FUNCTIONAL DNA NANOTECHNOLOGY

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



Florence Benn

Worcester College

University of Oxford

Trinity 2016

ABSTRACT

This thesis sets out to further the field of functional DNA nanotechnology through the design of novel functional DNA scaffolds, and investigates their applications and efficacy. The work presented here comprises two parts: The design of a chiral DNA nanotube that acts as a scaffold for motor motion and for an enzyme cascade; and the design of two different tetrahedral scaffolds for selection of a combination of three ligands, which together have a greater binding effect than the sum of the individual components.

It begins by proposing the design of a DNA origami nanotube which distinguishes between the inside and outside face of the tube at the design stage, which most previous designs reported do not. The previous designs in the literature result in a distribution of 50:50, of one face forming the inside surface on one tube and the same face forming the outside surface of a different tube. In the design presented in this thesis, this distinction results from making the tube chiral, which forces it to roll up in a predetermined manner. Chirality is introduced by varying the positions of staple crossovers and this process is explained. The chiral tubes may stack end-to-end to form long polymers, or exist in monomeric form with stacking suppressed, by inclusion of different sets of staples at the ends of the tubes. We confirm tube formation and right-handed chirality with AFM and CD respectively.

The efficacy of the tube as a scaffold for an enzyme cascade is tested and discussed in context of the wider field. No significant enhancement is observed when enzymes are tethered to the inside of the tubes, compared to when they are tethered to the outside or are free

in solution, although the same slight trend is always observed. Suggestions are made to better this experiment and further understand the underlying physics of such systems.

We propose using the tube as a scaffold for a DNA track, upon which a DNA motor may walk. DNA motors are introduced and we attempt to observe micron-scale, inter-tube motion within the confines of our origami tube. Initial experiments show the motor moving and we propose methods of fluorescent labelling via PAINT to better the experimental set-up for TIRF microscopy, which currently is limited by photobleaching.

The second part of this thesis proposes systems for selection of a combination of three ligands, which together have a greater binding effect than the sum of the individual components. Here we design two tetrahedral systems where either three ligands or three aptamers are brought together at a vertex of the tetrahedron to form a binding domain. The aptameric system allows for selection, amplification and reassembly of the strongest binders, because the functional and structural sequences are on one strand of DNA, following ligation. This design betters the initial tetrahedral system, where the coding/record strands for amplification are separate from the functional binding domain strands the ligands are attached to. This means it is not possible to reassemble this particular structure after amplification of the record strand.

ACKNOWLEDGMENTS

I am exceptionally thankful for the help and support I have had from my supervisor, Professor Andrew Turberfield. His initial faith towards me undertaking a DPhil is something I will always be grateful for. I have grown under his supervision and his patience with me over the years is admirable. It turns out he is *always* one step ahead, even when you thought he was not listening- now that is impressive! I would also like to thank Jon Bath as a colleague and good friend for all his help, collaboration and continual enthusiasm and level-headedness. He is a *fantastic* scientist, a wonderful human and a true kindred spirit - he has been a big part of making this possible and in shaping my time in DNA group.

Additionally, my undergraduate tutor Professor Steve Evans has provided fantastic support over many years and has always been great inspiration. I would have never got this far without his belief!

The environment in the lab in Oxford has been fantastic due to the people in DNA group. I am particularly grateful to Natalie Haley, Alex Lucas and Jon (again!) for their support during my year of illness. Also to Sam Tusk for keeping me on my toes with his sneaky lab tricks and wit. Thanks to Wenjing Meng and Robert Schreiber for being great sources of knowledge; Aiman Entwhistle, Adam Wollman, Carlos Sanchez, Helen Carstairs and Robert Machinek for providing a welcoming, cake- and tennis-filled environment when I started; Ibon Santiago Gonzalez- who started with me- for his wise, philosophical outlook on life; Celine Journot, Mike Boemo and Antonio Garcia Guerra for their fondue sets, comedy and endless energy; and to the newbies, Joel Spratt, Katie Young and Behnam Najafi for being lovely

people with a fresh outlook on the field. I wish you all the best for the future.

Thanks to EPSRC, BBSRC and Worcester College for keeping me afloat and fed during these years. The Stulz group in Southampton University have been useful collaborators for enzyme work in this thesis. Thanks to Dr Afaf El-Sagheer who provided us with click strands, and to Professor Lee Sweetlove who gave us interesting insights into enzyme cascades.

Outside of lab-life, the wonderful array of housemates from Rickyard Close over the years have all become great friends: Thanks go to them for providing excellent fun and many puzzle-filled evenings. Thanks to Andreas Kisling Harris who has assisted me in times of LaTeX-, Matlab- and Mathematica-crisis, which I am thankful for.

Thank you to all my family; my big bros for mocking me and loving me in equal parts; my uncles and aunts in Oxford for being 'just down the road'. Most importantly, I would like to thank my parents for their years of support, motivation, belief, and humour. They are the toughest and most interesting cookies I have come across, and are great examples of how a flexible approach to life is key.

CONTENTS

1 INTRODUCTION TO DNA AND DNA NANOTECHNOLOGY 1

1.1 DNA 1

1.1.1 The structure and properties of DNA 1

1.1.2 DNA as a building material 5

1.2 Beyond simple DNA nanostructures 12

1.3 A molecular biology toolkit 12

1.4 Observing DNA nanostructures 14

1.4.1 Scanning probe microscopy and the atomic force microscope 14

1.4.2 Fluorimetry 16

1.4.3 Total internal reflection fluorescence microscopy 17

1.4.4 DNA electrophoresis 18

1.5 Proposal for this thesis 20

2 DESIGN OF DNA NANOTUBES 21

2.1 DNA nanotubes 21

2.2 Design 26

2.2.1 Design of a 10-helix chiral tube 27

2.2.2 Design of a 20-helix chiral tube 35

2.3 Synthesis 37

2.4 Characterisation 38

2.4.1 Atomic Force Microscopy 38

2.4.2 Melting curve analysis 47

2.4.3 Circular dichroism 50

2.5 Conclusions 55

3 A ROLE FOR DNA NANOTUBES IN ENZYME CASCADES 57

3.1 Biological modulation: scaffolds and tubes 57

3.1.1 Organisation in cells 58
3.2 Mechanisms of channelling 60
3.2.1 Channelling: Perfect or leaky? 60
3.2.2 The mechanisms 60
3.2.3 Measuring channelling effects 62
3.3 Synthetic enzyme cascades 62
3.4 Modelling enzyme kinetics 67
3.5 Constructing a three-enzyme cascade 72
3.5.1 Biochemistry 73
3.5.2 Enzyme-DNA conjugation 74
3.5.3 Assembly of components 75
3.6 Characterisation 78
3.6.1 AFM 78
3.7 Results 81
3.7.1 Assays 81
3.8 Further work 86
3.9 Conclusions 88
4 MOTORS IN TUBES 89
4.1 Transport on the nanoscale 89
4.2 Aim 91
4.2.1 Track design 92
4.2.2 Track secondary structure 96
4.3 Fluorimetry; getting from end to end 97
4.3.1 Assembly of monomer and dimer tube tracks 97
4.3.2 Results and discussion 99
4.4 Microscopy; long range motion attempts 111
4.4.1 Tube alterations for imaging with microscopy 112
4.4.2 Preparing a flow channel for imaging 112
4.4.3 Optimisation of conditions for motor operation 113
4.4.4 Results and discussion 117

4.5 Summary 129

4.6 Future work 129

5 DNA TETRAHEDRA FOR SELECTION OF BINDERS 133

5.1 Introduction 133

5.1.1 Combinatorial assembly 133

5.1.2 Dynamic combinatorial chemistry 134

5.1.3 DNA-encoded combinatorial libraries 135

5.1.4 Click nucleic acid ligation 139

5.2 DNA-click tetrahedron for ligand selection 142

5.3 Progress towards a library 147

5.3.1 Results 147

5.3.2 Further work 149

5.4 Adapting the DNA tetrahedron for selection of co-operative aptamers 149

5.4.1 Introduction 149

5.4.2 The paranemic tetrahedron 156

5.4.3 Results 158

5.4.4 Further work 164

6 SUMMARY AND OUTLOOK 165

6.1 Closing remarks 168

A APPENDIX 171

A.1 Assembly protocols 171

A.1.1 From Chapter 2: Plasmid synthesis 171

A.1.2 From Chapter 5: Assembling clickable tetrahedra 173

A.1.3 From Chapter 5: Assembling paranemic structures 174

A.2 Methods 179

A.2.1 AFM 179

A.2.2 TIRF microscopy 179

A.2.3 Microscopy image analysis 181
A.2.4 Ensemble fluorimetric measurements 183
A.2.5 Circular dichroism of helical AuNP arrangements 183
A.2.6 TEM 184
A.3 Materials 185
A.3.1 Oligo synthesis and storage 185

BIBLIOGRAPHY 195

LIST OF FIGURES

Figure 1 The structure and properties of DNA. 4

Figure 2 Ned Seeman's DNA motifs. 6

Figure 3 The assembly of origami. 10

Figure 4 Examples of origami assemblies. 11

Figure 5 Tube designs from literature that do not differentiate between inside and outside faces of a tube in their design. 23

Figure 6 Building a scaffold strand for the 10-helix tube 29

Figure 7 Creating curvature in a 10- and 20-helix nanotube via staple asymmetry and moving crossover positions. 32

Figure 8 Crossover positions in a 10- and 20-helix nanotube 33

Figure 9 10-helix tube staple design and arrangememnt. 34

Figure 10 Controlled end-to-end stacking of tubes. 36

Figure 11 AFM of 10-helix tube. 40

Figure 12 AFM of 20-helix tubes 41

Figure 13 AFM of stacking tubes. 44

Figure 14 The distribution of tube lengths. 45

Figure 15 Summary of stack lengths 46

Figure 16 Melting curve analysis of the origami tubes. 49

Figure 17 AuNP positions for CD. 53

Figure 18 CD results. 54

Figure 19 Previous casacde assemblies with DNA origami scaffolds. 64

Figure 20 Hess' origins of activity enhancement in enzyme cascades on scaffolds. 71
Figure 21 3 enzyme cascade reaction. 74
Figure 22 Enzyme-DNA conjugation gel. 75
Figure 23 Enzyme positions for different tube cascade setups. 77
Figure 24 AFM images of enzymes inside the DNA nanotube. 79
Figure 25 Further AFM images of enzymes inside the DNA nanotube. 80
Figure 26 Enzyme cascade absorbance changes. 82
Figure 27 The burnt bridges motor. 91
Figure 28 The position of the track inside the chiral nanotube. 95
Figure 29 Schematics of a mid-track stator showing where the undesirable secondary structure arises. 96
Figure 30 Schematics of the monomer and dimer tube track with fluorophores at the beginning and end of the track. 99
Figure 31 Schematics of the fluorophores at beginning and end of track. 100
Figure 32 Comparison of motor motion along 5 and 26 stators. 102
Figure 33 Motor movement along a dimer tube. 105
Figure 34 A schematic of the FRET between the Cy3 on the motor and stators. 107
Figure 35 Depiction of the dimer track with a Cy3-labelled motor. 110
Figure 36 Enzyme cutting tests for the different buffer conditions. 116

Figure 37 Examples of typical fields of view of stacked tubes with Cy3-labelled tracks over multiple days. 121

Figure 38 The mean pixel brightness over time for the test sample and ocntrols. 122

Figure 39 Snapshots of a video with flashes that may indicate motor movement for the test sample. 123

Figure 40 Sequential snapshots of a video for negative control for observing direct motor movement in TIRF. 124

Figure 41 Sequential snapshots of a video for positive control for observing direct motor movement in TIRF. 125

Figure 42 Typical field of view for Cy3 labelled motor in TIRF. 128

Figure 43 Principles of the yoctoreactor. 137

Figure 44 Reddavide's dynamic combinatorial DNA library 138

Figure 45 Click chemistry reaction. 141

Figure 46 Diagrammatic representation a new 6-stranded DNA tetrahedron. 144

Figure 47 Experimental procedure showing combinatorial assembly process. 146

Figure 48 Stepwise assembly of the DNA tetrahedron. 147

Figure 49 Paranemic crossover motif. 152

Figure 50 Assembly process of paranemic tetrahedron. 155

Figure 51 Paranemic tetrahedron. 157

Figure 52 Paranemic tetrahedron assembly. 159

Figure 53 Possible arrangement of assembled paranemic tetrahedron. Gel shows assembly. 162

Figure 54 Testing the assembly efficiency of PX regions in a paranemic tetrahedron. 163
Figure 55 Method of conjugation, binding and selection for thrombin binding. 176
Figure 56 Aptamer complements gel 178

LIST OF TABLES

Table 1 Oligonucleotides for nanotubes 186
Table 2 Oligonucleotides for tetrahedra 193

1

INTRODUCTION TO DNA AND DNA NANOTECHNOLOGY

1.1 DNA

1.1.1 *The structure and properties of DNA*

“It has not escaped our notice" wrote James Watson and Francis Crick in their seminal paper, that the double helix structure “immediately suggests a copying mechanism for the genetic material." This elegant spiral, first drawn by Crick’s wife Odile, depicted what would become life’s most famous molecule [1, 2].

Watson and Crick worked together on studying the structure of DNA (deoxyribonucleic acid), the molecule that contains hereditary information for all of life’s cells. At that time Maurice Wilkins and Rosalind Franklin, both working at King’s College, London, were using X-ray diffraction to study DNA, and Crick and Watson used the pair’s findings in their own research [3].

In April 1953, they published the news of their discovery, for which they would later win the 1962 Nobel Prize in Medicine: the molecular structure of DNA was determined based on all its known features - the double helix. Their model set the stage for the rapid advances in molecular biology that continue to this day and marked the beginning of a new era of life sciences [4].

Today, DNA is a well-characterised polymeric molecule. It occurs as a single- or double-stranded molecule in nature (ss, ds respectively).

Our understanding of DNA as a genetic code has allowed progress in genomics and gene therapy. Its unique properties have also played a critical role in connecting molecular biology to materials science and engineering [5].

DNA is made of four types of nucleotides, which are linked covalently into a polynucleotide chain (a DNA strand). DNA nucleotides are composed of a five-carbon deoxyribose sugar to which a single phosphate group is attached [6]. The sugar-phosphate motif repeats along the length of the DNA strand forming a sugar-phosphate 'backbone'. The phosphate groups are negatively charged and give the backbone and the DNA an overall negative charge. The bases extend from the sugar and may be either adenine (A), cytosine (C), guanine (G), or thymine (T). DNA strands are inherently polar and the DNA molecule comprises two complementary, anti-parallel strands which are held together by hydrogen bonds [6]. Polarity is indicated either by a 5′ phosphate group or a 3′ hydroxyl group (named due to the location of the group on either the 5′ or 3′ carbon of the sugar ring) [7]. An N-glycosidic bond links the base to a carbon on the deoxyribose sugar [8]. The angle that this bond makes in relation to the base-base interface gives rise to one face of the base-pair being larger than the other face. This assymmetry creates a major and minor groove along the double helix structure.

So called 'base-pairs' arise between A and T (which form two hydrogen bonds) and G and C (which form three hydrogen bonds). Both A and G are purine molecules and comprise two aromatic rings, whereas T and C, the pyrimidines, have just one aromatic ring. The bulkier purines pair with the more compact pyrimidines, meaning that base pairs are of similar widths and the phosphate backbones are separated by the same distance along the length of dsDNA [6]. The bases are depicted in Figure 1 b. Neighbouring bases in the helix

are involved in non-covalent pi-pi orbital stacking interactions which arise between pi-orbitals in adjacent aromatic rings and play a major role in stabilising the helix [9].

DNA may be mechanically rigid or flexible. For naturally occurring, ds 'B-form' DNA, which wraps around itself forming a plectonemic coil, a persistence length of 50 nm is typical, which roughly corresponds to 150 base pairs or bps, whereas for ssDNA the persistence length is just a few nanometers, corresponding to a few bases [10, 11, 12]. A double helix spans approximately 2 nm in diameter, and each helical turn comprises 10.5 bps with a length of 3.6 nm (see Figure 1 a), which depends on the salt concentration of the solution and the sequence [5, 13]. B-form DNA has an intrinsic right-handed (RH) chirality, but 'Z-form' DNA also exists, which adopts a left-handed (LH) conformation. Z-form DNA is assembled from regions of alternating purines and pyrimidines such as GCGCGCGC. In this thesis we are primarily concerned with B-form DNA [14, 15].

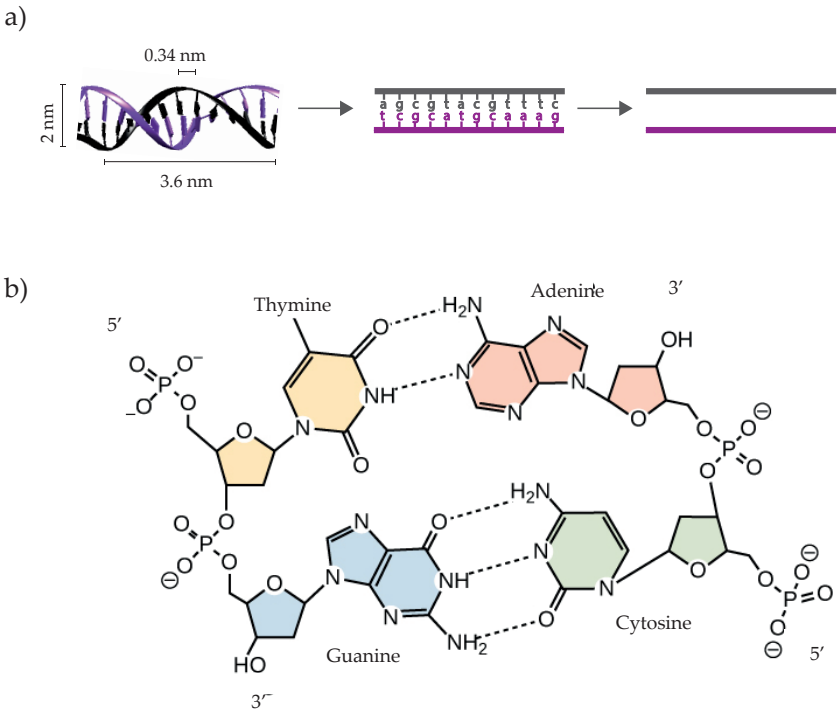


Figure 1: The structure and properties of DNA. a) DNA has a pitch of 0.34 nm per base and of 1.36 nm per helical turn. The width of a helix is roughly 2 nm. The DNA helix can be depicted as two lines with bases, or simply, as two parallel lines, but the three depictions are equivalent. b) Adapted from *http://devarchive.cnx.org/* [16] The bases of DNA are adenine, guanine, cytosine and thymine and comprise one or two aromatic rings depending on whether the base is a purine or pyrimidine.

1.1.2 *DNA as a building material*

In 1982 the field of biotechnology was focussed on improving areas such as plant genetics, microbial physiology and biochemical engineering of enzymes; areas for which the ability to manipulate the genetic DNA code was imperative [17]. Concurrently, Nadrian Seeman was imagining using DNA as a structural building material and envisioned working with it outside of its traditional biological context as a genetic material [18]. His idea, inspired by the Holliday junction found in molecular recombination, was to create mobile and immobile junctions of three or four DNA strands where one strand would bind to two other strands, as depicted in Figure 2 [19]. Duplexes radiate out in a cross shape about a central point. The strands would be synthetic and would require sequence design to prevent self-complementarity and to ensure that the junction formation is that with the lowest free energy. In the case of immobile junctions, branch migration seen in Holliday junctions would be prevented [19]. Seeman furthered his original idea by suggesting such motifs could come together to form large assemblies where a strand in each duplex of a junction has a ss-domain which is complementary to another (known as 'sticky ends'); 1998 saw the first robust experimental example of such an array [20].

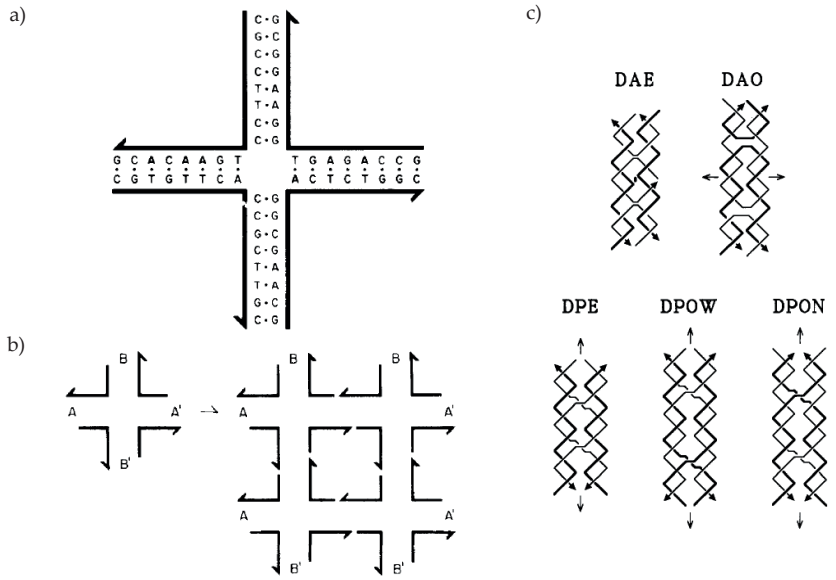


Figure 2: Ned Seeman's DNA motifs. Figures a and b are adapted from 'Nucleic acid junctions and lattices', N C Seeman, *Journal of theoretical biology*, **99**, 1982. Figure c is adapted from 'DNA double-crossover molecules', T. J. Fu, N. C. Seeman, *Biochemistry*, 1993, 32 (13), pp 3211-3220. a) An immobile nucleic acid junction b) Formation of a two-dimensional lattice from an immobile junction with sticky ends. A is a sticky end, and A' is complementary to it. A similar relationship exists between B and B'. Four of the monomeric junctions on the left are complexed in parallel orientation to yield the structure on the right. If the inter-junctional spacing is large enough, a ligase would be able to close the overlapped gaps to make the complex on the right a covalently bonded structure. c) Schematic drawings of the five different structural arrangements of double-crossover structures. The structures shown are named by the acronyms describing their basic characteristics. All names begin with D for double crossover. The second character refers to the relative orientation of their two double-helical domains, A for antiparallel and P for parallel. The third character refers to the number of helical half-turns between crossovers, E for an even number and O for an odd number. A fourth character is needed to indicate whether the crossover occurs at the major groove, denoted as the wide 'W' region, or the minor groove, denoted as the narrow 'N' region.

In addition to Holliday junctions, Seeman and Fu worked to create other synthetic analogues of small molecules involved in the molecular biology of recombination. These molecules are known as double-crossovers and involve two DNA links between helical domains, where a DNA strand crosses over to an adjacent helix [21]. The distance between crossover points must be specified as multiples of half-turns, in order to avoid torsional stress in these systems. Double crossover molecules (D) with parallel (P) or anti-parallel (A) configurations were created, as well as their respective subsets with even (E) and odd (O) numbers of half-helical turns between crossovers, and are all depicted in Figure 2 c. One example by Mitchell *et al.* demonstrates a system of DNA tiles which is designed to assemble 2D arrays [22]. However, the arrangement also gives rise to nanotubes, where sheets of tiles roll up on themselves. The tiles used here are 1D double, anti-parallel, even-odd (DAE-O) double-crossover motifs. As with Seeman's lattices, double crossover tiles with four single stranded sticky ends on each one were arranged such that two types of tile tessellate.

One of the main highlights of Seeman's DNA lattice work is the ease with which the structures form; they do so via self-assembly. Self-assembly provides an elegant method of building nanoscale structures and devices. This approach, which is employed throughout nature to form complex biological systems, exploits compatibility between distinct molecular building blocks such that it is energetically favourable for them to assemble into a defined structure. Consequently this method has the advantage of being more robust than top-down fabrication where the target structures are not necessarily thermodynamically favoured. Self-assembly also permits structures to be modified with atomic precision which can be difficult to achieve using top-down techniques. As well as drawing on biology for inspiration,

another advantage is that biological components are easily incorporated into bottom-up systems comprising synthetic components.

DNA lends itself to applications as a nanoscale building material for self-assembling structures for a number of reasons. As touched on above, the fact that the four bases pair with one another with a high degree of specificity make it possible to accurately design structures by programming the complementarity of component strands. Secondly, there exists a multitude of techniques from biochemistry already in practice with which to both manipulate and study DNA. Finally, the primary function of DNA as a tool for genetic storage can be used to add information content to nanostructures, for example, including instructions for functionality or labels for identification and replication.

Following Seeman's pioneering work, the field gained momentum in the 1990s along with innovations in chemical synthesis techniques of synthetic oligonucleotides [23]. In 2006 Paul Rothemund introduced his DNA origami technique which makes larger structures possible [24]. Instead of having many short strands coming together to base-pair and form a structure, DNA origami makes use of a several-kilobase-long strand (for example, viral genomes such as M13mp18, naturally occurring plasmids and cloned plasmids like pUC19 have all been used) known as a scaffold strand, to shape the assembly. To design an origami structure, one imagines first routing the scaffold strand to raster-fill a predetermined shape, for example a rectangle. In practice, the filamentous DNA is held in the pre-determined shape by numerous short synthetic oligos which base-pair with the scaffold and are known as staple strands. The sequence of the scaffold strand is already known (due to sequencing or published results), thus the sequence of the staple strands can be programmed accordingly to complement two or three distinct regions per staple on the scaffold

strand and therefore hold those regions in proximity to each other. Every staple holds different, unique regions of the scaffold together and so enables the initial shape imagined to be raster-filled with a double helix. By altering the sequences of the set of staple strands, different scaffold regions can be brought into proximity to raster-fill and assemble a new pattern- for example a circle, instead of a rectangle. We refer to the place where a staple base-pairs with one region of a scaffold and then crosses over to bind with another region of the scaffold, a staple crossover position. Staples may be different shapes but are commonly 'S'-shaped with three binding regions and two staple crossovers per staple, or 'U'-shaped with two binding regions and one staple crossover between them. The staple crossovers are points where DNA weaves between sets of parallel helices and so by determining when these jumps happen, one can control how the parallel helices are oriented in relation to one another and whether helices lie flat in a plane or whether each helix is at an angle to the last.

To assemble origamis, the short staple strands are mixed with, and in excess to, the scaffold strand and annealed together. Through design, the target structure is the most stable configuration and forms co-operatively via the annealing process [25]. Figure 3 describes the origami process further with an example of forming an origami tube. Subsequently, a wide variety of structures have since been created, including both one-, two- and three-dimensional shapes and objects, demonstrating proof-of-principle. A few such assemblies can be seen in Figure 4.

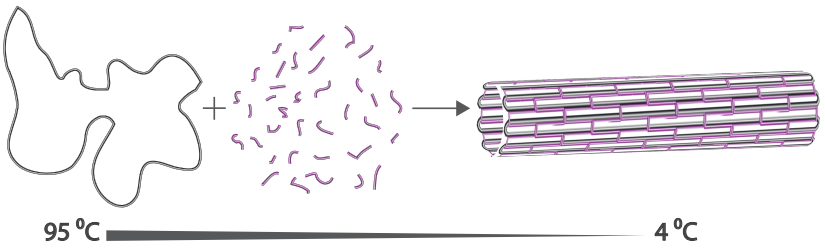


Figure 3: The assembly of origami. To start with, a target shape for the intended application is conceived, for example an origami tube. Next a scaffold-staple layout for the target shape is imagined and the appropriate set of staple sequences to build the tube are designed. The scaffold strand is depicted as black and may be circular or non-circular. The small purple strands are the staples and their path can be seen on the far right. Self-assembly reactions are prepared and subjected to a thermal annealing procedure. The staples are usually in at least 3 x excess to the scaffold strand during assembly. Heating denatures any secondary DNA structure and cooling encourages assembly of our designed shape.

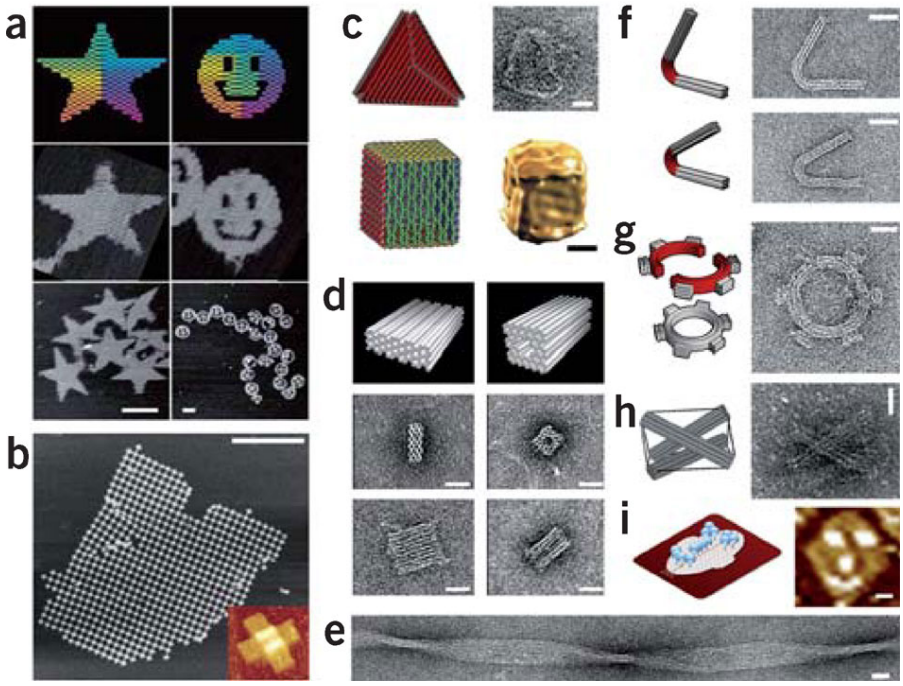


Figure 4: Image adapted from "A primer to scaffolded DNA origami", C. E. Castro *et al.*, Nature Methods **8**, 221-229 (2011) (a) Designs of single-layer DNA origami shapes (top) and AFM images of these objects (middle and bottom). The pointed star and the smiley face each have diameters of roughly 100 nm. Reprinted from [24]. (b) AFM image of crystalline DNA origami arrays formed from several hundred copies of a cross-shaped single-layer DNA origami object. Inset, image of a 100-nm-long cross-shaped origami monomer. Reprinted from [26]. (c) Container-like DNA origami objects (left) imaged with negative-stain TEM (top) and cryogenic TEM (bottom); reprinted from [27] (top) and [28] (bottom). (d) Design and images of multilayer DNA origami objects [29]. (e) Image of a multimeric multilayer DNA origami object with global twist deformation [30]. (f,g) Design and images of space-filling multilayer DNA origami objects such as bent bars (f) and a gear with square teeth (g) displaying custom curvature [30]. (h) Tensegrity prism created by combining multilayer DNA origami struts and ssDNA strings; reprinted from [31]. (i) Design and image of a single-layer DNA origami shape with site-directed protein attachments; reprinted from [32].

1.2 BEYOND SIMPLE DNA NANOSTRUCTURES

Advancing the field from just an academic discipline, the most recent and exciting innovations have been functional systems [33]. These include nanoscale polyhedra that function as molecular cages and delivery systems [22]; DNA-based biomimetic walkers that navigate tracks in response to DNA fuel [34]; the incorporation of functional nucleic acids (FNAs) into structures, such as aptamers and DNAzymes, to recognise specific targets [35]; and DNA-origami membrane channels that function as single molecule detectors [36], to name just a few. This list, and this discussion, is not (and does not try to be) exhaustive and so continues throughout this thesis, as and when relevant assemblies and technologies arise: in Chapter 4 we take a closer look at DNA motors as we endeavour to observe motor motion in a novel environment; in Section 5.4 of Chapter 5 we work towards incorporating aptamers into a 3D DNA tetrahedron for selection experiments; Chapter 2 focusses on creating a new DNA nanotube where we also discuss how 3D origamis are formed, and the thesis goes on to explore such a structure's potential uses.

1.3 A MOLECULAR BIOLOGY TOOLKIT

We have touched on the extent of the field of DNA nanotechnology, which has grown in breadth and depth, in inventiveness and functionality, over the past few decades. But such assemblies and explorations are only made possible by to the tools provided by molecular biology [5]. Key to creating DNA nanostructures is being able to dictate exactly which bits of DNA go where; to decide what is bound to what and where cuts should be made; being able to create multiple copies

of our design; to determine where functionalities should be placed, and so on.

When it comes to performing such modifications in the cell, the protagonist is the enzyme. Enzymes are a class of protein which are responsible for determining most of the chemical transformations that make and break covalent bonds in cells [6] [1]. They bind to a wide variety of ligands and catalyse only a very specific reaction, often working in teams to accomplish multiple alterations in quick succession. DNA nanotechnology employs a number of enzymes to realize specific molecular changes.

Endonucleases are the 'scissors' of the enzyme world, cutting the phosphodiester backbone of a polynucleotide. Restriction endonucleases cut at specific sites, called restriction sites, which are often palindromic sequences of roughly four to eight bases. By incorporating such sites into our designs we can create recognition sites for cutting. Such cuts often give rise to a specific sticky-end post-cut, which when applied to two or more DNA helices, creates 'recombinant' DNA; homologous puzzle-piece pairs to be stuck together later. Endonucleases allow the scaffold strand to be made longer, in Chapter 2.

In contrast, exonucleases work from either end of a polynucleotide chain to remove single nucleotides, step by step. In this thesis we use Exonuclease III which acts in the 3′ to 5′ direction, to digest strands we no longer require.

Ligase is used throughout this thesis as a DNA glue- sealing strands together where they have complementary sticky ends [5].

DNA polymerases allow for multiple copies of our strands to be made via replication. In the polymerase chain reaction (PCR) polymerases use one strand of DNA as a template upon which to assem-

1 However, the active parts of the ribosome is composed of catalytic ribose nucleic acids (RNA) and there are other ribozymes too, which are involved in splicing a special type of RNA (messanger RNA). Therefore RNA too might play a role in the future of nucleic acid nanotechnology

ble a complementary version. For such a reaction to occur, primers are introduced which are strands complementary to a small region of the template and provide a starting point for polymerases to assemble individual nucleotides from the hydroxyl group at the 3' end [6].

1.4 OBSERVING DNA NANOSTRUCTURES

1.4.1 *Scanning probe microscopy and the atomic force microscope*

Perhaps the most powerful observation technique in DNA nanotechnology is one that allows us to 'see' nano-assemblies. Atomic force microscopy (AFM) provides such a tool, which has added bonuses of high-resolution and versatility, in imaging modes, sample condition, substrate and in the information we obtain from it.

The field of scanning probe microscopy (SPM) emerged in 1982 with the advent of the scanning tunnelling microscope (STM) [37]. In STM, a sharp tip is raster-scanned over a surface and its distance is controlled using the tunnelling current from tip to surface as a control signal. Experiments are usually performed in vacuum and with low temperature requirements, which preclude most applications in biology [38].

In 1986 however, the invention of the AFM meant that SPM techniques could be applied to biomolecules too [39]. The inventors of the STM, Binnig and Gerber, applied the raster-scan technology of the STM to AFM operation, where a sharp probe detects surface topography by directly measuring forces between probe and surface [38]. Such technology does not require a conductive substrate and can operate under ambient conditions. The AFM also can be used to garner further information about the sample: it is capable of eval-

uating morphology, mechanical property, electrostatic charge, and even the chemistry of the molecular species by various AFM imaging modes [23]. Such modes may be employed in real-time with high-speed AFM, in air or in liquid.

The first mode available was that of contact mode. In contact mode the sample and cantilever are in constant contact during the scanning process which despite excellent resolution, also gives rise to large lateral forces which may be damaging. 1987 saw the introduction of a much more biologically-friendly non-contact mode in which a constant distance between sample and probe is maintained via a feedback loop regulated by laser deflection and reflection from the back surface of a canitlever, directed onto a quadrant photodiode [40]. Bettering this still, the dynamic modes of intermittent contact and tapping mode imaging arrived on the scence in 1993 [41]. The cantilever is excited by a piezoelectric crystal and oscillates a few nanometers above the sample surface. Interactions between the cantilever and the substrate are detected as changes in the oscillation amplitude and phase. Tapping mode combines the advantages of contact and non-contact imaging. Because cantilever contact is only intermittent, tip-limited resolution is maintained while ensuring low lateral forces are imposed on the sample. Forces perpendicular to the sample however are somewhat high due to capillary forces induced by a water layer on the sample surface in ambient conditions [38]. To prevent such destructive forces biological samples are often imaged in liquid, a technique first demonstrated in 1992 for contact mode and 1994 for tapping mode [42, 43]. This is advantageous because biological assemblies can be imaged in their native environment.

One problem encounted using AFM is imaging 3D structures. Although many 3D DNA nanostructures have been successfully observed using AFM (including real-time motion of DNA walkers [44]),

the strong forces that exist between DNA and mica can often flatten 3D structures, thus not allowing their topology to be observed [45, 46]. Tip convolution is another problem which may be encountered. Due to the non-zero width of an AFM tip (an ideal tip would be a delta function), one may experience tip convolution which limit the lateral resolution of non-planar structures. In some instances a feature on the surface may be smaller than the average width of the tip and when this condition is met, the feature images the tip instead of the tip imaging the feature- the topography displayed is that of the tip and not the feature. Therefore, for some assemblies a better choice for imaging may be electron microscopy (EM). This is particularly the case where assemblies are functionalised with metal components, ie. metal nanoparticles, due to the large contrast between metal and DNA in EM. EM is mentioned briefly in Appendix A.2.6 as it rarely features in this thesis and any images included are taken by collaborators.

1.4.2 *Fluorimetry*

Fluorescence resonance energy transfer, or FRET, describes an energy transfer mechanism between two molecules - which may be two fluororphores, or a fluorophore and a quencher - where energy is transferred from a donor molecule to an acceptor molecule. The donor is excited at its excitation wavelength and transmits the energy via a non-radiative process to the acceptor, whose excitation spectra overlaps with the emission spectra of the donor. The donor then returns to its ground state without emitting a photon [47]. If the acceptor molecule is a fluorophore then upon receiving energy via non-radiative transfer, the fluorophore may emit a photon at its characteristic wavelength. However, if the acceptor is a dark quencher molecule

the received energy may be dissipated as heat to its surroundings. The donor and acceptor molecules are known as FRET pairs. FRET efficiency decreases with the sixth power of inverse distance, due to dipole-dipole coupling, making FRET possible only when molecules are very close together [48]. FRET is sensitive to distances shorter than 10 nm and thus is applicable to our nanodevices, as will be seen in Chapter 4.

When monitoring controlled motion, it is desirable to be able to track a device's progress in real time and FRET provides a simple mechanism with which to do so. We can exploit such energy transfer relations when monitoring motion of DNA nanodevices which are able to move, by attaching quenchers and fluorophores to our moving devices and stationary components. Monitoring the changes in fluorescent signals in real time with a fluorimeter tells us about the relative positions of DNA components and can be translated into an understanding of how our devices move.

Such measurements are bulk, ensemble in nature. This means that the average behaviour of a sample is probed, where the measured intensity of fluorescence represents the sum of many individual fluorophores. Consequently we assume there is little divergence from the average behaviour when such experiments are conducted and that infrequent events can safely be ignored.

1.4.3 *Total internal reflection fluorescence microscopy*

Due to the ease with which DNA may be functionalised, fluorophores can be easily appended and multiple labelled strands may decorate an origami sample. Intercalating dyes which slot in between bases may also label our nanostructures. Both give rise to localised fluorescence which may be observed via fluorescence microscopy. A par-

ticularly useful version of fluorescence microscopy is total internal reflection fluorescence microscopy (TIRF). TIRF microscopy utilises a laser beam that is obliquely incident upon the substrate liquid interface (contained within a channel between glass slides) at an angle greater than the critical angle of refraction [49]. The beam is reflected entirely at the interface due the difference in refractive indeces between the glass and the aqueous buffer but an electromagnetic field known as an 'evanescent wave' penetrates into the sample channel. Put in a very simplified way, the evanescent wave has an intensity that decays exponentially with perpendicular distance from the surface [49]. This means that only a thin layer of a sample in the channel is illuminated adjacent to the glass slide surface, which has the benefit of high signal-to-noise ratio. Nanostructures can be tethered to this surface to ensure they remain in the TIRF field during imaging.

1.4.4 *DNA electrophoresis*

DNA electrophoresis is a method frequently used in this thesis to identify, purify and quantify nucleic acids and their assemblies. The technique involves passing DNA samples through a gel matrix by application of an electric field. The DNA migrates through the gel away from the anode and towards a cathode due to its inherent negative charge. Two gel types are used here: either agarose or polyacrylamide.

Agarose is an alternating copolymer which upon cooling from high temperatures adopts helical fibre bundles held in place by hydrogen bonds and junction zones. Upon application of an electric field, gel fibres and fibre bundles may re-orient in the electric field [50]. On the other hand, polyacrylamide gels are chemically cross-linked and formed by a reaction of acrylamide with a bifunctional crosslink-

ing agent such as N,N-methylenebisacrylamide (Bis) [51]. Polyacrylamide gels are polydisperse in structure. The composition of gels in this thesis is given in terms of a percentage; the total (w/v) concentration of agarose or acrylamide. For both, lower percentage gels are better for resolving very high molecular weight molecules, while much higher percentages are needed to resolve smaller ones. The median pore radius of a 1% agarose gel is roughly 100 nm [51]. Much smaller pore sizes in polyacrylamide mean that differentiation between a single base in a 1 kb molecule is possible [52]. However, it is not only pore size that affects the mobility of DNA through a gel. Electrostatic interactions between the sample and gel fibres may also affect motion of DNA, which is ultimately a molecular-mass dependent effect [51]. Other affecting factors include voltage, friction between migrating molecules and gel matrix, and the ionic strength of the buffer used [53] and it is a combination of all effects that determine the mobility of a nucleic acid or nanostructure through a gel. For example, a higher ionic strength of the buffer reduces the effective charge of the migrating DNA and may contribute towards a slower rate of migration, and vice versa.

Thus depending on our assembly's size, charge, conformation and functionalities, different structures will migrate through a gel at different speeds. We use this to our advantage to monitor step-wise assembly and to see whether stoichiometry is correct. Different sized structures can be purified by this method too, where for example staple strands migrate much more quickly than assembled origamis, due to molecular sieving effects [54].

1.5 PROPOSAL FOR THIS THESIS

The theme of this thesis is functional DNA nanotechnology. In Chapter 2 the design of a helical nanotube is presented which enables the tubes to fold up in a pre-determined manner. Such a design is used in Chapters 3 and 4 to provide a scaffold for functionality. In Chapter 3 enzymes are attached to the insides of tubes and their potential as a scaffold for a cascade is analysed in the context of the wider field. In Chapter 4 we explore the motion of a DNA motor along a DNA track which is built into the interior surface of the tube. We use both ensemble measurements and direct observation techniques via microscopy. Chapter 5 differs from the previous chapters in that it uses DNA tetrahedra instead of nanotubes as the scaffold for functionality. Here we set the stage for two different selection experiment systems for either strong binding chemical fragments or aptamers. Both systems aim to assemble an amplifiable and readable record DNA strand.

2

DESIGN OF DNA NANOTUBES

In this chapter the design of a helical nanotube is presented which enables the tubes to fold up in a pre-determined manner.

2.1 DNA NANOTUBES

A common goal of DNA nanotechnology and synthetic biology is to create a tool box of nano-components that can be drawn upon for a diversity of roles ranging from biophysical studies to nanofabrication. Nanotubes present themselves as being a vital component of this and numerous tube designs have been reported, each towards a different end. Tubes are desirable for many reasons; for use as containers [55], as addressable platforms [56], as templates for growth of functional devices [57], as well as simply mimicking biological filaments such as actin or microtubules. The examples here make use of a combination of the advantages granted by nanotubes made from DNA, including a high aspect ratio, ease of size variability and addressability.

Different approaches have been used to build DNA nanotubes. The original method forms tubes by connecting the edges of planar tiles. The tiles used in these instances are known as DX tiles [22, 58]. The DX motif was described in detail in Chapter 1 section 1.1.2 and can be seen in Figure 5 ai. When such motifs have sticky ends and are made from an odd number of half turns along their length, they can assemble into large 2D arrays which may roll into 3D tubes (where

the free energy gained from base-stacking is enough to overcome the activation energy of cyclisation, the latter being mechanical stress),

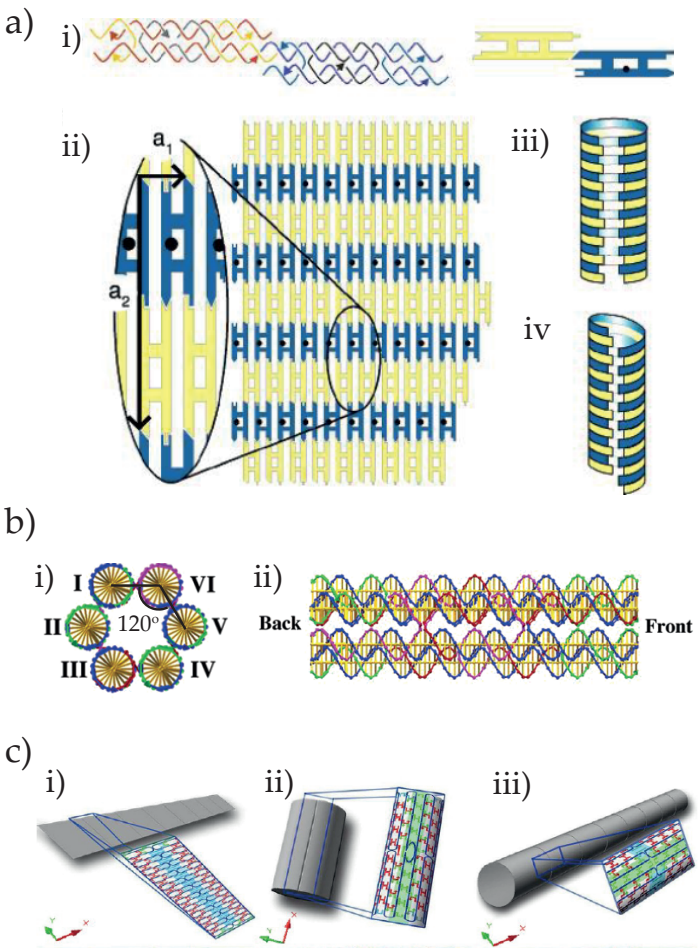


Figure 5: Tube designs from literature that do not differentiate between inside and outside faces of a tube in their design. **ai-iv)** Adapted from Mitchell *et al.* 2004 [22]. Self-assembly of DNA tiles into sheets and tubes. ai) Structure of the double crossover tiles: arrowheads mark the 3′ end of each oligonucleotide. The 6 nt single-stranded sticky ends on the yellow tile are complementary to those on the blue tile; complementary shapes on the schematic representations of the tiles indicate complementary sticky ends. A 5′ biotin label is represented by a black dot. a ii) The tiles tessellate to form extended two-dimensional arrays. The sheets can fold and close upon themselves to form tubes, producing either alternating rings (a iii) or nested helices (a iv) of blue and yellow tiles. **b i-ii)** Adapted from Mathieu *et al.* 2005 [59]. b i) A cross-sectional geometrical view of a six helix bundle in which the helices are indicated by Roman numerals. b ii) A geometrical side view of a six helix bundle in which the crossovers are separated by seven nucleotide pairs. The back and front are indicated as the ends to be used in self-assembly. Figures **ci-iii)** are adapted from Fu *et al.* 2013 [60]. ci) DNA origami nanoribbons, cii) Short DNA origami nanotubes, ciii) Long DNA origami nanotubes.

or form tubes with controlled diameters through sequence design [61]. Figure 5 a shows an example of a chiral nanotube formed from the DX tile motif, with sticky ends. There is no distinction however between which face of the origami forms the inside of the tube and which face will form the outside, instead a 50:50 distribution is produced.

Various designs of tubes make use of many single oligos as opposed to tile assemblies or the origami technique. One design creates a modular construction of rings with bracing strands between subsequent rings, to form ladder-like tubes [62]. The cross-section is either triangular or square depending on whether three or four helices are braced together parallel to the length of the tube. Similarly, Lo *et al.* have assembled a 'nanopeapod' with longitudinal variation in capsule size along its length. They also demonstrate size-selective encapsulation of gold nanoparticles inside the 'peadpod' followed by triggered release [55]. Another assembly of single oligos features a 'graphene-like' honeycomb sheet that folds into a tube, with the advantage that each hexagon is addressable [63].

Numerous DNA tube designs feature DNA helices collected in a bundle around a central axis. A six-helix bundle presented by Mathieu *et al.* seen from the side in Figure 5 bii, paves the way for many n-helix bundle designs [59]. In this design, three DX tiles are assembled adjacently, such that crossovers between helices arise every 1.5 helical turns, or 16 bps. To assemble a bundle of six helices into a hexagonal arrangement, each plane between pairs of helices must be rotated 120° with respect to the previous plane, as illustrated in Figure 5 bi. This is easily achieved in DNA with 10.5 bps per turn where seven bps corresponds to 2/3 of a turn (an exterior angle of 240°). Thus, an offset of 7 bps between adjacent crossovers establishes an interior angle of 120° between helices. Mathieu assemblies two lots of

three DX tiles with 120° between helices and binds them together to form the tube from these two halves.

Shawn Douglas advances the idea of designing origami structures using a hexagonal starting motif and simultaneously presents an open-source software package which aids design, including a pattern for a square starting motif [29]. The caDNAno package allows the authors to provide novel examples of 3D structures formed from honeycomb-pleated sheets of helices or square-pleated helices [64]. This has facilitated the construction of helical bundles for a multitude of functions, including origami tubes that serve as scaffolds for nano-plasmonic structures and a DNA rod which promotes alignment of membrane proteins in nuclear magnetic resonance structure determination [65, 66].

In addition to the bundle designs, other origami methods of tube formation have arisen. One example allows for combinatorial assembly of sections of a tube to form an enzyme cascade. This method forms a large origami ribbon (see Figure 5 ci) which can be connected along opposite edges to form a tube of circumference between 20 and 35 nm (Figure 5). Despite its merits, including the ability for sub-section re-arrangement, as in Figure 5, the tube does not provide a means to distinguish which surface of the origami will form the inside and which surface will form the outside of a tube. In order to side-step this dilemma, the authors functionalise both sides, leading to a distribution of half of the enzymes being on the inside and half outside [60]. Masayuki Endo *et al.*, form both right- and left-handed (RH and LH respectively) helical tubes from origami, but similarly cannot predict the outcome of assemblies or indeed which surface is the inner or outer one. They discover that their longer tubes form RH structures whilst their shorter ones assemble with LH chirality and

explain this in terms of bending stress due to helical stiffness and non-natural helical pitch [67].

We see from these examples that many of the tube designs previously presented (excluding the origami bundles) do not distinguish in their design between the inside face and the outside face of the tube, and none of the chiral tubes presented achieve this distinction either. Instead, a distribution of assemblies is created. This is not an efficient method of assembly if one requires a tube of a certain chirality with distinct inside and outside faces: a subsequent stage of separating the two families would be necessary as well as wasteful. Such assemblies that make this distinction might be required for numerous purposes including waterproofing tubes on one side, lining the inside of tubes, encapsulation, for uses involving the absorption of different chiral components of light, for creating a chiral path along a tube for a molecular motor, and so on.

2.2 DESIGN

When designing DNA tubes via the origami method we see from the previous section that there are two pre-existing options: one may first design an origami rectangle and connect two opposite edges to form a seam, in a two-step process, or design a bundle such that each helix is indistinguishable from the next. Here we present designs for both 10- and 20-helix chiral bundles which obey the latter strategy, where no seam is formed. However, where previous designs have not distinguished between the inside and outside tube faces, here we aim to do just that to produce a readily functional tube. This is done by introducing an intrinsic chirality into the origami.

Sometimes it is convenient to view 3D designs as flat and in two dimensions - particularly for design using caDNAno - and in these

cases the staples connecting the 'top' and 'bottom' of the projections are no different to staples connecting helices in the 'middle' of the 2D view. The 2D view may be termed a 'sheet' of origami for ease of description, although never during assembly is a planar structure formed before assembling into a tube; rather the tube assembles in one step.

In order to have prior knowledge of which side of an origami sheet becomes the interior of a tube, it is necessary to force it through design to assemble one way or the other. Traditionally to form a flat plane of origami, with a dihedral angle of 180° between adjacent helices, the staple crossover positions occur roughly every 1.5 helical turns, or 16 bases, using a square lattice caDNAno starting motif [24]. Work by a previous group member, Shelley Wickham, reduces the RH global twist expected to arise from the original Rothemund tile [44]. The global twist results from a slight under-winding of the helix at 10.7 bp/turn, but by deleting a few bps evenly spread across the structure, a more favourable 10.5 bp/turn is introduced [30, 13].

In our design we aim to create a global curvature such that it is energetically favourable for the origami to assemble with a RH helicity (as opposed to a distribution of RH and LH chiral tubes), and thus for the inside/outside surfaces to be distinguishable from one another both in the design stage and following formation.

2.2.1 *Design of a 10-helix chiral tube*

The first tube design is an origami forming 10 helices, with a length of 104 nm and a roughly 7 nm diameter from the centre of one helix to another. The resulting dihedral angle between two planes composed of three helices is 144°. The scaffold strand comprises a modified pUC19 plasmid with 336 bases inserted from a region of M13mp18

bacteriophage. The formation process is outlined in Figure 6 and described in greater detail in Appendix A.1.1.

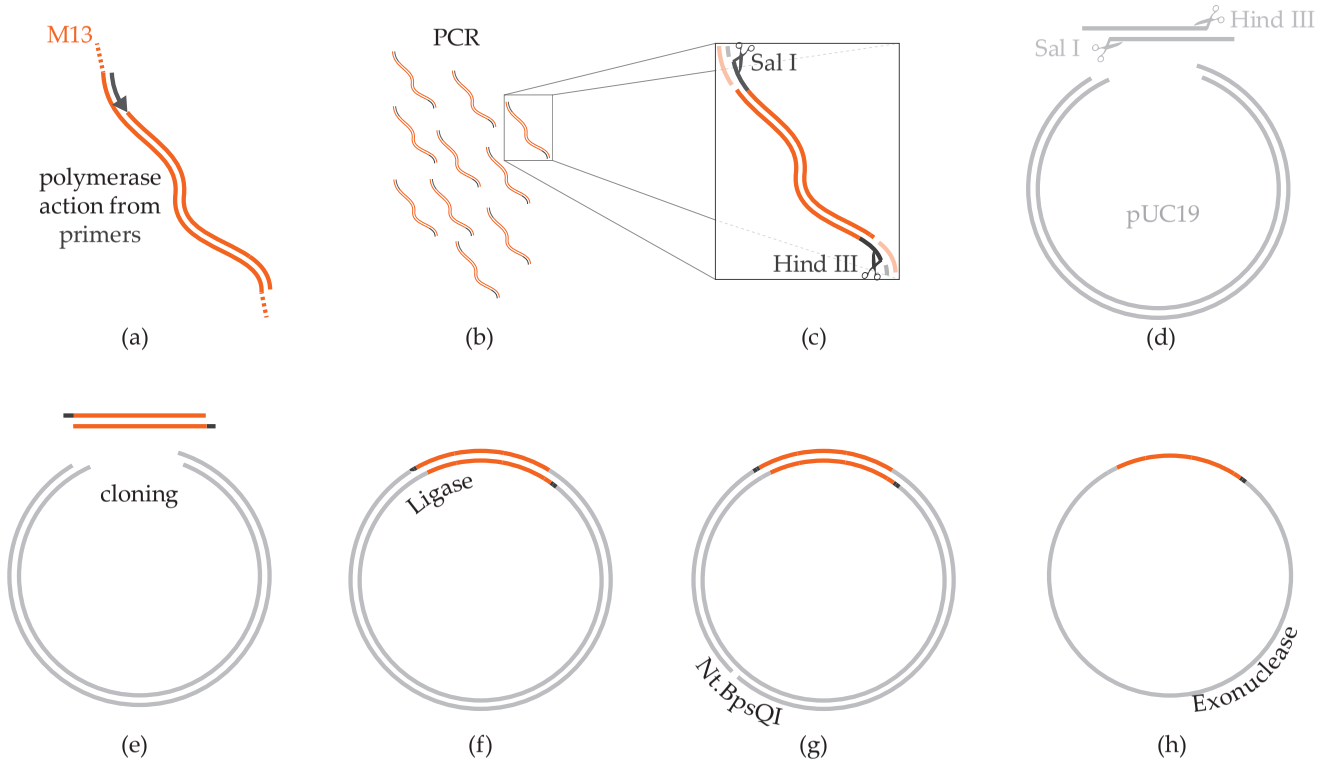


Figure 6: The steps involved in building a scaffold strand for the 10-helix tube. (a) Primers are designed to bind to either complementary end of a 348 bp region of M13mp18 for amplification (b) Polymerase action copies the region in 28 rounds of PCR. (c) The primers have Sal I and Hind III restriction sites to allow creation of sticky ends following the double digest of the PCR product. This is followed by a PCR clean-up to get rid of small waste DNA (d) The vector pUC19 also undergoes a double-digest at matching restriction sites and is dephosphorylated with *Antarctic phosphatase* to prevent self ligation (e) The PCR product is cloned into the cut vector (f) T4 DNA ligase catalyses the formation of a phosphodiester bond between 5′ phosphate and 3′ terminal OH of the restriction fragments. The recombinant plasmid is then transformed into competent bacterial cells (a DH5α derivative), a description of which can be found in Appendix A.1.1. (g) Nt.BspQI creates a break, or nick, after any nucleotide N upon recognising the sequence 5′...GCTCTTCN...3′ (h) Exonuclease III catalyses the stepwise removal of nucleotides from the 3′-OH and operates from the nick. Finally a phenol-chloroform extraction separates the plasmid from a mixture of molecules. Again this is described in more detail in Appendix A.1.1.

Normally in DNA origami, the binding domains of the staples, between crossover positions, retain a symmetry throughout the structure. For example, for S-shaped staples mentioned in the Introduction, the top domain is 8 bps, the middle domain is 16 bps and the bottom domain is again 8 bps. This means the breaks between neighbouring DNA strands in a row or helix, called nicks, and the crossover positions, occur in register between parallel helices across the structure. This is analogous to the 'running' brick pattern in a brick wall (see Figure 7 ai), where the join between two bricks side-by-side in a row matches the position of the join between two side-by-side bricks two rows below. In the case of U-shaped staples, the same symmetry is normally imposed by the two binding domains both being 16 bps in length. This again means that crossover positions happen the same distance from the end of the tube, every other helix, like the brick wall.

In the chiral tube design here however, curvature is imposed by the presence of asymmetric U-shaped staples, with two domains that differ in length by two bases. The top domain is 15 bps and the bottom is 17 bps and they are connected at a staple crossover position. This differs from the usual symmetric U-shaped staple where both domains consist of 16 bps. The staples tessellate such that the binding domains along the length of a helix alter between 15 bps and 17 bps. Such asymmetry shifts the staple-crossover position two bases further along and around a helix, in relation to that on the helix two-previous. Using the running brick wall analogy again, this means that where the inter-brick gaps were at the same position every other row, they are now shifted slightly further right in relation to the inter-brick gap two rows before. Due to the RH nature of DNA and the fact that DNA is a helix rather than a flat line, this shift results in each crossover position being two bases further around the RH helix, or 'into' the sheet

and thus the result of this over the entire structure is the sheet curving into the page in a RH manner, rather than out of the page in a LH manner. By shifting the position of the crossovers 2 bases further, relative to the position of the crossover two helices previous, a shift of 10 bps over the whole structure is introduced (10 helices/2) x 2 bps). There still exist 32 bps between crossovers on the same helix (15 + 17) and so no twist or strain within the individual helices or between adjacent helices is introduced by this asymmetry. Through this, the RH chirality of DNA is propagated in the overall tube structure. Figure 7 (ai-ii, bi-ii) illustrates how the common origami tile (the Rothemund tile [24]) is designed to be flat and shows how the design is adjusted in our case to introduce curvature for the 10-helix tubes, with U-shaped staples.

The tube may be assembled with non-modified staples at either end of the tube or with end-staples that have unpaired poly-thymine loops between the 15 and 17 bp domains. This either allows or disrupts non-specific end-to-end stacking of tubes (caused by base-stacking) respectively either forming long thread-like tubes or monomers [24].

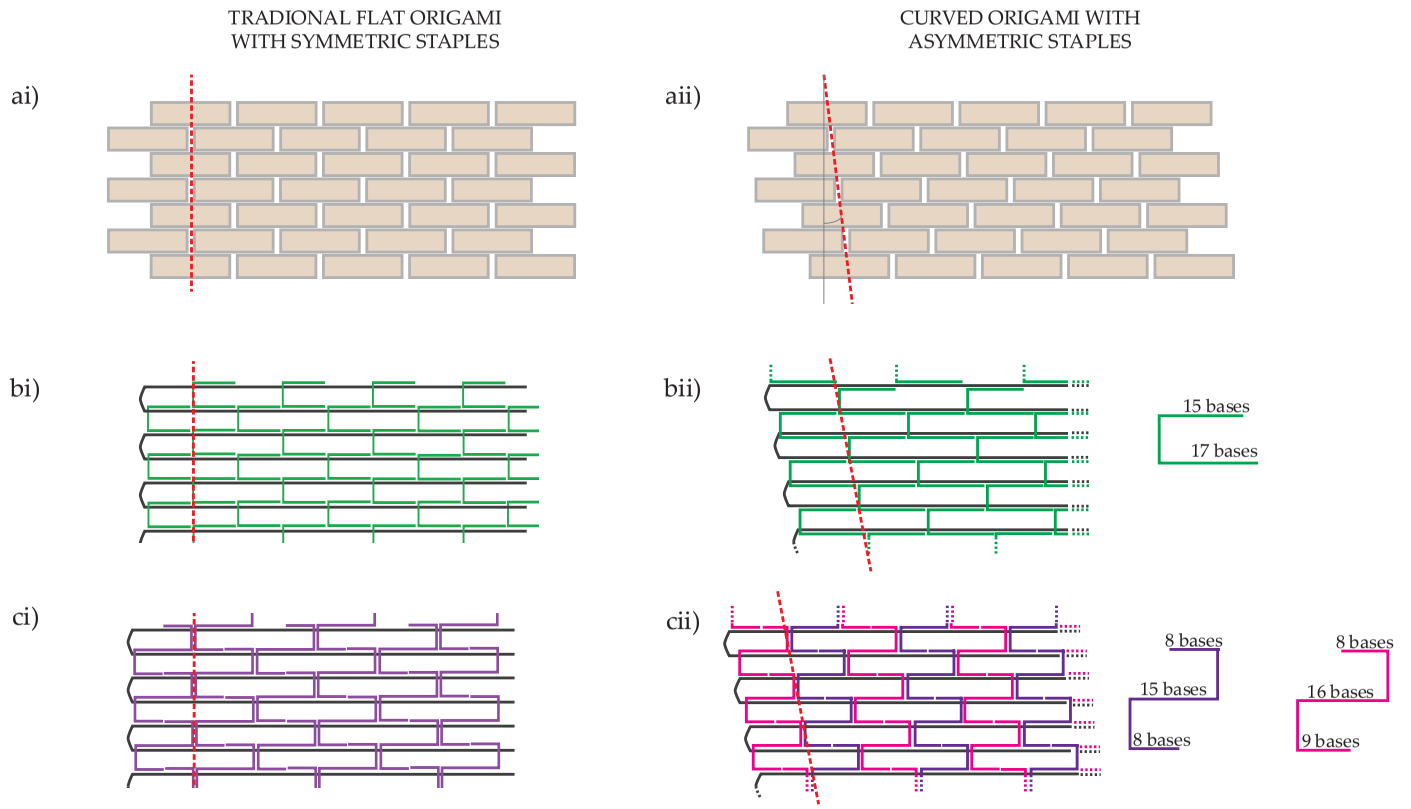


Figure 7: Creating curvature in a 10- and 20-helix nanotube via staple asymmetry and moving crossover positions. ai) The pattern that crossover and nick positions make in normal flat origami is analogous to the patterns made by bricks in a 'running' brick wall. aii) When the nick and crossover positions are shifted in our origami design it is analogous to moving each row of bricks in a wall slightly further right in relation to the previous row, such that the gaps between bricks on one row is shifted in relation to the row two-previous. bi) A flat sheet of origami made from U-shaped staples where the crossover positions are all in line between every other helix, shown by the vertical red dashed line. bii) The crossover positions can be shifted relative to that on the helix two-previous with U-shaped staples by introducing asymmetry in the binding domains- where one domain is shorter than the other. ci) For S-shaped staples, the story is similar. Where there is symmetry in the staple, the crossover and nick positions line up vertically and allow for flat origami. cii) shows the arrangement of staples to create the correct curvature for the tube. The staples with differing length binding domains give rise to the correct crossover positions and curvature.

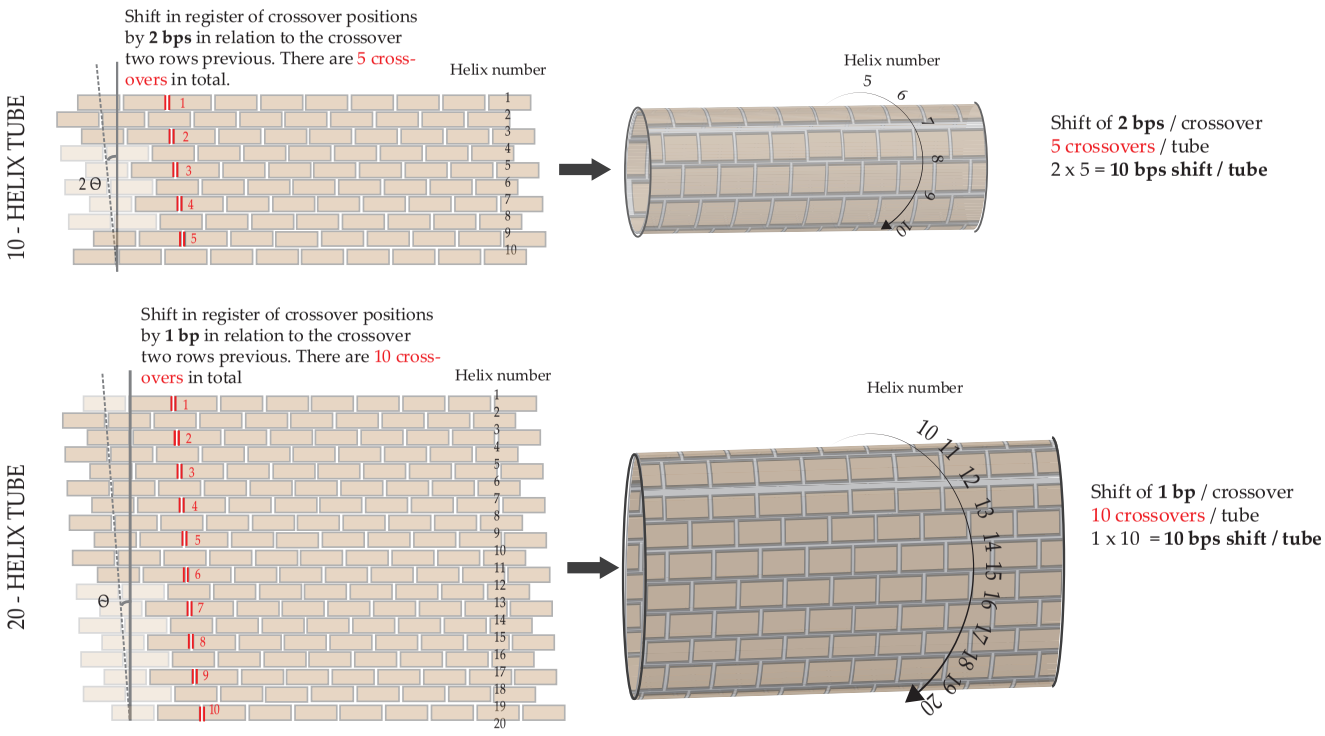


Figure 8: Crossover positions in a 10- and 20-helix nanotube. The lines between rows of bricks represent a double helix. The brick wall analogy is used here to show the crossover positions. There are 10 crossover positions for the 20-helix tube and 5 crossovers for the 10-helix tube. Crossovers occur *between* helices- thus there is one per pair of helices. For the 10-helix tube the crossover position is moved two bps further around the helix (in a RH manner) relative to the position of the previous crossover. For the 20-helix tube the crossover position is moved only 1 bp further around the helix (in a RH manner) relative to the position of the previous crossover. This means the angle of crossover positions made relative to normal crossover positions (parallel to the tube cross-section) in the 10-helix tube, is twice the angle made in the 20 helix tube. But there are 10 bps shifted across all crossover positions in both structures.

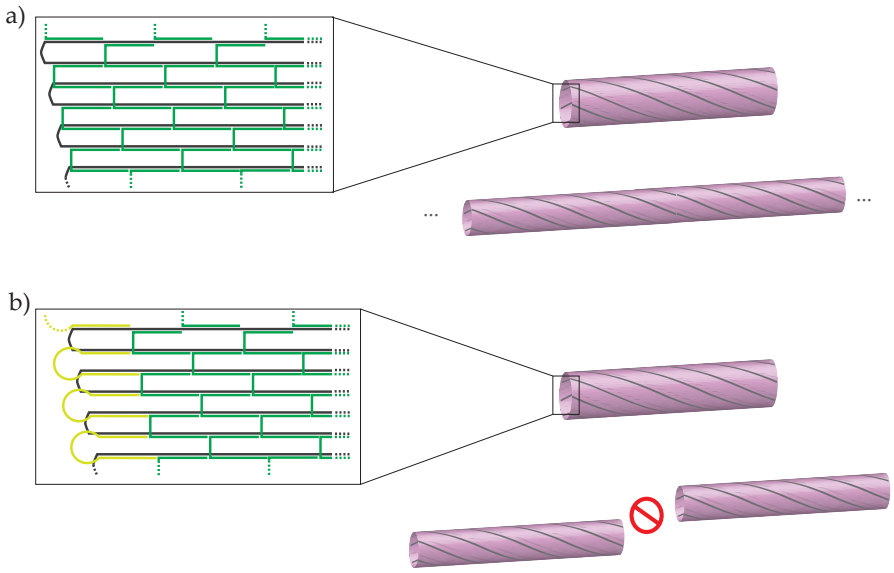


Figure 9: By including different sets of end staples, the tubes exhibit different behavoir. The normal U-shaped staples are in dark green and the special end staples in lime (a) Where there are no special staples at either end of the tube, they undergo non-specific end-to-end stacking which is energetically favourable for bases (b) By including special end staples with unpaired thymine loops present, the thymines repel each other and prevent stacking of bases and tubes.

2.2.2 *Design of a 20-helix chiral tube*

Here we introduce the design of a chiral tube with 20 helices which shares the same design principle as the 10-helix tube presented above. In order to give the 20-helix tube the same chirality as the 10-helix tube, the structure must be altered slightly. In the 10-helix design a shift in register of two bases per crossovers was introduced. Crossovers occur between five pairs of helices and so the resulting shifts in crossover positions across the entire structure consists of 10 bases in total (5 pairs x 2 bps), which is one complete helical turn per structure, for the 10 helix tube (see top row of Figure 8). To introduce a total shift of crossover positions over the entire structure of 10 bps for the 20-helix tube requires a shift in register of just one base further around the helix with respect to the previous crossover position, because crossovers occur between 10 pairs of helices instead of just five previously (see bottom row of Figure 8). This is achieved by having asymmetric S-shaped staples, instead of U-shaped staples. The S-shape staples are asymmetric and of alternating compositions and thus length along the tube, as seen in Figure 7 cii . Picking two neighbouring cross-sections of the tube, the staple domains consist of 8, 16, 9 bps - the magenta staples in Figure 7 cii - and 8, 15, 8 bps - the purple staples, such that there are still always 32 bps between crossovers on each helix ((8+8+16) and (9+8+15)). This is in contrast to the normal S-shaped staples where the domains are 8, 16 and 8 bps respectively.

The tube is assembled from 236 staples and an M13mp18 single-stranded bacteriophage scaffold. M13 is 7249 bases but the tube only requires 7057 bases to form, thus a small loop of 192 bases is formed at one end of the tube, but it is purely extrinsic to the design. The 20-helix tube has the advantage over the 10-helix tube in that the

scaffold strand is naturally occurring and does not require synthesis before assembly.

An additional advantage of the 20-helix tube is that it is designed such that stacking can be promoted through inclusion of specific end-staples which bind between tubes. The option for suppressing stacking is also included, but in place of poly-thymine loops, as with the 10-helix tube, the end-staples have eight unpaired poly-thymine tails instead. This has the same effect as in the 10-helix design and creates a steric repulsion between monomer tubes. Figure 10 illustrates how different end-staple configurations enable control over the lengths of the resulting structures.

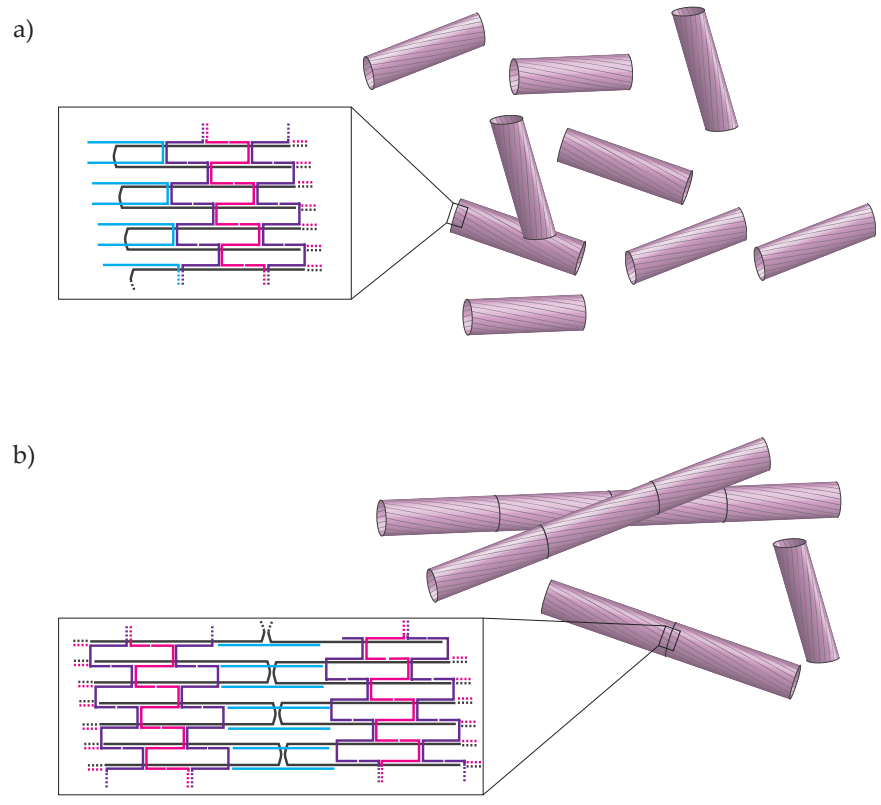


Figure 10: We can control whether the tubes form either monomers, where stacking is suppressed, or polymers, where specific stacking is promoted. (a) Inclusion of blue U-shaped staples, with unpaired poly-thymines at either end, suppress end-to-end stacking (b) The blue straight staples here are designed such that they complement the scaffold at either end of the same helix and the tubes polymerise.

2.3 SYNTHESIS

Both tube designs are assembled by annealing at 96° C followed by stepwise cooling to room temperature at a rate of -1° C min^{-1}. Previously it has been suggested that 3D origamis should be assembled over much longer time periods, for example 173 hours so that the cooling is slower [29]. However the quicker one-pot annealing time used here gave well assembled tubes when observed with AFM, as will be seen in the next section. This is probably due to the fact the 3D tube here is more like a 2D structure with periodic boundary conditions, compared to 3D origamis with multiples sheets and layers.

The $MgCl_2$ content of the buffer was varied to find a favourable composition for tube assembly. The optimal buffer comprised 50 mM NaCl, 10 mM $MgCl_2$ and 1 x Tris-EDTA and shall be referred to as tube buffer. The $MgCl_2$ concentrations tested were 1 mM, 5 mM, 10 mM, 50 mM and 100 mM. The ratio of monovalent- and divalent-cations is important in the formation of 3D origami structures. Divalent cations have be found to increase the amount of proper folding (but also increase unwanted aggregation) and monovalent cations have been found to antagonise structure compaction by competing with the binding of divalent cations [29, 68]; this disrupts folding but also disrupts aggregation.

The assembly of both the 10- and 20-helix tubes is a one-pot reaction where all the staples strands and the appropriate scaffold strand are mixed and annealed together. The ratio of scaffold:staple:end-staple strands is 1:3:5.

Following assembly the origami tubes are purified by passing though a series of three size-exclusion columns packed with Sepharcryl s-300 resin in 1 x tube buffer. The columns consist of tightly packed beads forming a porous layer. The small oligonucleotides can enter the me-

dia and are retained for longer than larger oligos, which pass through the column quickly. This is effective at separating small unbound staples from folded origamis.

To assemble the 20-helix stacking tube structures, the tubes are first assembled and purified in the absence of any end staples. Following this first purification, linking staples (1:1, linkers:origami) are added to the tube solution and incubated for 1 hour at 30 °C.

2.4 CHARACTERISATION

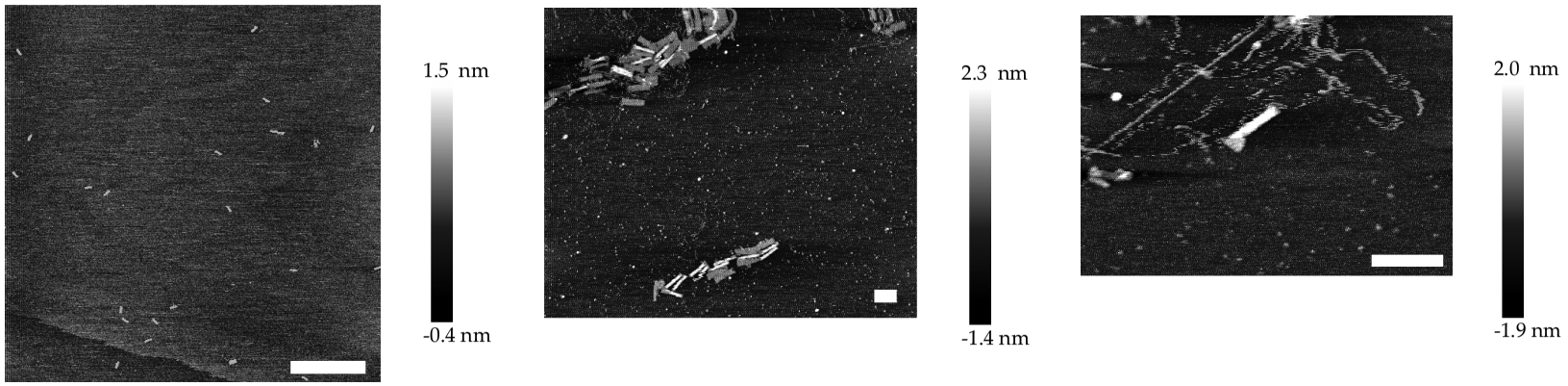
2.4.1 *Atomic Force Microscopy*

AFM was used to image the tube assemblies and was operated in tapping mode. 5 µl of origami is deposited onto atomically flat mica and imaged using a fluid cell. The cell holds a silcon nitride cantilever in place and this oscillates within a meniscus of 90 µl imaging buffer, comprising 1 x TAE, 12.5 mM $MgCl_2$ and 4 mM $NiCl_2$. Ni(II) binds DNA to mica tightly enough to allow imaging in solution with little movement of structures [46]. Appendix A.2.1 contains full details of the AFM method.

Figures 11 and 12 show topographic images of 10- and 20-helix tubes respectively, in monomeric form. In 11 b, the tubes appear to be aligned, both side-by-side and roughly end-to-end. This could be due to the presence of grooves, steps or edges on the mica surface and thus be an energetically favourable arrangement. It does not appear to be end-to-end base stacking however, as there are defined spaces between tubes and they are slightly offset from one another.

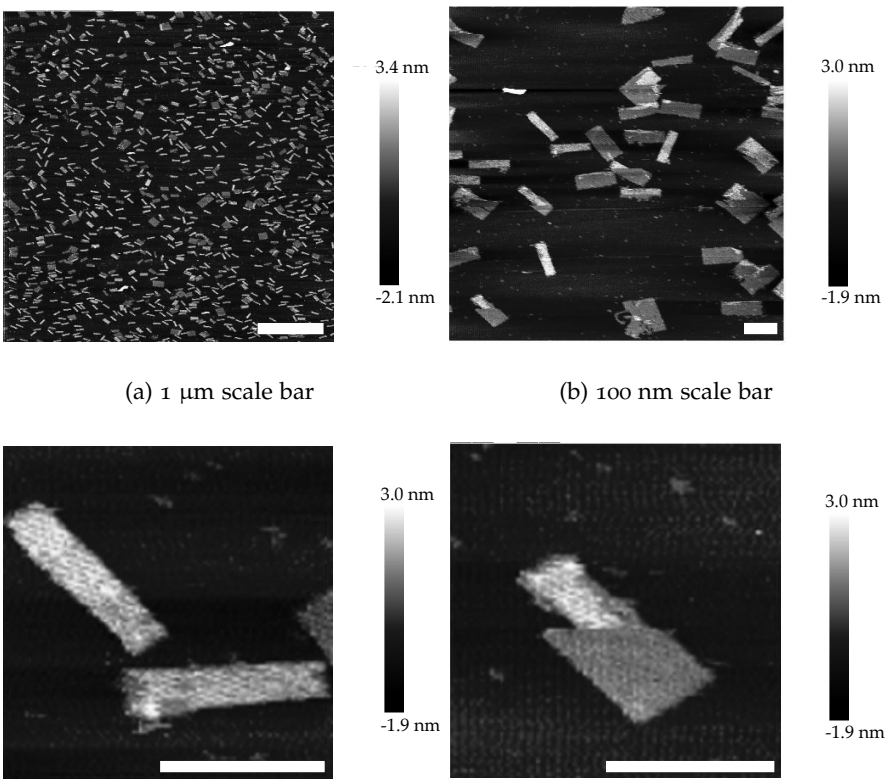
The well formed tubes are approximately half the width of the flattened single-layer sections of the tubes, which are often seen in the images. One possibility as to why these single layers are observed

during imaging may be as follows. As mentioned previously, nicks are breaks between DNA strands in a helix which prevent the helix being continuous. These are found between the ends of separate, neighbouring staples that bind with the scaffold on the same helix. Nicks give rise to areas of weakness due to discontinuity in the phosphate backbone of DNA. Thus the tube opening may be initiated by the tip at nick positions. To prevent tubes opening via this mechanism during AFM imaging, it is possible to increase their stability through ligation of staples at nicks. This was shown by O'Neill *et al.* to produce structures capable of surviving temperatures up to 70 °C and which remain closed during AFM imaging [69]. This would enable us to make an estimate of the number of well formed tubes - at present it is not possible to tell if the flat tile regions have opened due to the force of imaging, due to interactions between the mica, Ni(II) ions and DNA, or are mal-formed during assembly.



(a) Scale bar is 1 μm (b) Scale bar is 100 nm (c) Scale bar is 100 nm

Figure 11: The 10-helix tube in monomeric form, with the thymine-loop end-staples present. (a) 6 nM origami. This concentration is somewhat sparse, but it is possible to see well formed monomers that do not stack end to end. (b) 8 nM origami. Some origami tubes open up on the mica surface. It may also be due to forces of the tip exerted on the origami. (c) 8 nM origami. At the lower LH end of the tube pictured here, we see an example of how a tube can 'unzip' due to the cantilever force during imaging. Cross-sectional analysis confirms this by showing us the double layered origami is 3 nm high (equivalent to two flattened sheets) and that the single layer is 1.5 nm (which is what we expect for a single origami sheet) The double layers appear brighter than the single origami layer.



(a) 1 μm scale bar

(b) 100 nm scale bar

(c) 100 nm scale bar

(d) 100 nm scale bar

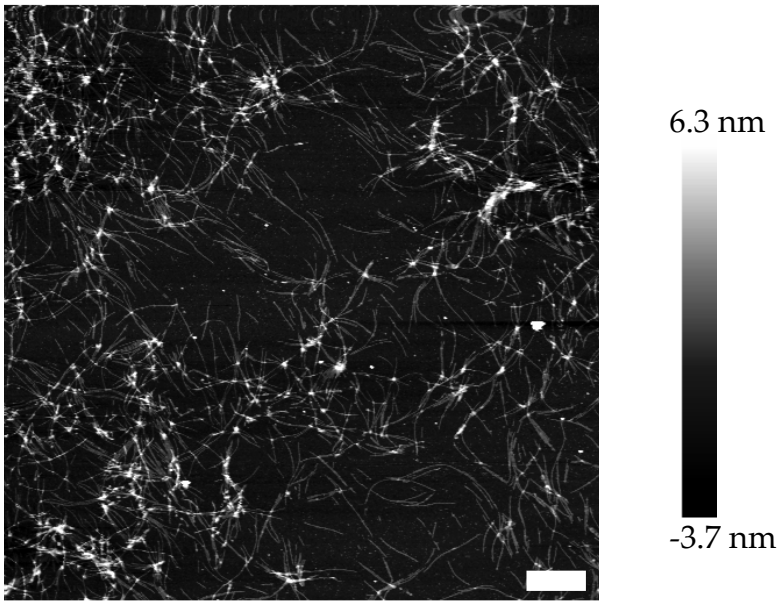
Figure 12: The 20-helix tubes form individual monomers with the inclusion of poly-T staple tails. The monomers are just over 100 nm in length. (a-b) The majority of tubes are well formed, some are flattened upon imaging. It is also possible to see the presence of mal-formed tubes or those which have formed dimers. (c) A closer view of the 20-helix tube shows us the origami lattice structure. (d) Again it is possible to open the tubes to a single layer by changing the the cantilever force.

Figure 13 a shows the non-specific base-stacking of the 10-helix tubes produces structures up 3 μm in length, imaged at a concentration of 20 nM origami. In comparison, the 20-helix tubes in Figure 13 b show structures up to about 2 μm, which are formed due to base-pairing between linking staples and scaffolds on separate structures. It is expected that the latter formation is more stable as well as more specific, due to base-pairing being stronger than the base stacking interaction.

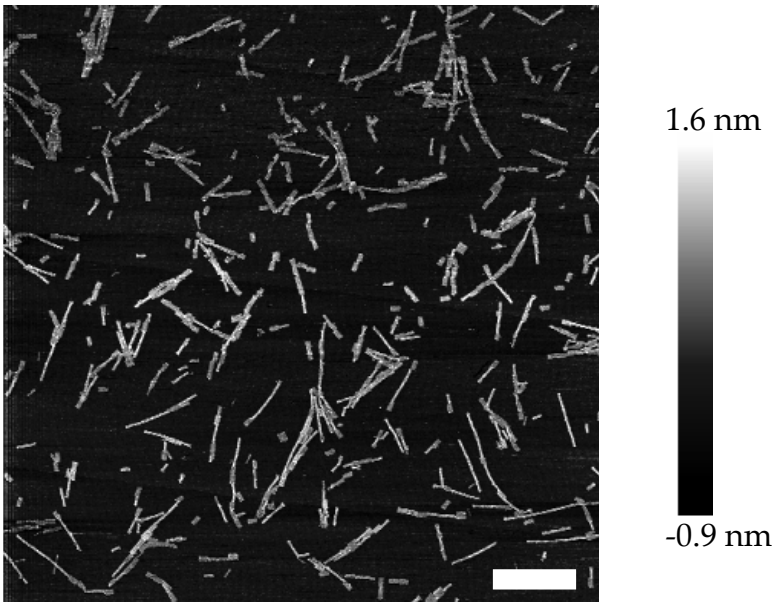
The lengths of the 20-helix, end-to-end stacked tubes are analysed further. Four AFM images of 7 μm frame size are first binarised. Using an in-built function in the software Wolfram Mathematica [70], the tube images are fitted with individual ellipses such that the short and long axes of the ellipses fit the short and long axes of the tubes, respectively. The component measurements of best-fit ellipses of the tubes are read out by the programme. Limits to this fitting programme are applied such that objects shorter than three-quarters of a tube length, and/or narrower than half a tube width are discounted. The resulting sizes are plotted in histograms in Figure 14. This rough analysis shows that using this method of polymerisation gives complexes which favour overall shorter stack lengths, between one and two tubes long. Stacks of over 10 tubes are found in some cases but are rare. This is what we might expect - as the polymer gets longer, the concentration of free monomer decreases so it is less and less likely that you add another one to a stack, but the probability of the polymer breaking at any point on its length stays the same. The Nobel laureate Flory, together with Schulz, described the relative ratios of polymers of different length after a polymerization process, based on their relative probabilities of occurrence [71]. However, when the number of monomers is large, this can be described by an exponential function, where the probability of forming longer stacks of tubes is lower than

forming shorter stacks. An exponential distribution is what we observe after our tube polymerisation process, where shorter lengths are favoured. Figure 15 summarises the distribution of stack lengths for the 4 histograms in Figure 14, by plotting the total frequency of stacks formed from a given number of tubes (the red points) and the blue fit shows the distribution of lengths fits an exponential function well, with an R-squared value of 0.992.

It would be interesting to change the polymerisation method and analyse the resulting lengths of structures formed. The method of polymerisation used here is not particularly efficient. The linking staples are the same for both ends of the tube, and so if the stochiometry of tube:end-staples is slightly in favour of the end-staples, both ends of a tube could be functionalised with staples, thus slowing or preventing polymerisation. This may be thought of as having two ends per tube activated for polymerisation, whereas ideally only one end would be activated per tube, ready to polymerise to an unactivated end of a second tube. By having 20 linking points with relatively short binding regions to each end of the tube, we hope to have overcome this in our design. A marked improvement on our design however, would be to form two different monomer tubes 'A' and 'B' that could only polymerise to form 'ABABAB...' [72, 26].

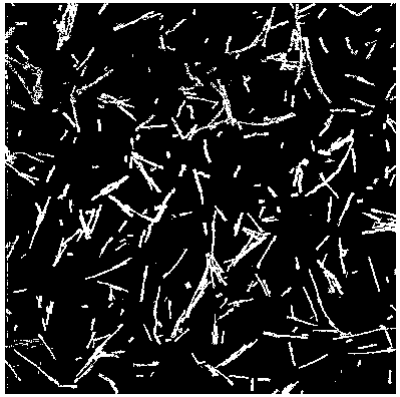


(a) 10-helix tubes. 1 μm scale bar



(b) 20-helix tubes. 1 μm scale bar

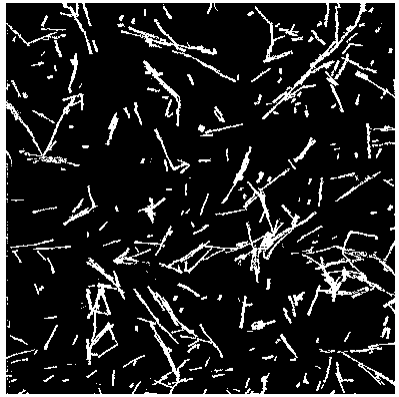
Figure 13: Both 10-H and 20-H tubes show evidence of end-to-end stacking, but through different means. (a) The 10-helix bundles show non-specific end-to-end base stacking. No special end-staples are included here and many tubes come together to minimise energy. (b) Surprisingly, the specific stacking encouraged by the inclusion of linking staples in the 20-H bundles, does not seem to produce such lengthy complexes as non-specific stacking. Here we see structures up to approximately 1.5 μm. The stacking is more precise than base stacking, giving rise to structures with a greater persistence length than in (a).



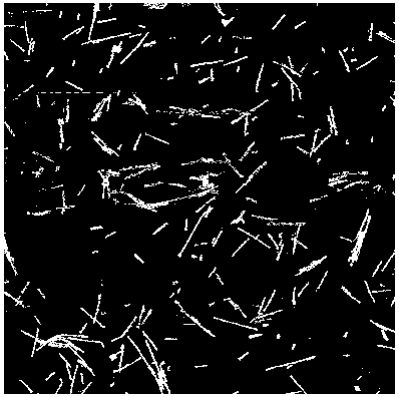
(a) Binarised AFM region 1. 10 nM origami, 7 µm frame



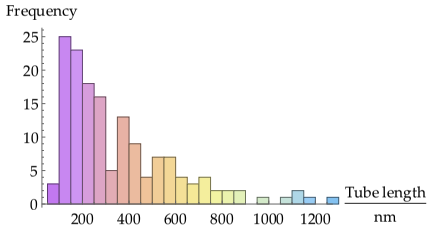
(b) Binarised AFM region 2. 10 nM origami, 7 µm frame



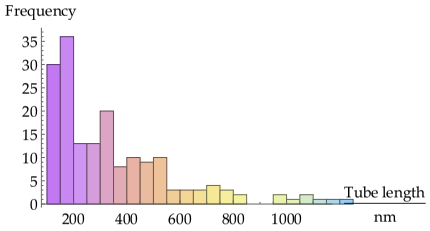
(c) Binarised AFM region 3. 10 nM origami, 7 µm frame



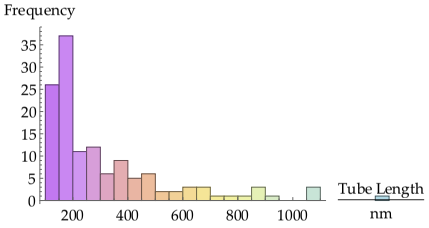
(d) Binarised AFM region 4. 10 nM origami, 7 µm frame



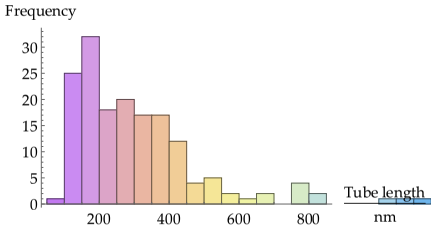
(e) Tube length distribution for region 1. Median length 277 nm.



(f) Tube length distribution for region 2. Median length 280 nm.



(g) Tube length distribution for region 3. Median length 224 nm.



(h) Tube length distribution for region 4. Median length 270 nm.

Figure 14: The distribution of tube lengths resembles a Flory-Schulz distribution of polymerisation, which describes how shorter tubes are favoured and more probable than longer tubes [73]. The AFM images are binarized and the dimensions of elliptical components read out- this is an automated process. The objects are then sorted to filter out those with a length shorter than 0.75 X tube length and a width narrower than 0.5 X tube width, which would be mal-formed tubes. The resulting distributions of lengths are plotted in (b), (d), (f) and (h).

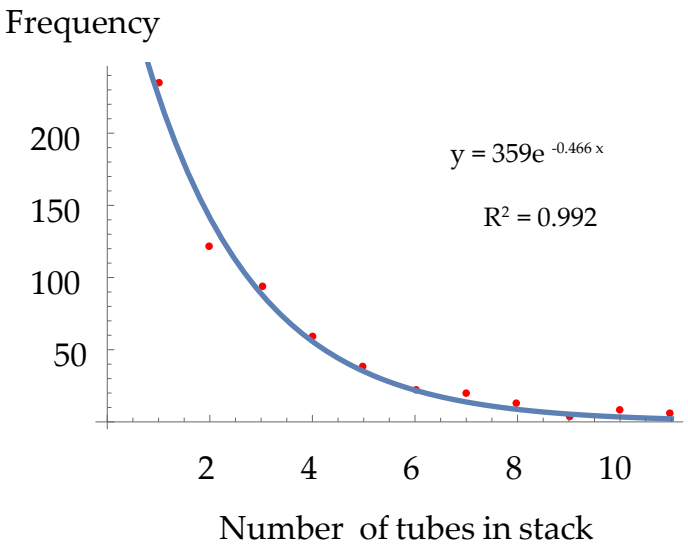


Figure 15: Summary of tube stack lengths. The exponential function describes the relation between probability of formation and the number of tubes present in a stack of tubes. The frequency here is the total number of observed stacks of a certain length. the lengths of the stacks can be thought of as the number of tubes in that stack, as is shown here.

2.4.2 *Melting curve analysis*

The melting temperatures of both tube types are determined via melting curve analysis. SYBR Green is a DNA intercalator which preferentially binds to dsDNA compared to ssDNA and fluoresces more brightly when bound to the former. By monitoring changes in fluorescence with temperature, when the dye is mixed with origami, it is possible to find the melting temperatures of the tubes. The point at which origami melts is best observed by plotting the negative differential of the fluorescence with temperature; this shows the change from double- to single-stranded DNA as a sharp peak in a curve. As SYBR Green is an intercalator (it slots between planar bases), it is important to use it sparingly so as to have little effect on base-pairing itself [74]. Here 0.7 pmol of origami is used with SYBR Green at a much lower working concentration, typically 1:10,000. The samples undergo two cycles of heating to 95 °C and slow cooling to 25 °C, as seen at the bottom of Figure 16. This takes place in a real-time PCR machine which allows fluorescence monitoring with varying temperature: as the temperature is increased, the amount of bound dye decreases.

For the 20-helix tube, melting temperatures of the origamis in the presence of varying amounts of formamide are found. This may allow for tube assembly via a different method from temperature annealling, where temperature sensitive components cannot be readily included. Formamide is a DNA denaturant which acts by disrupting the hydrogen bonds between bases. Formamide has been found to lower the melting temperature linearly by 0.6 °C per % formamide present in the buffer [75, 76]. However, here it would appear the melting temperature is lowered by approximately 1 °C per % formamide present in the buffer for samples containing 0-30% formamide. For the sample with 40% formamide, this relationship does not hold -

probably due to the fact the structure barely forms at such high concentrations of denaturant. The real-time PCR machine does not go to temperatures lower than 25 °C and so lower temperatures were not probed.

This data provides us with an understanding of the concentrations of formamide required to assemble the 20-helix tubes isothermally, by dialysing solutions with reducing concentrations of denaturant. Examples of assembly processes such as this have been conducted by Jungmann *et al.* and Sobczak *et al.*, and have been shown to be useful for temperature sensitive experiments involving origami [75, 74].

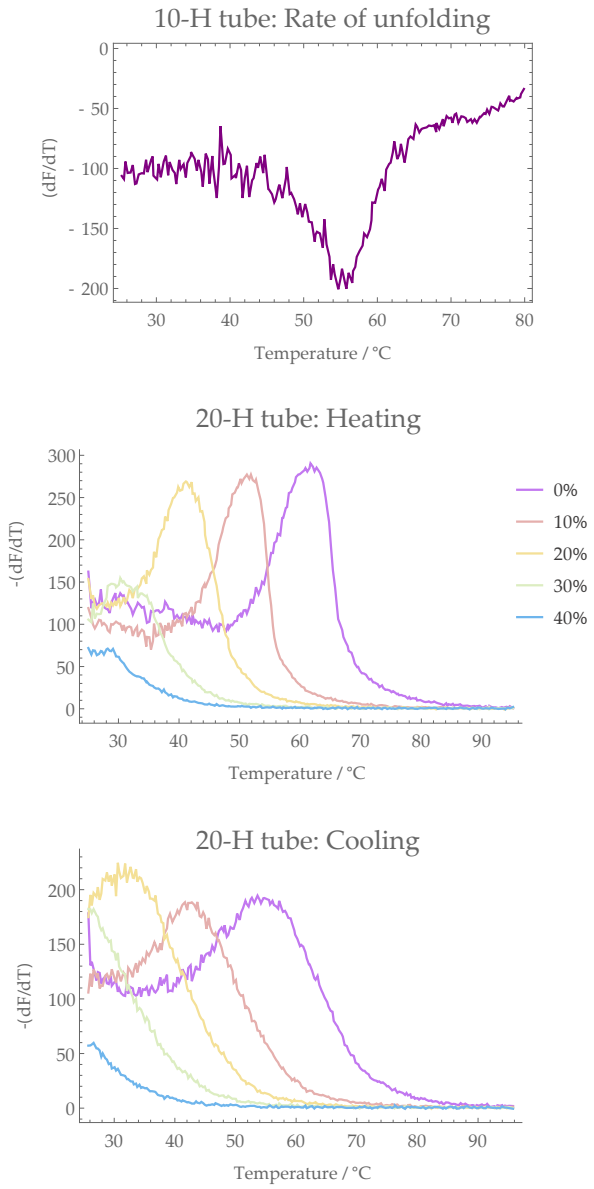


Figure 16: Melting curve analysis of the origami tubes. Top: the graph shows the rate of unfolding for the 10-helix tube. The dip in the differential of the fluorescence with temperature shows the melting point to be around 56 °C. The cooling data was not recorded for the 10-helix tubes. Middle: the graph shows one cycle of the heating data for varying amounts of formamide. Bottom: this graph shows data from one of the corresponding cooling cycles. The melting temperatures are taken from the heating data and are 28, 32, 42, 53 and 63 °C for samples in buffer containing 40, 30, 20, 10, and 0% formamide respectively.

2.4.3 *Circular dichroism*

We have so far presented the tube design and characterised the assembly in terms of AFM and melting curve analysis. However, crucial to our aim is producing a tube with predefined inside and outside surfaces. To prove successful formation in this regard, we exploit the fact that our origami structures can be functionalised with gold nanoparticles (AuNPs) which are comparable in size to the wavelength of light.

An AuNP can interact with the oscillating electromagnetic field of a light wave. Metals such as gold have conduction band electrons, commonly known as free electrons, and these undergo collective oscillations upon interaction with the incident light. The oscillations spread across the surface of the AuNP inducing a dipole between the negatively charged electrons and the ionic lattice. The maximum amplitude of this oscillation arises at a frequency known as the surface plasmon resonance frequency. Surface plasmon resonance (SPR) is accompanied by a strong absorption of incident light, and such strong absorbances are characteristic of the noble metals. The frequency of the SPR and the wavelength of absorbed light may be affected by the dielectric constant of the solution, the type of metal, and the size and shape of the particle [77].

The circular dichroism (CD) effect describes the difference in absorbance between the LH and RH components of circularly polarised light, after passing through an optically active solution. Such a solution is created here by decorating the 20-helix nanotubes with AuNPs evenly spaced along and around the tube surface to create a LH chiral arrangement, which may seen in Figure 17. This arrangement of AuNPs creates a coupling between SPRs on separate AuNPs, and its chirality gives rise to differential absorption of LH and RH light. The

collective oscillations (both the RH and LH component) are themselves made up of a longitudinal component (where the oscillations propagate in the direction of the z-axis of the helix) and a transverse component (where oscillations are in the x-y direction). These two components are excited by the corresponding components of the incident light. In very simplified terms, the LH propagation results in more of the LH component of light being absorbed than the RH component.

The characteristic CD signal for a solution containing LH chiral arrangements is a 'peak-dip', as opposed to the RH 'dip-peak' [65]. The total CD signal is the sum of signals for the transverse and longitudinal oscillations. In our case, the LH signal would be made up of a peak at approximately 520 nm from the z-axis contribution and a dip at roughly 555 nm from the x-y-axis contribution.

Observing a LH signal however, is dependent on the tube forming as designed. Figure 17 a shows the positioning of the AuNPs on a 2D representation of the origami. If the tube forms by such a sheet rolling into the page (as it is designed to, and described above along side the brick wall analogy) then the AuNPs end up on the outside of the tube and form a LH helix, as depicted in Figure 17 b . However, if the sheet rolls out of the page then a RH CD signal should be observed. An equal mixture of the two would produce no signal. The tubes themselves are designed to adopt the RH chirality of B-form DNA, but the AuNP arrangement is designed to be LH if the tube forms correctly.

Nine thiol-modified AuNPs are evenly spaced along the tube. The AuNPs are functionalised with multiple thiol-oligos (via a thiol-thiol reaction) consisting 16 thymine bases which form attachment sites for complementary regions of staples extending from the origami. The AuNPs are incubated with the tube following origami assembly and

purification. The resulting assemblies are analysed with transmission electron microscopy (TEM) and an example is shown in Figure 17 c. The AuNPs are 10 nm in diameter and according to the manufacturer, the SPR (and corresponding CD peak wavelength) should occur at 520 nm [78]. A range of wavelengths of light from 370 nm to 750 nm were sent through two samples; one sample of origami+AuNPs, and a control of AuNPs in solution, which contained the same amount of AuNPs as was added to the origami+AuNP sample. A plot of these, together with the background-subtracted data, is seen in Figure 18. The background-subtracted data is generated by subtracting the control curve from the origami+AuNP curve. DNA itself is chiral and so this ensures the signal we look at is resulting from the origami+AuNP structure alone and not due to the DNA attached to the AuNPs. The strong peak-dip signal proves the tubes form as they are designed to. The corresponding arrangement with the AuNPs on the inside of the tube was also made and measured but did not give rise to a CD signal. This is either because the AuNPs are big compared to the diameter of the tube and are preventing formation, or because the diameter of the helix is too small for this effect to be observable. TEM has not been carried out on these assemblies and so mal-formation cannot be ruled out in this case.

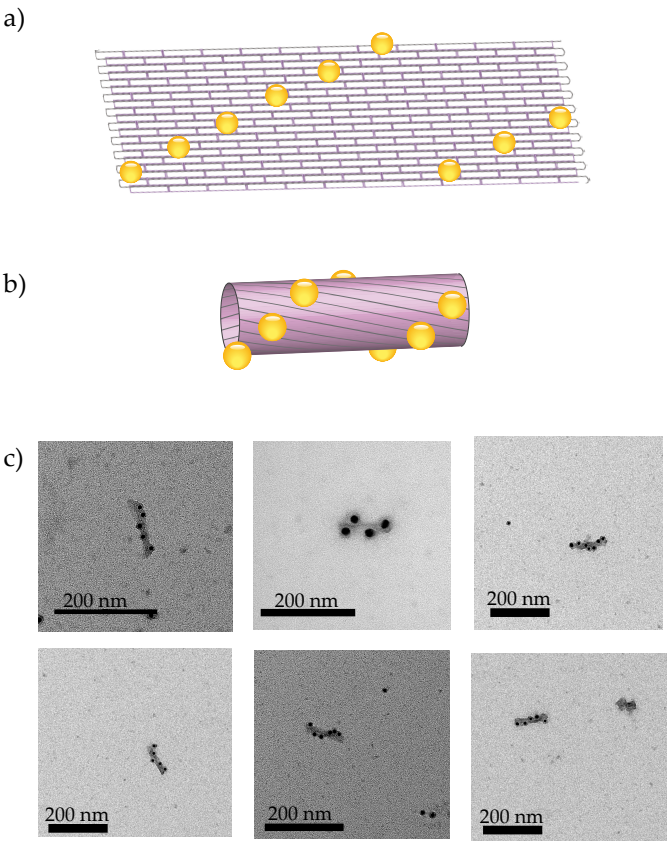


Figure 17: AuNPs are attached to nine places on the exterior surface of the nanotube, with a pitch of six AuNPs. (a) The placement of nanoparticles is such that they form a continuous helix about the exterior of the tube. The diagrammatic AuNPs are not drawn to scale. (b) If the tube assembles correctly, as is designed with in-built RH curvature, the AuNPs should form a LH arrangement. (c) The quality of the assemblies is checked with TEM and the resulting LH arrangement is visible. All AuNPs are 10 nm a part from the top row, middle image, where the AuNPs are 15 nm. Thanks to Robert Schreiber and Natalie Haley for taking the images and allowing me to use them here.

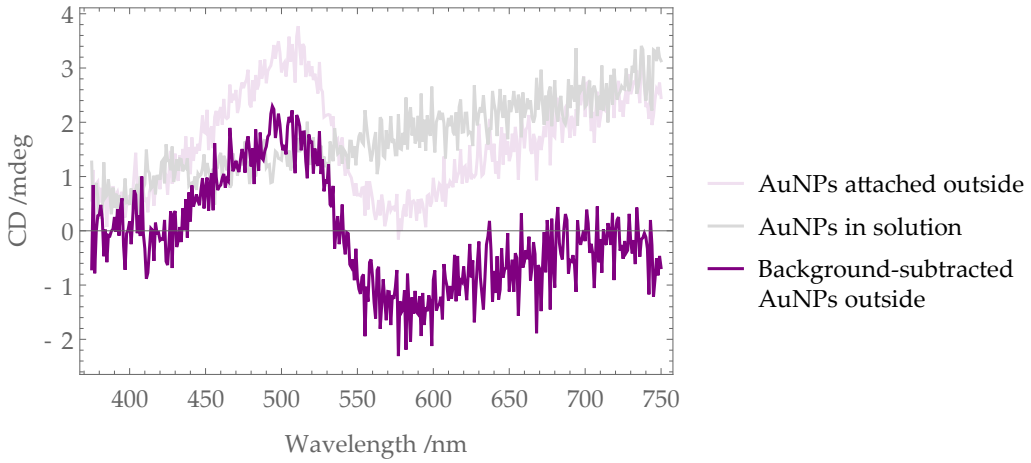


Figure 18: The CD signal of the 20-helix nanotubes exhibits the characteristic peak-dip attributed to a LH chiral arrangement of plasmonic nanoparticles. The dark purple curve shows the background-subtracted data created by subtracting the curve for the AuNPs+DNA in solution from the origami+AuNP curve, which are both shown in the background of the plot for reference. The CD signal is the difference in absorbance of LH and RH components of light (LCP and RCP respectively), $\Delta A = A_{LCP} - A_{RCP}$

2.5 CONCLUSIONS

In this chapter the designs of 10- and 20-helix chiral DNA nanotubes are presented, with the purpose of predefining internal and external surfaces of the tubes. The RH chirality of DNA is adopted in the overall tube design and introduced by shifting the register of crossover and nick positions from one helix to the next. The 20-helix tubes have the ability to stack in a controlled manner to form structures of approximately 15 tubes in length, although shorter structures of 2 tubes in total are favoured. The assemblies are characterised using AFM and the melting properties determined by monitoring the rate of change of fluorescence of an intercalating dye with temperature. Proof of design is achieved by observing a LH chiral signal resulting from the arrangement of plasmonic nanoparticles about the tube exteriors. The design and characteristics of the tubes have been consolidated and we may now move forward to explore areas for their practical application.

3

A ROLE FOR DNA NANOTUBES IN ENZYME CASCADES

In this chapter, three enzymes that form a substrate-product reaction cascade are attached to the inside and outside surfaces of 20H DNA nanotubes. The potential of the tube to act as a scaffold for such a cascade is analysed in the context of the wider field.

3.1 BIOLOGICAL MODULATION: SCAFFOLDS AND TUBES

The interior of a cell presents an highly crowded environment which is vital to many of its functions. For example, as regulation of pH, temperature and concentration is important in biochemical reactions, so is the effect of steric repulsion between biochemical components. The 'excluded volume effect of macromolecular crowding' encourages proximity and binding between bulky components allowing reactions where otherwise they might not occur [79]. The term 'crowded' is used to describe the interior, as opposed to 'concentrated', illustrating that no single species occurs at a high concentration, yet all together the molecular species occupy a significant fraction of the available cellular volume. The greater the number of molecules in a solution, the more the configurational entropy of the species is reduced. Thus a binding event leads to a reduction of the total free energy of the system, concomitantly increasing the volume available to all other species present [80]. Compact cellular arrangements result in reduced diffusion coefficients of components whilst increasing

their activity. Therefore the effect of crowding on the overall rate of a reaction will depend on the interplay between these two factors.

3.1.1 *Organisation in cells*

Excluded volume effects make structured domains and compartmentalisation essential for achieving high catalytic rates *in vivo*, such as is required for signalling and metabolic pathways. The importance of compartmental organisation of the cell has long been recognised; organelles and membranes create functionally and spatially distinct areas from the cytosol or the extracellular environment [6]. The more recent discovery of scaffold proteins has furthered our understanding of the coordination required for the cell to carry out its many distinct activities. Scaffold proteins are so named because of their role in coordinating the physical assembly of components involved in a network or pathway although there is not a specific definition in the literature [81]. Despite being conceptually related, scaffolds are diverse in both the pathways they regulate and in their structure [82]. It is also common for scaffolds to serve multiple functions, including but not limited to activation, localisation and protection of bound entities from sequestering side reactions [81].

Similarly to scaffolds, metabolons also play a role in flux regulation. The term 'metabolon' was coined by Srere to desribe a 'supramolecular complex of sequential metabolic enzymes and cellular structural elements' [83]. Metabalons are different to scaffolds in that a scaffold holds enzymes together and a metabolon is an enzyme complex, but both provide examples of enzymes operating together and in proximity to each other in the cell. Metabolons might be transient, fragile or stable and are held together only by non-covalent interactions and so may be assembled as and when they are needed [84]. The compact

arrangement of intracellular species assists in metabolon assembly, and due to the concentration of enzyme sites often exceeding that of metabolites, pathways can progress via direct transfer of intermediates from one enzyme to the next. This bypasses dissociation and diffusion in the aqueous environment and leads to an increased flux through successive enzymes in a cascade [85]. This process is called channelling and will only occur when the enzymes involved are sufficiently proximate. Substrates that are channelled benefit from enhanced reaction rates, preclusion from competitive or scavenging reactions that take place in the bulk, possibly protection from solvation, and in some instances, require a reduced amount of substrate [86].

The significance and existence of channelling effects *in vivo* has long been a topic for debate. This debate was prompted by Ovadi's paper in which he discusses the physiological significance of channelling [87, 88]. To illustrate one area of issue- which arose in two instances- a channelling effect was inferred upon finding an enhanced reaction activity between fused scaffold proteins [89, 90]. The findings were later refuted and the results explained entirely in terms of free diffusion, but instead the calculations were carried out with newly characterised kinetic constants, whose alteration during the fusion process had not been accounted for in the original work [91, 92]. Additionally Jacobson and Wojcieszyn postulate that small molecular diffusion in the cytomatrix occurs at nearly aqueous rates and may be rapid enough to overcome loss of efficiency due to three-dimensional diffusion [93]. They propose this as a mechanism of transport for small molecules in cytoplasm, although they are far from certain as they also propose a counter-argument to this.

Next, the various mechanisms of channelling will be discussed, which are applicable to both *in vivo* and *in vitro* systems.

3.2 MECHANISMS OF CHANNELLING

3.2.1 *Channelling: Perfect or leaky?*

There are two overriding classes of mechanism for channelling of substrates between enzymes: the first is perfect or direct transfer, and the second describes a proximity or 'leaky' transfer [94]. In both instances, the intermediate is prevented from entering the bulk phase [95, 94]. The latter mechanism is found in more general circumstances and is defined to be not so efficient and enzymes in this instance are usually located in high surface densities or complexed into large aggregates sometimes involving the membrane. In perfect transfer, the substrate is passed from one enzyme in an enzyme-enzyme complex immediately to another in the same complex and often involves a molecular feature to assist such transfer. It is only since the advent of advanced X-ray techniques for crystal structure determination that proposed mechanisms involving molecular features have been elucidated. Prior to such techniques, the challenge of isolation and purification of metabolons and other large complexes meant that experimental evidence for molecular channelling mechanisms often lacked a strutural basis [87, 96].

3.2.2 *The mechanisms*

TUNNELS Over the past two decades there have been multiple discoveries of enzyme-tunnel-enzyme structures [97, 98, 99, 100]. The inaugural identification was of a 2.5 nm tunnel linking the active sites on α catalytic and β catalytic subunits of the bi-enzyme structure tryptophan synthase, found in *Salmonella tryhimurium*. In the final two steps of trytophan synthesis, the α subunit catalyses the reversible

formation of indole and glyceraldehyde-3-phosphate from indole 3-glycerolphosphate, following which, the β subunit irreversibly catalyses the formation of tryptophan and water from the indole intermediate with serine [101]. Two α and two β subunits come together to form a tetramer $\alpha_2\beta_2$. Within this, the diffusion of intermediates along a hydrophobic channel is mediated by configurational changes of proteins initiated via ligand binding. The resulting diffusion is fast and thus reaction rates and affinity for substrates is increased by one to two orders of magnitude [101, 102].

ELECTROSTATICS Electrostatics is another mechanism which is responsible for substrate guiding and channelling between active sites. Upon solving the structure of the bi-enzyme complex dihydrofolate reductase-thymidylate synthase, Knighton *et al.* proposed that attractive interactions between the charged substrate and surface might be responsible for the channelling effect previously observed [103, 104]. There is no tunnel between the two active sites which are separated by 4 nm, suggesting that the intermediate substrate is released from the interior of the protein. However, the surface between these active sites contains positively charged residues which interact with the negatively charged dihydrofolate intermediate, preventing complete release into the bulk and thus assisting in channelling. Theories such as this one have also been supported by simulations and modelling [105].

In addition to the molecular tunnel and electrostatics examples presented here, there exists an abundance of naturally occurring channelling aides; ordered metabolon arrangements located in membranes are commonly found in plants, as well as more architecturally intricate synthetic arrangements, which will be presented in the next section [106, 107, 108].

3.2.3 *Measuring channelling effects*

In order to observe and quantify channelling, most methods focus on an indirect form of measurement by monitoring its effects. The transient or lag time, τ, lends itself as a clear way to identify channelling. The transient time is that required to reach the steady-state flux of a cascade intermediate. In the instance of 'perfect' channelling τ approaches zero, but with a 'leaky' channel as τ increases, the effectiveness of channelling decreases [105]. Another method which enables recognition of channelling involves setting up either a competing or inhibiting reaction in the bulk phase. This means that it is possible to quantify the extent to which a channel is leaky. Where the competing reaction sequesters the enzyme cascade intermediate, if the channelling effect is substantial then little should be lost. However, in the absence of channelling, a significantly reduced amount of intermediate successfully makes it to the next enzyme in the cascade and the challenging bulk reaction succeeds in scavenging [109, 96].

It is evident that biology has come up with a tremendous array of possibilities for ensuring fast, efficient and highly specific transport of reactants. This is also a challenge faced by chemists and biotechnologists wishing to design syntheses with similar attributes as those produced by evolution. Next, we analyse and discuss the various synthetic designs that attempt to mimic biological channelling.

3.3 SYNTHETIC ENZYME CASCADES

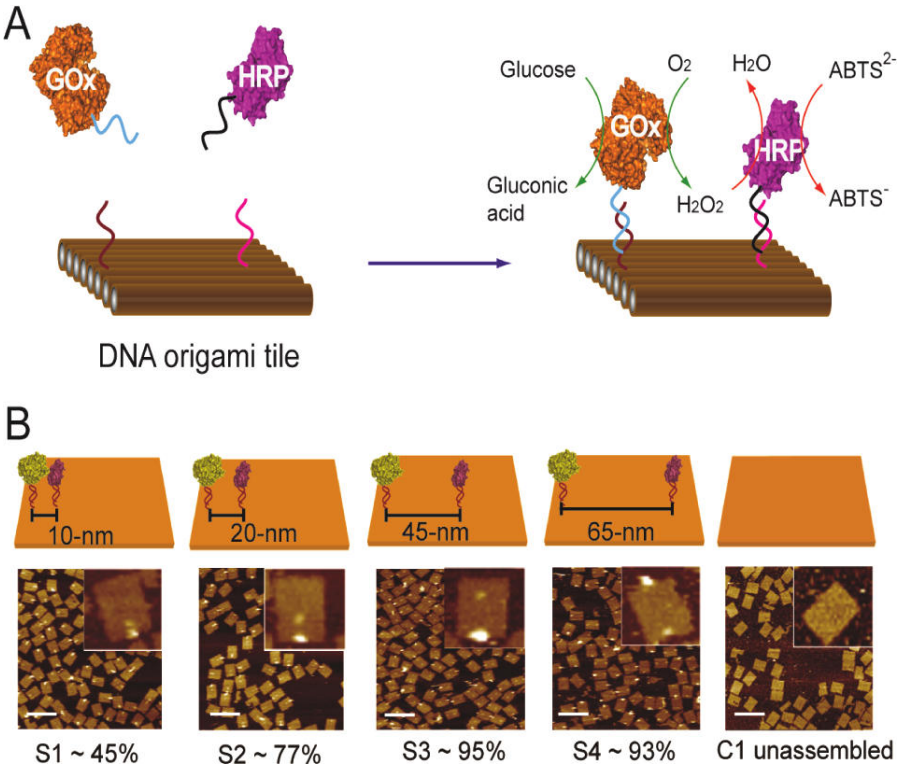
Since the early twentieth century, enzymes have been isolated and used in all walks of life; from the food industry to the textile industry, in pharmacology, in chemical synthesis, for biosensing and in biofuel production [110]. Manipulating enzyme arrangements to

our own effect has the advantage that enzymes from different organisms can be brought together to create novel cascades. In addition to the many advantages listed in the previous section for evolved biological cascades, synthetic biological cascades also benefit from mild reaction conditions whereas classic synthesis often takes place in polluting solvents [110]. Most synthetic scaffolds have so far focused on increasing the throughput of a cascade via proximity. For some such techniques, random co-immobilisation is sufficient; enzymes have been co-encapsulated inside coated carbon nanotubes, cross-linked to specially prepared surfaces or cross-linked to one another directly [111, 112, 113]. In other cases, enzymes are located more precisely via site-specific protein tagging and may be subsequently attached to scaffolds that have so far included proteins, DNA, nano-containers and micro-fluidic devices [114, 115, 116, 117]. Here we shall take a closer look at the DNA scaffolds presented previously.

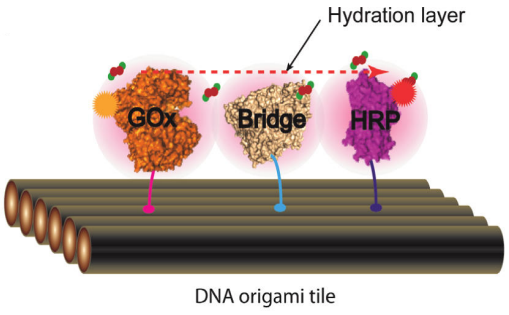
The initial steps towards creating DNA-directed multi-enzyme complexes were first taken by Niemeyer *et al.* They chose to work with the enzymes NADPH flavin mononucleotide oxidoreductase (NFOR) and Luciferase (Luc), which catalyze two consecutive reaction steps. To improve regioselective coupling of the DNA with protein moieties they used a recombinant biotinylated form of both enzymes (bNFOR and bLuc). This enabled each enzyme to bind with a streptavidin-functionalised single DNA strand. When the single strands associated to form a double helix, the enzymes were located next to each other with a high degree of specificity [116].

A few years later in 2005, there were two examples of DNA tile structures with a greater number of protein attachment sites. The first demonstrated that the programmability of DNA can enable a large, repeating array of proteins to be assembled, and again made use of a biotin-streptavidin interaction to hold the proteins in place [118].

The second example is the first all-DNA scaffold and cleverly creates periodic guanine-aptamer regions along the side of a triple-crossover tile [119]. The G-aptamer is unimolecular 15 base quadruplex that binds to the protein thrombin with nanomolar affinity [120]. Rinker *et al.* also use thrombin binding aptamers and investigate the effect of distance changes in binding via a pincer mechanism [121].



(a) Tile figure from Fu *et al.*[122]



(b) Bridge figure from Fu *et al.*[122]

Figure 19: HRP/GOx are assembled with a controllable inter-enzyme distance on an origami tile. The assemblies are characterised by atomic force microscopy (AFM). Upon introduction of a bridging molecule tethered between the two enzymes, an even greater enhancement in activity is reported [122].

The latter two examples highlight advances in protein-DNA binding techniques that paved the way for Wilner *et al.* to progress this field to involve more complex biomolecular structures and networks [56]. In this case the pioneering idea from Niemeyer is revisited and the reactivity of both a two-enzyme and enzyme-cofactor cascade is controlled. The enzymes glucose oxidase (GOx) and horseradish peroxidase (HRP), as well as the NAD^+ co-factor, are assembled on a novel hexagonal DNA scaffold. The spacing between the enzymes and/or cofactor were varied such that they were on adjacent interlocking hexagons (and thus separated by 13 nm) or tethered such that two bare hexagons were in between creating a larger spacing (corresponding to a 33 nm gap). An increase in reaction rate of 20% is reported for the adjacent set-up compared to when there is a larger separation. This increase is attributed to a mechanism analagous to that of the leaky channel described earlier, creating a lower local concentration at the second enzyme/cofactor when the separation is largest. However the author infers that there is no electrostatic channelling mechanism occurring here, after finding no increased reaction rate in the presence of the foreign calf thymus DNA, and that the effect is due to spacial organisation and proximity alone. This is rationalised in terms of an increase in flux of the diffusing intermediate on the second enzyme/cofactor, where distances are smaller compared with larger separations [56].

Following the report of synthetic protein scaffolds that spatially recruit metabolic enzymes *in vivo*, an RNA-based scaffold for the spatial organisation of *in vivo* bacterial metabolism was also engineered [123, 124]. This exciting revelation demonstrates increased hydrogen output as a function of scaffold architecture for one- and two-dimensional RNA constructs. Again, the authors portray flux con-

straints as being responsible for the increase and do not offer any further discussion on this point [124].

Gaining inspiration from nature, Fu *et al.* mimic a 'swing-arm' intermediate-channelling mechanism on a DNA tile [108, 122]. The two-enzyme cascade of HRP and GOx is used again here and the effects of both separation and bridging on the overall activity is explored. Similarly to Rinker *et al.*, an enhanced activity was reported for closely spaced enzymes (this time 10 nm was optimal). Following such a result the authors thought it possible to enhance the activity observed at longer interenzyme distances by placing a protein bridge between the enzymes. To test this, the 'bridge-based' cascade was assembled (as seen in Figure 19 b) in which a noncatalytic protein was inserted between GOx and HRP (which were 30 nm apart), with the aim of connecting the protein hydration shells and facilitating the surface-limitted diffusion of hydrogen peroxide. The inclusion of a non-catalytic molecular bridge or swing-arm produced an even greater enhancement. This led Fu to conclude that dimensionally limited diffusion of intermediates across connected protein surfaces contributed to the enhancement in activity for closely spaced GOx/HRP assemblies, which is in contrast to the findings of Rinker *et al.* [121, 122].

During the lifetime of this D.Phil, more architecturally complex cases have been divulged. These include a few designs to allow for the activation and deactivation of an enzyme cascade based around a reconfigurable DNA tweezers-type structure [125, 126]. In these assemblies, actuation of enzymatic function is achieved by switching the tweezers between open and closed states [125]. Fu *et al.*[1], introduced aspects of combinatorial assembly into their version of the HRP/GOx cascade. They progressed their scaffold from a tile to a DNA tube and described the observed activity enhancement as a re-

1 Yanming Fu, not Jinglin Fu of the swing-arm scaffold design!

sult of a synergy arising from both the stabilising effect of the DNA nanostructure and the enclosing effect of the confined nanotube [60].

Here we see that there is some disagreement as to the mechanisms responsible for enhancements reported with synthetic DNA scaffolded systems, just as there is disagreement about natural channelling mechanisms.

The most progress made in better understanding such processes, in relation to the field of DNA nanotechnology, has been by the Hess group. They explore the origins of activity enhancement through modelling and we shall go on to look at their findings in the following.

3.4 MODELLING ENZYME KINETICS

Perhaps the most well-known equation in biochemistry is that formulated and named after German biochemist Michaelis and Canadian physician Menten. The Michaelis-Menten model is based on their study of the enzyme invertase (Inv) which catalyses the hydrolysis of glucose into fructose and sucrose [127]. A substrate S may bind to an enzyme E forming an intermediate complex with it, I. The intermediate is subsequently converted to some product P and is released to leave the enzyme unaltered.

$$S + E \underset{k_-}{\overset{k_+}{\rightleftharpoons}} E \cdot I \underset{k_{rev}}{\overset{k_{cat}}{\rightleftharpoons}} P + E$$

The constants k_+, k_-, k_{cat} and k_{rev} express the rate constants for different stages of the reaction. Such a model allows the rate of product production, v, to be derived in terms of the substrate concentration:

Let V_0 be the initial velocity of the reaction,

$$V_0 = k_{cat}[E.I] \tag{1}$$

The maximum velocity occurs when all enzymes are saturated,

$$\begin{aligned} [E.I] &= [E]_0 \\ V_{max} &= k_{cat}[E]_0 \end{aligned} \tag{2}$$

In steady-state, the rate of formation of E.I equals the rate of consumption of E.I, where,

$$\begin{aligned} \text{Rate of formation of } E.I &= k_+[E][S], \\ \text{Rate of consumption of } E.I &= k_-[E.I] + k_{cat}[E.I] \end{aligned} \tag{3}$$

So in the steady state,

$$k_-[E.I] + k_{cat}[E.I] = k_+[E][S] \tag{4}$$

Rearranging slightly, we get,

$$(k_- + k_{cat})/k_+ = [E][S]/[E.I] \tag{5}$$

Grouping the kinetic constants together by introducing K_M,

$$\begin{aligned} K_M &= (k_- + k_{cat})/k_+, \\ \text{where } [E] &= [E]_0 - [E.I] \\ K_M &= ([E]_0 - [E.I])[S]/[E.I] \end{aligned} \tag{6}$$

Solving for $[E.I]$,

$$[E.I] = [E]_0[S]/(K_M + [S]) \tag{7}$$

Substituting equation 7 into equation 1, for V_0,

$$V_0 = \frac{d[P]}{dt} = V_{max}\frac{[S]}{K_M + [S]} = k_{cat}[E]_0\frac{[S]}{K_M + [S]} \tag{8}$$

The reaction rate asymptotically approaches V_{max} which is only reached when all substrate is enzyme-bound. The enzyme has initial concentration of $[E]_0$. K_M, the Michaelis constant, is the concentration of substrate at which the the rate is half V_{max}. The Michaelis-Menten model holds true under the following assumptions:

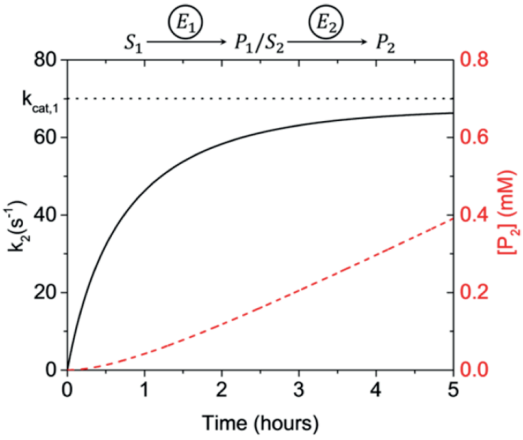
- Rapid equilibrium; equilibration must happen faster than the time scale on which the product is formed,
- Free ligands; free diffusion implies the law of mass-action is applicable,
- Steady-state; this assumes that the conversion of substrate to product is instantaneous .

The steady state of a cascade (scaffolded or otherwise) is determined by the V_{max} of the slowest enzyme [128]. It has been proposed however, that the presence of a scaffold increases the initial throughput of the cascade via the enchanced flux mechanism touched on ealier [122, 56, 129]. Idan and Hess test the extent of this proposal by simulating random walks of a substrate molecule starting at a distance d from a model-sphere enzyme placed at the centre of a spherical container [129]. The model finishes when the substrate hits either the wall of the container or the enzyme, and those paths that reach the enzyme (as opposed to the wall) are described as having taken a 'direct' path. For a cascade involving two or more enzymes, if the starting concentration of the substrate for the second enzyme is zero ($[S]_{2,0} = 0$) the flux from the direct path will dominate over the indirect flux for a short time. As $[S]_2$ increases however, the indirect path flux will dominate. The time at which this changeover occurs can be approximated as follows,

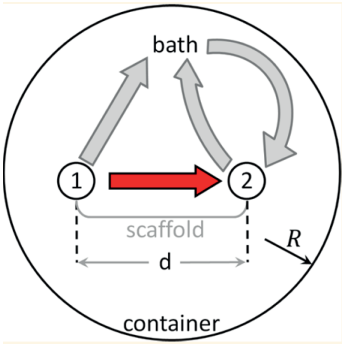
$$t \approx \frac{v}{4\pi D_{substrate} d} \tag{9}$$

the derivation of which is expressed in detail in the paper by Idan and Hess [129]. $D_{substrate}$ is the diffusion constant of the substrate. This approximation does not depend on the enzyme properties, but instead shows how a scaffold increases initial throughput on a time scale that depends on the volume of the container v, the inter-enzyme distance d and on the diffusivity of the substrates. The larger the container, the shorter the time scale over which the scaffold has an effect.

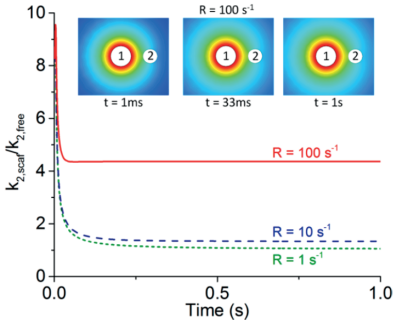
Idan and Hess derive three simple models that aim to explain the experimental results that have come before, some of which are described above. In their aggregation model they simulate a toroidal ribbon of enzymes (analagous to Wilner's ribbon structures [56]) where a high density of enzymes is approximated as one enzyme with a high turnover rate. It is found that a scaffold might be beneficial if only to provide a large number of target enzymes in the vicinity of the preceding enzyme. They also show that in the presence of a sequestering side reaction, the concentration of the intermediate substrate is only appreciable in a small volume around the first enzyme (see figure 20c). The ratio of reaction rates, $\frac{k_{2,scaff}}{k_{2,free}}$, increases with increased sequestration indicating that a scaffold can provide a sustained boost to throughput.



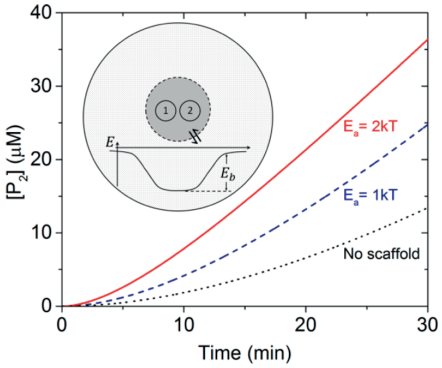
(a) Figure from Idan and Hess.[129]



(b) Figure from Idan and Hess.[129]



(c) Figure from Idan and Hess.[129]



(d) Figure from Idan and Hess.[129]

Figure 20: Idan and Hess' figures describe the enzyme cascade systems they investigated. Figures reprinted from 'Origins of activity enhancement in enzyme cascades on scaffolds', *ACS Nano.* 2013 7; (10), 8658-65. (a) Reaction rate and final product concentration of a cascade of two enzymes in the absence of a scaffold. Initially, the reaction rate (black solid line) increases linearly and the product concentration (red dashed line) quadratically with time. On a time scale of hours, the system reaches a steady state and the reaction rate reaches the maximum reaction rate (here limited by the constant production rate of enzyme 1) (b) The molecules released from enzyme 1 can be conceptually separated into a 'direct' stream toward enzyme 2 (red) and a stream into the surrounding bath (grey) (c) Reaction-diffusion simulations for enzyme pairs with a sequestering reaction. A higher sequestering reaction rate decreases the lifetime of the intermediate substrate. The concentration of the intermediate substrate is only significant in a small volume around the first enzyme (d) Simulations of the compartmentmodel for different barrier energies. For a spherical compartment with a 20 nm radius, an energy barrier on the order of a few kT increases production rate compared to a non-scaffold system. This increase is evident on time scales of minutes and hours. [129].

In their third and final model, they approximate compartmentalisation of a scaffold by imagining enclosing the enzymes within a barrier that is semi-permeable. The probability of a substrate molecule crossing the barrier is inversely proportionate to the probability of it instead being bound; the latter being expressed by the Arrhenius factor $e^{\frac{E_{bind}}{k_B T}}$, where $k_B T$ is a scaling factor for the energy in systems such as this and E_{bind} is the binding energy between the substrate and scaffold. From this model they could show that the scaffold compartmentalisation has the effect of increasing the local concentration of substrate molecules and thus increases the production rate of the second enzyme. Figure 20 shows the models in more detail.

The take-home message of this paper is that a scaffold may contribute to an enhanced production rate, but only over a certain (potentially short) time scale. It is therefore interesting to consider and recognise on which time scale future scaffolds have effect and to study the dynamics accordingly. Our aim is to construct a three-enzyme cascade by localising enzymes within our origami tubes and compare their effectiveness to a system with the enzyme localised on the outside of the tubes. The effectiveness of the systems should be analysed in light of the findings of the Hess group.

3.5 CONSTRUCTING A THREE-ENZYME CASCADE

Here we present the design and results from a 3-enzyme cascade localised inside and outside the 20H origami nanotube. The work here is a collaboration between myself and the Stulz group in Southampton University. The Stulz group were responsible for functionalising the enzymes with DNA strands which are complementary to extended staple regions on the DNA tube.

3.5.1 *Biochemistry*

As described in the previous section, many enzyme cascades presented previously have made use of the enzymes HRP and GOx, and thus the experimental procedures involving these are well understood. We set out to create a 3-enzyme cascade localised in and on a DNA tube and to observe any change in effects between configurations. For ease, we have chosen to use these enzymes in combination with Invertase (Inv), which can be added to the beginning of the cascade.

GOx is a globular dimeric protein and has dimensions 9.5 nm x 8.5 nm [130, 122]. Inv's dimensions are roughly 10.6 nm x 11.9 nm [131] and HRP's dimensions are roughly 5.7 nm x 3.7 nm [122]. All enzymes should therefore fit within the tube which has a diameter of approximately 20 nm.

Inv catalyses the hydrolysis of sucrose to glucose. Glucose is the substrate for GOx which catalyzes the oxidation of β-D-glucose into D-glucono-1,5-lactone, which subsequently hydrolyzes to gluconic acid. The glucose oxidation step also produces hydrogen peroxide (H_2O_2). This acts as a substrate for HRP and mediates the reduction of blue-green 2'2'-azino-bis[3-ethylbenzthiazoline-6-sulphonic-acid] ($ABTS^{2-}$) to lilac $ABTS^{-}$, $\lambda \approx 420$ nm [56, 112]. A diagram of the simplified biochemistry is depicted in Figure 21. Therefore observing the colorimetric change by monitoring absorbance at $\lambda = 420$ nm wavelength, we indirectly probe the reaction progress.

The starting substrates required are ABTS solution, sucrose solution and the enzymes Inv, GOx and HRP. The substrates are prepared in phosphate buffered saline (PBS).

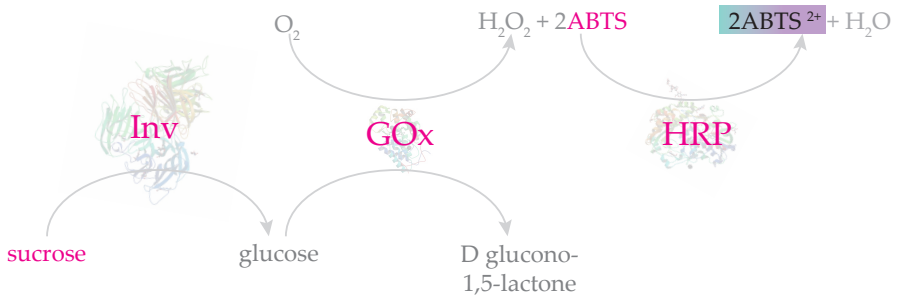


Figure 21: 3 enzyme cascade reaction. The pink labels indicate the substances needed at the start of the cascade, which are included in the cuvette at the beginning of the reaction. The reaction diagram is not comprehensive, but includes relevant substances. Invertase catalyses the hydrolysis of sucrose to glucose and fructose. Glucose oxidase catalyzes the oxidation of β-D-glucose into D-glucono-1,5-lactone, which then hydrolyzes to gluconic acid.

3.5.2 *Enzyme-DNA conjugation*

To attach the enzymes to the origami tube, they are functionalised with 20 base strands which are complementary to unique staple extensions protruding from the origami tube. Initially the DNA strands must be attached to their respective enzymes. The DNA is first modified to include a thiol at the 3′ end. The 3′-thiol-modified single stranded oligonucleotide (10 equiv) in 0.1 M PBS, (pH 8.0, 100 µL) is mixed with 0.2 M 1.4-dithiothreitol (10 µL) and incubated for 3 hours. Excess DTT is removed by gel-filtration chromatography using Nap Column (GE Healthcare). Enzyme (1 equiv) is dissolved in 0.1 M PBS buffer (pH 7.4, 250 µL) and incubated for 1 hour with N-[ε-maleimidocaproyloxy]succinimide ester (sulfo-EMCS, 6 equiv) (Thermo Fisher Scientific). Excess sulfo-EMCS is removed by washing, and is filtered using an Amicon 0.5 filter (10 or 30 kDa). The purified DNA and enzyme solutions are combined and incubated for 2 hours at room temperature in PBS. The conjugates are purified from excess DNA on a size exclusion column (Superdex75 10/300) during high performance liquid chromatography (HPLC), in a buffer of 50

mM HEPES, 75 mM NaCl, pH 7.5. To check the DNA-enzyme conjugates are correctly formed, they are run on a Tris-acetate gel for 50 mins at 150 V, shown in Figure 22. There is a slight shift in the band to a slower running position when the DNA is attached compared to when the enzyme is free.

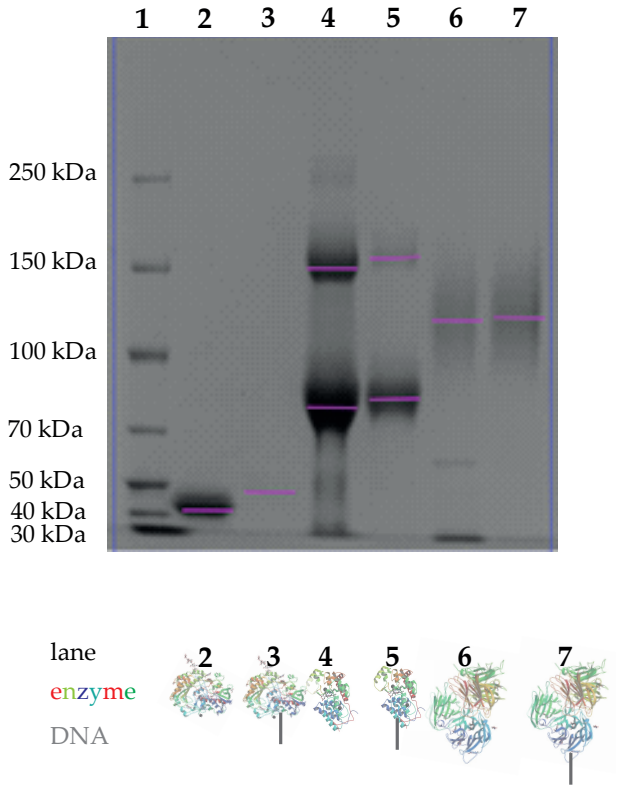


Figure 22: Enzyme-DNA conjugation gel. This gel was made and run by my collaborator Gabriella Marth at Southampton University. The pink lines indicate the band position from the peak of a fitted curve. NuPAGE Tris-Acetate gel of enzyme-DNA conjugates. Lane 1 marker, lane 2 HRP, lane 3 HRP-DNA, lane 4 GOx, lane 5 GOx-DNA, lane 6 Inv, lane 7 Inv-DNA. Denaturing conditions: 3-8% gel with a constant voltage of 150 V for 50 mins.

3.5.3 *Assembly of components*

The DNA nanotubes are assembled as described in Chapter 2 by thermal annealing. The staples at the end of the tube are the set that prevent end-to-end stacking by the inclusion of poly-T tails. Three

staples in the tube are replaced with staples that have 3′ extensions of 20 unique bases with four extra unpaired thymine bases close to the origami surface to provide flexibility. The extensions are complementary to the oligos which are conjugated to the enzymes and are such that only Inv-DNA binds to the first extension, only GOx-DNA binds to the second and only HRP-DNA to the third. These three staples are separated by approximately 11 nm each.

In addition, the same tube configuration is assembled but with the enzymes staples positioned on the outside of the tube as opposed to the inside. This is achieved by moving the nick position of the three staples and their respective neighbours by half a turn each. This positions the extensions on the outside of the tube such that when the enzyme-DNA conjugates are added, they are attached to the outside.

To form the control version, the tube is also assembled without any staple extensions. The enzyme-DNA conjugates are added in the same ratio and are thus present in solution, but are not specifically attached to the tube.

After origami assembly, all the tubes (inside, outside, control) are purified using 3 x micro-biospin columns packed with Sepharcryl S-300, as described previously in Chapter 2.

Following purification, the concentration of each tube type is found by measuring the absorbance of DNA at 260 nm in a UV-vis spectrometer.

GOx-DNA is mixed with the DNA origami tubes in a ratio of 1:3, tube:enzyme-DNA, in a total volume of 90 µl in 1 x tube buffer (pH 8.3) and annealed from 37 °C to 4 °C. HRP-DNA and Inv-DNA are then mixed with the tube (again in a ratio of 1:3, tube:enzyme-DNA) and the solution (total volume now 100 µl) is annealed from 37 °C to 4 °C over a gradient of -1 °C per minute in a thermal cycler. GOx is added to the tubes first due to its position as the middle enzyme in

the cascade; it may otherwise be blocked from getting into position by either Inv or HRP.

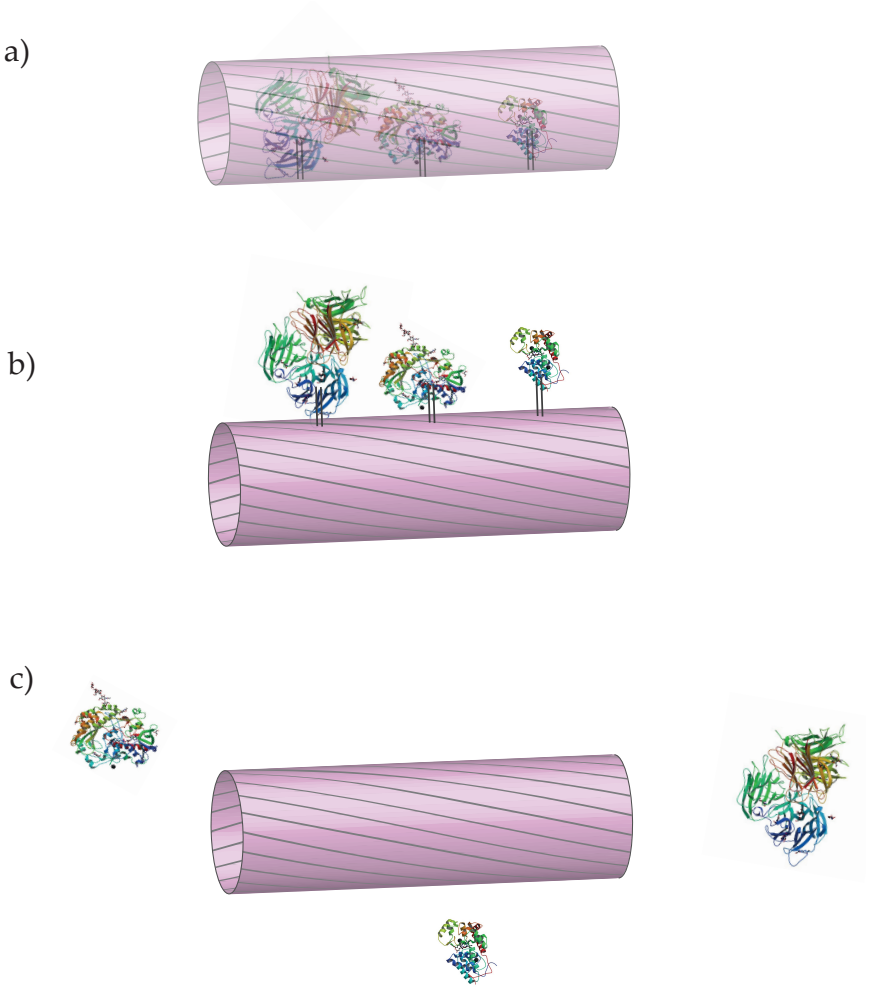


Figure 23: Enzyme positions for different tube cascade set-ups. a) The enzymes are positioned inside the tube with a complementary staple binding to the DNA strand on the DNA-enzyme conjugate. In b) the enzymes are positioned on the outside of the tube and are held on the tube in the same manner as before. In c) the enzymes are not attached to the tube because the staples do not have the complementary extension.

3.6 CHARACTERISATION

3.6.1 *AFM*

The tube and enzyme assemblies are visualised using AFM to check their formation before assays are run.

AFM images of the tubes with different enzyme combinations inside are taken. Upon attempting AFM of the three enzymes in their positions in the centre of the tube, it was not possible to separate them and they appeared as one bright spot, and it proved difficult to differentiate between bound enzymes inside and unspecific 'blobs'. Instead, the 20 bp attachment regions for the enzymes are moved apart to different staples, purely for the purpose of imaging, so that one is at the LH end, one in the middle and one at the RH end of the tube.

In Figure 24 (top) only HRP is included at the LH end. It is possible to see in the cross-sectional traces in b (top) that there is a bump of around 5 nm which may correspond to the size of HRP with a DNA origami layer, where HRP is 5.7 nm x 3.7 nm.

Figure 24 a (bottom) shows an example of the tube that has been opened up from repeated imaging with the AFM tip, with three enzymes present. It is possible to see that both Inv and GOx (in the green and red circles respectively) are slightly larger than HRP (in the blue circle), as expected.

Figure 25 top and bottom show another example of a tube with just HRP attached at one end.

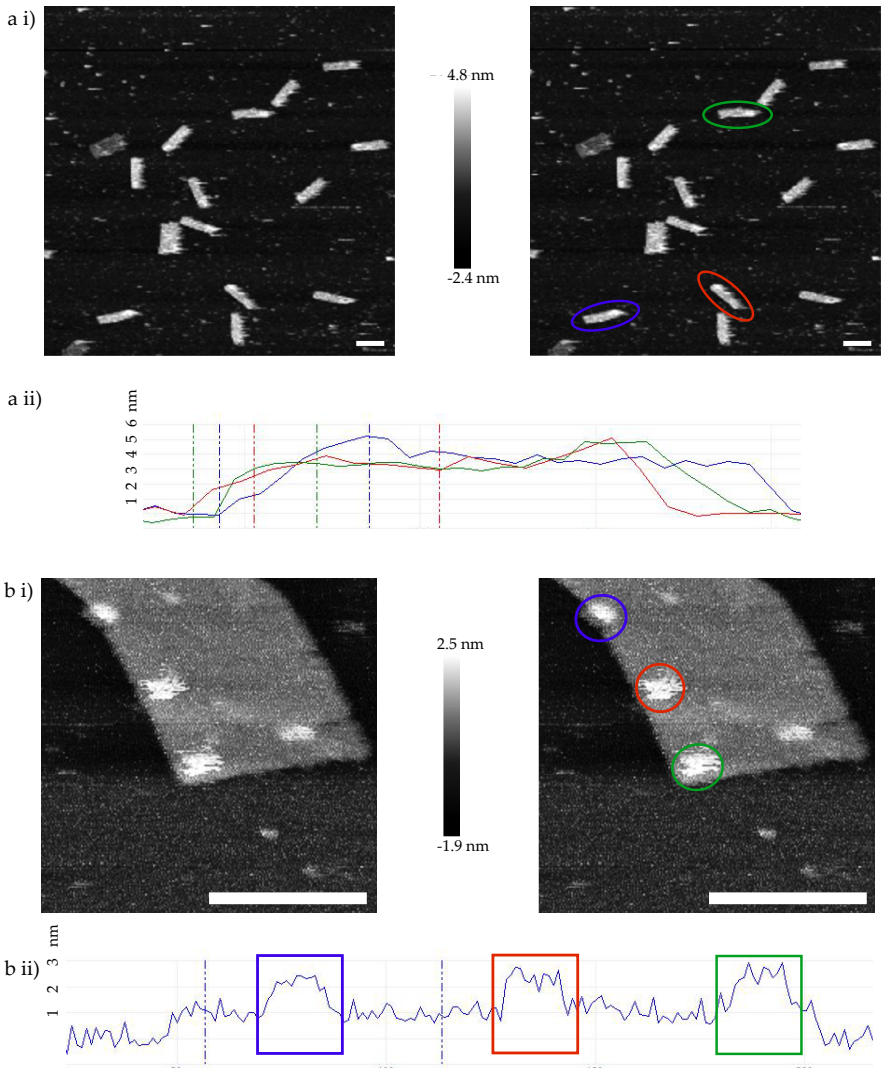


Figure 24: AFM images of enzymes inside the DNA nanotube. a i) HRP is included at the LH end of the tube. The scale bar is 100 nm, b i) Three enzymes present in the separated-staple imaging version so that the enzymes are clear and distinct. The enzymes are 30 nm apart and the first enzyme is located at the tube end. The tube has been opened up due to repeated imaging in a single place. The scale bar is 60 nm. Both a ii and b ii show the cross-sections of the highlighted tubes (for a) and enzymes (for b) drawn in blue, red and green. Where only HRP is present we see a bump at one end of the tube of around 5 nm in height, which could correspond to a double origami layer plus the short axis of HRP. Where the three enzymes are present the cross-sections appear quite similar in size but are clearly larger features than the single origami layer.

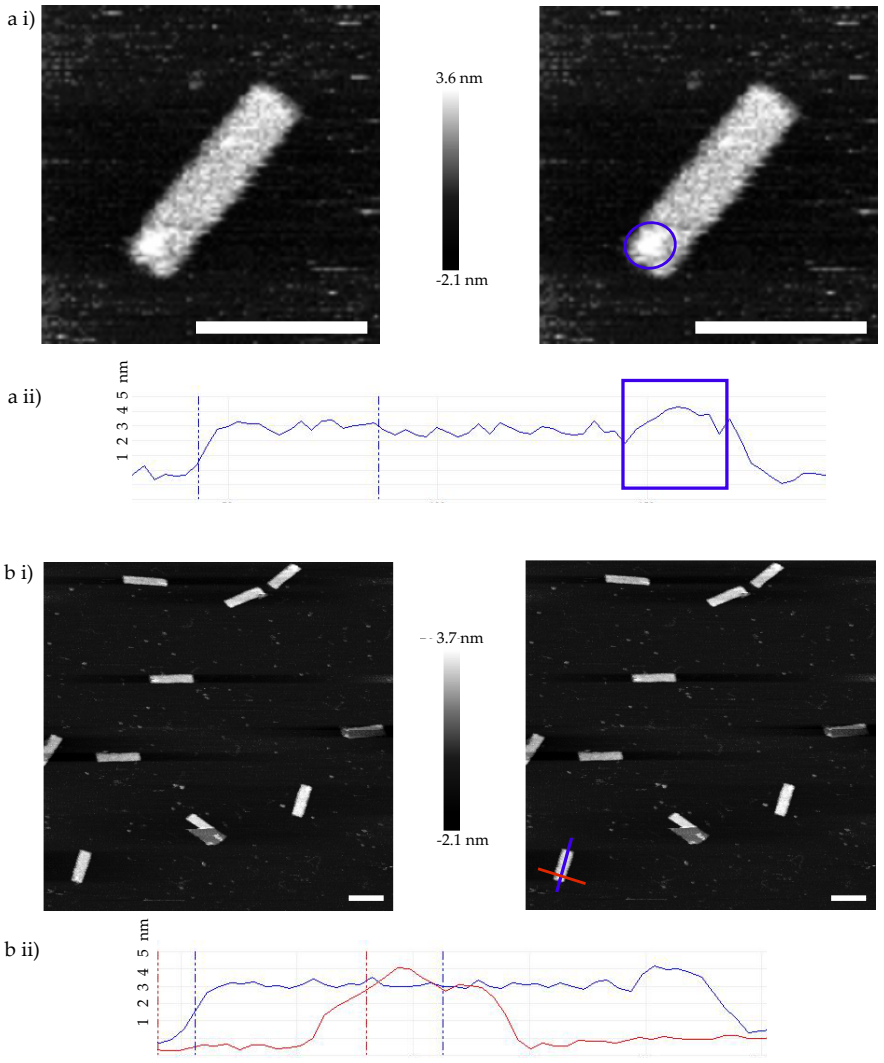


Figure 25: Further AFM images of enzymes inside the DNA nanotube. The AFM scale bar is 100 nm in both a i and b i. Both: show an example of a tube with just HRP attached. The cross-sections so either a highlighted enzyme bump in (a ii) or two examples of tube cross-sections perpendicular to each other (b ii).

3.7 RESULTS

Following conformation of enzymes inside tubes on the AFM, we may now proceed with the cascade experiments.

3.7.1 *Assays*

The enzyme-DNA+origami assemblies are added to 300 µl 20 mM ABTS solution and 1 mL PBS. Just before the reaction is ready to run, this mixture is added to 300 µl 113 mM sucrose solution in a quartz cuvette. The reaction concentrations are thus 3.5 mM ABTS and 19.9 mM sucrose. There are 0.2 pmol DNA tube per sample and 0.6 pmol enzyme-DNA conjugates. The absorbance at λ = 420 nm is monitored for changes in absorbance of ABTS.

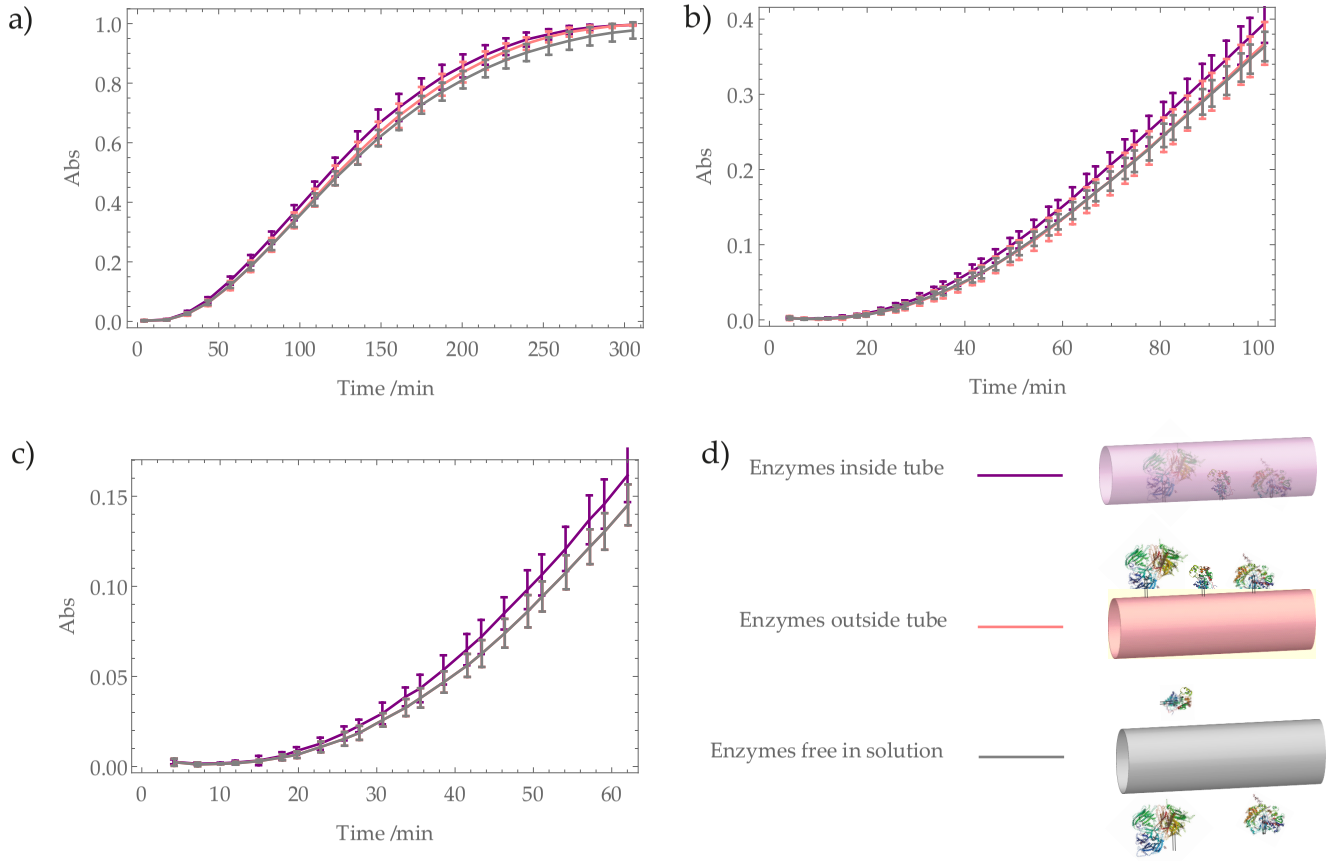


Figure 26: Enzyme cascade absorbance changes. The purple lines indicate enzymes attached on the inside of the tube, the pink lines indicate enzymes attached on the outside and the grey lines indicate enzymes not attached to the tube, which are instead free in solution. a-c) The data is the average of three repeats and the absorbance values are normalised from zero and one. The samples have 0.2 pmol origami tube, 0.6 pmol enzyme-DNA, the sucrose concentration is 19.9 mM and the ABTS is 3.53 mM. The total reaction volume is 1.7 mL. The error bars are the are the coefficient of variation, found by taking the standard deviation of the three repeats. a) shows the entire time frame, b) shows the first 100 mins and c) shows the initial 60 mins. Figure d is the legend and shows the assemblies for reference.

All of the assays carried out feature the same initial lag before reaching steady-state. Looking at modelling enzyme kinetics in section 3.4, one may expect the initial reaction rate to be proportional to the starting substrate concentration, for a simple 1-enzyme system, where the substrate flow is not spatially confined. Such a model however assumes the enzyme is operating in its optimal environment. However, this is not the case for Inv in our set-up, and all three of the enzymes are tethered too. Inv favours an acidic environment but here the reaction takes places in slightly alkaline PBS (pH 7.4). This may account for the lag at the start of the reactions, where slow operation of Inv creates a bottle-neck beyond which substrate for GOx must first accumulate before this second enzyme may proceed. We monitor the color change of ABTS produced in the final step of the cascade, and because Inv is slow, we measure its rate of converting sucrose to glucose, where HRP and GOx are faster.

For each experiment that is run, a family of three curves is produced; a tube with enzymes attached inside, a tube with enzymes attached outside, and a control, with enzymes free in solution with the tubes. In each experiment the amounts of enzyme-DNA conjugate, origami, sucrose and ABTS added are the same for each cuvette and are made from the same stocks and divided up.

Figure 26 shows three curves where the origami tube amount is 0.2 pmol and the DNA-enzyme conjugate is 0.6 pmol, in all instances. The curves are averages of three repeats for the same experimental conditions and they are normalised to the range zero to one. The curves in Figure 26 a, b and c show the averages over different time scales. They are plotted together with their respective error bars which are the standard deviation.

There is a very slight difference between the three curves. The experiments always display the same behaviour even though the differ-

ence is slight: where the enzymes are tethered inside the tubes, the lag before steady state is always slightly shorter- especially noticeable when looking over the shorter time-scales - than for enzymes tethered on the outside of the tube or for enzymes in solution with the tubes. Bearing in mind Hess' discussion, it is worth considering the initial throughput of the cascade on its own, where Hess suggests a scaffold only has a beneficial effect in the early stages of the assay, when the flux from the direct path will dominate the contribution of the flux from the bulk.

For the outside tethered enzymes and the control, over the first 60 mins the lags appear the same and the curves are on top of each other. In this time frame the two curves are both most distinct from the curve for enzymes inside the tube, although again the differences are slight. The enclosed space of the tube has a *very* small benefit in reducing lag time over the enzymes being localised on the outside of the tube and in solution.

The benefit we see from our tube is much smaller compared to the benefit recorded by previous scaffolds in literature (such as Fu's and Wilner's [122, 56]). In previous systems only two enzymes, HRP and GOx, are tethered in a cascade but here we introduce a third, Inv. As mentioned above, Inv is not operating at its optimal rate and thus its activity is slowed down, contributing to the lag in the kinetics. In previous two-enzyme cascades involving these enzymes, HRP is the slower enzyme but the lag corresponds to accumulation of hydrogen peroxide produced by GOx. But by introducing Inv (in unfavorable conditions) into the system we slow the initial reaction down and change the limiting step. Therefore it may be possible that by introducing Inv we change the kinetics of the system such that we cannot compare between previous results and ours. Previous experiments also used different concentrations of starting substrate,

origami and enzyme and in all cases the previous experiments used higher amounts - for example, Fu *et al.* use 1 nM tiles and we use 0.1 nM tubes. It would have been beneficial for us to have tested the same experimental conditions as previous experiments with two enzymes, and to characterise the two-enzyme cascade for our system thoroughly, before having moved on to the three enzyme cascade. Then it may be possible to see if the difference in our results was due to dynamics intrinsic to the extended cascade or to reaction conditions.

Our experiment is sensitive to channelling of both hydrogen peroxide and glucose, which are the inter-enzyme substrates. We cannot rely on the lag before steady-state being entirely due to slow Inv: it is also possible that glucose is not channelled but hydrogen peroxide is, in which case a lag due to the slow accumulation of glucose in bulk solution would effect all experiments approximately equally. It is necessary to conduct more experiments here, introducing competing side reactions for both hydrogen peroxide and glucose, to determine if and where channelling is occurring. It would be advantageous to deliberately disrupt any channelling by introducing a competing side reaction which competes for either glucose or hydrogen peroxide. If the reduced lag time is due partly to surface interactions of one substrate or the other then the presence of a competing side-reaction should have less of an effect on enzymes free in solution as it does on enzymes attached to tubes, when comparing to the same set-up without a competing side reaction. Otherwise, the differences between enzymes in solution and enzymes on tubes in the presence and in the absence of a side-reaction should be similar between both cases.

It may be that the DNA scaffold creates a virtual compartment due to weak attractions between polar substrate molecules and charged scaffold and that this is greater when the enzymes are entirely en-

closed by the surface of a tube rather than just sitting on the surface. This would mean substrates were not lost to the bulk inside the tube as readily as if the enzymes were on the outside of the tube. Such attractions may result in this slight initial throughput increase. This is in accordance with Hess' discussion where lag time is reduced due to increased time in proximity of the enzymes, brought about by attractions between substrates and scaffold.

3.8 FURTHER WORK

It would be beneficial to run more repeats of this experiment to reduce the size of the error bars and determine further how significant the small differences between scaffold geometries and free-enzyme control are. As the error bars are standard deviations, we cannot regard overlapping error bars as being statistically insignificant, yet we also cannot regard no overlap as being statistically significant. Thus it would be beneficial to determine the significance. Experiments at differing enzymes and origami concentrations were run although no repeats were achieved for each different concentration and so is not used in this discussion. It would be highly beneficial to investigate further the effect of changing the enzyme and origami amount and see if the trend follows what may be indicated by our results.

Originally it was hoped that the same assays could be run with a 'waterproofed' version of the origami tube. It would be interesting to see how waterproofing changes the scaffold properties. It may be possible to use Lipofectamine 3000 (a lipid-based transfection agent) to coat the tube non-specifically, or a more sophisticated approach would be to tether cholesterol molecules to the outside of the tube and wrap the origami in a membrane. With our experiment, it would have been possible for substrate molecules to pass through the holes in the

lattice of the DNA origami and so the tube wall is not a perfect boundary. Our set-up does however provide charged surfaces to which substrates may be attracted. Thus it would be of interest to see if, with a 'waterproof' wall, the efficacy of the cascade is bettered, or perhaps worsened if enhancements are due to surface charge-substrate attractions.

In order to determine the underlying physics of why initial lag times are reduced with scaffolds, particularly whether it is due to charged scaffold-substrate attractions, it would be advantageous to deliberately disrupt the process by introducing a competing side reaction, as briefly mentioned in the previous section, which competes for intermediate substrates. If the reduced lag time is due partly to surface interactions then the presence of a competing side-reaction in bulk should have less of an effect on enzymes free in solution as it does on enzymes attached to tubes, when comparing to the same set-up without a competing side reaction. Otherwise, the differences between enzymes in solution and enzymes on tubes in the presence and in the absence of a side-reaction should be similar between both cases.

Information could be gained from comparing different origami scaffolds, for example tubes of varying diameter and tiles. There are also many variations of just enzymes and tube topology that would be interesting to investigate. For example, to compare activities of set-ups where the first and the last enzyme are tethered to the tube and the middle enzyme is free in solution, and all combinatorial arrangements of such, including switching the orders of enzymes in the cascade. It would also be of interest to make a tube where the scaffold selected for different orders of enzymes, i.e. HRP, GOx, Inv or Inv, GOx, HRP etc. This would create a record of which enzyme order

had been assembled which could be selected for, read via sequencing and amplified by transformation.

3.9 CONCLUSIONS

In conclusion, our results show a very small difference between curves for enzymes tethered inside tubes, outside tubes and enzymes free in solution with tubes, indicating that enzymes tethered inside an origami tube provides a small benefit in reducing lag time. For all experiments individually (sets of three curves; inside, outside, control), the qualitative trend is always the same, pointing towards the enzymes tethered inside the tube giving rise to a *very slight* initial increase in activity, which is in accordance with Hess' discussion where scaffolds may provide attractive surfaces for charged substrates. There is plenty of room for further investigation of these assemblies, some of which have been outlined in the previous section.

4

MOTORS IN TUBES

In this chapter we explore the motion of a DNA motor along a DNA track which is built into the interior surface of the origami nanotube. We use both ensemble measurements and direct microscopic observation.

4.1 TRANSPORT ON THE NANOSCALE

Biology contains a realm of processes which require cells to have miniature motors. Mini-motors are involved in tasks such as protein folding and unfolding, transportation of organelles and molecules (including creating gradients to enable transport), and in cellular movement [132].

Unlike the conventional motors we meet day-to-day, on the cellular scale Brownian motion is prominent and random thermal fluctuations have a role to play in influencing molecular motor movement. As Oster and Wang put it, at such small length scales 'You would not need to even pedal your bicycle: you would simply attach a ratchet to the wheel preventing it from going backwards and shake yourself forward' [133]. In a thought experiment by Feynman however, and later rigorously confirmed by Magnasco, a Brownian ratchet[1] fails because the lever, preventing the ratchet turning backwards, will also undergo Brownian motion and thus is displaced backwards as often

1 The Brownian ratchet imagines a ratchet attached to a paddle, which is small enough to be driven by random thermal fluctuations in a bath of constant temperature. A lever prevents the ratchet from turning backwards and thus may be thought of as a perpetual motion machine.

as it is driven forwards [134, 135]. This highlights the importance of the part of the phrase above 'shake **yourself** forwards': although thermal fluctuations may assist in movement of components, it is not enough to just rely on them as a driving force - another source of energy is required. This comes in the form of breaking intermolecular bonds which allows chemical energy to be converted to work.

In the case of the commonly known linear molecular motors kinesin and dynein (which transport cargoes by walking along microtubules in cells) and nucleic acid motors such as DNA and RNA polymerases, the hydrolysis of nucleoside triphosphates (NTPs) to nucleoside diphosphates (NDPs) provides driving energy.

Synthetic analogues of linear motors have been created, including those made entirely from DNA operating on nucleic acid tracks, but their energy source comes from breaking base-pairs, from phosphodiester backbone hydrolysis or from both ATP and DNA hydrolysis [136, 137, 138, 139, 44, 140, 141, 142]. In some all-DNA systems, the track is made from single strands which act as individual steps or 'stators' for the oligo motor strand to traverse. These may be linked and anchored via another oligo, an origami tile or indeed operate free in solution [141]. In 2004, two bipedel motors were presented on such tracks: one operated analogously to an 'inch-worm' and the other demonstrated the 'hand-over-hand' mechanism by which kinesin walks [143, 144].

In 2005, a single-footed motor termed the 'burnt-bridges' motor was demonstrated by Bath *et al.* and is the walker design that shall be used in this chapter [138]. This system enforces unidirectionality by permanently altering the track after the motor has passed through. When the motor is bound to a stator the duplex formed contains a short sequence which is recognised by a restriction endonuclease. In the presence of such an enzyme the stator strand is nicked, leav-

ing an 8-nucleotide segment to melt off and diffuse away, exposing the complementary single-stranded region of the motor. This energetically unfavourable exposure initiates stepping of the motor to the next fully intact stator, where the process repeats [138]. The mechanism for movement is illustrated in Figure 27.



Figure 27: The burnt bridges motor. The motor stepping is initiated by additon of an enzyme which cuts the stator at the restriction site. The short top diffuses away and exposes a toe-hold for strand displacement. The motor finds the next stator via a random walk and unzips from the stator it is on. This process repeats along the track.

4.2 AIM

In this chapter we aim to assemble and observe the burnt-bridges motor walking inside the 20-helix DNA nanotube. We also hope to provide evidence of the motor crossing between origami tubes that are stacked end-to-end as a means towards achieving long-range motion. Such a system would enable observation of controlled motion within a confined space, analogous to the cellular environment.

To observe such an assembly, one may approach it in two ways: either by directly monitoring the motor motion, through carriage of a fluorescent cargo, or by indirectly observing the motor by watching the effect it has on the labelled track as it passes along the nanotube. In an ideal situation we would observe the motor directly in real-time using TIRF microscopy. To prepare for this, it is necessary to check simple motor movement within a tube; whether crossing

between origamis is feasible; whether the motor can move down a double stranded track; and whether it can operate with a bulky cargo attached. These initial tests are performed using fluorimetric ensemble measurements in solution. We then move on to test the operating conditions in the TIRF microscopy environment and attempt to test the initial motor set-up. Indirect observation along the track is then attempted followed by an initial attempt at direct observation using TIRF.

The work in this chapter was a collaboration with Alexandra (Alex) Lucas and Natalie Haley. The strategy for this project and the majority of the experiments were made jointly. However, some parts of the project were specifically accomplished by one person as follows. The tube was designed solely by myself prior to this project (and is described in Chapter 2). The assembly of the tubes for the experiments was done by myself, and I provided the AFM images of the structures. Schematics of the 3D tubes are from my original images and in some places have been adapted by Alex Lucas. Natalie Haley provided the microscopy expertise on the microscope. She was my experimental partner when Alex left. Natalie also ran the gels that tested enzyme cutting in the microscopy conditions. Alex Lucas analysed the microscopy images and wrote Mathematica programs, under advice from myself and Natalie Haley. I adapted and built on the line drawings from a drawing Alex Lucas originally made.

4.2.1 *Track design*

The stators for the burnt-bridges motor are anchored to the inside of the nanotube. As explained in Chapter 2, the position of the nicks is such that they point towards the inside of the nanotube when it is assembled. Stators are extensions of the staples and therefore they

appear at the nick positions. Consequently, all stators should be on the inside of the tube. However, it is known that long single strands of DNA, when attached to origami, may indeed thread through the holes present in the lattice structure from one side of the origami to the other [145]. Recent research by Alex Lucas in our group suggests a possible strategy to prevent molecular threading on origamis which we have applied to our stators here [146]. The base of the stator closest to the origami surface is made double stranded by extending the neighbouring staple at the nick position so that it forms the stator complement for 8 bps. This stator is termed 'braced' and through bracing single strands by making them double-stranded, the probability of threading is thought to be reduced due to their reduced flexibility. The braced staple bases are depicted in Figures 27 and 28 as grey dsDNA regions.

There are 21 stators per tube and they are arranged as illustrated in Figure 28 b. The path defined by the stators is chiral, as a result of the chirality of the tube, and has a total pitch of one helical turn per tube, which can be seen in Figure 28 c. This means that when the tubes are stacked end-to-end the final stator on the first tube lines up with the first stator on the next tube, and so on. The gap between each stator, known as the step size, is approximately 6 nm but is larger for an inter-tube step, where it is 12 nm.

In different experiments, stators are prepared in slightly differing conditions, with some parts double stranded and others single stranded, but each case will be described explicitly where appropriate. The stator comprises three regions which may or may not be utilised: the bracing region at the base (seen in grey in Figures 28 and 29), the 24 nucleotide motor region above (seen in green) and a 20 nucleotide region for a fluorophore-functionalised strand to bind to (in brown in Figure 29). The bracing and motor, and the motor and fluorophore

regions, are separated by two unpaired A bases and one unpaired T base respectively, providing enough flexibility to allow the motor to reach across and step to the next stator.

The motor is forced to start its journey from the first stator in a tube by annealing the first stator and its motor-complement together, separately from the tube assembly. The ratop of the motor:stator here is 0.95:1, so that there are no excess motors in solution and therefore motors binding to non-start stators should not be a problem. The tube is assembled in the absence of the initial stator-staple, which is feasible due to the robustness of origami assembly. This gap is then filled post-assembly by adding in the annealed complex to the tube mix, a method employed and shown to work by Wickham *et al.* [44]. The initial stator should only fill the gap that is left for it. This ensures that there is one motor per tube and that the motor starts at the first stator. The effects of diffusion could mean that the motors dissociate from the initial stators and bind to a stator further along the track, dissociated motors in these instances would start in various locations along the tube, but the method described reduces this likelihood. It has also been shown that dissociation and diffusion of a motor from the start stator to a stator further down the track is a very slow process compared with enzyme-driven motor motion and therefore we do not need to test this again here[44]. The final stator on the track is indicated as such by the presence of a sequence mismatch in the restriction site of the motor region. Consequently, when the motor binds to this stator, the enzyme does not recognise a restriction site and so the motor remains bound and stationary.

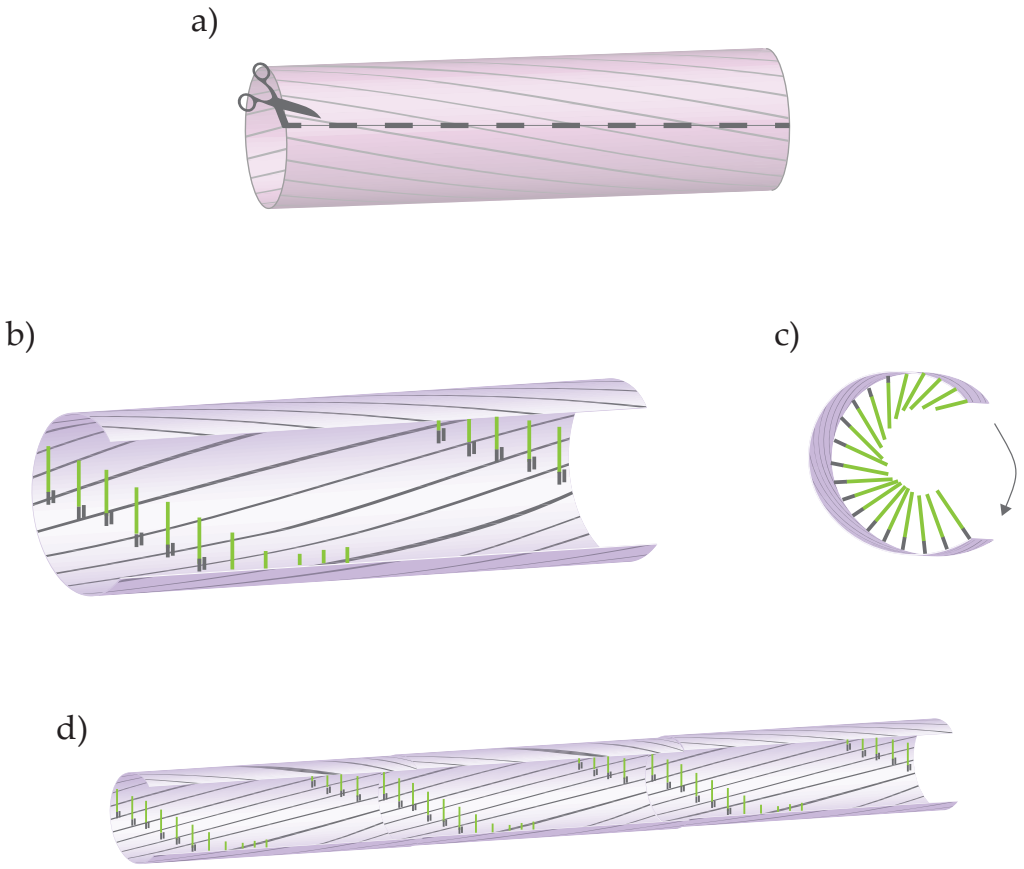


Figure 28: The position of the track inside the chiral nanotube. The stators are single stranded and therefore are flexible when the motor is not bound. (a) If we imagine cutting the tube open, we would see a track assembled as is shown in (b). (c) shows the stators on the inside of the tube. In (d) we see how the track lines up between tubes that are stacked end-to-end, to allow continuous movement.

4.2.2 *Track secondary structure*

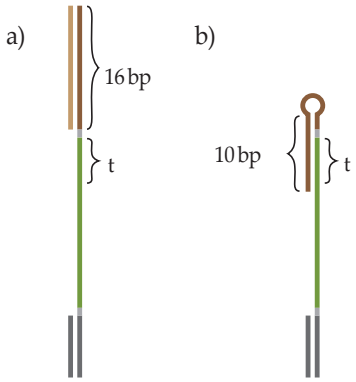


Figure 29: Schematics of a mid-track stator showing where the undesirable secondary structure arises. The stator has a top extension for fluorophore-strand binding. a) The most stable structure when the complementary strand for the extension is included is that which we desire. b) The undesirable structure. The stator creates a hairpin that covers the 6 base toehold (t) for the motor to initiate strand displacement with.

Following completion of experiments in this stage of the joint project, we discovered the presence of undesirable secondary structure in the stators. The top of the fluorophore-strand binding region complements part of the motor binding region on the stator, which makes it possible to create a hairpin structure comprising a 10 bp stem. The result of such a structure is likely to make motor progress slower than if it were not present: The stem must first melt away before the motor can step to the next stator. The probabilities of each structural arrangement were probed using NuPack [147] (a programme which aids in design and analysis of nucleic acid structures), and the correct motor-bound configuration is favoured over the hairpin loop (the free energy of secondary structure with all strands present is -38.4 kcal/mol, and is -15.0 kcal/mol in the absence of the complement to the top region, which shows that even in the absence of the complement, the bad secondary structure is much less likely to form). Figure 29 shows the two possible configurations of the stator.

4.3 FLUORIMETRY; GETTING FROM END TO END

In this section we test the motor movement within the tube and attempt to answer the questions, "Is crossing between origamis with a 12 nm step feasible?", "Can the motor can move down a double stranded track?", "Will it operate with a bulky fluorescent cargo attached?", "Does the enzyme diffuse into the tube to enable motion in the first place?".

Here we use fluorimetry to probe ensemble averages of the motor motion within tubes. At this stage, this is a simple experimental set-up with which to establish the basic behaviour of the system, before progressing onto single-molecule observations.

4.3.1 *Assembly of monomer and dimer tube tracks*

In this section we work with monomers and dimers of tubes.

Monomeric tubes are formed as described in Chapter 2, by including end-staples in the tube mix with poly-T tails to suppress stacking. In this section, all monomeric tubes have five stator steps, A_0-A_4. The tube is first assembled in the absence of the initial stator A_0.

Dimers are formed from what will be termed tube A and tube B, where A is the first tube from which the motor starts walking, and B follows A. A is assembled with only the LH end poly-T tail staples and not the RH end poly-T tail staples, to prevent stacking at the LH end and allow for stacking at the RH end with tube B. The tube is also assembled in the absence of the initial stator A_0. When fully assembled, this A tube has 21 stator steps going the full length of the tube, A_0-A_{21}. Here, the B tube only has five steps (B_0-B_4), such that when it attached to tube A in the dimer, there are a total of 26 steps for the motor to traverse. B is assembled with the RH end poly-T tail

staples to suppress stacking from that end and allow stacking with the A tube at the LH end.

For the fluorimetric measurements, the tubes are all annealed at 54 nM in 1 x NEB 2.1 buffer with a 3 x excess of staples, a 5 x excess of stators and a 6 x excess of fluorophore strands, as and when required (to allow for monitoring of motion). Any linker staples, necessary to join tubes A and B in the dimer configuration, or poly(dT)-end staples for suppressing stacking in all cases, are added in a 5 x excess to the tube.

The motor, M, is pre-annealed with A_0 (to make A_0.M), the initial fluorophore F_0 and any other additional strands unique to the experiment, by heating to 95°C and cooling to 20°C at a rate of 7.5 °C/min (see each section for the specific details of each experiment). A_0 is at 5 μM in 1 x NEB 2.1 buffer. A, B and the pre-annealed A_0.M complex are all individually purified.

For the monomer, the pre-annealed complex is loaded at the beginning of the track by incubating it at 37 °C for 1 hr with the purified A tubes, which are missing the staple for A_0, in a ratio of 0.95:1, A_0:tube. For the dimer, the pre-annealed A_0.M complex is added to purified A tubes (with a stators A_1-A_{21}) and purified B tubes, in an A:B ratio of 1:1, and incubated for 1 hr at 37 °C, with 5 x linker staples. The linker staples are designed to join the tubes in a specific register so that the stators line up correctly between tubes. The final dimer structure is then purified again to remove excess linking staples. The purification method is the same as previously described via three passes through S-300 Sepharcryl-packed microbiospin columns. Examples of monomer and dimer designs are shown in Figure 30.

Before experiments may begin, the purified assemblies must have KOAc added to a final concentration of 50 mM to enable operation of the enzyme.

4.3.2 *Results and discussion*

4.3.2.1 *Basic motion inside tubes*

We first examine the motor motion on an A tube comprising five stator steps in total and compare it to the motor walking 26 steps; 21 in tube A and then a further five in the B tube. The first stator, A_0 has its fluorophore-complement region labelled with strand F_0 which is functionalised with 5′ TET fluorophore. The final stator, A_4 for the monomer and B_4 for the dimer, has its fluorophore-complement region labelled with strand F_{end} which is functionalised with 5′ Cy5 fluorophore - the set-ups are illustrated in Figure 30. The motor strand for this experiment (M_Q) is functionalised with a 3′-quencher so as to bring it within proximity of the fluorophore when both functionalised strands are bound to a stator. The quencher Iowa Black absorbs in the red spectral range and dissipates the absorbed energy as heat into its surroundings. TET (Tetrachlorofluorescein) is related to the dye fluorescein and emits with a peak wavelength of 539 nm, making the Iowa Black quencher suitable for blocking its emission. The same goes for the blocking of the Cy5 emission which is centred around a 668 nm peak[148].

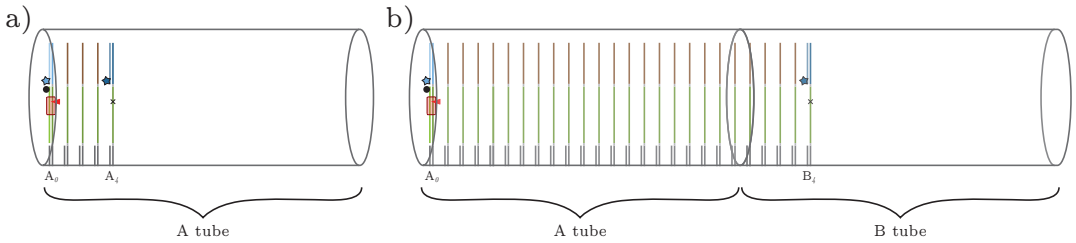


Figure 30: Schematics of the monomer and dimer tube track with fluorophores at the beginning and end of the track. The motor carries a quencher which reduces intensity of the fluorophore fluorescence when in close proximity. The tube that the motor is loaded onto is referred to as an 'A tube'; any subsequent tubes are referred to as 'B tubes'. a) The monomer has 5 stators: A_0-A_5. b) The dimer has 26 stators, A_0-A_{20} and B_0-B_4.

The $M_Q.A_0.F_0$ complex is assembled separately from the tube to ensure there is at most one motor per tube, as described above, with ratios 0.9:1:3 M_Q:A_0:F_0 with A_0 at 5 µM in NEB buffer 2.1 (see Appendix A.2.4). This complex is then added to the tube assembly, in a ratio of 0.9:1 A_0:tube and incubated for 1 hour at 37 °C . The start stator A_0 has a different extension to the other stators so that the excess F_0 does not bind to any of the other stators.

Motor motion is observed by monitoring changes in fluorescence intensity of both F_0 and F_{end}. When M_Q is bound to a stator with a fluorophore, the fluorescence emission is quenched and the intensity is lowered. An illustration of the stator and motor with fluorophore and quencher respectively are shown in Figure 31.

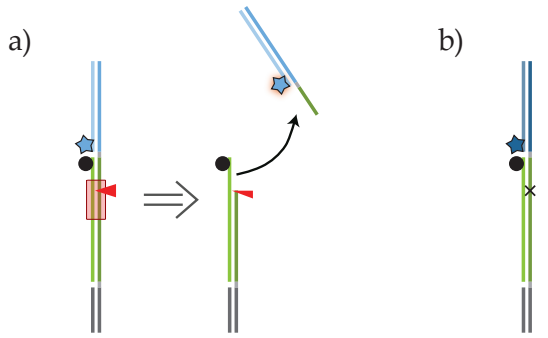


Figure 31: Schematics of the fluorophores at a) the beginning on A_0(F_0) and b) end of the track (F_{end}). The stators at the beginning and end of the track have unique extensions at the top, that are complementary to a TET-labelled strand (TET_0) and a Cy5-labelled strand ($Cy5_{end}$) respectively. The motor is depicted here has a 3' quencher. Upon addition of enzyme, A_0 is cut by the nicking enzyme at the restriction site indicated by the triangle and the labelled top diffuses away. The end stator has a mismatch in the restriction site, indicated by the cross, meaning the enzyme does not cut this stator and the motor remians bound.

Upon the addition of 10 units of nicking enzyme, Nt.BbvCI (10,000 units/ml), we see from Figure 32 b that where the fluorescence of F_0 was suppressed, as the motor steps away from A_0, the intensity increases. The rate at which the intensity increases is the same for both the dimer and monomer tracks, demonstrating that the enzyme

proceeds in solution to A_0 at the same rate, and operates with the same efficiency for both monomer and dimer.

However, the rates at which the motor reaches the end of each track differ. The dimer track is 21 steps longer than the monomer track and thus the rate of quenching of F_{end} is much slower for the dimer than it is for the monomer.

The changes in fluorescence intensity of F_0 indicate enzyme cutting and the melting off of $A_0.F_0$ duplex top. And so the intensity changes do not indicate the motor stepping away, only that it becomes possible for it to do so. To normalise the fluorescence level we must do so with respect to the initial level F_0. We can also normalise F_{end} by adding in excess motor to the assemblies after the signal has flattened out. In doing so it is possible to learn what proportion of motors successfully reached the end of the track during the given time. Looking at the bottom graph in 32 b, we see that roughly $2/3$ of the number of motors that got to the end of the five-step track made it to the end of the 26-step track.

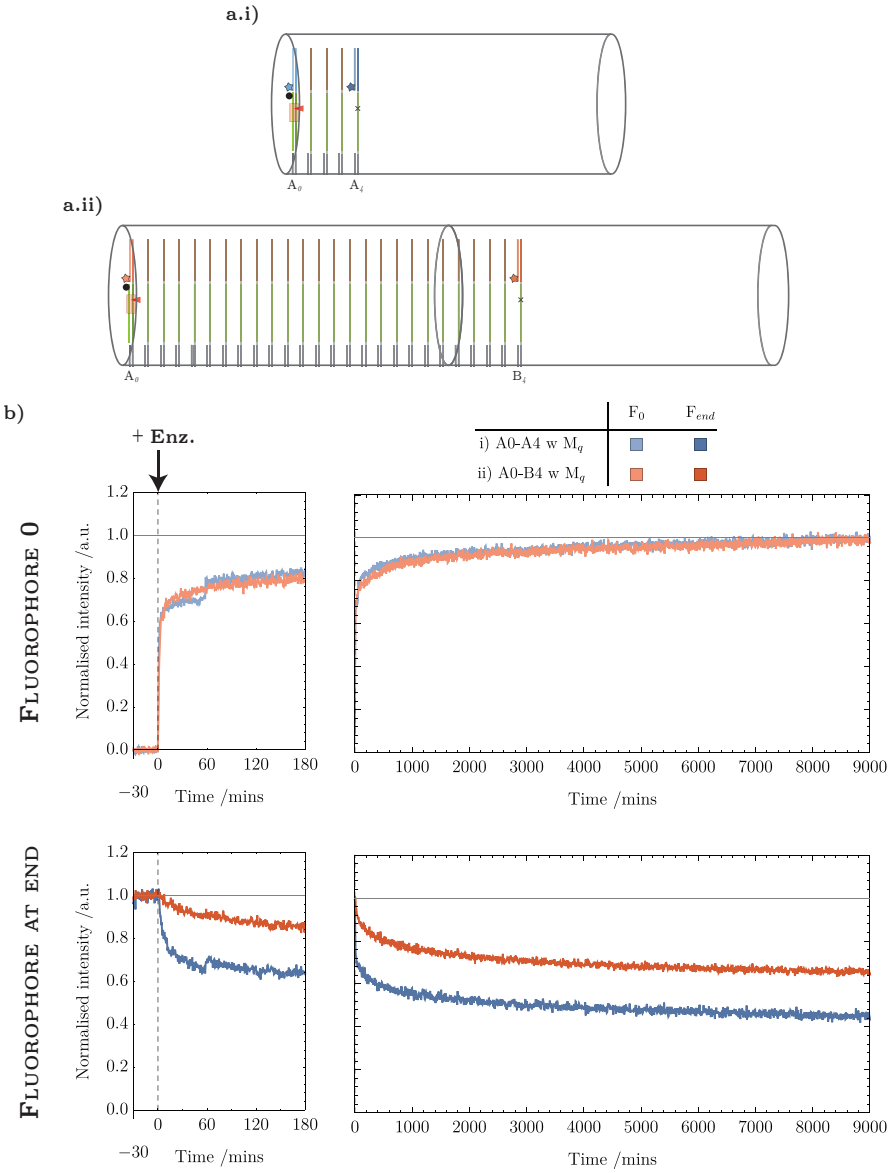


Figure 32: Comparison of motor motion along 5 and 26 stators a) Depictions of the tracks for i) a 5-stator track on a monomer tube; and ii) a 26-stator track that crosses between two connected tubes (a dimer). b) The normalised fluorescence traces of the fluorophore at the beginning and end of the track (F_0 and F_{end}). Motion is initiated by the addition of the nicking enzyme at t = 0 mins. The F_0 signal is normalised with respect to its initial level, as zero (averaged over 30 mins before t = 0 mins), and the level just before excess motor is added is defined as 1. The F_{end} signal is normalised with respect to its initial level too, also averaged over 30 mins before t = 0 mins, and defined as 1, before any motor is added. The signal is normalised with respect to the final level, after excess motor is added, which would represent the signal given off if all stator tops had been cut off except the end stator, and all end-stators had a motor bound. This signal should be the most quenched F_{end} signal.

4.3.2.2 *Motion along a less flexible, double-stranded track*

The size of a single step between stators is too small to be resolved with super-resolution techniques. But as mentioned above, we would like to observe the motion of the motor along the tracks using TIRF microscopy. One way in which to do this is to label every stator on the track with a strand functionalised with a fluorophore and observe changes in intensity as the motor walks and cuts off the fluorophore-labelled tops. The fluorophore-strand binding domain on the stators are 20 bps. When the motor is bound to a stator it gives rise to a full double helix, from the base of the stator next to the origami to the top of the stator, with a length of 17 nm. This length is comparable to the diameter of the nanotube which is approximately 18 nm. The inclusion of unpaired bases between binding domains is therefore essential to provide flexibility for motor motion from one stator to the next. Here we test whether motion is possible with such a track, whether there is enough flexibility to ensure the motor is not slowed down considerably compared to its motion along an unlabelled, entirely single-stranded track, and whether the motor can reach across the larger step size for stators between tubes.

The same set-up is used here as above- the 26-stator track made from the AB tube dimer, together with the $M_Q.A_0.F_0$ complex with ratios 0.9:1:3, M_Q:A_0:F_0. F_0 and F_{end} are TET and Cy5 respectively, as before. This time however, we bind the top fluorophore domain of all other stators to an unlabelled version of their complementary strand. A 5 x excess of this strand is added during the incubation of the $M_Q.A_0.F_0$ complex with the assembled dimer tubes. Subsequent purification removes any excess, unbound strands.

Just as before in Section 4.3.2.1, we see evidence of the motor stepping away from A_0 by the increase in fluorescence of F_0 and reaching B_4 by the reduction of intensity in F_{end}. The rates of increasing

and decreasing intensities are also very comparable to that which we observed when the track was entirely single-stranded. Therefore it seems that this double-stranded labelling technique does not interfere with normal operation of the motor.

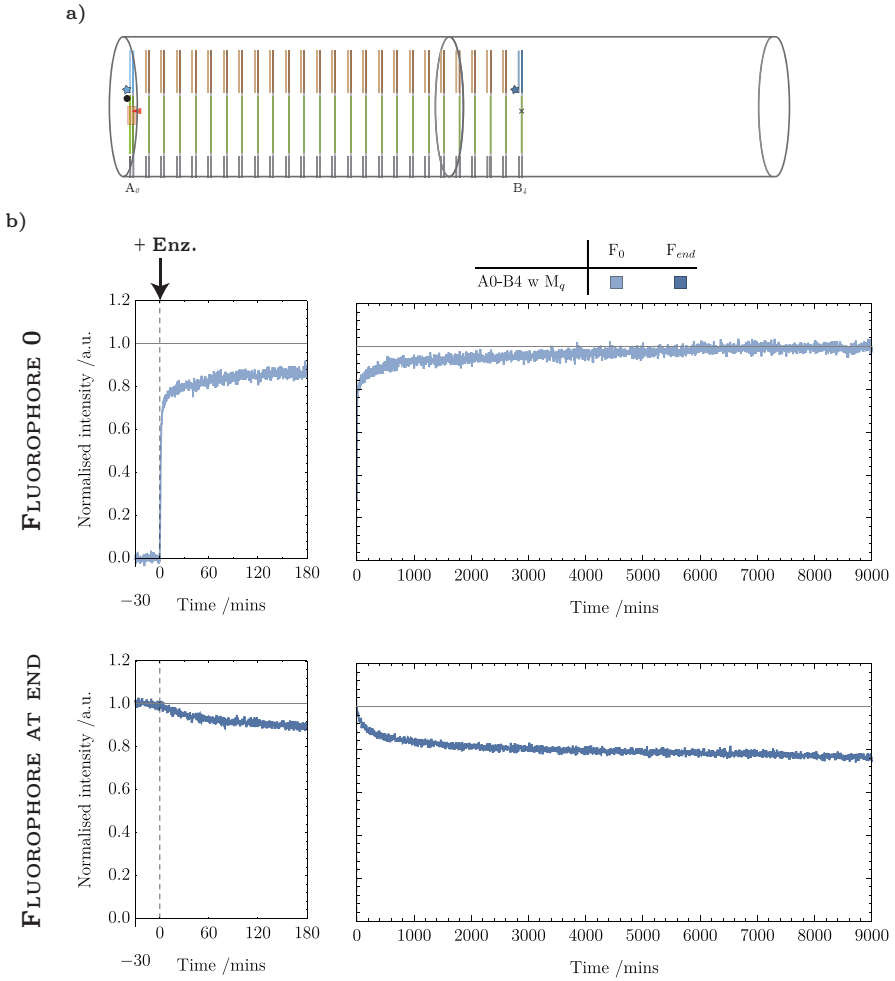


Figure 33: Motor movement along a dimer tube. The stators have double-stranded domains above the motor-binding site which allow us to test whether the motor will be able to move along a fluorescently labelled track in TIRF a) Depiction of the 26-stator track. b) The normalised fluorescence traces of the fluorophore at the beginning and end of the track (F_0 and F_{end}). The motor motion is initiated by the addition of the nicking enzyme at t = 0 mins. The F_0 signal is normalised with respect to its initial level and the level just before excess motor was added. The F_{end} signal is normalised with respect to its initial and final levels after excess motor is added.

4.3.2.3 *Transporting a fluorescent cargo*

As we saw in Chapter 1 Section 1.4.2, FRET describes an energy transfer mechanism between two fluorescent molecules, where energy is transferred from a donor fluorophore to an acceptor fluorophore.

Here FRET pairs of fluorophores (as opposed to a fluorophore and quencher) are used to test whether the motor can transport a large fluorescent cargo along the inside of the tube. Again we are testing an experimental set-up that will be used later in TIRF microscopy, to enable direct motor observation. In order to observe the motor directly, we have designed it to carry a cargo of two identical strands ($Cy3_M$) which are functionalised with Cy3 fluorophores at both their 5′ and 3′ ends. The 3′ end of the motor strand is extended to create two identical domains which are complementary to a dual-labelled Cy3 strand; this motor strand is called M_{Cy3}. The design of this experiment is illustrated in Figure 35 a. By having four fluorophores we hope to increase the intensity of the motor and make observation via TIRF microscopy easier. Figure 34 shows the FRET processes that take place in this experiment.

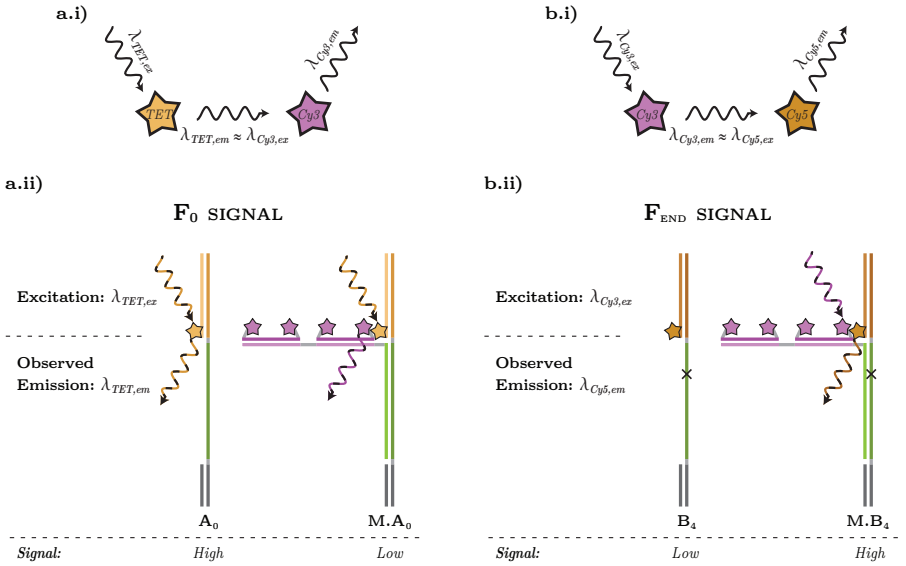


Figure 34: A schematic of the FRET between the Cy3 on the motor and a) the TET on the first stator and b) the Cy5 on the final stator. This allows the motor's presence to be detected at both ends of the track. Cy3 and TET and Cy3 and Cy5 are FRET pairs, trasferring energy between each other and suppressing each other's emission. When TET is next to Cy3 the TET emission is suppressed and when Cy3 is next to Cy5 the Cy3 emission is supressed. Schematic used with permission from Alex Lucas [146].

As above, stator A_0 is labelled with TET as F_0 and stator B_4 is labelled with Cy5 as F_{end}. Cy3 forms a FRET pair with both TET and Cy5. The motion of the motor is observed by monitoring the fluorescence interplay between TET, Cy3 and Cy5, which is outlined in Figure 34. The dimer tube is again employed here.

The motor complex is formed by annealing together M_{Cy3}:$Cy3_M$:A_0:F_0 in the ratio of 0.9:2.25:1:3 and it is purified and added to the tube post-assembly.

We begin by exciting F_0 with $\lambda_{TET,ex} = 539nm$ and monitoring its emission at $\lambda_{TET,em} = 522nm$. When the motor is bound to A_0, the FRET interaction means that no emission is observed as the energy is transferred to Cy3 and it emits instead at $\lambda_{Cy3,em} = 550nm$. However as the motor carries its Cy3 cargo away from A_0, the FRET interaction decreases and eventually is prohibited; thus we observe an emission at $\lambda_{TET,em}$. Simultaneously, to monitor motors reaching the end of the track at B_4, Cy3 is excited with $\lambda_{Cy3,ex} = 564nm$ and we look for a Cy5 emission at $\lambda_{Cy5,em} = 648nm$. When the motor is attached to B_4 the energy of an excited Cy3 is transferred to Cy5, whose emission is observed at $\lambda_{Cy5,em}$.

Thus, to observe successful passage of the bulky fluorescent cargo we look for an increase in intensity of F_0 and F_{end} emission. This is what we observe in Figure 35 b upon addition of the enzyme. Similar characteristics as the previous ensemble measurements are displayed, where the motor steps away from A_0 quickly and reaches B_4 more slowly, due to the many steps it must traverse. The rates are similar to the rates in the previous ensemble measurements. The fluorescence signals are normalised with respect to the level on the initial plateau just before enzyme is added (averaged over 20-30 mins) and final level, defined by the final plateau before excess motor is added. This time such normalisation does not give an indication of how many

motors are reaching the end of the track, because we have not normalised with respect to the level after excess motor addition, due to insufficient data being collected. Despite this, the graph does show that it is possible for the bulky motor to move inside the tube.

For all FRET results here, it would have been beneficial to do extra controls to gather more quantitative data from these experiments, for example, to find what the F_{end} signal looks like when the motor is attached to the penultimate stator. However, because the fluorophores are cut off during motor motion it makes it difficult to do such controls. At this stage, we only require qualitative results to determine whether or not the motor can move in such a set-up.

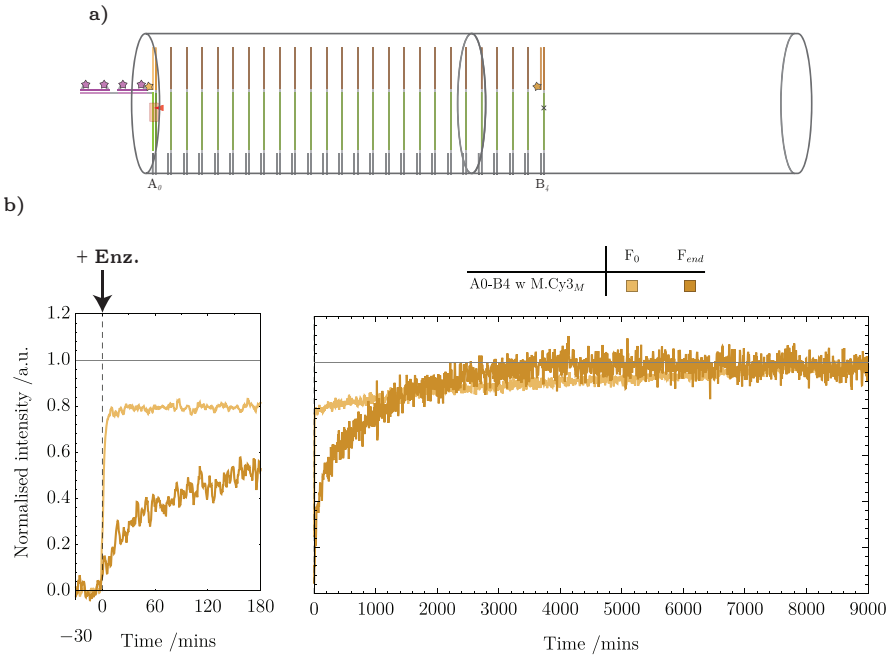


Figure 35: Depiction of the dimer track with a Cy3-labelled motor. b) The normalised fluorescence traces of the fluorophores at the beginning and end of the track (F_0 and F_{end}). The motion is initiated by the addition of the nicking enzyme at t = 0 mins. The fluorescence signals are normalised with respect to their initial level and final levels, before enzyme was added and before excess motor was added respectively.

4.4 MICROSCOPY; LONG RANGE MOTION ATTEMPTS

In this section we use TIRF microscopy to illuminate a thin layer of a sample and attempt to observe the motion of a molecular motor through DNA nanotubes. The initial attempt is indirect, i.e. the effect the motor has on a fluorescently labelled tube is measured. A direct attempt is then made where the motion of the fluorescently-labelled motor itself is monitored. TIRF allows reduction of the background noise in such circumstances. In both instances we aim to observe long-range motion of a motor across a number of nanotubes stacked end-to-end. In order to have a clear image, it is important the motors remain in focus and thus stay within the narrow TIRF layer. To provide such assurance, the tubes are altered slightly to allow tethering to a surface.

End-to-end stacks are assembled from multiple B tubes with an A tube at one end acting to cap the stack. A tubes are assembled as described in Section 4.3.1, but at 20 nM instead of 54 nM, with poly-T end-staples suppressing stacking at the LH end of the tube. In this case however, the B tubes are assembled at 20 nM with 100 nM linking staples included (promoting stacking at both the LH and RH ends of the B tubes) and left for at least 12 hours to join together. The A tubes, with the pre-annealled $M.A_0$ complex attached, and the B tubes are purified and added together in a ratio of 1:20, A:B. The tubes are bound to one another by incubating for at least 1 hour at 37 °C. In doing so we aim to have an A tube, with a motor present, at the start of most B stacks.

Where the tube is labelled, it is done with a green fluorescent nucleic acid binding dye called YOYO-1 Iodide. YOYO-1 is the dimer of the asymmetrical cyanine dye oxazole yellow, YO, and emits at a wavelength of 509 nm [149]. The YOYO is excited using a wide-field

excitation mercury lamp in conjunction with a 'wideband interference blue' (WIB) excitation filter which allows wavelengths of 460-495 nm to pass. YOYO staining is described in more detail in Appendix A.2.2.1. Where Cy3 is used, such as on the dual-labelled Cy3 motor, or as track labels on the fluorescent binding domains of stators, it is illuminated with a 532 nm laser and the beam is passed through a Cy3/Cy5 filter cube.

4.4.1 *Tube alterations for imaging with microscopy*

The nanotubes are examined after being placed into a flow channel between two glass slides, as described in the proceeding section. To hold them in place at the surface of the glass slide, they are tethered using biotin-streptavidin interactions. During the design stage for this origami, seven nick positions evenly spaced along the tube are moved five bases around their respective helices, to position them on the outside of the nanotube instead of on the inside. This is done by shortening one staple and lengthening its neighbour by half a helical turn. One staple ending at each newly-positioned nick is then extended to include a 16 nucleotide domain that complements a biotinylated strand. The biotinylated strand is added to the annealing mix in a 1.2 x excess over normal staples, such that each tube segment is functionalised with five biotinylated anchors .

4.4.2 *Preparing a flow channel for imaging*

A flow channel is created via an adapted protocol from Machinek 2014 [150]. The channel is set vertically between two glass slides and is horizontally positioned between two pieces of parafilm set 1 mm apart. The channel volume (which varies between 10 and 20 µl) pro-

vide a space for the sample to be contained within and observed. Each addition to the channel is drawn through using capillary action and blotting paper. To provide control over the tubes with their biotinylated attachments, the glass slide is treated as follows. Firstly one channel-volume (1 CV) of 1/10,000 300 nm latex beads is added to provide a reference point for the slide surface when focusing the microscope. The beads are incubated for 1 min. The channel is next treated with 1 CV bovine serum albumin (BSA) for 1 min to prevent non-specific adhesion to the glass surfaces. The BSA solution contains biotinylated-BSA (biotin-BSA) in a ratio of 4:1 BSA:biotin-BSA, (2 mg/ml:0.5 mg/ml). 1 CV of 1 mg/ml streptavidin creates stationary binding sites on the slide surface, for the biotinylated tubes upon attaching to the biotin-BSA. The tube samples are the final addition to the channel which now comprises a passivated surface with some tubes specifically bound on the surface and reference beads interspersed. Between each addition, the channel is washed with 2 CV of origami buffer to get rid of unbound species. Protein incubation steps last for 5 mins whilst the sample is incubated for 2 mins only. Finally the channel and contents are treated with anti-photobleaching buffer (details in Appendix A.2.2.2) and sealed with Valap to prevent evaporation and oxidation from the channel ends.

4.4.3 *Optimisation of conditions for motor operation*

Imaging buffer in previous studies was not the same buffer as we use here [150]. For example, for our imaging experiments we add the fluorescent YOYO-1 dye to the buffer in some cases, as well as anti-photobleaching buffer. Therefore it is necessary to check whether the cutting enzyme can operate in such an environment before carrying out microscopy experiments. YOYO-1 is an intercalator and has

been known to cause slight unwinding in DNA and so here we test the extent of its effect on our structures [151]. We also use an anti-photobleaching buffer and thus test for it affecting our structures too.

Three buffers are compared: microscopy buffer (1 x TAE, 12.5 mM $MgCl_2$, 50 mM NaCl), microscopy buffer and 1/100 YOYO-1 Iodide, and the anti-photobleaching buffer with 50 mM NaCl and 12.5 mM $MgCl_2$. Cutting is tested by adding 0.5 µl Nt.BbvCI enzyme (10,000 units/ml) to different complexes and incubating for 1 hour at 37 °C and running the products (and their respective reference complexes without enzyme) on a polyacrylamide gel with loading dye. The dye contains the protein denaturant sodium dodecyl sulfate (SDS), to denature the enzyme and prevent it from binding to the DNA in the gel. Each lane contains 18 µl sample with 2 µl 10 x the respective buffer, and enzyme as specified. The stator concentration was 100 nM for all of the samples. The gels are stained with Sybr gold and scanned for Cy3 and Sybr gold fluorescence.

Differing versions of a motor-bound stator are tested: a unlabelled motor and an unlabelled stator (M.A, annealed in a M:A ratio of 1.2:1), an unlabelled motor and a labelled stator ($M.A.Cy3_A$, in a $M:A:Cy3_A$ ratio of 1.2:1:1), a motor labelled at both the 3′ and 5′ end, bound to an unlabelled stator ($M_{3',5'}.Cy3_M.A$, in a $M_{3',5'}:Cy3_M:A$ ratio of 1.2:2.4:1), and a dual-labelled motor at the 3′ end bound to an unlabelled stator ($M_{2x3'}.Cy3_M.A$, in a $M_{2x3'}:Cy3_M:A$ ratio of 1.2:2.4:1).

The enzyme cuts successfully in both origami and anti-photobleaching buffer; the cut stator tops are highlighted by green ellipses in Figure 36, where they are not present in their respective lanes without enzyme. Cutting in the YOYO-1 buffer is less obvious due to the reduced signals from both Cy3 and Sybr gold. Being an intercalator too, Sybr gold is a direct competitor with YOYO-1 for DNA binding and so it is likely less is bound here, reducing the intensity of the signal.

Cy3 and YOYO-1 have overlap in their emission and absorption spectra and thus can act as FRET pairs [152]. It is possible therefore that the signal of Cy3 is reduced due to a FRET interaction resulting from proximity between it and YOYO-1. Despite this, we still see cutting of the stator tops, highlighted by the blue ellipses in Figure 36.

The gels also tell us how YOYO-1 staining affects our structures. Compared to the buffers without YOYO-1, there are fewer $Cy3_M$ strands bound to both $M_{5',3'}$ and $M_{2x3'}$ in the presence of the dye. This indicates the smaller, less strongly-bound parts of the structures are affected by the presence of YOYO-1. The $Cy3_M$ strand is 10 bases and thus would benefit from an increased G-C rich content if redesigning in the future. This would provide stronger binding. It is possible to see some of the motor-bound stators have one, two or no $Cy3_M$ strands present- this is seen as multiple bands in the respective lanes.

Figure 36 includes diagrams of the structures contained in each gel lane. The faint parts of the diagrams are not present in the lane but are included in the diagram to show how the bold coloured sections bind together to create the working nanostructure. For example in (i), the only bold part is the green motor strand, but the faint regions show how this motor strand binds to the stator and the position of the fluorophore strand above it.

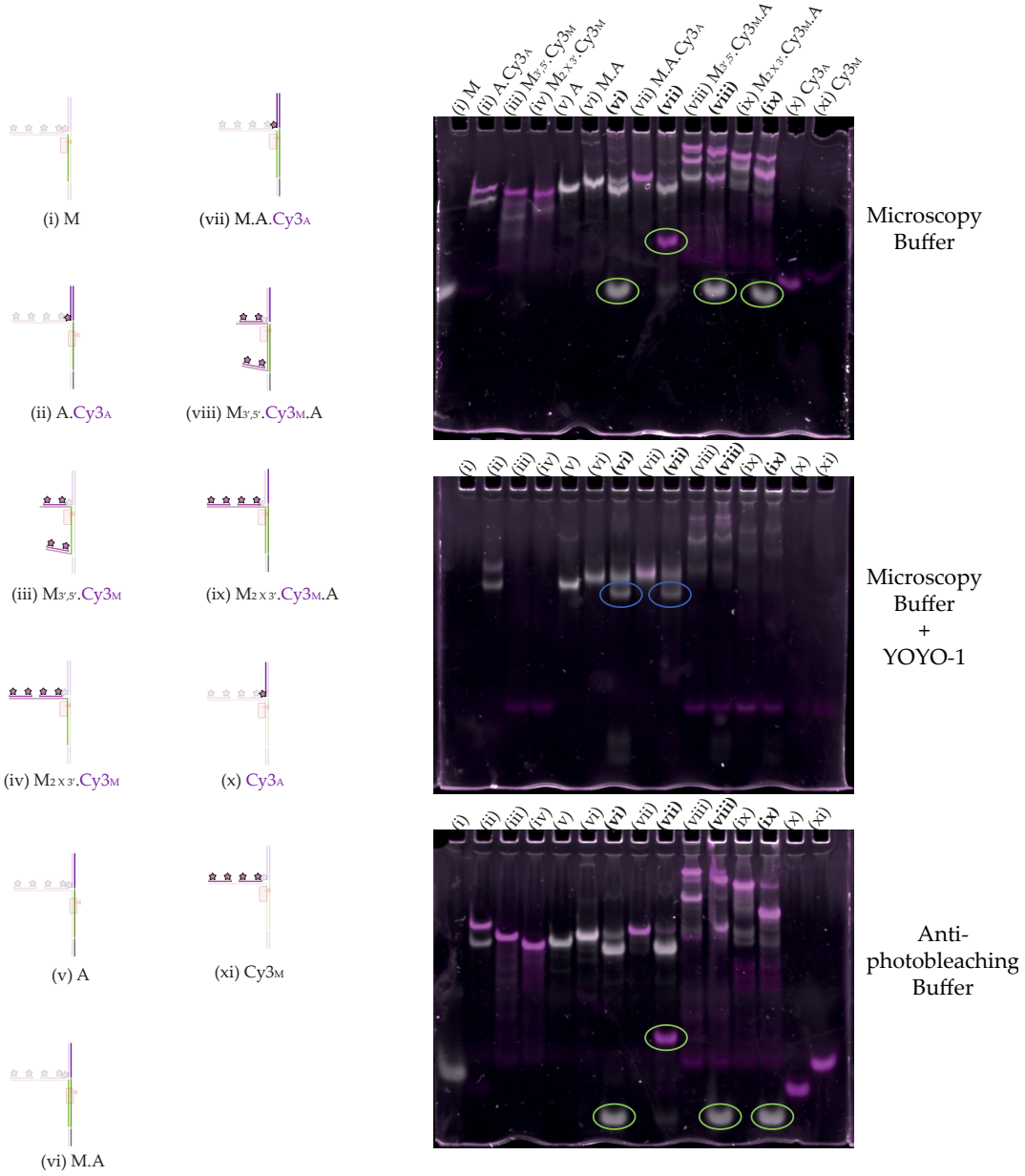


Figure 36: Enzyme cutting tests for the different buffer conditions used in the microscopy experiments, together with schematics illustrating what is contained in each gel lane: a) microscopy buffer, b) microscopy buffer + 1/100 YOYO-1, c) anti-photobleaching buffer + 50 mM NaCl + 12.5 mM $MgCl_2$. The gels are scanned for Sybr Gold (white signal) and Cy3 (pink signal). Green highlights the tops of anchorages cut free by the enzyme; where this is less obvious however, blue ellipses highlight the cut motor-stator complex. Lanes and schematics i-v | reference samples of M, A.$Cy3_A$, $M_{3',5'Cy3}$.$Cy3_M$, $M_{2x3'Cy3}$.$Cy3_M$ and A (stator). Lanes and schematics vi-ix | the consecutive pairs of lanes contain samples of annealed M.A, M.A.$Cy3_A$, $M_{3',5'Cy3}$.$Cy3_M$.A and $M_{2x3'Cy3}$.$Cy3_M$.A; the second lane of each pair (**vi-ix**) is incubated with enzyme too. Lanes and schematics x-xi | reference samples of $Cy3_A$ and $Cy3_M$. The gels are 15% 29:1 polyacrylamide and were run at 200V for 120 min. Sample details: enzyme digestion used 0.5 µl of Nt.BbvCI (10,000 units/ml) per 20 µl sample, and was incubated at 37 °C for 1 hour.

4.4.4 *Results and discussion*

4.4.4.1 *Motor motion via track observations*

In section 4.3.2.2 the motion of the motor was tested and verified along a double stranded track using ensemble measurements. This design enables fluorescent labelling of the track and has a domain at the top of the stators that binds to a functionalised Cy3 strand ($Cy3_A$). As the motor passes along the track and the enzyme cuts successive stators, the fluorescent-labelled domains diffuse away into solution. Where we start with a localised region of fluorescence along tubes with intact tracks, as the motor walks and the signal becomes less localised, the fluorescent labels are no longer bound to the tubes on the surface and we instead see fluorescence from labelled fragments that come into and go out of TIRF focus.

Initially the brightness over time of two labelled track samples is monitored where one sample has enzyme added to it (forming the 'test sample' aliquot) and the other does not (forming the 'negative control' aliquot, so motors should not be able to step to the next stator). Following this experiment the positive control was also set up and observed. Here, excess motor is added such that every stator is bound to a motor, and the extreme case of all stators being cut can be compared to the negative control and test sample. The positive control does not require the motor to be able to move down the track, but instead allows us to check the enzyme can cut the stators, and to monitor the reduction in fluorescence signal from $Cy3_A$ diffusing into solution.

For all three samples, the $M.A_0$ complex is pre-annealed in a ratio of 0.9:1, $M:A_0$, with A_0 at 5 μM. For the negative control and test sample, the tubes are annealed at 20 nM with all stators present and the $Cy3_A$ strand is included in the annealing process in a ratio of 1:1

$Cy3_A$:stators. The $M.A_0$ complex is added to the A tubes (lacking the start stator) in a ratio of 0.9:1 A_0:tube and incubated for 1 hour at 37 °C. The A and B tubes are then purified separately before they are combined in a ratio of 1:20, A:B, with 100 nM linker staples. Excess linker staples are purified away. The negative control and the test sample are put into separate flow channels (20 µl of 20 nM tubes) and washed through with 2 CVs of anti-photobleaching buffer each. The test sample had an addition of 0.5 µl Nt.BbvCI enzyme present during this wash.

For the positive control, where every stator should be bound to a motor and thus may be cut, the assembly process is mostly the same. The only difference here is that instead of the $Cy3_A$ strands being included during the annealing process of the tubes (as they were for the negative control and test sample), they are instead incubated with pre-assembled A tubes during the $M.A_0$ incubation step, and in a separate incubation step for the B tubes (where no motor is incubated). The samples are then purified and added together as above in a 1:20 ratio of A:B, with 100 nM linker staples. Excess linker staples are purified away. 20 µl of 20 nM positive control tube is added to a flow channel and washed with anti-photobleaching buffer containing 0.5 µl Nt.BbvCI enzyme, as with the test sample. All flow channels are sealed with Valap to prevent evaporation.

The first images of the test sample, the negative control and positive control are taken roughly 15 mins after the addition of anti-photobleaching buffer (with, without and with enzyme, respectively). Immediately after the first images of the positive control are taken, its buffer is replaced with anti-photobleaching buffer containing enzyme and an excess of motor strand, to encourage all stator tops to be cut (thus making it the positive control). The motor excess is 10 x the stator amount, where we assume all tubes in the flow channel

have stuck, thus 10 x the stator amount is conservative and may be more than this in reality. The final motor concentration is 4.2 μM.

Figure 37 shows typical images of the samples taken on day 1 (the day of assembly) and for a number of days following. In between imaging, the slides containing the flow channels and samples are wrapped in foil to protect from the light and photobleaching, and incubated at 37 °C.

The images from day 1 for the test sample (37 a, 'Expt') are less bright than the negative control sample (37 b), as can be seen in Figure 38 where the mean pixel brightness over time is plotted for all three samples. It could be that the motor has already moved and 'burnt' the track behind it and that this is responsible for the reduced fluorescence signal seen here. It would be beneficial to have taken images before and after enzyme addition to provide a reference image for uncut tracks. For the positive control, images are taken pre- and post-enzyme addition which allows us to make such a comparison (Figure 37 c). There is a sharp decrease in the sample brightness once excess motor is added, suggesting the enzyme successfully cuts the labelled top of the stators in the presence of the motor. This also suggests that the dimming of the samples seen over the days is not entirely due to photobleaching alone.

For all samples the localised signals on the surface get dimmer with time. As mentioned, this may well be in part due to photobleaching. However, from watching the samples in real-time under TIRF, as time went on, not only did the localised signals get dimmer but flashes of Cy3 fluorescence also became apparent. On inspection of videos for the test sample it appeared that the number of flashes increased with time.

It was thus proposed that the increase in Cy3 flashes may be due to an increase in the number of cut labelled stator tops freely diffusing

in the flow channel, and that the flashes may be attributed to these coming into and going out of TIRF focus. Sequential snapshots of the samples are given in Figures 39, 40 and 41. For the positive control in Figure 41 the flashes are more numerous than for the test sample in 39. This may be related to the slower rate of cutting in the test sample and the lower sample density compared with the positive control. There are very few flashes at all in the negative control that did not contain enzyme (on inspection of the video).

The flashes are observed and quantified by plotting the median-subtracted version of the microscopy image. The median value for individual pixels over 10 images is found. The median is subtracted from each pixel value and then divided by the median. This is then normalised with respect to its 0.001^{th} and 0.999^{th} quantile values (the minimum and maximum values are sensitive to outliers). This method indicates whether a pixel in a certain frame is unusually bright or unusually dark, relative to its median intensity over the period of observation, and is a best attempt at picking out flashes from static images.

The bright flashes are then detected (and are highlighted with ellipses) using Mathematica's in-built Laplacian of the Gaussian (see Appendix A.2.3) for edge-detection on the normalised images [70]. The total numbers of detected spots in the videos for the main sample, the negative control and the positive control are 58, 3 and 595 respectively.

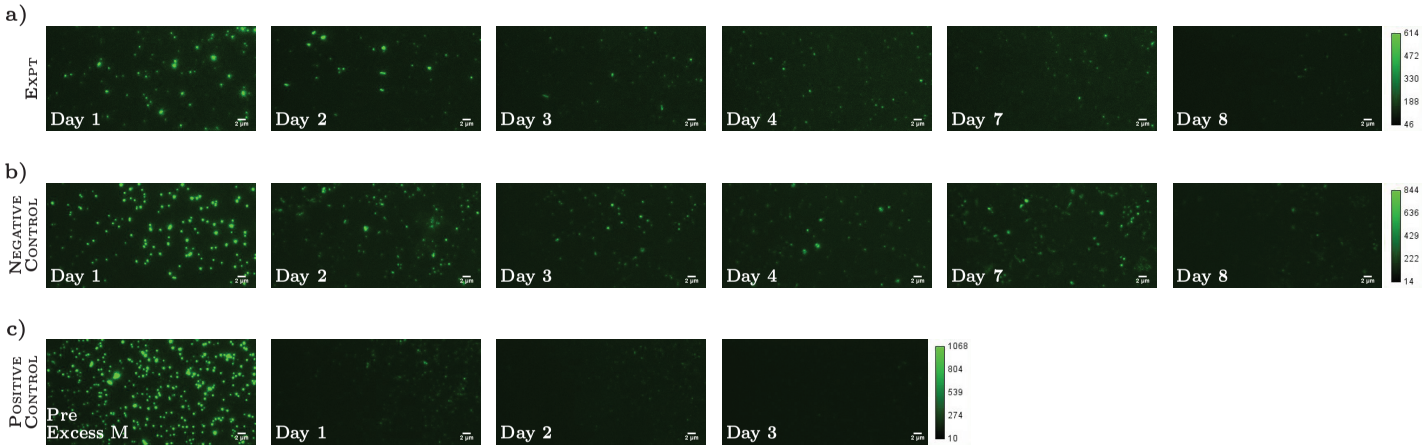


Figure 37: Examples of typical fields of view of stacked tubes with Cy3-labelled tracks over multiple days. A motor is loaded onto an A tube that is able to cap the front of stacked B tubes. The tubes have full tracks that are labelled with $Cy3_A$ that binds to stator tops either by a-b) inclusion of $Cy3_A$ in the annealing of the tube or c) incubating the annealed tubes with $Cy3_A$. a) The test sample has enzyme in the buffer to initiate motor motion and results in the $Cy3_A$ tops being cut free. b) The negative control without enzyme - here photobleaching is the only effect reducing the fluorescence. c) The positive control has excess motor and enzyme in the buffer; the motor strands bind to the entire track and are cut quickly by the enzyme. Images for the positive control are also taken prior to adding the excess motor and enzyme. All the images for each sample have the same brightness/contrast settings and have a single appropriate calibration bar. Microscopy details: Cy3/Cy5 filter, illuminated by the 532nm laser, 100 ms exposure, 100 EM gain.

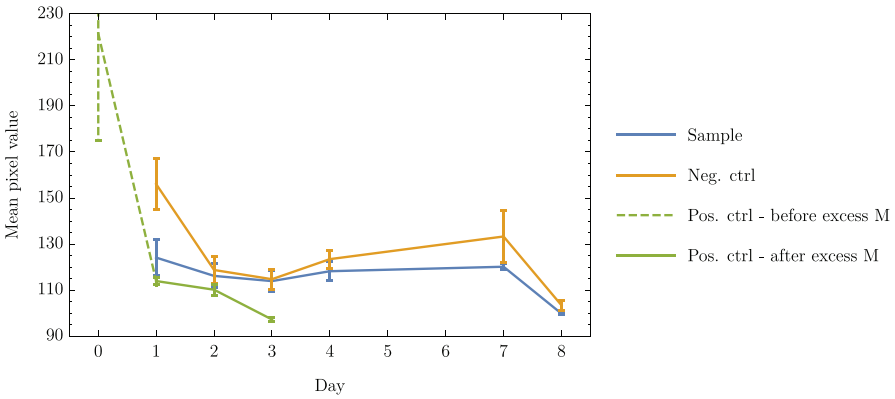


Figure 38: The mean pixel brightness over time for the test sample, the negative control and the positive control. This is averaged over 5-6 images taken each day per sample. The standard deviation provides the error bars. The dashed line and data point at Day 0 for the positive control indicates data taken before addition of excess motor.

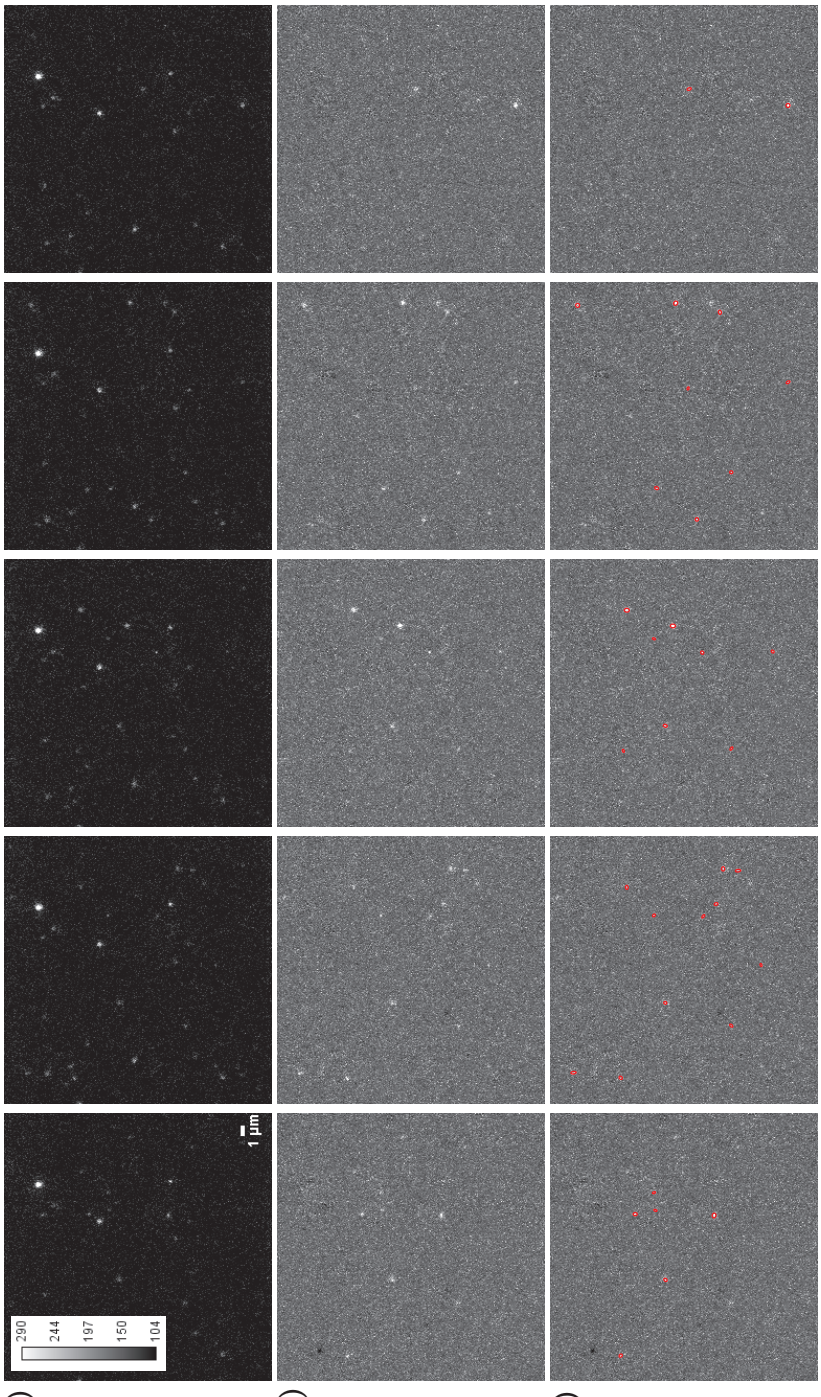


Figure 39: a) Sequential snapshots of a video for the test sample with normal amount of enzyme added (i.e. not the positive or negative control), separated by 100 ms, going from left to right. Microscope details: Cy3/Cy5 filter, illuminated by the 532 nm laser, 100 ms exposure, 100 EM gain. The calibration and scale bar are appropriate for all images. b) Median subtracted images of the corresponding image in (a). Bright spots indicate unusually bright pixels. c) Detected spots in the median subtracted images are highlighted here by overlaying the images with their fitted ellipses (there are very few; they are small ellipses. A total number of 58 detected spots are found over all 10 images in the video.

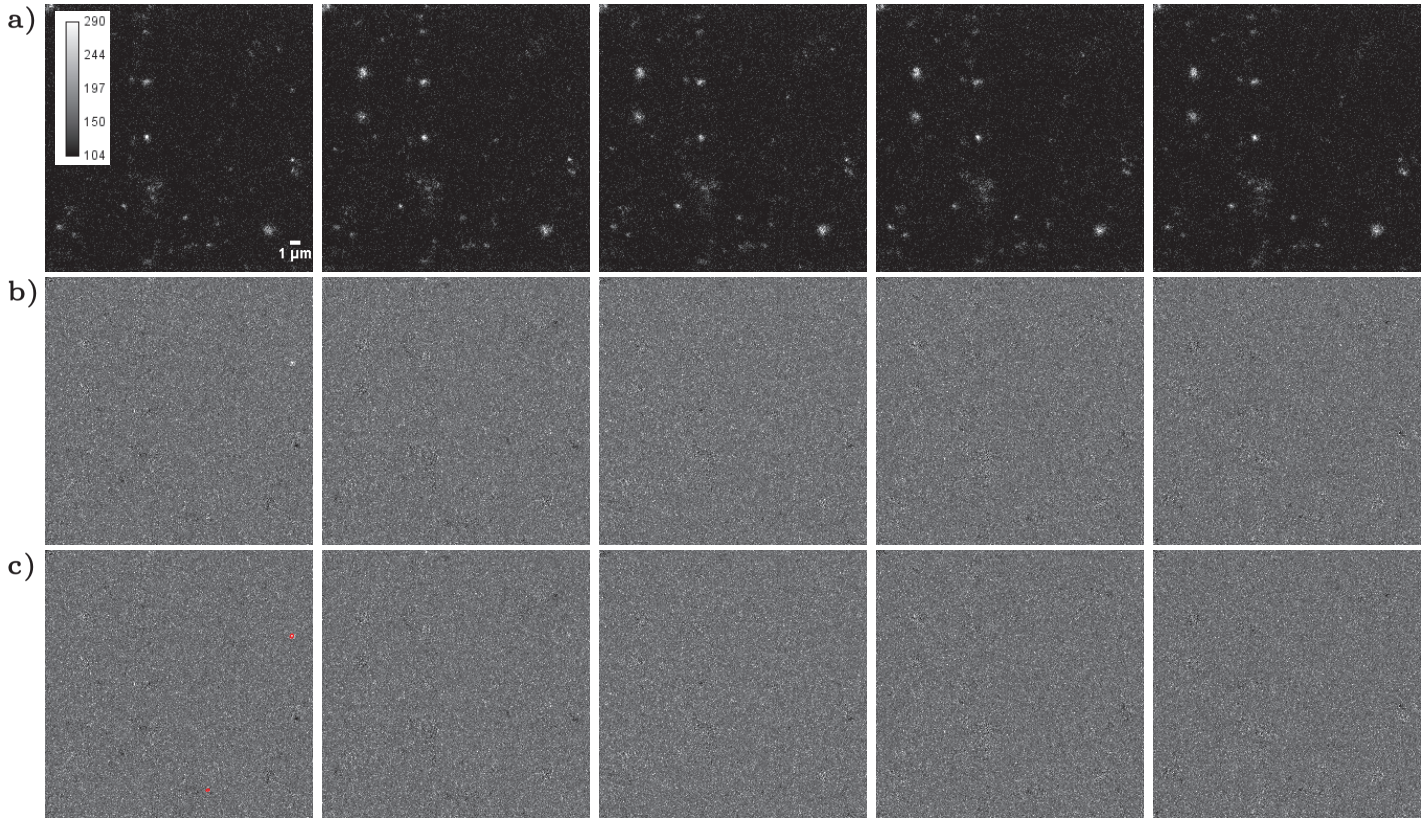


Figure 40: a) Sequential snapshots of a video for the negative control, separated by 100 ms, going from left to right. Microscope details: Cy3/Cy5 filter, illuminated by the 532 nm laser, 100 ms exposure, 100 EM gain. The calibration and scale bar are appropriate for all images. b) Median subtracted images of the corresponding image in (a). Bright spots indicate unusually bright pixels. c) Detected spots in the median subtracted images are highlighted here by overlaying the images with their fitted ellipses. A total number of 2 detected spots are found over all 10 images in the video (they are hard to see as they are small ellipses)

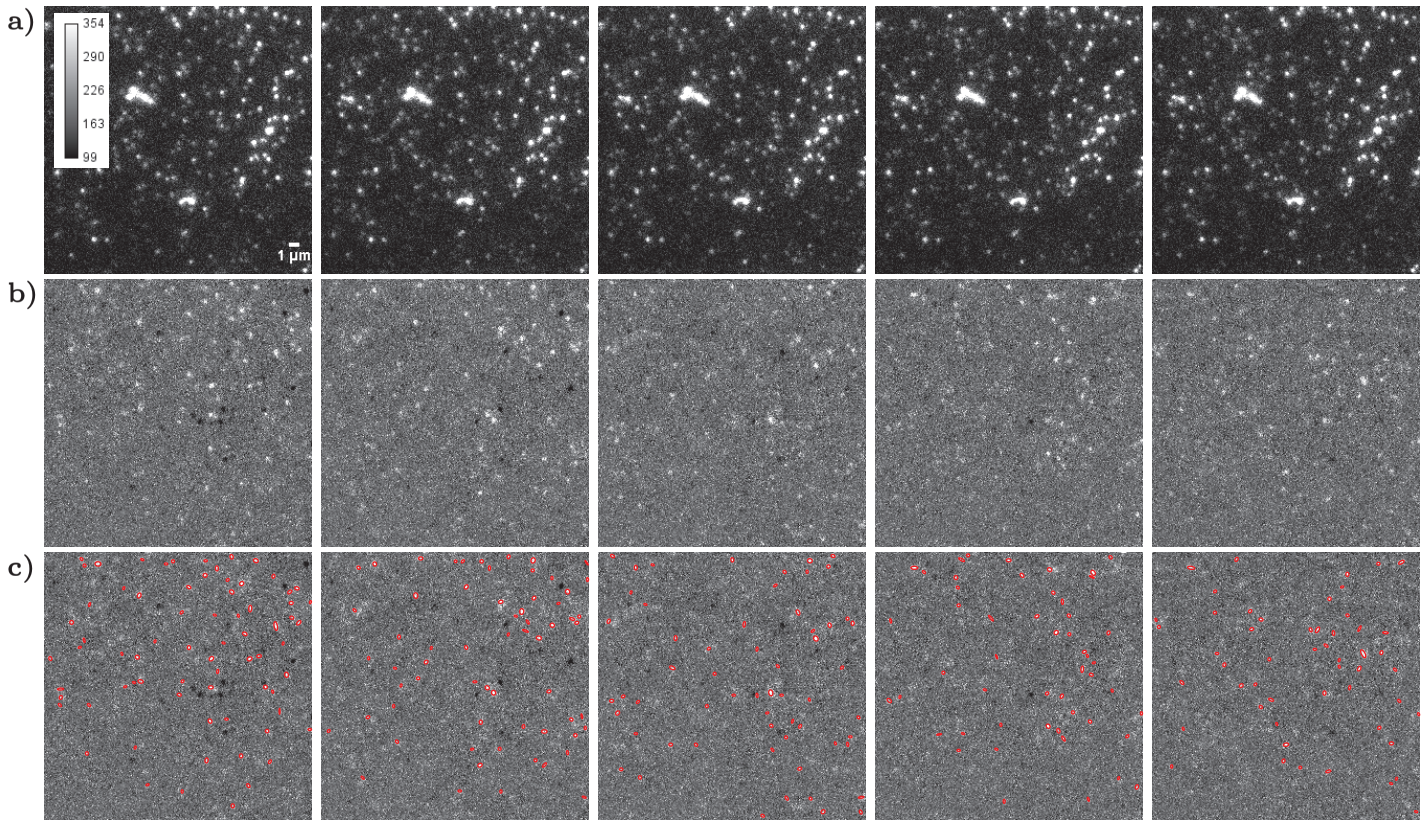


Figure 41: a) Sequential snapshots of a video for the positive control, separated by 100 ms, going from left to right. Microscope details: Cy3/Cy5 filter, illuminated by the 532 nm laser, 100 ms exposure, 100 EM gain. The calibration and scale bar are appropriate for all images. b) Median subtracted images of the corresponding image in (a). Bright spots indicate unusually bright pixels. c) Detected spots in the median subtracted images are highlighted here by overlaying the images with their fitted ellipses. A total number of 595 detected spots are found over all 10 images in the video.

To summarise this section, we have seen that all the samples suffer from photobleaching despite the presence of anti-photobleaching buffer. We have also seen the samples with enzyme added have flashing spots that are potentially indicative of diffusion of cut stators due to motor movement. However, it would be beneficial to observe motor motion directly to be certain of motion.

4.4.4.2 *Attempting direct observation of a motor inside a nanotube*

In section 4.3.2.3, motor motion with four Cy3 fluorophore cargos was tested and observed using fluorimetry. We now move on to attempt observation of such a motor using TIRF. The motor is labelled with two dual-labelled Cy3 strands ($Cy3_M$) with the aim of creating as bright a signal as possible. In the initial experiment here, we look to detect the motor but do not yet attempt to observe motor motion.

The extended motor M_{Cy3} is pre-annealed to $Cy3_M$ and A_0 with a ratio of 0.9:2.25:1, M_{Cy3}:$Cy3_M$:A_0 (A_0 at 5 μM). A and B tubes are annealed at 20 nM separately. The A tube anneal includes LH-end poly-T staples and the stators for the track excluding A_0. The B tube anneal includes 100 nM of linker staples to promote stacking. The labelled motor-A_0 complex is added to the A tubes and incubated at 37 °C for 1 hour to load the motor, with A_0 at 5 x excess to the tubes. The A and B tube assemblies are purified and combined in a ratio of 1:20, A:B, by incubating at 37 °C for 1 hour, encouraging capping of the A tube upon the B tube stacks. The sample is prepared in the flow channel as described above, including anti-photobleaching buffer, but enzyme is not included.

Upon observation using TIRF, the samples were found to photobleach within 40 seconds of imaging despite the presence of anti-photobleaching buffer. The mean pixel values over time are plotted for three fluorescent points in Figure 42, which illustrates the changes

in brightness. The spots are all seen to bleach in up to three steps of equivalent sizes. Such distinct steps are characteristic of individual fluorophores bleaching and are likely attributed to the Cy3 labels on the motor. Four steps were not seen and this may be due to fast photobleaching of a fluorophore prior to observation or due to unsaturated $Cy3_M$ sites on the motor strand. Despite this, it is unlikely that this signal arises from random fluorescent signals on the slide surface.

Due to the quick bleaching times, compared with the time it takes for a motor to transverse a track (possibly around 9000 seconds according to ensemble measurements), it would be difficult but not impossible for such a set-up to be used to detect long range motion of the motor on a tube. To reduce the rate of photobleaching, snapshots of motor position would need to be taken infrequently and the sample protected from light in between observations. Instead, a different labelling method could be effected which does not suffer the same photobleaching problems as the Cy3 and may be explored in the future (see Section 4.6 for suggestions). This experiment confirms it is possible to label and observe our motor, and with further experimental refinements this may be promising for future experiments.

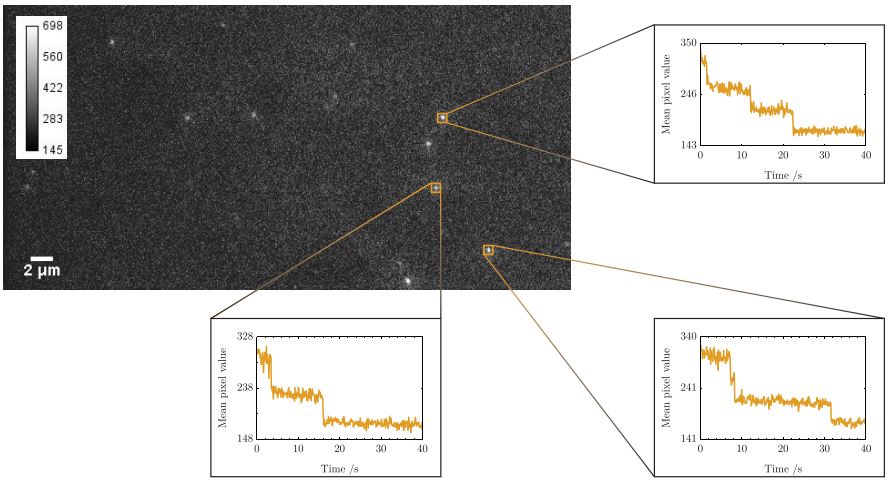


Figure 42: Labelled motors (with up to four Cy3 fluorophores) are added to tubes by pre-annealing to A_0. This motor-A_0 complex is then incubated with the A tubes (who are missing the staple for the first stator). The A and B tube assemblies are purified and added together in an A:B ratio of 1:20 and incubated. This is the first image in a video of a typical microscopy field of view. The tubes are not stained and thus are not visible here. Bright spots photobleach in multiple, roughly uniform steps. The intensity traces of three such spots are shown as insets. The mean pixel value is found for the highlighted square region around each spot in each image for a video. Microscopy details: Cy3/Cy5 filter, illuminated with a 532 nm laser, 100 ms exposure, 200 EM gain.

4.5 SUMMARY

Here we have carried out the preliminary experiments in observing molecular motor motion along a track in a DNA nanotube. Towards observing such an assembly, we have approached the challenge in two ways: by directly monitoring the motor motion, through it carrying a fluorescent cargo, and by indirectly monitoring the motor by watching the effect it has on a labelled track as it passes along the nanotube.

We have worked towards testing conditions for direct motor observation in real-time using TIRF microscopy. Simple motor movement within a tube was initially tested via ensemble measurements. We have shown it is possible for the motor to cross between the origami tubes; that the motor can move down a double stranded track; and that it can operate at normal pace with a bulky fluorescent cargo attached. We have proposed that Cy3 flashes during TIRF measurements may be indicative of motor motion. More work is required to achieve direct motor motion detection via TIRF over long distances, especially to limit the effect of photobleaching.

4.6 FUTURE WORK

In section 4.4.4.1 indirect observation of the motor moving was attempted by observing the changes in track fluorescence over time. The changes in brightness of the labelled-track occur on a much shorter time scale than the separation of these images. Future work would therefore benefit from providing a reference image of the sample before enzyme is added (as was done with the positive control sample). Subsequent images should then be taken more frequently for the first hour and less frequently after that, as photobleaching

allows. A microscopy set-up that will allow addition of enzyme during an imaging session, without removing the slide from the stage, is currently being developed. This will allow for images to be taken immediately after enzyme addition.

As mentioned in section 4.4.4.2 it may be necessary to use a different labelling method to observe long range motor motion, due to negative photobleaching effects. An alternative may be to use DNA-PAINT (DNA points accumulation for imaging in nanoscale topography) [153]. PAINT makes use of short fluorescently-labelled oligos which bind transiently to their complement. The PAINT strands exist in excess in solution and frequently interchange between a bound state and a freely diffusing state. In TIRF, only the the surface layer is illuminated and thus only the strands bound at a given time experience photobleaching. These are subsequently exchanged for other, non-bleached strands in solution, demonstrating how PAINT reduces limitations previously imposed by photobleaching. Here, DNA-PAINT could be used to label the motor or the stators of the track. One may still anticipate problems with PAINT; for example it is possible that the PAINT label may not be able to diffuse freely into and out of the tube, this may mean that the stators do not have access to the reservoir of unbleached strands. We have seen the enzyme can diffuse into a dimer and monomer tube (in Section 4.3) but the results for the longer stacks are not conclusive.

It may also be possible to create a track-motor system where the track lights up only once the motor has passed. This might involve the stator having a short double-stranded region which templates a stator comprised of two strands; one with a fluorophore and the other with quencher. The short strand would need to be displaced upon motor arrival. The enzyme cutting would result in the quencher strand being cut in half and both parts diffusing away, allowing the fluo-

rophore to fluoresce. This would provide another method of observing motor motion.

5

DNA TETRAHEDRA FOR SELECTION OF BINDERS

In Chapter 3 we briefly touch on how our DNA nanotube design may be modified to include a gene strand. Such a strand would enable assembly and testing of different enzyme combinations in a cascade, and provide a readable record of the assemblies.

Here we present two systems to develop a gene for selection experiments, based around a DNA tetrahedron. The first design separates the functional region from coding region by having them on different complementary strands. The second design is a single-stranded tetrahedron where DNA for recording and functional DNA are present within the same strand.

Both designs use elements of combinatorial assembly and have options to include strands that can be ligated together using click-chemistry. Click chemistry advice was provided by Tom Brown and Afaf El-Sagheer from the Brown group, now based in Oxford.

5.1 INTRODUCTION

5.1.1 *Combinatorial assembly*

Both drug discovery (based on screening large libraries) and biochemistry require development of compounds or molecules with specific function, but often discovery can be a lengthy process if every potential candidate is synthesized in multiple steps and analysed individually. Combinatorial approaches thus surpass such methods generat-

ing large libraries of compounds simultaneously by assembling every possible combination of a discrete set of chemical or biological building blocks [154]. Libraries can then be screened against a target using a high-throughput method, without the need for purification. Parallel synthesis is a method that is often employed to form combinatorial libraries, where discrete sets of compounds are prepared simultaneously in either physically-separate reaction vessels or on insoluble resin-based supports [155]. Despite the merits of this technique regarding large-scale synthesis, the properties of the resulting, largely stable, compounds must be thoroughly analysed for suitability, which can be time-intensive regardless of high-throughput screening [156].

5.1.2 *Dynamic combinatorial chemistry*

A dynamic combinatorial library (DCL) describes a 'one-pot', solution-phase, combinatorial library in which all components are in equilibrium. Constituent building blocks form reversible interactions and can continuously interchange with one another, offering the opportunity to access any conceivable, dormant library member. Library composition adapts in response to the experimental conditions to ensure that the overall Gibbs free energy is minimised [157]. Because the distribution of library members is typically governed by thermodynamics, altering the environment can affect their stability and produce changes in the composition of libraries. Library composition can therefore be adjusted in response to a variety of selection pressures.

One such selection pressure may be the introduction of an external template into the system, such as a ligand or biomacromolecule. This alters the distribution of the library members in favour of the ones that give rise to the strongest interaction between them and the target molecule. This DCL approach is advantageous over traditional

combinatorial methods because it is self-screening: it allows for the selection of new species with specific function directly, without further synthesis. However the technique still requires the chemist to identify the product at the end of the experiment so mixtures can be hard to deal with.

5.1.3 *DNA-encoded combinatorial libraries*

Techniques such as phage display and RNA-display use vast library sizes for selection and amplification. The power of these techniques lies in having a biological label or 'gene', which enables identification and amplification using biological techniques, such as PCR. Phage display is a technique for the production and screening of novel proteins and polypeptides. A gene fragment is inserted into a gene responsible for the surface protein of a bacteriophage. The new protein appears in the surface coating of the phage, where it can be manipulated and tested for biological activity [158]. Similarly in RNA display, proteins are associated with their messenger RNA (mRNA) progenitor. In this way certain proteins with a specific binding affinity may be identified by their mRNA [159].

Consequently DNA has been used to both broaden the repertoire of and add information content to chemical libraries where, in analogy with phage display, chemical moieties (phenotypes) are covalently attached to single DNA strands (genotypes). Two kinds of approaches are employed to make DNA-encoded libraries: in 'DNA-recorded chemistry' an oligo acts solely as a label for species identification within a pre-assembled library, whereas in 'DNA-directed chemistry', DNA plays a role in both library assembly and identification [156]. The latter method provides scope for combinatorial assembly using DNA.

For example, a subcategory of DNA-directed chemistry called encoded self-assembling chemical (ESAC) libraries aims to develop small molecule combinations for protein binding [160]. Pharmacophores are conjugated to DNA strands that contain two domains: one domain is a sequence for hybridisation to other DNA strands, and the other domain is a distal sequence that acts as a label for the fragment, for its identification. Target molecules have successfully been used to template the assembly of duplexes and triplexes that bring effective binding combinations of fragments into proximity with each other. ESAC library experiments have led to novel bidentate binders, and have also been used for affinity maturation of a lead fragment [161, 162].

In 2009 the Danish company Vipergen reported using a three-way junction to perform combinatorial DNA-templated synthesis of encoded chemical libraries [163]. This system combines the ESAC library method with elements of David Liu's work on DNA-templated synthesis whilst making use of three-way DNA junctions with hairpin regions [164]. Oligonucleotides were functionalised with peptides and amino-acids and contained domains for both self-assembly and identification. A library of pentapeptides formed following combinatorial self-assembly and chemical reactions between the building blocks at the centre of the junction. The identification region is made up separate sections of DNA which each code for one of the amino-acids assembled, and altogether provide a 'gene strand' record of the sequence of amino acids, which can be read out. Their so-called 'yoctoreactor' can be seen in Figure 43.

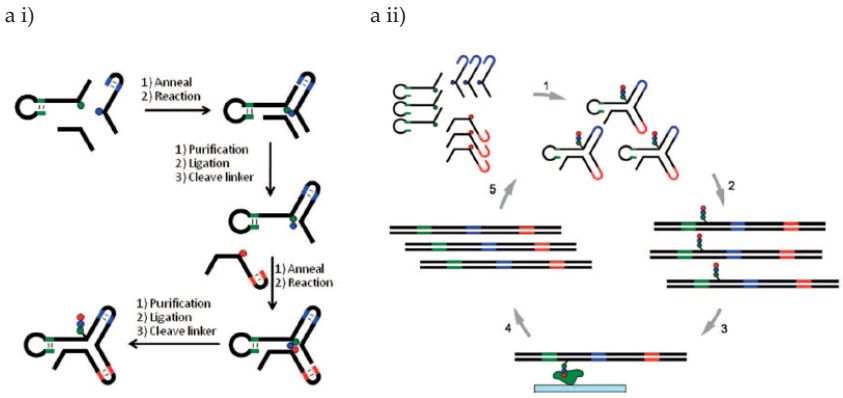


Figure 43: Diagram adapted from Hansen *et al.* Journal of the American Chemical Society, 2009, 131, 3, 1322-7 [163]. An approach to DNA-directed library synthesis, using a DNA-encoded library. The diagram shows the principles of the 'yoctoreactor'. The black moieties represent constant assembly sequences, and the colored moieties represent codons unique for the individual BBs (colored circles). (a) Stepwise assembly and reactions forming the yoctoreactor library. (b) Key elements of the yoctoreactor molecular evolution cycle: step 1, combinatorial library assembly (annealing, chemical reactions, and DNA ligations); step 2, reactor dismantling by primer extension; step 3, affinity selection by target binding and washing; step 4, elution and PCR amplification of enriched DNA library (optionally, the composition of the library is monitored by mass DNA sequencing); step 5, enriched display library formation by rolling translation.

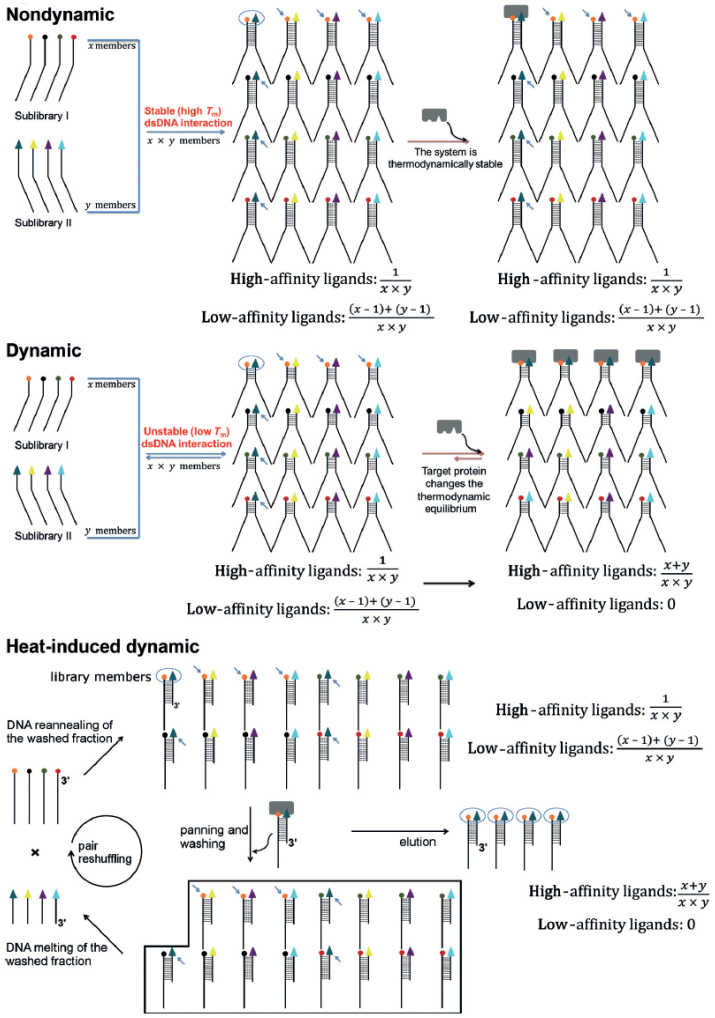


Figure 44: Diagram adapted from Reddavide *et al.* Angewandte Chemie, 2015, 54, 27, 7924-8 [165]. The principles of DNA-encoded dynamic combinatorial chemical library (EDCCL) technology and heat-induced DNA-encoded dynamic combinatorial chemical library (hi-EDCCL) technology. In the EDCCL method, thermodynamic instability causes pair reshuffling. Under selection pressure from a protein target, the constituents are shifted to the generation of potent bidentate binders. This shift results in an increase in the high-affinity-ligand signal and a decrease in noise from low-affinity ligands. When the oligonucleotides in the supernatant are melted, the two sub-libraries are reformed, and the building blocks are free to self-assemble again to give the high-affinity binders. The authors defined this concept as pair reshuffling. After some cycles of reshuffling and incubation in the resin, the high-affinity binders can be eluted and give a higher signal as compared to that after a single cycle of incubation as in the standard dual-pharmacophore selection. The relative abundances of high- and low-affinity bivalent ligands, under dynamic and non-dynamic conditions, are shown as the probability of their occurrence in the library before and after the addition of the target protein. Circles indicate the'correct pair' (potent binders) of pharmacophores in the library, whereas arrows indicate pharmacophores 'wrongly paired' in the library.

More recently, Reddavide *et al.* have demonstrated a DNA-encoded dynamic combinatorial chemical library [165]. This has a double advantage of sensitive hit identification from vast numbers of compounds (provided by DNA), and enhanced enrichment through shifting equilibrium towards generating strong binders (provided by dynamic assembly). In order to make a dynamic DNA library, the interactions of DNA must be weak enough such that they are reversible, yet strong enough to stabilise the structure upon creation of a strong ligand combination. Here, this is enabled by short complementary regions [165]. The working of this library are seen in Figure 44.

These examples demonstrate the power of DNA as a functional nanomaterial for chemical application, making full use of its dual capacity; as a means of information storage and as a robust building material allowing for templated assembly. So far, few and simple assemblies have been presented but these pave the way for more interesting structures for DNA combinatorial assembly.

5.1.4 *Click nucleic acid ligation*

Click nucleic acid ligation is the technology which enables us to carry out selection experiments here.

Solid-phase phosphoramidite synthesis is an efficient and widely-used high-throughput process for producing synthetic oligos. However, for synthesis of DNA strands exceeding 150 bases in length, the accumulation of sequence errors renders this method unsuitable [166, 167, 168]. In response to this, the exciting field of click nucleic acid ligation promises to provide a simple method for synthesising long oligos via entirely chemical methods [169]. In addition to using click for synthesis of long synthetic DNA molecules, click is also be-

ing applied in an increasing variety of roles; where enzymatic ligation is unsuitable and also for non-canonical DNA structures [170, 171].

Here 'click' is used to describe the very fast and efficient copper (I)-catalysed [3+2] azide-alkyne cycloaddition (CuAAC) reaction depicted in Figure 45, however the term click may be used to describe any reaction that simply and reliably joins small molecules together under mild reaction conditions [172, 173]. The CuAAC reaction was chosen for nucleic acid ligation because azides and alkynes are small, specific and easily appended to nucleic acids without disturbing their biophysical properties. The reaction between unactivated 5'-azides and 3'-alkynes is slow despite being very exothermic, thus templating with splints, bringing the azide and alkyne groups into proximity, and adding Cu (I) for catalysis, greatly accelerates the reaction. In order to take full advantage of click nucleic acids it is desirable for the triazole linkage to be biocompatible, therefore permitting *in vivo* applications. The Brown group have made the greatest efforts towards mimicking the natural phosphodiester linkage and Figure 45 shows the evolution of their click linkage. Their first generation linkage was successfully read by PCR but resulted in a loss of one nucleotide at the click site following amplification. The presence of a rigid amide bond or the absence of the 5'-methylene group and 3'-oxygen may make this backbone an incompatible substrate for polymerases [174]. The second generation linkage however, addresses these issues and presents a truly biocompatible DNA linkage that is functional both *in vitro* and in *Escherichia coli* [175].

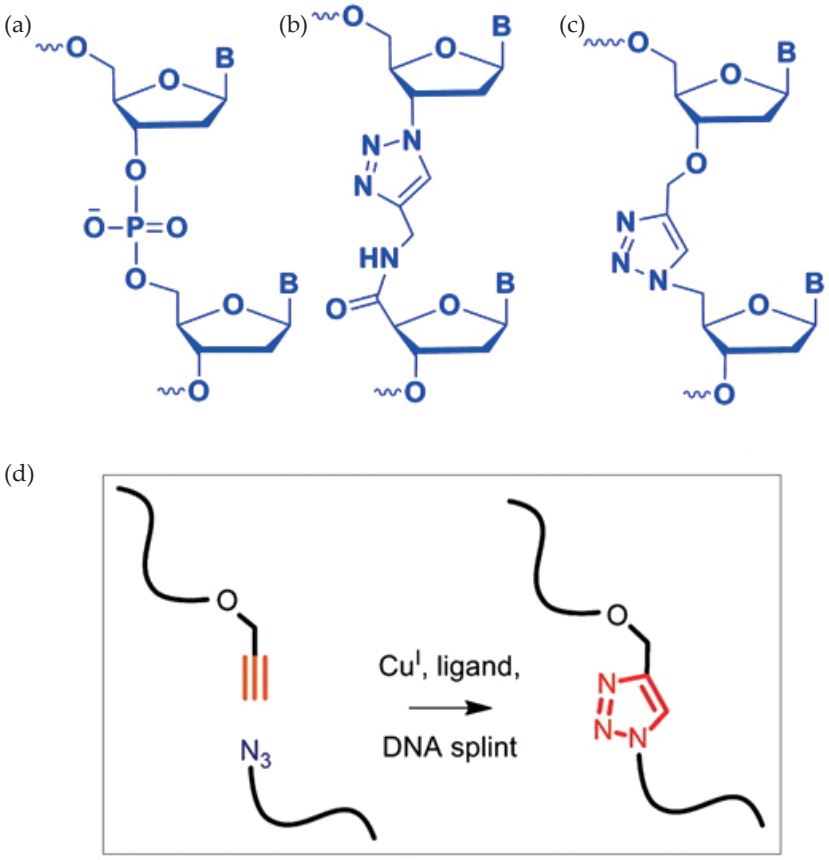


Figure 45: Diagram adapted from Afaf H. El-Sagheer *et al.* PNAS 2011, 108, 11338-11343 [175]. (a) Canonical DNA; (b) first-generation triazole DNA analogue [174]; (c) second generation biocompatible triazole analogue; (d) click ligation to produce triazole DNA.

5.2 DNA-CLICK TETRAHEDRON FOR LIGAND SELECTION

This project set out to create a new way for selecting protein binders using a library of 3D DNA scaffolds. Click nucleic acid ligation was designed to be included to create a record of the strongest binders. Originally it was thought that this work would be a collaboration with Dr. Wenjing Meng who was developing a library of chemical fragments of which a specific combination of three would bind and inhibit the protein tyrosine kinase. The fragments she was designing were to be used in this project.

The overall steps involved in such a process are; (i) DNA scaffold design, (ii) fragment preparation - this is done separately regardless of the DNA work, as this is pure organic synthesis, and (iii) selection. Stages (i) and (ii) were carried out, but despite a small library of fragments for selection being made, two reasons prevented continuation and completion of step (iii). Firstly, non-selective binding was observed on most solid supports chosen (in the cases where non-selective binding was not observed, nothing bound at all including DNA that was supposed to bind) - this would have given a high background noise. Secondly, very little DNA was recovered from the selection experiments and thus did not provide enough colonies to ensure a decent statistic calculation (for example, for a library of four, fewer than 10 usable colonies were grown on the agar). Both of these reasons made it infeasible to continue down this route.

However, the design of a 3D DNA scaffold library was under way at this point. There is scope for this design to be continued with a different method of fragment library selection, or in fact for a different binding/selection purpose altogether. The design will therefore be presented in general terms when it comes to the library, selection

and binding stages, and is not described for the selection of a specific protein.

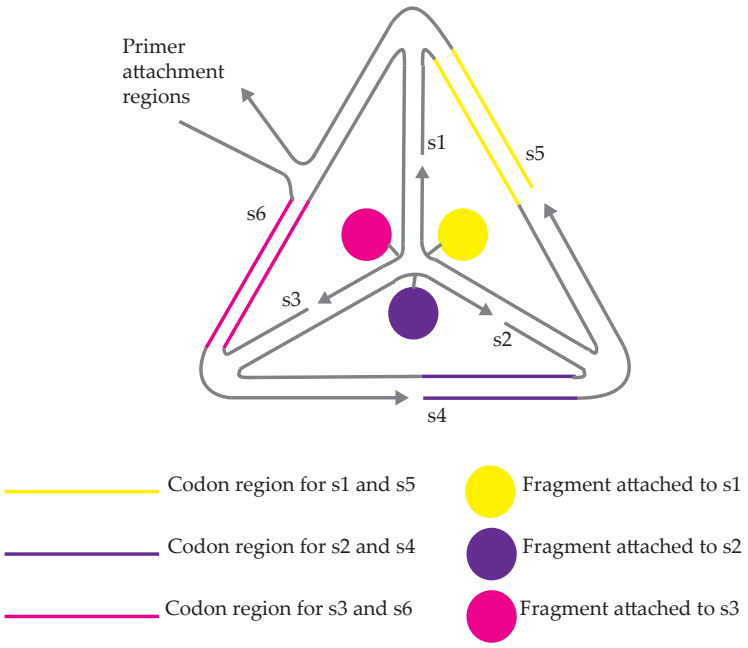


Figure 46: Diagrammatic representation a new 6-stranded DNA tetrahedron, with fragments appended at a vertex. Assembly brings together different combinations of fragments. Codon colours specify which fragment they code for. The labels s1-s6 indicate strands 1-6.

This project aimed to create an elegant way of selecting protein binders assembled from a library of fragments using a DNA scaffold. Towards this, a tetrahedron comprising six strands has been designed and assembled. The traditional four-stranded DNA tetrahedron, used previously in the Turberfield group, has a DNA strand running around the edge of each face [176]. Instead, here one of the four strands is divided into three smaller strands and two of these have primer attachment regions for subsequent amplification (s5 and s6 in Figure 46). The 63-base strands labelled in positions one, two and three in Figure 46 are functionalised with one of the fragments, which are brought together at a vertex when the tetrahedron is assembled. These strands also have a 14-base codon region which is unique to the associated fragment.

The shorter strands in positions four, five and six contain one of the 14-base codon complements and have modified 5′-azide and 3′-alkyne ends allowing for click ligation between them at two sites following templated assembly.

The first step in this system is to assemble a combinatorial library of tetrahedra from a set of strands that allows for all permutations of position with functionality, such that any combination of three ligands can be brought together in any of the three positions. This is illustrated in step one of Figure 47.

Once assembled, the strands in positions four, five and six (see steps one and two of Figure 47) of the effective library members can be ligated together using click to produce a gene strand containing two second generation triazole linkages. This is a biocompatible record of the ligand types and positions that have assembled. Next, the library mixture is introduced to the protein to select for the strongest binders. PCR can then be performed, amplifying the gene strand and finally, this can be sequenced to tell us about the frag-

ment types and positions that have been selected for. The latter steps described correspond to steps three and four shown in Figure 47.

The advantage of using a 3D scaffold to form the library is that it is possible to separate the functional binding region (the vertex formed between strands one, two and three) from the recording region (the gene strand formed from strands four, five and six).

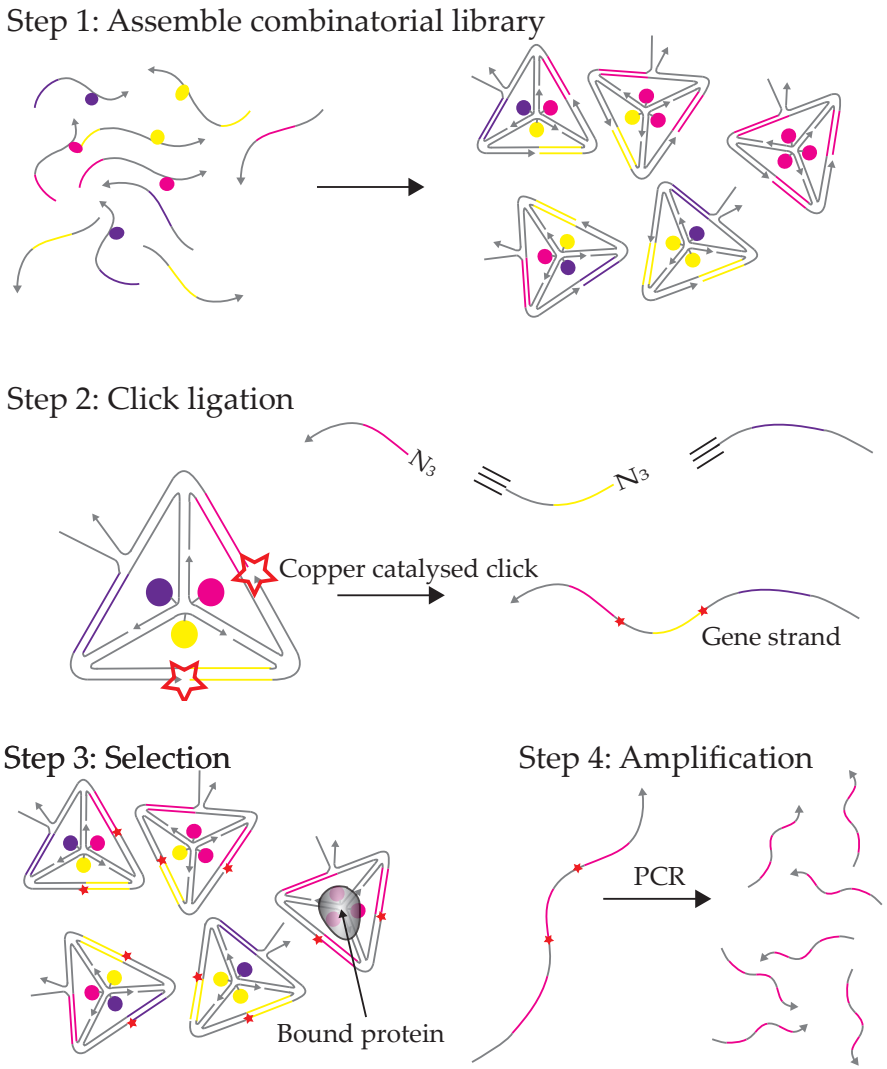


Figure 47: Experimental procedure showing combinatorial assembly process, in which selection comes after the assembly and click steps. For Step 3, the protein only binds to a specific fragment arrangement and so only those tetrahedra are selected for and amplified. If a library member consists of n fragments and each fragment has m different types, the recording strand required will be mxn rather than more traditional m^n.

5.3 PROGRESS TOWARDS A LIBRARY

5.3.1 *Results*

Initially, one version of the new tetrahedron was assembled by annealing at 95°C followed by step-wise cooling to room temperature. Instead of using the whole library of strands with multiple codon regions at this point, a tetrahedron with known, distinct codon regions is used to check assembly. The assembly was characterised using native polyacrylamide gel electrophoresis (PAGE).

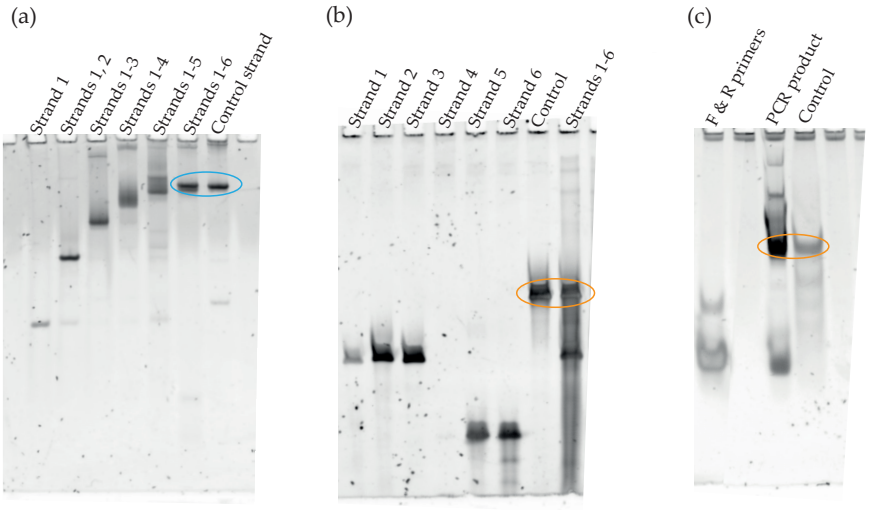


Figure 48: (a) Stepwise assembly of the DNA tetrahedron, 10% 19:1 acrylamide PAGE, 180 V, 55 minutes, 1 pmol DNA, (b) denaturing PAGE of the clicked tetrahedron product, 10% PAGE 19:1 acrylamide, 1 x TBE, 7 M urea, 20 W, 20 minutes, 1.3 pmole DNA, (c) 10% PAGE of the product following PCR of the assembled and PAGE-purified click strand, F and R indicate forward and reverse primers respectively. In (a) the assembled tetrahedron and the control tetrahedron (made with a ready-formed gene strand) are highlighted in the blue ellipse, and in (b) and (c), highlighted in the orange ellipses is the gene strand, showing successful click and amplification. Again, the control gene strand is bought from IDT ready-assembled and is used as a reference for the click product.

The click ligation conditions for the codon strands were optimised to achieve the best efficiency for this system. The reaction requires Cu(I) for catalysis which is provided from the reduction of Cu(II), in the form of copper sulphate, by magnesium ascorbate. A ligand is also present in the reaction to prevent oxidation of Cu(I) back to Cu(II) [168]. Thus by varying the amounts of copper, magnesium ascorbate and ligand, the efficiency and yield of the reaction can be significantly improved or degraded. Degassing of every component at each stage is required to prevent oxidation of the copper. Click experiments were characterised using denaturing PAGE. Following the click reaction, a new band on the gel appears indicating successful ligation as is evident in Figure 48 b. A full description of the protocol and assembly conditions for click are presented in the Appendix A.1.2.

It is possible to see that the reaction is not completely efficient, but also that side products form resulting from the multiple click sites in this system; they can just be seen in the smearing in Figure 48 b (the smearing is due to the denaturing buffer containing urea). The untemplated reactions between strands five and six, and between four with itself, are possible but are much slower than the templated reaction. This is why the desirable product is the most abundant band. Before amplification, this band is cut out from the denaturing gel, and only this is amplified which avoids amplification of an undesired product.

The click strand is successfully amplified in 28 rounds of PCR using Taq-polymerase. The resulting gel band matches that of a control gene strand of the expected click product as seen in 48 c, demonstrating the correct product is amplified.

5.3.2 *Further work*

The next step for this project is to implement the process with the full library of strands, to allow for the combinatorial selection experiment to proceed. This will first require the conjugation of fragments to functionalise the strands. Further characterisation of the tetrahedron assembly with the fragments present should then be carried out. The fine experimental details of the selection experiment will also need to be addressed and different methods of separation following templated assembly should be tested, to recover the effective library members.

In general, this current approach creates a static combinatorial library and the initial assembly of library members is irreversible. It would however be elegant to alter the design such that the library becomes dynamic and that members are created upon introduction of the target, shifting the equilibrium towards favourable binders. This could be done in a similar way as is presented by Reddavide *et al.* and mentioned above, by shortening the complementary regions between strands one, two and three [165].

5.4 ADAPTING THE DNA TETRAHEDRON FOR SELECTION OF CO-OPERATIVE APTAMERS

5.4.1 *Introduction*

The clickable tetrahedron in the first half of this chapter allows for the labelling and subsequent identification of strong binding combinations of three ligands. However, it does not allow for the original structure (or a copy of it) to be reassembled after identification from the clicked strand, and the labelling 'codon' sequence is on a sepa-

rate strand of DNA to the ligand-labelled strands. A natural progression from this tetrahedron is to imagine a tetrahedron formed as the secondary structure of a single strand of DNA. This has the advantage that it can be cloned and thus replicated as described by Figure 50. In 2009 Hao Yan's group presented a tetrahedron that could be formed from three oligonucleotides and ligated together to form a single strand; this was also inspiration for the design presented below [177]. Such designs require clever routing of DNA through the structure, to respect its topology. A motif known as the paranemic crossover enables such routing and is presented below. This section also looks at the inclusion of specific sequences of DNA, which allow for binding of proteins, into our structure to create functionality.

5.4.1.1 *Paranemic crossover motif*

A paranemic crossover (PX) motif describes another way of forming a four-stranded unorthodox DNA junction, similar to the double crossover motif described in Section 1.1.2 of Chapter 4 and in Section 1.1.2 of the introductory chapter. Such a motif is formed when strands of the same polarity in two co-axial helices cross over to the other helix, as illustrated in Figure 49. When each strand reaches the central axis of the motif it crosses to the next helix, resulting in the period of the motif being twice that of a regular double helix [178].

If one examines the part of the PX molecule flanking the central dyad axis, crossovers occur with alternating frequencies, depending on whether a major groove or a minor groove flanks the axis; these different spacings are indicated as W (for wide, or major groove) or N (for narrow, or minor groove) [178]. A minor-groove separation of five bases together with a major-groove separation of six, seven and eight bases produces a stable motif [179, 178]. Paranemic assemblies have been demonstrated in both RNA and DNA, but the smallest 'minimal'

motif has only been formed using RNA [180]. This minimal motif is made from three half-helical turns of RNA with just one crossover spanning the major groove. Instead of forming the junction with four separate strands of DNA, it is possible to form a PX with two closed dumbell strands, demonstrating that the interaction between the strands is entirely Watson-and-Crick based and not plectonemic [178]. This implies the motif formation is reversible as no bases or secondary structure have to be destroyed when the component strands are separated. The ability to form reversible tertiary interactions may be of use for DNA nanotechnology and self-assembling structures.

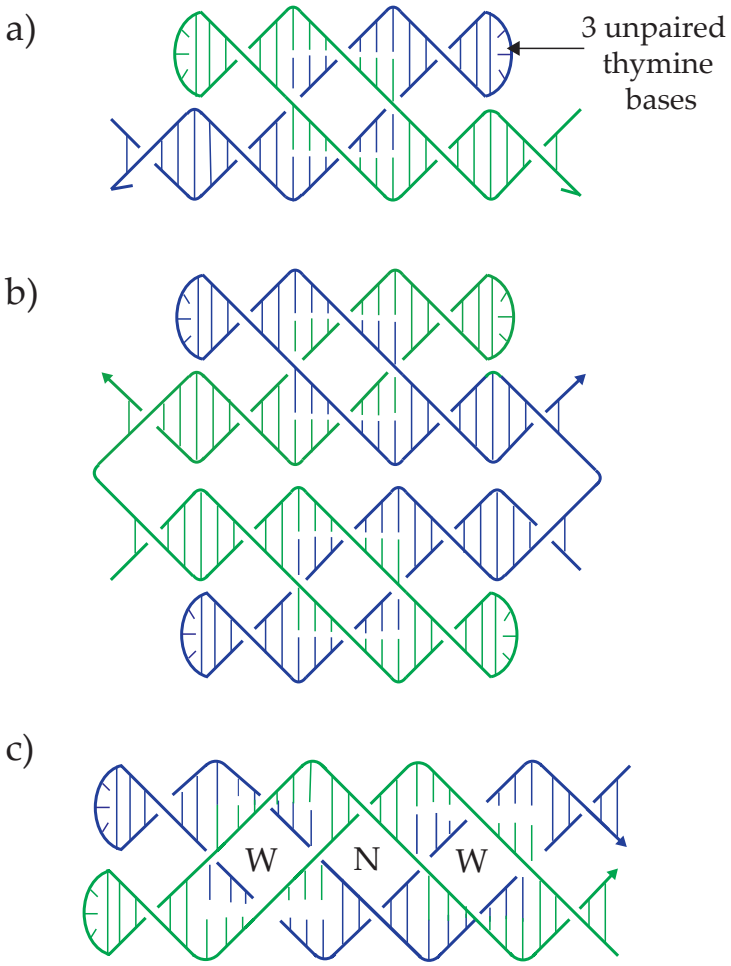


Figure 49: Paranemic crossover motif. The backbone of the PX motifs describes the region of the strand that hybridizes to itself, whereas the crossover sequence is a region that hybridizes to a separate strand. (a) Minimal PX motif. The three unpaired thymine bases are indicated by the arrow. They are present in each loop to allow flexibility for the strand to fold around. (b) Dimer of minimal motif linked by unpaired thymine bases. (c) PX motif with two cross-overs, drawn showing the major and minor grooves indicated by W and N for wide and narrow, respectively. This example is drawn to include the major and minor grooves, for ease of depiction they are not drawn in a or b.

5.4.1.2 *Nucleic acid aptamers for thrombin binding*

Nucleic acid aptamers are single-stranded DNA/RNA strands with binding affinity for small-molecule targets, proteins and cells. The molecular recognition properties of aptamers make them analogous to antibodies but have the added advantages of being robust and easily engineered [181]. Although some exist naturally, the real benefit of aptamers is that they can be developed for specific targets that did not previously have any known binders. Aptamers have great potential for biomedical and biotechnological applications, including recognition and therapeutics, and molecular sensing, respectively [182, 35].

Two groups independently and simultaneously came up with methods for discovering new aptamers. Gold and Tuerk reported 'systematic evolution of ligands by exponential enrichment' termed SELEX [183], whilst Szostack and Ellington described an 'in vitro selection method' [184]: both terms are used widely and interchangeably. The method begins with library synthesis of strands containing a random-sequence region. This is followed by multiple rounds of selection, under increasingly stringent conditions, and then amplification, which has the aim of enriching the library with the strongest binders.

Using this method, aptamers have been selected for many different targets and purposes. Thrombin, a protease involved in haemostasis, has two exosites for which aptamers of both DNA and RNA have been developed. The aptamers for the fibrinogen binding exosite were found to contain a highly-conserved 15-mer sequence that strongly binds to thrombin, with dissociation constants of 3 nM to 450 nM [185, 186]. The 15-mer thrombin binding aptamer (TBA) forms when guanine bases associate to generate G-quartets and these quartets assemble to form a four-stranded non-canonical DNA structure known as a G-quadruplex. Formation of G-quadruplexes requires the presence of monovalent cations, such as potassium. These bind in the cen-

tral cavity between the carbonyl oxygens and stabilize the structure [187]. The second thrombin exosite aptamer, HD22-27mer, is a duplex-quadruplex structure widely used as a sensor due to its higher affinity than TBA [188].

Amongst the many aptamer systems that have been developed for thrombin detection, are those that make use of thrombin bivalency [121]. One such system describes how the nucleic acid linker between two aptamers need not be passive, but can itself be evolved to enhance the overall binding of the structure [189]. This suggests that multivalency in general, may lead to improved binding.

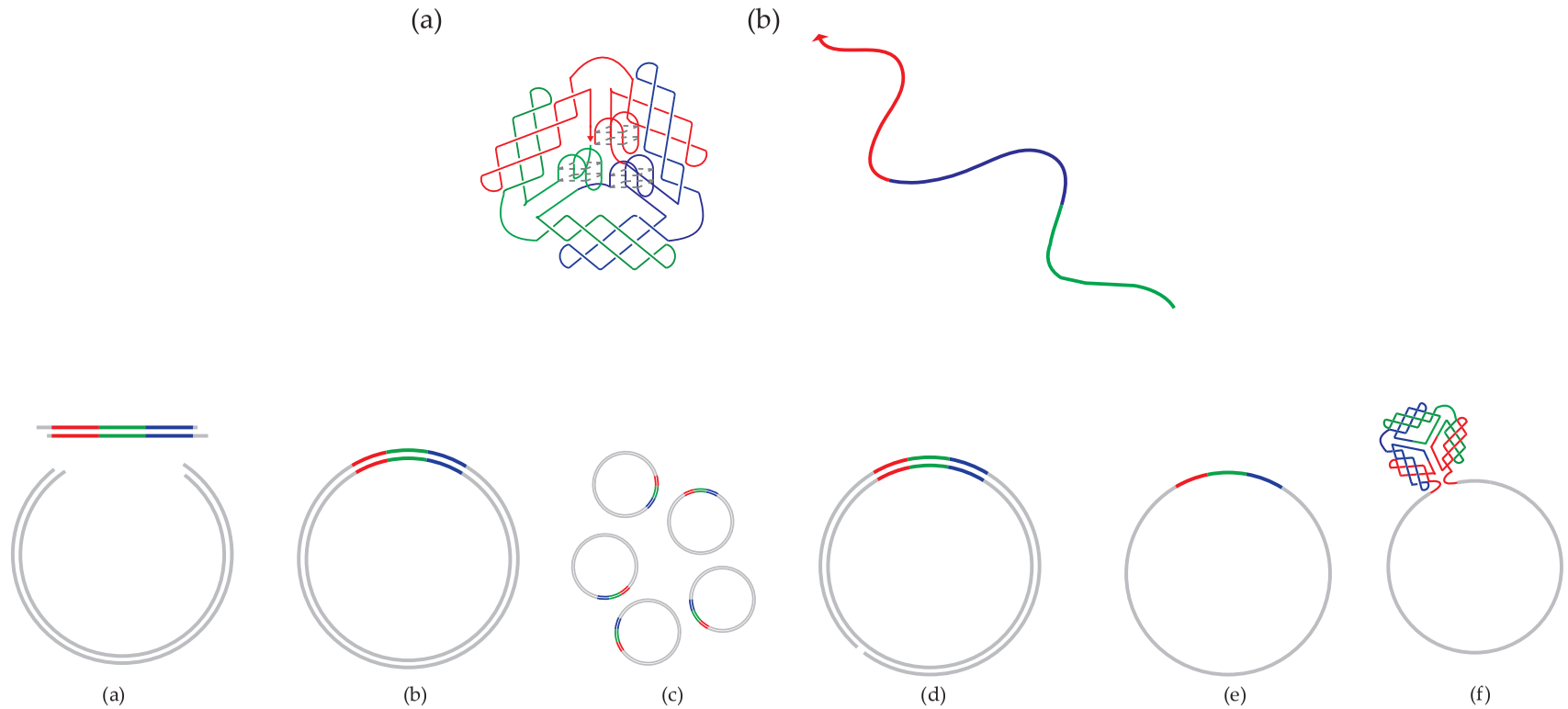


Figure 50: (a) The paranemic tetrahedron consists of three strands which can be ligated at either two or three points to create a single linear or circular gene strand. The paranemic tetrahedron is drawn here as double-lines, which respresent double-helix regions. The bases are not drawn on in these instances for clarity of the position of aptamers and for PX regions. The apatamers are illustrated here as thrombin-binding aptamers but may be any sequence, for example one selected for enhancd binding. From hence forth, the paranemic tetrahedron will be drawn without the aptamers to ensure clarity of the diagrams. Please refer back to this diagram if in doubt as to where the aptamers are situated, (b) This diagram shows how upon ligation of the tetrahedron at two points, a continuous linear strand is formed which may be amplified and read via sequencing, (c i-iv) shows how it is possible to amplify the gene strand using *E. coli* by incorporating restriction sites into the ends of the gene strand. This allows for double digestion of the vector and strand, ligation into a plasmid, transformation into bacteria and replication. A similar approach may be used to produce random aptamers by binding random sequences in loops to a ss-vector+gene-strand plasmid and using DNA polymerase to create a circle which may be inserted into and amplified by bacteria.

5.4.2 *The paranemic tetrahedron*

Here we present a tetrahedron consisting of three ligatable oligonucleotides that form a continuous strand following ligase action. Each of the three constituent strands have two regions for minimal paranemic motif assembly and an aptamer region. Upon successful assembly, the three aptamer regions are brought together at one vertex of a tetrahedron, analagous to the clickable tetrahedron fragments described above. This structure is known as the paranemic tetrahedron and is shown in Figure 50 a.

Originally the aim was to develop a multivalent aptamer binding complex for thrombin, using the two known aptamers TBA and HD22 in conjunction with a third that would be developed through rounds of selection experiments. Here we would hope to see that multivalency gives rise to a binding which is greater than the sum of the three individual aptamers. This is one example of how the paranemic tetrahedron may be used, and the initial steps taken toward incorporating thrombin binding are presented in the Appendix. However, this system may in theory be used for the binding and selection of any protein or complex and so the remainder of this chapter presents the paranemic tetrahedron binding in general terms and focusses instead on structure design and assembly.

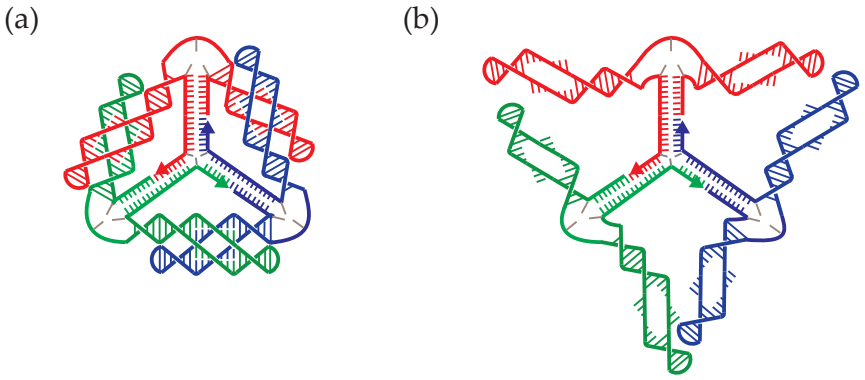


Figure 51: Paranemic tetrahedron (a) A well-formed paranemic tetrahedron with all three PXs bound. Each side of the tetrahedron is 21 bases or two helical turns. The grey bases are unpaired thymine bases included for flexibility. there are also unpaired thymine bases in the loops of the PX regions, but they are not drawn so as not to overcrowd the diagram. (b) The structure may resemble this if the PXs do not bind. There are 12 bases involved in the PX however, thus energetically they should bind.

5.4.3 *Results*

In order to be the correct size to make up the edge of the traditional DNA tetrahedron, the PX module needs to be the smallest possible 'minimal' motif [180], with an extra half-turn at one end. This means the total length of each edge of the tetrahedron is 21 bases, which is the same as the clickable tetrahedron.

The initial step is to characterise the minimal motif formation because prior to this, it has only been demonstrated in RNA, as described in section 5.4.1.1. The sequence of the minimal RNA motif previously reported by Afonin *et al.* was assembled in DNA rather than RNA, substituting uracil for thymine bases [180]. Both strands involved in the RNA motif have G-C rich regions at either end and form a stable backbone. The DNA version assembles successfully confirming that such a structure is stable. However, upon trying other sequences of the minimal motif, not all are found to bind successfully suggesting that its formation is sequence dependent. Following assembly of carefully designed minimal-motif crossovers, with G-C rich backbone sequences, three orthogonal motifs for use in the paranemic tetrahedron were chosen. The strands are shown in figures 51 and 50 in red, green and blue and are known as R, G and B respectively and the strand sequences are all shown in Appendix A.3. The assembly method is described in the following paragraph.

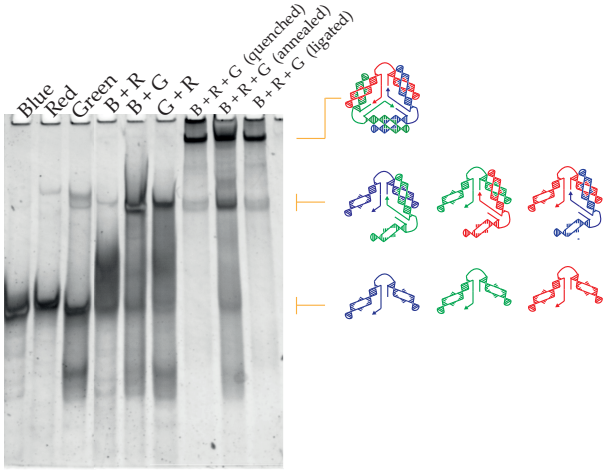


Figure 52: Paranemic tetrahedron assembly. The paranemic tetrahedron is drawn here with double parallel lines respresenting double-helical regions. The bases are not drawn on in these instances for clarity of the position of PX regions. Please refer to Figure 51 for the positions of bases. The first three lanes show individual strands B, R and G. In addition to G, there is a faster band which may be ss-G, having not all formed with the designed secondary structure. The middle three strands show pairwise combinations of R, G and B and it is possible to see how the binding at this step is not entirely efficient due to the smears and unbound strands left over. The final three lanes compare the tetrahedron: one set of DNA is assembled via quenching and is then subject to ligation at three points (creating a circular exonuclease-resistant structure), followed by exonuclease action; in the next lane the DNA is assembled by annealing; in the final lane the DNA is annealed together, ligated and then exonucleased to digest any uncircularised product. The gel is 10% PAGE, 29:1 acrylamide and was run at 210 V for 71 minutes.

Figure 52 shows the formation of the paranemic tetrahedron. Each strand is either added together with a stochiometric ratio of 1:1:1 R:G:B and annealed at 96°C followed by step-wise cooling to room temperature, or each strand is individually heated to 96°C and immediately quenched on ice, then the strands are mixed and incubated together for one hour at 37°C. The default method is quenching and is the method of assembly for all single, pairwise and tetrahedron combinations in the gel, apart from the tetrahedron that is annealed and is labelled as such. It is also the method employed to assemble the minimal motifs described in the previous paragraph. The assembly takes place in 1 x TE buffer with 15 mM $MgCl_2$. Both the single R and G strands show dimer formation which runs to a similar position as the pairwise combinations of strands, as might be expected. Not all of G forms the structure indicated in the accompanying diagram; the presence of a faster band indicates some of G may be entirely single-stranded, or only partially formed into the PX structure. The pairwise combinations in the middle three lanes show that the assembly of a PX between two strands is not entirely efficient, with leftover half-PX motifs and ss-G strand too. This is carried over into the formation of the three-PX tetrahedron which additionally shows leftover pairwise combinations of strands. However, the efficiency seems greater for the paranemnic tetrahedron assembly compared with pairwise assemblies, indicted by the dimmer bands of leftovers in the final three lanes.

Figure 51 b shows how the paranemic tetrahedron may exist if the PX regions are not formed successfully. It is not possible to determine which structure from Figure 51 (a or b) is the final band in the gel assembly in 52, without further investigation. Figure 53 shows how by binding a Cy5-labelled version of the R paranemic motif (Cy5R) to B, it is possible to see if the correct closed version is formed upon ad-

dition of the R+G complex. If the tetrahedron assembles with closed PXs then Cy5R can be inferred to have been displaced, but if not, Cy5R remains bound to B and the PX between G and B may be open or closed.

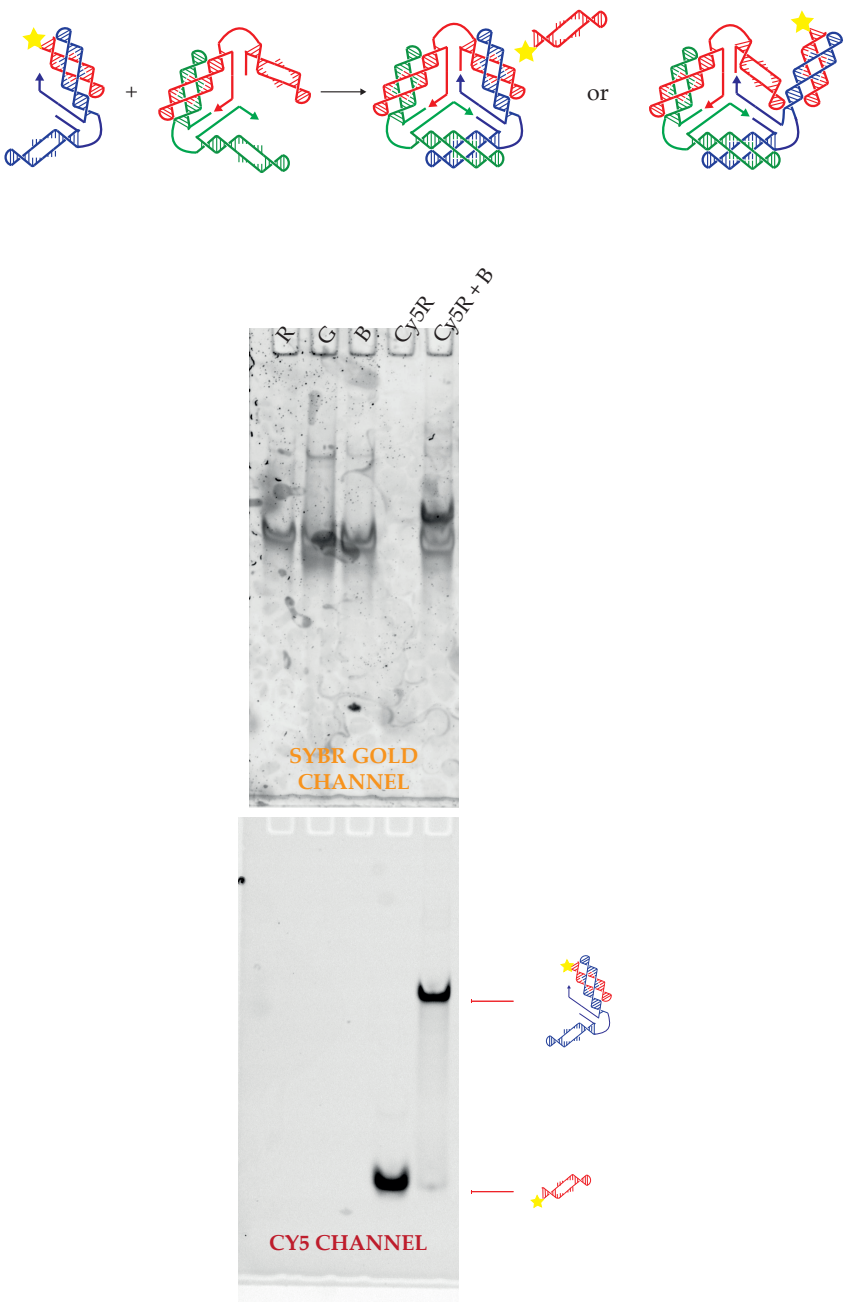


Figure 53: Please note the paranemic tetrahedron is drawn here with double parallel lines respresenting double-helical regions. The bases are not drawn on in these instances for clarity of the position of PX regions. Please refer to Figure 51 for the positions of bases. Top: Diagrammatic representation of what may happen when we use Cy5R to check the PX region assembly. The final tetrahedron structure may be well formed with fully displaced Cy5R (left), or it may be still bound or partially bound with a mixture of closed/open PX regions (right). Gel: In lane 5 of the gel we see that all the Cy5R is used up making Cy5RB duplex, indicating the PX forms well in this instance. DNA stuck in the well may be aggregate structures formed from interactions between many different motifs. The gel is 10% TBE and was run for 35 minutes at 230 V.

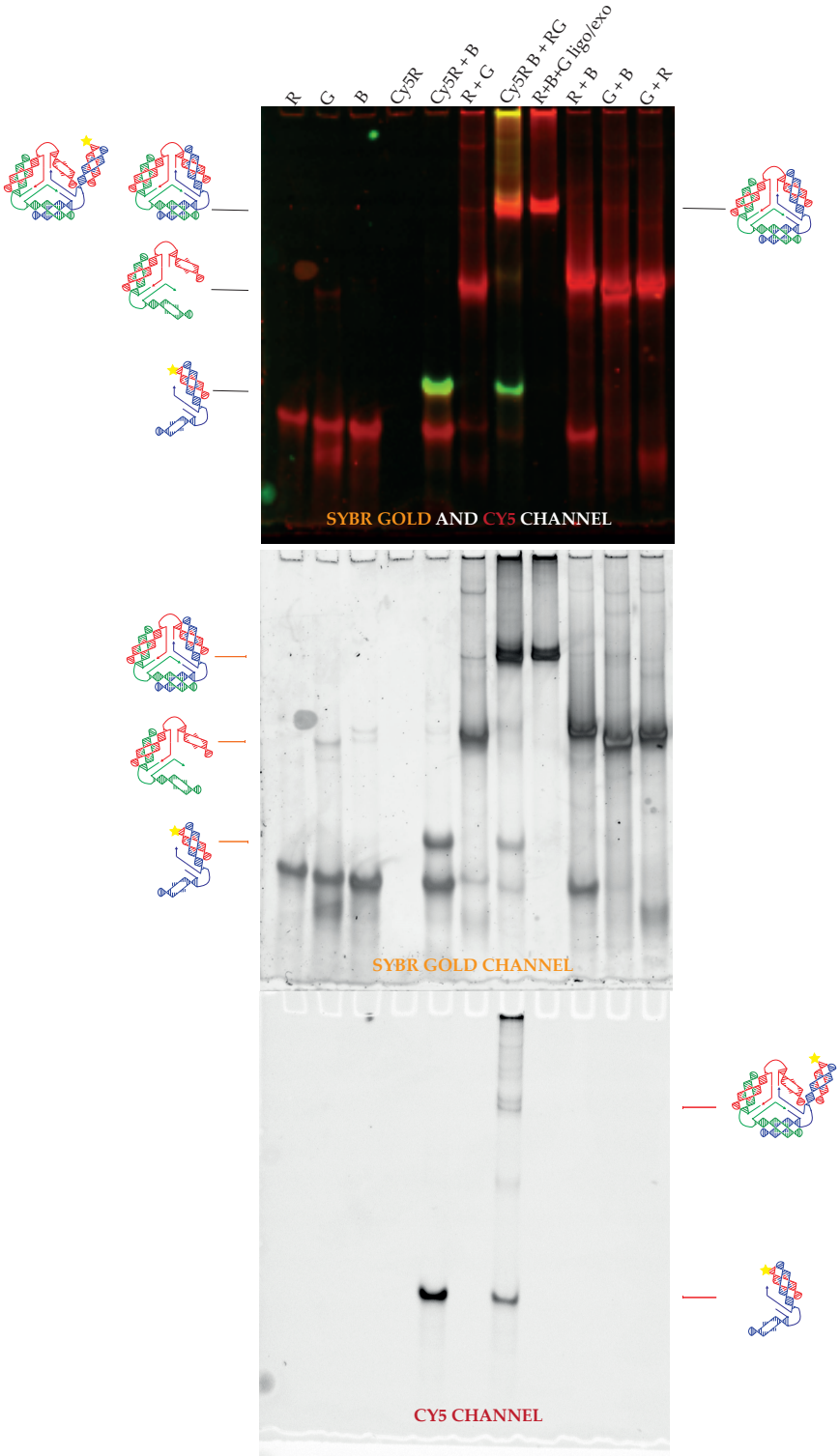


Figure 54: Testing the assembly efficiency of PX regions in a paranemic tetrahedron. Please note again, the paranemic tetrahedron is drawn here with double parallel lines respresenting double-helical regions. The bases are not drawn on in these instances for clarity of the position of PX regions. Please refer to Figure 51 for the positions of bases. The gel was run for 71 minutes at 210 V and is 10% TBE. The product in the 'Cy5RB + RG' lane runs to the same plce as the RGB complex that has been ligated and exonucleased. In this instance there are three ligation sites and it forms a continuous circle which is nuclease resistant.

Looking at the Cy5 channel in Figure 53 we see that the PX between Cy5R and B forms well; all the Cy5 in the lane is used up and the only binding between Cy5R and B is paranemic. The gel in Figure 54 also includes pairwise assemblies of strands R, G and B for reference and is run for a longer time than the prior gel. Looking at the Cy5 channel of Figure 54, the Cy5RB+RG lane has a very small amount of Cy5R still bound to B in the tetrahedron and there is some still bound in Cy5RB duplex form. Comparing this lane with the same lane in the Sybr Gold channel, the band corresponding to the tetrahedron is much darker than in the Cy5 channel, indicating that only a very small amount of Cy5R remains bound (some is displaced and most ends up in aggregates). This implies that some of the DNA forms the correct tetrahedron structure with closed PXs, but that some PXs may be unbound, giving rise to the partially open structure illustrated at the top of Figure 53.

5.4.4 *Further work*

We leave this project in a similar position as the clickable tetrahedron. Here we have set the scene for aptamer selection using a single-stranded 3D structure by designing, assembling and testing the feasibility of such a structure. However, there is still great scope and opportunity for developing this concept further and applying it to a selection system.

6

SUMMARY AND OUTLOOK

In this chapter we summarise the achievements of the work presented in this thesis.

This thesis set out to further the field of functional DNA nanotechnology through the design of novel functional DNA scaffolds, and investigate their applications and efficacy. The work presented here comprises two parts: the design of a chiral tube that acts as scaffold for motor motion and for an enzyme cascade; and the design of two different tetrahedral scaffolds for selection of a combination of three ligands, which together have a greater binding effect than the sum of the individual components.

In Chapter 1 we introduce the DNA molecule and see how it is used as a building material instead of a genetic material. We give a brief overview of DNA nanotechnology and learn how origami structures are designed and assembled.

In Chapter 2 we begin looking at previous designs of DNA nanotubes and acknowledge that many designs fail to distinguish, at the design stage, which face of an origami assembles to form the inside or outside face of the tube. We also look at how flat and curved origamis are formed. Next, we present the design of a DNA chiral nanotube, designed to assemble in a pre-determined manner, such that the scientist has knowledge of which face of the origami will form the inside and which face will form the outside of the tube. The chirality is introduced by changing the positions of staple-crossovers. Both 10-helix and 20-helix designs are presented and the option for end-to-end tube stacking is introduced by the inclusion of different

sets of 'end-staples'. The formation of tubes is confirmed by AFM and the chirality is confirmed by CD. There is scope to improve the design by adapting the method of end-to-end stacking such that two adjacent tubes each have only one end activated for polymerisation, instead of both.

In Chapter 3 we take a look at how nature uses compartmentalisation and structural arrangements to streamline cellular processes, where otherwise they are slow or would never occur. Many DNA origami designs have drawn inspiration from the cell in this respect and have assembled components upon synthetic scaffolds to find an enhanced activity compared to untethered arrangements. Often, the underlying physics is not fully understood or explained, although guesses are made about the underlying mechanism. We look to a paper by Hess *et al.* [129] which attempts to understands such assemblies through modelling. They note that scaffolds only have a beneficial effect early on in the assay, where the flux from the direct path (enzyme-to-enzyme) dominates over the flux from the bulk solution (enzyme-bulk-enzyme). They also indicate that charged scaffolds, such as DNA, may provide a surface to which oppositely charged substrates are attracted, such that inter-enzyme substrates are not lost to the bulk and stay in the direct flux path for longer. We use our origami tube to test its efficacy as an enzyme scaffold and see if tethering enzymes inside the tube reduces the lag time compared to enzymes tethered outside of the tube and enzymes free in solution. Qualitatively, we see a slight reduction in lag time for enzymes inside the tubes, compared to enzymes outside and free enzymes. However, we have not determined whether the difference is significant or not, thus further experiments and analyses would shed light on this. In addition to our set-up, there are many more set-ups that could be explored. Particularly ones including competitive side-reactions, which

may give an indication as to the underlying physics of our system. The tube design may be furthered by the inclusion of a combinatorial region of scaffold which may select for different enzyme assemblies.

In Chapter 4 we use our chiral tubes to tether DNA tracks for a burnt-bridges motor. We aim to assemble a system with a motor that can undergo controlled, inter-origami motion within a confined space (inside the tube). This is a demonstration of synthetic, directed transport of components in a separate compartment to the bulk solution and draws inspiration from biology, where compartmentalisation is used in transport. Ensemble measurements test the experimental set-up for the TIRF microscopy measurements that follow. The ensemble measurements show the motor can move inside the tube, and thus that the enzyme enters inside, that the motors can step between the origami tubes, and that the motor can transport a bulky fluorescent cargo. The TIRF results are only very initial but do indicate that the motor is moving. This is seen from observations of a fluorescent track, where the fluorophores becomes delocalised upon enzyme cutting and motor motion. It is thought that flashes in a video are the cut-off fluorophore strands coming into and going out of TIRF focus, indicating enzyme cutting. A decrease in fluorescence over time for samples of motor and enzyme (test sample), and excess motor with excess enzyme (positive control), is thought to result from enzyme cutting (indicating motor motion), not only photobleaching. This is found upon comparison with a negative control sample, which did not undergo such a decrease in brightness over time when no enzyme is added. These experiments could be continued to observe direct motor motion using PAINT, which would reduce the negative effect of photobleaching.

Chapter 5 uses a different functional scaffold to the previous chapters - a tetrahedron instead of a tube. The chapter outlines the designs

and assembly protocols for two selection systems. Both designs aim to select a combination of three ligands, which together produce an enhanced binding effect greater than the sum of the individual components. The first tetrahedron is designed to use click chemistry to ligate three strands running around the base, which have sequences that code for the ligands assembled at the vertex. Ligation produces a 'gene' strand which can be read via sequencing, to learn the nature of the assembled ligands. Selection experiments were not carried out but this system would allow for a read-out of the strongest binders, following selection. This tetrahedron system was furthered to produce a second tetrahedral selection system, made entirely from DNA. This system combines functional apatamer regions into the structure itself (where before the functional ligands were on different strands to the coding region), such that when three strands of DNA are ligated together the whole structure is formed from one DNA strand. This makes use of PX motifs to obey the topology of DNA and allow such a structure to fold. Again, no selection experiments were carried out, but this design paves the way for an interesting 3D DNA selection system, made from one strand, that can be both read via sequencing and amplified and replicated in *E. coli*.

6.1 CLOSING REMARKS

Functional DNA nanotechnology is expanding the field of structural DNA nanotechnology from just an academic, proof-of-principle discipline to one whose applications are far-reaching and potentially beneficial.

For functional systems to become a reality, the cost of producing high-quality DNA, which has decreased in recent years, will have to decrease further to make it a competitive technology.

Many DNA systems have been proposed as potential vehicles for drug delivery. However, careful thought is still required within the field to consider how programmability, detecting and reporting capabilities, and molecular motion devices, may all be brought together to form a conceptually novel drug delivery device, which should be possible. This would offer something more sophisticated than the introduction of static DNA vesicles, as is practiced today [190].

An article by Krishnan and co-workers outlines the anticipated molecular, cellular and organismal pathways that DNA nanodevices will encounter when they are introduced into higher organisms [191, 190]. The trend of DNA nanostructures being introduced into cells, makes work such as this crucial if the field is to progress along this path, where devices not only deliver to cells but also interact and re-program cells.

Although macro-scale self-replication is becoming something of a trend, notably with the increasing affordability of 3D printers, for self-replication of synthetic devices on the nanoscale, progress is much slower. One example aims to make progress in this area: DNA tiles contain instructions for replication of daughter and grand-daughter tiles [192]. However, this process still requires human interjection and assistance. For DNA nanotechnology to be truly clever, it would be exciting to see more effort in this area of the field.

A

APPENDIX

A.1 ASSEMBLY PROTOCOLS

A.1.1 *From Chapter 2: Plasmid synthesis*

PCR OF M13MP18 REGION 5 µl of both 1 µM Forward and Reverse primers are added to 2.38 µl M13mp18 template, 25 µl 2 X *Taq* Master Mix (New England Biolabs) in 50 µl total. 28 rounds of PCR are carried out with an annealing temperature of 55 °C.

DIGESTION Restriction double digests of both the PCR fragment and the pUC19 vector are carried out, followed by PCR clean-up kit. Digests occur for 1 hour at 37 °C followed by heat-killing for 20 mins at 80 °C for HindIII and SalI, in NEB buffer 4. The template strand is dephosphorylated using *Antarctic phosphatase* at 37 °C for 1 hour and heat killing after at 65 °C for 5 mins.

LIGATION 1 µl ligase is added to 6 µl cleaned PCR product and 2 µl vector in 1 X ligase buffer and incubated at 16 °C overnight.

TRANSFECTION Competent DH5α *E. Coli.* cells are thawed on ice for half an hour. This is followed by a heat shock at 42 °C in a water bath for 30 seconds. The cells are left to recover on ice for 5 minutes. 200 µl of SOC medium is added to the tube and mixed. 100 µl of the mix is then plated out onto an agar selection plate and spread

with plastic beads. The plate is incubated overnight at 37 °C. The plasmid should contain the ampicillin resistance gene and only if this is properly incorporated should it grow on an agar+ampicillin plate.

TRANSFORMATION LB am ampicillin broth is prepared by adding 50 ml LB and 50 µl of ampicillin. 2 ml broth is added to falcon tubes and separate colonies added to individual tubes by dabbing a pipette tip into a colony and waggling in the surface of the broth. Loosely screw on caps and incubate at 37 °C overnight.

ALKALINE-LYSIS 'MINI-PREP' The DNA needs to be recovered from the cells and so the cells must be disrupted. Using a Qiagen Mini-Prep kit: Burst cells using alkaline solution to release cell contents. Neutralise the lysate by adding acid, which also denatures proteins in cells so they become insoluble. Centrifuge lysate to separate out protein components (discard pellet). Purify remaining lysate in a column, which binds only plasmid DNA and elute with elution solution.

The resulting recombinant plasmid DNA may be checked on an agarose gel by performing a drop out of the insert previously added, using Hind III and Sal I.

NICKING 20 units of Nt.BspQI is added to the plasmid in the presence of 1 X NEB 3.1 buffer. This is incubated at 50 °C for one hour and followed by heat-killing at 80 °C for 20 mins.

NUCLEASE ACTION 1 µl Exonuclease III is added to the nicked plasmid and incubated at 37 °C for one hour and heat-killed at 70 °C for 20 mins.

PHENOL-CHLOROFORM EXTRACTION Phenol lyses cells with aqueous solution appearing at the top and organic solution appearing at the bottom. Chloroform prevents phenol retaining the aqueous phase. Components are separated by their different immicibilities in liquids that differ in density. Add 80 µl phenol to nicked+exonucleased plasmid. Vortex and centrifuge for 1 min. Place aqueous phase in clean eppendorf. Add 40 µl phenol and 40 µl chloroform and vortex and spin for 1 min. Take aqueous phase and add to another clean eppendorf. Add 80 µl chloroform, vortex, spin for 1 min. Keep aqueous phase.

A.1.2 *From Chapter 5: Assembling clickable tetrahedra*

For the click reaction to proceed it requires Magnesium ascorbate (as a reducing agent), a copper catalyst and a ligand. There are 2 click sites per tetrahedron.

MAKING THE CATALYST SOLUTION 1.2 µl 100 mM magnesium ascorbate solution (at 10 Cu equivalent per click site) is added to the 0.036 mg ligand (7 Cu equivalent per click site), the resulting solution was degassed by bubbling through nitrogen/argon gas. Then 0.12 µl 100 mM $CuSO_4$ stock solution (1 Cu equivalent per click site) was added to this, where Cu was at 12 nmol. The resulting solution was degassed again.

DNA SOLUTION 12 nmol of clickable tetrahedra are assembled from strands of 30 µM concentration in 100 µl in 1 x TE and 15 mM $MgCl_2$ buffer. The tetrahedron were annealed by heating to 95 ° and cooling to room temperature over roughly 15 minutes. The DNA solution was degassed.

COMBINING The catalyst solution was added to DNA solution. For one click point 20 equivalent of Cu was used.

The reaction mixture was left at room temperature for two to three hours.

Then the reaction mixture was passed through NAP size exclusion column. The resulting structures were purified by denaturing gel.

DENATURING GEL The gel solution is 21 g urea, 5 mL 10 x TBE (Tris/Borate/EDTA) and 12.5 mL 40% 19:1 Acrylamide which are mixed together and topped up to 50 mL with Milli-Q water. To this 5 mL click solution (as above) 50 µl APS (ammonium persulphate) and 5 µl TEMED (Tetramethylethylenediamine).

The gel is cast, set and pre-run at 20 W for 40 mins in TBE buffer. The wells are washed out with TBE before being loaded with the samples containing formamide, and loading buffer with urea:DNA, 1:1.

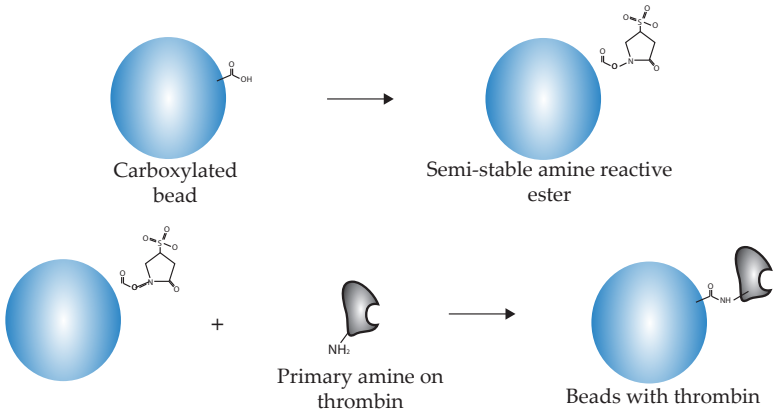
A.1.3 *From Chapter 5: Assembling paranemic structures*

This project originally set out to use structured libraries for the selection of new co-operative aptamers, and for thrombin binding in particular. Here we shall present the initial experimental steps taken towards finding a multivalent thrombin binder using a DNA scaffold. More specifically we set out to evolve a random 15-mer that, in conjunction with the TBA (for the fibrinogen exosite), produces a complex that exhibits an enhanced binding.

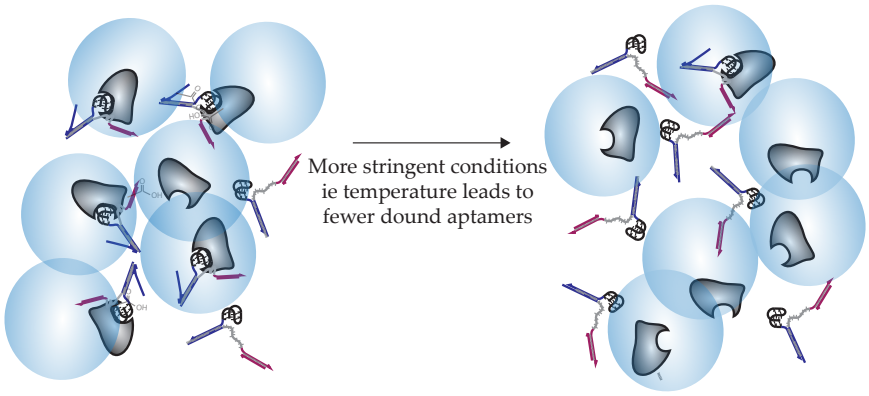
In order to establish binding conditions, we began by assembling a simple structure: it comprised a duplex of aptamers and operates as described in figure 55. We call the strand including the random 15-mer R and the strand with the G-quadruplex, G.

The first step in the experiment is to conjugate thrombin to magnetic beads which are functionalised with a carboxylate group. EDC is a zero-length carbodiimde cross-linker that conjugates primary-amines directly to carboxyl groups. The presence of sulfo-NHS during conjugation helps to form a more stable amine-reactive ester. Following conjugation, the next step is to perform a binding experiment by incubating the assembled aptamer duplex with thrombin-beads under progressively more stringent conditions, using pH or temperature. The supernatant containing the unbound structures can be pipetted off whilst using a magnet to keep the thrombin-bead complexes and binders stationary. PCR can then be performed amplifying the effective binders, either directly on the beads, or following a strand-displacement reaction that separates the R strand.

Step 1: Attach thrombin to magnetic beads



Step 2: Perform a binding experiment



Step 3: Amplify strong binders

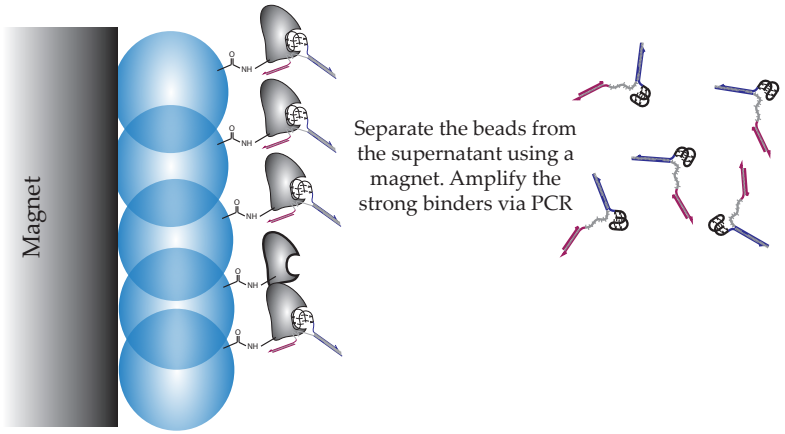


Figure 55: Method of conjugation, binding and selection for thrombin binding.

Initially the experiment was carried out in the absence of beads, to determine whether the aptamer-complex would bind to thrombin. The binding buffer was optimised and was tris-based with 5 mM KCl, 140 mM NaCl and 1 mM MgCl2. No G+R binding was found for thrombin concentrations below 0.4µm. Following optimisation, the same experiment was carried out but this time with thrombin conjugated to the magnetic beads. The experimental steps were followed on a gel. The aptamer complexes were incubated with the thrombin-beads and then the supernatant containing unbound complexes was removed (supernatant 1). A displacement experiment was carried out on the bound complexes using the complement of R (DS) to displace the random strand (supernatant 2). Finally the G-quadruplex strand was displaced from thrombin using its complement (Gcomp) (supernatant 3). Regarding the gel however, no complexes remain bound following incubation despite using the same conditions that gave rise to binding in the absence of beads. This suggests that the conjugation of thrombin to the beads was not successful and that the thrombin was lost during washing. Conjugation was repeated with varying conditions but still to no avail.

To overcome this problem, a biotinylated form of thrombin was next used that should bind very strongly and efficiently to streptavidinated magnetic beads. The biotin-streptavidin interaction is also very robust in a wide range of conditions, including extremes of pH and temperature. In this instance however, non-specific binding was prevalent despite increasing the stringency of reaction conditions. More work is required here to make this specific thrombin selection set-up feasible and is unfortunately beyond the scope of this project.

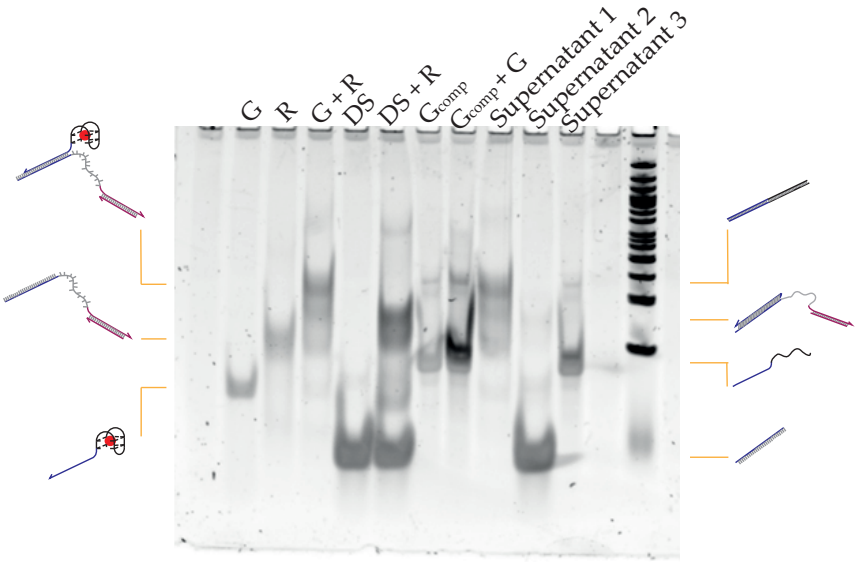


Figure 56: Thrombin aptamer binding gel. G is the G-quadruplex strand, R is the strand containing the random sequence for new aptamer selection, which has a complementary region to bind to G. G+R shows a shift band where the strands have bound. DS is the displacing strand, which displaces strand R from being bound to strand G initiated by a toe-hold. DS+R shows some R strands binding to DS in excess DS. G_{comp} complements the G-quadruplex strand in an attempt to be able to displace it from being bound to thrombin. G_{comp}+G shows that all G present is bound by this due to lack of G band. However the shift seen in the G_{comp} position is only slight. Supernatants 1, 2 and 3 show the different stages of strand addition to the thrombin molecule. In supernatant 1 G and R are bound. Upon addition of DS, supernatant 2 shows G is displaced and R binds to DS (a very faint band). Upon addition of the G_{comp} strand in step 3, supernatant 3 has the G_{comp}+G band, showing the G strand excess bound but not necessarily that it has been displaced from the thrombin molecule.

A.2 METHODS

A.2.1 *AFM*

The mica disk is attached to a glass slide with Vaseline and is then cleaved using Selotape until a complete disk is visible on the tape, obtaining an atomically flat surface. 5 µL of purified origami sample (between 10 and 20 nM) is deposited onto the fresh surface. 90 µL of imaging buffer is added to form a meniscus.

A cantilever is set in place on the fluid cell, the laser aligned onto the tip labelled 'B' and can then be used to approach the meniscus. Near to the meniscus, a further 40 µl of buffer is pipetted upwards towards the optical surface of the fluid cell so this fluid meets the meniscus and creates an hourglass shape. The cantilever may then approach the surface for imaging.

All AFM images were captured using a Veeco Dimension 3100 and Bruker SNL-10 cantilevers. The AFM was used in tapping mode. Nanoscope Scanning Probe Microscopy software is used for imaging and in analysis.

A.2.2 *TIRF microscopy*

All the microscopy experiments in Chapter 4 were performed using an Olympus IX-81 inverted microscope. An Andor iXon EMCCD camera recorded the images. An Olympus 100 x, 1.4 NA magnification oil-immersion objective was used along with a 1.6 x microscope internal magnification. Emission was passed through a MAG Biosystems Dual View (DV-2) device that splits the signal into two channels, which are filtered separately with a 565-20 nm and a 665-20 nm filter. The first 565 nm channel was used for the sample observations. Its signal was

projected onto one 512 px x 256 px half of the 512 px x 512 px camera chip. The 665 nm channel was used with the Cy3/Cy5 filter to observe the beads with the mercury lamp, to compare against sample locations. For imaging Cy3 fluorescence, the samples were illuminated with a OZ Optics 50 mW 532 nm laser fibre-coupled (single-mode, manufacturer-stated 20 mW at the fiber end) to an Olympus module for objective-based TIRF. For imaging the YOYO-1 Iodide fluorescence, the samples were illuminated with a mercury lamp (Prior Scientific, Lumen 200). Andor iQ and μManager software were used as imaging and microscope control software. The flow channels were prepared according to Lucas, 2016 [146].

A.2.2.1 *YOYO-1 Iodide staining*

YOYO-1 Iodide (1 mM stock in DMSO; Molecular Probes, Invitrogen) is a 1.27 kilo-dalton molecule. It has a 3,200-fold increase in fluorescence when bound to ds-DNA compared to ss-DNA [193]. A 1/10 dilution of the YOYO-1 was added to the DNA tube samples in a 1:10 volume ratio. The samples were purified 10-15 mins later with S300 size-exclusion resin.

A.2.2.2 *Anti-photobleaching buffer*

The microscopy buffer was 1x TAE, 12.5mM $MgCl_2$. The anti-photobleaching buffer was adapted by a colleague Samuel Tusk from Selvin and Ha 2008 and Cooper et al. 2013 [194, 195]. The buffer was made immediately prior to use. The recipe for it is as follows: 276 μl glucose buffer (0.8% w/v D-glucose), 2.5 μl gloxy solution, 1.25 μl Trolox (100 mg/m in DMSO), 0.78 μl NBA (100 mg/ml in DMSO), 2.25 μl COT (2% in DMSO). The final DMSO concentration is 1.5%. Glucose buffer was made as: 1 x TAE, 12.4 mM $MgCl_2$, 0.8% glucose. 100 x gloxy solution was made according to: 20 μl of 0.2 mg/ml catalase, 80 μl mi-

croscopy buffer + 10 mg glucose oxidase. This was mixed by tapping, centrifuged at high speed for 1 min and the supernatant collected and stored at 4 °C. For an enzyme-containing anti-photobleaching buffer, the 276 µl glucose buffer was replaced by 273.17 µl glucose buffer and 2.83 µl KOAc or NaCl (5 M). 2.5 µl of Nt.BbvCI (10,000 units/ml) was then added to 47.5 µl of the new anti-photobleaching buffer.

A.2.3 *Microscopy image analysis*

In Chapter 4 Section 4.4.4.1, we use Mathematica to detect flashes in microscopy images as spots in still snapshots. The method is provided by Alex Lucas and adapted from Lucas 2016 [146].

The analysis is done in Wolfram Mathematica 10.2 using the Laplacian of the Gaussian for edge detection [70], normalising and then binarising the images. Spots are detected from the binarised images using the 'ComponentMeasurements' function. 'ComponentMeasurements' gives axes values for an ellipse fitted to each spot. This function is used to highlight the detected spots in images for comparison on visual inspection. The spot size is defined as the number of pixels the spot comprises. Very large or small spots are filtered out with respect to minimum and maximum sizes or widths, to avoid noise and aggregates.

A.2.3.1 *Laplacian of the Gaussian*

The change in intensity at an edge in a binarised image is characterised by a peak in the first derivative and hence a zero-crossing in the second derivative. This way shapes in images can be detected. This is implemented using the 'LaplacianGaussianFilter' function in Mathematica. An input image (I(x; y)) is first convolved with a Gaussian kernel (G(x; y; r)) to smooth out noise, according to:

$$G(x;y;r) = \frac{1}{2\pi r^2} e^{-\frac{x^2+y^2}{2r}} \tag{10}$$

The Laplacian of the Gaussian is then found according to:

$$f(x,y) = \nabla^2[I(x,y) * G(x,y;r)] \tag{11}$$

r is chosen depending on the size of the object that is being fitted: If it is too small it will not smooth out the noise, but if it is too big it will smooth out the objects to be detected.

A.2.3.2 *Flashes in the videos*

The median-subtracted images are created by applying the following equation to each pixel of the image, where p is the pixel value, p_{median} is the median pixel value for that pixel over the entire video, and p_{ms} is the new median subtracted pixel:

$$p_{ms} = \frac{p - p_{median}}{p_{median}} \tag{12}$$

Chapter 4 Section 4.4.4.1 requires the detection of the bright regions in the median subtracted images that correspond to flashes in the videos. The median subtracted images for all 3 samples are normalised together with respect to the 0.001^{th} and 0.999^{th} quantile values. The Laplacian of the Gaussian filter is used with $r = 3$. The output images are then normalised between the minimum and maximum value of all the images. The images are then binarised with a set cut-off of 0.42. Only spots with sizes between 5 and 70 pixels are considered.

A.2.4 *Ensemble fluorimetric measurements*

In Chapter 4, purified origami samples were made up to 150 µl with a final buffer of either; a) replicate CutSmart: 10 mM MgOAc, 20 mM Tris-acetate, 50 mM KOAc; b) NEB buffer 2.1: 10 mM $MgCl_2$, 1 x Tris-EDTA, 50 mM NaCl.

The method was adapted from Alex Lucas [146]. The samples were placed in 140 µl quartz cuvettes (Starna Scientific). The cuvettes were washed thoroughly with Milli-Q water and ethanol using a vacuum pump before use. The fluorescence was measured in a Cary Eclipse Fluorescence Spectrometer (Varian). The temperature of the samples was held at 37 °C by a water bath circulating through the fluorimeter. Samples were mixed with a pipette after 5 mins of acclimatising to temperature. 400 µl of mineral oil (Sigma-Aldrich) was added to stop sample evaporation (the effects of which can otherwise be seen after about 1-1.5 hours), whilst allowing other solutions to be added to the sample during the experiment. The fluorimeter emission/excitation settings for the fluorophores used were: Cy5 648 nm/668 nm, Cy3 550 nm/564 nm, TET 522 nm/539 nm, TEX615 596 nm/613 nm. The slit widths were 5 nm and 10 nm for the excitation and emission respectively.

A.2.5 *Circular dichroism of helical AuNP arrangements*

The 10 nm AuNPs were already conjugated to single-stranded DNA (19 poly-Ts) by post doc Robert Schreiber who was working with them for another project: he kindly allowed me to use some of his. A complete description of the conjugation process is available in the Supplementary Information of the paper by Kuzyk *et al.* [54]

Before adding the AuNPs to the origami, the unbound thiolated DNA strands must be removed. This is achieved by spinning for 5 min at 14000 rcf through a 100 kDa MWCO centrifugal filter (Amicon Ultra, Millipore) followed by additional four spinning steps (5 min, 14,000 rcf) with 400 µl 1 x TAE buffer added to the centrifugal filter before each step.

The DNA origami tubes are purified from excess staples using agarose gel electrophoresis. The correct band is cut-out and then squeezed between two parafilm-covered glass slides, relinquishing the liquid containing the nanostructures. The gel is 1.5 % agarose gel in 0.5 x TBE buffer with 11 mM $MgCl_2$.

25 µL 300 nM 10 nm AuNPs are added to 25 µl purified origami tubes and 150 µl 1 X origami buffer and left on a shaker at room temperature for 24 hours. Gel purification is used again to separate unbound AuNPs from the AuNPs+origami.

A stronger CD signal is achieved following silver-enhancement (Ag-enhancement) of the AuNPs. 15 µl of Ag-enhancement mix (A+B+C of HQ Silver Enhancement, Nanoprobes, USA) is added to the AuNPs+origami.

A.2.6 *TEM*

Transmission electron microscopy (TEM) can obtain both high-resolution images and diffraction information from a single sample and the magnification can be as high as 10^6. High energy electrons (the accelerating voltage ranges from 125kV to 3MV) are electromagnetically focused and transmitted through the sample. The type of interaction the electrons have with that sample determine the information obtained. The electrons that undergo elastic scattering give rise to a diffraction pattern. However, when primary electrons are scattered inelastically from the sample electrons, at regions of varying density

(or when in solid state, grain boundaries, defects and dislocations), absorption and scattering effects lead to spatial variation in the intensity of the transmitted electrons. When the sample and TEM are set up for Bragg diffraction, the electron beam incident from the top is diffracted from sets of vertical planes in the sample and consequently two beams emerge from the bottom of the sample. If an aperture is placed around the direct beam the image formed is called a bright-field image. Similarly, if the aperture intercepts the diffracted beam, the image formed is called a dark-field image. These images show diffraction contrast. The image sin this thesis were taken using bright-field TEM.

A.3 MATERIALS

A.3.1 *Oligo synthesis and storage*

Most oligonucleotides were purchased from Integrated DNA Technologies at a synthesis scale of 25-100 nmol concentration. Most strands were lypohilised and were resuspended to 100 μM concentration in 1 x TE or milli-Q water. Fluorescently labelled strands were additionally HPLC purified. Origami well plates came with strands pre-suspended in TE buffer at 100 μM concentration.

DNA strands that had been adapted for click-chemistry were kindly provided by Afaf El-Sagheer from ATD Bio and were used at working concentrations of either 3 or 1 μM.

Working concentrations for individual tubes of 1 μM were prepared for all other strands and working stocks of lower concentration for staple strands were also prepared. All working stocks were kept in the fridge, apart from click strands, which were kept frozen. Well-plates were also kept frozen.

Table 1: Oligonucleotides for nanotubes

NAME	SEQUENCE
Left end polyT 1	TTTTTTTTAGCTTTCATCAATGAGACGGGCAACTTTTTTTT
Left end polyT 2	TTTTTTTTCGAACGAGTAGACCTTCCTGTAGCCTTTTTTTT
Left end polyT 3	TTTTTTTTTTATGCGATTTTTTCCCAATTCTGTTTTTTTT
Left end polyT 4	TTTTTTTTAATTGTATCGGTATTGTGAATTACCTTTTTTTT
Left end polyT 5	TTTTTTTTGATATTCACAAAAAAGGAGCCTTTTTTTTTTT
Left end polyT 6	TTTTTTTTCTTTTTAAGAAAGACGATTGGCCTTTTTTTTTT
Left end polyT 7	TTTTTTTTAGTAATTCTGTCTCTTACCGAAGCCTTTTTTTT
Left end polyT 8	TTTTTTTTATAACGGATTCCGACAAAAGGTAATTTTTTTT
Left end polyT 9	TTTTTTTTCTTCTGACCTGAATCGGGAGAAACATTTTTTTT
Left end polyT 10	TTTTTTTTAGCTGATTGCCCCAGAGATAGAACCTTTTTTTT
Right end polyT 1	TTTTTTTTCACGACGTTGAGCTAAACAGGTTTTTTTT
Right end polyT 2	TTTTTTTCGGAGACAGTCAGGTTTTCCCAGTTTTTTTT
Right end polyT 3	TTTTTTTGCTTTAAACAGTGTGAGAAAGGCTTTTTTTT
Right end polyT 4	TTTTTTTTACCAAGCGCGAACCCCCTCAAATTTTTTTTT
Right end polyT 5	TTTTTTTTACCGCCACCCTCCCAGCGATTATTTTTTTTT
Right end polyT 6	TTTTTTTTAGCCAGCAAAATGAGGTTTAGTTTTTTTTT
Right end polyT 7	TTTTTTTTGATTAGTTGCTATTGGGAATTAGTTTTTTTT
Right end polyT 8	TTTTTTTTGTTATATAACTACCTTAAATCAATTTTTTTT
Right end polyT 9	TTTTTTTTATAATACATTTGGCTTAGGTTGGTTTTTTTT
Right end polyT 10	TTTTTTTTAGGCCGATTAAGCCGTCAATAGTTTTTTTT
Polymerising 1	TGAGACGGGCAACTCACGACGTTG
Polymerising 2	GGTTTTCCCAGAGCTTTCATCAA
Polymerising 3	CCTTCCTGTAGCCCGGAGACAGTCA
Polymerising 4	GTGAGAAAGGCCGAACGAGTAGA
Polymerising 5	TTCCCAATTCTGGCTTTAAACAGT
Polymerising 6	CCCCCTCAAATTTATGCGATTTT
Polymerising 7	ATTGTGAATTACCACCAAGCGCGAA
Polymerising 8	CCAGCGATTATAATTGTATCGGT
Polymerising 9	AAAGGAGCCTTTACCGCCACCCTC
Polymerising 10	GAGGTTTAGGATATTCACAAA
Polymerising 11	GACGATTGGCCTTAGCCAGCAAAAT
Polymerising 12	TTGGGAATTAGCTTTTTAAGAAA
Polymerising 13	TCTTACCGAAGCCGATTAGTTGCTA
Polymerising 14	CCTTAAATCAAAGTAATTCTGTC
Polymerising 15	CGACAAAAGGTAAGTTATATAACTA
Polymerising 16	GCTTAGGTTGGATAACGGATTC
Polymerising 17	ATCGGGAGAAACAATAATACATTTG
Polymerising 18	GCCGTCAATAGCTTCTGACCTGA
Polymerising 19	CAGAGATAGAACCAGGCCGATTAA
Polymerising 20	AGCTAAACAGGAGCTGATTGCCC
Non-addressable Invertase	CCTCCCTTTAACGGGGTCAGTGTTTCTGTA
Non-addressable GOx	AATCACCGCGTATAAACAGTTATTGTCGTC

Non-addressable HRP	TTATTAGCCGGAACCTATTATTCGCCCTCAT
Addressable Invertase 18	CCTCCCTTTAACGGGGTCAGTGTTTCTGTATTTGTTGAGGATGAGGAGTGGAG
Addressable GOx	AATCACCGCGTATAAACAGTTATTGTCGTCTTTCTTCACTCTCACTTTAACAC
Addressable HRP	TTATTAGCCGGAACCTATTATTCGCCCTCATTTTCTCTCCATCCTACTCCTCAC
Non-addr. Inv neighbour	TGGGATTTTTGCAGGGAGTTAACCGGATAT
Non-addr. GOx neighbour	TTTCCAGATCGTCACCCTCAGCATGACCTTC
Non-addr. HRP neighbour	AGTTAGCGAACGAGGGTAGCAAGAACGGTG
Non-stator 1	TGCTTTCGCAAAAAAAAGGCTCCACAAATAAA
Non-stator 2	AAAATCTAAACGAGTAGTAAATTGAACAGCTTG
Non-stator 3	TCTGAATTCACCACCAGAGCCGCCGGAAACCG
Non-stator 4	ACCAGAACTACCGTTCCAGTAACAACTAAA
Non-stator 5	GAACTGGCAATATCAGAGAGATAATAAGTCCT
Non-stator 6	TGAGCGCTATGATTAAGACTCCAGAACCGC
Non-stator 7	CGAGCATGAGCCAACGCTCAACAGAAGAAAACA
Non-stator 8	CAGTATAATAGAAACCAATCAAAGAATTAA
Non-stator 9	TACCTTTTTTGTTTGGATTATACTAGGTGAGG
Non-stator 10	AATCCTGATTAATGGAAACAGTAAGCCTGT
Non-stator 11	ACGCTGAGCAAACTATCGGCCTTGAGAAAGGA
Non-stator 12	GAAGAACTAGCCAGCAGCAAATCATATTCC
Non-stator 13	CGCTGGCATTCCACACAACATACGGTGCCGGA
Non-stator 14	GCTCACAAAGTGTAGCGGTCACATACTTCT
Non-stator 15	TGCGCAACATTGCCTGAGAGTCTTAATACTTT
Non-stator 16	ATCAGGTCTGTTGGGAAGGGCGAATCATGGT
Non-stator 17	TAAAAATTGCAAAGCGGATTGCATTGCCAGAG
Non-stator 18	AGTCAGAATTTAGAACCCTCATACCGGAGAG
Non-stator 19	CGTCCAATGATAAATTGTGTCGATAAAACGAA
Non-stator 20	CATCGCCTACTGCGGAATCGTCAACCATAAA
Non-stator 21	TTTGACCCAGAACCGCCACCCTCTGTATCACC
Non-stator neighbour 1	CTCATTATATTTCAACTTTAATCTTATCAGCT
Non-stator neighbour 2	TGACAGGAAAGCCAGAATGGAATCACGTTG
Non-stator neighbour 3	ATACCGATCGAATAATAATTTTTAGCGCAGTC
Non-stator neighbour 4	AATTGAGCAATAATAACGGAATGAGCCGCC
Non-stator neighbour 5	ATGGCTTTGAGCCACCACCCTCAACCCAAAA
Non-stator neighbour 6	TAATTGAAAAATAATATCCCATAGGGTAAT
Non-stator neighbour 7	AGTATGTTCCTGAACAAAGTCAGCCTAATTTA
Non-stator neighbour 8	ATGGAAGTACATTTAACAATTTATTCTTAC
Non-stator neighbour 9	GTCTTTCCATATGCGTTATACAACATTTGAAT
Non-stator neighbour 10	ATCCAGAATTAACACCGCCTGCATCAATAT
Non-stator neighbour 11	AATATATGAGATGATGGCAATTCAACAGTGCC
Non-stator neighbour 12	AGCATAAAAAGCGAAAGGAGCGGCCTGAGTA
Non-stator neighbour 13	AAAGCATCTAATAACATCACTTGGCGCTAGGG
Non-stator neighbour 14	CAAGAGAAAAAGCGCCATTCGCTGTTATCC
Non-stator neighbour 15	AACCACCAGTTTCCTGTGTGAAATCATTCAGGC
Non-stator neighbour 16	ATTAAGAGAAGCCTTTATTTCACAAAGGCT

Non-stator neighbour 17	GGCCTCTTATTTTTGAGAGATCTAACGCAAGGA
Non-stator neighbour 18	ACCTGCTCAGTAAAATGTTTAGACTATTAT
Non-stator neighbour 19	ATGCAATGTCAGGTCTTTACCCTGACTGGATAG
Non-stator neighbour 20	TCTGAATTCACCACCAGAGCCGCCGGAAACCG
Non-stator neighbour 21	CATTGAATACAAAGTACAACGGAGACACTCATC
Non-biotin 1	ATGGTTTAACCAGTCAGGACGTGTCTGGAA
Non-biotin 2	ACACCAGCGTTAATAAAACGAAAACATGTT
Non-biotin 3	CATTCAGTTTACAGGTAGAAAGAGCTTAAT
Non-biotin 4	TCATTACCTAGGAATACCACATTTGATAAGA
Non-biotin 5	ATCAAGAATAACGCCAAAAGGACAGGATTA
Non-biotin 6	AACCAGGCTCGATGAACGGTAATCAATCGGTT
Non-biotin 10	AACGCCAGAATCACCATCAATATGGTAAAGAT
Non-biotin neighbour 1	GTTTCATGATACATTTCGCAAAAGGAACGC
Non-biotin neighbour 2	TTAAATATAGCTATATTTTCATTTTTGTTA
Non-biotin neighbour 3	TGCTGAATAAGGTGGCATCAATTCGTTAATA
Non-biotin neighbour 4	GGTCATTCATTAACATCCAATAAAGATTGT
Non-biotin neighbour 5	GAGAGTACAAAGAATTAGCAAAGGTTGATA
Non-biotin neighbour 6	GTACCAAAAATTCGAGCTTCAAAGTAACCCTCG
Non-biotin neighbour 10	TCAAAAGGTCAGAAAACGAGAATGTAAATATT
Stator 1	TTTACACGATCTACATTGGAACCTCAGCCCAACTAACATTCACTCCACAATGCTTTCGCAAAAAAAGGCTCCACAAATAAA
Stator 1 i	TTTACACGATCTACATTGGAACCTCAGCCCAACTAACATCGTCACTCCACAATGCTTTCGCAAAAAAAAGGCTCCACAAATAAA
Stator 2	TTTACACGATCTACATTGGAACCTCAGCCCAACTAACATTCACTCCACAAAAAATCTAAACGAGTAGTAAATTGAACAGCTTG
Stator 3	TTTACACGATCTACATTGGAACCTCAGCCCAACTAACATTCACTCCACAATCTGAATTCACCACCAGAGCCGCCGGAAACCG
Stator 4	TTTACACGATCTACATTGGAACCTCAGCCCAACTAACATTCACTCCACAAACCAGAACTACCGTTCCAGTAACAACTAAA
Stator 5	TTTACACGATCTACATTGGAACCTCAGCCCAACTAACATTCACTCCACAAGAACTGGCAATATCAGAGAGATAATAAGTCCT
Stator 6	TTTACACGATCTACATTGGAACCTCAGCCCAACTAACATTCACTCCACAATGAGCGCTATGATTAAGACTCCAGAACCGC
Stator 7	TTTACACGATCTACATTGGAACCTCAGCCCAACTAACATTCACTCCACAACGAGCATGAGCCAACGCTCAACAGAAGAAAACA
Stator 8	TTTACACGATCTACATTGGAACCTCAGCCCAACTAACATTCACTCCACAACAGTATAATAGAAACCAATCAAAGAATTAA
Stator 9	TTTACACGATCTACATTGGAACCTCAGCCCAACTAACATTCACTCCACAATACCTTTTTTGTTTGGATTATACTAGGTGAGG
Stator 10	TTTACACGATCTACATTGGAACCTCAGCCCAACTAACATTCACTCCACAAAATCCTGATTAATGGAAACAGTAAGCCTGT
Stator 11	TTTACACGATCTACATTGGAACCTCAGCCCAACTAACATTCACTCCACAAACGCTGAGCAAACTATCGGCCTTGAGAAAGGA
Stator 12	TTTACACGATCTACATTGGAACCTCAGCCCAACTAACATTCACTCCACAAGAAGAACTAGCCAGCAGCAAATCATATTCC
Stator 13	TTTACACGATCTACATTGGAACCTCAGCCCAACTAACATTCACTCCACAACGCTGGCATTCCACACAACATACGGTGCCGGA

Stator 14	TTTACACGATCTACATTGGAACCTCAGCCCAACTAACATTCACTCCACAAGCTCACAAA GTGTAGCGGTCACATACTTCT
Stator 15	TTTACACGATCTACATTGGAACCTCAGCCCAACTAACATTCACTCCACAATGCGCAACA TTGCCTGAGAGTCTTAATACTTT
Stator 16	TTTACACGATCTACATTGGAACCTCAGCCCAACTAACATTCACTCCACAAATCAGGTCT GTTGGGAAGGGCGAATCATGGT
Stator 17	TTTACACGATCTACATTGGAACCTCAGCCCAACTAACATTCACTCCACAATAAAAATTG CAAAGCGGATTGCATTGCCAGAG
Stator 18	TTTACACGATCTACATTGGAACCTCAGCCCAACTAACATTCACTCCACAAAGTCAGAAT TTAGAACCCTCATACCGGAGAG
Stator 19	TTTACACGATCTACATTGGAACCTCAGCCCAACTAACATTCACTCCACAACGTCCAATG ATAAATTGTGTCGATAAAACGAA
Stator 20	TTTACACGATCTACATTGGAACCTCAGCCCAACTAACATTCACTCCACAACATCGCCTA CTGCGGAATCGTCAACCATAAA
Stator 21	TTTACACGATCTACATTGGAACCTCAGCCCAACTAACATTCACTCCACAATTTGACCCA GAACCGCCACCCTCTGTATCACC
Stator brace 1	CTCATTATATTTCAACTTTAATCTTATCAGCT
Stator brace 2	TGACAGGAAAGCCAGAATGGAATCACGTTG
Stator brace 3	ATACCGATCGAATAATAATTTTTAGCGCAGTC
Stator brace 4	AATTGAGCAATAATAACGGAATGAGCCGCC
Stator brace 5	ATGGCTTTGAGCCACCACCCTCAACCCAAAA
Stator brace 6	TAATTGAAAAATAATATCCCATAGGGTAAT
Stator brace 7	AGTATGTTCCTGAACAAAGTCAGCCTAATTTA
Stator brace 8	ATGGAAGTACATTTAACAATTTATTCTTAC
Stator brace 9	GTCTTTCCATATGCGTTATACAACATTTGAAT
Stator brace 10	ATCCAGAATTAACACCGCCTGCATCAATAT
Stator brace 11	AATATATGAGATGATGGCAATTCAACAGTGCC
Stator brace 12	AGCATAAAAAGCGAAAGGAGCGGCCTGAGTA
Stator brace 13	AAAGCATCTAATAACATCACTTGGCGCTAGGG
Stator brace 14	CAAGAGAAAAAGCGCCATTCGCTGTTATCC
Stator brace 15	AACCACCAGTTTCCTGTGTGAAATCATTCAGGC
Stator brace 16	ATTAAGAGAAGCCTTTATTTCACAAAGGCT
Stator brace 17	GGCCTCTTATTTTTGAGAGATCTAACGCAAGGA
Stator brace 18	ACCTGCTCAGTAAAATGTTTAGACTATTAT
Stator brace 19	ATGCAATGTCAGGTCTTTACCCTGACTGGATAG
Stator brace 20	CACCCTCAAAGAATACACTAAAATTTGTAT
Stator brace 21	CATTGAATACAAAGTACAACGGAGACACTCATC
Biotin staple 1	ATGGTTTAACCAGTCAGGACGTGTCTGGAAGTTCTCCTCTCCACCACCT
Biotin staple 2	ACACCAGCGTTAATAAAACGAAAACATGTTTTAACTCCTCTCCACCACCT
Biotin staple 3	CATTCAGTTTACAGGTAGAAAGAGCTTAATTGCTCTCCTCTCCACCACCT
Biotin staple 4	TCATTACCTAGGAATACCACATTTGATAAGAGGTCTCCTCTCCACCACCT
Biotin staple 5	ATCAAGAATAACGCCAAAAGGACAGGATTAGAGACTCCTCTCCACCACCT
Biotin staple 6	AACCAGGCTCGATGAACGGTAATCAATCGGTTGTACCTCCTCTCCACCACCT
Biotin staple 10	AACGCCAGAATCACCATCAATATGGTAAAGATTCAACTCCTCTCCACCACCT
Plain motor	CGATGTTAGTTGGGCTGAGGTTCC
Double-labelled Cy3	/5Cy3/TTTCATCACACTCTT/3Cy3Sp/

Cy3 motor both ends	GAGTGTGATGATTGCATGTTAGTTGGGCTGAGGTTCCTTGAGTGTGATGA
Cy3 motor one end	GCATGTTAGTTGGGCTGAGGTTCCTTGAGTGTGATGATTTGAGTGTGATGA
Motor 3′ quencher	CGATGTTAGTTGGGCTGAGGTTCCTTT/3IAbRQSp/
TET strand	/5TET/CGAAGTGTAGGGAGTGGGAT
TEX615 strand	/5TEX615/GAGGGTTGAGTGAGATAGAG
Cy5 strand	TGGGTAGTGGAAGTAAAAGG/3Cy5Sp/
Cy3 stator top comp.	/5CY3/TATGTAGATCGTGTAAA
Stator 1i for TET	ATCCCACTCCCTACACTTCGTGGAACCTCAGCCCAACTAACATCGTCACTCCACAATGC TTTCGCAAAAAAAAGGCTCCACAAATAAA
Stator 5 for Cy5	CCTTTTACTTCCACTACCCATGGAACTTCAGCCCAACTAACATTCACTCCACAAGAACTG GCAATATCAGAGAGATAATAAGTCCT
Inside dimer norm ext1	GGGGATGTAGCTGATAAATTAATGTATTTTAA
Outside dimer norm ext1	TGCGGGAGGAAGCCCGAAAGACTTAAAATAGC
Core staple 1	AAATTAATGGTTAGAACCTACCACCGAACGAA
Core staple 2	CTAAAATATCCTGAGAAGTGTTTCTTTGACGA
Core staple 3	TCCTCATTAGGTTGAGGCAGGTCAAGTAAGCAG
Core staple 4	TCAAAATCTTCGACAACTCGTATTGAAGGTTAT
Core staple 5	AACAATGACAAAGTTACCAGAAGCCAGCAT
Core staple 6	CGTGCCAGGCCGTAAAGCACTAACCTACATT
Core staple 7	TTTTGTTGTTGGTGTAGATGGGAAACCTGT
Core staple 8	GTCCACTATTAAAGAATTCCAGTTTGGAACAA
Core staple 9	ACAGTATCAGCCCCAAAAACAGGAATCATACA
Core staple 10	TCAACAGTCACCGAGTAAAAGAGTATGCGCCG
Core staple 11	CATCAAAATAACAACCCGTCGGTGGGCGCC
Core staple 12	ACAGGAAACAGAAGATAAAACAGTCTGAATA
Core staple 13	GGTAGCTCGCTATTACGCCAGCGAGGATCC
Core staple 14	ACAGTGCCGAACCAGAGCCACCACTACATACA
Core staple 15	CACGCAAATCAATATCTGGTCAAAGTTTGA
Core staple 16	AGATTAGAAGGGATTTTAGACAGGTTAGAATC
Core staple 17	GGTTCCGAAATCGGCAGCGGTCCACGCTGGTT
Core staple 18	AACAGGGGCAACATATAAAAGAATAATCAA
Core staple 19	AGGGAAGAGTGTAAAGCCTGGGGTGCACTCCA
Core staple 20	AAGAAGATACGTAAAACAGAAATTGATAGCCC
Core staple 21	TCACTGCCGAGCCCCCGATTTACCATTGCA
Core staple 22	TCAAATAATCCGGTATTCTAAGAGCGTCTT
Core staple 23	TCCAGAGCGGTAAATATTGACGAAACCATC
Core staple 24	TTGAGATTCAAATCAACGTAACATATTCGGTC
Core staple 25	AACGCCTGAGGACTAAAGACTTCCGAACTG
Core staple 26	GCGTTTTAACATTATGACCCTGGGAGCAAA
Core staple 27	GGGGTAATCATGTTACTTAGCCGGAATACGTA
Core staple 28	GTGAACCAGAGGCGGTTTGCGTATATTCTCCGT
Core staple 29	AAATCCCTTATAAATCTCTATCATTGCAGCAA
Core staple 30	ACCGCGCCTTAATGGTTTGAAATTCCCTTAGA
Core staple 31	ATATAAGTGGAAACGTCACCAATGGAAATTATT

Core staple 32	CGTGGACTCCAACGTCATAGGGTTGAGTGTTG
Core staple 33	ACCAGAAGCTATTAATTAATTTACCGACCG
Core staple 34	CATTAAAGTTACCAACGCTAACGAACGCGAGG
Core staple 35	CAACATTAGAATAAGGCTTGCCCTGACAACAA
Core staple 36	TGATTATCTGAGTGAATAACCTTGAATAAAC
Core staple 37	TTATTTGCGATGAAACAAACATCTAGGGCT
Core staple 38	GGAATTGAGTTGCGCCGACAATGACGAGAA
Core staple 39	CGGTCAGTACAATATTACCGCCAGGAGCTTGAC
Core staple 40	AGACGCTGCGTTATTAATTTTAAGTTGGCAAA
Core staple 41	AATCGCAACGAACCTCCCGACTAATTTTAT
Core staple 42	AATAAAAGGCACAGACAATATTTCAGATGAA
Core staple 43	GCTGAGGCTGCTAAACAACTTTCAGTACTGGT
Core staple 44	AGGGTGGTAGAGAGGGGCGATGCGACCAGT
Core staple 45	GGCAAGGCCTTTAATTGCTCCTTTCAACTAAT
Core staple 46	GCAGATACGTAATCTTGACAAGAAAGGCCGCTT
Core staple 47	ATCAGAAAGGCCTCAGGAAGATCGCCTAATG
Core staple 48	TGAGCGAGATAATTCGCGTCTGGTTTAGTTT
Core staple 49	AACACTGACCATTAAACGGGTAAAACGAGGC
Core staple 50	CACCCTCAAAGAATACACTAAAATTTGTAT
Core staple 51	AAACCAAGTCATAATTACTAGAAAACATAAATC
Core staple 52	TGGCCCTGTTTTCTTTTCACCAGCATTAAATG
Core staple 53	GTACTCAGCACCAGTAGCACCATTCACCGAC
Core staple 54	GCTTTGAATAACAGTACCTTTTACAAGCGTAAG
Core staple 55	ATCCTTGAATCATTTTGCGGAACAAAATATCA
Core staple 56	TAAAACATCAATCGTCTGAAATGGTTTTTGGG
Core staple 57	GCCAGCTTCAATCATATGTACCCCATTAAGCAA
Core staple 58	GAAGGCTTTATTTTAGTTAATTTCCTTAGATTA
Core staple 59	GAGAGGCTGTCAATCATAAGGGAATTTCATGAG
Core staple 60	AGAGCGGGTAAAACGACGGCCAGTAAGTTGGGT
Core staple 61	CCTGAATCGTGAATTATCACCGTACCATTAG
Core staple 62	TTTTGAAGTATGTAAATGCTGATTACCTTTTT
Core staple 63	AATCAGCTACGGCGGATTGACCGGCCAACGC
Core staple 64	TTAGTATCTTATCATTCCAAGAAAACATAAA
Core staple 65	CATAGCTCACCCGCCGCGCTTACTGTCCAT
Core staple 66	ACCGGAATACCGCACTCATCGAAAAATGAA
Core staple 67	AGTAGTAGTTTGCGGATGGCTTAGATTCATCAG
Core staple 68	GATAGCAGGATAAGTGCCGTCGATTTCAGGG
Core staple 69	AATAGCAGGAATAAGTTTATTTTAGCCCCC
Core staple 70	TGTGATAATTATTTTCATCGTACAAATAAG
Core staple 71	AGTGAGGCTGAAAGGAATTGAGAAATCCTT
Core staple 72	GCCTATTTGTTTGCCATCTTTTCAACGCAAAG
Core staple 73	CACCCTCATGATGATACAGGAGTACAGTTTC
Core staple 74	AGAGGCAAGAGCCACCACCCTCATGAGGGTTG
Core staple 75	ATAAGCAATCTGCCAGTTTGAGCGTTGCGC

Core staple 76	AGGCTTTGTAGCATTCCACAGACATGAAACAT
Core staple 77	CTACAGGGCGAGCTCGAATTCGTATCGGTGCG
Core staple 78	CAAGGCCATAGCCCGGAATAGGAGAACCGC
Core staple 79	TTGACGCTCGCCATTAAAAATATATCAAAA
Core staple 80	ACCTAAATCAATAGCAAGCAAATAAAATAAA
Core staple 81	AATCCTGTTTGATGGTTGCCCCAGCAGGCGAA
Core staple 82	CGTTTTAGGACAAAGAACGCGAGATGAATTTA
Core staple 83	TCAAAAACCTGAGTAATGTGTAGATATTCA
Core staple 84	AGGAAACGTTAAGCCCAATAATAAAATGCAGAA
Core staple 85	TAAAGCCTCAAACTCCAACAGGTATTACGAGG
Core staple 86	GAAGTTTGTTTCGTCACCAGTACCTCAAGAG
Core staple 87	AAAGGGCGAAAAACCGAAAAGAATAGCCCGAG
Core staple 88	TTGCGGGACGTTAGTAAATGAATCCTTGAGTA
Core staple 89	GTTTTCATAAGAGGCTGAGACTCAAACTAC
Core staple 90	AGCATGTTCCGGCACCGCTTCTGAGCCGGA
Core staple 91	ACCGTTCTGCTGCAAGGCGATTGCCAAGCT
Core staple 92	CCTAAAGGCGCTTTCCAGTCGGGCGCATCGTA
Core staple 93	AAGGATTAGCCTTTAGCGTCAGACCCAGCGCC
Core staple 94	TACAGACCGAGCAACACTATCACGAACCAG
Core staple 95	GTCGAGGTCTGCATTAATGAATCGTAATGGGA
Core staple 96	TAGGTCACAAAATTCGCATTAAATTTGGGGCGC
Core staple 97	GAGCTGAAATAATGCTGTAGCTCCTAACGGAA
Core staple 98	TCAAGTTTGGATTAGCGGGGTTTGTACCGT
Core staple 99	ACCGTGCAATATTTAAATTGTAAACTACTAAT
Core staple 100	AGTGAGCCCGGCGAACGTGGCGCTGGTAAT
Core staple 101	GAACAAGAGAATCGCCATATTTAACCTGAGCAA
Core staple 102	CATTGGCCTTTAATGCGCGAACAAAGAAAT
Core staple 103	CAATAGAATTTTGTTTAACGTCAGAACAAGCA
Core staple 104	CAAACAAATAGGTCTGAGAGACGCAAATCC
Core staple 105	AGCGGAGCACGCATAACCGATAAAGCTGCT
Core staple 106	TAAAGGTGAAGCGCATTAGACGGGTAATCGGCT
Core staple 107	ACCAGGCGCACCGTAATCAGTAGCCGATTGAG
Core staple 108	GGGAACAACATTTTTTAACCAATTGGTCAATA
Core staple 109	GCGGGGATCACCCAAATCAAGTATTATTTA
Core staple 110	TGAGCCATTTTTGCACCCAGCTACTGCGGGAGG
Core staple 111	GTAACATTAAACATAGCGATAGATCTTCTG
Core staple 112	AAAGACAAATTATTTATCCCAATCGGAATCATT
Core staple 113	GAAAGTATTCGGCATTTTCGGTCATGTCACAAT
Core staple 114	ACCAACTTACGACGATAAAAACCCAAATATC
Core staple 115	TTGATTAGACCTTGCTGAACCTCAAGAAACC
Core staple 116	TGCCCGAAAGAAGAGTCAATAGAAACTTTT
Core staple 117	GGCGAATTTTTTCAGGTTTAACGTTTGAATGG
Core staple 118	ATGCCACTGCCCAATAGGAACCCATTGCTCAGT
Core staple 119	AATACGTGGGACATTCTGGCCAATTCACCGCC

Core staple 120	TGCATGCCAACGTGCTTTCCTCGAACGGTA
Core staple 121	AACCTCCGAGGATTTAGAAGTATTACAACTAAT
Core staple 122	CAGTAATACAACATGTTCAGCTGAGCAAGA
Core staple 123	CTATTAGTAGATTCACCAGTCACAGCCCACTAC
Core staple 124	CAGCCATAAGGGCGACATTCAACGACAGAA
Core staple 125	CTGAACACAGCAAACGTAGAAAACGGAACCG
Core staple 126	GCAGACGTTTGCAAAAGAAGTTTCAAAAAG
Core staple 127	TGCGTAGAATTCATTTCAATTACAACGCCA
Core staple 128	TATACAGTACCAAGTTACAAAATTTCGAGC
Core staple 129	GGAGGGAACTAATTTGCCAGTTACCAGATATA
Core staple 130	TTTACCAGTGAAAGAGGACAGATCGGCTACAG
Core staple 131	AAACGATTAATTCATATGGTTTATGTAGCGC
Core staple 132	CCATCGCCTGAGAATAGAAAGGAAGCGTCATAC
Core staple 133	CCACCAGAACGCTCATGGAAATAATCGGAAC
Core staple 134	AAATCGTCGGAGCGGAATTATCATGAAAAATCT
Core staple 135	CCGGGTACCGCGTACTATGGTTGTTATAATC
Core staple 136	ATAGCAAACGAAGGCACCAACCAATCCGCG
Core staple 137	AGCCGTTTATAAGGCGTTAAATAAGCTTCTGT
Core staple 138	CAGCATCGGTAACGATCTAAAGTTATGCCCCCT
Core staple 139	GCACGTATTGCAGGTCGACTCTATGGCGAAAG
Core staple 140	ACCGGAAGCAGAGCATAAAGCTAGTAAAACT
Core staple 141	ACACCACGCCTTTACAGAGAGAATCGGGTATT
Core staple 142	ACAATAAAAGAGAATATAAAGTACGCCTGATT
Core staple 143	ACATGTAATTTATCAACAATAGACCCACAAG
Core staple 144	AATAAGTTCAGAGCCGCCACCCTCTTATTACGC
Core staple 145	CGCCAGAATCTTTAGGAGCACTAAGACTTTA
Core staple 146	AAAATCTCAGGTGAATTTCTTAGGCTTGAG
Core staple 147	GGGGAAAGTAACTCACATTAATTGGGGACGACG
Core staple 148	AACCCTCAATTAACCGTTGTAGCAGCTGCGCGT
Core staple 149	ACCTGTTTGCAACTAAAGTACGGTTGGGAAGA
Core staple 150	CGCGCCTGTTTAGGCAGAGGCATTCGCGCAGA
Core staple 151	GACCATTATCCATATAACAGTTGAAAGAACTGG
Core staple 152	CATAGTAAAGGCGCATAGGCTGGCGCGAAAGA

Table 2: Oligonucleotides for tetrahedra

NAME	SEQUENCE
Tetrahedron strand 1	TAAGAAACAGAACCCATACCAGAGAATAAAGAACTACCCTACAAGCCCGCCGAAACA AGAACA
Tetrahedron strand 2	GTAGGGTAGTAAGGGTAAGAAGTAAAGTATCATAAAGGGTCACAAATATCAAACGGCG GGCTT
Tetrahedron strand 3	TGACCCTTTAACCCGCCTCACCCTCCACCCTACTGTTTCTTATGTTCTTGTTATTGATATTT G
Tetrahedron strand 4	TTCTTACCCTACTTTATTCTC
Tetrahedron strand 5	TGGTATGGGTAAGGGTGGAGGATAATTTTGATATGGT

Tetrahedron strand 6	TAATAGCGACGATTTAGTGAGGCGGGAGATACTTTAC
FP09 01primer	TTTTTTGTCGACTAATAGCGACGATTTA
R08 03primer	AAAAAAGCATGCGCAATCATATCAAAAT
R08 02primer	CAACCCGCATGCATCATATCAAAAT
Paranemic red strand	GAGAACACGCAGCCCGAAACATCTCGCTTCGGCGAGCACCAGCGGGCAGCACGCGCGC GAGTAGAGCGCTTCGGCGCTGTGAGGGCGCGCGTGCAGCGTGTTCTCGTAGTCCCTCA
Paranemic green strand	ATTTGCGCACAGATGTCTACTCCCGTGTTCGCACGGCCTCACACATCACGGGCTGCCCGT GTGACGGCCTTCGGGCCGGGTTTAGGGCAGCCCGAGTGCGCAAATGAAACACGGGA
Paranemic blue strand	GTTTATTGCGACTAGATCACACGAGGCTTCGGCCTCTAAACCTCTAGACGTGCAGCGTA TGTTTAGCGCTTCGGCGCTCTGGTGACGCTGCACGACGCAATAAACTTGTGATGCCA

BIBLIOGRAPHY

[1] J D Watson and F H C Crick. Molecular structure of nucleic acids, 1953.

[2] A Rutherford. *Creation: The Origin of Life / The Future of Life*. Penguin UK, 2013.

[3] BBC History. Historic figures: Crick and Watson (1916-2004), 2014.

[4] Nobelprize.org. The Nobel Prize in Physiology or Medicine 1962, 2014.

[5] D Luo. The road from biology to materials. *Materials Today*, 6(11):38–43, 2003.

[6] B Alberts, A Johnson, J Lewis, M Raff, K Roberts, and P Walter. The Structure and Function of DNA. In *Molecular Biology of the Cell*. Garland Science, New York, 4th edition, 2002.

[7] H Lodish, A Berk, and S L Zipursky. Section 4.1 Structure of Nucleic Acids. In *Molecular Cell Biology*. W. H. Freeman, New York, 4 edition, 2000.

[8] ATDBio. Nucleic Acid Structure.

[9] S L Price and A J Stone. The electrostatic interactions in van der Waals complexes involving aromatic molecules. *Journal of Chemical Physics*, 86:2859, 1987.

[10] B Tinland, A Pluen, J Sturm, and G Weill. Persistence Length of Single-Stranded DNA. *Macromolecules*, 30(19):5763–5765, 1997.

[11] S B Smith, Y Cui, and C Bustamente. Overstretching B-DNA: The Elastic Response of Individual Double-Stranded and Single-Stranded DNA molecules. *Science*, 271:795–798, 1996.

[12] M Vologodskaia and A Vologodskii. Contribution of the intrinsic curvature to measured DNA persistence length. *Journal of molecular biology*, 317(2):205–213, 2002.

[13] S Geggier and A Vologodskii. Sequence dependence of DNA bending rigidity. *Proceedings of the National Academy of Science*, 107(35):15421–15426, 2010.

[14] V I. Ivanov, L E Minchenkova, E E Minyat, and M D Schyolkina. The B to A transition of DNA in solution. *Journal of biomolecular structure & dynamics*, 87(4):817–833, 1974.

[15] A Rich, A Nordheim, and A H Wang. The chemistry and biology of left-handed Z-DNA. *Annual Review of Biochemistry*, 53:791–846, 1984.

[16] Devarchive.cnx.org. Determining the Structure of DNA.

[17] A Bull, G Holt, and M Lilly. Biotechnology: International Trends and Perspectives. *Organisation für wirtschaftliche Zusammenarbeit und Entwicklung (OECD)*, pages 1–78, 1982.

[18] N C Seeman. Nucleic acid junctions and lattices. *Journal of theoretical biology*, 99(2):237–247, nov 1982.

[19] R Holliday. A mechanism for gene conversion in fungi. *Genetical Research*, (5):282–304, 1964.

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

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

[22] J C Mitchell, J R Harris, J Malo, J Bath, and A J Turberfield. Self-assembly of chiral DNA nanotubes. *Journal of the American Chemical Society*, 126(50):16342–16343, 2004.

[23] C Fan. *DNA Nanotechnology*. Springer Berlin Heidelberg, 2013.

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

[25] K E Dunn, F Dannenberg, T E Ouldridge, M Kwiatkowska, A J Turberfield, and J Bath. Guiding the folding pathway of DNA origami. *Nature*, 525(7567):82–86, 2015.

[26] W Liu, H Zhong, R Wang, and N C Seeman. Crystalline two-dimensional DNA-origami arrays. *Angewandte Chemie - International Edition*, 50(1):264–267, 2011.

[27] Y Ke, J Sharma, M Liu, K Jahn, Y Liu, and H Yan. Scaffolded DNA origami of a DNA tetrahedron molecular container. *Nano Letters*, 9(6):2445–2447, 2009.

[28] E S Andersen, M Dong, M M Nielsen, K Jahn, R Subramani, W Mamdouh, M M Golas, B Sander, H Stark, C L P Oliveira, J S Pedersen, V Birkedal, F Besenbacher, K V Gothelf, and J Kjems. Self-assembly of a nanoscale DNA box with a controllable lid. *Nature*, 459(7243):73–6, 2009.

[29] S M Douglas, H Dietz, T Liedl, B Hogberg, F Graf, and W M Shih. Self-assembly of DNA into nanoscale three-dimensional shapes. *Nature*, 459(7245):414–418, 2009.

[30] H Dietz, S M Douglas, and W M Shih. Folding DNA into Twisted and Curved Nanoscale Shapes. *Science*, 325(5941):725–730, 2009.

[31] T Liedl, B Hogberg, J Tytell, D E Ingber, and W M Shih. Self-assembly of three-dimensional prestressed tensegrity structures from DNA. *Nature Nanotechnology*, 5(7):520–524, 2010.

[32] B Saccà, R Meyer, M Erkelenz, K Kiko, A Arndt, H Schroeder, K S Rabe, and C M Niemeyer. Orthogonal protein decoration of DNA origami. *Angewandte Chemie - International Edition*, 49(49):9378–9383, 2010.

[33] T Torring and K V Gothelf. DNA nanotechnology: a curiosity or a promising technology? *F1000prime reports*, 5(May):14, jan 2013.

[34] R A Muscat, J Bath, and A J Turberfield. A programmable molecular robot. *Nano letters*, 11(3):982–7, mar 2011.

[35] N K Navani and Y Li. Nucleic acid aptamers and enzymes as sensors. *Current opinion in chemical biology*, 10(3):272–281, jun 2006.

[36] M Langecker, V Arnaut, T G Martin, J List, S Renner, M Mayer, H Dietz, and F C Simmel. Synthetic lipid membrane channels formed by designed DNA nanostructures. *Science (New York, N.Y.)*, 338(6109):932–6, nov 2012.

[37] G Binnig, H Rohrer, C Gerber, and E Weibel. Surface studies by scanning tunneling microscopy. *Physical review letters*, 49(1):57–61, 1982.

[38] R Jungmann, M Scheible, and F C Simmel. Nanoscale imaging in DNA nanotechnology. *WIRES Nanomedicine and Nanobiotechnology*, (4):66–81, 2012.

[39] G Binnig and C F Quate. Atomic Force Microscope. *Physical Review Letters*, 56(9):930–933, 1986.

[40] Y Martin, C C Williams, and H K Wickramasinghe. Atomic force microscope-force mapping and profiling on a sub 100 nm scale. *Journal of Applied Physics*, 61(10):4723–4729, 1987.

[41] Q Zhong, D Inniss, K Kjoller, and V B Elings. Fractured polymer/silica fiber surface studied by tapping mode atomic force microscopy. *Surface Science*, 290:L688–L692, 1993.

[42] H G Hansma, J Vesenka, C Siegerist, G Kelderman, H Morrett, V Elings, C Bustamante, and P K Hansma. Reproducible Imaging and Dissection of Plasmid DNA Under Liquid with the Atomic Force Microscope. *American Association for the Advancement of Science*, 256(5060):1180–1184, 2016.

[43] P K Hansma, J P Cleveland, M Radmacher, D A Walters, P E Hillner, M Bezanilla, M Fritz, D Vie, H G Hansma, C B Prater, J Massie, L Fukunaga, J Gurley, and V Elings. Tapping mode atomic force microscopy in liquids. *Applied Physics Letters*, 64(13):1738–1740, 1994.

[44] S F J Wickham, M Endo, Y Katsuda, K Hidaka, J Bath, H Sugiyama, and A J Turberfield. Direct observation of stepwise movement of a synthetic molecular transporter. *Nature Nanotechnology*, 6(3):166–169, 2011.

[45] Y He, T Ye, M Su, C Zhang, A E Ribbe, W Jiang, and C D Mao. Hierarchical self-assembly of DNA into symmetric supramolecular polyhedra. *Nature*, 452(7184):198—-U41, 2008.

[46] H G Hansma and D E Laney. DNA binding to mica correlates with cationic radius: assay by atomic force microscopy. *Biophysical journal*, 70(4):1933–1939, 1996.

[47] V Helms. Principles of Computational Cell Biology. In *Principles of Computational Cell Biology*, chapter Structure, page 202. Wiley-VCH, 2008.

[48] L Stryer. Fluorescence energy transfer as a spectroscopic ruler. *Annual Review of Biochemisiry*, 47(5):819, 1978.

[49] D Axelrod. Cell-substrate Contacts Illuminated by Total-Internal Reflection Fluorescence. *Journal of Cell Biology*, 89(9):141–145, 1981.

[50] N C Stellwagen. Electrophoresis. *Electrophoresis*, 30(1):1–14, 2010.

[51] D L Holmes and N C Stellwagen. The electric field dependence of DNA mobilities in agarose gels: a reinvestigation. *Electrophoresis*, 11(1):5–15, 1990.

[52] W W Grody, R M Nakamura, F L Kiechle, and C Strom. *Molecular Diagnostics: Techniques and Applications for the Clinical Laboratory*. Academic Press, 2009.

[53] G Lucotte and F Baneyx. *Introduction to Molecular Cloning Techniques.* Wiley-Blackwell, 1993.

[54] A Kuzyk, R Schreiber, Z Fan, G Pardatscher, E M Roller, A Högele, F C Simmel, A O Govorov, and T Liedl. DNA-based self-assembly of chiral plasmonic nanostructures with tailored optical response. *Nature*, 483(7389):311–314, 2012.

[55] P K Lo, P Karam, F A Aldaye, C K McLaughlin, G D Hamblin, G Cosa, and H F Sleiman. Loading and selective release of cargo in DNA nanotubes with longitudinal variation. *Nature chemistry*, 2(4):319–28, apr 2010.

[56] O I Wilner, Y Weizmann, R Gill, O Lioubashevski, R Freeman, and I Willner. Enzyme cascades activated on topologically programmed DNA scaffolds. *Nature Nanotechnology*, 4(4):249–254, 2009.

[57] H Yan, S H Park, G Finkelstein, J H Reif, and T H LaBean. DNA-templated self-assembly of protein arrays and highly conductive nanowires. *Science (New York, N.Y.)*, 301(5641):1882–4, sep 2003.

[58] P W K Rothemund, A Ekani-Nkodo, N Papadakis, A Kumar, D K Fygenson, and E Winfree. Design and characterization of programmable DNA nanotubes. *Journal of the American Chemical Society*, 126(50):16344–16352, 2004.

[59] F Mathieu, S Liao, J Kopatsch, T Wang, C Mao, and N C Seeman. Six-helix bundles designed from DNA. *Nano Letters*, 5(4):661–665, 2005.

[60] Y Fu, D Zeng, J Chao, Y Jin, Z Zhang, H Liu, D Li, H Ma, Q Huang, K V Gothelf, and C Fan. Single-step rapid assembly of DNA origami nanostructures for addressable nanoscale bioreactors. *Journal of the American Chemical Society*, 135(2):696–702, jan 2013.

[61] Z Li, M Liu, L Wang, J Nangreave, H Yan, and Y Liu. Molecular behavior of DNA origami in higher-order self-assembly. *Journal of the American Chemical Society*, 132(38):13545–13552, 2010.

[62] F A Aldaye, P K Lo, P Karam, C K McLaughlin, G Cosa, and H F Sleiman. Modular construction of DNA nanotubes of tunable geometry and single- or double-stranded character. *Nature Nanotechnology*, 4(6):349–352, 2009.

[63] O I Wilner, R Orbach, A Henning, C Teller, O Yehezkeli, M Mertig, D Harries, and I Willner. Self-assembly of DNA nanotubes with controllable diameters. *Nature Communications*, 2, 2011.

[64] 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 Research*, 37(15):5001–5006, 2009.

[65] A Kuzyk, R Schreiber, Z Fan, G Pardatscher, E Roller, A Högele, F C Simmel, A O Govorov, and T Liedl. DNA-based self-assembly of chiral plasmonic nanostructures with tailored optical response. *Nature*, 483(7389):311–314, 2012.

[66] S M Douglas, J J Chou, and W M Shih. DNA-nanotube-induced alignment of membrane proteins for NMR structure determination. *Proceedings of the National Academy of Sciences of the United States of America*, 104(16):6644–6648, apr 2007.

[67] M Endo, S Yamamoto, T Emura, K Hidaka, N Morone, J E Heuser, and H Sugiyama. Helical DNA origami tubular structures with various sizes and arrangements. *Angewandte Chemie - International Edition*, 53(29):7484–7490, 2014.

[68] R Owczarzy, B G Moreira, Y You, M A Behlke, and J A Wälder. Predicting stability of DNA duplexes in solutions containing magnesium and monovalent cations. *Biochemistry*, 47(19):5336–5353, 2008.

[69] P Neill, P W K Rothemund, A Kumar, and D K Fygenson. Sturdier DNA nanotubes via ligation. *Nano Letters*, 6(7):1379–1383, 2006.

[70] WolframMathematica. LaplacianGaussianFilter, 2008.

[71] P J Flory. Molecular Size Distribution in Linear Condensation Polymers. *Journal of the American Chemical Society*, 58(7):1877–1885, 1936.

[72] R Jungmann, M Scheible, A Kuzyk, G Pardatscher, C E Castro, and F C Simmel. DNA origami-based nanoribbons: assembly, length distribution, and twist. *Nanotechnology*, 22(27), 2011.

[73] Flory, P J. *Principles of Polymer Chemistry*. Cornell University Press, 1953.

[74] JP J Sobczak, T G Martin, T Gerling, and H Dietz. Rapid Folding of DNA into Nanoscale Shapes at Constant Temperature. *Science*, 338(6113):1458–1461, 2012.

[75] R Jungmann, T Liedl, T L Sobey, W Shih, and F C Simmel. Isothermal assembly of DNA origami structures using denaturing agents. *Journal of the American Chemical Society*, 130(31):10062–3, aug 2008.

[76] R D Blake and S G Delcourt. Thermodynamic effects of formamide on DNA stability. *Nucleic Acids Research*, 24(11):2095–2103, 1996.

[77] G Mie. A contribution to the optics of turbid media, especially colloidal metallic suspensions. *Annual of Physics*, 25:377–445, 1908.

[78] SigmaAldrich. Gold Nanoparticles: Properties and Applications. 2016.

[79] K Richter, M Nessling, and P Lichter. Macromolecular crowding and its potential impact on nuclear function. *Biochimica et Biophysica Acta - Molecular Cell Research*, 1783(11):2100–2107, 2008.

[80] R J Ellis. Macromolecular crowding: obvious but under appreciated. *Trends in Biochemical Sciences*, 26(10):597–604, 2001.

[81] Andrey S Shaw and Erin L Filbert. Scaffold proteins and immune-cell signalling. *Nature reviews. Immunology*, 9(1):47–56, 2009.

[82] M C Good, J G Zalatan, and W A Lim. Scaffold proteins: hubs for controlling the flow of cellular information. *Science*, 332(6030):680–6, may 2011.

[83] P A Srere. The metabolon. *Trends in Biochemical Sciences*, 10(3):109–110, 1985.

[84] Y. H Percival Zhang. Substrate channeling and enzyme complexes for biotechnological applications. *Biotechnology Advances*, 29(6):715–725, 2011.

[85] D. K. Srivastava and Sidney A. Bernhard. Metabolite Transfer via Enzyme-Enzyme Complexes. *Science*, 234(4780):1081–1086, 1986.

[86] D E Atkinson. *Cellular Energy Metabolism and Its Regulation*. Academic Press, New York, 1977.

[87] J Ovadi. Physiological significance of metabolic channelling. *Journal of Theoretical Biology*, 152(1):1–22, 1991.

[88] J Ovadi. Physiological significance of metabolite channeling: author's response to commentaries. *Journal of theoretical biology*, 152:135–141, 1991.

[89] C Lindbladh, M Rault, C Hagglund, W C Small, K Mosbach, L Bülow, C Evans, and P a Srere. Preparation and kinetic characterization of a fusion protein of yeast mitochondrial citrate synthase and malate dehydrogenase. *Biochemistry*, 33:11692–11698, 1994.

[90] P Ljungcrantz, H Carlsson, M O Månsson, P Buckel, K Mosbach, and L Bülow. Construction of an artificial bifunctional enzyme, beta-galactosidase/galactose dehydrogenase, exhibiting efficient galactose channeling. *Biochemistry*, 28(22):8786–92, 1989.

[91] H Pettersson and G Pettersson. Kinetics of the coupled reaction catalysed by a fusion protein of β-galactosidase and galactose dehydrogenase. *Biochimica et Biophysica Acta - Protein Structure and Molecular Enzymology*, 1549(2):155–160, 2001.

[92] H Pettersson and G Pettersson. Kinetics of the coupled reaction catalysed by a fusion protein of β-galactosidase and galactose dehydrogenase. *Biochimica et Biophysica Acta - Protein Structure and Molecular Enzymology*, 1549(2):155–160, 2001.

[93] K Jacobson and J Wojcieszyn. The translational mobility of substances within the cytoplasmic matrix. *Proceedings of the National Academy of Sciences of the United States of America*, 81(November):6747–6751, 1984.

[94] H O Spivey and J M Merz. Metabolic compartmentation. *BioEssays : news and reviews in molecular, cellular and developmental biology*, 10(4):127–30, 1989.

[95] M K Geck and J F Kirsch. A novel, definitive test for substrate channeling illustrated with the aspartate aminotransferase/-malate dehydrogenase system. *Biochemistry*, 38(25):8032–8037, 1999.

[96] I Wheeldon, S D Minteer, S Banta, S C Barton, P Atanassov, and M Sigman. Substrate channelling as an approach to cascade reactions. *Nature Chemistry*, 8(4):299–309, 2016.

[97] J M Krahn, J Hyun Kim, M R Burns, R J Parry, H Zalkin, and J L Smith. Coupled formation of an amidotransferase interdomain ammonia channel and a phosphoribosyltransferase active site. *Biochemistry*, 36(37):11061–11068, 1997.

[98] J B. Thoden, H M. Holden, G Wesenberg, F M Raushel, and I Rayment. Structure of carbamoyl phosphate synthetase: A journey of 96 Å from substrate to product. *Biochemistry*, 36(21):6305–6316, 1997.

[99] X Huang and Frank M Raushel. Enzyme Catalysis and Regulation: Restricted Passage of Reaction Intermediates through the Ammonia Tunnel of Carbamoyl Phosphate Synthetase Restricted Passage of Reaction Intermediates through the Ammonia Tunnel of Carbamoyl Phosphate Synthetase. *Journal of Biological Chemistry*, 275(34):26333–26240, 2000.

[100] J Carere, S McKenna, M S Kimber, and S Y K Seah. Characterization of an aldolase-dehydrogenase complex from the cholesterol degradation pathway of Mycobacterium tuberculosis. *Biochemistry*, 52(20):3502–3511, 2013.

[101] C C Hyde, S A Ahmed, E A Padlan, E W Miles, and D R Davies. Three-dimensional structure of the tryptophan synthase $\alpha 2\beta 2$ multienzyme complex from Salmonella typhimurium. *The Journal of Biological Chemistry*, 263(33):17857–71, 1988.

[102] M Q Fatmi, R Ai, and C E A Chang. Synergistic regulation and ligand-induced conformational changes of tryptophan synthase. *Biochemistry*, 48(41):9921–9931, 2009.

[103] D R Knighton, K Chen-Chen, E Howland, and C A Janson. Structure of and kinetic channelling in bifunctional dihydrofolate reductase-thymidylate synthase. *Nature*, 1(3):186–194, 1994.

[104] T D Meek, E P Garvey, and D V Santi. Purification and characterization of the bifunctional thymidylate synthetase-dihydrofolate reductase from methotrexate-resistant Leishmania tropica. *Biochemistry*, 24(3):678–686, 1985.

[105] A H Elcock and G A Huber. Electrostatic channeling of substrates between enzyme active sites: Comparison of simulation and experiment. *Biochemistry*, 36(51):16049–16058, 1997.

[106] B S Winkel. Metabolic Channeling in Plants. *Annual Review of Plant Biology*, 55(1):85–107, 2004.

[107] K Jørgensen, A V Rasmussen, M Morant, A H Nielsen, N Bjarnholt, M Zagrobelny, S Bak, and Birger L Møller. Metabolon formation and metabolic channeling in the biosynthesis of plant natural products. *Current Opinion in Plant Biology*, 8(3):280–291, 2005.

[108] M Smolle, A E Prior, A E Brown, A Cooper, O Byron, and H G Lindsay. A new level of architectural complexity in the human pyruvate dehydrogenase complex. *Journal of Biological Chemistry*, 281(28):19772–19780, 2006.

[109] M Trujillo, R G K Donald, D S Roos, P J Greene, and D V Santi. Heterologous expression and characterization of the bifunctional dihydrofolate reductase-thymidylate synthase enzyme of Toxoplasma gondii. *Biochemistry*, 35(20):6366–6374, 1996.

[110] F. Kazenwadel, M. Franzreb, and B. E. Rapp. Synthetic enzyme supercomplexes: co-immobilization of enzyme cascades. *Analytical Methods*, 7(10):4030–4037, 2015.

[111] X Deng, T W Eyster, Y Elkasabi, and J Lahann. Bio-orthogonal polymer coatings for co-presentation of biomolecules. *Macromolecular Rapid Communications*, 33(8):640–645, 2012.

[112] L Zhu, R Yang, J Zhai, and C Tian. Bienzymatic glucose biosensor based on co-immobilization of peroxidase and glucose oxidase on a carbon nanotubes electrode. *Biosensors and Bioelectronics*, 23(4):528–535, 2007.

[113] C Mateo, A Chmura, and S Rustler. Synthesis of enantiomerically pure (S)-mandelic acid using an oxynitrilase-nitrilase bienzymatic cascade: A nitrilase surprisingly shows nitrile hydratase activity. *Tetrahedron Asymmetry*, 17(3):320–323, 2006.

[114] C You, S Myung, and Y H Zhang. Facilitated substrate channeling in a self-assembled trifunctional enzyme complex. *Angewandte Chemie - International Edition*, 51(35):8787–8790, 2012.

[115] S F M van Dongen, M Nallani, J J L M Cornelissen, R J M Nolte, and J C M van Hest. A three-enzyme cascade reaction through positional assembly of enzymes in a polymersome nanoreactor. *Chemistry (Weinheim an der Bergstrasse, Germany)*, 15(5):1107–14, jan 2009.

[116] C M Niemeyer, J Koehler, and C Wuerdemann. DNA-directed assembly of bienzymic complexes from in vivo biotinylated NAD(P)H:FMN oxidoreductase and luciferase. *ChemBioChem*, 3(2-3):242–245, 2002.

[117] C R Boehm, P S Freemont, and O Ces. Design of a prototype flow microreactor for synthetic biology in vitro. *Lab on a chip*, 13(17):3426–32, 2013.

[118] Sung Ha Park, Robert Barish, Hanying Li, John H. Reif, Gleb Finkelstein, Hao Yan, and Thomas H. LaBean. Three-helix bundle DNA tiles self-assemble into 2D lattice or 1D templates for silver nanowires. *Nano Letters*, 5(4):693–696, 2005.

[119] Y Liu, C Lin, H Li, and H Yan. Aptamer-directed self-assembly of protein arrays on a DNA nanostructure. *Angewandte Chemie - International Edition*, 44(28):4333–4338, 2005.

[120] R F Macaya, P Schultze, F W Smith, J A Roe, and J Feigon. Thrombin-binding DNA aptamer forms a unimolecular quadruplex structure in solution. *Proceedings of the National Academy of Sciences*, 90(8):3745–3749, 1993.

[121] S Rinker, Y G Ke, Y Liu, R Chhabra, and H Yan. Self-assembled DNA nanostructures for distance-dependent multivalent ligand-protein binding. *Nature Nanotechnology*, 3(7):418–422, 2008.

[122] J L Fu, M H Liu, Y Liu, N W Woodbury, and H Yan. Interenzyme Substrate Diffusion for an Enzyme Cascade Organized on Spatially Addressable DNA Nanostructures. *Journal of the American Chemical Society*, 134(12):5516–5519, 2012.

[123] J E Dueber, G C Wu, G R Malmirchegini, T S Moon, C J Petzold, A V Ullal, K L J Prather, and J D Keasling. Synthetic protein scaffolds provide modular control over metabolic flux. *Nature*, 27(8):5–8, 2009.

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

[125] M Liu, J Fu, C Hejesen, Y Yang, N W Woodbury, K Gothelf, Y Liu, and H Yan. A DNA tweezer-actuated enzyme nanoreactor. *Nature Communications*, 4:2127, 2013.

[126] Y Hu, F Wang, C Lu, J Girsh, E Golub, and T Willner. Switchable Enzyme/DNAzyme Cascades by the Reconfiguration of DNA Nanostructures. *Chemistry*, 20(49):16203–9, dec 2014.

[127] L Michaelis and M L Menten. Die Kinetik der Invertinwirkung. *Biochem Z*, 49(February):333–369, 1913.

[128] O Idan. Modeling Nanoscale Transport Systems. 2014.

[129] O Idan and H Hess. Origins of activity enhancement in enzyme cascades on scaffolds. *ACS Nano*, 7(10):8658–8665, 2013.

[130] S Libertino, V Aiello, A Scandurra, M Renis, and F Sinatra. Immobilization of the enzyme glucose oxidase on both bulk and porous SiO2 surfaces. *Sensors*, 8(9):5637–5648, 2008.

[131] J Sanz-Aparicio and A Polo. Structure of invertase from Schwanniomyces occidentalis. *Protein Data Bank*, 2009.

[132] C Bustamante, Y R Chemla, N R Forde, and D Izhaky. Mechanical processes in biochemistry. *Annual Review of Biochemistry*, 73(1):705–748, 2004.

[133] G Oster and H Wang. How Protein Motors Convert Chemical Energy into Mechanical Work. In M Schliwa, editor, *Molecular Motors*, chapter 8, pages 207–226. Wiley-VCH, Weinheim, 2004.

[134] R P Feynman. Vol. 1. In *The Feynman Lectures on Physics*, chapter 46. Addison-Wesley, Massachusetts, 1963.

[135] M O Magnasco. Molecular combustion motors. *Physical Review Letters*, 72(16):2656—-2659, 1994.

[136] T Cha, J Pan, H Chen, J Salgado, X Li, C Mao, and J H Choi. A synthetic DNA motor that transports nanoparticles along carbon nanotubes. *Nature Nanotechnology*, 9(1):39–43, 2014.

[137] T Omabegho, R Sha, and N C Seeman. Coordinated Legs. *Science*, 324(April):67–71, 2009.

[138] J Bath, S J Green, and A J Turberfield. A free-running DNA motor powered by a nicking enzyme. *Angewandte Chemie - International Edition*, 44(28):4358–4361, 2005.

[139] M Liber, T E Tomov, R Tsukanov, Y Berger, and E Nir. A Bipedal DNA Motor that Travels Back and Forth between Two DNA Origami Tiles. *Small (Weinheim an der Bergstrasse, Germany)*, 11(5):568–75, 2015.

[140] S F J Wickham, J Bath, Y Katsuda, M Endo, K Hidaka, H Sugiyama, and A J Turberfield. A DNA-based molecular motor that can navigate a network of tracks. *Nature Nanotechnology*, 7(3):169–173, 2012.

[141] Y Chen, M Wang, and C Mao. An autonomous DNA nanomotor powered by a DNA enzyme. *Angewandte Chemie - International Edition*, 43(27):3554–3557, 2004.

[142] 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, 2004.

[143] J Shin and N A Pierce. A synthetic DNA walker for molecular transport. *Journal of the American Chemical Society*, 126(35):10834–10835, 2004.

[144] W B Sherman and N C Seeman. A precisely controlled DNA biped walking device. *Nano Letters*, 4(7):1203–1207, 2004.

[145] N Wu, D M Czajkowsky, J Zhang, J Qu, M Ye, D Zeng, X Zhou, J Hu, Z Shao, B Li, and C Fan. Molecular threading and tunable molecular recognition on DNA origami nanostructures. *Journal of the American Chemical Society*, 135(33):12172–5, aug 2013.

[146] A E Lucas. *Dynamic DNA motors and structures*. Oxford University, 2016.

[147] J N Zadeh, C D Steenberg, J S Bois, B R Wolfe, M B Pierce, A R Khan, R M Dirks, and N A. Pierce. NUPACK: Analysis and Design of Nucleic Acid Systems. *Journal of computational chemistry*, 32(16):170–173, 2011.

[148] Integrated DNA Technologies. Fluorophores Modifications, 2016.

[149] S Matsuura, J Komatsu, K Hirano, H Yasuda, K Takashima, S Katsura, and A Mizuno. Real-time observation of a single DNA digestion by lambda exonuclease under a fluorescence microscope field. *Nucleic acids research*, 29(16):E79, 2001.

[150] R R F Machinek. Control and Observation of DNA Nanodevices, 2014.

[151] B Kundukad, P S Doyle, and J Yan. Effect of YOYO-1 on the mechanical properties of DNA. *Soft Matter*, 10(48):9721–9728, 2014.

[152] T Zal and N R J Gascoigne. Photobleaching-corrected FRET efficiency imaging of live cells. *Biophysical journal*, 86(6):3923–39, 2004.

[153] R Jungmann, C Steinhauer, M Scheible, A Kuzyk, P Tinnefeld, and F C Simmel. Single-molecule kinetics and super-resolution microscopy by fluorescence imaging of transient binding on DNA origami. *Nano letters*, 10(11):4756–61, nov 2010.

[154] Borman, S. Combinatorial chemists focus on small molecules, molecular recognition, and automation. *American Chemical Society*, (Chemical & Engineering News), 1996.

[155] O Ramström and J Lehn. Drug discovery by dynamic combinatorial libraries. *Nature reviews. Drug discovery*, 1(1):26–36, jan 2002.

[156] R E Kleiner, C E Dumelin, and D R Liu. Small-molecule discovery from DNA-encoded chemical libraries. *Chemical Society reviews*, 40(12):5707–5717, dec 2011.

[157] P T Corbett, J Leclaire, L Vial, K R West, J Wietor, J K M Sanders, and S Otto. Dynamic combinatorial chemistry. *Chemical reviews*, 106(9):3652–711, sep 2006.

[158] G P Smith. Filamentous Fusion Phage : Novel Expression Vectors that Display Cloned Antigens on the Virion Surface. *Science*, 228(4705):1315–1317, 2016.

[159] R W Roberts and J W Szostakt. RNA-Peptide Fusions for the in vitro Selection of Peptides and Proteins. *Proceedings of the National Academy of Sciences*, 94(23):12297–12302, 2016.

[160] S Melkko, J Scheuermann, C E Dumelin, and D Neri. Encoded self-assembling chemical libraries. *Nature biotechnology*, 22(5):568–74, may 2004.

[161] S Melkko, Y Zhang, C E Dumelin, J Scheuermann, and D Neri. Isolation of high-affinity trypsin inhibitors from a DNA-encoded chemical library. *Angewandte Chemie*, 46(25):4671–4, jan 2007.

[162] J Scheuermann, C E Dumelin, S Melkko, Y Zhang, L Mannocci, M Jaggi, J Sobek, and D Neri. DNA-encoded chemical libraries for the discovery of MMP-3 inhibitors. *Bioconjugate chemistry*, 19(3):778–85, mar 2008.

[163] M H Hansen, P Blakskjaer, L K Petersen, T H Hansen, J W Højfeldt, K V Gothelf, and N J V Hansen. A yoctoliter-scale

DNA reactor for small-molecule evolution. *Journal of the American Chemical Society*, 131(3):1322–7, jan 2009.

[164] X Li and D R Liu. DNA-templated organic synthesis: nature's strategy for controlling chemical reactivity applied to synthetic molecules. *Angewandte Chemie (International ed. in English)*, 43(37):4848–4870, sep 2004.

[165] F V Reddavide, W Lin, S Lehnert, and Y Zhang. DNA-Encoded Dynamic Combinatorial Chemical Libraries. *Angewandte Chemie*, 54(27):7924–8, 2015.

[166] H Caruthers. Chemical Synthesis of DNA and DNA Analogues. (21):278–284, 1991.

[167] C B Reese. Oligo- and poly-nucleotides: 50 years of chemical synthesis. *Organic & biomolecular chemistry*, 3(21):3851–68, nov 2005.

[168] A H El-Sagheer and T Brown. Click Nucleic Acid Ligation: Applications in Biology and Nanotechnology. *Accounts of Chemical Research*, 45(8):1258–1267, 2012.

[169] P Kocalka, A H El-Sagheer, and T Brown. Rapid and efficient DNA strand cross-linking by click chemistry. *Chembiochem : a European journal of chemical biology*, 9(8):1280–5, may 2008.

[170] R Kumar, A El-Sagheer, J Tumpane, P Lincoln, L M Wilhelmsson, and T Brown. Template-directed oligonucleotide strand ligation, covalent intramolecular DNA circularization and catenation using click chemistry. *Journal of the American Chemical Society*, 129(21):6859–64, may 2007.

[171] A H El-Sagheer and T Brown. New strategy for the synthesis of chemically modified RNA constructs exemplified by hairpin and hammerhead ribozymes. *Proceedings of the National*

Academy of Sciences of the United States of America, 107(35):15329–34, aug 2010.

[172] V V Rostovtsev, L G Green, V V Fokin, and K B Sharpless. A stepwise huisgen cycloaddition process: copper(I)-catalyzed regioselective "ligation" of azides and terminal alkynes. *Angewandte Chemie (International ed. in English)*, 41(14):2596–9, jul 2002.

[173] H C Kolb, M G Finn, and K B Sharpless. Click Chemistry: Diverse Chemical Function from a Few Good Reactions. *Angewandte Chemie (International ed. in English)*, 40(11):2004–2021, jun 2001.

[174] A H El-Sagheer and T Brown. Synthesis and polymerase chain reaction amplification of DNA strands containing an unnatural triazole linkage. *Journal of the American Chemical Society*, 131(11):3958–64, mar 2009.

[175] A H El-Sagheer, A P Sanzone, R Gao, A Tavassoli, and T Brown. Biocompatible artificial DNA linker that is read through by DNA polymerases and is functional in Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America*, 108(28):11338–11343, 2011.

[176] R P Goodman, R M Berry, and A J Turberfield. The single-step synthesis of a DNA tetrahedron. *Chemical Communications*, (12):1372–1373, 2004.

[177] Z Li, B Wei, J Nangreave, C X Lin, Y Liu, Y L Mi, and H Yan. A Replicable Tetrahedral Nanostructure Self-Assembled from a Single DNA Strand. *Journal of the American Chemical Society*, 131(36):13093–13098, 2009.

[178] Z Shen, H Yan, T Wang, and N C Seeman. Paranemic crossover DNA: a generalized Holliday structure with applications in nanotechnology. *Journal of the American Chemical Society*, 126(6):1666–1674, feb 2004.

[179] C H Spink, L Ding, Q Yang, R D Sheardy, and N C Seeman. Thermodynamics of forming a parallel DNA crossover. *Biophysical journal*, 97(2):528–538, jul 2009.

[180] K A Afonin, D J Cieply, and N B Leontis. Specific RNA Self-Assembly with Minimal Paranemic Motifs. (3):455–463, 2008.

[181] S D Jayasena. Aptamers: an emerging class of molecules that rival antibodies in diagnostics. *Clinical chemistry*, 45(9):1628–50, sep 1999.

[182] H Khan and S Missailidis. Aptamers in oncology : a diagnostic perspective Aptamers in oncology : a diagnostic perspective. *Gene Ther Mol Biol*, 12:111–128, 2008.

[183] C Tuerk and L Gold. Systematic Evolution of Ligands by Exponential Enrichment: RNA Ligands to Bacteriophage T4 DNA Polymerase. *Science*, 249:505–510, 1990.

[184] A D Ellington and J W Szostak. In vitro selection of RNA molecules that bind specific ligands. *Nature*, 346:818–822, 1990.

[185] L C Bock, L C Griffin, J A Latham, W H Vermaas, and J J Toole. Selection of single-stranded DNA molecules that bind and inhibit human thrombin. *Nature*, 355:564–566, 1992.

[186] J J. Li, X Fang, and W Tan. Molecular Aptamer Beacons for Real-Time Protein Recognition. *Biochemical and Biophysical Research Communications*, 292(1):31–40, mar 2002.

[187] A Sosic, A Meneghello, E Cretaio, and B Gatto. Human thrombin detection through a sandwich aptamer microarray: interaction analysis in solution and in solid phase. *Sensors (Basel, Switzerland)*, 11(10):9426–9441, jan 2011.

[188] I R Krauss, A Pica, A Merlino, L Mazzarella, and F Sica. Duplex-quadruplex motifs in a peculiar structural organization cooperatively contribute to thrombin binding of a DNA aptamer. *Acta Crystallographica Section D: Biological Crystallography*, 69(12):2403–2411, 2013.

[189] K M Ahmad, Y Xiao, and H T Soh. Selection is more intelligent than design: improving the affinity of a bivalent ligand through directed evolution. *Nucleic acids research*, 40(22):11777–11783, dec 2012.

[190] Editor. DNA nanotechnology 2.5. *Nature Nanotechnology*, 10(9):729–729, 2015.

[191] S Surana, A R Shenoy, and Y Krishnan. Designing DNA nanodevices for compatibility with the immune system of higher organisms. *Nature nanotechnology*, 10(9):741–7, 2015.

[192] T Wang, R Sha, R Dreyfus, M E Leunissen, C Maass, D J Pine, P M Chaikin, and N C Seeman. Self-replication of information-bearing nanoscale patterns. *Nature*, 478(7368):225–228, 2011.

[193] Rye, H S. Stable fluorescent complexes of double-stranded DNA with bis-intercalating asymmetric cyanine dyes: properties and applications. *Nucleic acids research*, 20(11):2803–2812, 1992.

[194] P R Selvin and T Ha. Single-molecule Techniques: A Laboratory Manual. page 507. 2008.

[195] D Cooper, H Uhm, L J Tauzin, N Poddar, and C F Landes. Photobleaching lifetimes of cyanine fluorophores used for single-molecule Forster resonance energy transfer in the presence of various photoprotection systems. *ChemBioChem*, 14(9):1075–1080, 2013.