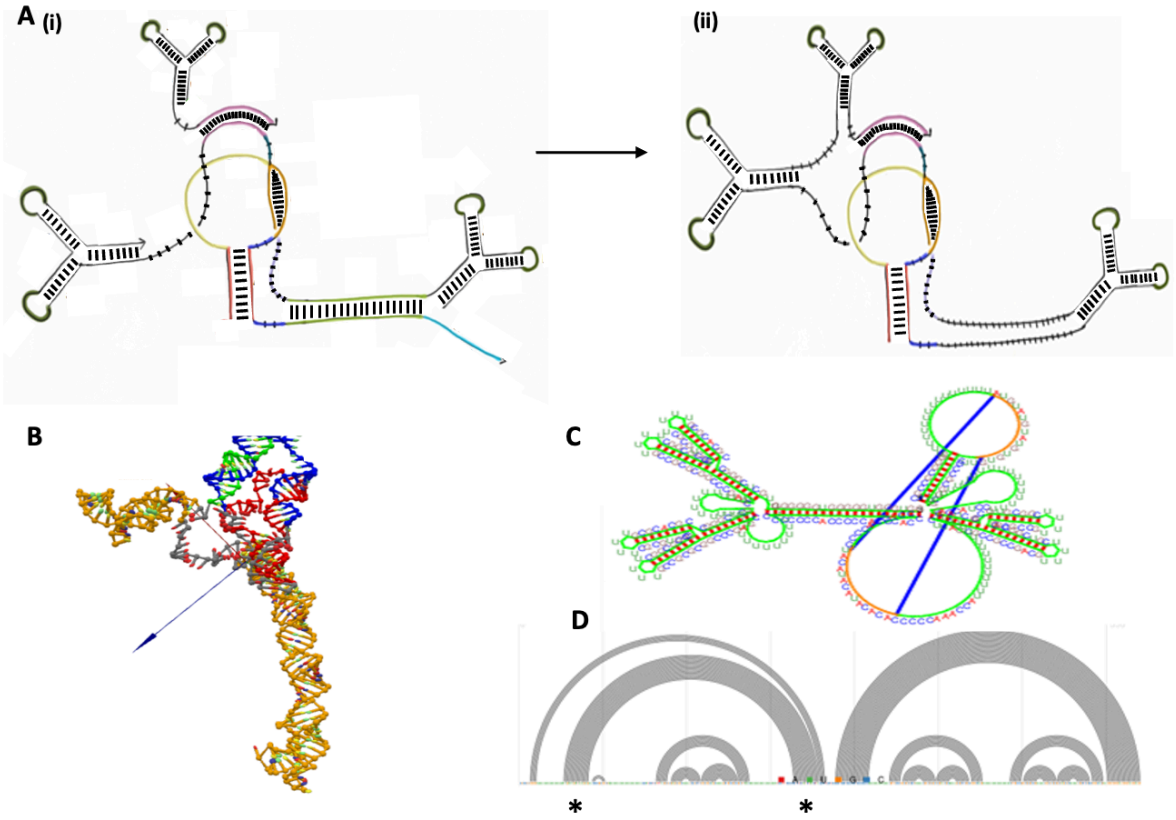


*Figure 3.17: Overview of the RNA single-knit model. A) Schematic displaying the conversion of the single-knit model from three separate components (i) to a single-stranded design (ii). Here the hairpin is attached to the thread as well as the lock strand. B) Single-stranded single-knit structure modeled in oxRNA. The hairpin region is coloured red, the thread regions are coloured in yellow and the lock region is coloured blue. Co-transcriptional folding predictions are generated in C) KineFold as a schematic and D) CoFold as an arc plot. In the KineFold model, there is a single pseudoknot-like interaction predicted exactly where the hairpin and thread interaction region is as depicted by the blue lines. The Co-Fold model also predicts this interaction as indicated by the *s. The Co-Fold plot is generated using Turner* [116, 117] *thermodynamic parameters.*

3.7.2 Converting the DNA multi-knit structure to RNA

Converting the multi-knit DNA structure to a RNA multi-knit design was not as successful as the single-knit model. As shown in a representative folding pathway depicted in Figure 3.18A, hairpins 1, 2, 3 and 4 do not fold correctly. The next four lowest free-energy structures predicted were all similar. In the helix graph shown in Figure 3.18B, the helical folding fluctuates consistently, representing the unwanted secondary interactions occurring in the structure. The sequence for the RNA folding pathway shown in Figure 3.18 was generated by converting the thymine bases in the multi-knit DNA design to uracil. Due to the complex interactions of RNA, additional optimisation of the sequence is required to form a successful multi-knit structure out of single long RNA strand. This step is out of the scope of my thesis as I would need to remodel the whole design, but this molecular architecture would pose as an interesting topic for future research.

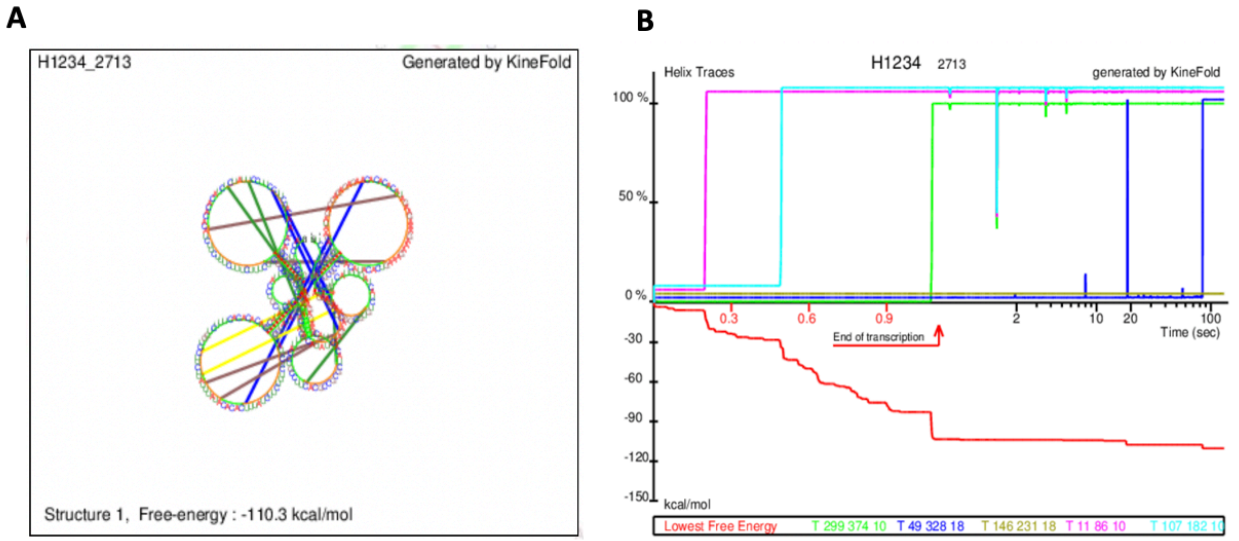


Figure 3.18: A) Folding pathway and B) helical graph for hairpins 1, 2, 3 and 4 in the multi-knit RNA structure as generated by KineFold. A) The hairpin structures are trapped in complex secondary structures as represented by the various crossover lines between the hairpins. In B) the helical traces fluctuate, which may represent unwanted binding events. The simulation was run in batch mode to generate the full folding pathway.

3.8 Chapter Summary and Outlook

In this Chapter, I went through the topic of knitting and explained how this technique can be used to assemble nucleic acid nanostructures. I then showed the designs for both single-knit and multi-knit DNA structures. In the single knit model, three separate components come together in a specific order to produce a knitted or non-knitted reaction. In the multi-knitted model, the assembly is designed to occur through single-stranded folding. I used various software to optimise the design of the structures as well as to predict their potential folding pathways.

For my single-knit HTL model, I first showed how I optimised the structural parameters in oxView and oxDNA. This included the size of the hairpin loop and the length of the interaction domain between the hairpin and thread. Additionally, I characterised the looping dynamics in short oxDNA simulations to ensure that the interaction was consistent and stable. Once the design is optimised, I detail how the system can achieve either a knitted or non-knitted structure based on the order each component is added into the reaction. I added additional components to aid analyse of these reactions by gel electrophoresis.

As for my multi-knitted structure, I showcase various design options from stacked hairpins to conjoining hairpins. To achieve the most optimal design for *in vivo* assembly, I created a single-stranded multi-knitted design. Although the design is simple in terms of containing only a single molecule, the kinetics of the folding pathway pose a challenge in predicting the secondary structures. With the help of Dr. Jonathan Bath, we have modelled the kinetics of a simple Plus-Minus multi-knitted structure. These simulations showed that a knitted structure with each row containing loops that form from the same direction (PMPM_MPMP), resulted in the least twisting of the overall lattice. This falls in line with the concept of garter stitches in a knitted fabric, which also utilises the same looping in each row and creates a knitted garment that curls less than a knitted structure that alternates stitch patterns. This simplified design was then given sequence specificity and specialised for rolling circle amplification. In a similar manner as the single-knit model, this multi-knit model was optimised in oxView and oxDNA and given appropriate sequences as generated by NUPACK. The minimal multi-knit design is too long to synthesise as a single-strand, as canonically used for RCA. Instead, it was synthesised as two double-stranded DNA templates. I then overview the molecular assembly methods that can be used to circularise this template. The successful folding pathways of the hairpins in the multi-knit structure are predicted in KineFold and CoFold. Finally, I attempt to convert the DNA knitted designs to RNA based structures. The single-knit model was successful in adopting RNA parameters, but the single-stranded multi-knit RNA design requires further optimisation. The next step is to analyse these models experimentally.

Chapter 4

Analysing the DNA Knitted Structures

4.1 Introduction

In this Chapter, I will analyse the designed sequences showcased in the previous chapter for my single-knit and multi-knit DNA structures. Analysis will be conducted by gel electrophoresis for both structures, and atomic force microscopy for my multi-knitted design. For the single-knit design, the HTL model, I optimised the signal of the strands on polyacrylamide gels characterised each control strand. I then investigate the knitted and unknitted reactions, screening for all the DNA structures with SYBRGold as well as utilising strategic fluorescent strands. For my multi-knit design, I investigate the circularisation of the two half templates and show the polymerised product on a polyacrylamide gel. To elucidate more information of the folded structure, I utilise atomic force microscopy to image the topology of the multi-knitted DNA structures. Finally, I finish with a discussion on what improvements can be made to the system as well as future experiments to confirm the exact looping pattern of the multi-knitted structure.

4.2 Author Contributions

All gel electrophoresis experiments were conducted by Catherine Fan. Initial AFM imaging of the constructs was carried out by Catherine Fan under the guidance of Dr. Casey Adam (Contera group). All remaining AFM, and subsequent TopoStats analysis, was carried out by Catherine Fan under the guidance of Dr. Tom Catley (Pyne group) as part of the Royce Partner Equipment Access Schemes.

4.3 Analysing the Single DNA Knit Stitch Structure

4.3.1 Controls for the HTL structure

To find the optimal electrophoresis conditions to run my samples, I tested control strands on 15% and 10% acrylamide gels, running them for 30 and 45 minutes as depicted in Appendix Figure B.5. I found that the 10% gels completed after 45 minutes using 90 V were ideal for visualising the separation of the distinct bands. Additionally, a short period of quenching — heating the sample up to 95 °C for 30 seconds and then cooling on ice for 30 seconds — for the hairpin, thread and lock strands improved the signal of the bands and reduced well aggregation. In particular, quenching the thread in 10 mM $MgCl_2$ showed improved resolution of the thread band as well as the hairpin and thread combination (Appendix Figure B.6). As shown in Figure 4.1, the hairpin, thread and lock strands can be visualised in wells 1, 2 and 3, respectively. Controls for the toehold displacement strand, invader strand 1, invader strand 2, invader strand 3 and invader strand 4 are shown in wells 9-13, respectively. When combined together the hairpin and thread strands form a

band shifted higher than either individual bands, as indicated by the purple arrow in well 4. In comparison, when the thread and lock are combined, as seen in well 8, the band is at a similar level to the band of just the thread strand. As the kinetics of Watson-Crick base pairing occurs rapidly [146] the strands are only incubated together for 10-15 minutes. When the hairpin, thread and lock are added together in that order, the resulting structure runs higher in well 5. This band is more obvious when combined with the toehold displacement strand in well 6 as indicated by the pink arrow. You can notice a faint band where the red arrow is in well 6, which may represent unknitted hairpins bound to the toehold displacement strand. Finally when the hairpin, thread, lock, and toehold displacement strand are combined in that order with invader strand 1 — hence the hairpin and thread are only associated together if they are knitted— there is a faint band remaining as indicated by the orange arrow that might represent the single-knitted structures. The yield of knitted structures is quite low as this band is much fainter than the bands below indicating that the non-knitted structures were predominant in this reaction, resulting in separation of the components upon displacement.

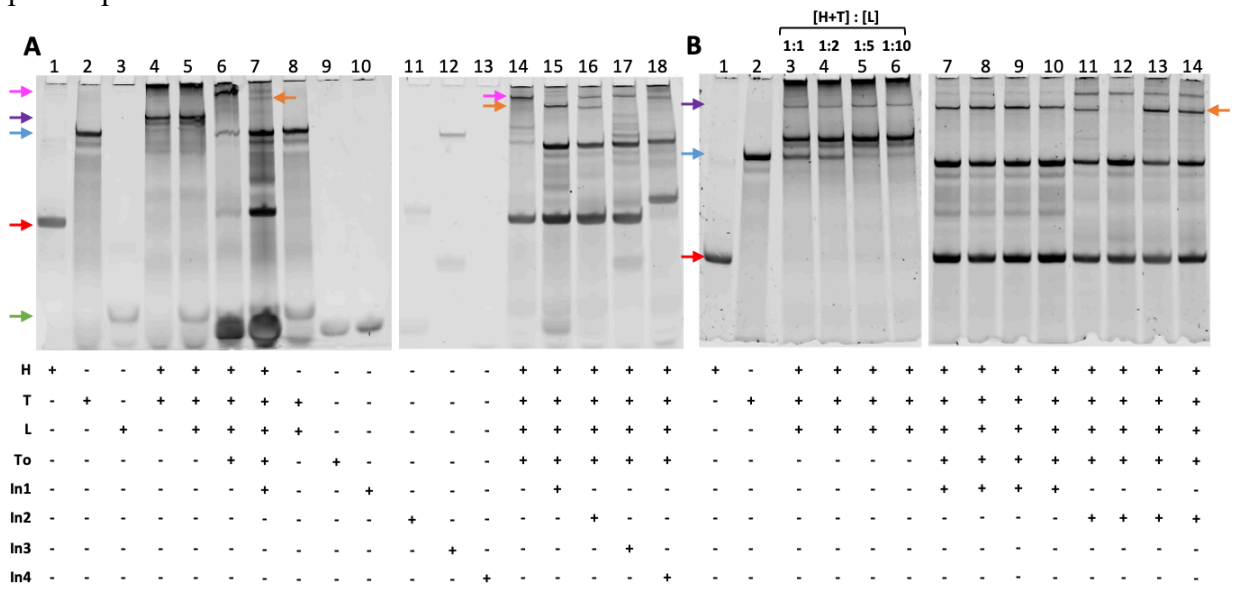


Figure 4.1: Control strands for the HTL single-knit structure. A) Individual and combined strands. Lanes 1, 2, 3, 9 and 10 contain the hairpin, thread, lock, toehold displacement strand and invader strand, respectively. Lanes 4, 5, 6, 7 and 8 contain combinations of the hairpin plus thread, hairpin plus thread plus lock, hairpin plus thread plus lock plus displacement strand, and hairpin plus thread plus lock plus displacement strand plus the invader strand. B) Titration of the lock strand when combined with the hairpin and thread. Wells 3-6 test varying ratios of the lock strand in comparison to the hairpin and thread combined ranging from 1:1, 1:2, 1:5 and 1:10. Wells 7-10 utilise the samples from wells 3-6 with the addition of invader strand 1. Wells 11-14 also utilise the samples from wells 3-6 with the addition of invader strand 2. The band for the lock strand is indicated by the green arrow. The band for the hairpin strand is indicated by the red arrow. The main band for the thread strand is indicated by the blue arrow. The purple arrows indicate the hypothesised main band for the hairpin and thread strands combined. The pink arrows indicate the hypothesised main band for the hairpin, thread, lock and toehold displacement strands combined. The orange arrows indicate the hypothesised main band for the knitted structure. The components in each lane is indicated by a + as listed below the 10% polyacrylamide gels.

In wells 15-18 of Figure 4.1, I compare the use of different invader strands — whose displacement interactions are overviewed in Appendix Figure B.7 — to displace the C-C* domain between the hairpin and thread strands. Well 15 shows the use of invader strand 1, which shifts the hairpin-thread-lock-toehold band in well 14 (pink arrow) down as indicated by the orange arrow. In comparison, the use of invader strand 3 in well 17 and invader strand 4 in well 18 show the same upper band as well 14. The displaced hairpin strand is a bit higher in well 18 as invader strand 4 is the longest and would make the hairpin-invader structure larger, and therefore run slower in the gel. In well 16, where invader strand 2 was used, both the pink and orange arrow bands appear. I believe the lower of these two bands (the orange arrow) represents the knitted structure as the displaced E-E* and C-C* domains would make the thread more flexible while running through the gel, therefore there should be a change in band shift from the size of hairpin-thread-lock-toehold structure. I continue examining the HTL model with invaders 1 and 2, as both have the shifted lower band.

In Figure 4.1B, I investigate the concentration of lock strand needed in comparison to the concentration of the hairpin and thread combined together. Once again, wells 1 and 2 show the control hairpin and thread strands, respectively. Wells 3-6 have increasing ratios of the hairpin and threads strands to varying lock strand amounts. It appears that with more lock strands, the more the thread strands are trapped with the hairpin strands as the blue arrow band decreases in signal with increased lock concentration. Wells 7-10 and 11-14 take the samples from wells 3-6 and add in the toehold displacement strand as well as either invader 1 for the first four wells or invader 2 for the last four wells. Although there are increased concentrations of lock, there isn't a difference with the band signal of the potential knitted structure. Well 12 seems to be slightly different from the other invader 2 displaced structures, but this may just be an outlier. With this information, I chose to continue on with using a lock strand ratio of two to one with the hairpin-thread combination, and chose to use invader strand 2 as the displacement strand for the C-C* domain. These conditions seemed to provide the best results of the knitted structure, but to elucidate more, I will use fluorescent control strands (on the hairpin and lock) to confirm which bands have which strands when analysed on a polyacrylamide gel.

As shown in Figure 3.5B, the hairpin strand is modified to have a 5′ Cy5 molecule and the lock strand is modified to have a 5′ FAM molecule. I examined these strands in the HTL model in Figure 4.2, where gel A displays the SYBRGold signal (black/white) as representative of all DNA strands, gel B shows the Cy5 signal (red) representing all structures associated with hairpin strands, and gel C represents all structures associated with the lock strands using a FAM signal (green). In this figure, I once again explore varying ratios of combined hairpin-thread samples to different concentrations of lock strand. I also investigated a decreased lock concentration of half of the hairpin-thread combined concentration in well 5 of all three gels, which do not have strong signals of the HTL band (purple arrow). It appears again that a ratio of one to two between the hairpin-thread and the lock strands is best, as almost all of the hairpin-thread combinations are associated with the lock strands. This is compared to the 1:5 ratio in the eighth wells of the gels, which creates a new lower band as indicated by the dark green arrow. An interesting aspect to note in these gels, as well as in the previous gels, is the aggregation in the wells highlighted by the blue arrow in the gels of Figure 4.2.

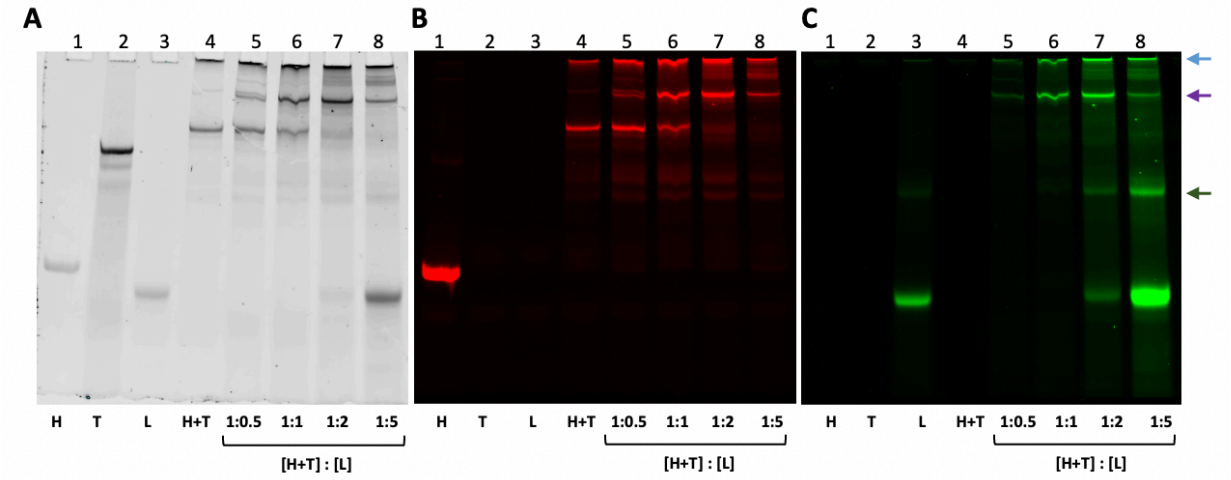


Figure 4.2: Titrating lock strand with the hairpin and thread. All three 10% polyacrylamide gels show the same samples under different filters: A) SYBRGold, B) Cy5 and C) FAM. Wells 1-3 show the control strands for the hairpin, thread and lock, respectively. Well 4 combines the hairpin with the thread. Wells 5-8 compare different ratios of the hairpin-thread to the lock strands. The purple arrow indicates the potential knitted structure. The blue arrow indicates the aggregation of sample in the wells. The dark green arrow indicates a new band created by excess lock strands.

4.3.2 Relieving the well aggregation

Although quenching the control strands has improved the overall signal of each band, there is still substantial aggregated signal in the wells of the polyacrylamide gels for my HTL structure. This aggregation was the most for the lock strands. To relieve this aggregation, I investigated quenching the lock strand in different concentrations of $MgCl_2$, as well as looked at the quenching speed in these buffers (Figure 4.3). Mg^{2+} or other cations are commonly used during the annealing of DNA to screen the negative charges on the nucleic acid backbone, allowing the strand to bind together. With DNA origami structures, the concentration of $MgCl_2$ used is normally around 12µM. For my buffer test shown in Figure 4.3A, I used either 10 mM (wells 1-4) or 100 mM (wells 5-7) $MgCl_2$ or I used MilliQ water containing no cations at all (wells 8-9). Although the different concentrations of $MgCl_2$ had varying success with relieving the well aggregation, strands quenched in MilliQ water had the most consistent success with the least amount of aggregation in the wells. This may be due to the fact that the lock strand is shorter compared to the hairpin and thread strands. Additionally, cations may cause too much condensing of the DNA leading to accumulation in the wells. Although wells 8 and 9 had a similar amount of signal in the wells, well 9 had more overall signal in the lock strand (green arrow) while the signal in well 8 is split between the correctly folded lock strand as well as an intermediate structure. The difference in well aggregation is highlighted when comparing the gels in Figure 4.3B and C. Both gels run the same knitting reactions, except gel B contains lock strands quenched in $MgCl_2$ while gel C contains locks quenched in water. There is substantially more well aggregation in the structures associated with the $MgCl_2$ quenched locks as seen in increased FAM signal at the top of the gels (blue arrow). Therefore, for the remaining HTL reactions, I quenched all the lock strands for one minute in water.

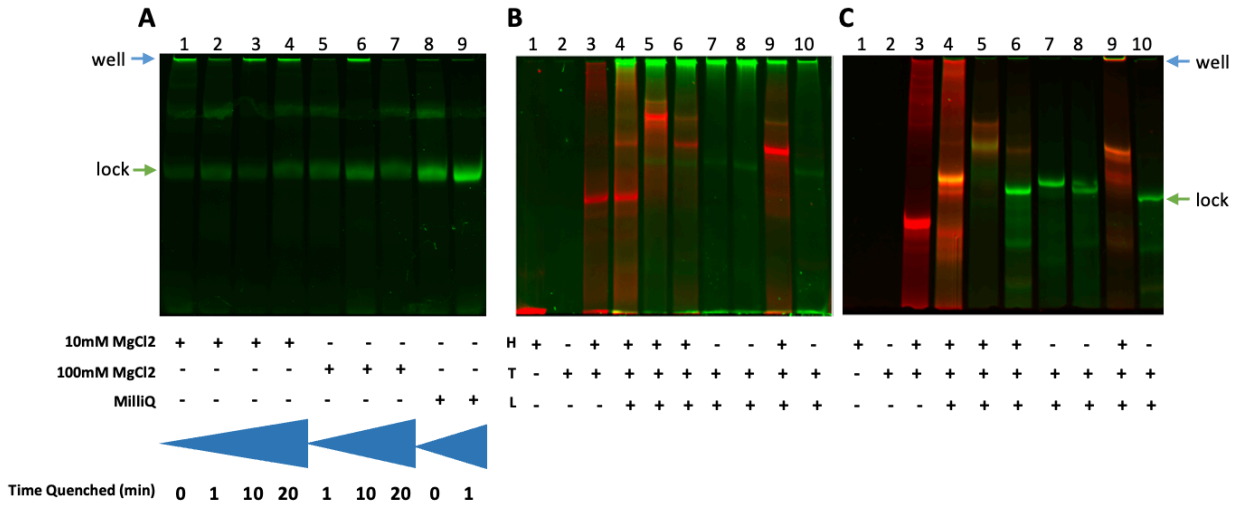


Figure 4.3: Relieving the well aggregation of the HTL structure. A) Quenching lock strands in either 10 mM or 100 mM $MgCl_2$ or in water for either 0, 1, 10 or 20 minutes as indicated by the blue gradient bar. Gels B) and C) show the same HTL reactions except B) utilised lock strands quenched in 10 mM $MgCl_2$ while C) uses lock strands quenched in water. The difference in well aggregation is indicated at the blue arrows. The lock strand is indicated by the green arrows. Samples are run on 10% polyacrylamide gels.

4.3.2 PAGE analysis of knitted and non-knitted reactions

Combining all the gel electrophoresis optimisations made so far for my HTL structure, results in the improved knitted (lanes 3-6) and non-knitted (lanes 7-9) reactions shown in the gel in Figure 4.4. Here, the hairpin (lanes 1) and thread (lane 2) controls are quenched in $MgCl_2$ while the lock strand is quenched in water. The quenched lock strand is also added at double the concentration of the hairpin-thread combination. After adding the toehold displacement strands, invader strand 1 is used for the displacement of the C-C* domain. The resulting bands are clear with minimal well aggregation. The red bands represent the hairpin or hairpin-associated structures, such as the hairpin-thread or hairpin-toehold displacement strand. The green band represents the lock strand or lock-associated structures such as the lock-thread. The yellow bands represent when the hairpin and lock strands are associated together. These yellow bands are likely to be the knitted structures. Two experimental replicates, with the same results are shown in Appendix Figure B.8.

As discussed back in Section 3.5.2 (Figure 3.6), when the components are added together in the order of hairpin-thread-lock-toehold-invader strand then the resulting structure is most likely to be knitted. Whereas if the components are added with the lock and thread first, preventing the thread from looping into the hairpin, the addition of the displacement strands causes the hairpin and thread strands to separate as no knitted structures were achieved. Additionally, whether the toehold displacement strand or invader strand is added first into the "knitted" HTL reaction does not change the results of the knitted structures as depicted in Appendix Figure B.9. Both reactions maintained the knitted structure as a yellow band on a polyacrylamide gel.

The intermediate stages for the knitted reaction with the hairpin-thread-lock (lane 4) and hairpin-thread-lock-toehold (lane 5) are more ambiguous in terms of determining which band is the main structure, as there are various overlapping red and green bands. Whereas the combination of all the components (lane 6) shows a very distinct yellow band as indicated where the orange arrow is. As this band is maintained as a yellow signal, the hairpin and lock strand must be associated together, which if displaced by the toehold displacement and invader strands only occur if the structure is knitted and locked in place. On the other hand, when all the displacement strands are added to the non-knitted reaction, the resulting structure reverts back to separate thread-lock-invader and hairpin-toehold displacement strand structures.

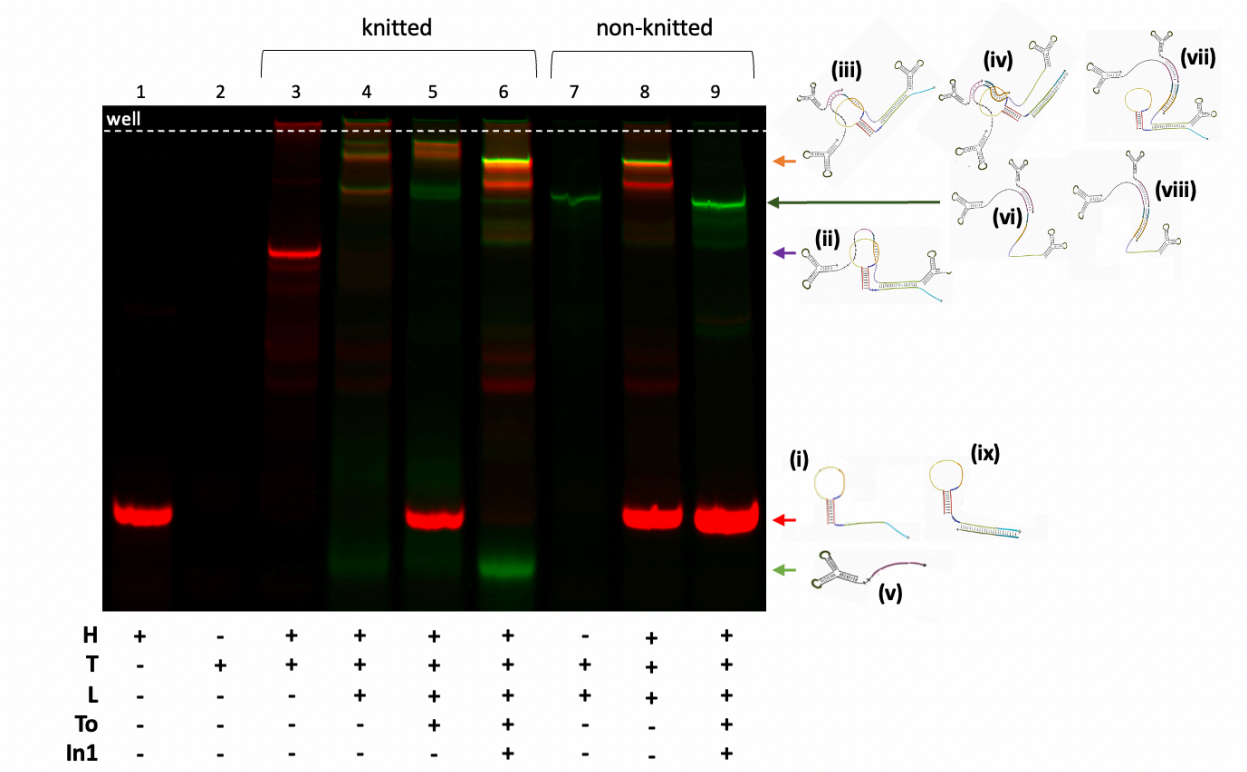


Figure 4.4: Knitted and non-knitted multi-knit reactions by 10% PAGE. Lanes 1-6 represent the knitted reaction. Lane 1 contains just the hairpin, with the band and schematic (i) indicated at the red arrow. Lane 2 contains the thread strand, there is not signal as there is no fluorophore on this strand. Lane 3 contains the hairpin and thread strand, with the main band and schematic (ii) indicated at the purple arrow. Lane 4 contains the hairpin, thread and lock strands, with the main structure and schematic (iii) indicated at the orange arrow. Lane 5 contains the hairpin, thread, lock and toehold displacement strands. Lane 6 contains the hairpin, thread, lock, toehold displacement strand and invader strand, with the main knitted structure and schematic (iv) indicated at the orange arrow. The excess of lock strands (v) is released after this reaction. Lanes 7-9 represent the non-knitted reaction. Lane 7 contains the thread and lock, with the band and schematic (vi) indicated at the dark green arrow. Lane 8 contains the thread, lock, invader and hairpin strands, with the main band and schematic (vii) indicated at the orange arrow. Lane 9 contains the thread, lock, invader, hairpin and toehold displacement strands, with the non-knitted thread-lock band indicated at the dark green arrow (viii) and the non-knitted hairpin-toehold displacement band indicated at the red arrow (ix). The dashed white line denotes the wells.

4.4 Analysing the Multi-knit DNA Structure

4.4.1 Analysing the multi-strand multi-knit DNA structure by PAGE

With the success of analysing my single-knit DNA model through gel electrophoresis, I moved on to see if I could achieve the same results with my multi-knit design. I first attempted to analyse the multi-component structure as discussed in Figure 3.7C. This design contains three hairpin structures that bind together, then three thread loops that attach to the hairpin loops and finally three lock strands that secure the knitted structures. Each component binds to each other in the order of hairpin A, hairpin B, hairpin C, thread C, thread B, thread A, lock A, lock B and lock C. Figure 4.5 depicts the polyacrylamide gels for the consecutive addition of each component. Gel A shows the hairpin interactions. The hairpin A band highlighted in pink in lane 1 of this gel is shifted higher when combined with hairpin b in lane 4. This is a sign of successful binding. Additionally this band shifts even higher when combined with hairpin C and then thread C as shown in lanes 7 and 8, respectively. This could also mean successfully binding, but it is a bit unclear as most of the sample is aggregated in the well as highlighted by the pink boxes.

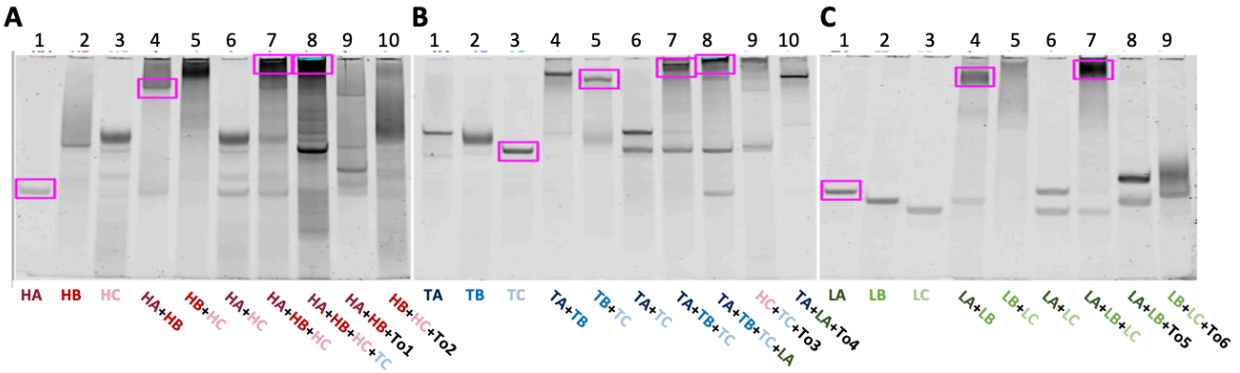


Figure 4.5: PAGE analysis of the multi-stranded multi-knitted DNA structure. A) Depicts the hairpin interactions, where lane 1 contains hairpin A (HA), lane 2 contains hairpin B (HB), lane 3 contains hairpin C (HC), lane 4 contains hairpin A with hairpin B, lane 5 contains hairpin B with hairpin C, lane 6 contains hairpin A with hairpin C, lane 7 contains hairpin A with hairpin B and hairpin C, lane 8 contains hairpin A with hairpin B, hairpin C and thread C, lane 9 contains hairpin A with hairpin B and toehold displacement strand 1 (To1), and lane 10 contains hairpin B with hairpin C and toehold displacement strand 2 (To2). B) Depicts the thread interactions, where lane 1 contains thread A (TA), lane 2 contains thread B (TB), lane 3 contains thread C (TC), lane 4 contains thread A and thread B, lane 5 contains thread B and thread C, lane 6 contains thread A and thread C, lane 7 contains thread A, with thread B and thread C, lane 8 contains thread A, with thread B, thread C and lock A, lane 9 contains hairpin C with thread C and toehold displacement strand 3 (To3), and lane 10 contains thread A with lock A and toehold displacement strand 4 (To4). C) Depicts the lock interactions where lane 1 contains lock A (LA), lane 2 contains lock B (LB), lane 3 contains lock C (LC), lane 4 contains lock A and lock B, lane 5 contains lock B and lock C, lane 6 contains lock A and lock C, lane 7 contains lock A, with lock B and lock C, lane 8 contains lock A, with lock B and the toehold displacement strand 5 (To5), and lane 9 contains lock B, with lock C and the toehold displacement strand 6 (To6). The pink boxes highlight the regions of interest. Samples are run on 10% polyacrylamide gels.

In a similar manner, the thread strands also bind correctly as depicted in gel B. Here, thread strand C (lane 3) shifts higher when combined with thread B (lane 5) and even higher when thread A (lane 7) and hairpin C (lane 8) are added. Gel C shows similar results with the lock strands. In this gel, lock A (lane 1) shifts higher when combined with lock B (lane 4) and even higher when lock C is added (lane 7).

Although the components appear to bind correctly to one another, it is unclear if the thread strands are looped into their appropriate hairpin strands. To elucidate the knitted structure, fluorophores are once again added to the structure in a similar fashion as the single-knit model. As shown in the schematic of Figure 4.6A, hairpin A is modified with a Cy5 molecule and lock A is modified with a FAM molecule. When analysing the hairpin-thread-lock reaction compared to the thread-lock-hairpin reaction of the A components in Figure 4.6C, you can see that neither are "knitted" when the displacement strands are added. In lane 9, the thread-lock-hairpin reaction is expected to not knit as the lock prevents the thread from looping into the hairpin. But the hairpin-thread-lock reaction in lane 6, which should have formed a knit, did not. These reactions are highlighted by the pink boxes, which both do not show any overlapping signal of Cy5 and FAM, indicating that these components have disassociated. A reason this structure may not have worked with this type of analysis, is that the thread strands lack any dumbbell ends, therefore when not attached to any of the other thread strands, the looped thread can fully pull through the hairpin loop when the interaction domain is displaced, even if the appropriate lock strand is bound. In the case of the reactions in gel C, once invader strand 3 displaces the interaction domain between hairpin A and thread B as indicated by the *, then thread A can pull through the hairpin loop as its ends are free, whether or not the lock strand was bound.

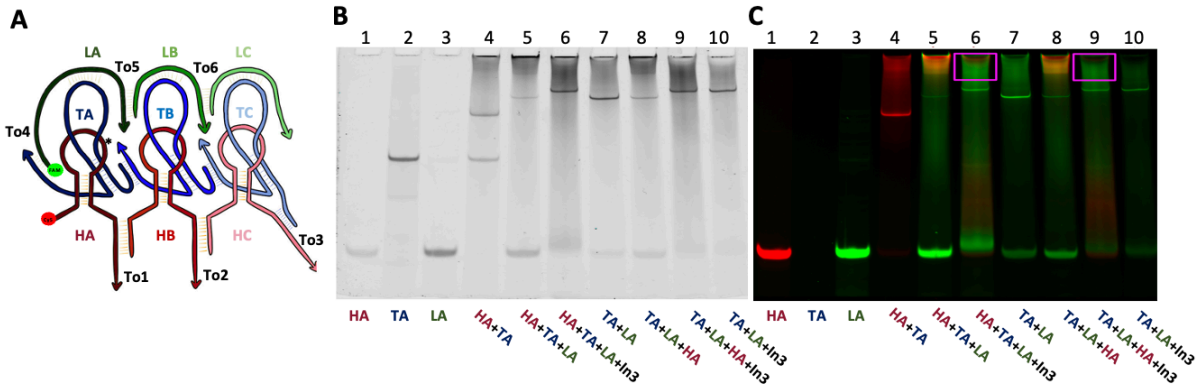


*Figure 4.6: Overview of the unsuccessful knitting of the multi-stranded multi-knit structure. A) Schematic of the conjoining hairpin, thread and lock model. Toehold displacement strands bind to the regions indicated by ToX and the * represents the region invader strand 3 would bind and displace. B) SYBRGold stained samples in a 10% polyacrylamide gel and C) shows the same samples under the Cy5 and FAM filters. The pink boxes highlight regions of interest.*

This multi-strand system could be analysed if all the components are added at once. A caveat to this approach is that by the time three components are added together, the band is either very close to the well or aggregated in the well. Therefore adding all nine components would not be decipherable. Therefore, a single-stranded approach is potentially more appropriate for analysis.

4.4.2 Analysing the single-stranded multi-knit DNA structure by PAGE

As the multi-stranded multi-knitted DNA structure was too difficult to analyse by PAGE, I moved onto investigating the unimolecular approach. This assembly method utilised the design previously discussed in Figure 3.11. As mentioned, the template strand is split into two double-stranded sequences of 415 and 408 base pairs for ease of synthesis. These strands are shown in the polyacrylamide gel in lanes 1 and 2 of Figure 4.7A as highlighted by the blue box around the 400 bp ladder mark.

These structures are then assembled together and circularised by golden gate assembly (lane 3). This assembled structure runs on the gel faster, at around the 300 bp ladder mark, compared to the individual linear components. The faster run of the circular template has been seen with similar rolling circle templates before [147]. The circularisation of the template is confirmed by exonuclease III digestion (lane 4), which should degrade any linear strands. Finally the nicking enzyme, Nt.BbvCI is added to create an attachment site for the polymerase. In these reactions, Phi29 polymerase is used. Similar results are shown with the use of the BstIII polymerase, as shown in Appendix Figure B.10, except that BbvCI digestion did not produce any additional bands apart from the signal accumulated within the wells of the BstIII reactions.

Once again, I've added complementary fluorescent oligonucleotides to my structure to aid visualisation to the samples in the polyacrylamide gel in Figure 4.7A. As depicted in the schematic in Figure 4.7B, I've designed a Cy5 modified strand to bind the linker region between hairpins 2 and 3, as well as a FAM modified strand to bind the linker region between hairpins 6 and 7. The result of adding these fluorescent strands to the polymerised product is shown in lanes 5 and 6 of Figure 4.7A. The fluorescent strands will only bind to the single-stranded polymerised product, and not to the double stranded template sequence, therefore the bands highlighted in the blue boxes are likely the multi-knitted structure.

The amplified sample is aggregated in the well which is common for RCA product. To further aid the PAGE analysis, I've added an invader strand (InB) that will bind the Nt.BbvCI recognition site at the beginning of each repeated 8 hairpin motif. With this domain now double-stranded, the BbvCI enzyme can digest this region, resulting in various multimers of the minimal motif. This can create structures of 8 hairpins, 16 hairpins, 24 hairpins and so on depending on where the enzymes cut the structures. The results of this reaction are shown in lanes 7 and 8 of Figure 4.7A. In lane 7, you can see a new band just beneath the well aggregation. Although I was expecting multiple bands representing the different sized multimers of the minimal motif, it is interesting to produce at least one new fluorescent band that can run through the gel.

Although I have assembled the rolling circle template, it is still unclear whether the polymerised product is knitted or not. I attempted to investigate the melting temperature of the polymerised product through RT-qPCR, but as the sample is very heterogenous with various lengths of amplified structures as well as the original templates, the resulting data has no clear melting curve as depicted in Appendix Figure B.11. The only method to clearly analyse the multi-knitted structures will be through microscopy.

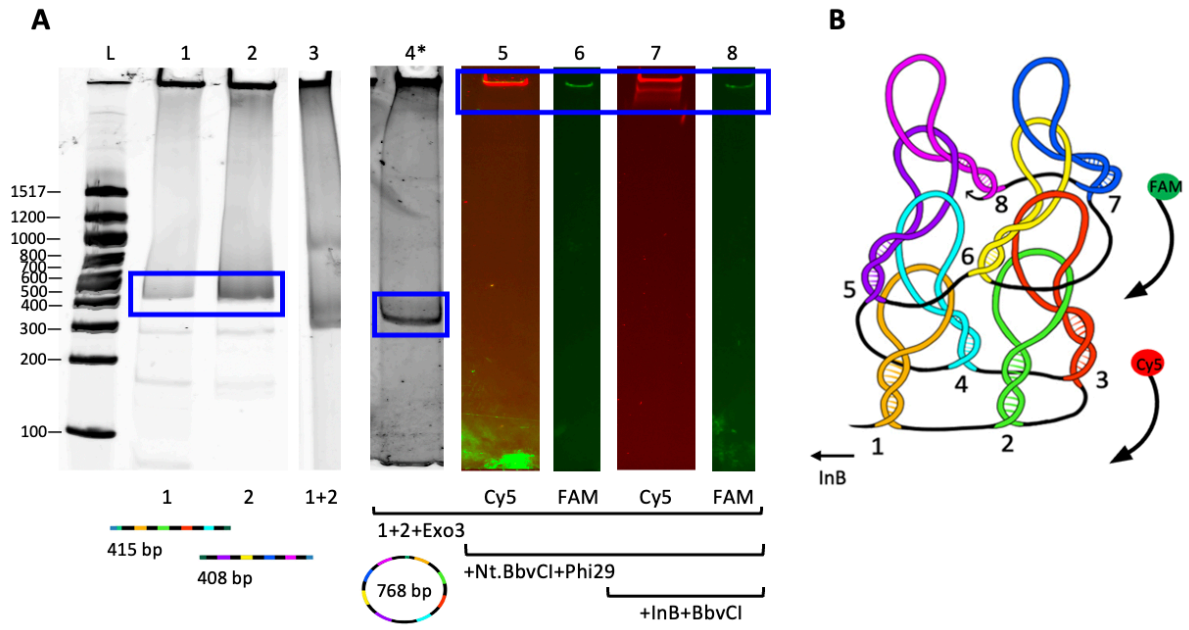


Figure 4.7: Overview of the single-stranded multi-knit structure. A) Components of the same 10% polyacrylamide gel show the molecular assembly of the DNA template and the resulting rolling circle amplified product. A 100 bp DNA ladder from NEB is used as a reference. The ladder and samples in lanes 1-4 are stained with SYBRGold. Lane 4 has increased contrast compared to the other lanes to better see the band. Lanes 5 and 7 show the Cy5 signal of the RCA product while lanes 6 and 8 show the FAM signal from respective complementary fluorescent oligonucleotides. Blue boxes highlight the main bands. B) Schematic of the single-stranded multi-knitted design.

4.4.3 Analysing the single-stranded multi-knit DNA structure by AFM

When deciding on which microscopy technique to use to image my multi-knitted DNA structures — whether that be transmission electron microscopy, scanning electron microscopy or cryoelectron microscopy, which all have been used to image DNA nanostructures [52, 148] — I chose atomic force microscopy (AFM) as this method best visualises the topology of flat objects at nanometer-scale spatial resolution. In particular, high-speed AFM, which is a force-based imaging technique that utilises a mechanical probe controlled by piezoelectric elements to scan the surface of a sample is effective at imaging the landscape of DNA nanostructures [149, 150]. This technique has been used to visualise the major and minor grooves of double-stranded DNA [151] as well as image individual hairpins on DNA origami structures [21].

To image my multi-knitted DNA nanostructures by AFM, I followed the advice of Dr Casey Adams, who at the time had experience imaging neurons with the Cypher AFM in the Contera lab [152]. Additionally, I used the materials leftover from previous members of my lab who have used AFM to image DNA nanostructures [153]. One important remnant used were the AC40 Biolever mini cantilevers, which are short AFM probes that are effective at responding quickly to shape changes of biomolecules in liquid. To prepare my rolling circle product for AFM, I immobilised the sample onto atomically flat mica disks using 12 mM $MgCl_2$, 4 mM $NiCl_2$ and 1X TE buffer. The added divalent cations in this buffer aid the flattening of the negatively charged DNA

nanostructures on to the mica. To stick the DNA nanostructures onto the mica, the sample is diluted to under 200 ng/µL in the imaging solution and incubated on the mica for two minutes. Afterwards, 50 µL more of the buffer is added on top of the sample.

When imaging my unpurified RCA polymerised product using this initial protocol, I was able to obtain micrographs of structures that could potentially be my multi-knitted DNA structure, as seen in Figure 4.8. In micrograph B, which captures a wide scope of the structures, you can see that majority of the structures are ribbon-like with distinct patterning. When imaging at a higher resolution, as shown in micrographs C and D, the pattern alternates between a height of 2 nm (yellow) and 1 nm (orange), which matches the height of double-stranded DNA and single-stranded DNA, respectively. The double-stranded regions could represent the knitted domains of the structure, while the single-stranded areas could represent the single-stranded linker domains between the hairpins.

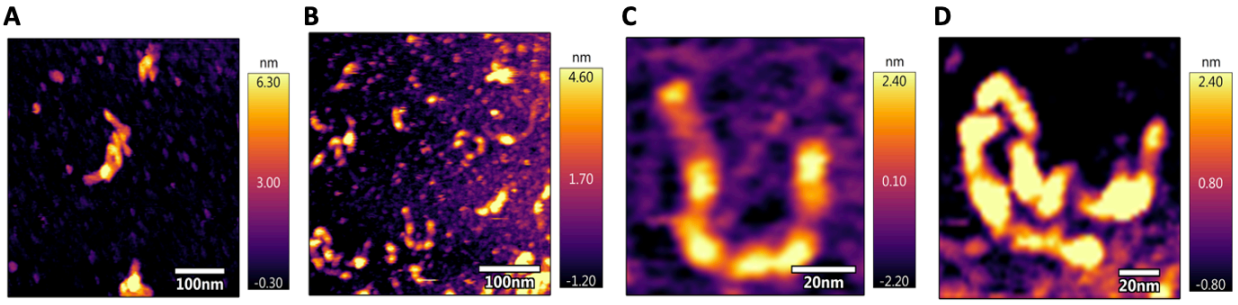


Figure 4.8: AFM micrographs of the single-stranded multi-knitted DNA structure. The coloured scale grades the height of the structure in nm. Scale bars are provided on the bottom right of each micrograph. Each image has been flattened and optimised for analysis.

4.5 Improving the AFM Resolution of the Multi-knit Nanostructures

4.5.1 Optimising the adsorption of the nanostructures to the mica

After achieving my initial AFM images, I wanted to see if I could improve the protocol to image at a higher resolution to resolve the looping pattern of my multi-knitted DNA nanostructures. So far, it is unclear if these asymmetrically patterned ribbons are knitted or not, therefore higher resolution of the structures that can show the inner-linkage of the hairpins within each other is needed.

The first improvement step was to see if I could optimise the absorption buffer to flatten out the nanostructures in a better position on the mica disks to visualise the looped hairpin regions. My original protocol was constructed from the advice of AFM users within or near my lab and was not optimised for such a flexible structure. Therefore, I explored alternative protocols used to adsorb DNA to mica. One of these protocols came from Alice Pyne's group, who have a strong focus on single-molecule studies of the DNA double helix. Their protocol utilises 20mM HEPES buffer instead of TE and negates the use of magnesium chloride [124]. The sample is diluted to 10 ng/µL in the buffer and 20 µL is incubated on the mica disk for 30 minutes. Two to three wash

steps, —exchanging the buffer — is carried out with a final top up of 40 µL of the buffer to the sample. Another protocol I found that is similar to my original adsorption method uses an increased concentration of 20mM $NiCl_2$ to image DNA G-quadruplexes [123]. The rest of this protocol follows the same steps as my original one.

A comparison of these protocols on how well they flatten my DNA multi-knitted structures onto mica disks is shown in Figure 4.9. All three protocols produce a long ribbon like structure around 100-200 nm in length, which share a similar repeating pattern of single-stranded and double-stranded DNA. Using my original protocol results in the blurring of the double-stranded and single-stranded regions as seen in the micrographs in part A. On the other hand, the protocol used in the micrographs in part C show clearer distinction of the different heights, but the overall structures are very linear and compact. This may suggest that the DNA is condensing too much due to the increased cation concentration. The Pyne protocol shown in the micrographs in part B of the figure show the best version of the potential knitting, as you can see a zig-zag pattern that could correspond to the hairpins looping back and forth.

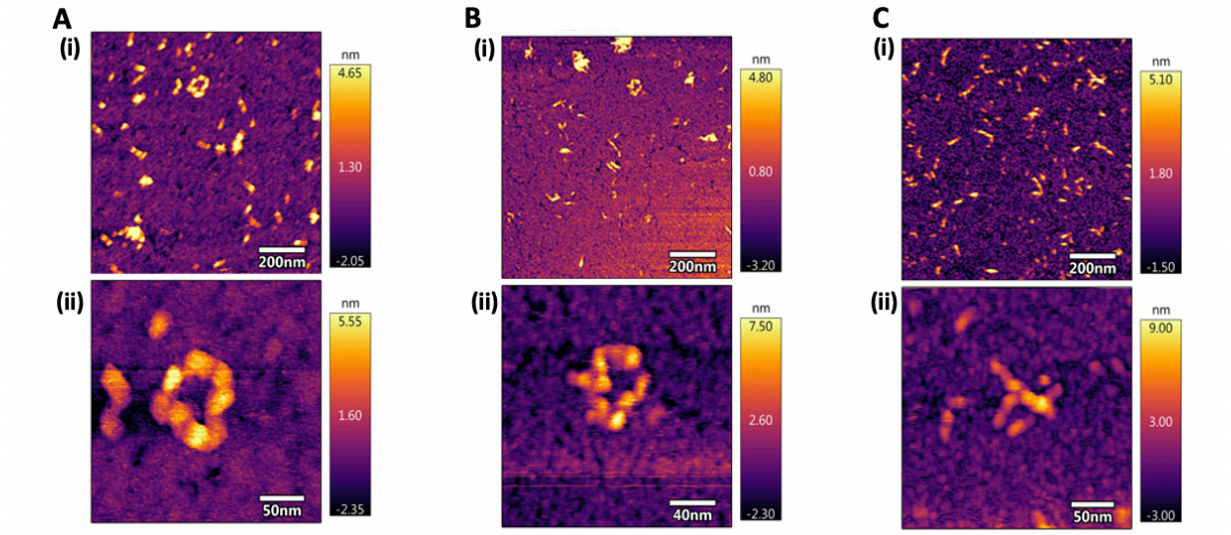


Figure 4.9: Comparing different AFM adsorption protocols. A) 12 mM $MgCl_2$, 4 mM $NiCl_2$ and 1X TE buffer discussed in Section 4.4.3. B) 3 mM $NiCl_2$ and 20 mM HEPES pH 7.4 as described [124]. *C) 10 mM $MgCl_2$, 20 mM $NiCl_2$ and 1X TE buffer as described in* [123]. *(i) Images display a wide field of view, whereas (ii) images show a zoomed up micrograph of a structure of interest. The coloured scale grades the height of the structure in nm. Scale bars are provided on the bottom right of each micrograph. Each image has been flattened and optimised for analysis.*

4.5.2 Imaging the multi-knit structure with specialised equipment and software

As I was able to achieve the best resolution images utilising an adsorption protocol from the Pyne group, I went to visit their lab in person to image my samples under their guidance. Through the Royce Institute Equipment Access Scheme, I was able to train under Dr. Tom Catley in the Pyne group and use the Bruker FastScan-XR AFM as well as their specialised software. In addition to their high-speed microscope, I was also able to use the compatible Bruker FastScan-D cantilevers.

These probes push the limits for high-resolution imaging of fragile biological samples, such as DNA, without deforming the samples' fundamental atomic structures.
For the adsorption protocol, we used a similar method as the Pyne protocol mentioned in the previous section. We utilised varying concentrations of $MgCl_2$ as well as changed the incubation times, but the success of the structures flattening correctly to the mica appeared to be arbitrary. In Appendix Figure B.12 it can be noted that the same absorption protocol is used, but differing results occur with how the samples appear on the mica. Particularly, the DNA can vary in how condensed they appear to form against the mica.

Finally, to analyse and trace the DNA structures in my AFM images, I used the software TopoStats pioneered by the Pyne group [126]. TopoStats is a Python toolkit for the automated editing and analysis of AFM images and works by creating pixel masks containing the positions of the molecules. The software has been used to characterise topologically complex DNA conformations. In addition to identifying DNA minicircles, DNA origami rings and pore forming proteins and various other biological nanostructure, TopoStats has shown promise in analysing DNA catenanes and knots, which is very similar to the looping architecture of my multi-knitted DNA nanostructures.

4.5.3 Analysis of the template strands

Using the optimised imaging buffer and protocol from the Pyne group, I was able to produce AFM images of the rolling circle amplification templates for my multi-knit DNA structure. Each half template is around 400 bp, which can be visualised in Figure 4.10A and B part (i), and traced by TopoStats as highlighted in blue in part (ii). In part C of this Figure, the distribution of 32 of these strands is shown, with the average length identified as 148 nm.

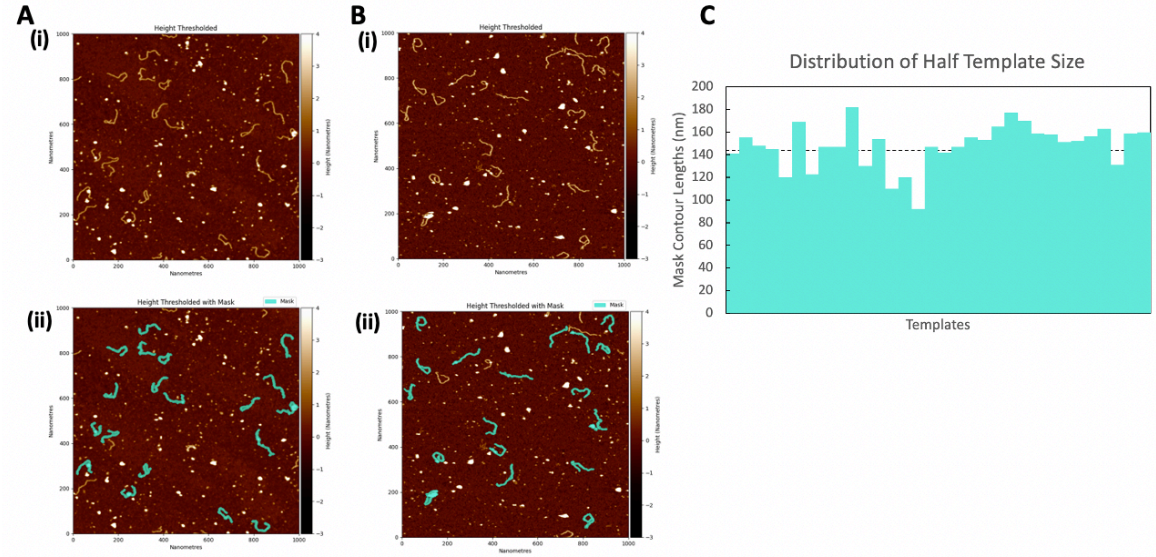


Figure 4.10: AFM analysis of the RCA half template strands. Micrographs for a 1000 x 1000 nm region of sample is shown in A) and B), without (I) or with (ii) TopoStats tracing shown in blue. A scale bar representing the height of the structures from the mica is given in nm. C) Distribution of the TopoStats tracing of the half templates. 32 strands are traced with the average length being 148 nm as indicated by the dashed line.

When imaging the golden gate assembled templates, as represented in Figure 4.11, you can see a mixture of half templates (strand 1), full linear templates (strand 5) and full circularised templates (strand 3). The half template strands range from lengths 133 to 153.7 nm, while the full linearised strand 5 is 280 nm. This corresponds to nearly double the length of the half templates. The circularised full template also measures around 257 nm. A bit of the length from the half templates are cut off due to there being primer regions to allow for amplification of the half templates. The primer sites occur before and after the restriction enzyme sites, and are therefore digested out during assembly. Additionally, there seems to be a smaller circular structure (strand 6) that is around 150 nm in length. This may be a half template that circularised. All these strands appear to have a uniform height of 2 nm, representing double stranded DNA.

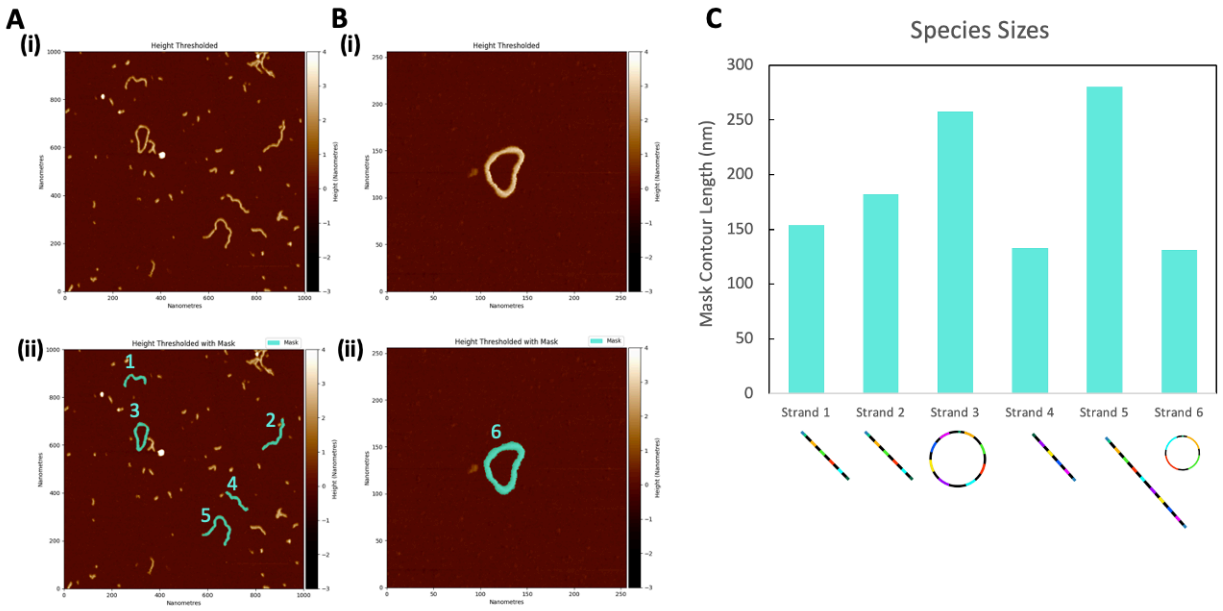


Figure 4.11: AFM analysis of the golden gate assembled templates. A) and B) show different samples of the same golden gate assembled reaction. A) shows a range of half and full templates within a 1000 x1000 nm region, whereas B) shows a unique circularised half template zoomed in at 250 x 250 nm. The Topostats tracing is shown in blue in (ii). A scale bar representing the height of the structures from the mica is given in nm. C) Depicts a graph of the traced templates and their approximate length.

It is important to note that my golden gate assembled sample is very heterogenous, containing a mixture of partial or full templates in either linear or circularised forms. Ideally, I would need to purify the circularised full templates for rolling circle amplification, but as the concentration of the circularised structures is so low at this point after various molecular cloning steps, I proceed to amplify the whole sample. This should still allow for various multi-knitted structures to form as the half template with the Nt.BbvCI site as well as the full linear template can still produce a knitted structure, only forming 4 or 8 hairpins, respectively, instead of continuous repeats with the circular template.

4.5.4 Analysis of the single knit hairpin

After imaging the template strands and ensuring correct circularisation, I analysed the individual eight hairpins (separate oligonucleotides) of my multi-knitted design to see if I could resolve their structures by AFM. As discussed previously in Figure 3.11 and 3.12, Each hairpin has a stem of 10 bp and a loop of 56 nt with two distinct interaction domains. As these hairpins are small with such a flexible long single-stranded loop region that they will be difficult to flatten onto the mica. As shown in Figure 4.12, I was able to image individual structures with a double stranded region, most likely representing the stem, as well as a densely packed region that most likely shows the single stranded loop region twisted in on itself. These individual structures are approximately 10 nm high and 4 nm wide (Figure 4.12A), which corresponds to the dimensions of my hairpins. Alongside the AFM micrographs, I also ran the eight hairpins on a polyacrylamide gel (Figure 4.12D), where hairpins 1, 2,4, and 6 run at the same level, just under the 100 bp benchmark of the DNA ladder. Hairpins 3, 5, 7 and 8 run at a higher level, just under the 200 bp point of the ladder, with the latter three forming smears that might represent various conformational forms.

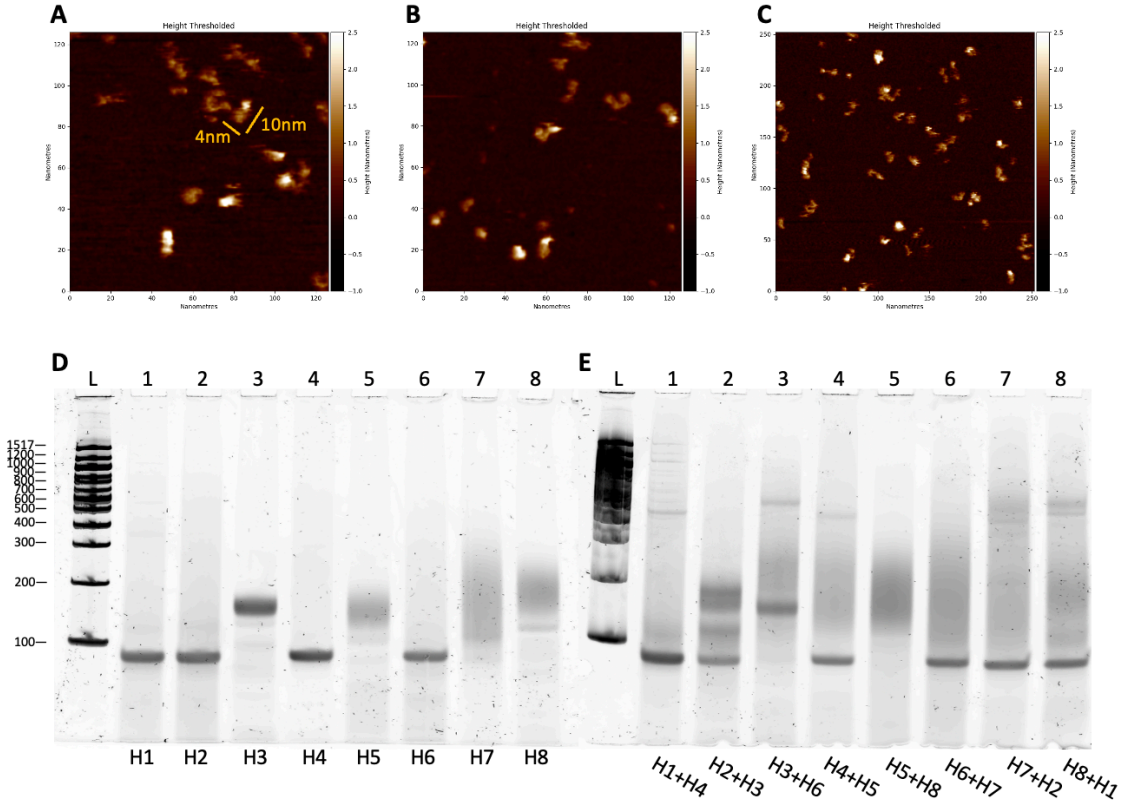


Figure 4.12: Analysis of single DNA hairpins. A-C) Show AFM micrographs of Hairpin 1 (H1), Hairpin 2 (H2) and Hairpin 3 (H3). A-B) show a 120 x 120 nm area while C) shows a 250 x 250 area. The approximate dimensions of a single hairpin are given in A). A scale bar representing the height of the structures from the mica is given in nm. D-E) show the 10% PAGE for the individual hairpins (1-8) as well as combinations of stacked hairpins. In E) lane 1 contains hairpin 1 and hairpin 4, lane 2 contains hairpin 2 and hairpin 3, lane 3 contains hairpin 3 and hairpin 6, lane 4 contains hairpin 4 and hairpin 5, lane 5 contains hairpin 5 and hairpin 8, lane 6 contains hairpin 6 and hairpin 7, lane 7 contains hairpin 7 and hairpin 2, and lane 8 contains hairpin 8 and hairpin 1. A 100 bp DNA ladder (L) is labelled and samples are stained with SYBRGold.

In addition to imaging the individual hairpins, I attempted to analyse two stacked hairpin interactions. The bands for these structures run on a polyacrylamide gel is shown in Figure 4.12E. Here, it can be seen that only the interactions between hairpin 1 and 4 (lane 1), hairpins 3 and 6 (lane 3), hairpins 4 and 5 (lane 4), hairpins 7 and 2 (lane 7) and hairpins 8 and 1 (lane 1) form a clear higher band that would likely represent the two hairpins stacked together. Unfortunately, when moving to image these hairpin pairs, the structures were not able to flatten on the mica and therefore no clear images were captured.

4.5.5 Analysis of the multi-knit structures

With the templates and hairpins characterised as best as possible, we moved onto imaging the final rolling circle product. As seen in Figure 4.13, there are long ribbon-like structures similar to those seem in my original AFM micrographs, but instead of repeats of high and low regions of height, there is better resolution of the actual looping structures. In particular, there are these unique rounded structures (indicated by the pink arrows), similar to the formation of the leaves of a clover. These curved regions may potentially represent the intermediate stages of the hairpin structures, with the stems binding and unbinding transiently. As the height of these strands is mostly around 1 nm, the structure is most likely made from a single strand of DNA. And where the height increases, most likely represents regions of inter-looped hairpins.

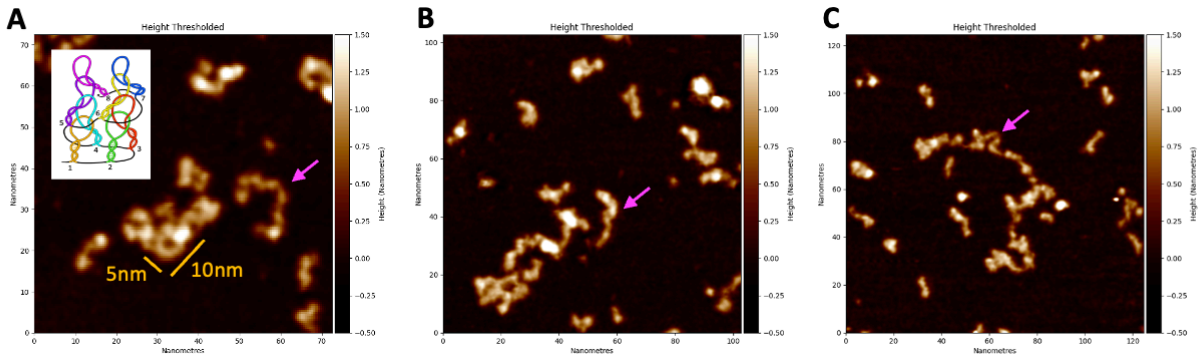


Figure 4.13: AFM micrographs of the multi-knitted DNA structures. A) is the most zoomed in at 70 x 70 nm, while B) is 100 x 100 nm and C) is 120 x 120 nm. The dimensions of the individual loops are approximated in A, where a single looped hairpin is about 10 nm tall and 5 nm wide. A schematic of the full structure is represented in the top left corner as reference. Structures of interest, such as the clover-like rounded motifs are highlighted by the pink arrows. A scale bar representing the height of the structures from the mica is given in nm.

Additional to the clover like motifs, the overall structure does seem to form continuously growing looped regions with the dimensions of two hairpins in each row. This is best seen in Figure 4.13A, where the lengths of an approximate individual hairpin are indicated in yellow and a schematic of the 8 hairpins together is shown. Here, the structure appears to start with a base of two hairpins, but only appears to continuously loop in from the left side hairpins. The concept of DNA knitting should allow for mis-looped regions as long as the rest of the row is looped. But it appears that not

all the hairpin stems form completely, therefore the next row of knitting does not have a loop structure to loop into if the previous row is not knitted through. This hindrance may be fixed by making the hairpin stems longer and therefore more stable.

The rolling circle template was designed to allow for continuously growing repeats of the 8 hairpin motif. But when looking at these structures in the given micrographs, as well as in wider fields of view shown in Appendix Figure B.13, it appears that there are not many substantially long structures as expected from rolling circle amplification. Mentioned previously, my final rolling circle product is very heterogenous as the initial assembled template sample also contains various stages of half and full templates, resulting in different possible lengths of hairpins being amplified. When examining the wide view micrographs of the template samples, the full circle templates were much rarer than the half templates or linear full templates, therefore it is likely that many partial motifs are being produced. Additionally, if the looping regions in the polymerised are too complex, it is possible that the polymerase detach earlier before it can transcribe multiple rounds of the structure.

It is promising to see a final structure that consistently produces loop-like and rounded structures. This may mean that there is a preferred folding pathway for the production of 8 hairpins from the template. The issue appears to be that not all of the hairpins fold into each other. This factor would not be a problem in a larger multi-knit DNA nanostructure, as the "knitting" design is programmed to continue looping from a mis-looped hairpin as long as there are other successfully knitted hairpins to maintain the structure of each row. This behaviour is inspired by knitted fabrics, which still can uphold their overall structure despite mistakes in the knitting. As my structure only has two hairpins in each row, to minimise the sequence for rolling circle amplification, this advantage is not effective as there are not enough supporting hairpins in each row to mitigate the impact of non-knitted hairpins. Therefore, a redesign for a larger multi-knit structure is needed. Ideas for larger DNA knitted structures for future experimentation is discussed in Section 7.2.1

4.5.6 Potential discussions about looping analysis

An interesting aspect to TopoStats is that is can trace the directionality of the DNA strand as well as decipher the top and bottom strands in a DNA crossing (unpublished work by the Pyne group). In particular, this software determines the direction of DNA crossings based on the Gaussian smoothed masks of the molecules, where the skeletons of the structures are biased towards the higher DNA backbones and false branches are pruned out based on their heights and lengths. This is potentially useful for determining the directionality of the looping structures in my multi-knit DNA design.

Discussing with Dr. Tom Catley and the rest of the Pyne group, it is not likely that the current version of TopoStats can handle the complexity of my mutli-knit structure, as the double-stranded and single-stranded regions are too close together and are irregular in shape. Although this software is not yet advanced enough to trace my asymmetrical multi-knitted structures, as they contain various double-stranded and single-stranded regions, TopoStats is still useful for the image processing of my rolling circle amplified samples as well as for analysing the template structures and individual hairpins. Using TopoStats, I can trace the approximate lengths of regular shaped

structures, such as my linear and circular templates. Hopefully, in the future TopoStats can be used to characterise a structure as complicated as the DNA multi-knitted assembly.

4.6 Chapter Summary and Outlook

In this Chapter, I went through the gel electrophoresis analysis of both my single-knit and multi-DNA designs and analysed the atomic force microscopy of the rolling circle reactions. The single-stranded HTL structure appears to be successfully knitted, but it is still unclear whether the final multi-knitted structure is fully knitted or just bound together in a formation without looping the hairpins within each other in a knitted fashion.

For my single-knit DNA design, I worked through the optimisation of the hairpin, thread and lock strands for PAGE electrophoresis. This required the quenching of the control strands, particularly in $MgCl_2$ for the hairpin and thread strands, and in MilliQ water for the lock strands. This relieved the aggregation in the wells, allowing for more distinct bands to show in the gel. Next, I worked out the concentrations needed for each component to form a knitted or non-knitted structure.

To further analyse the knitted structure, I added fluorophores to the hairpin and lock strands to confirm they are associated together. Once all components are added together, whether in the order of the hairpin-thread-lock or thread-lock-hairpin, the displacement strands are then mixed in, resulting in either a knitted or non-knitted reaction. A knitted reaction is shown on a polyacrylamide gel as a maintained yellow band, whereas the non-knitted reaction results in separate thread-lock strands and hairpin-displacement strands.

As for my multi-knit DNA design, I first attempted to analyse the multi-component version containing conjoining hairpins, threads and locks. In a polyacrylamide gel, this structure was too close to or aggregated into the wells of the gel, making analysis difficult. Therefore, I moved onto analysing my single-stranded multi-knitted structure created by rolling circle amplification. For this design, I first had to check the assembly and circularisation of the template strands, which is confirmed by both PAGE and AFM. When checking the RCA product on a gel, it was aggregated in the wells, but a bound fluorescent oligonucleotide strand ensured that this final structure was single-stranded. I utilised the RCA polymerase Phi29 for these reactions, as the amplified product had more of a reaction to BbvCI digestion compared to when the polymerase BstIII was used.

To further analyse my multi-knitted DNA structure, I used atomic force microscopy to visualise the hairpin looping. My first attempts at imaging the structures resulted in blurry ribbon-like structures with alternating double-stranded and single-stranded regions. To further resolve the design, I worked with the Pyne group in Sheffield to optimise the adsorption and imaging protocols of my sample. With these experiments, I was able to see distinct looping shapes made from a long single-stranded piece of DNA. It is still unclear though, whether these rounded structures are actually looped into each other or just overlapping one another. It would require further analysis to confirm if the structure is correctly knitted. Additionally, it appears that the adsorption protocol is not consistent as we would get various forms of condensed or immobilised samples of the mica using the same imaging buffers. Further optimisation of this step would be helpful when conducting large rounds of imaging.

Another concern from the final AFM images, is that the resulting structures are not very long. The rolling circle template was designed to create long ribbons of knitted structures, intended to aid visualisation of the design. With the shorter structures, it is harder to the confirm the directionality of the looping. Therefore, it would be useful to design a template that allows for more hairpins in each motif or to improve amplification from the RCA polymerase. A few examples of potential larger multi-knit designs are discussed in future research directions in Chapter 7.

Another reason for the shorter RCA products is that the template strands were not purified as the concentration after golden gate assembly is very low. The concentration of the fully circular templates was also low in comparison to the half or full linear templates. Therefore, developing a way to purify or obtain large quantities of the full circularised template will also help optimise imaging of the multi-knit structures. One method I used in the attempt to obtain more circular templates was through the use of the gene editing system, Cre-loxP [154]. This system uses flanking loxP genes around your gene of interest that then circularises the sequence into a plasmid by Cre enzymes. This circular plasmid can then be transformed into *E. coli* cells, which can be purified to achieve high quantities of the circular plasmid. I attempted to use this system with my sequence as overviewed in Appendix Figure B.14 but the resulting sequence purified from the bacterial cells contained multiple mutations. This may be caused by the complex repetitive nature of my multi-knit DNA design. An alternative purification or amplification step would therefore be needed to obtain higher concentrations of the fully circularised template.

Chapter 5

Intracellular Delivery of DNA Origami CryoEM Tags

5.1 Introduction

In this Chapter, I will introduce a new technique to deliver *in-vitro* assembled complex DNA origami into mammalian cells. I will explain the use of both commercial and homemade versions of a cell confinement chamber to create transient pores in the cell membrane of MF10A cells and HEK293T cells to allow for large complex structures to enter the cells. In particular, I focus on delivering asymmetrical DNA origami cryoEM tags to bind GFP-expressing intracellular membrane proteins within the cell. I will show preliminary data using fluorescent confocal microscopy to show promising delivery and intracellular membrane tagging of the DNA nanostructures using cell confinement.

5.2 Author Contributions

The initial DNA origami signposts were provided by Dr. Emma Silvester, a previous student in the Turberfield group. The origami structures were later assembled by Catherine Fan. Preliminary experiments with MCF10A cells were conducted in the Baum lab under the guidance of Dr. Zaw Win. All later experiments carried out with HEK293T cells were conducted by Catherine Fan. The inducible HEK293T cell line expressing eGFP-cd63-BMP7 was created and given by Dr. Antonio Garcia Guerra (Turberfield group).

5.3 Confinement Chambers

5.3.1 Using confinement as an intracellular delivery method

As discussed in Sections 1.3.2 and 1.3.3, there are various methods to deliver nucleic acids into the cell, but many are limited due to the sizes of the nanostructures that can be delivered, and if successful, usually result in low cellular uptake. Most of these methods rely on receptor-mediated endocytosis which can trap the nanostructures within the endosomes, therefore they require an additional escape mechanism. One promising technique to bypass size restrictions as well as endosomal uptake is through cell squeezing. I first came upon this technique being used for the cytosolic delivery of DNA nanopegboards into HeLa cells [155]. Here, a solution containing the DNA nanopegboards was flowed through a microchannel with a constriction region smaller than the cell diameter. This confinement creates temporary pores in the membrane of the HeLa cells, allowing the DNA nanopegboards in the surrounding fluid to diffuse into the cell. Through this method, they were able to deliver 80% of nanopegboards into the cytosol, which was comparable to the 70% delivered through an endocytic pathway. Additionally, this confinement method could

be used to target the delivered nanopegboards to the nucleus through a biotinylated nuclear localisation signal peptide. Alongside the use as intracellular delivery tools for DNA nanostructures, microfluidic devices have also been developed for real-time monitoring of cancer cells exposed to labeled dendrimer nanoparticles [156]. With this delivery system, cell viability as well as internalisation efficiency in either static or dynamic conditions can easily be characterised.

Unlike electroporation, which has been shown to cause genetic disruptions to primary human cells, microfluidic cell squeezing causes minimal aberrant transcriptional responses allowing for more therapeutic potential *in vivo* [157]. This 2018 study found that using a microfluidic device preserves the functionality of sensitive unstimulated human T cells and human stem cells, while electroporation of the same cell lines caused the dysregulation of key genes, functional pathways and disease markers. Similar to the previous microfluidics study, here they used a narrow channel to apply a specific pressure to the cell membranes, without causing risk to cell viability.

In contrast to the microfluidics experiments that flow the cells through narrow microchannels, another method that involves mechanical membrane disruption is the application of an external force, such as a coverslip or column, to squeeze the cells. This technique has been carried out recently in 2022 to improve nuclear import of large DNA nanostructures in various mammalian cell lines [158]. In particular, they applied gentle compression through a coverslip that resulted in approximately 50% nuclear accumulation without severe endosomal entrapment, cytotoxicity, or long-term membrane damage. Compression of the DNA nanostructures into the cells through this method was therefore successful in mechanotransduction-mediated intranuclear delivery.

As a delivery method, cell confinement is able to deliver large DNA nanostructures directly into the cytosol or even the nucleus of the cell, in intact condition with minimal risk to cell viability and functionality. This method of delivery can either be conducted through microfluidic microchannels or with a simple compression system such as pressing a coverslip over the cells. Both techniques use the concept of confining the cells just enough to create transient pores in the cell membrane without rupturing them. Each system will require its own parameters for compression height and time, but overall, cell confinement appears to be a straightforward and effective method for intracellular delivery.

5.3.2 Commercial 4Dcell confiners

Most of the microfluidic methods for intracellular delivery require the engineering of a bespoke device with specific microchannels and flow rates, whereas a simple confinement chamber only considers the height of compression on the cells. One confinement system I found commercially, are the 4Dcell chambers as depicted in Figure 5.1A [159]. These 4Dcell confiners come in either a dynamic system that utilises a vacuum pump to control the height and speed of confinement (Figure 5.1B), or a static confiner that manually compresses the cells to the height of the micropillars (Figure 5.1C). Both methods homogeneously confine cells between two parallel surfaces to a defined micrometer height, using specific polydimethylsiloxane (PDMS) or acrylamide-based micropillars attached to a PDMS piston. The confinement height can be fine-tuned to sub-micrometer resolution to achieve the optimal compression for any cell type.

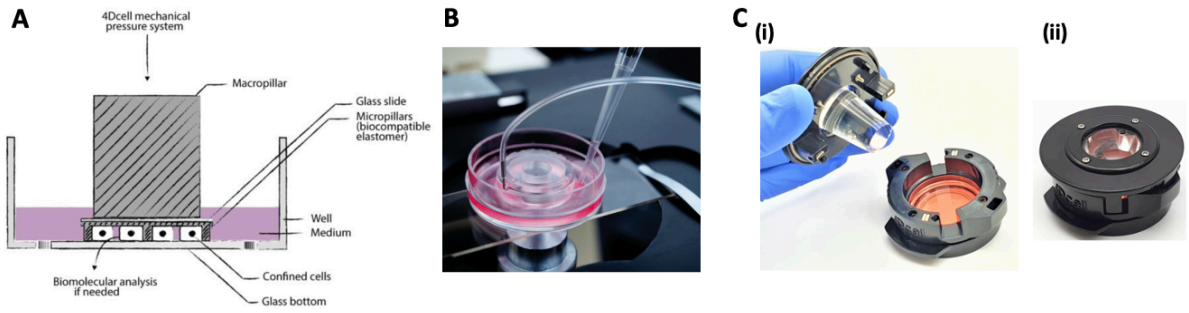


Figure 5.1: Overview of the 4Dcell cell confiners. A) Schematic of the confinement process. Biocompatible micropillars are attached to a PDMS piston which applies pressure onto the cells adhered to a petri dish [159]. *B) Dynamic and C) static cell confiners. C (i) Depicts the confine in an open state and (ii) is when the confiner is closed. The images were taken from 4Dcell.com.*

The dynamic cell confiners in particular allows for imaging of cells before, during and after confinement, as you can control the pressure through the vacuum pump while imaging As the lid of the device does not need be moved during confinement, you can track single cells throughout the process. In Figure 5.2 you can see the tracking of the same set of cells through the process of confinement. These images were taken while I visited the Baum lab, where they study the use of pressure onto their cells to eludicate the affects to transcription by mechanotransduction. Therefore, the cells in Figure 5.2C are almost completely squashed. For my cell delivery experiments, I want to find a middle ground where the cells are perturbed enough to allow intracellular delivery without completely breaking the cell membranes.

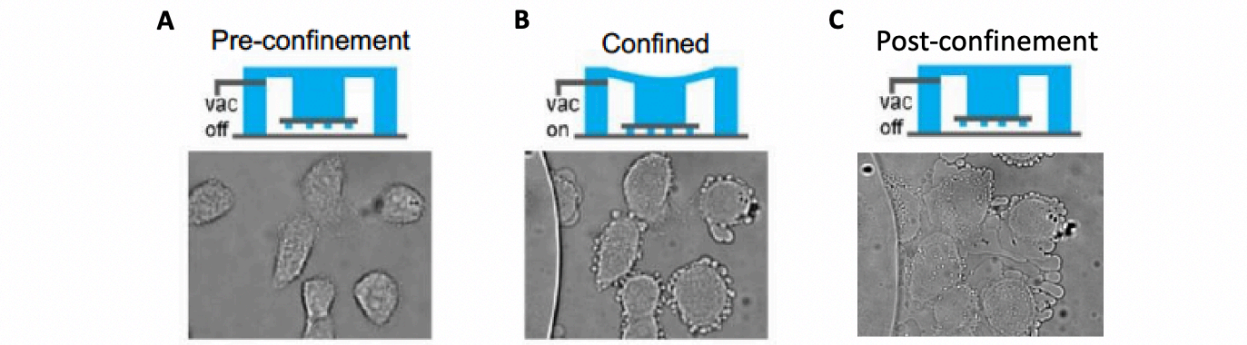


Figure 5.2: MCF10A cells compressed by a 4Dcell dynamic cell confiner. A) Cells unconfined, B) cells under confinement and C) cells released from confinement. Cells were confined to a height of 5 µm. Images were taken at 40X magnification in brightfield. Images were provided with permission from the Baum Lab.

The 4Dcell confinement chambers have been used to study various cellular applications, including exploring the relationship between the cell cytoskeleton and cell polarity [160], cell migration behaviour under confinement [161], cancer cell mechanical properties [162], cell separation [163], and label-free recognition of specific cells based on deformability [164]. The cell confinement

devices provide targeted insight into the effect of cell deformation and mechanical stresses on cell behaviour. One application these commercial cell confiners have not been used in yet is as a tool for intracellular delivery of nanoparticles.

5.3.3 Homemade cell confiners

In addition to the commercial cell confiners, you can make your own simplified versions using generic lab equipment. While visiting the Baum lab, I was taught the protocol to make homemade cell confinement chambers using a 12-well plate as shown in Figure 5.3. The micropillars are made from acrylamide based on a set mold and the columns are made out of PDMS. This system can be customised for different cell sizes and cell counts. The protocol for this setup is based on Le Berre's methods for the 4Dcell chambers [159].

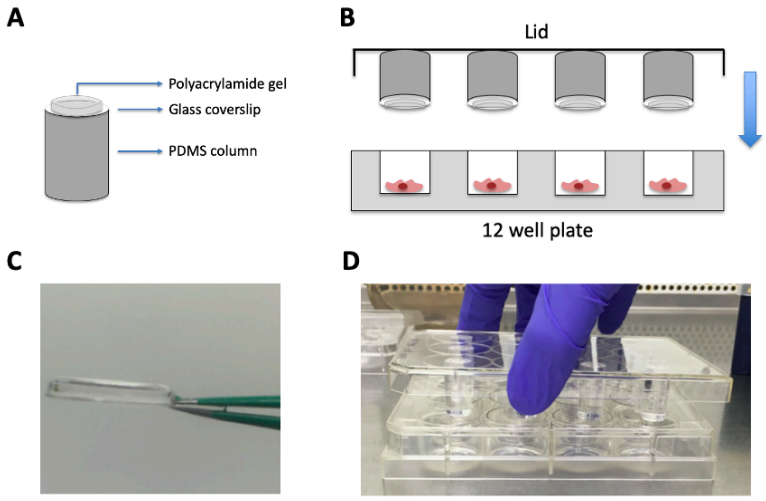


Figure 5.3: Homemade cell confinement made out of a 12-well plate, acrylamide micropillars, glass coverslips and PDMS columns. A) Shows a schematic of a single column and its components, while B) shows a schematic of the full well plate. C) Depicts an image of the polyacrylamide micropillars and D) depicts the full set up. Images taken from the Baum lab.

5.4 Utilising DNA Origami cryoEM Tags for Intracellular Delivery

5.4.1 DNA origami signposts as cryoelectron tomography tags

As an interesting nanostructure to deliver into cells using the cell confinement method, I chose DNA origami cryoelectron tomography tags developed by a previous student in the Turberfield lab, Dr. Emma Silvester [72]. These nanostructures were designed with an asymmetric 96-helix honeycomb lattice "sign" and a 6-helix bundled "post" with a RNA aptamer for binding a target protein Figure 5.4A. Additionally, these structures were annealed with Cy5-conjugated strands within the sign head.

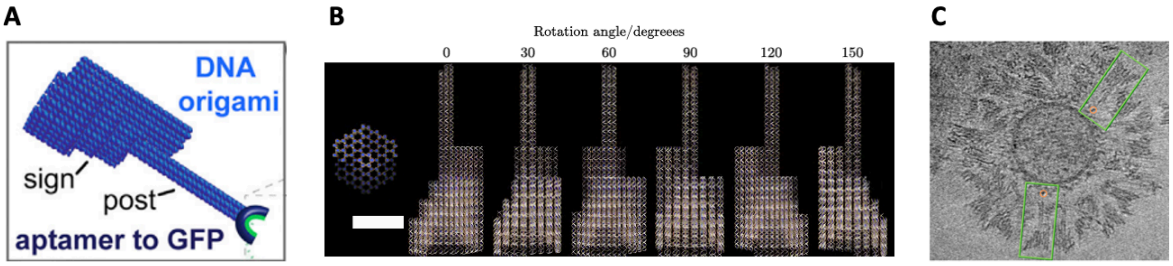


Figure 5.4: DNA origami signposts for cryoelectron tomography. Figures adapted from [72]. *A) Schematic of the signpost. B) Digital rendering of various angles of the nanostructures. C) Tomograph slice of aptamer-conjugated DNA signposts bound to MoMLV.*

Dr. Emma Silvester was unable to intracellularly deliver her DNA signposts into mammalian cells, as the constructs were too big and complex for conventional delivery methods such as electroporation. To show the binding capacities of her DNA signposts, she utilised RNA aptamers to bind to YFP-tagged Env in the envelope of murine leukemia virus (MLV) as seen in Figure 5.4C. Additionally she was able to bind the outer membrane of cells expressing sfGFP-gB [72]. The pattern of binding was similar to the binding observed with α-GFP antibodies conjugated to Alexa-647. With the success of the DNA signposts binding to the outer membranes of cells, and clear distinction through cryoEM, these structures show promise to bind the intracellular architecture if an efficient delivery technique can be achieved.

5.4.2 Tagging intracellular proteins with DNA signposts

To tag an intracellular protein, the same DNA signposts were used as the ones designed by Dr. Emma Silvester. These nanostructures included a GFP aptamer at the end of the "post" side of the structure. I worked with Dr. Antonio Garcia Guerra, a postdoctoral researcher in the Turberfield group, to obtain a doxycycline inducible HEK293T cell line expressing eGFP-cd63-BMP7. This cell line produces GFP on an integral membrane protein known to be predominantly associated with MVB/late endosomes and exosomal membranes [165, 166]. For future experimentation, the tagging of an intracellular protein further within the cell, such as any nuclear component, would support the use of these DNA signposts as intracellular labels.

5.5 Optimising Confinement for Intracellular Delivery

5.5.1 Experimental design

Preliminary experiments were conducted in the Baum lab, under the guidance of Dr. Zaw Win using MCF10A cells and the 4Dcell dynamic cell confiner. The results of these initial tests are depicted in Figure 5.5. Here, the cells were imaged live within the cell confiner. As there was no appropriate cell membrane dye to use at the time, we utilised bright-field to analyse the cell locations and to visualise the intactness of the cell membranes. Figures 5.5 A, B and C represent a

region of unconfined cells after two hours of incubation with 10 nM of DNA signposts. Part B of the figure shows the lowest slice in a confocal stack, at the bottom of the well. Here, you can see that a few of the Cy5-bound DNA origami (red) are stuck to the bottom of the wells. In part C of the figure, you can see the middle slice of the same confocal stack. This region represents approximately the inner portion of the cells. An accumulation of Cy5 signal in part C, as indicated by the yellow arrows, does not overlap with any region of cells visualised in part A or B, therefore no fluorescent DNA signposts are associated with these unconfined cells.

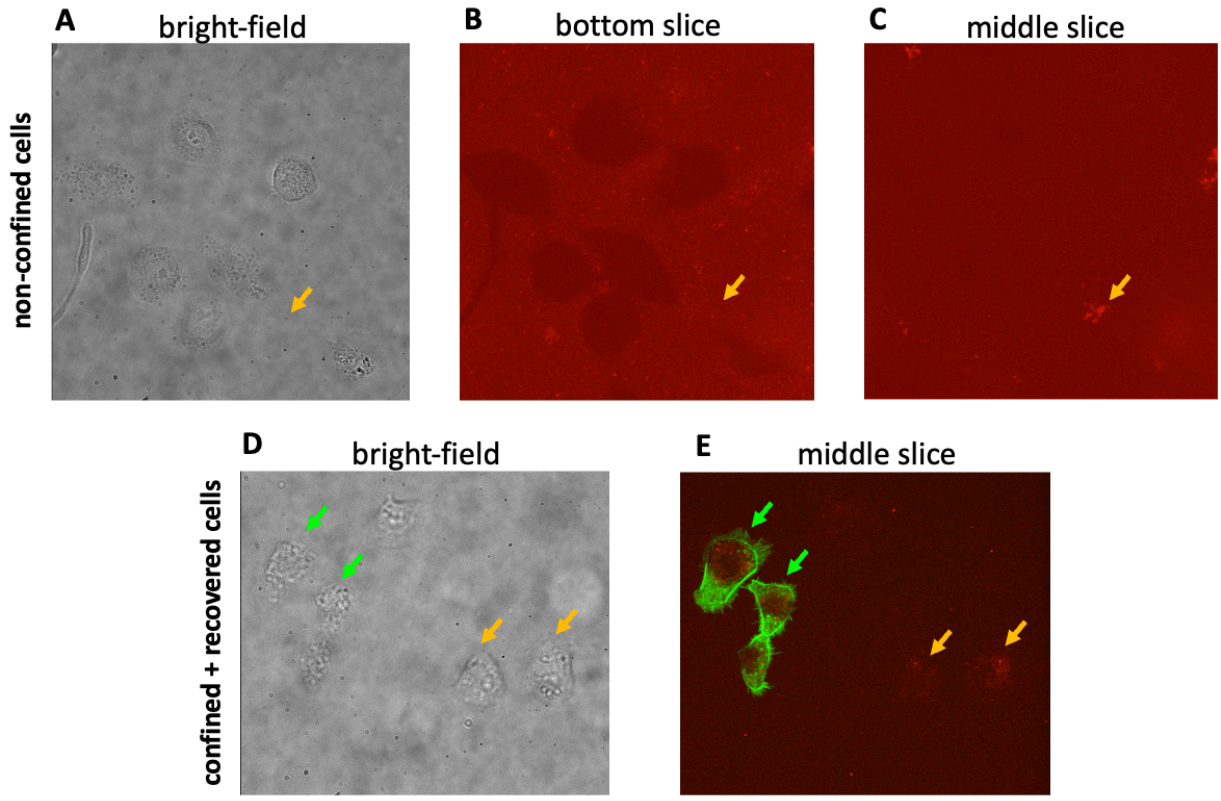


Figure 5.5: MC10A cells under confinement from a dynamic 4Dcell chamber. A-C) Show the non-confined cells incubated with 10 nM of DNA signposts for two hours, while D-E) depict the cells confined from 1 hour with the DNA signposts and recovered for another hour with the DNA signposts. A) and D) depict the cells in brightfield, were as B), C), and E) show confocal slices of fluorescent signals from the cells. In particular, the Cy5-bound DNA signposts are shown in red with a highlighted accumulation of signal indicated by the yellow arrows, while the actin filaments are shown in green due to phalloidin staining as highlighted by the green arrows. Images are taken at 100X magnification with oil immersion.

In contrast to the unconfined cells in Figure 5.5, the cells that are confined under the 4Dcell chambers show significant cellular uptake of the fluorescent DNA signposts as seen in part E of the figure. Here, it appears that the two cells in the bottom right as highlighted by the yellow arrows— you can see the same cells in bright-field in part D of the figure — have more fluorescent nanostructures than any of the non-confined cells. You can also see fluorescence in the cells on the left, but these cells are stained with the actin dye phalloidin (highlighted by the green arrows)

which likely means that their cell membranes have been completely ruptured. This broken membranes theory falls in line with the fact that these cells are full of the DNA signposts as these nanostructures can easily flow into the dead cells. In comparison, the two cells on the right were not stained by the phalloidin, indicating that their cell membranes have recovered from the confinement and are still intact. Therefore, the only way these cells obtained the DNA signposts is if there were transient pores created by the cell confiners to allow for the temporary diffusion of the DNA signposts into the cell.

With this preliminary study, it appears that an hour is an appropriate amount of time to confine the cells and another hour is needed for recovery of the membrane. I therefore, continue with these timeframes for all following experiments. A schematic of the experimental setup is shown in Figure 5.6. Cells are imaged at the start (T0) as a control, then confined for an hour and imaged again (T60), and finally the cells recover for another hour (T120) and are washed and replenished with media free of any DNA signposts. The cells are imaged at this final point, with the only DNA origami signposts associating with the cells being those that have been taken up intracellularly during confinement. A control sample of cells that do not undergo confinement is also imaged at these time points with the same concentration of DNA signposts, 10 nM, included in the media throughout.

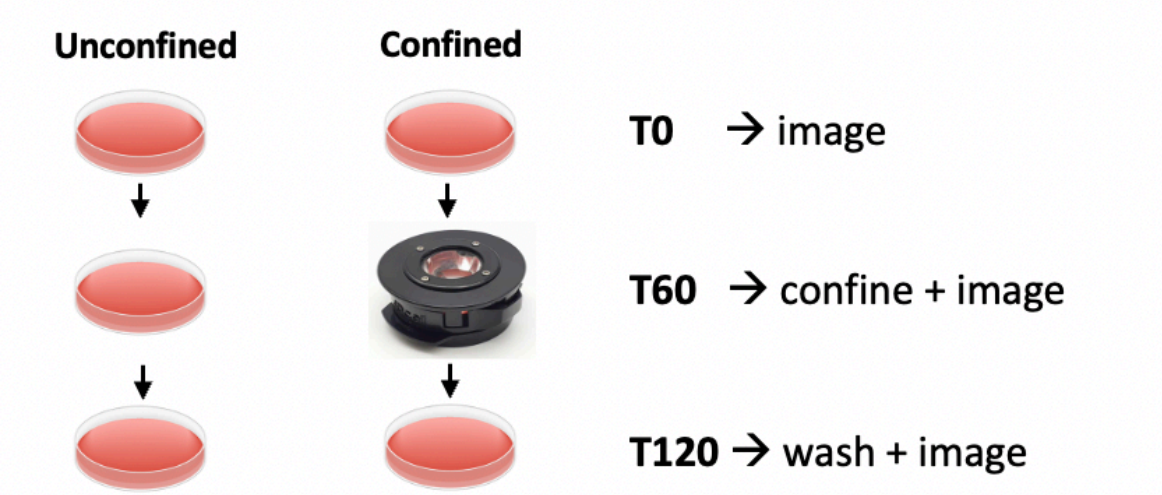


Figure 5.6: Schematic for the experimental setup of cell confinement using a 4Dcell static chamber. One petri dish containing the cells is left as a control (unconfined) while the other is confined for one hour, then released from confinement and allowed to recover for another hour. These confined cells are then washed with PBS and replenished with media free of any DNA signposts, so that any remaining signal from the DNA origami is associated with those delivered into the cell. Images are taken at these three time points, before, during and after confinement. Images are also taken for the control sample at similar timepoints.

This general protocol can be used for any cell line. For the HEK293T cells expressing inducible eGFP-cd63-BMP7, these cells were seeded the night before with approximately 30, 000 cells in a 35 mm petri dish. Then 500 ng/mL of doxycycline is added to start induction of the GFP fused proteins. On the day of the experiment, the cells follow the same protocol as shown in Figure 5.6.

In fluorescent confocal analysis, the green channel is used to identify the expressed eGFP-cd63-BMP7, and the red channel is used for the Cy5-bound DNA signposts. No membrane dye is required as the eGFP-cd63-BMP7 proteins represent the intracellular membrane.

5.5.2 Finding the optimal confinement height and pressure

In addition to the timeframe for confinement, another important parameter to optimise for intracellular delivery is the confinement height and subsequent pressure. For these studies, I used the homemade 12-well static confinement system as shown in Figure 5.7. In part A of the figure, bright-field analysis shows the MCF10A cells after an hour of confinement at 0 kPa (i), 5 kPa (ii), 20 kPa (iii) and 30 kPa (iv). You can see here that 20 and above kPa completely crushes the cells, therefore a confinement height with a pressure of around 5 kPa is most appropriate for maintaining cell viability. A more expansive overview of this experiment is shown in Appendix Figure B.15.

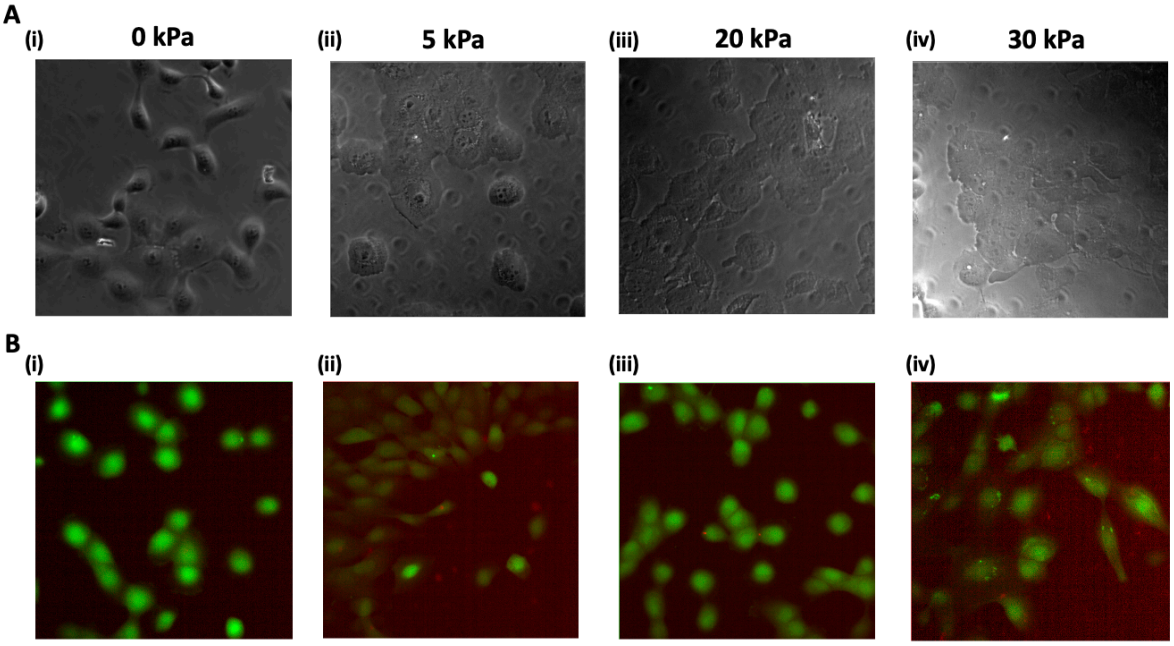


Figure 5.7: Testing out various confinement pressures on MC10A cells. A) showcases bright-field images of cells after confinement, while B) shows a different region of cells after confinement using fluorescence confocal microscopy. All images represent cells after an hour of confinement as well as an hour of recover. (i) is a control at 0 kPa, (ii) is after 5 kPa is applied, (iii) is after 20 kPa is applied and (iv) is after 30 kPa is applied to the cells. In B) the Cy5-bound DNA signposts are shown in red and a CellTracker Green CMFDA dye represents the cell membrane in green. . Images are taken at 100X magnification with oil immersion.

In part B of the same figure, you can see the confocal fluorescence microscopy analysis of another region of cells under the specified confinement pressures. The cells were allowed to recover for an hour after confinement and were washed with media free of DNA signposts. Additionally, the cell membrane dye CellTracker Green CMFDA stains the live cells green. In Figure 5.7B(ii), there are more DNA signposts (in red) associated with the green-stained cells than in part (iv) where the

red fluorescence mostly accumulates in the squashed unstained cells. This means that in part (iv) that the cells have been ruptured allowing the flow of DNA signposts into the dead cells, whereas in part (ii) the cells have had the DNA signposts intracellular delivered during confinement and were able to repair their membranes. In contrast, the cells in part (i) do not have any red fluorescence associated with any of the cells, therefore the DNA signposts do no passively diffuse into the cell without mechanical interference.

5.6 Intracellular delivery and tagging of DNA signposts

5.6.1 Tagging an intracellular membrane protein

After obtaining the optimised parameters for the time and height of confinement necessary for intracellular delivery of DNA signposts into mammalian cells without complete rupturing of the cell membranes, I then moved on to confining HEK293T cells expressing eGFP-cd63-BMP7 to allow for intracellular tagging of the intracellular membrane with the DNA signposts. This experiment was carried out with a 4Dcell static cell confiner with micropillar heights of 4 μm. The experimental setup is the same as discussed in Section 5.5.3. The fluorescence confocal microscopy results of this experiment are shown in Figure 5.8. In part A of this figure, the unconfined cells that have been incubated with 10 nM of DNA signposts is shown. You can see here that the membrane (green) is intact for all these cells, as no confinement was carried out, but there is also no association of the DNA signposts (pink) with any of the cells.

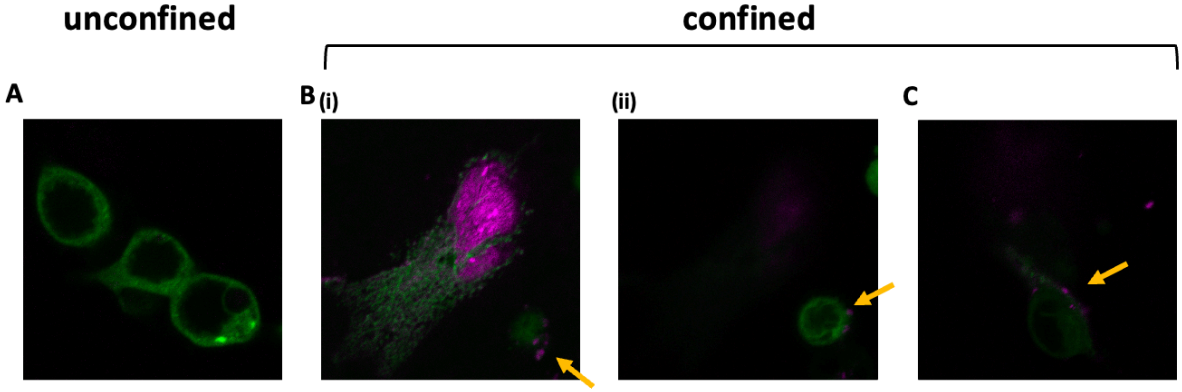


Figure 5.8: Analysis of intracellular membrane tagging of the DNA signposts after confinement. A) Cells without confinement after incubation with 10 nM DNA signposts for two hours. This is a middle slice of a representative confocal stack. B) and C) depict cells after confinement for an hour and recovery for another hour in media containing 10 nM DNA signposts. The cells were washed and replenished without the DNA signposts. B) shows two slices in the same confocal stack where (i) is the top slice for the bottom right cell and (ii) is a mid-way slice. You can see there are fluorescent DNA signpost (in pink) associated with the membrane in both these slices as highlighted by the yellow arrows, indicating internalisation and tagging of the DNA signposts. In C) which is a middle slice of another confocal stack, you can see DNA signposts associated with the intracellular membrane as highlighted by the yellow arrow. In all the images, the HEK293T cells are expressing eGFP-cd63-BMP7, which appears as green as representative of the cell membrane, and the DNA signposts are represented in pink. Images are taken at 100X magnification with oil immersion.

In contrast to the unconfined cells, the confined cells shown in part B and C of this figure show binding of the DNA signposts with the intracellular membrane of the cells as indicated by the yellow arrows. In part B, the top view of the bottom right cell is shown (i) as well as a middle slice (ii) of the confocal stack, representing the approximate mid-way point of the cell. In both cases, you can see small accumulation of the DNA signposts within the green-stained intracellular membrane compartment. This is also the case with any cell shown in part C. In part B(i), there is also a larger cell on the left, that has been completely squashed by the confinement pillars as represented by the flattened membrane. This cell has accumulated massive amounts of the DNA signposts throughout the cell as the membrane is most likely ruptured allowing the DNA origami to flush into the cytosol.

5.6.2 Issues encountered and improvements required

Although the microscopy images in the previous section show promising membrane tagging of intracellularly delivered DNA origami signposts in mammalian cells through cell confinement, there are still various improvements that can be made to confirm this activity. Further characterisation of the cellular uptake of the DNA signposts is necessary. This can be done by quantifying the yield of intracellularly delivered nanostructures. Additionally, I would have liked to compare this delivery method with the efficiency of other conventional intracellular delivery techniques, such as electroporation and lipofection. I could also check for compartmentalisation of the internalised DNA signposts using endosomal and lysosomal dyes.

I experienced various issues throughout these experiments, particularly at the end. I was only able to conduct the experiment twice, before the 4Dcell static chamber started to break the petri dishes when fully closed. I believe the PDMS columns were swelling up after the uses, and therefore the column was changing the height of confinement. The columns are advertised to allow for continuous use as long as they are cleaned between experiments, but this does not seem to be the case for my cell confiner. To counteract this problem, it would have been better if I made my own PDMS columns. This can be done if a mold is made for the exact height required. Therefore, I would be able to use a new column for each experiment. Alternatively the use of the dynamic 4Dcell cell confiners would have also solved this problem as the vacuum pump allows for the fine-tuning of the confinement height. In a previous microfluidics study discussed in Section 5.3.1, real-time uptake experiments showed that dynamic conditions for microfluidic delivery of nanoparticles into cancer cells was more efficient than in static conditions 156]. It would be interesting to make a direct comparison of the static versus the dynamic commercial cell confiners.

Alongside these improvements, the future directions of this work would be to see these experiments carried out with various other DNA nanostructures. Other reconfigurable nanodevices or nanostructures with unique intracellular functionalities would be interesting to observe. Additionally, if further characterisation of the DNA origami signposts is carried out with the cell confiners, it would be ideal to image these structures by cryoelectron tomography to see the distinct binding of the DNA tags to intracellular proteins of interest, and one day be used to characterise a new protein.

5.7 Chapter Summary and Outlook

In this Chapter, I have discussed the idea of using microfluidics and cell confinement as an intracellular delivery tool for introducing complex DNA nanostructures into mammalian cells. I have shown the experimental design and a few optimised parameters for using either a commercial or homemade cell confiner. In particular, I focused on delivering DNA origami cryoEM tags into MCF10A cells as well as a HEK293T cell line expressing eGFP-cd63-BMP7. I used confocal fluorescence microscopy to study the confinement and cellular uptake of the cells. In the absence of a membrane stain, I used bright-field to characterise the cell membranes before, during and after confinement.

In my preliminary experiments using the 4Dcell dynamic cell confiners in the Baum lab, I was able to intracellularly deliver DNA signposts into MCF10A cells. As there was no appropriate cell membrane dye available, I used bright-field to visualise the location of the cells and phalloidin to stain the ruptured cells. Only the ruptured cells and the gently confined cells were able to take up the DNA signposts into the cytosol.

I then used a homemade 12-well cell confiner, also while working in the Baum lab, to test out various gel stiffnesses and therefore confinement pressures on MC10A cells. Here, it was shown that confinement pressures above 20 kPa completely crushed the cells, whereas a confinement pressure closer to 5 kPa was ideal for creating temporary pores in the cell membrane. This confinement pressure is then used for subsequent experiments.

Finally, I was able to work with a HEK293T cells line expressing eGFP-cd63-BMP7 to fully test out the binding capabilities of the DNA origami signposts. As these nanostructures have an RNA aptamer to bind to GFP, once delivered into the cytosol, they can bind to the intracellular membrane proteins expressing GFP. I was able to show binding of the intracellular membrane protein by the DNA signposts after confinement. Further characterisation of intracellular delivery efficiency and comparison to other conventional delivery methods is still required. In the end, I experienced a few issues involving the static 4Dcell cell confiners, with either the materials breaking or the cells not sufficiently being confined. It would be interesting to see these experiments continue on with the dynamic cell confiners in future work.

Chapter 6

Bistable DNA Hairpins as a Reconfigurable Mechanism

6.1 Introduction

In this chapter, I explore the dynamics of reconfigurable short DNA nanostructures that exist in two energetically favourable states. I have designed a trio of three DNA hairpin strands that can switch between conformations through toehold-mediated strand displacement. Additionally, I will compare the change of free energy parameters between NUPACK 3.0 and NUPACK 4.0. Briefly, I will discuss how the study on the dynamics of my reconfigurable DNA hairpins can expand to be used to design larger, more complex DNA nanostructures for intracellular activity.

6.2 Author Contributions

All sequences were designed in NUPACK by Catherine Fan. The gel electrophoresis analysis of the hairpins was also conducted by Catherine Fan. Discussion of the dynamics of the reconfiguration of the system was carried out between Catherine Fan, Professor Andrew Turberfield, and another student in the Turberfield lab, Joakim Bohlin.

6.3 Using NUPACK to design my DNA hairpins

6.3.1 Introduction to NUPACK

NUPACK is an online interface within the Beckman Institute at Caltech [107, 108, 109], that can be used for the analysis and design of nucleic acid structures. This software allows for the study of complex and test tube ensembles containing arbitrary numbers of interacting strand species. Additionally, the NUPACK algorithms can be used alongside Python to create more computationally heavy nucleic acid designs and reaction calculations.

In NUPACK, DNA secondary structures are defined as a set of base pairs corresponding to Watson-Crick pairs. Complex ensemble combines all of the secondary structures for an arbitrary number of interacting RNA or DNA strands. A polymer graph representation of a secondary structure is constructed by ordering the strands around a circle, drawing the backbones in succession from 5′ to 3′ around the circumference with a nick between each strand and drawing straight lines connecting the paired bases. A secondary structure is pseudoknotted if all the strand orderings within the polymer graph contain crossing lines. In NUPACK, pseudoknots are excluded from the structural ensemble due to the complexity of the design [167]. A NUPACK reaction can also become more complex by creating a test tube ensemble, which is a dilute solution containing a set of strand species at specified concentrations, that interact to form a set of complex species.

For each non-pseudoknotted secondary structure, the free energy is estimated in NUPACK as the sum of the empirically determined free energies of the constituent loops plus a strand association penalty. These loop free energies are modeled for the different loop types, including the hairpin, interior loop, multi-loop, and exterior loop. The free energy parameters in NUPACK are temperature-dependent and are based on various studies on the thermodynamics of DNA and RNA [110, 111, 116, 117].

In NUPACK, the equilibrium probability is generated to calculate the probability of a base pair forming at equilibrium within an ensemble. Additionally the complex ensemble defect quantifies the equilibrium number of incorrectly paired nucleotides over the ensemble. The average number of nucleotides that are incorrectly paired at equilibrium relative to the specified secondary structure are evaluated over the Boltzmann-weighted ensemble of non-pseudoknotted secondary structures. Finally a normalised complex ensemble defect is calculated to represent the equilibrium fraction of incorrectly paired nucleotides evaluated over the ensemble relative to the target structure. A normalised complex ensemble defect near 0% is closest to the ideal target structure.

The NUPACK web app contains three pages for analysis, design and utilities. The analysis page analyses the equilibrium base-pairing properties of one or more test tube ensembles and subsidiary complex ensembles. Additionally, the melting curves of the structures can be generated here. The design page creates the sequences for one or more test tube ensembles and subsidiary complex ensembles. Sequence design is performed subject to diverse user-specified sequence constraints including composition, complementarity, pattern prevention and biological constraints. And finally, the utilities page combines all the functions by analysing, designing and preparing figures for a single complex ensemble. Alongside the use of the web interface, NUPACK 4.0 is also a C++ library distributed as a Python package. As a Python package, NUPACK can be used to generate more customisable and complex sequence designs and analyses.

6.3.3 Designing my DNA hairpins in NUPACK

To study the basics of reconfigurable nucleic acid design, I focused on creating the simplest DNA structure that could switch between two conformations. Therefore, I created various DNA hairpins designs that can switch between an A or B domain hairpin stem. The simplest hairpin was designed with the A and B domains each being 8 nt long as shown in Figure 6.1. As the strand order is A-B-A*-B*, where the * domain is a complementary sequence, the resulting hairpin can exists in either a B stem (blue) hairpin with a loop domain of A* (Figure 6.1A) or a A stem (red) hairpin with a loop domain of B (Figure 6.1B). This “Short Stem Hairpin” design can be switched between the A and B stem conformations and back again, using toehold mediated strand displacement of an 8 nt C domain (pink) or 8 nt E domain (green). Binding of the C domain (Figure 6.1A) displaces the B stem, switching the structure to the A stem state (Figure 6.1B), whiling binding of the E domain (Figure 6.1C) displaces the A stem, switching the structure to the B stem state (Figure 6.1D). The initial toehold strands are removed by secondary toehold displacement strands as shown in parts B and D of this figure. The C* displacement strand is removed through binding of the D (yellow) domain, while the E* displacement strand is removed through binding of the F (light blue) domain. As both A and B hairpin conformations are designed to have equal free energies, once the displacement strands are removed the structure is maintained in its new state.

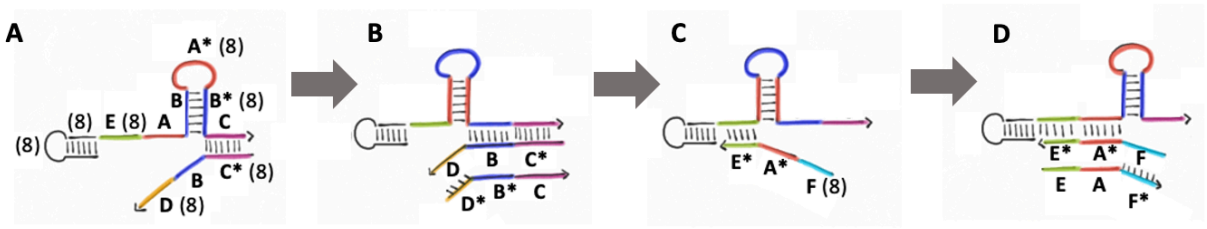


Figure 6.1: Schematics of the short stem hairpin design. The A (red) and B (blue) hairpin stems are 8 nt each, resulting in short stemmed hairpins in either conformation. Toehold displacement strands switch the structure from A) B stem hairpin into B) A stem hairpin. The strands bind to either the C (pink) or E (green) domains of the hairpin. Additional displacement strands are used to remove the initial displacement strands. The reaction can go from A) to B) to C) to D) back to A) or the other way around depending on the displacement strands used or through transient self-interactions. The structure also contains a 8 bp hairpin motif on the 5' end of the structure to aid analysis as the two structures run asymmetrically through the gel.

In Figure 6.2, a "Long Stem Hairpin" design is shown with the same domains as the "Short Stem Hairpin", except that the A and B domains are now 18 bases long to increase the stability of the hairpin stems. It is interesting to note that in this longer stem system, the hairpin loop is also increased to 18 nt, which may result in additional secondary structures.

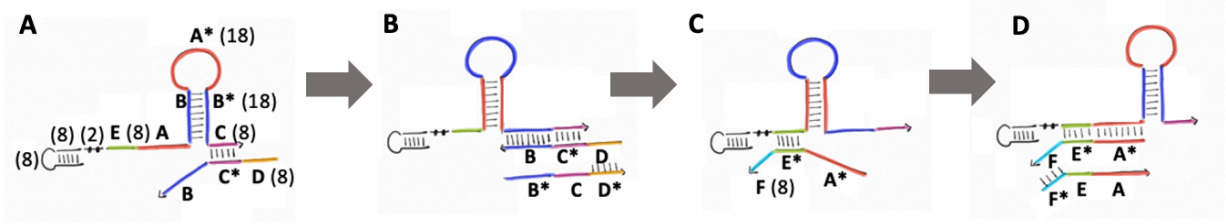


Figure 6.2: Schematics of the long stem hairpin design. The A (red) and B (blue) domains are 18 nt each, resulting in longer stemmed hairpins in either conformation. Toehold displacement strands switch the structure from A) B stem hairpin into B) A stem hairpin. The strands bind to either the C (pink) or E (green) domains of the hairpin. Additional displacement strands are used to remove the initial displacement strands. The reaction can go from A) to B) to C) to D) back to A) or the other way around depending on the displacement strands used or through transient self-interactions. The structure also contains a 8 bp hairpin motif of the 5' end to add gel electrophoresis analysis.

An alternative system to the "Long Stem Hairpin" is shown in Figure 6.3, where the A and B domains are kept at 8 nt each and a new 10 nt α domain (teal) is added to the sequence to increase the hairpin stem lengths to 18 bp without increasing the hairpin loop length. An α domain is placed

between the A and B domains, while an α* domain is placed between the B and A* domains, and a final α domain exists between the A* and B* domains. This design is termed the "Extended Long Stem Hairpins". An α* domain is also added to the toehold displacement strands to unfold the hairpin stems in either conformation.

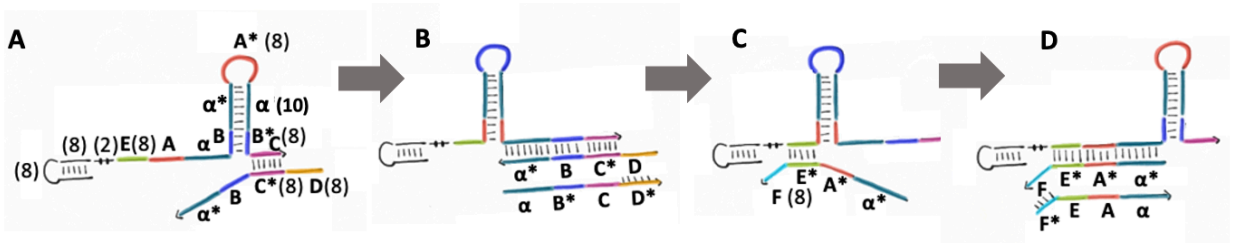


Figure 6.3: Schematics of the extended long stem hairpin design. The A (red) and B (blue) domains are 8 nt long and additional flanking 10 nt α domains (teal) are added in between the A and B domains to increase the hairpin stems to 18 bp without increasing the length of the hairpin loops. Toehold displacement strands switch the structure from A) B stem hairpin into B) A stem hairpin. The strands bind to either the C (pink) or E (green) domains of the hairpin. Additional displacement strands are used to remove the initial displacement strands. The reaction can go from A) to B) to C) to D) back to A) or the other way around depending on the displacement strands used or through transient self-interactions. The structure also contains a 8 bp hairpin motif of the 5′ end to add gel electrophoresis analysis.

Each of these three hairpin designs contains an additional hairpin with an 8 bp stem and 8 nt loop, which has been placed arbitrarily on the 5′ of the strand. This motif allows the strand to run distinctly on a polyacrylamide gel as the A and B stemmed hairpins are now asymmetrical. Initially, when designed without this additional hairpin motif, the two conformations were mirrors of each other and therefore ran at exactly the same length on a polyacrylamide gel.

As mentioned, the hairpin designs require two energetically favourable states to allow the system to exist in two different conformations. This equilibrium energy landscape can be created in NUPACK by choosing domain sequences that allow the A and B stem conformations to have similar free energies of their respective secondary structures. This is shown in Figure 6.4A where the Short Stem Hairpin design has a free energy of -17.73 kcal/mol in the A stem conformation (i) and a free energy of -18.12 in the B stem conformation (ii). The normalised ensembled defect for either conformation is also selected to be as similar as possible at 29% for the A state and 19% for the B state. These values are relatively equal.

In Figure 6.4B, the Long Stem Hairpin design is shown with a free energy of -30.71 kcal/mol for the A stem conformation and -31.20 kcal/mol for the B stem conformation. The normalised ensemble defects are 31 % and 36%, respectively. In part C of this figure, the predicted secondary structure of the Extended Long Stem Hairpin design is shown with a free energy of -32.39 kcal/mol in the A stem conformation and -32.45 kcal in the B stem conformation. The normalised ensemble defects are 29 % and 34%, respectively.

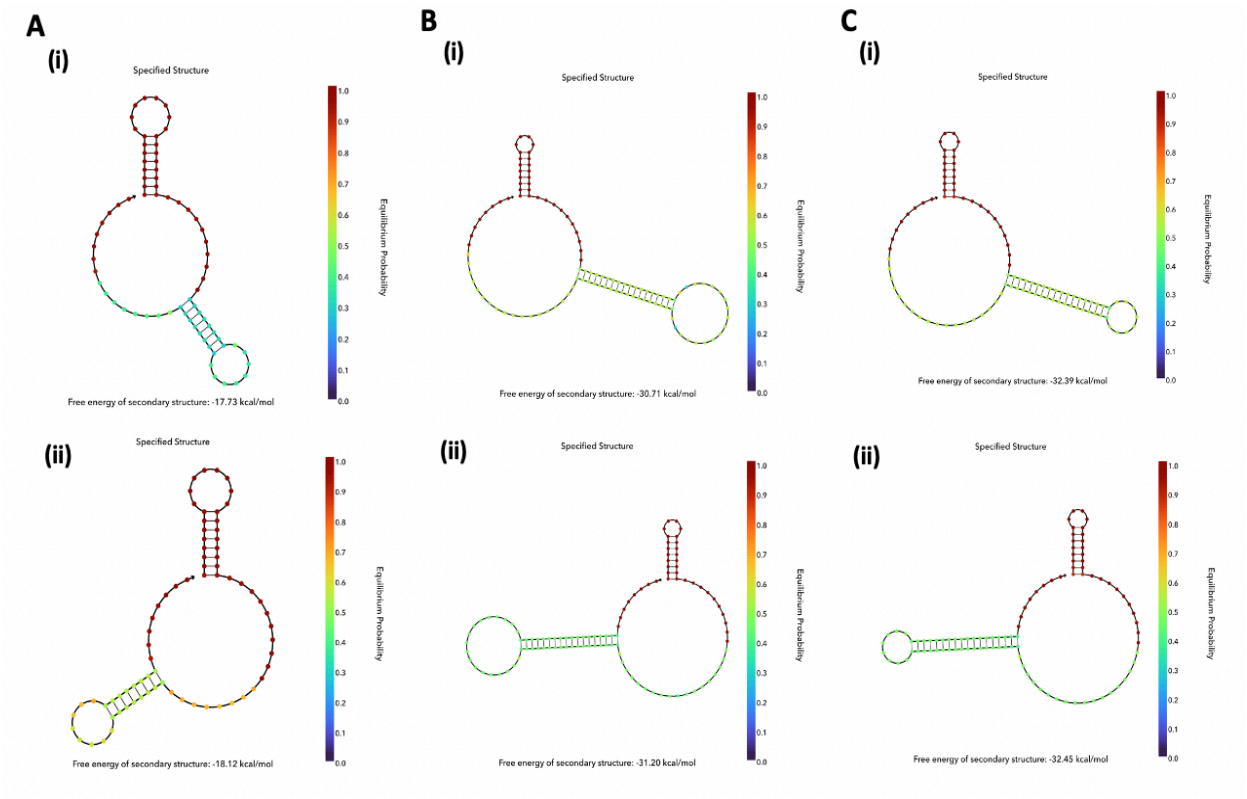


Figure 6.4: Predicted secondary structures for the three reconfigurable DNA hairpin designs as predicted by NUPACK 3.0. A) Short stem hairpin design in the (i) A stem conformation and (ii) B stem conformation. B) Long stem hairpin design in the (i) A stem conformation and (ii) B stem conformation. C) Extended long stem hairpin design in the (i) A stem conformation and (ii) B stem conformation. The bases are coloured based on their equilibrium probability to exist in that state, where red is likely and blue is less likely.

6.3.3 Comparing NUPACK 3.0 and NUPACK 4.0

When I first started these experiments, NUPACK 3.0 [107] was the canonical version of the software, whereas now NUPACK 4.0 [108, 109] is the updated version. The new NUPACK web app runs on a scalable hybrid cloud computer cluster, and also utilises slightly different free energy parameters. Although the thermodynamic parameters previously described by SantaLucia Jr and Hicks [111] are the same, in NUPACK 3.0, a G · T was treated as a wobble pair instead of classifying G and T as a mismatch as done in NUPACK 4.0. NUPACK 3.0 also prohibited ensembles to contain a wobble pair as a terminal base pair in an exterior loop or a multiloop. Therefore, terminal mismatch free energies in exterior loops and multiloops are replaced by two dangle stacking — formed with at most two adjacent unpaired bases — free energies. NUPACK 3.0 had dangle stacking options without coaxial stacking, whereas NUPACK 4.0 allows for the stacking of a base pair with an adjacent base pair within an exterior loop or multiloop.

These changes in the software affect the free energy calculations, resulting in different predicted normalised ensemble defects for the exact same sequences designed using NUPACK 3.0. These changes can be seen in Figure 6.5. In part A of this figure, the changes to the Short stem hairpin design predictions is shown, where the A stem conformation now has a normalised ensembled defect of 28% (i) whereas the B conformation is 17%. Therefore, the B stem state is predicted to be slightly more stable with the new free energy parameters, and may have preference in forming compared to the A stem state. In part B of the figure for the Long stem hairpin design, the A stem conformation now has a normalised ensembled defect of 43% (i) whereas the B conformation is 23%. This new predicted also favours the B stem state more than when predicted by NUPACK 3.0. Finally, for the Extended long stem hairpin design shown in part C of the figure, the normalised ensemble defects are similar to those predicted in NUPACK 3.0. The A stem conformation has a new predicted 30% ensemble defect, while the B stem conformation is predicted to be 33%, therefore both states have similar probabilities to form than the other designs.

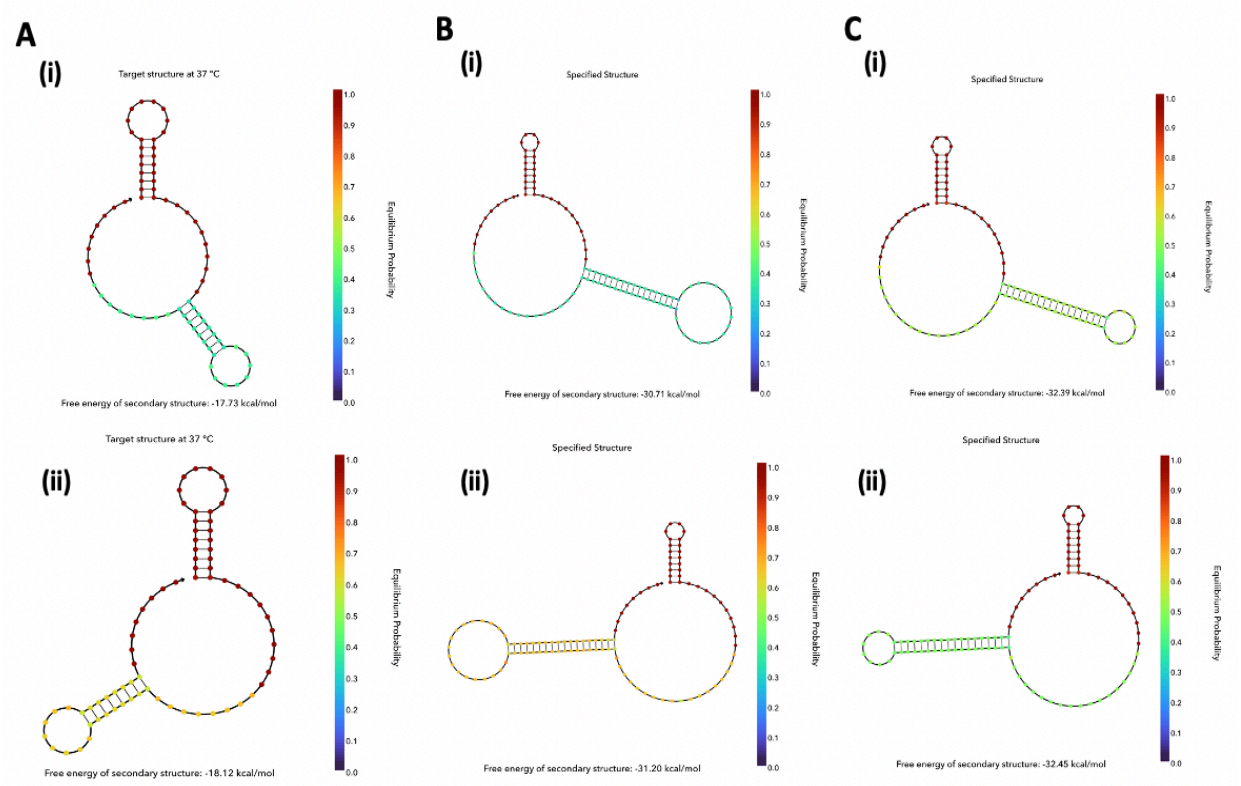


Figure 6.5: Predicted secondary structures for the three reconfigurable DNA hairpin designs as predicted by NUPACK 4.0. A) Short stem hairpin design in the (i) A stem conformation and (ii) B stem conformation. B) Long stem hairpin design in the (i) A stem conformation and (ii) B stem conformation. C) Extended long stem hairpin design in the (i) A stem conformation and (ii) B stem conformation. The bases are coloured based on their equilibrium probability to exist in that state, where red is likely and blue is less likely.

6.4 Analysing the Bistable Hairpin Designs

6.4.1 Reaction parameters with each hairpin design

Using the predicted secondary structures from NUPACK 3.0 at the time, I ordered the sequences to be synthesised by Integrated DNA Technologies (IDT), and analysed the reconfigurable reactions by polyacrylamide gel electrophoresis (PAGE). A 15% polyacrylamide gel was used, which was run for 90 minutes at 90 V. In cases where the gel would heat up due to the applied current, the tank was placed in a cold room (4 °C) to prevent denaturation of the samples.

For each reconfigurable design, a hairpin sequence with two favourable free energy states, termed the AB strand, is synthesised. Additionally, toehold mediated displacement strands are also synthesised to switch the AB hairpin into either the A or B stem conformation, and secondary displacement strands are synthesised to remove the first strands from the hairpins. Control strands where also synthesised, where the structure is designed to exist only in the A stem state or the B stem state. These control strands are used as a reference point for the size of the two conformational shapes.

Before combining any of the components, the AB hairpins were quenched for 30 seconds at 95 °C and then 30 seconds on ice. This allows the structure to reform into its equilibrium state, which should be both the A stem and B stem conformations. This prompt quenching reaction should prevent any kinetic traps from occurring within these bistable hairpins. To combine the toehold displacement strands with the hairpins, I annealed the strands together by bringing the temperature of the reaction down from 95 °C to 15 °C at a rate of -1 °C per min.

6.4.2 Short stem hairpins

An abridged PAGE analysis of the reconfiguration of the short stem hairpin design is shown in Figure 6.6, with the full gel available in Appendix Figure B.16 and an experimental replicate shown in Appendix Figure B.17. In these figures, you can see that the AB hairpin exists as a single band in lane 1 of the gel that runs in between the sizes of the control bands in lanes 4 and 5. Therefore, the short stem AB hairpin most likely exists in an intermediate state than in either the A or B stem conformations. This is further proven after toehold-mediated strand displacement in lanes 2 and 3 of the gel, where the AB hairpins should have been switched into the A or B stem conformations, respectively, but instead they maintain a single band at the same height on the gel as the original AB hairpin.

One reason for this intermediate state, is that the hairpin stems in either conformation are only 8 bp long. This length may not be stable enough to maintain the switched states and the resulting hairpins may be transiently switching between conformations on its own. A more stable hairpin structure, with a longer stem, is required for this system. It is difficult to confirm whether the NUPACK 3.0 or NUPACK 4.0 predicted ensemble defects were more accurate for this Short stem hairpin design as the desired conformations of the structures are not achieved.

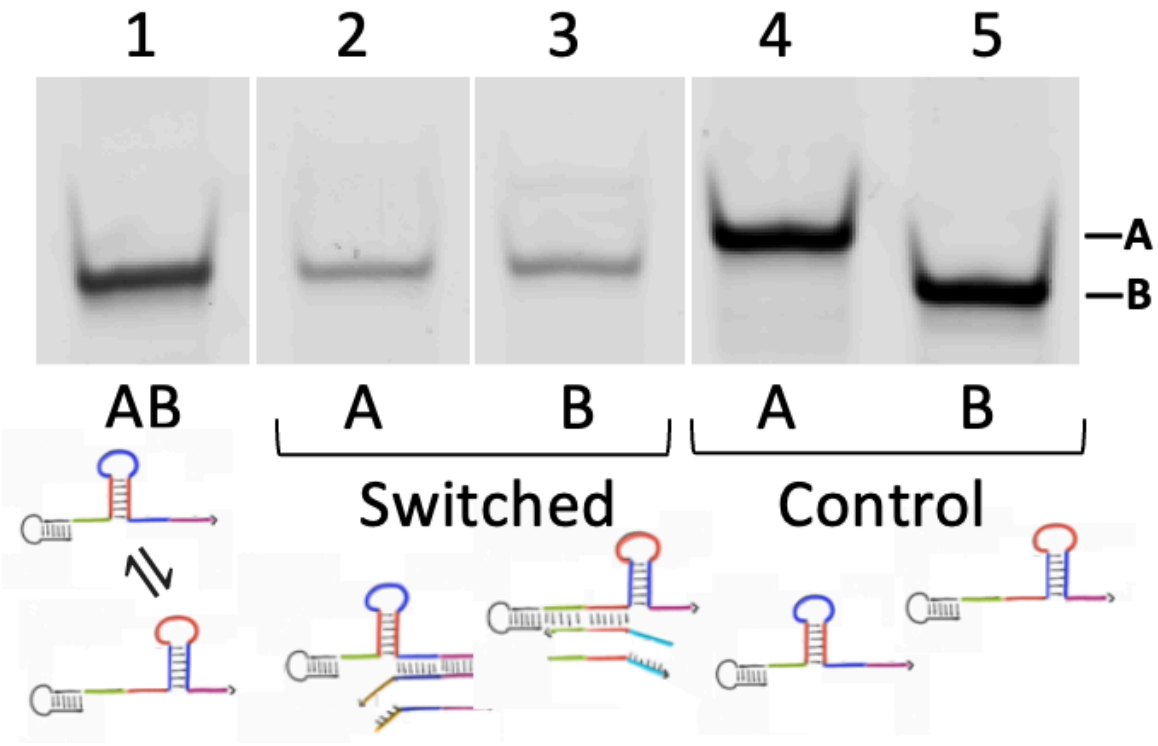


Figure 6.6: PAGE analysis of the short stem hairpin design. Lane 1 represents the AB hairpin. Lane 2 represents the AB hairpin switched into the A stem conformation after toehold-mediated strand displacement (TMSD). Lane 3 represents the AB hairpin switched into the B stem conformation after TMSD. Lanes 4 and 5 show the controls for the A stem hairpin and B stem hairpin, respectively. The samples are stained with SYBRGold. The lanes have been stitched together from the same representative gel. Schematics of the existing hairpins in each state is shown below. Samples are run on 10% polyacrylamide gels.

6.4.3 Long stem hairpins

As the short stem hairpin design failed to stably switch between two conformational states, I designed a hairpin with a longer stem to increase the stability of the structure in either the A or B stem states. An abridged PAGE analysis of the reconfiguration of the long stem hairpin design is shown in Figure 6.7, with the full gel available in Appendix Figure B.18 and two experimental replicates shown in Appendix Figure B.19. In contrast to the short stem design, the long stem hairpins are shown to exist in both the A and B conformations as depicted in lane 1. It is interesting to note that the B conformation is preferred to the A conformation, with 92.35% preferring the B state as calculated from four separate reactions. This outcome is consistent with the NUPACK 4.0 predictions for the normalised ensemble defects, where the B conformation was predicted to have double the probability of forming compared to the A conformation.

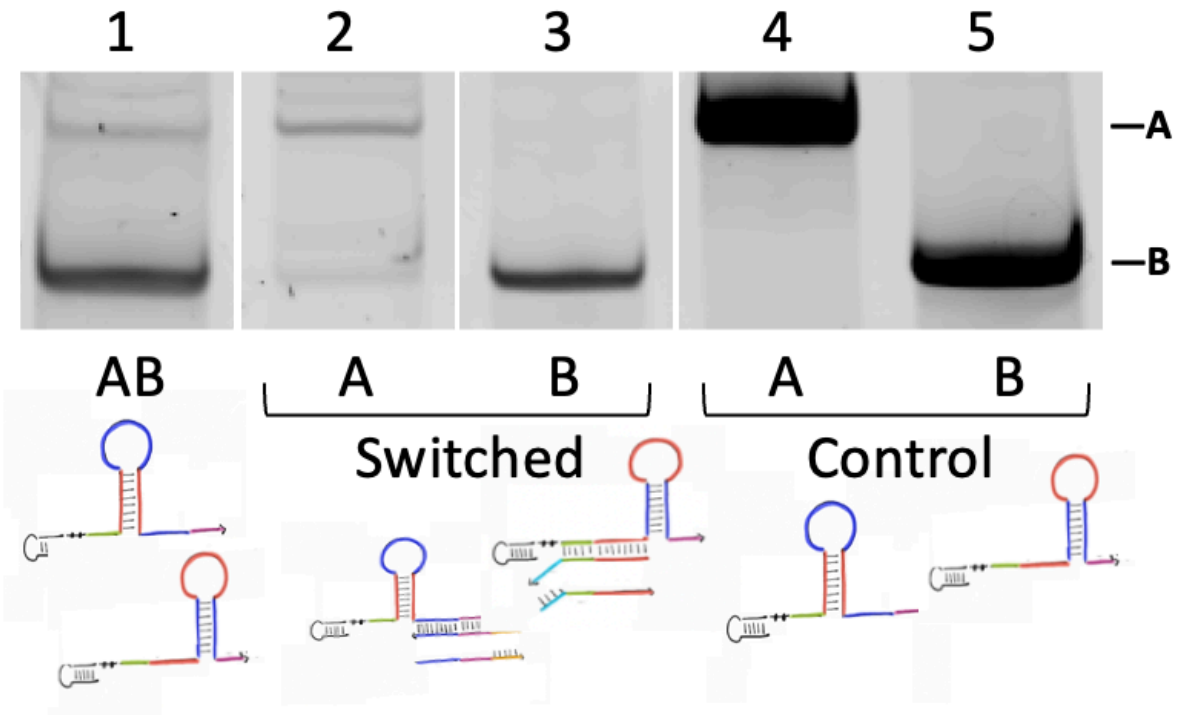


Figure 6.7: PAGE analysis of the long stem hairpin design. Lane 1 represents the AB hairpin, 7.65% of which exists in the A stem state and 92.35% exists in the B stem state. Lane 2 represents the AB hairpin switched into the A stem conformation after toehold-mediated strand displacement (TMSD) where 70.57% of structures exist in the A stem state and 29.43% remain in the B stem state. Lane 3 represents the AB hairpin switched into the B stem conformation after TMSD where 94.9% of structures exist in the B stem state and 5.1% remain in the A stem state. The band intensity in the A or B conformation is quantified from four gel replicates of separate reaction samples. Lanes 4 and 5 show the controls for the A stem hairpin and B stem hairpin, respectively. The samples in the 10% polyacrylamide gel are stained with SYBRGold. The lanes have been stitched together from the same representative gel. Schematics of the existing hairpins in each state is shown below.

Despite the B stem confirmation possessing a better predicted normalised ensemble defect, after toehold mediated strand displacement, majority of the structures can exist in the A stem conformation as shown in lane 2 of Figure 6.5. Here, 70.57% of the structures are in the A state. When the hairpin is switched to the B conformation, 94.9% of the structures maintain the B stem state as shown in lane 3.

6.4.4 Extended long stem hairpins

Although the long stem hairpin design was successful in switching conformations, the probability to switch between the states was not equal. This may be due to the longer hairpin loop region resulting from the increased A and B domains, which can potentially result in unwanted secondary

interactions not predicted in NUPACK. To increase the hairpin stems without increasing the hairpin loops, I designed an extended long stem hairpin structure where the A and B domains remain as 8 nt each, and additional 10 nt long α domains is added in between the A and B sequences to only increase the hairpin stem length. An abridged PAGE analysis of the reconfiguration of the extended long stem hairpin design is shown in Figure 6.8, with the full gel available in Appendix Figure B.20 and two experimental replicates shown in Appendix Figure B.21. In this new system, the AB hairpin exists in either conformations at a better proportion than in the long stem hairpin system, but the B conformation is still preferred with 71.99% of the structures existing in the B stem state as seen in lane 1 of the gel. It is interesting to note in this system that the AB hairpins have decreased stability compared to the long stem hairpin design when switched in either the A or B conformations. As seen in lanes 2 and 3 of the gel in Figure 6.6, the A switched reaction only has 38.2% of structures in the A stem state and the B switched reaction only has 72.58% of structures in the B stem state.

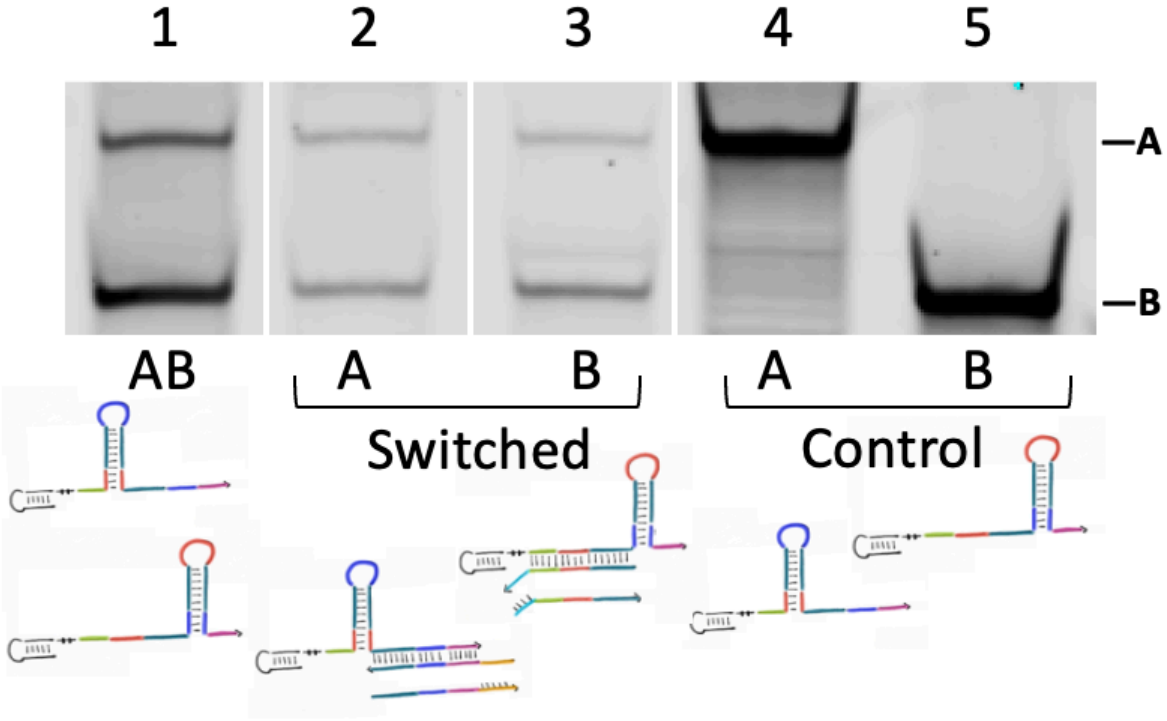


Figure 6.8: PAGE analysis of the extended long stem hairpin design. Lane 1 represents the AB hairpin where 28.01% of structures exist in the A stem state and 71.99% exist in the B stem state. Lane 2 represents the AB hairpin switched into the A stem conformation after toehold-mediated strand displacement (TMSD) where 38.2% of structures exist in the A stem state and 61.8% remain in the B stem state. Lane 3 represents the AB hairpin switched into the B stem conformation after TMSD where 72.58% of structures exist in the B stem state and 27.42% remain in the A stem state. The band intensity in the A or B conformation is quantified from four gel replicates of separate reaction samples. Lanes 4 and 5 show the controls for the A stem hairpin and B stem hairpin, respectively. The samples in the 10% polyacrylamide gel are stained with SYBRGold. The lanes have been stitched together from the same representative gel. Schematics of the existing hairpins in each state is shown below.

Despite being predicted by both NUPACK versions 3.0 and 4.0, to have similar normalisation ensembled defects in the A and B stem conformations, the analysis in Figure 6.8 shows that the switch conformations are not entirely stable. This phenomenon may be due to pseudoknot interactions which cannot be predicted by NUPACK. Specifically, it could be that the additional α domain is flexible enough to displace the α-α* domain in the hairpin stem as depicted in Figure 6.9. These types of interactions show the limitations of NUPACK as a nucleic acid structural analysis software. Therefore, it would be interesting to model the dynamics of the extended long stem hairpins in simulation, such as through the coarse-grained modelling program, oxDNA.

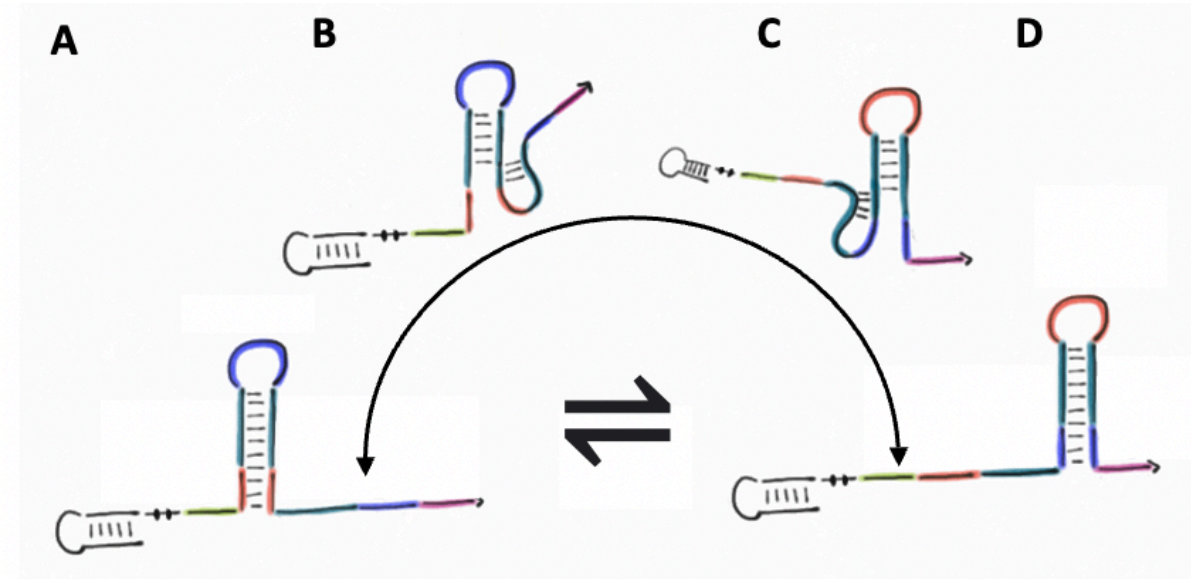


Figure 6.9: Schematic of the potential reconfigurable dynamics of the extended long stem hairpin. A) The A stem hairpin may switch into the B stem hairpin in D) through self-displacing interactions of the α domain (in teal) as seen in B) and C).

6.4.5 Mirrored extended long stem hairpins

An interesting aspect to the extended long stem hairpins is that the B stem state is preferred even when switched into the A stem state. This is seen in lane 2 of Figure 6.8 where 61.8% of the structures are in the B state, even though the displacement strands where used to switch the system into the A stem state. One theory for this preference is that in the potential self-interaction dynamics discussed in Figure 6.9, the α domain closest to the 3′ prime end of the strand has more flexibility compared to the α domain closest to the 5′ end of the strand due the hairpin motif on the 5′ end. Therefore, the 3′ end α domain can better displace the α-α* interaction in the A stem, spontaneously switching the hairpin into the B conformation, compared to the 5′ end α domain, which may struggle to self-displace the α*-α interaction in the B stem.

To test this theory, I created a mirrored version of the extended long stem hairpin as shown in Figure 6.10A, with a full gel available in Appendix Figure B.22 with an experimental replicate

shown in Appendix Figure B.23. In this the design, the additional 8 bp hairpin motif has been moved from the 5′ end of the sequence to the 3′ end. Comparable to the other hairpin designs, this mirrored model was predicted in NUPACK to have similar free energies, at -32.46 kcal/mol in the A state and -32.52 kcal/mol in the B state, as well as similar normalised ensemble defects, at 29% in the A state and 35% in the B state as shown in Figure 6.10B.

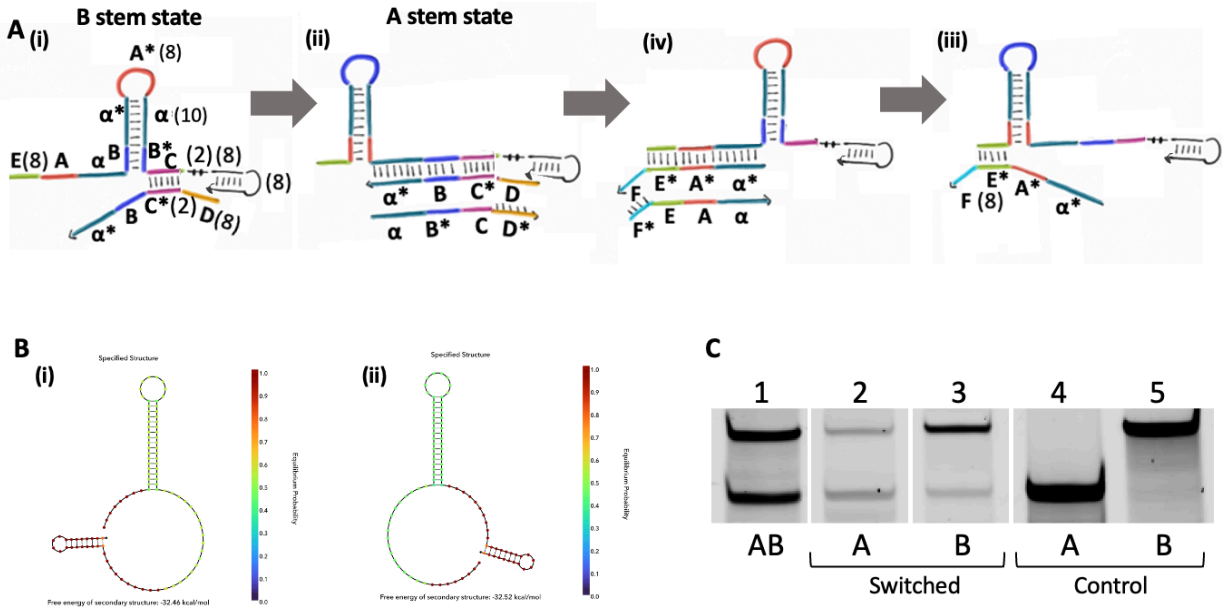


Figure 6.10: Overview of the mirrored extended long stem hairpin design. A) Schematic of the reconfigurable dynamics of the hairpin. (i) Shows the B stem conformation which can be switched into the A stem conformation (ii) and back (iv). B) NUPACK 4.0 predicted secondary structures. (i) The A stem conformation is predicted to have a free energy of -32.46 kcal/mol and a normalised ensemble defect of 29%. (ii) The B stem conformation is predicted to have a free energy of -32.52 kcal/mol and a normalised ensemble defect of 35%. C) PAGE analysis of the reconfigurable switching. Lane 1 represents the AB hairpin where 54.31% of structures exist in the B stem state and 45.69% exist in the A stem state. Lane 2 represents the AB hairpin switched into the A stem conformation after toehold-mediated strand displacement (TMSD) where 73.08% of structures exist in the A stem state and 26.92% remain in the B stem state. Lane 3 represents the AB hairpin switched into the B stem conformation after TMSD where 67.6% of structures exist in the B stem state and 32.4% remain in the A stem state. The band intensity in the A or B conformation is quantified from two gel replicates of separate reaction samples. Lanes 4 and 5 show the controls for the A stem hairpin and B stem hairpin, respectively. Samples are run on a 10% polyacrylamide gel.

If my theory is correct, the mirrored extended long stem hairpin should have preference for the A stem state even when switched into the B conformation. When analysing the reaction by PAGE (Figure 6.10C), it appears that this is not the case. Although the A state is now preferred at 73.08% when switched into the A conformation (lane 2), the A state is not preferred when the hairpin is

switched into the B conformation, with only 32.4% of structures displacing into the A conformation (lane 3). Therefore, there must be other unknown dynamics involved in the reconfigurations of the extended long stem hairpins.

6.5 Chapter Summary and Outlook

In this Chapter, I showcased the reconfigurable dynamics of three bistable hairpins. I designed a short stem hairpin, a long stem hairpin and an extended stem hairpin and analysed the toehold-mediated conformation switching of these structures by gel electrophoresis. Each hairpin should have two favourable energetic conformations, the A stem or B stem states. The results of these experiments reinforced the reliable predictability of NUPACK for generating simple secondary structures, but also revealed potential self-interactions outside the realm of the NUPACK algorithms.

For my short stem hairpin design, it was shown that the 8 bp stem was too short to stably maintain either the A or B stem conformations. In the gel, the unperturbed hairpin existed in an intermediate stage between the A and B hairpin controls. When the toehold displacement strands where added, the hairpin stayed in this intermediate state.

To create a more stable hairpin, I designed a longer 18 bp stem hairpin. This design was able to switch in between the A and B conformations based on the toehold displacement strands added. The system appeared to prefer the B stem state, which was also predicted by NUPACK. The unequal probability of conformation switching may be due to the extended hairpin loop region, which may be causing unwanted secondary interactions.

To improve on the longer stem hairpin design, I created an extended long stem hairpin that utilises α domains to increase the length of the hairpin stem to 18 bp, while maintaining a 8 nt hairpin loop. This system showed the most interesting results, as both the A and B conformations where preferred no matter which toehold displacement strands were used. My theory for this phenomenon is that the α domains may be self-displacing the α-α* interactions in either the A or B stems, resulting in spontaneous conformational switching without the aid of external strands.

When analysing this gel bands for the extended long stem hairpin design, although both the A and B conformational bands appear, the B stem state is slightly more preferred. I hypothesised this may be cause by the additional 8 bp hairpin motif used to make the structures asymmetrical to aid analysis on a polyacrylamide gel. This motif was arbitrary placed at the 5′ end of the structure, which may cause the α domain closest to 5′ end to have less flexibility, and therefore displace the B stem state less than α domain closest to the 3′ end which doesn't have any associated secondary structures. Therefore, I moved the hairpin motif to the 3′ end of the structure to see if the A conformation was preferred. Although this mirrored extended long stem hairpin had improved preference of the A stem state when switched to the A conformation by TMSD, it did not have increased preference to the A state when switched to the B conformation by TMSD as I hypothesised. Therefore, other dynamic interactions must be taking place to result in the B conformation preference in the extended long stem hairpin design. This may be due to sequence preference, but it will be hard to analyse this in NUPACK as this interaction is most likely

pseudoknot-based. Instead, the coarse-grained modelling software oxDNA may be more useful for simulating the dynamic interactions of my reconfigurable hairpins.

In oxDNA, I would be able to reconstruct the hairpins with the same sequences used in my gel electrophoresis analyses. I can then compare my PAGE results with the behaviour of the simulated reactions. For the short stem hairpins, I would expect them to transiently open and close the hairpin stems in either conformation, without the need of an external mutual trap. On the other hand, I anticipate that the long stem hairpin will predominately stay in the B stem conformation. The use of a mutual trap may force the long stem hairpins into the A stem state. As for the extended long stem hairpins, it will be interesting to see how the structure switches between conformations, and confirm if the reconfiguration occurs due to self-displacing interactions from the α domains. I would expect some sort of spontaneous switching of the extended long stem hairpins without the need of mutual traps.

My three bistable hairpin designs reveal interesting reconfigurable dynamics of short DNA strands, particularly with the hairpin motif. These hairpins can potentially be used in larger nucleic acid nanostructures as a reconfigurable tool. The applications for these hairpins, particularly within the cell, can include biosensors that can switch conformations upon reaching a specific cellular target, reconfiguration of the overall structure to aid intracellular delivery of the nanostructure, or even become dynamic fluorescence probes. By incorporating a split aptamer sequence in the hairpin design, I could have the hairpins switch on and off fluorescence in regards to a dynamic reaction within the cell. More work is required to achieve these applications. Future research on these reconfigurable hairpins should focus on modelling the system in oxDNA as well as optimising the domain sequences to fine-tune the reconfigurability of the hairpins.

Chapter 7

Conclusion & Future Research Directions

7.1 Summary and Discussion

Nucleic acids offer a molecular engineering technology that allows for *de novo* design of a complex and diverse set of biomolecular systems. The structure and characteristics of DNA and RNA provides a platform for precisely controlling bottom-up self-assembly on the nanometer scale with programmability and specificity of Watson-Crick base pairing, as well as facile and affordable oligosynthesis, and an abundance of nucleic acid manipulation techniques. With applications in smart therapeutics, drug delivery systems and as tools for molecular biology that operate within living cells, many of the strengths of nucleic acid nanostructures play well the intracellular landscape. One limiting factor to approaching these applications, is that most nucleic acid nanotechnologies are created *in vitro* as it is easier to control assembly conditions as well as characterise the final constructs. To counteract this movement, there needs to be more interest of cell biologists in this biophysical and biochemical heavy field. Intracellular construction or delivery of nucleic acid nanostructures opens up this technology to a plethora of fields in biophysics, diagnostics, nanoparticle and protein assembly, nucleic acid computing and storage, biomolecule structure determination, drug delivery and synthetic biology.

In this thesis, I investigate three approaches to bringing nucleic acid nanostructures closer to *in vivo* application. By focusing on assembling unimolecular nucleic acid nanostructures, my first method of knitting with nucleic acids allows for the co-transcriptional folding of nucleic acid nanostructures, which can one day be carried out completely *de novo* in the cell. My second approach towards *in vivo* DNA nanostructures, is using cell confinement as an intracellular delivery tool for *in vitro* assembled complex DNA nanostructures. Finally, I discuss the idea of reconfigurable DNA nanostructures, using bistable DNA hairpins as a model, which can one day be incorporated into larger dynamic DNA nanostructures.

In Chapter 3, I designed both single-knit and multi-knit DNA nanostructures. I used NUPACK to design and characterise the thermodynamic stability of the sequences. Additionally, I used KineFold and CoFold to predict the folding pathway of the single-stranded structures. These software were particularly useful in modelling the pseudoknot-like interactions in the multi-knit structure. For my single-knit model, I create a HTL design consisting of a hairpin, thread and lock strand. Depending on the order each component is added into a reaction, the resulting structure is either “knitted” with the thread looped into the hairpin or “non-knitted”, where the thread is not associated with the hairpin. To optimise the potential for the thread to loop into the hairpin, I used oxDNA to simulate various HTL sequence designs and found that a hairpin loop of 33 nt with an interaction domain of 15 nt to bind to the thread strand, resulted in consistent “knitting”. Additionally, I added toehold displacement strands as well as fluorophores to the structure to aid the gel electrophoresis analysis in Chapter 4. For the multi-knit model, I designed a few versions including a stacked hairpin design as well as conjoining HTL model. With the help of Dr. Jonathan Bath, we simulated simple multi-knitted structures as a Plus-Minus model. Here, we found that a

knitted structure with each row containing loops that form from the same direction had the least torsional twisting and favourable dynamics. Using this concept, I added sequence specificity to the multi-knit structure and designed a DNA template for rolling circle amplification. The simplest RCA motif of the multi-knit structure includes 4 rows with 2 hairpins each. This 8 hairpin system was optimised in oxDNA and KineFold given sequences as designed in NUPACK. I also explain the molecular assembly method to circularise the DNA RCA template from two dsDNA templates, as my sequence was too repetitive to order as a single strand. I wrap up with chapter by attempting to convert my DNA knitting design into a RNA knitting system. I was able to use the DNA parameters to create an RNA single-knit design, which correctly folded in KineFold and oxRNA. But my multi-knit DNA structure failed to translate into an RNA based sequence. Further design optimisation is needed to one day create RNA multi-knitted nanostructures.

In Chapter 4, I analysed the single-knit and multi-knit DNA designs from Chapter 3. For the single-knit HTL model, I ran the components on a polyacrylamide gel to study the knitting reaction. I first had to optimise the conditions required to achieve the best signal of the strands on the gel and reduce any well aggregation. This included quenching the lock strand in MilliQ water, as well as finding the ideal stoichiometry for each component. With the fluorescent strands, I was able to show that the hairpin-thread-lock-invader-toehold displacement strand order resulted in a knitted structure, while the thread-lock-invader-hairpin-toehold displacement strand order did not achieve knitting. As for my multi-knit model, I attempted to analyse the conjoining HTL design by PAGE. This system consisted of adding three hairpins, three threads and three locks to make three rows with three knitted structures each. Unfortunately, it was unclear whether this system could form knitted loops as the bands for the structures were too close to the well of the gel. Therefore, I moved to my single-stranded rolling circle amplification method for producing multi-knit DNA. As my RCA template was in two linear pieces, I just had to circularise the structure using golden gate assembly proteins. I then used a nicking enzyme to create an attachment point for the polymerase, Phi29. I analysed this molecular assembly and the final RCA product on a polyacrylamide gel. It is clear that the template was circularised, as I used an exonuclease to remove any non-circular structures and a band still remained. Unfortunately though, it appears that the yield of the circular templates is low. Additionally, it is unclear if the final RCA product is knitted or not as the structures are all aggregated in the gel wells. Therefore, I used atomic force microscopy to analyse the structures. Through AFM, particularly under the guidance of the Pyne lab, I was able to image the circular templates, individual hairpins, as well as resolve structures that appear to have looped hairpins in a final “knitted” structure. Future experimentation is required to determine if the loops are topologically “knitted” into each other.

In Chapter 5, I showcased cell confinement as a new intracellular delivery tool for DNA origami. Using the commercial dynamic and static 4Dcell confiners, as well as homemade 12-well static version, I was able to show that at lower pressures around 5kPa, the cells can be compressed enough to create transient pores in the membrane to allow for the diffusion of DNA origami structures into the cytosol. After compression for an hour, the cells can recover their cell membranes and continue to grow normally. Through the use of confocal fluorescence microscopy, I was able to show the intracellular targeting of DNA origami signposts to eGFP-cd63-BMP7 an intracellular membrane component, after mammalian cells were compressed with the static 4Dcell confiner.

Finally, in Chapter 6, I analysed the reconfigurable dynamics of three bistable DNA hairpins by gel electrophoresis. These reconfigurable DNA hairpins were designed in NUPACK to have two energetically favourable states as either an A stem hairpin or a B stem hairpin. The structures can be switched into either conformation through toehold-mediated strand displacement. I also used these designs to compare the free energy predictions from NUPACK 3.0 compared to the newer NUPACK 4.0. I designed a short stem hairpin which formed an intermediate structure on a polyacrylamide gel, even when switched into either state. This structure most likely does not have a stable enough hairpin stem and transiently switches into both conformations. My longer stem hairpin design was able to successfully form either conformation was switched by the displacement strands. Here, NUPACK 4.0 was better able to predict the slight preference for the B stem conformation. To reduce the hairpin loop size from the long stem hairpin design, I created an extended long stem hairpin, which utilised α domains to increase the hairpin stem length without increasing the hairpin loop size. These extended stem hairpins showed an interesting ability to form both the conformations no matter which displacement strands were used. I expect that these hairpins are able to spontaneously self-displace their hairpin stems through the α domains. These hairpins would be interesting candidates to model the dynamic reactions of in the coarse-grained modelling system, oxDNA. They would also work well in larger DNA nanostructures as tools for reconfiguration.

7.2 Outlook and Directions for Future Research

The future directions for cell confinement as an intracellular delivery mechanism as well as the reconfigurable DNA hairpins have been discussed in their respective chapters. In this section, I will expand on the capabilities of DNA knitting as a novel unimolecular nucleic acid assembly method by highlighting the directions for future research within this field. For this thesis, my work on DNA knitting is primarily carried out *in vitro* to allow for easier characterisation of the folded structures, as well as allowing for easier modifications to the preliminary designs. As the concept is intended for intracellular self-assembly, it is important to shift the focus of DNA knitting to RNA knitting and work on the co-transcriptional folding of the design. Additionally, by taking inspiration from the knitted fabrics, there is plethora of new knitting patterns that can be used to make complex nucleic acid nanostructures.

7.2.1 Larger multi-knit nanostructures

My multi-knit DNA nanostructures were limited in design size to only 4 rows of 2 hairpins wide in each motif, due to the constraints of rolling circle amplification. It would be interesting to create larger DNA multi-knit structures, whether that is wider structures with more hairpins in each row, or longer ribbons of multiple knitted hairpins. An example of a 3 by 4 knitted structure is shown in Figure 7.1A and an example of a 2 by 8 knitted structure is shown in Figure 7.1B. Both of these designs would make resolving the hairpin loop structures by AFM a lot more clear as there are more “knitted” interactions to observe.

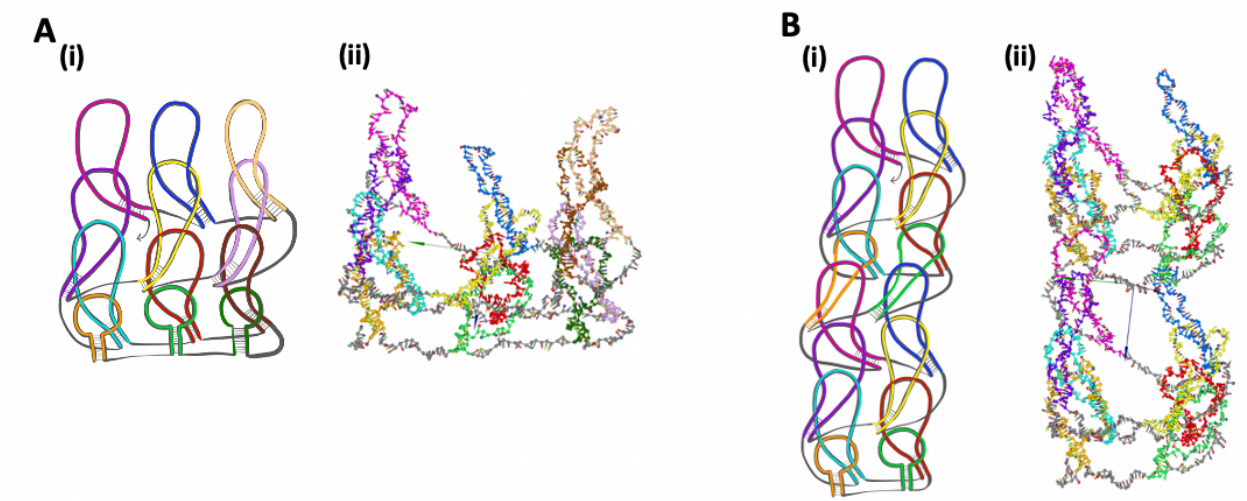


Figure 7.1: Larger DNA multi-knit designs. A) (i) Schematic of 3 by 4 design with 4 rows of 3 hairpins each and the corresponding oxDNA model (ii). B) (i) Schematic of 2 by 8 design with 8 rows of 2 hairpins each and the corresponding oxDNA model (ii). Each hairpin is highlighted in a unique colour.

The main hurdle in creating these larger multi-knit structures is designing the RCA template with these added repeats. The template will most likely need to be synthesised into 3 linear dsDNA sequences, which will require more molecular assembly steps to circularise. One issue I had with my current 2 half template design was the low yields of the circularised templates after molecular cloning. I would therefore need a more efficient method to isolate concentrations of the circular template for these larger DNA nanostructures. This can potentially be done using plasmid DNA and transforming the templates into bacterial cells.

In addition to creating a larger multi-knitted structure by rolling circle amplification, another approach to designing a bigger structure is by attaching the hairpins to an assembled DNA origami structure. Not only can the DNA origami act as a base structure, but it can potentially aid assembly by preventing overtwisting of the structure and allows for more hairpins attachment point. A large DNA origami base can also help with visualisation as DNA origami is easily recognisable by AFM and TEM. With the help of Dr. Erik Benson, I briefly modelled two rows of DNA knitted hairpins on a simple 6-helix DNA origami tube, as seen in Figure 7.2.

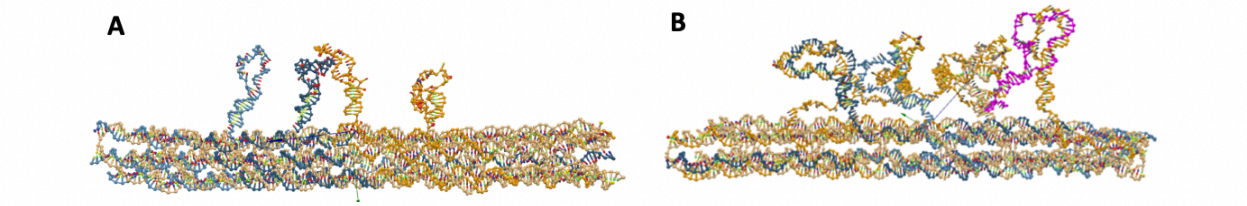


Figure 7.2: Knitting DNA on a DNA origami tube. A simple 6-helix bundle DNA origami tube is used as a base for DNA knitting. A) 4 starting hairpins are added to the DNA origami tube. B) A second row of hairpins is knitted into the first row of hairpins. One of the knitted hairpins is highlighted in pink. These models were made in oxDNA.

7.2.2 Modifications to the sequences

In addition to modifying the overall size of the multi-knit DNA nanostructures, it would be interesting to see the effects of changing smaller aspects of the design sequence. In particular, I would like to see how adding mismatches to the C domain can affect the knitting interactions. In both my single-knit and multi-knit DNA structures, the C domain is quite long at 15 or 17 nt, respectively. This was designed to improve stability of the looping, as well as to better position the thread strand to loop into the hairpin strand as simulated through oxDNA. Adding mismatches into this region can still position the thread correctly, but also may reduce any torsional strain, as shown in Figure 7.3A, that the longer C-C* interaction may cause. Due to the overtwisting of the knitted hairpins, I had to design my structures with a lot of flexibility in mind to counteract this movement. If the mismatches can help reduce this strain, I may be able to reduce the sizes of the hairpin loops as well as the linker spaces between each hairpin, which may finetune the mechanical and elastic characteristics of the multi-knit DNA nanostructures.

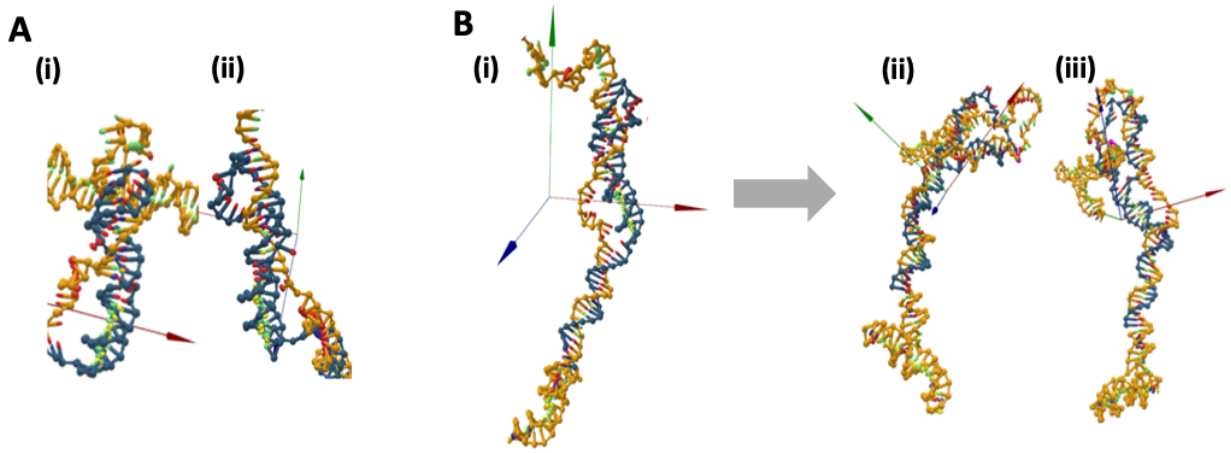


Figure 7.3: Adding in mismatches into the C domain of the single-knit DNA model to counteract torsional strain. A) Examples of overtwisted hairpins with C domain lengths of (i) 12 nt and (ii) 14 nt. Neither structure is able to loop the thread (dark blue) into the hairpin (yellow) due to the overtwisting. B) (i) An overtwisted 14 nt C domain is fixed by adding a single mismatch in the 10th base pair of the C-C interaction. (ii) and (iii) show alternating view points of the improved hairpin, which has a thread (dark blue) looped into the hairpin (yellow). These models were made in oxDNA.*

I briefly analysed the effects of adding a single mismatch in the 10th base pair of the C-C* domain in my oxDNA single-knit models as shown in Figure 7.3B. In part (i) of this figure, a hairpin structure with a C domain of 14 nt is overtwisted by the thread strand. By adding a mismatch into the same sequence, the model was able to counteract the torsional strain and the thread was able to loop into the hairpin (ii and iii). This mismatch analysis was only conducted in my single-knit hairpin-thread model. It would be interesting to see how mismatches can change the behaviour of a multi-knitted DNA structure.

Another interesting change to my multi-knit DNA structures would be the removal of the hairpin stems. I included the hairpins in the original design to aid the looping of the structure as a hairpin

is already a pre-formed loop. If instead, the structure "knits" without the hairpin motif, this may increase the kinetics of knitting as the strand would not have to form each hairpin first. I have tentatively designed an enhanced image of a 8 row multi-knit structure where rows 2-4 and 6-8 do not have the hairpin stems, as seen in Figure 7.4B. Rows 1 and 5 contain hairpin stems as row 1 acts as the base of the structure and row 5 is just a repeat of row 1. If all rows were to be without hairpins, another form of a base — potentially the DNA origami tube shown in Figure 7.3 — is required to prevent the loops from unravelling at the bottom. The side view of the structure with less hairpin stems (Figure 7.4A(ii)) shows more of a zig-zag pattern commonly seen in knitted fabrics compared to the design with hairpin stems (Figure 7.4B(ii)). Therefore, the multi-knit DNA design without hairpin stems may be more true to the nature of knitting, and may provide more flexibility of the overall structure.

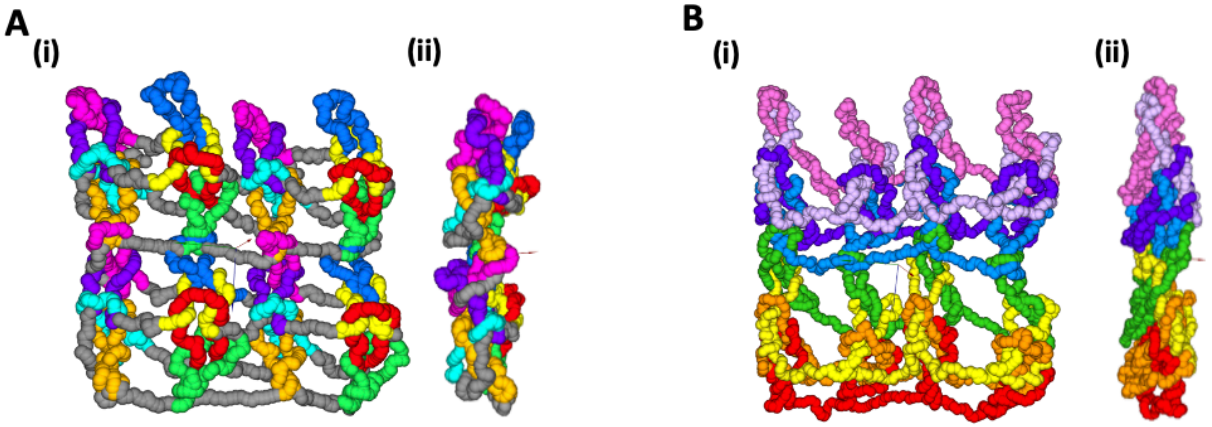


Figure 7.4: Removing the hairpin stems from the multi-knit DNA design. A) Front (i) and side (ii) views of a 4 by 8 multi-knit structure without hairpins stems in rows 2-4 and 6-8. Each row is highlighted in a different colour. B) Front (i) and side (ii) views of a 4 by 8 multi-knit structure with hairpin stems. The models are made in oxDNA with the enhanced nucleoside spheres and connecters to aid visualisation.

In addition to these modifications to the multi-knit DNA nanostructure, it will be interesting to see what this technique can transform into in the future. If the assembly process can be finely controlled, a plethora or knitting patterns can be produced. I have only discussed flat sheets of multi-knitted nucleic acid nanostructures in this thesis, but the DNA knitting method has the potential to form three-dimensional tubes as well as spheres if taking inspiration from the world of knitting.

7.2.3 RNA Knitting

I briefly discussed RNA knitting in Chapter 3, where I worked on the designs of both single-knit and multi-knit RNA nanostructures. For simplicity, I converted my DNA knitting sequences to RNA for these designs. Therefore, they were not optimised for single-stranded RNA assembly, as additional kinetic traps need to be considered. As the nucleic acid knitting technique was initially

conceptualised as a method for *de novo* assembly of nucleic acids within the cell, the next big steps for DNA knitting would be to create RNA knitted nanostructures. These RNA nanostructures could then be tested using *in vitro* transcription. The use of T7 RNA polymerase allows for efficient transcription of the nanostructures from a DNA template. The transcribed product can then be analysed by AFM in a similar manner to my RCA multi-knit DNA nanostructures. Once *in vitro* characterisation is achieved, the next goals would be *in vivo* co-transcriptional folding inside bacterial and mammalian cells.

7.2.4 Simulations of these nucleic acid knitted structures

In addition to the design and experimental analysis of the single-knit and multi-knit DNA and RNA nanostructures, these structures pose as interesting designs to study *in silico*. Dynamic models of the design can be created, similar to the Plus-Minus simulations conducted in Chapter 3. Furthermore, it would be interesting to simulate the folding pathway of the knitted structures through modeling the co-transcriptional folding. In addition to the temperature and salt concentrations used in these models, further complex parameters such as the confined space of the simulation and the polymerase speed can be manipulated to control the folding pathway. For the simulation space, a larger area would allow for more flexibility of the knitting strand, increasing the potential for looping compared to if the structure is simulated in a tight box. As for polymerase speed, faster transcription may avoid kinetic trapping, although a slow polymerase may be useful to form complex secondary structures.

References

[1] J. M. Stewart and E. Franco. *Self-assembly of large RNA structures: learning from DNA nanotechnology*. DNA and RNA Nanotechnology, 2(1)January 2016.

[2] J. D. Watson and F. H. C. Crick. *Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid*. Nature, 171(4356):737–738, April 1953.

[3] D. M. Gray and R. L. Ratliff. *Circular dichroism evidence for G·U and G·T base pairing in poly[r(G-U)] and poly[d(G-T)]*. Biopolymers, 16(6):1331–1342, June 1977.

[4] G. Varani and W. H. McClain. *The G·U wobble base pair*. EMBO Rep, 1(1):18–23, July 2000.

[5] M. H. F. Wilkins, A. R. Stokes, and H. R. Wilson. *Molecular Structure of Nucleic Acids: Molecular Structure of Deoxypentose Nucleic Acids*. Nature, 171(4356):738–740, April 1953.

[6] R. E. Franklin and R. G. Gosling. *Molecular Configuration in Sodium Thymonucleate*. Nature, 171(4356):740–741, April 1953.

[7] C. A. Hunter and J. K. M. Sanders. *The nature of .pi.-.pi. interactions*. J Am Chem Soc, 112(14):5525–5534, July 1990.

[8] I. Vakonakis and A. C. LiWang. *N1···N3 Hydrogen Bonds of A:U Base Pairs of RNA Are Stronger than Those of A:T Base Pairs of DNA*. J Am Chem Soc, 126(18):5688–5689, May 2004.

[9] M. Swart, C. Fonseca Guerra, and F. M. Bickelhaupt. *Hydrogen Bonds of RNA Are Stronger than Those of DNA, but NMR Monitors Only Presence of Methyl Substituent in Uracil/Thymine*. J Am Chem Soc, 126(51):16718–16719, December 2004.

[10] J. Plavec, W. Tong, and J. Chattopadhyaya. *How do the gauche and anomeric effects drive the pseudorotational equilibrium of the pentofuranose moiety of nucleosides?*. J Am Chem Soc, 115(21):9734–9746, October 1993.

[11] Y. M. Vianney, J. Jana, N. Schröder, and K. Weisz. *Nucleic Acid Structure and Biology*. , in *Nucleic Acid Biology and its Application in Human Diseases*, Singapore: Springer Nature Singapore, 2023, :1–34, .

[12] N. B. LEONTIS and E. WESTHOF. *Geometric nomenclature and classification of RNA base pairs*. RNA, 7(4):S1355838201002515, April 2001.

[13] E. Winfree, F. Liu, L. A. Wenzler, and N. C. Seeman. *Design and self-assembly of two-dimensional DNA crystals*. Nature, 394(6693):539–544, August 1998.

[14] P. W. K. Rothemund, N. Papadakis, and E. Winfree. *Algorithmic Self-Assembly of DNA Sierpinski Triangles*. PLoS Biol, 2(12):e424, December 2004.

[15] R. Holliday. *A mechanism for gene conversion in fungi*. Genet Res (Camb), 89(5–6):285–307, December 2007.

[16] N. C. Seeman. *Nucleic acid junctions and lattices*. J Theor Biol, 99(2):237–247, November 1982.

[17] J. Chen and N. C. Seeman. *Synthesis from DNA of a molecule with the connectivity of a cube*. Nature, 350(6319):631–633, April 1991.

[18] R. P. Goodman, R. M. Berry, and A. J. Turberfield. *The single-step synthesis of a DNA tetrahedronElectronic supplementary information (ESI) available: stoichiometry control. See http://www.rsc.org/suppdata/cc/b4/b402293a/*. Chemical Communications, (12):1372, 2004.

[19] Y. Zhang and N. C. Seeman. *Construction of a DNA-Truncated Octahedron*. J Am Chem Soc, 116(5):1661–1669, March 1994.

[20] N. C. Seeman and H. F. Sleiman. *DNA nanotechnology*. Nat Rev Mater, 3(1):17068, November 2017.

[21] P. W. K. Rothemund. *Folding DNA to create nanoscale shapes and patterns*. Nature, 440(7082):297–302, March 2006.

[22] T. J. Fu and N. C. Seeman. *DNA double-crossover molecules*. Biochemistry, 32(13):3211–3220, April 1993.

[23] C. Mao, W. Sun, and N. C. Seeman. *Designed Two-Dimensional DNA Holliday Junction Arrays Visualized by Atomic Force Microscopy*. J Am Chem Soc, 121(23):5437–5443, June 1999.

[24] D. Reishus, B. Shaw, Y. Brun, N. Chelyapov, and L. Adleman. *Self-Assembly of DNA Double-Double Crossover Complexes into High-Density, Doubly Connected, Planar Structures*. J Am Chem Soc, 127(50):17590–17591, December 2005.

[25] Y. He *et al. Hierarchical self-assembly of DNA into symmetric supramolecular polyhedra*. Nature, 452(7184):198–201, March 2008.

[26] S. M. Douglas, A. H. Marblestone, S. Teerapittayanon, A. Vazquez, G. M. Church, and W. M. Shih. *Rapid prototyping of 3D DNA-origami shapes with caDNAno*. Nucleic Acids Res, 37(15):5001–5006, August 2009.

[27] Y. Ke, L. L. Ong, W. M. Shih, and P. Yin. *Three-Dimensional Structures Self-Assembled from DNA Bricks*. Science (1979), 338(6111):1177–1183, November 2012.

[28] B. Wei, M. Dai, and P. Yin. *Complex shapes self-assembled from single-stranded DNA tiles*. Nature, 485(7400):623–626, May 2012.

[29] R. Carlson. *The changing economics of DNA synthesis*. Nat Biotechnol, 27(12):1091–1094, December 2009.

[30] W. M. Shih, J. D. Quispe, and G. F. Joyce. *A 1.7-kilobase single-stranded DNA that folds into a nanoscale octahedron*. Nature, 427(6975):618–621, February 2004.

[31] C. Lin, M. Xie, J. J. L. Chen, Y. Liu, and H. Yan. *Rolling-Circle Amplification of a DNA Nanojunction*. Angewandte Chemie International Edition, 45(45):7537–7539, November 2006.

[32] C. Lin, X. Wang, Y. Liu, N. C. Seeman, and H. Yan. *Rolling Circle Enzymatic Replication of a Complex Multi-Crossover DNA Nanostructure*. J Am Chem Soc, 129(46):14475–14481, November 2007.

[33] Z. Li *et al. A Replicable Tetrahedral Nanostructure Self-Assembled from a Single DNA Strand*. J Am Chem Soc, 131(36):13093–13098, September 2009.

[34] M. M. Ali *et al. Rolling circle amplification: a versatile tool for chemical biology, materials science and medicine*. Chem Soc Rev, 43(10):3324, 2014.

[35] C. Geary, P. W. K. Rothemund, and E. S. Andersen. *A single-stranded architecture for cotranscriptional folding of RNA nanostructures*. Science (1979), 345(6198):799–804, August 2014.

[36] E. K. S. McRae *et al. Structure, folding and flexibility of co-transcriptional RNA origami*. Nat Nanotechnol, 18(7):808–817, July 2023.

[37] C. Geary, G. Grossi, E. K. S. McRae, P. W. K. Rothemund, and E. S. Andersen. *RNA origami design tools enable cotranscriptional folding of kilobase-sized nanoscaffolds*. Nat Chem, 13(6):549–558, June 2021.

[38] D. Han *et al. Single-stranded DNA and RNA origami*. Science (1979), 358(6369)December 2017.

[39] H. Lee, E. Popodi, H. Tang, and P. L. Foster. *Rate and molecular spectrum of spontaneous mutations in the bacterium* Escherichia coli *as determined by whole-genome sequencing*. Proceedings of the National Academy of Sciences, 109(41)October 2012.

[40] E. Benson *et al. DNA rendering of polyhedral meshes at the nanoscale*. Nature, 523(7561):441–444, July 2015.

[41] B. Yurke, A. J. Turberfield, A. P. Mills, F. C. Simmel, and J. L. Neumann. *A DNA-fuelled molecular machine made of DNA*. Nature, 406(6796):605–608, August 2000.

[42] T. Mayer, L. Oesinghaus, and F. C. Simmel. *Toehold-Mediated Strand Displacement in Random Sequence Pools*. J Am Chem Soc, 145(1):634–644, January 2023.

[43] J. H. Reif. *The design of autonomous DNA nano-mechanical devices: Walking and rolling DNA*. Nat Comput, 2(4):439–461, 2003.

[44] P. Yin, H. Yan, X. G. Daniell, A. J. Turberfield, and J. H. Reif. *A Unidirectional DNA Walker That Moves Autonomously along a Track*. Angewandte Chemie International Edition, 43(37):4906–4911, September 2004.

[45] F. A. Aldaye and H. F. Sleiman. *Modular Access to Structurally Switchable 3D Discrete DNA Assemblies*. J Am Chem Soc, 129(44):13376–13377, November 2007.

[46] L. E. Baltierra-Jasso, M. J. Morten, and S. W. Magennis. *Sub-Ensemble Monitoring of DNA Strand Displacement Using Multiparameter Single-Molecule FRET*. ChemPhysChem, 19(5):551–555, March 2018.

[47] A. Singh, G. Patel, and S. S. Patel. *Twinkle-Catalyzed Toehold-Mediated DNA Strand Displacement Reaction*. J Am Chem Soc, November 2023.

[48] E. Torelli, B. Shirt-Ediss, S. A. Navarro, M. Manzano, P. Vizzini, and N. Krasnogor. *Light-Up Split Broccoli Aptamer as a Versatile Tool for RNA Assembly Monitoring in*

Cell-Free TX-TL Systems, Hybrid RNA/DNA Origami Tagging and DNA Biosensing. Int J Mol Sci, 24(10):8483, May 2023.

[49] E. S. Andersen *et al. Self-assembly of a nanoscale DNA box with a controllable lid*. Nature, 459(7243):73–76, May 2009.

[50] A. E. Marras *et al. Cation-Activated Avidity for Rapid Reconfiguration of DNA Nanodevices*. ACS Nano, 12(9):9484–9494, September 2018.

[51] E. M. Willner *et al. Single-Molecule Observation of the Photoregulated Conformational Dynamics of DNA Origami Nanoscissors*. Angewandte Chemie International Edition, 56(48):15324–15328, November 2017.

[52] E. Benson, R. C. Marzo, J. Bath, and A. J. Turberfield. *A DNA molecular printer capable of programmable positioning and patterning in two dimensions*. Sci Robot, 7(65)April 2022.

[53] A.-K. Pumm *et al. A DNA origami rotary ratchet motor*. Nature, 607(7919):492–498, July 2022.

[54] L. Jaeger and A. Chworos. *The architectonics of programmable RNA and DNA nanostructures*. Curr Opin Struct Biol, 16(4):531–543, August 2006.

[55] A. Chworos *et al. Building Programmable Jigsaw Puzzles with RNA*. Science (1979), 306(5704):2068–2072, December 2004.

[56] C. J. Delebecque, A. B. Lindner, P. A. Silver, and F. A. Aldaye. *Organization of Intracellular Reactions with Rationally Designed RNA Assemblies*. Science (1979), 333(6041):470–474, July 2011.

[57] G. L. Lukacs, P. Haggie, O. Seksek, D. Lechardeur, N. Freedman, and A. S. Verkman. *Size-dependent DNA Mobility in Cytoplasm and Nucleus*. Journal of Biological Chemistry, 275(3):1625–1629, January 2000.

[58] I. Schoen, H. Krammer, and D. Braun. *Hybridization kinetics is different inside cells*. Proceedings of the National Academy of Sciences, 106(51):21649–21654, December 2009.

[59] J.-M. Williford, J. L. Santos, R. Shyam, and H.-Q. Mao. *Shape control in engineering of polymeric nanoparticles for therapeutic delivery*. Biomater Sci, 3(7):894–907, 2015.

[60] C. H. Li, W. Y. Lv, Y. Yan, F. F. Yang, S. J. Zhen, and C. Z. Huang. *Nucleolin-Targeted DNA Nanotube for Precise Cancer Therapy through Förster Resonance Energy Transfer-Indicated Telomerase Responsiveness*. Anal Chem, 93(7):3526–3534, February 2021.

[61] Q. Jiang *et al. DNA Origami as a Carrier for Circumvention of Drug Resistance*. J Am Chem Soc, 134(32):13396–13403, August 2012.

[62] C. Wiraja *et al. Framework nucleic acids as programmable carrier for transdermal drug delivery*. Nat Commun, 10(1):1147, March 2019.

[63] S. Shi *et al. Targeted and effective glioblastoma therapy via aptamer-modified tetrahedral framework nucleic acid-paclitaxel nanoconjugates that can pass the blood brain barrier*. Nanomedicine, 21:102061, October 2019.

[64] D. Y. Tam *et al. Penetrating the Blood–Brain Barrier by Self-Assembled 3D DNA Nanocages as Drug Delivery Vehicles for Brain Cancer Therapy*. ACS Appl Mater Interfaces, :acsami.0c02957, June 2020.
[65] X. Yang *et al. Effect of tetrahedral DNA nanostructures on LPS-induced neuroinflammation in mice*. Chinese Chemical Letters, 33(4):1901–1906, April 2022.
[66] H. Lee *et al. Molecularly self-assembled nucleic acid nanoparticles for targeted in vivo siRNA delivery*. Nat Nanotechnol, 7(6):389–393, June 2012.
[67] Q. Zhang *et al. DNA Origami as an* In Vivo *Drug Delivery Vehicle for Cancer Therapy*. ACS Nano, 8(7):6633–6643, July 2014.
[68] G. Zhu *et al. Self-assembled, aptamer-tethered DNA nanotrains for targeted transport of molecular drugs in cancer theranostics*. Proceedings of the National Academy of Sciences, 110(20):7998–8003, May 2013.
[69] F. A. Aldaye, W. T. Senapedis, P. A. Silver, and J. C. Way. *A Structurally Tunable DNA-Based Extracellular Matrix*. J Am Chem Soc, 132(42):14727–14729, October 2010.
[70] N. Stephanopoulos *et al. Bioactive DNA-Peptide Nanotubes Enhance the Differentiation of Neural Stem Cells Into Neurons*. Nano Lett, 15(1):603–609, January 2015.
[71] R. Jungmann, M. S. Avendaño, J. B. Woehrstein, M. Dai, W. M. Shih, and P. Yin. *Multiplexed 3D cellular super-resolution imaging with DNA-PAINT and Exchange-PAINT*. Nat Methods, 11(3):313–318, March 2014.
[72] E. Silvester *et al. DNA origami signposts for identifying proteins on cell membranes by electron cryotomography*. Cell, 184(4):1110-1121.e16, February 2021.
[73] J. Schnitzbauer, M. T. Strauss, T. Schlichthaerle, F. Schueder, and R. Jungmann. *Super-resolution microscopy with DNA-PAINT*. Nat Protoc, 12(6):1198–1228, June 2017.
[74] N. Oleksiievets *et al. Fluorescence lifetime DNA-PAINT for multiplexed super-resolution imaging of cells*. Commun Biol, 5(1):38, January 2022.
[75] T. L. Fisher, T. Terhorst, X. Cao, and R. W. Wagner. *Intracellular disposition and metabolism of fluorescently-labled unmodified and modified oligouncleotides microijjected into mammalian cells*. Nucleic Acids Res, 21(16):3857–3865, 1993.
[76] M. Amarzguioui. *Tolerance for mutations and chemical modifications in a siRNA*. Nucleic Acids Res, 31(2):589–595, January 2003.
[77] J. Watts, G. Deleavey, and M. Damha. *Chemically modified siRNA: tools and applications*. Drug Discov Today, 13(19–20):842–855, October 2008.
[78] Q. Mei *et al. Stability of DNA Origami Nanoarrays in Cell Lysate*. Nano Lett, 11(4):1477–1482, April 2011.
[79] C. E. Castro *et al. A primer to scaffolded DNA origami*. Nat Methods, 8(3):221–229, March 2011.
[80] N. Joseph *et al. Biodistribution and function of coupled polymer-DNA origami nanostructures*. Sci Rep, 13(1):19567, November 2023.
[81] N. Ponnuswamy *et al. Oligolysine-based coating protects DNA nanostructures from low-salt denaturation and nuclease degradation*. Nat Commun, 8(1):15654, May 2017.

[82] J. Rejman, V. Oberle, I. S. Zuhorn, and D. Hoekstra. *Size-dependent internalization of particles via the pathways of clathrin- and caveolae-mediated endocytosis*. Biochemical Journal, 377(1):159–169, January 2004.

[83] J. M. Meacham, K. Durvasula, F. L. Degertekin, and A. G. Fedorov. *Physical Methods for Intracellular Delivery: Practical Aspects from Laboratory Use to Industrial-Scale Processing*. SLAS Technol, 19(1):1–18, February 2014.

[84] M. M. C. Bastings *et al. Modulation of the Cellular Uptake of DNA Origami through Control over Mass and Shape*. Nano Lett, 18(6):3557–3564, June 2018.

[85] R. Agarwal, V. Singh, P. Jurney, L. Shi, S. V. Sreenivasan, and K. Roy. *Mammalian cells preferentially internalize hydrogel nanodiscs over nanorods and use shape-specific uptake mechanisms*. Proceedings of the National Academy of Sciences, 110(43):17247–17252, October 2013.

[86] J. J. Fakhoury, C. K. McLaughlin, T. W. Edwardson, J. W. Conway, and H. F. Sleiman. *Development and Characterization of Gene Silencing DNA Cages*. Biomacromolecules, 15(1):276–282, January 2014.

[87] E. Neumann, M. Schaefer-Ridder, Y. Wang, and P. H. Hofschneider. *Gene transfer into mouse lyoma cells by electroporation in high electric fields.* EMBO J, 1(7):841–845, July 1982.

[88] M. A. Badding, J. D. Lapek, A. E. Friedman, and D. A. Dean. *Proteomic and Functional Analyses of Protein–DNA Complexes During Gene Transfer*. Molecular Therapy, 21(4):775–785, April 2013.

[89] A. Chopra, S. Krishnan, and F. C. Simmel. *Electrotransfection of Polyamine Folded DNA Origami Structures*. Nano Lett, 16(10):6683–6690, October 2016.

[90] J. A. Kretzmann, A. Liedl, A. Monferrer, V. Mykhailiuk, S. Beerkens, and H. Dietz. *Gene-encoding DNA origami for mammalian cell expression*. Nat Commun, 14(1):1017, February 2023.

[91] J. Mikkilä *et al. Virus-Encapsulated DNA Origami Nanostructures for Cellular Delivery*. Nano Lett, 14(4):2196–2200, April 2014.

[92] M. R. Capecchi. *High efficiency transformation by direct microinjection of DNA into cultured mammalian cells*. Cell, 22(2):479–488, November 1980.

[93] M. Yamaizumi, A. L. Horwich, and F. H. Ruddle. *Expression and Stabilization of Microinjected Plasmids Containing the Herpes Simplex Virus Thymidine Kinase Gene and Polyoma Virus DNA in Mouse Cells*. Mol Cell Biol, 3(4):511–522, April 1983.

[94] K. R. Folger, E. A. Wong, G. Wahl, and M. R. Capecchi. *Patterns of Integration of DNA Microinjected into Cultured Mammalian Cells: Evidence for Homologous Recombination Between Injected Plasmid DNA Molecules*. Mol Cell Biol, 2(11):1372–1387, November 1982.

[95] C. Chiappini *et al. Biodegradable silicon nanoneedles delivering nucleic acids intracellularly induce localized in vivo neovascularization*. Nat Mater, 14(5):532–539, May 2015.

[96] D. Hachim *et al. Polysaccharide-Polyplex Nanofilm Coatings Enhance Nanoneedle-Based Gene Delivery and Transfection Efficiency*. Small, 18(36)September 2022.

[97] A. Angelin, O. Kassel, S. Rastegar, U. Strähle, and C. M. Niemeyer. *Protein-Functionalized DNA Nanostructures as Tools to Control Transcription in Zebrafish Embryos*. ChemistryOpen, 6(1):33–39, February 2017.

[98] P. L. Felgner *et al. Lipofection: a highly efficient, lipid-mediated DNA-transfection procedure.* Proceedings of the National Academy of Sciences, 84(21):7413–7417, November 1987.

[99] X.-L. Guo, D.-D. Yuan, T. Song, and X.-M. Li. *DNA nanopore functionalized with aptamer and cell-penetrating peptide for tumor cell recognition*. Anal Bioanal Chem, 409(15):3789–3797, June 2017.

[100] Z. Xia *et al. Tumor-Penetrating Peptide-Modified DNA Tetrahedron for Targeting Drug Delivery*. Biochemistry, 55(9):1326–1331, March 2016.

[101] J. Gilleron *et al. Image-based analysis of lipid nanoparticle–mediated siRNA delivery, intracellular trafficking and endosomal escape*. Nat Biotechnol, 31(7):638–646, July 2013.

[102] G. Sahay *et al. Efficiency of siRNA delivery by lipid nanoparticles is limited by endocytic recycling*. Nat Biotechnol, 31(7):653–658, July 2013.

[103] L. Liang *et al. Single-Particle Tracking and Modulation of Cell Entry Pathways of a Tetrahedral DNA Nanostructure in Live Cells*. Angewandte Chemie International Edition, 53(30):7745–7750, July 2014.

[104] C. Lin, S. Rinker, X. Wang, Y. Liu, N. C. Seeman, and H. Yan. *In vivo cloning of artificial DNA nanostructures*. Proceedings of the National Academy of Sciences, 105(46):17626–17631, November 2008.

[105] M. Li *et al. In vivo production of RNA nanostructures via programmed folding of single-stranded RNAs*. Nat Commun, 9(1):2196, June 2018.

[106] J. L. Litke and S. R. Jaffrey. *Highly efficient expression of circular RNA aptamers in cells using autocatalytic transcripts*. Nat Biotechnol, 37(6):667–675, June 2019.

[107] J. N. Zadeh *et al. NUPACK: Analysis and design of nucleic acid systems*. J Comput Chem, 32(1):170–173, January 2011.

[108] M. E. Fornace, N. J. Porubsky, and N. A. Pierce. *A Unified Dynamic Programming Framework for the Analysis of Interacting Nucleic Acid Strands: Enhanced Models, Scalability, and Speed*. ACS Synth Biol, 9(10):2665–2678, October 2020.

[109] M. E. Fornace, J. Huang, Newman Cody T., N. J. Porubsky, M. B. Pierce, and N. A. Pierce. *NUPACK: Analysis and Design of Nucleic Acid Structures, Devices, and Systems*. ChemRxiv, (This content is a preprint and has not been peer-reviewed)2022.

[110] J. SantaLucia. *A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics*. Proceedings of the National Academy of Sciences, 95(4):1460–1465, February 1998.

[111] J. SantaLucia and D. Hicks. *The Thermodynamics of DNA Structural Motifs*. Annu Rev Biophys Biomol Struct, 33(1):415–440, June 2004.

[112] A. Xayaphoummine, T. Bucher, and H. Isambert. *Kinefold web server for RNA/DNA folding path and structure prediction including pseudoknots and knots*. Nucleic Acids Res, 33(suppl_2):W605–W610, July 2005.

[113] A. Xayaphoummine, T. Bucher, F. Thalmann, and H. Isambert. *Prediction and statistics of pseudoknots in RNA structures using exactly clustered stochastic simulations*. Proceedings of the National Academy of Sciences, 100(26):15310–15315, December 2003.

[114] S. S. Sastry and B. M. Ross. *Nuclease Activity of T7 RNA Polymerase and the Heterogeneity of Transcription Elongation Complexes*. Journal of Biological Chemistry, 272(13):8644–8652, March 1997.

[115] J. R. Proctor and I. M. Meyer. *C o F old : an RNA secondary structure prediction method that takes co-transcriptional folding into account*. Nucleic Acids Res, 41(9):e102–e102, May 2013.

[116] D. H. Mathews, J. Sabina, M. Zuker, and D. H. Turner. *Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure*. J Mol Biol, 288(5):911–940, May 1999.

[117] Z. J. Lu, D. H. Turner, and D. H. Mathews. *A set of nearest neighbor parameters for predicting the enthalpy change of RNA secondary structure formation*. Nucleic Acids Res, 34(17):4912–4924, October 2006.

[118] E. Poppleton, J. Bohlin, M. Matthies, S. Sharma, F. Zhang, and P. Šulc. *Design, optimization and analysis of large DNA and RNA nanostructures through interactive visualization, editing and molecular simulation*. Nucleic Acids Res, 48(12):e72–e72, July 2020.

[119] J. Bohlin *et al. Design and simulation of DNA, RNA and hybrid protein–nucleic acid nanostructures with oxView*. Nat Protoc, 17(8):1762–1788, August 2022.

[120] E. Poppleton, R. Romero, A. Mallya, L. Rovigatti, and P. Šulc. *OxDNA.org: a public webserver for coarse-grained simulations of DNA and RNA nanostructures*. Nucleic Acids Res, 49(W1):W491–W498, July 2021.

[121] C. A. Schneider, W. S. Rasband, and K. W. Eliceiri. *NIH Image to ImageJ: 25 years of image analysis*. Nat Methods, 9(7):671–675, July 2012.

[122] K. Bose, C. J. Lech, B. Heddi, and A. T. Phan. *High-resolution AFM structure of DNA G-wires in aqueous solution*. Nat Commun, 9(1):1959, May 2018.

[123] P. J. Haynes, K. H. S. Main, B. Akpinar, and A. L. B. Pyne. *Atomic Force Microscopy of DNA and DNA-Protein Interactions*. , 2022, :43–62, .

[124] D. Nečas and P. Klapetek. *Gwyddion: an open-source software for SPM data analysis*. Open Physics, 10(1)January 2012.

[125] J. G. Beton *et al. TopoStats – A program for automated tracing of biomolecules from AFM images*. Methods, 193:68–79, September 2021.

[126] A. Verma, M. Verma, and A. Singh. *Animal tissue culture principles and applications*. , in *Animal Biotechnology*, Elsevier, 2020, :269–293, .
[127] F. Baletto, J. P. K. Doye, and R. Ferrando. *Evidence of Kinetic Trapping in Clusters of C60 Molecules*. Phys Rev Lett, 88(7):075503, January 2002.
[128] C. P. Royall and S. R. Williams. *The role of local structure in dynamical arrest*. Phys Rep, 560:1–75, February 2015.
[129] K. E. Watters, E. J. Strobel, A. M. Yu, J. T. Lis, and J. B. Lucks. *Cotranscriptional folding of a riboswitch at nucleotide resolution*. Nat Struct Mol Biol, 23(12):1124–1131, December 2016.
[130] P. P. Dennis and H. Bremer. *Differential rate of ribosomal protein synthesis in Escherichia coli B/r*. J Mol Biol, 84(3):407–422, April 1974.
[131] M. G. Izban and D. S. Luse. *Factor-stimulated RNA polymerase II transcribes at physiological elongation rates on naked DNA but very poorly on chromatin templates.* J Biol Chem, 267(19):13647–55, July 1992.
[132] T. C. Jarvis, J. W. Newport, and P. H. von Hippel. *Stimulation of the processivity of the DNA polymerase of bacteriophage T4 by the polymerase accessory proteins. The role of ATP hydrolysis.* J Biol Chem, 266(3):1830–40, January 1991.
[133] I. du monde arabe à P. Musée d'art et d'histoire à Genève. *Tissus d'Égypte: témoins du monde arabe, VIIIe. - XVe. siècles. Collection Bouvier, Exposition 1993-1994*. 1993.
[134] Julianna Photopoulos. *Spinning a sustainable fashion revolution: meet the physicists turning wood into clothes*. , *physicsworld, IOP Publishing*, 2023.
[135] S. G. Markande and E. A. Matsumoto. *Knotty knits are tangles on tori*. February 2020.
[136] K. Murasugi. *Knot Theory & Its Applications*. Boston, MA: Birkhäuser Boston, 2008.
[137] K. Murasugi. *Knot Theory in Molecular Biology*. , in *Knot Theory & Its Applications*, Boston, MA: Birkhäuser Boston, 2008, :267–283, .
[138] A. Legendre, E. Angel, and F. Tahi. *Bi-objective integer programming for RNA secondary structure prediction with pseudoknots*. BMC Bioinformatics, 19(1):13, December 2018.
[139] M. Taufer *et al. PseudoBase++: an extension of PseudoBase for easy searching, formatting and visualization of pseudoknots*. Nucleic Acids Res, 37(Database):D127–D135, January 2009.
[140] S. Sheikh, R. Backofen, and Y. Ponty. *Impact of the Energy Model on the Complexity of RNA Folding with Pseudoknots*. , 2012, :321–333, .
[141] R. B. Lyngsø and C. N. S. Pedersen. *RNA Pseudoknot Prediction in Energy-Based Models*. Journal of Computational Biology, 7(3–4):409–427, August 2000.
[142] Y. Zhang, S. Xiao, and H. Liang. *Avoiding Kinetic Trapping in the Self-Assembly of DNA-Functionalized Gold Nanoparticles by Using an Enthalpy-Mediated Strategy*. Macromolecules, 56(9):3454–3463, May 2023.
[143] J. E. Bird, J. Marles-Wright, and A. Giachino. *A User's Guide to Golden Gate Cloning Methods and Standards*. ACS Synth Biol, 11(11):3551–3563, November 2022.

[144] V. S. Anand and S. S. Patel. *Transient State Kinetics of Transcription Elongation by T7 RNA Polymerase*. Journal of Biological Chemistry, 281(47):35677–35685, November 2006.

[145] J. Zenk, C. Tuntivate, and R. Schulman. *Kinetics and Thermodynamics of Watson–Crick Base Pairing Driven DNA Origami Dimerization*. J Am Chem Soc, 138(10):3346–3354, March 2016.

[146] D. Lubrich, J. Bath, and A. J. Turberfield. *Design and assembly of double-crossover linear arrays of micrometre length using rolling circle replication*. Nanotechnology, 16(9):1574–1577, September 2005.

[147] M. Pilo-Pais, S. Goldberg, E. Samano, T. H. LaBean, and G. Finkelstein. *Connecting the Nanodots: Programmable Nanofabrication of Fused Metal Shapes on DNA Templates*. Nano Lett, 11(8):3489–3492, August 2011.

[148] A. Rajendran, M. Endo, and H. Sugiyama. *Single-Molecule Analysis Using DNA Origami*. Angewandte Chemie International Edition, 51(4):874–890, January 2012.

[149] M. Endo and H. Sugiyama. *Single-Molecule Imaging of Dynamic Motions of Biomolecules in DNA Origami Nanostructures Using High-Speed Atomic Force Microscopy*. Acc Chem Res, 47(6):1645–1653, June 2014.

[150] H. Kominami, K. Kobayashi, and H. Yamada. *Molecular-scale visualization and surface charge density measurement of Z-DNA in aqueous solution*. Sci Rep, 9(1):6851, May 2019.

[151] C. Adam, C. Kayal, A. Ercole, S. Contera, H. Ye, and A. Jerusalem. *Action of the general anaesthetic isoflurane reveals coupling between viscoelasticity and electrophysiological activity in individual neurons*. Commun Phys, 6(1):174, July 2023.

[152] T. R. Wilks *et al. "Giant Surfactants" Created by the Fast and Efficient Functionalization of a DNA Tetrahedron with a Temperature-Responsive Polymer*. ACS Nano, 7(10):8561–8572, October 2013.

[153] H. Kim, M. Kim, S.-K. Im, and S. Fang. *Mouse Cre-LoxP system: general principles to determine tissue-specific roles of target genes*. Lab Anim Res, 34(4):147, 2018.

[154] K. K. S. A. A. A. B. J. K. A. T. R. R. K. N. C. Nikolov PM. *Cytosolic delivery of large supramolecular protein complexes arranged on DNA nanopegboards*. bioRXiv, 2017.

[155] M. R. Carvalho, F. R. Maia, J. Silva-Correia, B. M. Costa, R. L. Reis, and J. M. Oliveira. *A semiautomated microfluidic platform for real-time investigation of nanoparticles' cellular uptake and cancer cells' tracking*. Nanomedicine, 12(6):581–596, March 2017.

[156] T. DiTommaso *et al. Cell engineering with microfluidic squeezing preserves functionality of primary immune cells in vivo*. Proceedings of the National Academy of Sciences, 115(46)November 2018.

[157] H. Li *et al. Intranuclear Delivery of DNA Nanostructures via Cellular Mechanotransduction*. Nano Lett, 22(8):3400–3409, April 2022.

[158] M. Le Berre, E. Zlotek-Zlotkiewicz, D. Bonazzi, F. Lautenschlaeger, and M. Piel. *Methods for Two-Dimensional Cell Confinement*. , 2014, :213–229, .

[159] C. R. Terenna *et al. Physical Mechanisms Redirecting Cell Polarity and Cell Shape in Fission Yeast*. Current Biology, 18(22):1748–1753, November 2008.
[160] G. Faure-André *et al. Regulation of Dendritic Cell Migration by CD74, the MHC Class II-Associated Invariant Chain*. Science (1979), 322(5908):1705–1710, December 2008.
[161] J. Guck *et al. Optical Deformability as an Inherent Cell Marker for Testing Malignant Transformation and Metastatic Competence*. Biophys J, 88(5):3689–3698, May 2005.
[162] S. M. McFaul, B. K. Lin, and H. Ma. *Cell separation based on size and deformability using microfluidic funnel ratchets*. Lab Chip, 12(13):2369, 2012.
[163] S. C. Hur, N. K. Henderson-MacLennan, E. R. B. McCabe, and D. Di Carlo. *Deformability-based cell classification and enrichment using inertial microfluidics*. Lab Chip, 11(5):912, 2011.
[164] J.-M. Escola, M. J. Kleijmeer, W. Stoorvogel, J. M. Griffith, O. Yoshie, and H. J. Geuze. *Selective Enrichment of Tetraspan Proteins on the Internal Vesicles of Multivesicular Endosomes and on Exosomes Secreted by Human B-lymphocytes*. Journal of Biological Chemistry, 273(32):20121–20127, August 1998.
[165] R. B. Koumangoye, A. M. Sakwe, J. S. Goodwin, T. Patel, and J. Ochieng. *Detachment of Breast Tumor Cells Induces Rapid Secretion of Exosomes Which Subsequently Mediate Cellular Adhesion and Spreading*. PLoS One, 6(9):e24234, September 2011.
[166] R. M. Dirks, J. S. Bois, J. M. Schaeffer, E. Winfree, and N. A. Pierce. *Thermodynamic Analysis of Interacting Nucleic Acid Strands*. SIAM Review, 49(1):65–88, January 2007.

Appendix

A Sequences

A.1 Chapter 3 and 4

Single-knit HTL design

Hairpin

\5Cy5\CGCCGCGGGGTTTTTTTTTTTTTTTTATGTATGTATGTGTGTTCCCCGCGGCGTT
CTCTCTCTATTCTGCTCTTTTTCCTACCAC

Thread

GAATCGCGCGTGCGGCTTTTGCCGCACGGCGGCTGCTTTTGCAGCCGCCGCGATTCA
AAGAGCAGAATAGAGAGAGACCAAACCCCCACACATACATACATCCCCACCACCCC
ACCCACCCCCACCCCCCCCTTTTTTTTTTTTCGAGCCCGGCGCGCCGTTTTCGGCGCGC
GCGTGCGCTTTTGCGCACGCCGGGCTCG

Lock

/56FAM/CGAGCCCGGCGCGCCGTTTTCGGCGCGCGCGTGCGCTTTTGCGCACGCCGG
GCTCGTTGGGGGGGTGGGGGTGGGTGG

Toehold displacement strand

GTGGTAGGAAAAAGAGCAGAATAGAGAGAGT

Invader stand 1

GGTGGTGGGGATGTATGTATGTGTG

Invader stand 2

ATGTATGTATGTGTGGGGGTTTGGT

Invader stand 3

ATGTATGTATGTGTGGGGGTTTGGTCTCTCTCTATTCTGCTCTTT

Invader stand 4

CACACATACATACATAAAAAAAAAA

Multi-knit conjoining DNA design

HairpinA

\5Cy5\TTTTTTTTTTGCGCCTCCTGTTTTTTTTTTTTTTTTTTTTTTTTTTGGTGGGTGTG
GTGTGTGTTCAGGAGGCGCTTTTTTTTTT

HairpinB

TTTTTTTTTTGCGCTGGAACTTTTTTTTTTTTTTTTTTTTTTTTTTGTGTGTGTGTGTGT
GTTTGTTCCAGCGCTTTTTTTTTT

HairpinC

TTTTTTTTTTCGACAAATGCTTTTTTTTTTTTTTTTTTTTTTTTTTGTGTGTGTGGTGGT
GGTTTGCATTTGTCGTTTTTTTTTT

ThreadA

TTTTTTTTTTCGCACCGCGGTTCACACACCACACCCACCACCACCACCAAGAGAGAT
ATAAGCAGATTTTTTTTTTTTTTTCCGCGGTGCGTTTTTTTTTT

ThreadB

TTTTTTTTTTGTTCACGACGTTACACACACACACACACACACCACACCCACCCACCC
ACCACACCATTTTTTTTTTTTTTTTCGTCGTGAACTTTTTTTTTT

ThreadC

TTTTTTTTTTCTCAGACAAGTTACCACCACCACACACACCACCCACCATAAGCAGCG
ATCGAATATTTTTTTTTTTTTTTTTCTTGTCTGAGTTTTTTTTTT

LockA

/56FAM/TCTCCTCCTCCCTCTCTTTTTTCTTTTTTTTTTTTTTCTGCTTATATCTCTCTTTT
TTTTTTTCACCAAACAACCCAACCACCCAACT

LockB

TTGGGTTGTTTGGTGTTTTTTTTTTTGGTGTGGTGGGTGGGTTTTTTTTTTTGGGAGGT
GTGGTGAGGAGGAGGAGG

LockC

CTCACCACACCTCCCTTTTTTTTTTATATTCGATCGCTGCTTTTTTTTTTTTTTTCTCCC
TCCCTACCCACAACACC

Toehold displacement strand 1 (between HA and HB)

CAACCACAACACCCACCACACCCAC

Toehold displacement strand 2 (between HB and HC)

GGTTGGTGGTTGGGATGGGTGGGTG

Toehold displacement strand 3 (between HC and TC)

AAACAAACAACCCACCCACACACCC

Toehold displacement strand 4 (between TA and LA)

TCCTCCTCTCTCTCCTCCTCCCTCT

Toehold displacement strand 5 (between LA and LB)

AGTTGGGTGGTTGGGTTGTTTGGTG

Toehold displacement strand 6 (between LB and LC)

CCTCCTCCTCCTCACCACACCTCCC

Invader1 (between TC and HC)

ATGGTGGGTGGTGTGTGTGGTGGTGGT

Invader2 (between TB and HB)

GGGTGTGGTGTGTGTGTGTGTGTGTGT

Invader3 (between TA and HA)

TGGTGGTGGTGGTGGGTGTGGTGTGTG

Multi-knit DNA design for RCA

Underlined regions indicate the adaptor regions as generated by Twist Biosciences for primer recognition and polymerase amplification.

Template 1

CAATCCGCCCTCACTACAACCGCGTCTCCAGTGTCCTCAGCTTCCGTCCATCGGCGC
GGCGTTAGAATAGAGGAGAGAGAACACACACATAGAAGAAAGAAAGAAAGTCACA
TACATCGCCGCGCCGTCCTGTTCATTTTATCATCTGTGCGGCGCGTTAGGAGAAGAA
AGAAGAAATATACACATAATAGAAGAAAGAGAATACATTCACATCGCGCCGCACTT
CCGTCCATTTTATTCTCTGTGATATGCGTTATTCTCTTTCTTCTATTTCACACATATTT
ATTCTTCCTGGTCTTTCACACACATCGCATATCACTCTCGTCCATTTTACTATCTGCG
CGCCGGCTTCTTTCTTTCTTTCTTCTACACACACATCTCTATTCCTTCCTTCTACACAC
ACATGCCGGCGCGCTTCCGTCCATTAGTGTGAGACCCTACTCTGGCGTCGATGAGGG
A

Template 2

CAATCCGCCCTCACTACAACCGGGTCTCTAGTGTTTCATTCTCTCCGAATGGGCTTAG
AAGGAAGGAATAGAGTCACGCACATAGATAGAAGAAAGGAATTGATACACATGCC
CATTCGGTCCTATCCATTTTATCCTCCGCCGCAGCGCTTAAGACCAGGAAGAATAAA
CACAGACATAAGAAGAAGAAAGAAAGTCAGACACATGCGCTGCGGCTCCTGTCCCT
TTCATTCTCTGGCTCAGGGCTTCTTTCTTTCTTCTTCTTATACACAGATTTCTTCTTTCT
TCTCCTACATACGCATGCCCTGAGCCTGCCGTCCATTCTATTCTCCGCTTGGTCCGTT
ATTCCTTTCTTCTATCTAGACACACATTCTCTCTCCTCTATTCTTCAGACACATCGGA
CCAAGCTCTCGTCCACCAGTGTGAGACGCTACTCTGGCGTCGATGAGGGA

Fluorescent strand 1

/5Cy5/AGAGAATAAAATGGACGGAA

Fluorescent strand 2

/56-FAM/AGAGAATGAAAGGGACAGGA

BbvCI invader strand

ACACTAGTCTTCATGGACGG

A.2 Chapter 6

Short stem hairpin (SSH)

CGCGAGCGTCCTGCATCGCTCGCGTTTCTCTTCTCTTCTTCCGGAAAAGGGGAAGAA
GCCTTTTCCTTACCTCC

SSH toehold displacement strand A1

GGGAGTAGGGAAGAAGAGAAGAGA

SSH toehold displacement strand A2

TCTCTTCTCTTCTTCCCTACTCCC

SSH toehold displacement strand B1
GGGAGTAGGGAGGTAAGGAAAAGG

SSH toehold displacement strand B2

CCTTTTCCTTACCTCCCTACTCCC

SSH control A

CGCGAGCGTCCTGCATCGCTCGCGTTTCTCTTCTCTTCTTCCACCGGCTAGGAAGAA
GCCTTTTCCTTACCTCC

SSH control B

CGCGAGCGTCCTGCATCGCTCGCGTTTCTCTTCTCTTCTTCCGGAAAAGGACCGGCT
ACCTTTTCCTTACCTCC

Long stem hairpin (LSH)

CGGTACCGGTTACGGTACCGTTTCTTTCCTCCCCCCCCCATTCTTCTTTGTAAGGTGGA
AGAGGGGGAAAGAAGAATGGGGGGGGCCCCCTCTTCCACCTTACCTTCCCCC

LSH toehold displacement strand A1

AAAGAAGAATGGGGGGGGAGGAAAGAGGGGGGGG

LSH toehold displacement strand A2

CCCCCCCCTCTTTCCTCCCCCCCCCATTCTTCTTT

LSH toehold displacement strand B1

AGAGGGGAGGGGGAAGGTAAGGTGGAAGAGGGGG

LSH toehold displacement strand B2

CCCCCTCTTCCACCTTACCTTCCCCCTCCCCTCT

LSH control A

CGGTACCGGTTACGGTACCGTTTCTTTCCTCCCCCCCCCATTCTTCTTTACCCCCCCCCC
CCCCCCCCAAAAGAAGAATGGGGGGGGCCCCCTCTTCCACCTTACCTTCCCCC

LSH control B

CGGTACCGGTTACGGTACCGTTTCTTTCCTCCCCCCCCCATTCTTCTTTGTAAGGTGGA
AGAGGGGGAAAAAGAGAGAGGAGGCGCCCCCTCTTCCACCTTACCTTCCCCC

Extended long stem hairpin (ELSH)

GCAGATCGTTCATTTCCGATCTGCTTTCTTTTTCCCCTCTTTCCCACTCTCTTCGGGAA
AAGGAAGAGAGTGGGAAAGAGGGCCCACTCTCTTCCTTTTCCCCTTCCCTC

ELSH toehold displacement strand A1

GAAGAGGGGGGAAGGAAGGAGGGAGGTAGGTAAG

ELSH toehold displacement strand A2

CTTACCTACCTCCCTCCTTCCTTCCCCCCTCTTC

ELSH toehold displacement strand B1

AGAGGGGAGGGGGAAGGTAAGGTGGAAGAGGGGG

ELSH toehold displacement strand B2

CCCCCTCTTCCACCTTACCTTCCCCCTCCCCTCT

ELSH control A
GCAGATCGTTCATTTCCGATCTGCTTTCTTTTTCCCCTCTTTCCCACTCTCTTCTCCCT
ATTGAAGAGAGTGGGAAAGAGGGCCCACTCTCTTCCTTTTCCCCTTCCCTC

ELSH control B

GCAGATCGTTCATTTCCGATCTGCTTTCTTTTTCTCCCTATTCCCACTCTCTTCGGGAA
AAGGAAGAGAGTGGGAAAGAGGGCCCACTCTCTTCCTTTTCCCCTTCCCTC

Mirrored extended long stem hairpin (MELSH)

CCTCCCTCCTTCCTTCCCCCCTCTTCGTAAGGTGGAAGAGGGGGGAAGGAAGCCCCC
TCTTCCACCTTACCTTCCCCCTTGTCTCGCGATCTCGCGAGAC

MELSH control A

CCTCCCTCCTTCCTTCCCCCCTCTTCAAAACCCAGAAGAGGGGGGAAGGAAGCCCCC
TCTTCCACCTTACCTTCCCCCTTGTCTCGCGATCTCGCGAGAC

MELSH control B

CCTCCCTCCTTCCTTCCCCCCTCTTCGTAAGGTGGAAGAGGGGGAAAACCCACCCCC
TCTTCCACCTTACCTTCCCCCTTGTCTCGCGATCTCGCGAGAC

A.3 Primer sequences

Forward primer for Template 1 and 2 of my multi-knit DNA structures

CAATCCGCCCTCACTACAACCG

Reverse primer for Template 1 and 2 of my multi-knit DNA structures

TCCCTCATCGACGCCAGAGTAG

Forward primer for adding the loxP+BamHI sequence to my multi-knit DNA template

AGTGAATTCGAGCTCGGTACCCGGGATAACTTCGTATAGCATACATTATACGAAGTT
ATCCTCAGCTTCCGTCCATCGG

Reverse primer for adding the loxP+BamHI sequence to my multi-knit DNA template

GCCTGCAGGTCGACTCTAGAGGATCATAACTTCGTATAATGTATGCTATACGAAGTT
ATGTGGACGAGAGCTTGGTCC

B Supporting Figures

B.1 Chapter 3

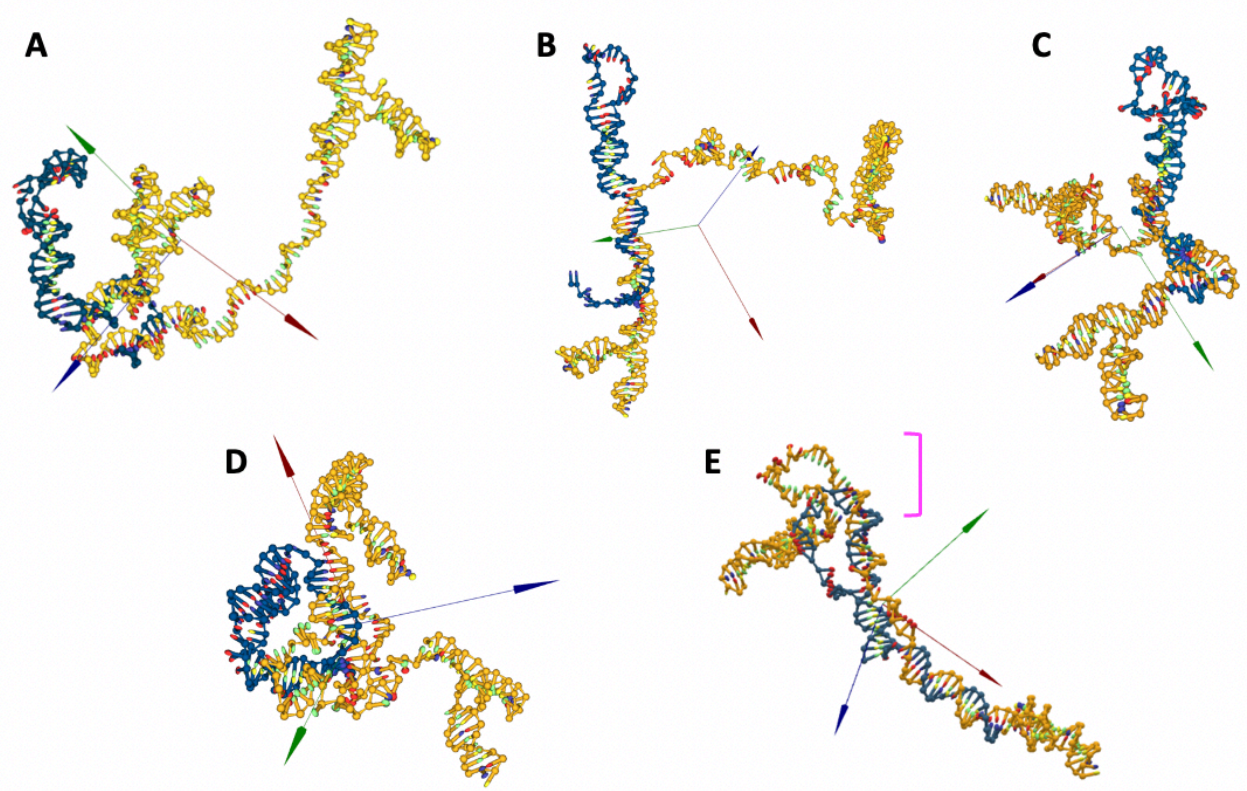


Figure B.1: Characterising increasing lengths of B domain in the HT model. A) 12 nt, B) 13 nt, C) 14 nt, D) 15 nt and E) 16 nt. Only the 16 nt B domain thread (dark blue) has successfully looped into the hairpin (yellow) as highlighted by the pink bracket in E). Simulations were run in oxDNA for 5e7 steps.

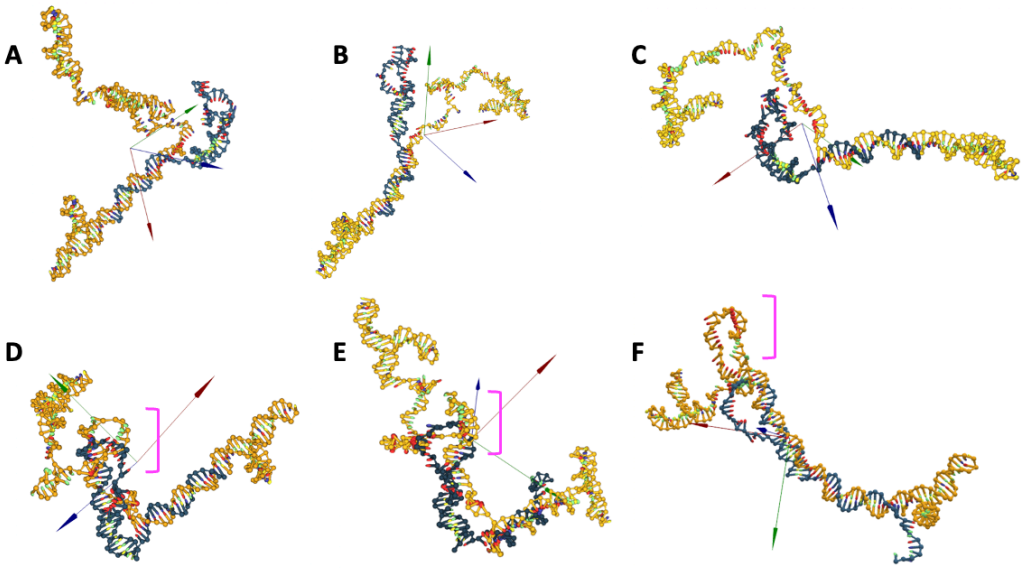


Figure B.2: Characterising increasing lengths of C domain in the HT model. A) 12 nt, B) 13 nt, C) 14 nt, D) 15 nt, E) 16 nt and F) 17 nt. A-C) Show no looping of the thread (yellow) into the hairpin (dark blue). D-F) show looping as indicated by the pink bracket. Simulations were run in oxDNA for 5e7 steps.

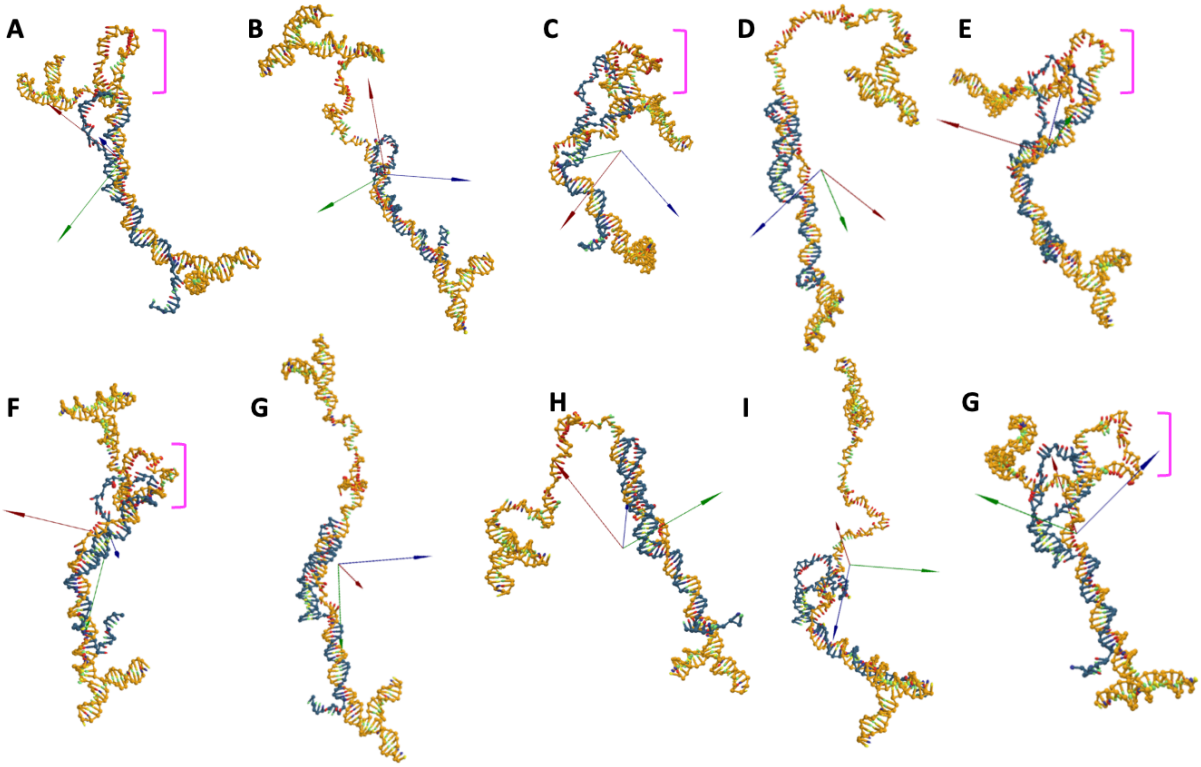


Figure B.3: Quantifying the looping efficiency of the 15 nt C-domain HTL model. Ten simulations of the design were run for 1e8 steps in oxDNA. A, C, E, F and G) were successful in successfully looping the thread (dark blue) into the hairpin (yellow) as highlighted by the pink bracket.

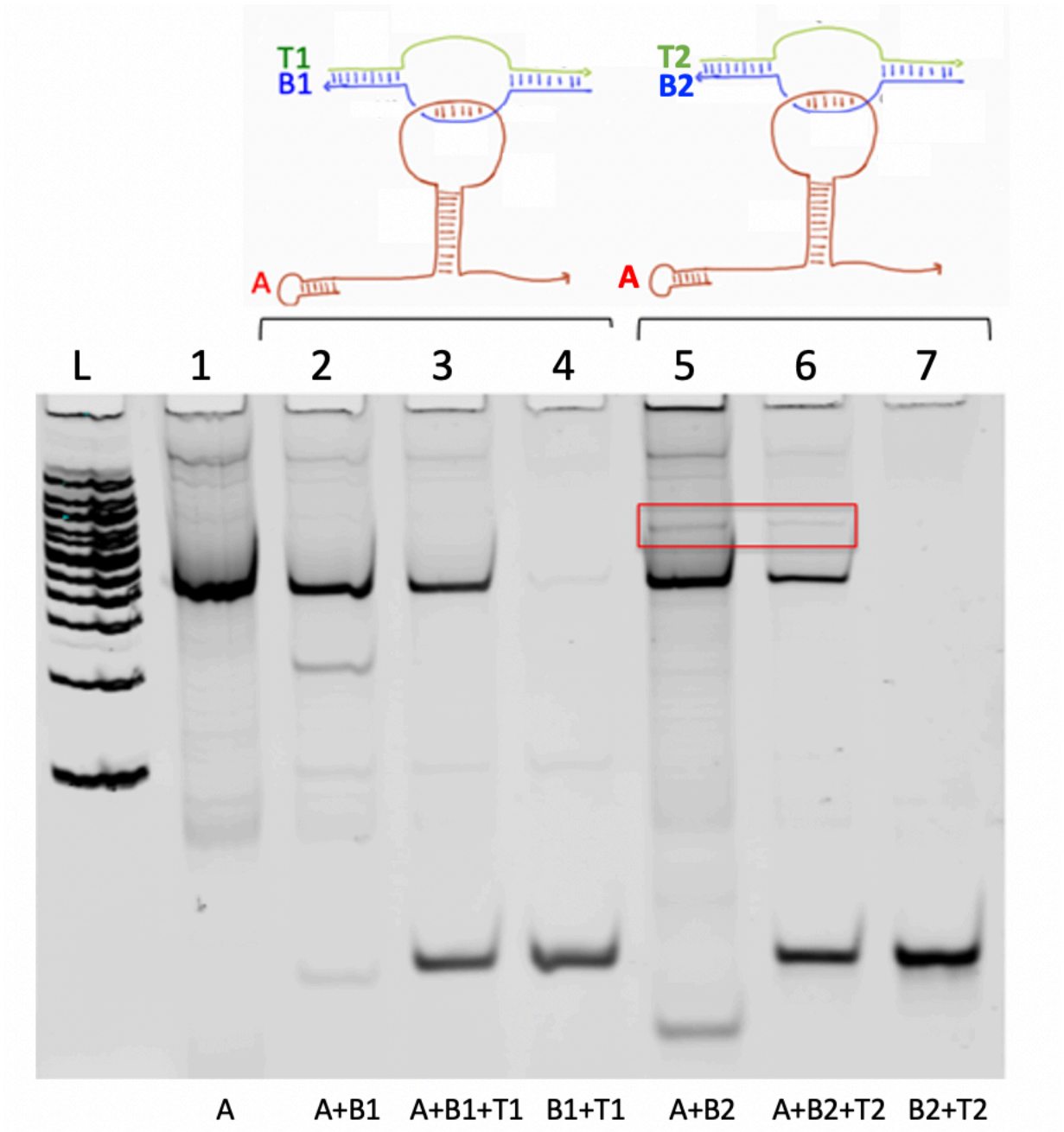


Figure B.4: An alternative lock strand design using two conjoining lock pieces. PAGE analysis of the alternative lock system contain a top (green) and bottom (blue) strand that binds the hairpin loop of hairpin A (red). Lane 1 contains hairpin A. Lanes 2-4 show one set of lock sequences: lane 2 contains hairpin A and the bottom lock strand 1, lane 3 contains hairpin A with the bottom lock strand 1 and top lock strand 1, and lane 4 contains the bottom lock strand 1 and top lock strand 1. Lanes 5-6 show another set of lock sequences: lane 5 contains hairpin A and the bottom lock strand 2, lane 6 contains hairpin A with the bottom lock strand 2 and top lock strand 2, and lane 7 contains the bottom lock strand 2 and top lock strand 2. The potentially locked structure is highlighted by the red box in lanes 5 and 6. A 100 bp DNA ladder is used (L) and samples are stained for SYBRGold. Samples are run on a 10% polyacrylamide gel.

B.2 Chapter 4

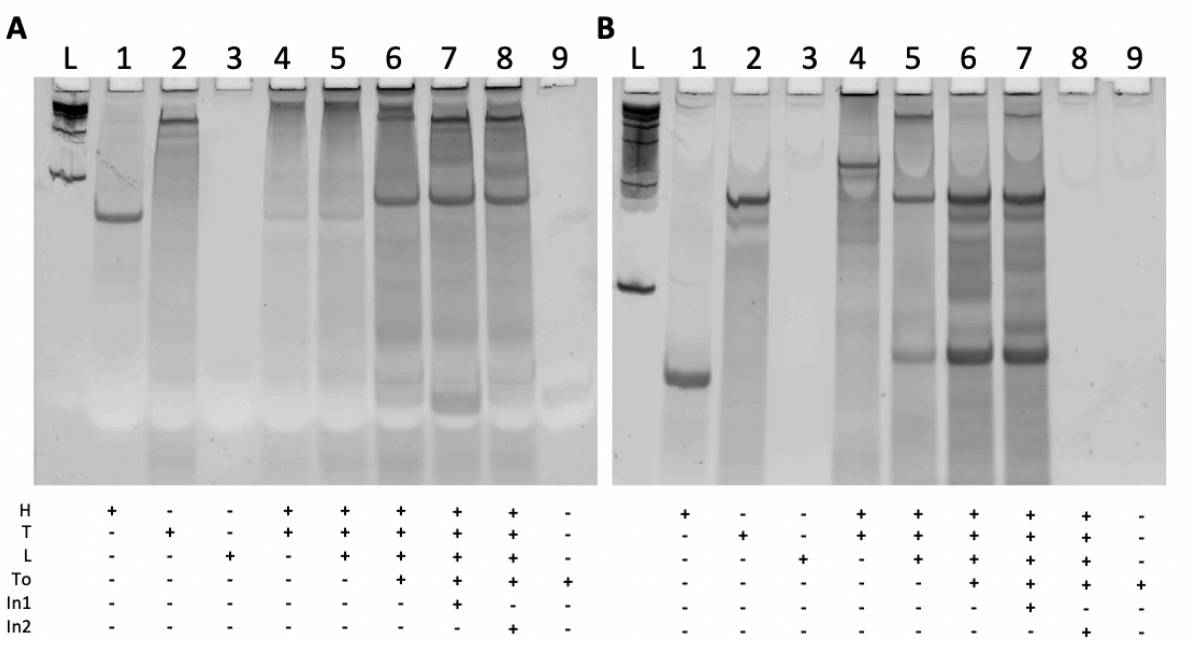


Figure B.5: Comparing polyacrylamide gel percentages for resolving the HTL model. A) 15% polyacrylamide and B) 10% polyacrylamide. Both gels are run for 45 minutes at 90 V. Lane 1 contains the hairpin, lane 2 contains the thread, lane 3 contains the lock, lane 4 contains the hairpin and thread, lane 5 contains the hairpin, thread and lock, lane 6 contains the hairpin, thread, lock and toehold displacement strand, lane 7 contains the hairpin, thread, lock, toehold displacement strand and invader strand 1, lane 8 contains the hairpin, thread, lock, toehold displacement strand and invader strand 2, and lane 9 contains the toehold displacement strand. The samples are stained for SYBRGold. A 100 bp DNA ladder is used in each L lane.

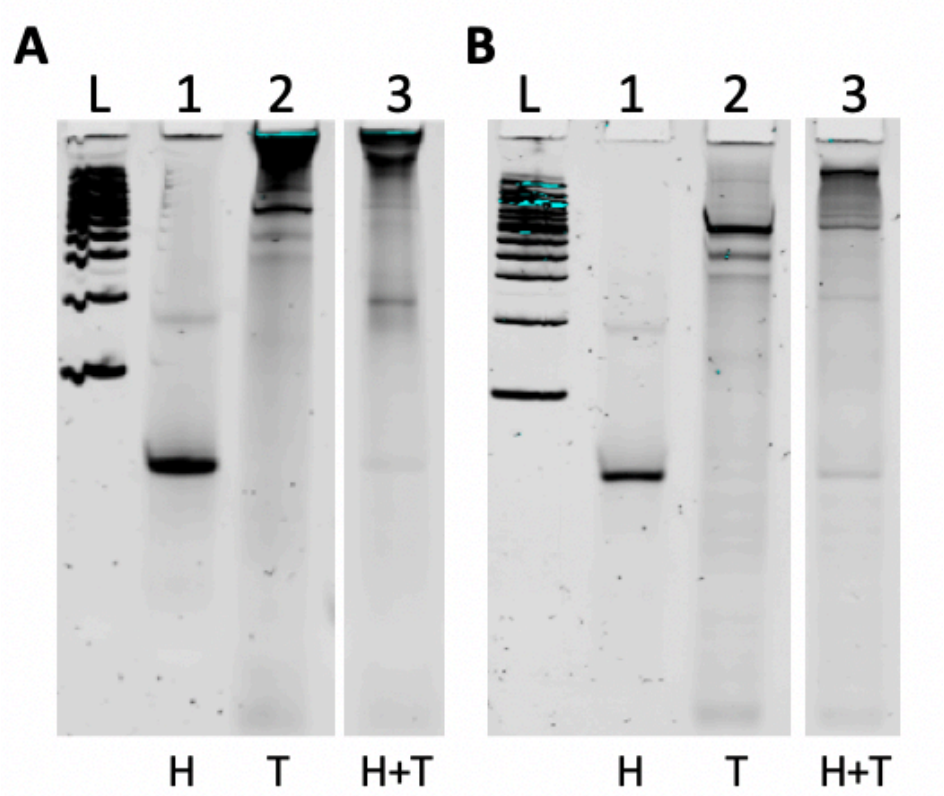


Figure B.6: Effect of quenching the thread strands for the HTL model. A) Non-quenched threads compared to B) quenched threads. Lane 1 contains the hairpin, lane 2 contains the thread, and lane 3 contains both the hairpin and thread combined. The samples are stained for SYBRGold. A 100 bp DNA ladder is used in each L lane. Samples are run on 10% polyacrylamide gels.

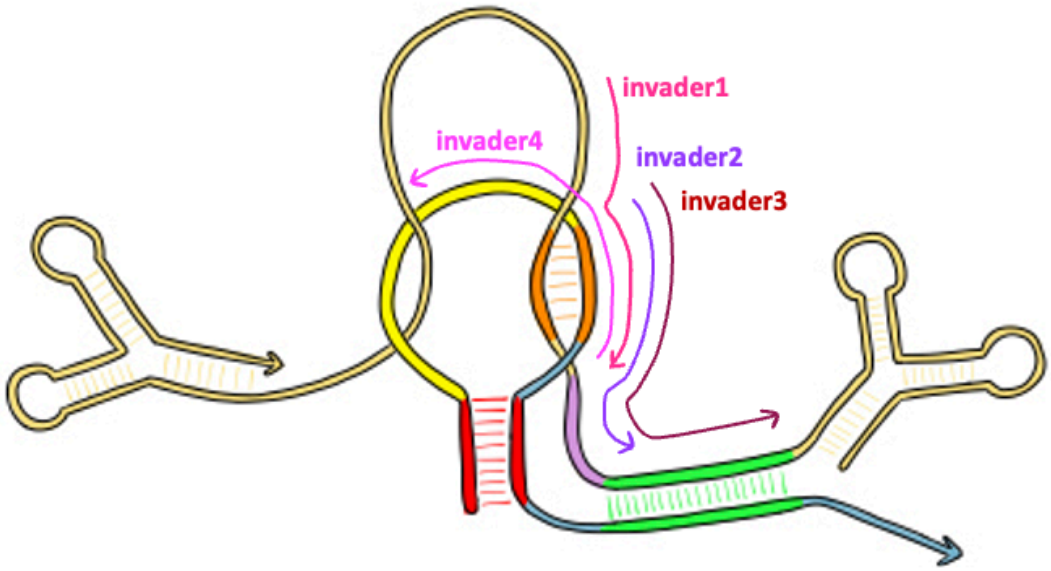


Figure B.7: Schematic of the various invader strands designed for the HTL model. Invader1 (dark pink) binds the L and C domains of the thread. Invader2 ((purple) binds C* and K domains of the thread. Invader3 (maroon) binds C*, K and E domains of the thread. Invader4 (light pink) binds C and B domains of the hairpin.*

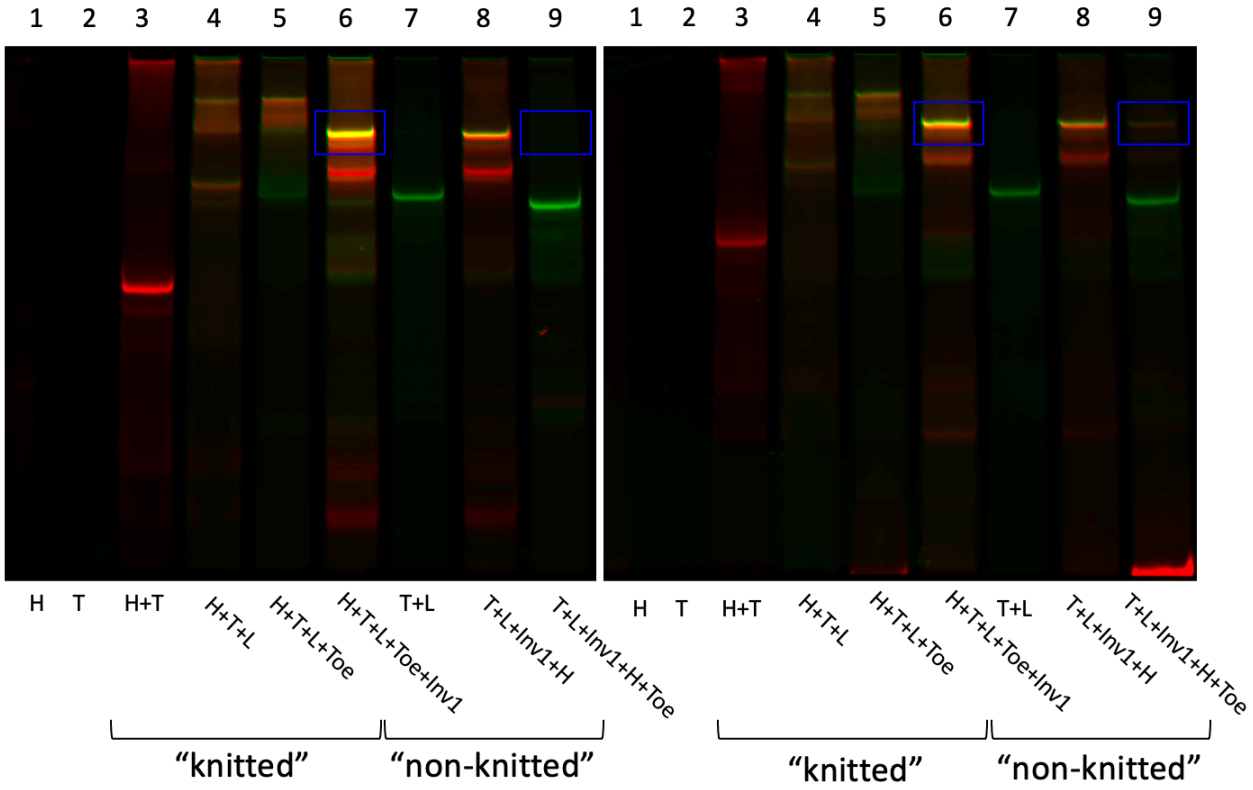


Figure B.8: Two experimental replicates of the single-knit reaction using different samples. The two gels shown the same set up using different biological samples. Lanes 1 and 2 contain the hairpin and thread, respectively. Lanes 4-6 depict the "knitted" reaction: lane 4 contains the hairpin with thread and lock, lane 5 contains the hairpin with thread, lock and toehold displacement strand, and lane 6 contains the hairpin with thread, lock, toehold displacement strand and invader strand. Lanes 7-9 depict the "non-knitted" reaction: lane 7 contains the thread and lock, lane 8 contains the thread, lock, invader strand and hairpin, and lane 9 contains the thread, lock, invader strand, hairpin and toehold displacement strand. Both gels show the same results. The blue boxes in lanes 6 highlighted the conserved "knitted" structure whereas the blue boxes in lane 9 show that the "non-knitted" reaction return to its original components. Filters for the Cy5 and FAM signals were used. Samples are run on 10% polyacrylamide gels.

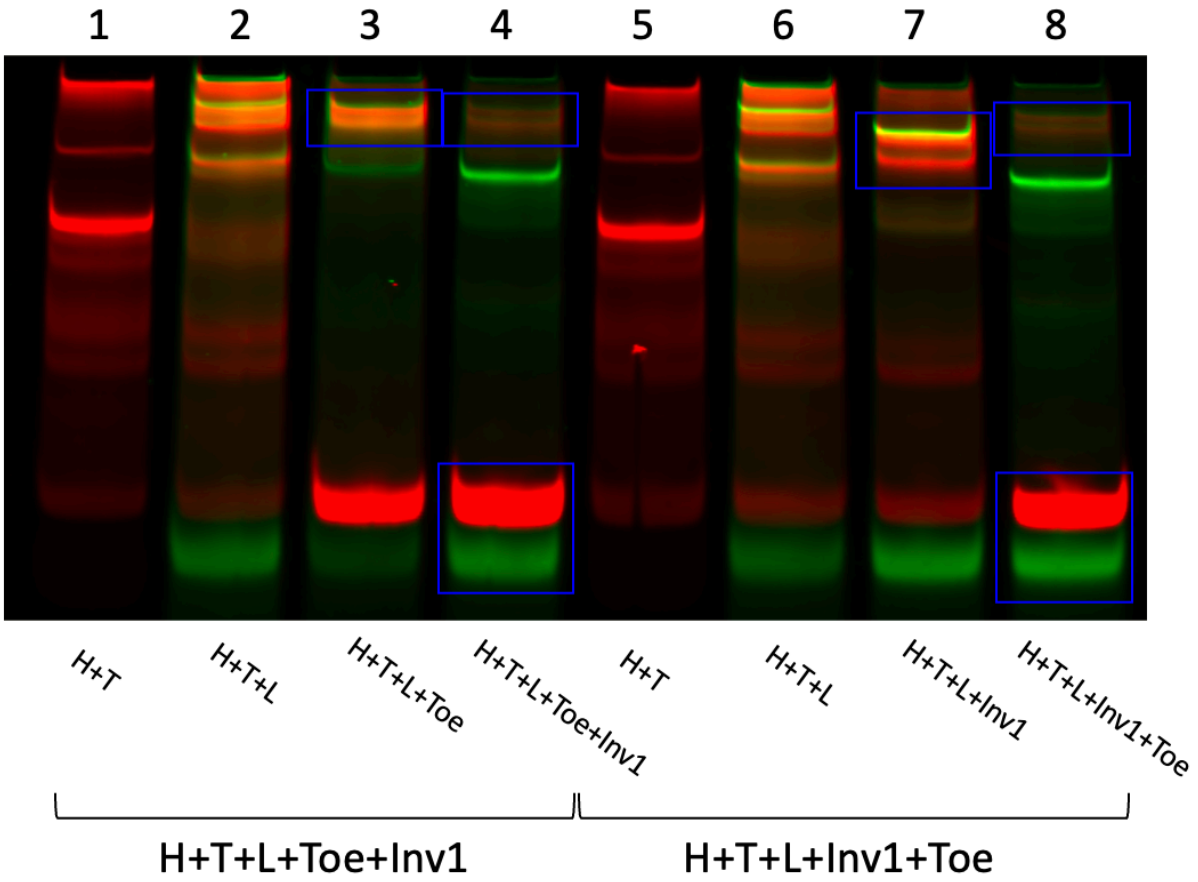


Figure B.9: Comparing the order of displacement strand addition in the HTL model. PAGE analysis of the hairpin, thread and lock combined with either the toehold displacement strand first as shown in lanes 1-4, or the invader strand first as shown in lanes 5-8. Lanes 1 and 5 contain the hairpin and thread. Lanes 2 and 6 contain the hairpin with thread and lock. Lane 3 contains the hairpin with thread, lock and toehold displacement strand. Lane 4 contains the hairpin with thread, lock, toehold displacement strand and invader strand. Lane 7 contains the hairpin with thread, lock and invader strand. Lane83 contains the hairpin with thread, lock, invader strand and toehold displacement strand. Both reactions result in maintained "knitted" structure as highlighted by the blue boxes at the top, albeit at a low concentration. Blue boxes at the bottom of the gel highlight the free hairpin and lock structures that appear after displacement, suggesting that the knitting reaction was not efficient. Filters for the Cy5 and FAM signals were used. Samples are run on a 10% polyacrylamide gel.

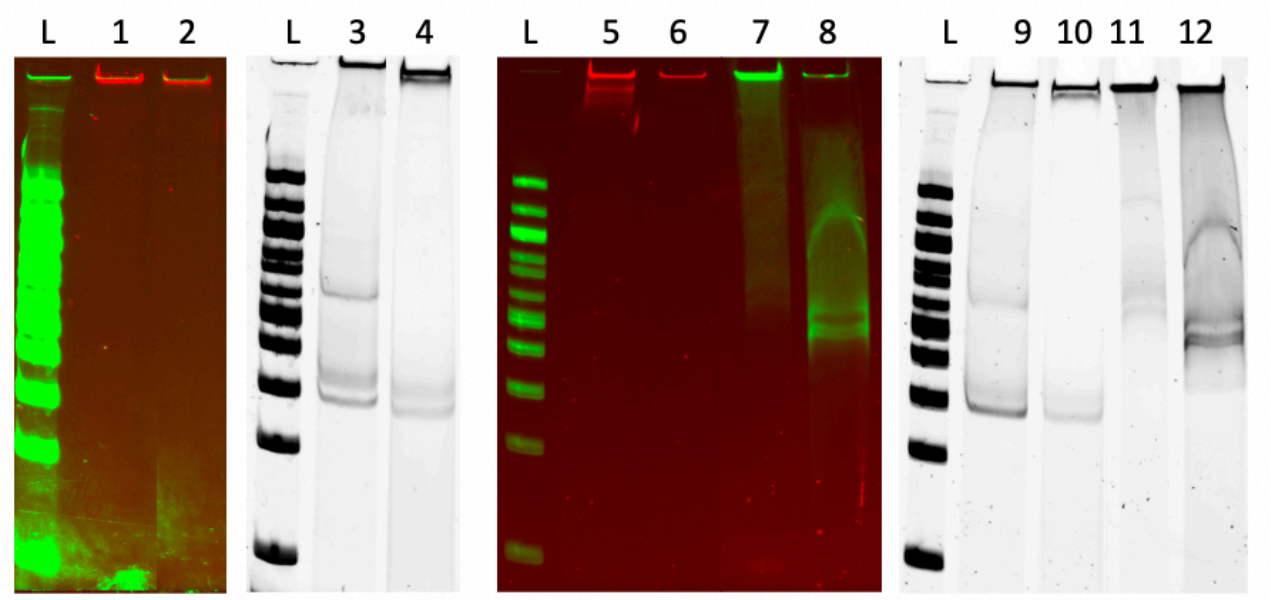


Figure B.10: Comparing Phi29 and Bst3 as rolling circle amplification polymerases for my multi-knit DNA design. Lane 1 contains the circularised and nicked template with Phi29, while Lane 2 uses Bst3. Both lanes utilise the fluorescent Cy5 and FAM probes. Lanes 3 and 4 are the respective previous lanes stained with SYBRGold. Lane 5 uses the product in lane 1 and binds the invader strand for BbvCI digestion. Lane 6 uses the product in lane 2 and binds the invader strand for BbvCI digestion. Lane 7 contains the two half templates combined at the BsaI restriction site while lane 8 circularises the template at both the BsaI and BsmBI sites. Lanes 9, 10, 11 and 12 are the previous four lanes stained with SYBRGold. A 100 bp DNA ladder (L) is used as reference. Samples are run on 10% polyacrylamide gels.

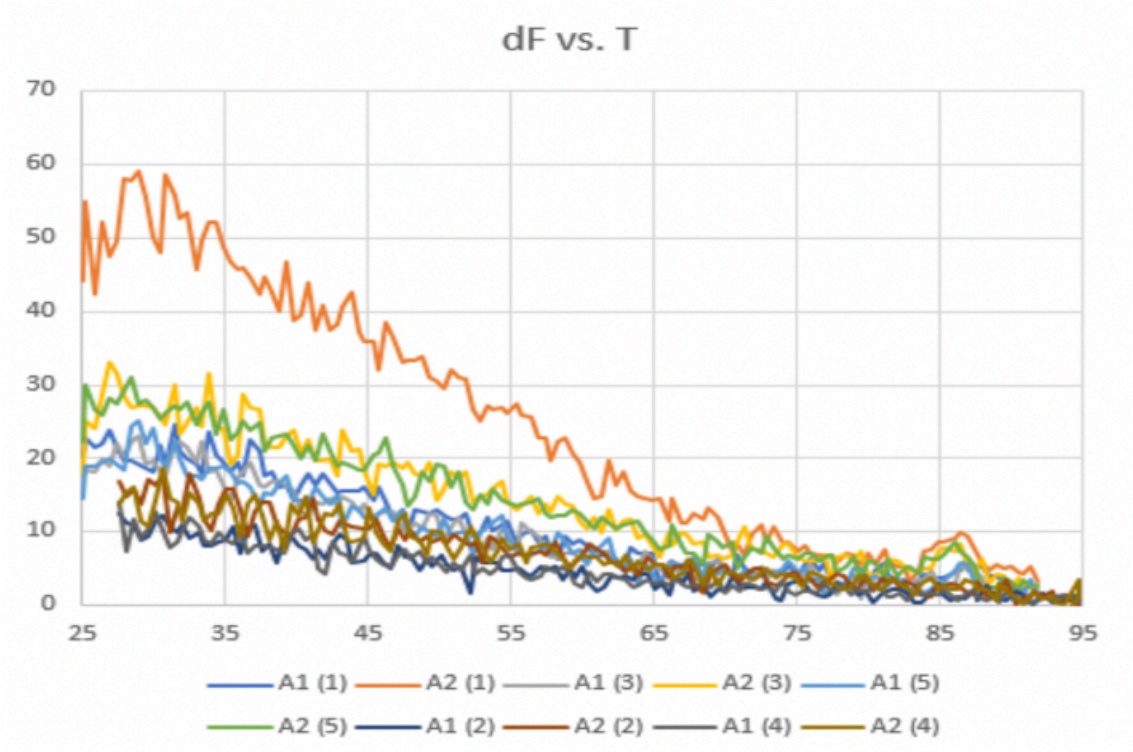


Figure B.11: Change in fluorescence with temperature of the Phi29 (A1) and Bst3 (B2) amplified multi-knit DNA nanostructures. The numbers 1-5 refer to the segment of the RT-qPCR protocol, where 1 represents the first cooling stage to 25 °C for 30 minutes, followed by 2 which represents the first heating stage to 95 °C for30 minutes. 3 and 5 are subsequent cooling stages while 4 is another heating stage.

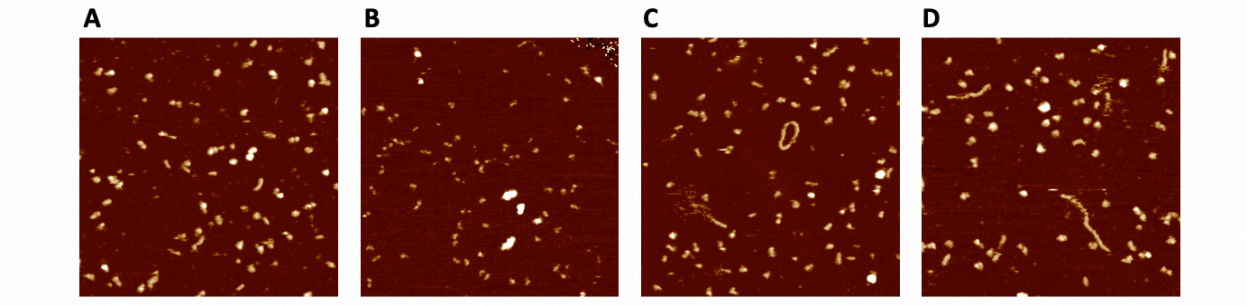


Figure B.12: Varying results from the same AFM protocol. Using the same Pyne group adsorption protocol lead to inconsistent immobilisation. As seen in A) and B) there are a few clumped regions which are likely the larger DNA templates that have condensed too much. On the other hand you can see clear circularised and linear templates in C) and D). These micrographs are 1000 by 1000 nm.

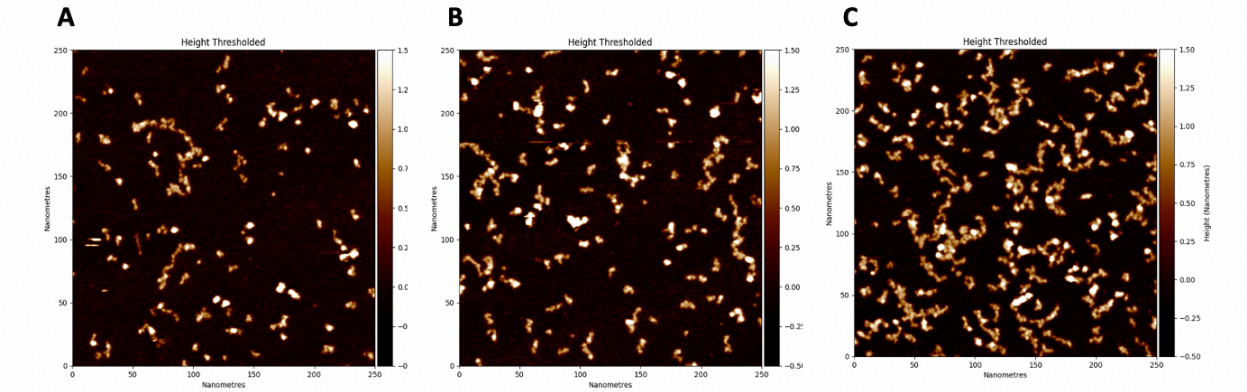


Figure B.13: Wide view (250 by 250 nm) AFM micrographs of the multi-knit DNA nanostructures. A-C) are different points of view. The height of the structures is given on the right.

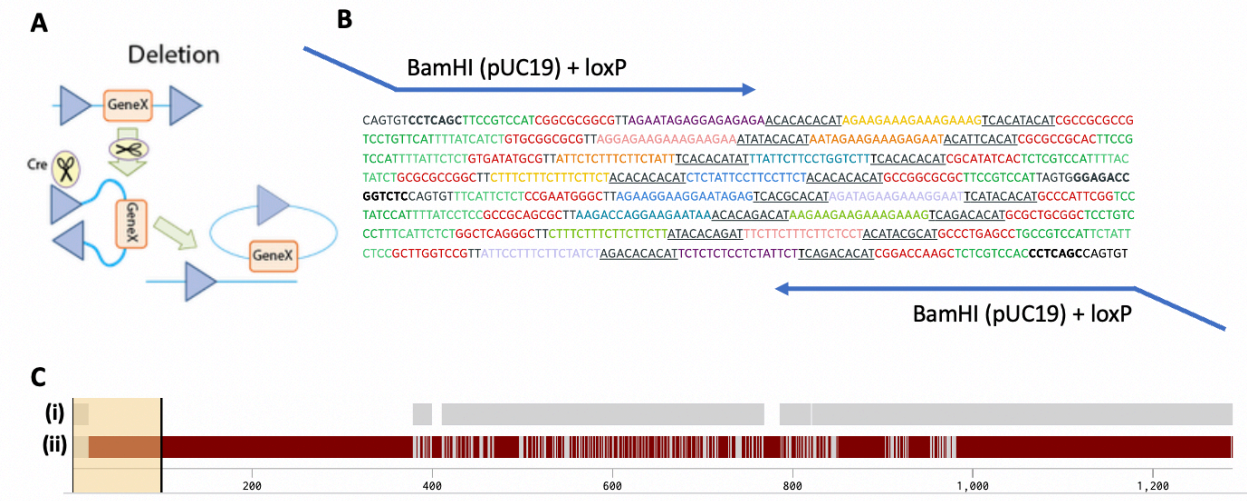


Figure B.14: Overview of Cre-loxP system to circularise and amplify my RCA multi-knit DNA template. A) Schematic of the Cr-loxP system. A sequence (GeneX) is inserted between two loxP sequences. The Cre enzyme then circularises the sequence around the loxP sequence. Figure is taken from addgene.com. B) Schematic of my RCA multi-knit template with primers to include the loxP sequence as well as a BamHI recognition site for pUC19 insertion on the 5' and 3' ends of the sequence. C) Sequence alignment after circularising and transforming the plasmid containing the multi-knit template into E. coli cells. (i) Original template sequence aligned against (ii) sequenced plasmid as amplified within the cell. Regions of grey are shared sequences, whereas the red regions represent the mutated bases. Alignment was conducted in Benchling.com.

B.3 Chapter 5

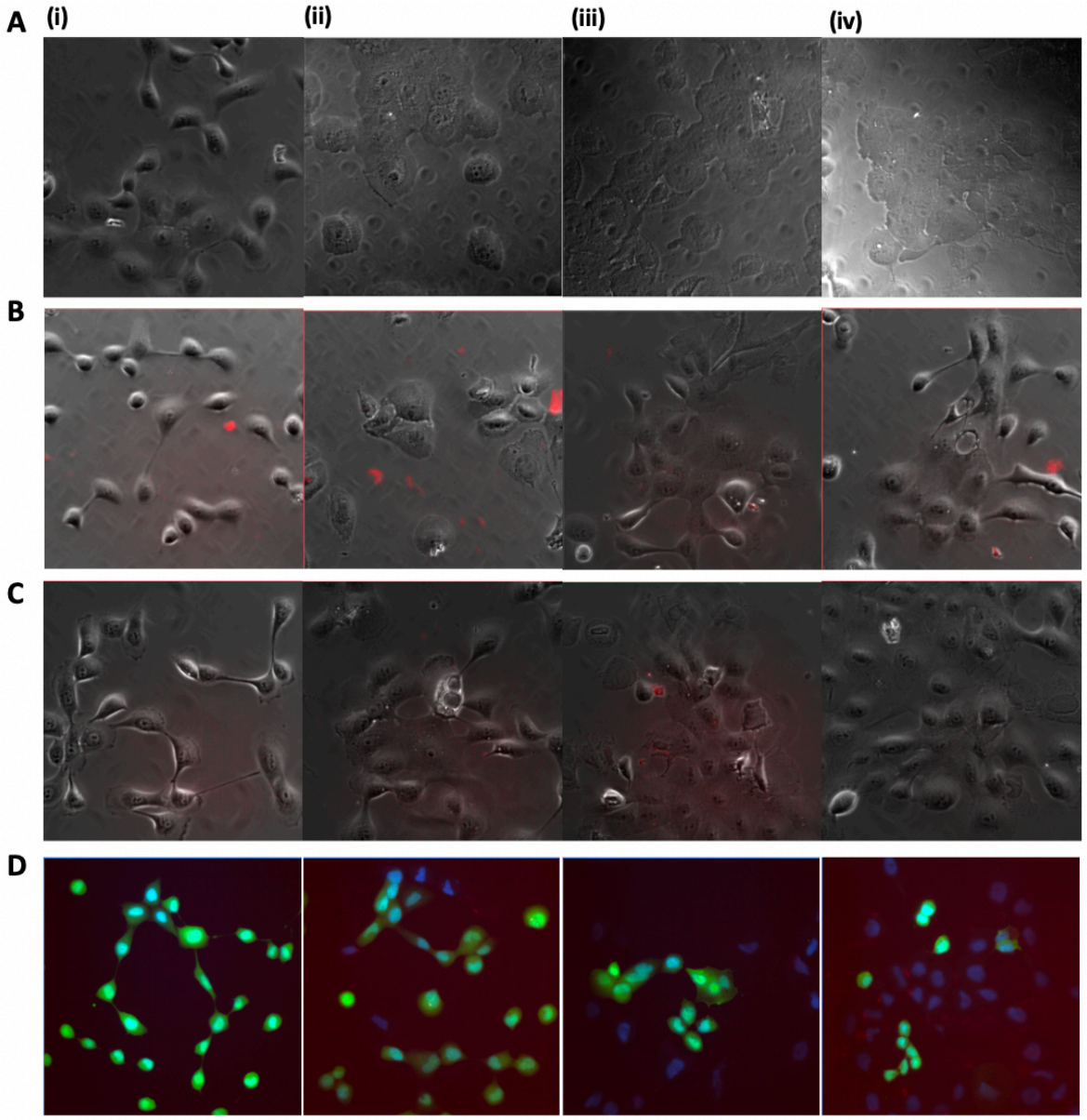


Figure B.15: Testing out various confinement pressures on MC10A cells A) during confinement, B) immediately after confinement, C) 120 minutes after confinement and D) 160 minutes. (i) is a control at 0 kPa, (ii) is after 5 kPa is applied, (iii) is after 20 kPa is applied and (iv) is after 30 kPa is applied to the cells. A-C) showcases bright-field images of cells while B) shows the Cy5-bound DNA signposts (red), a CellTracker Green CMFDA dye representing the cell membrane (green) and the nucleic stained with Hoechst (blue). Images are taken at 100X magnification with oil immersion.

B.3 Chapter 6

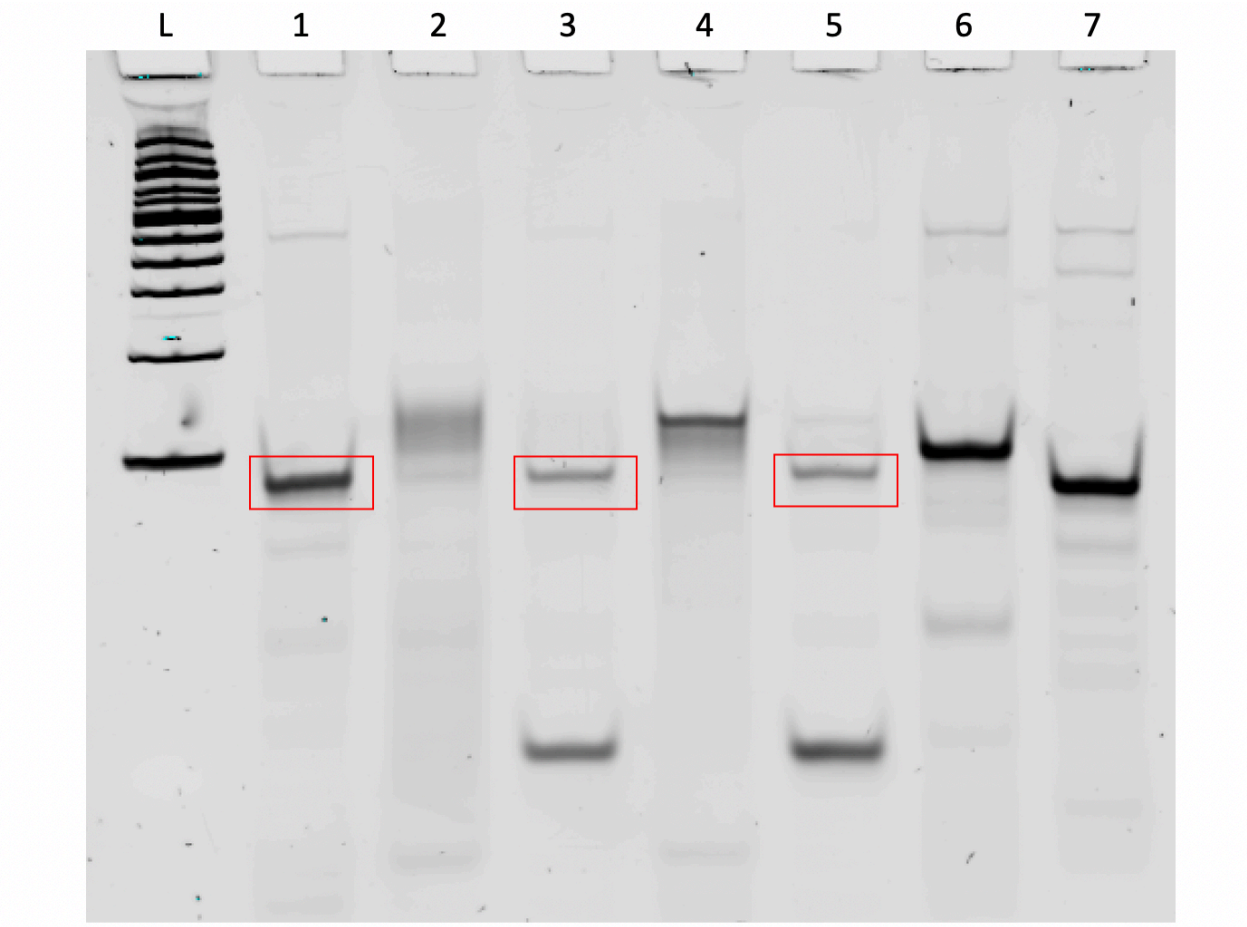


Figure B.16: Full PAGE analysis of the short stem hairpins. Lane 1 represents the AB hairpin with equal equilibrium probability for the A or B stem conformation. Lane 2 represents the AB hairpin combined with the toehold displacement strand A1 to switch the structure into the A stem conformation. Lane 3 shows the addition of toehold displacement strand A2 to remove toehold displacement strand A1. Lane 4 represents the AB hairpin combined with the toehold displacement strand B1 to switch the structure into the B stem conformation. Lane 5 shows the addition of toehold displacement strand B2 to remove toehold displacement strand B1. Lane 6 shows the control structure for the A stem conformation. Lane 7 shows the control structure for the B stem conformation. Red boxes highlight the AB, switch A state and switched B state. A 100 bp DNA ladder (L) is used. The samples are run on a 10% polyacrylamide gel and stained with SYBRGold.

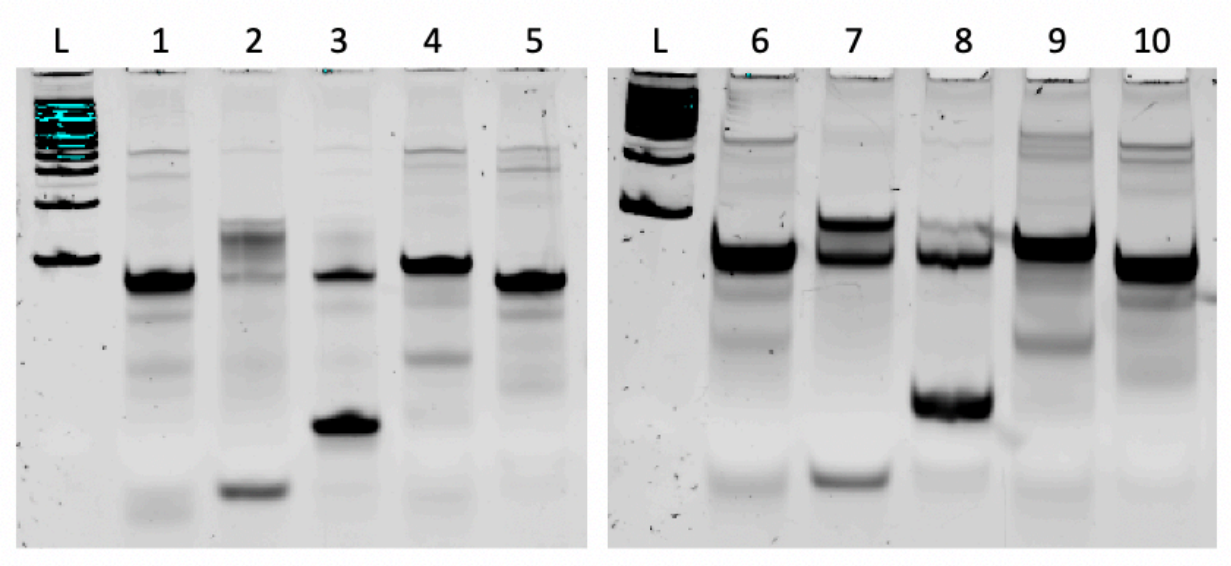


Figure B.17: Experimental replicates for the short stem hairpin design full PAGE analysis. Lanes 1 and 6 contain the AB hairpin, lane 2 contains the AB hairpin with the displacement strand (TA1) to switch the hairpin into the A conformation, lane 3 contains the sample in lane 2 with the addition of the displacement strand (TA2) to remove TA1, lanes 4 and 9 contain the B control, lanes 5 and 10 contain the B control, lane 7 contains the AB hairpin with the displacement strand (TB1) to switch the hairpin into the B conformation, lane 8 contains the sample in lane 7 with the addition of the displacement strand (TB2) to remove TB1. A 100 bp DNA ladder is used as reference (L). The samples in the 10% polyacrylamide gel were stained for SYBRGold.

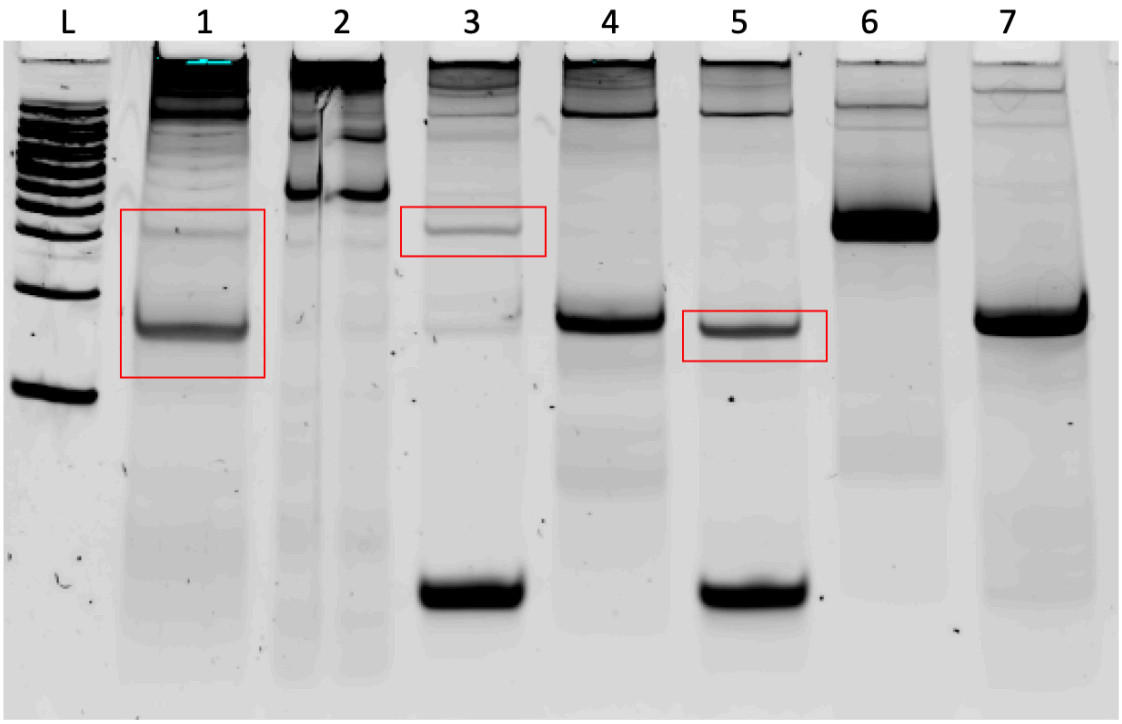


Figure B.18: Full PAGE analysis of the long stem hairpins. Lane 1 represents the AB hairpin with equal equilibrium probability for the A or B stem conformation. Lane 2 represents the AB hairpin combined with the toehold displacement strand A1 to switch the structure into the A stem conformation. Lane 3 shows the addition of toehold displacement strand A2 to remove toehold displacement strand A1. Lane 4 represents the AB hairpin combined with the toehold displacement strand B1 to switch the structure into the B stem conformation. Lane 5 shows the addition of toehold displacement strand B2 to remove toehold displacement strand B1. Lane 6 shows the control structure for the A stem conformation. Lane 7 shows the control structure for the B stem conformation. Red boxes highlight the AB, switch A state and switched B state. A 100 bp DNA ladder (L) is used. Samples are run on a 10% polyacrylamide gel and stained with SYBRGold

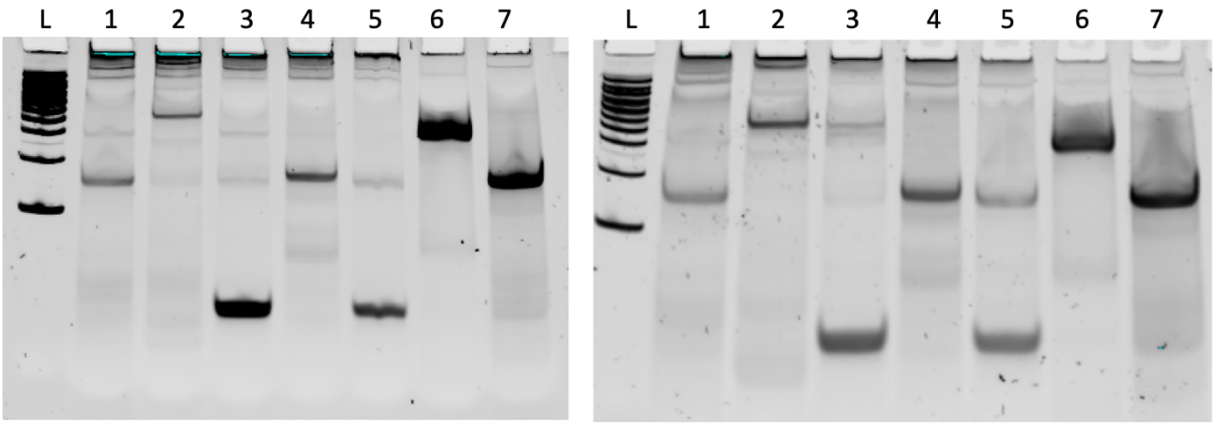


Figure B.19: Two experimental replicates for the long stem hairpin design full PAGE analysis. Lane 1 contains the AB hairpin, lane 2 contains the AB hairpin with the displacement strand (TA1) to switch the hairpin into the A conformation, lane 3 contains the sample in lane 2 with the addition of the displacement strand (TA2) to remove TA1, lane 4 contains the AB hairpin with the displacement strand (TB1) to switch the hairpin into the B conformation, lane 5 contains the sample in lane 4 with the addition of the displacement strand (TB2) to remove TB1, lane 6 contains the A control and lane 7 contains the B control. A 100 bp DNA ladder is used as reference (L). The samples in the 10% polyacrylamide gel were stained for SYBRGold.

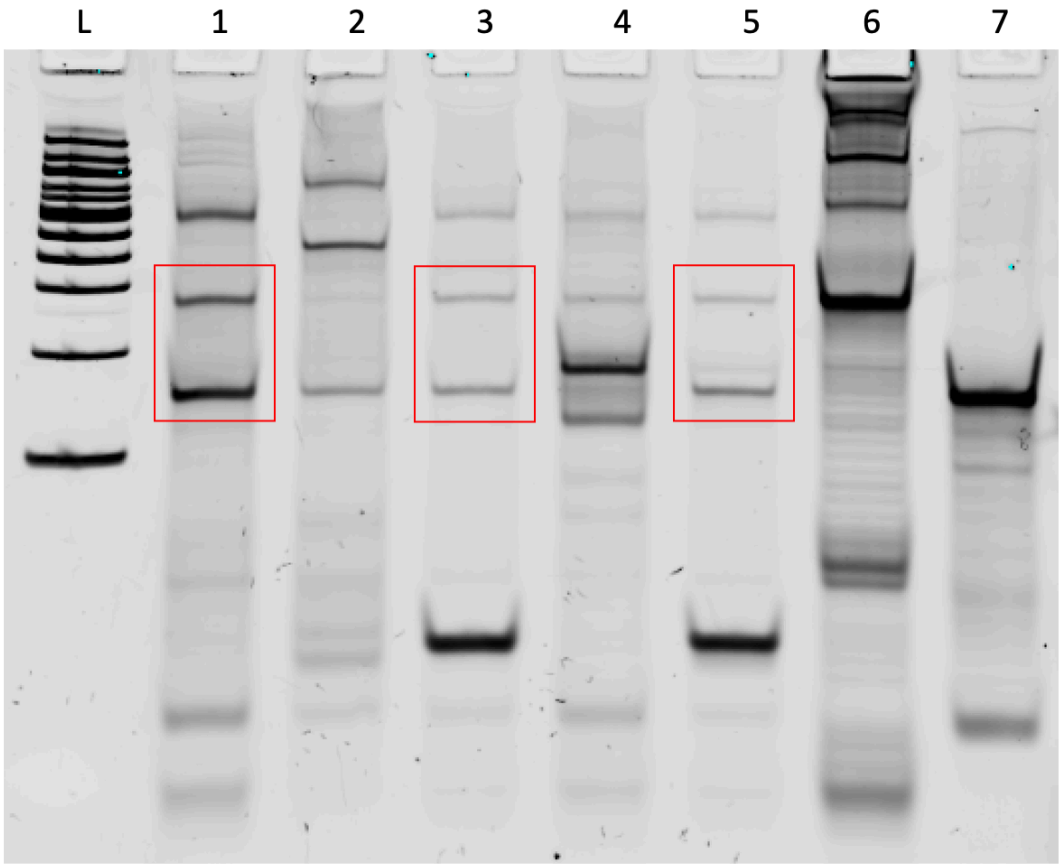


Figure B.20: Full PAGE analysis of the extended long stem hairpins. Lane 1 represents the AB hairpin with equal equilibrium probability for the A or B stem conformation. Lane 2 represents the AB hairpin combined with the toehold displacement strand A1 to switch the structure into the A stem conformation. Lane 3 shows the addition of toehold displacement strand A2 to remove toehold displacement strand A1. Lane 4 represents the AB hairpin combined with the toehold displacement strand B1 to switch the structure into the B stem conformation. Lane 5 shows the addition of toehold displacement strand B2 to remove toehold displacement strand B1. Lane 6 shows the control structure for the A stem conformation. Lane 7 shows the control structure for the B stem conformation. Red boxes highlight the AB, switch A state and switched B state. A 100 bp DNA ladder (L) is used. Samples are run on a 10% polyacrylamide gel and stained with SYBRGold.

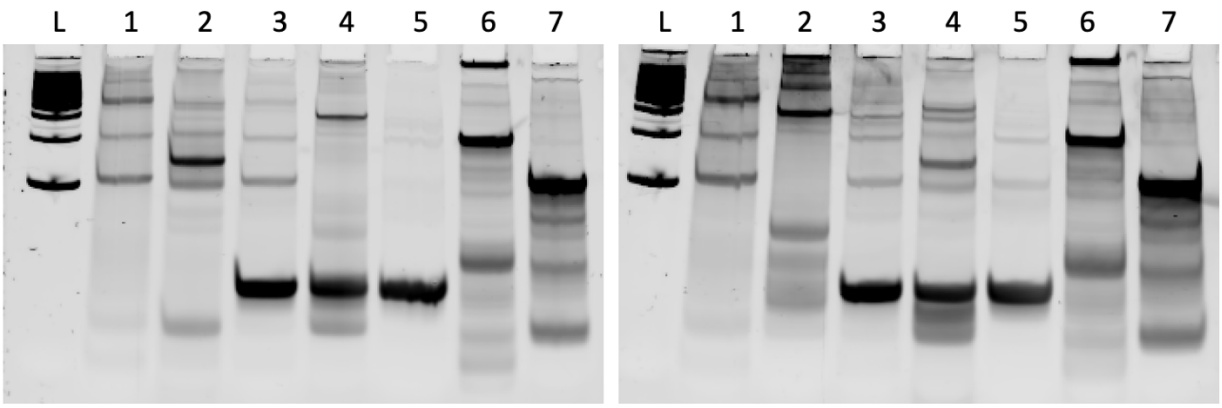


Figure B.21: Two experimental replicates for the extended long stem hairpin design full PAGE analysis. Lane 1 contains the AB hairpin, lane 2 contains the AB hairpin with the displacement strand (TA1) to switch the hairpin into the A conformation, lane 3 contains the sample in lane 2 with the addition of the displacement strand (TA2) to remove TA1, lane 4 contains the AB hairpin with the displacement strand (TB1) to switch the hairpin into the B conformation, lane 5 contains the sample in lane 4 with the addition of the displacement strand (TB2) to remove TB1, lane 6 contains the A control and lane 7 contains the B control. A 100 bp DNA ladder is used as reference (L). The samples in the 10% polyacrylamide gel were stained for SYBRGold.

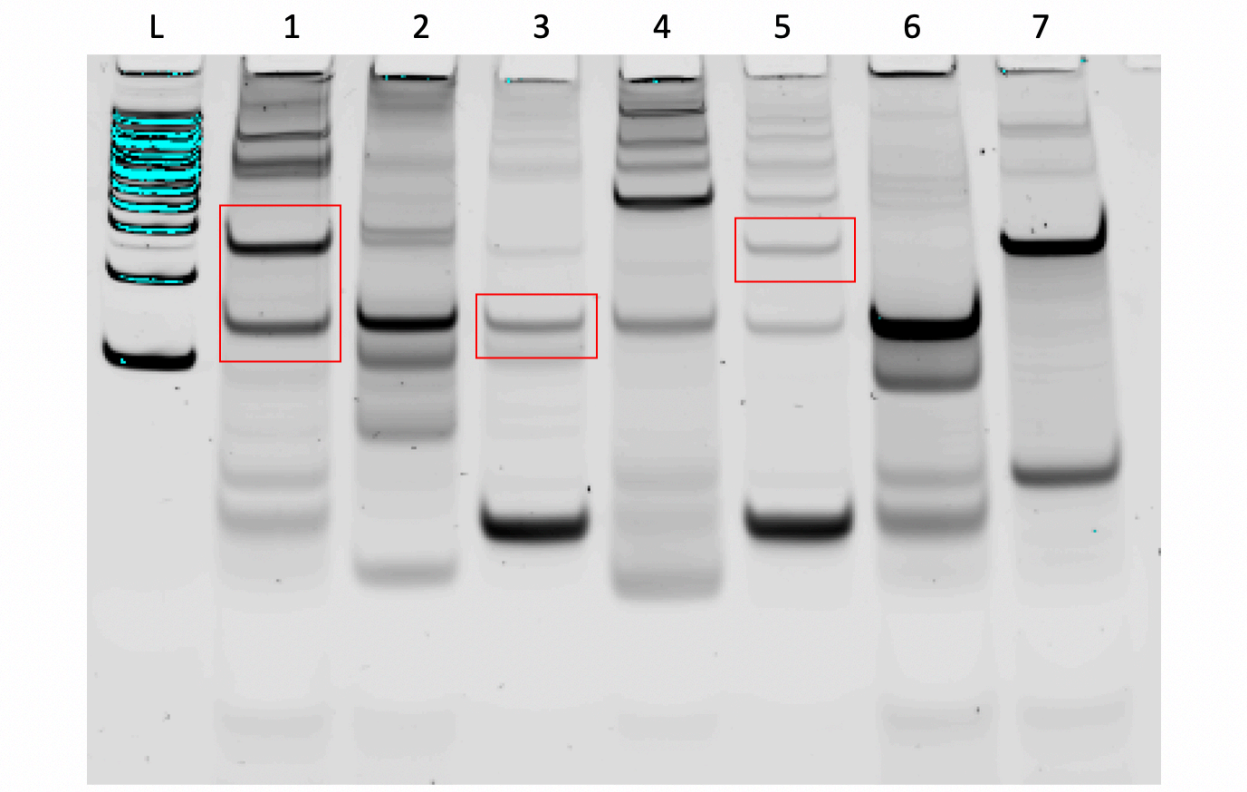


Figure B.22: Full PAGE analysis of the mirrored extended long stem hairpins. Lane 1 represents the AB hairpin with equal equilibrium probability for the A or B stem conformation. Lane 2 represents the AB hairpin combined with the toehold displacement strand A1 to switch the structure into the A stem conformation. Lane 3 shows the addition of toehold displacement strand A2 to remove toehold displacement strand A1. Lane 4 represents the AB hairpin combined with the toehold displacement strand B1 to switch the structure into the B stem conformation. Lane 5 shows the addition of toehold displacement strand B2 to remove toehold displacement strand B1. Lane 6 shows the control structure for the A stem conformation. Lane 7 shows the control structure for the B stem conformation. Red boxes highlight the AB, switch A state and switched B state. A 100 bp DNA ladder (L) is used. Samples are run on a 10% polyacrylamide gel and stained with SYBRGold.

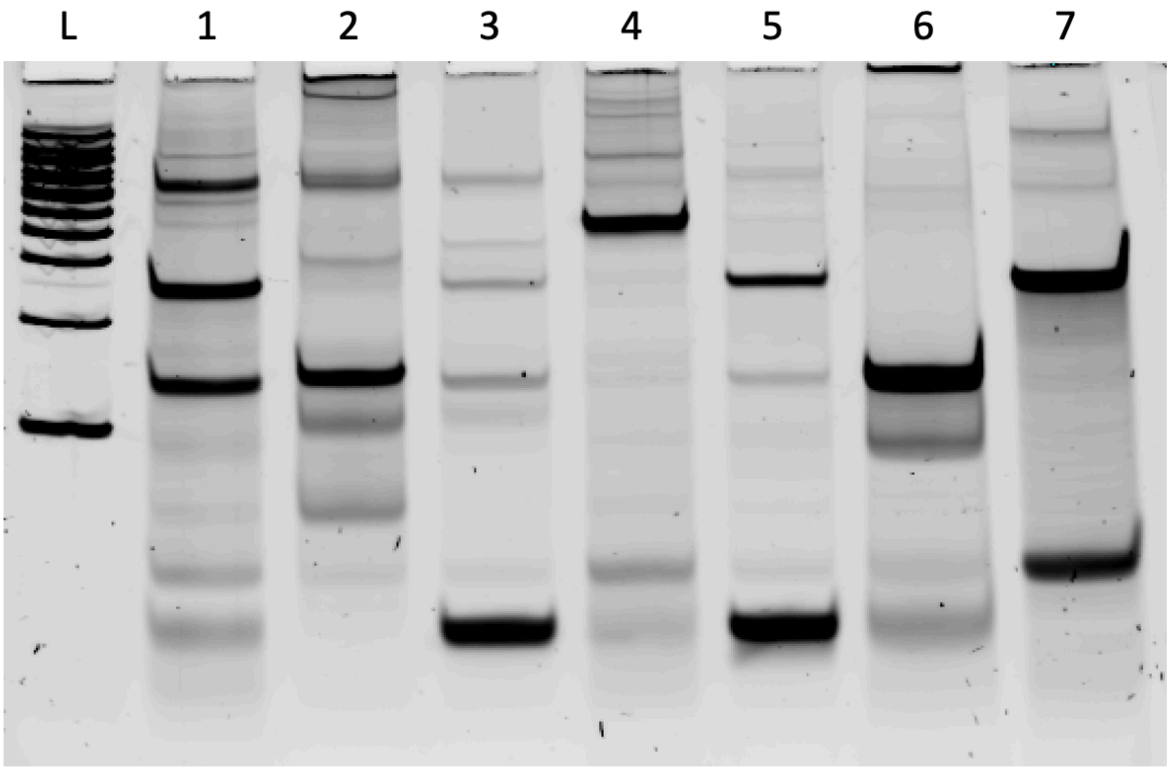


Figure B.23: Experimental replicate for the mirrored extended long stem hairpin design full PAGE analysis. Lane 1 contains the AB hairpin, lane 2 contains the AB hairpin with the displacement strand (TA1) to switch the hairpin into the A conformation, lane 3 contains the sample in lane 2 with the addition of the displacement strand (TA2) to remove TA1, lane 4 contains the AB hairpin with the displacement strand (TB1) to switch the hairpin into the B conformation, lane 5 contains the sample in lane 4 with the addition of the displacement strand (TB2) to remove TB1, lane 6 contains the A control and lane 7 contains the B control. A 100 bp DNA ladder is used as reference (L). The samples in the 10% polyacrylamide gel were stained for SYBRGold.