Folding Pathways of DNA Nanostructures



Katherine Gwyneth Young

Trinity College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Michaelmas Term 2018

FOLDING PATHWAYS OF DNA NANOSTRUCTURES

Katherine Young, Trinity College

Abstract of thesis submitted for the degree of Doctor of Philosophy,

University of Oxford, Michaelmas Term 2018

DNA origami is a robust technique for the production of DNA nanostructures, whereby a long scaffold strand is folded through interactions with many short staple strands. This thesis describes my work on two DNA origami-based systems. One is an origami tile whose folding can be kinetically-controlled through the use of 'helper strands', which are designed to bind to the scaffold. This allows the tile to be folded rapidly at temperatures below the tile melting temperature, which is not otherwise possible due to the formation and persistence of off-target staple-template interactions at such temperatures. Omission of selected helpers changes the folding pathway of the tile; this may allow the use of helper strands to 'prime' a template to preferentially fold with a specific staple set in a mixture of different staple sets.

The second system is a novel 'origami-like' structure in which the equivalent of the scaffold and staples are both double-stranded structures which interact through DNA T-motifs. This differs from conventional origami in that: a) binding domain sequences are reused, and b) the interactions which govern folding (T-motifs) are less stable than the staple-template hybridisations used in conventional origami. The structure I designed uses only four unique binding domains (seven copies of each) and two types of 'linker' (equivalent to staples), which means there are a large number of potential structures which could form. Despite this, folding of the target structure was seen at room temperature, suggesting that off-target interactions can be reversed at these temperatures to allow the structure to reach its minimum free energy. Changing the identity of only one of the linkers results in folding of a different target structure. These 'T-motif origami' structures can be used to investigate the minimum number of unique binding domains required to specify a structure, and for applications which require reconfigurability.

Both systems were characterised using gel electrophoresis and AFM imaging.

Acknowledgements

I would like to thank my supervisor, Prof. Andrew Turberfield, for his guidance, ideas and the feedback which helped shape this thesis. Thank you to my second (accidental) supervisor, Dr Jon Bath, for his continual help and patience in steering me in the right direction. I am immensely grateful for his help in turning things around when I had a 'crisis of confidence' in late 2017-early 2018, and this thesis is much stronger for the work I've achieved since then.

Behnam Najafi has provided the simulation results included in this thesis, which have greatly enhanced our understanding of both the helper-origami and T-motif systems. I thank him both for these contributions and for his patience when explaining his work to me.

Thank you to the other members of the Turberfield group, past and present, including Ibon Santiago, Florence Benn, Céline Journot, Antonio Garcia Guerra, Natalie Haley, Seham Helmi, Juan Jin, Sam Tusk, Joel Spratt, Robbie Oppenheimer and Emma Silvester, for being great company and lovely to work with. Our lab tech Sanna Piippo-Henderson also deserves a big thanks for keeping us organised and well-stocked.

I would like to thank my college advisor Louis Mahadevan, who has been a great source of advice, both academic and personal, during the course of my DPhil.

To my friends in Oxford and elsewhere, particularly Gemma, Nicola, Cathi, Susie, Kate and Leonie, thank you for all of your support/encouragement/food/hugs. It made such a big difference during the writing of this thesis.

Last but by no means least, I would like to thank my parents (Pam and Dick) for

all of their support and help, in the many forms they have provided it, throughout my DPhil but particularly this year. Thanks especially for providing me with an excellent 'writing retreat' (5 stars on TripAdvisor), and to my mum for proofreading this thesis.

My funding was provided by the EPSRC through the Life Sciences Interface Doctoral Training Centre.

To my parents

Contents

1 Introduction **11**

1.1 DNA 11

1.1.1 DNA as a building material 14

1.1.2 Strand displacement 16

1.2 Folding of DNA nanostructures 17

1.2.1 Early DNA nanostructures 18

1.2.2 DNA origami 19

1.2.3 Single-stranded tiles 21

1.3 Applications of DNA nanostructures 23

1.4 Future perspectives and challenges 25

2 Folding of a DNA Origami Nanostructure using 'Helper' Strands 29

2.1 Tile design 32

2.2 Helper strand design 35

2.3 Experimental design 38

2.3.1 Atomic force microscopy (AFM) 40

2.3.2 Gel electrophoresis 47

2.3.3 Temperature-controlled real-time fluorescence meas- urements 51

2.3.4 Simulations 55

2.4 Helper strands improve isothermal origami tile folding 56

2.4.1 Helper strands improve isothermal tile folding at 60°C-65°C . 57

2.4.2 Temperature-dependent transition from helper-coated template to fully-folded tile . . . 61
2.4.3 Effect of incubation time . . . 70
2.4.4 Further evidence of the role of helpers . . . 72
2.4.5 Isothermal folding of DNA origami . . . 74
2.4.6 Discussion . . . 81
2.5 Omitting selected helper strands further improves folding . . . 86
2.5.1 Nucleation . . . 86
2.5.2 Propagation . . . 88
2.5.3 Using helpers to fold multiple origami structures from a mixture of different staple sets . . . 99
2.5.4 Why does omitting helpers at the seam favour tile folding? . . . 101
2.6 Conclusions and further work . . . 104

3 Characterisation of T-motifs 106

3.1 Methods . . . 109
3.1.1 Use of distance markers in gel electrophoresis . . . 109
3.1.2 Construction of binding curves . . . 111
3.1.3 Controls . . . 116
3.1.4 Strand sequences, lengths and concentrations . . . 116
3.1.5 T-motif simulation results . . . 119
3.2 Results . . . 120
3.2.1 Possible sources of experimental error . . . 123
3.2.2 Analysis . . . 128
3.3 Summary . . . 133

4 Building Nanostructures with T-motifs 134

4.1 Testing individual linker binding . . . 137
4.1.1 Hierarchical linker binding . . . 138
4.1.2 Studying geometry . . . 142

4.2 T-motif origami-like structure . . . 147
4.2.1 Design . . . 147
4.2.2 Results . . . 151
4.3 Conclusions and further work . . . 170
4.3.1 Further work . . . 171

5 Conclusions **173**
5.1 Future directions . . . 174

A Materials and Methods **176**
A.1 Sources of materials and definitions . . . 176
A.2 DNA origami design and synthesis . . . 177
A.2.1 Design . . . 177
A.2.2 Template synthesis . . . 178
A.2.3 Origami preparation . . . 180
A.3 T-motif and T-motif origami design and synthesis . . . 180
A.3.1 Design . . . 180
A.3.2 Sample preparation . . . 182
A.4 Gel electrophoresis . . . 185
A.4.1 Analysis . . . 186
A.5 Temperature-controlled real-time fluorescence measurements . . . 187
A.5.1 Analysis . . . 188
A.6 Atomic force microscopy (AFM) . . . 188
A.6.1 Processing of AFM images . . . 189
A.6.2 Quantification of origami AFM images . . . 190
A.7 DNA quantification . . . 191

B DNA sequences **192**
B.1 DNA Origami (Chapter 2) . . . 192
B.1.1 Staple sequences . . . 193
B.1.2 10-offset helper set . . . 196

B.1.3 8-offset helper set 200
B.2 Single T-motifs (Chapter 3) 205
B.3 U-shaped T-motif structures (Chapter 4, section 4.1) 207
B.4 T-motif origami (Chapter 4, section 4.2) 211

C Program codes **213**

C.1 Measuring relative electrophoretic mobility using distance markers (Fiji macro) 213
C.2 Single T-motif binding curve fitting (Matlab) 215

Acronyms and Definitions

AFM	atomic force microscope/microscopy
bp	base pair(s)
DNA	deoxyribonucleic acid
FRET	Förster or fluorescence resonance energy transfer
(m)RNA	(messenger) ribonucleic acid
nt	nucleotide(s)
PCR	polymerase chain reaction

In this thesis, **'annealing'** or **'thermal annealing'** refers to heating a sample of DNA strands to 90 or 95°C (so that all hybridisation interactions are melted), before slowly cooling to room temperature (20 or 25°C). In principle, this allows the strands to reach thermodynamic equilibrium if done sufficiently slowly. For the experiments in this thesis, the protocol used was heating to 95°C and then cooling to 20°C in increments of 0.1°C, at a rate of 1°C/min (see Appendix A.1).

I have used **'mis-binding'** to refer to the formation of off-target interactions during DNA nanostructure folding, particularly those between staples and template during DNA origami folding. Off-target refers to those which are not present in the target structure.

Many of the results shown in this thesis are of gel electrophoresis experiments. Strand concentrations quoted for these experiments refer to the concentration in the sample before loading buffer was added and the sample loaded onto the gel. For all gels (both agarose and polyacrylamide), a 0.2× volume of 5× loading buffer was added to the sample, and then 10 μL of sample loaded onto the gel.

Unless otherwise stated, all figures in this thesis are original.

Chapter 1

Introduction

DNA nanotechnology utilises the Watson-Crick base-pairing of DNA to design nanoscale structures and dynamic machines which form by self-assembly. Since this idea was first proposed and used by Ned Seeman to construct junctions made of DNA in the early 1980s [64, 126, 129], the field has been growing rapidly in terms of both the range of structures available and their potential applications. The development of 'DNA origami' by Paul Rothemund in 2006 [122] has been instrumental in this growth, and it remains the most robust [58], and thus one of the most widely used [56], methods for building DNA nanostructures.

The work presented in this thesis looks at two 2D DNA nanostructures: 1) an origami tile which can be folded using a novel method (Chapter 2); and 2) an 'origami-like' structure, where the use of different interactions to those present in conventional origami results in novel properties (Chapters 3 and 4).

1.1 DNA

DNA (deoxyribonucleic acid) is a naturally-occurring polymer, found in the nuclei of the cells of living organisms (or in bacteria), where it acts as the genetic material. A strand of DNA consists of many subunits called nucleotides (nt) joined in a chain. Each of these nucleotides consists of a deoxyribose sugar, a phosphate group, and a nitrogenous base, as shown in Figure 1.1a [81]. The 3'-hydroxyl group of the sugar

moiety of one nucleotide reacts with the phosphate group of another nucleotide to form a covalent phosphodiester bond ('ligation'), thus forming a strand of DNA with a 'sugar-phosphate backbone' (Figure 1.1b). Each strand of DNA has a chemical polarity which depends on the carbon atoms of the sugar moieties at either end, as shown in Figure 1.1b. DNA sequences are conventionally written starting from the 5' end.

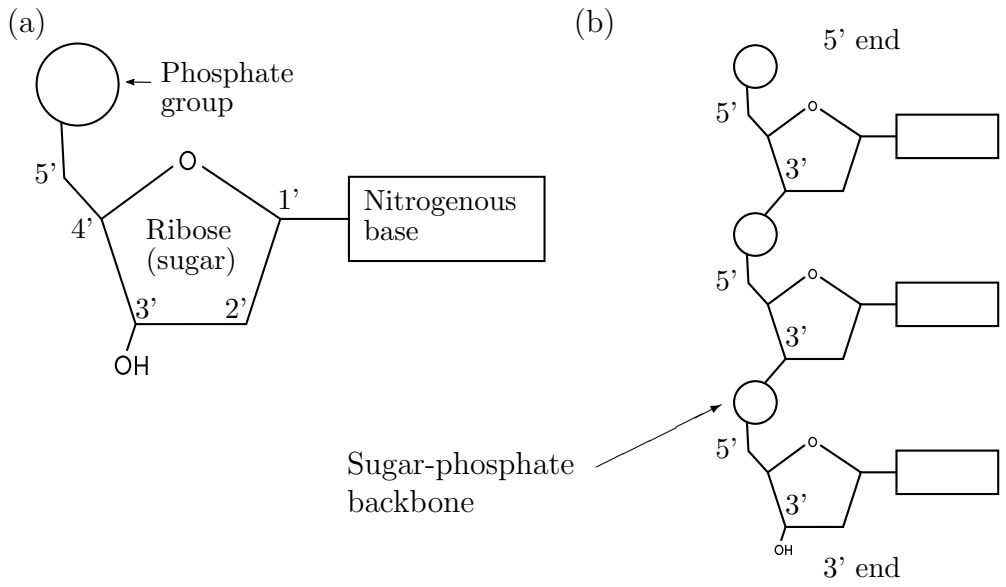
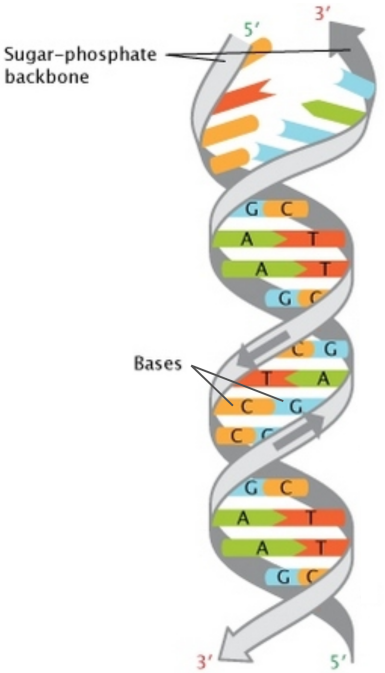


Figure 1.1: Structure of a single nucleotide (a), and a strand of DNA formed from a chain of nucleotides (b). In DNA, the nucleotides are specifically 'deoxyribonucleotides', since the sugar moiety is deoxyribose (the deoxy- prefix refers to the lack of a 2'-hydroxyl group). (a) The carbon atoms of the sugar are numbered as shown. (b) The ends of a DNA strand are named depending on the carbons of the sugar moieties of the terminal nucleotides, as shown.

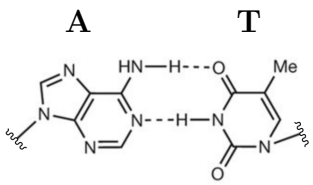
There are four different types of nitrogenous bases which occur in Nature: adenine (A), thymine (T), cytosine (C), and guanine (G) [149]. These define the sequence of a DNA strand. The bases of one DNA strand can form hydrogen bonds with those of another strand ('base-pairing') to form a double-helix structure as shown in Figure 1.2a [149]. This base-pairing (predominantly) follows specific rules: A can only pair with T, and G can only pair with C (known as Watson-Crick base-pairing) [149]. As shown in Figure 1.2b, the A-T interaction consists of two hydrogen bonds, and so is inherently weaker than the G-C interaction (three hydrogen bonds)

[115]. If two DNA strands have the correct sequences of bases to form a series of base pairs (bp) between them, they are said to be 'complementary' (or 'reverse complements' of each other due to the opposite polarities; see below), and the formation of base-pairing interactions between the two strands is known as hybridisation. Breaking these interactions is know as denaturation or melting. DNA strands bind in an anti-parallel fashion: i.e. the two strands have opposite 5'-to-3' directions as shown in Figure 1.2a [149], which also shows how the strands' directions are commonly depicted using arrows. Base-pairing can also occur between complementary regions of a single DNA strand, although this does not commonly occur in Nature.

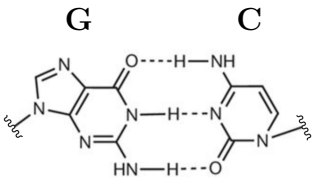
(a) Double helical structure of DNA



(b) Base-pairing



Adenine-thymine base pair



Guanine-cytosine base pair

denotes where base joins sugar-phosphate backbone

Figure 1.2: Base-pairing between two DNA strands results in the formation of a double-helix. (a) The double helix: bases are located on the inside of the helix, with the sugar-phosphate backbones of the two strands on the outside. Strands bind in an anti-parallel manner (i.e. they have opposite 5' to 3' directions). Reproduced with permission from [113]. (b) Detail of base-pairing within the double helix: A pairs with T, and G with C. The G-C interaction involves formation of an extra hydrogen bond compared to A-T, so is stronger. Adapted from [115], reprinted by permission from Springer Nature: Nature Reviews Chemistry. Copyright 2017 Macmillan Publishers Limited.

In aqueous solution (as is the case for both the DNA found in cells and for the

experiments described in this thesis) and in a 'relaxed' state (lacking strain, which can be caused by cellular processes such as transcription), double-helical DNA is found in the 'B-form' which is shown in Figure 1.3a [113]. The structure of B-form DNA was first elucidated by Watson and Crick in 1953 [149]; it is a right-handed double-helix with a diameter of 2 nm [149] and (in solution) ∼10.4 base pairs per turn [146], with each base pair occupying a distance of 0.34 nm in the direction of the helical axis [149]. There is a (larger) major groove and a (smaller) minor groove, which result from the positions at which the bases are joined to the backbone [115]. Although the interaction between the two strands is defined by the hydrogen bonds formed between bases (base-pairing), the stability of the double-helix also comes from stacking interactions between adjacent base pairs [155], as shown in Figure 1.3b. These result from the strong van der Waals interactions between the flat bases. In addition, the formation of a DNA double-helix minimises the exposure of the hydrophobic bases to the aqueous environment, whilst maximising the exposure of the charged phosphate groups of the backbone, which contributes to its stability [155]. Stacking interactions are able to occur at sites where there is a 'nick' (discontinuity) in the DNA backbone between bases [155], which we will see in section 1.2.2.1 is important in origami folding.

1.1.1 DNA as a building material

The Watson-Crick base-pairing described above allows the design of DNA strands which interact with each other through complementary 'domains'. A domain is a section of sequence on one DNA strand which is complementary (and so is able to bind) to a section on a different strand or a different region of the same strand. The high fidelity and predictability of these interactions allow complex DNA nanostructures to be built from large numbers of different DNA strands, where each strand may have multiple domains to allow interaction with a number of other strands [127]. For example, Figure 1.4a shows four DNA strands which each have two domains, allowing them to bind to two of the other strands. The structure formed (a

(a) Atomic model of B-form DNA (b) Stacking between bases

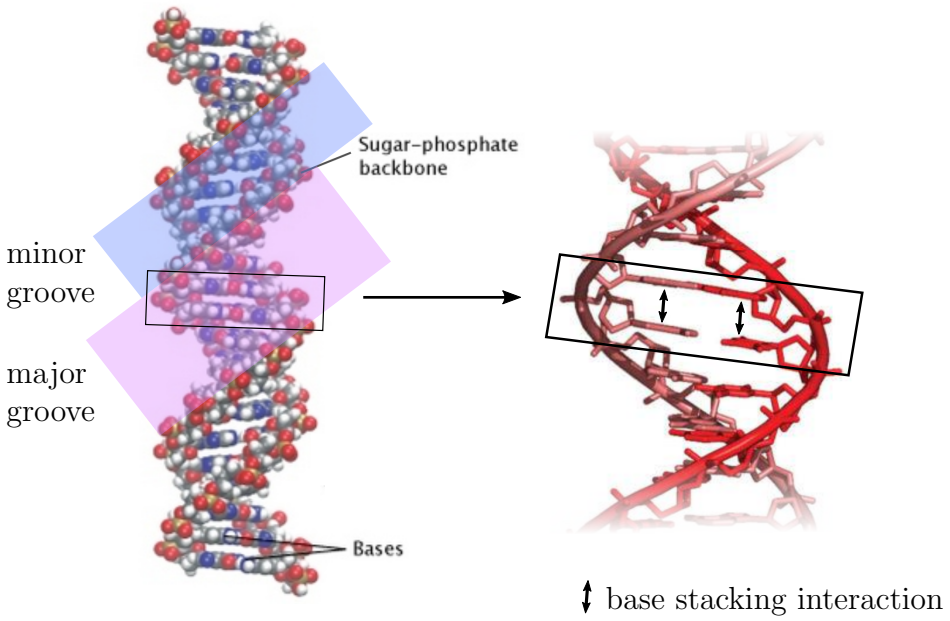


Figure 1.3: 3D structure of B-form DNA. (a) Atomic model showing major and minor grooves. Reproduced with permission from [113]. (b) In addition to base-pairing, base-stacking interactions (shown here by arrows) also contribute to the stability of the double helix [155]. Reproduced from [117] by permission of Oxford University Press.

cross shape) can be described by the domain-domain interactions it contains. The interactions in DNA nanostructures are often described at the domain-level, which is a level of abstraction above describing the individual base-pair interactions.

The final structure formed is determined by the strand sequences and thus interactions between the DNA strands, and so is a 'self-assembly' process. This does not necessarily mean that the same set of strands will always form the same structure; altering the conditions of folding can lead to different products due to the presence of kinetic traps which lead to misfolded structures (see section 1.2). In well-folded structures, the predictability of these interactions means that DNA nanostructures have precise spatial 'addressability' [148], i.e. the positions of each strand within the structure can be inferred from the design, which is important for many of their applications (section 1.3).

Figure 1.4a shows how four DNA strands can be designed to form a cross shape; immobile junction points such as this are used extensively in DNA nanostructures

e.g. in the tile shown in Figure 1.4b [20, 126].

(a) Cross shape formed from four DNA strands (b) Double-crossover tile

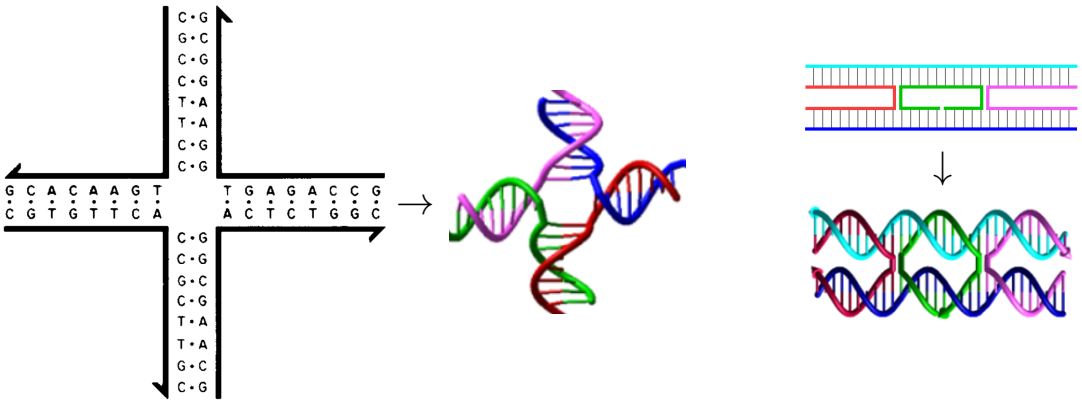


Figure 1.4: Designing DNA strands to self-assemble. (a) Four strands which form a cross shape. The sequences used lead to the formation of an immobile junction point. (b) Two such junctions are combined to form the 'double-crossover' or DX tile. (a) left: reprinted from [126], copyright (1982), with permission from Elsevier. (a) right and (b) bottom: adapted with permission from [161].

Commonly, once the target structure has been designed, the sequences of the strands required to build this shape are determined computationally. This allows minimisation of unwanted interactions between the strands, so that the target structure represents the free energy minimum [131]. The strands can then be artificially synthesised; a relatively low-cost process thanks to recent technological developments. To form the target product, the strands undergo a slow thermal annealing (see Definitions section of this thesis), and the target structure is folded under conditions of thermodynamic equilibrium.

1.1.2 Strand displacement

In addition to base-pairing, DNA has other physical properties which make it amenable for use as a building material for nanostructures. In particular, the ability of DNA to undergo strand displacement is used extensively in dynamic DNA systems (such as DNA motors, sensors, and DNA computing devices). This thesis will predominantly focus on static DNA structures, in which strand displacement is

often important to allow correct folding of the target structure.

Figure 1.5 shows the process of toehold-mediated strand displacement, in which an invader strand is able to displace the incumbent strand by binding to a single-stranded toehold region (blue) on the target strand. The example shown is slow to reverse as the invader-template duplex is more stable than the incumbent-template duplex (due to hybridisation of the toehold region). The process can be made reversible if the incumbent strand is able to bind to a different toehold on the target, in which case the process is known as toehold exchange.



Figure 1.5: Toehold-mediated strand displacement. Initially, the incumbent and target strands are hybridised via a complementary domain. The invader strand is complementary both to the displacement domain (black) of the target and to an adjacent 'toehold' domain (blue), allowing it to displace the incumbent to form a more stable duplex. Displacement is initiated by binding of the invader to the toehold, followed by branch migration. Adapted from [89], reprinted by permission from Springer Nature: Nature Communications. Copyright 2014 Macmillan Publishers Limited.

Blunt-end displacement reactions (i.e. where the toehold length is zero) are also possible but occur more slowly [118, 137, 164]. In more complex systems, where the toehold region is not adjacent to the displacement region, 'remote toehold' displacements can occur [39]. These types of strand displacement reactions will be relevant when discussing folding of DNA origami nanostructures in section 1.2.2.1.

1.2 Folding of DNA nanostructures

As described above, DNA nanostructures are usually designed so that all binding domains have a unique sequence, and so 'off-target' interactions (i.e. interactions which are not present in the target structure) between strands are minimised. Despite this, the occurrence of unwanted interactions is still possible, especially during

folding of complex structures where there are many binding domains. If these interactions are not reversed during the folding process, this is referred to as a 'kinetic trap' or 'kinetic arrest', as they prevent the structure from reaching its free energy minimum (the target structure). In nanostructures with a complex topology, such a trap can also be reached by forming the correct interactions but in the wrong order [73, 95].

Thus, the kinetic accessibility of the target structure must be considered in order to produce a high yield of correctly-folded structures [59]. This involves studying the order in which interactions form during folding (i.e. the 'folding pathway') and using this knowledge to design a system which favours the formation of desired interactions (in the correct order) and minimises unwanted interactions. This has been achieved in different ways and thus there are a range of successful strategies for building DNA nanostructures. These are detailed in the following sections (only static nanostructures are discussed as these are the primary focus of this thesis).

1.2.1 Early DNA nanostructures

Some of the earliest DNA structures were small polyhedra constructed from many short DNA strands, including a cube [21] and a truncated octahedron [166]. In such systems, the yield of well-folded structures is very sensitive to stoichiometry; the probability of off-target interactions is greatly increased when the constituent strands are at different concentrations to one another. To minimise these unwanted interactions, stepwise protocols were used which included purification steps to control the concentration of substructures taken forward to the next stage. Similarly, whilst small 2D tiles (like that shown in Figure 1.4b) could be formed in a one-pot reaction, assembling these into arrays again required multiple stages [110]. Unfortunately use of a multistep procedure invariably limits the yield. More recently, simple 3D structures [42, 43] and repetitive 2D and 3D crystals [86, 169] have been successfully produced in one-pot reactions (the latter through careful purification and control of the strand concentrations). However, the requirement for perfect

stoichiometry limits the complexity of structures which can be formed in this way [122].

1.2.2 DNA origami

An alternative to using many short DNA strands is to use one long strand (a 'scaffold' or template strand) which is folded through interactions with shorter strands. This method was used to create tile arrays [156] and an octahedron [132] (the latter also extensively relied on hybridisation between different regions of the scaffold strand) and is the basis of the DNA origami technique developed by Rothemund [122]. In DNA origami, a circular single-stranded scaffold or template strand (traditionally M13mp18, a commercially-available bacteriophage genome) is folded through its interactions with many short strands known as staples (Figure 1.6). Staples bind to two or more different domains on the template, thus bringing them into close proximity. The final structure is solely determined by these template:staple interactions.

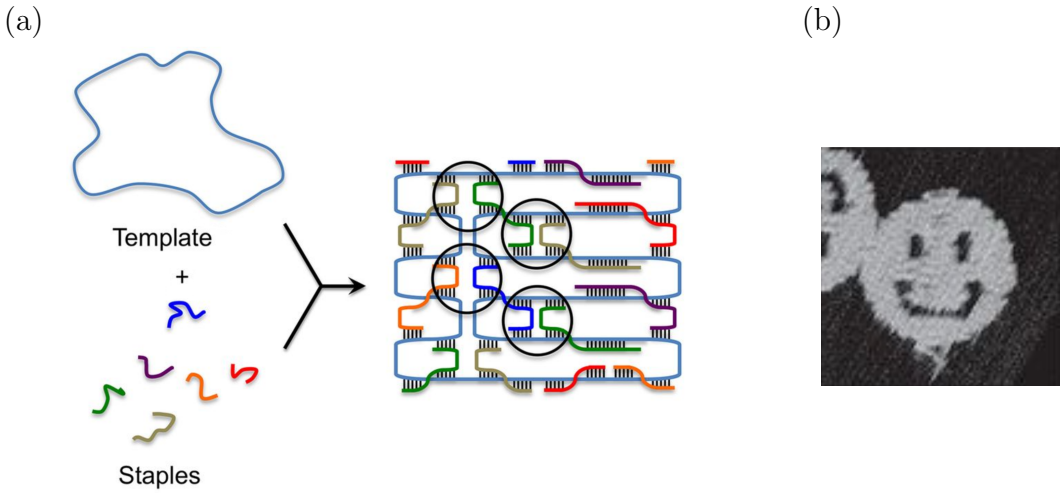


Figure 1.6: DNA origami. (a) Schematic of DNA origami folding; circles show sections of the template which have been brought into close proximity through staple binding. Traditionally, the template strand is the 7249 nt M13mp18 viral genome, which is folded into various shapes using hundreds of unique staple strands (staples are not usually reused within the same structure). Adapted with permission from [53]. (b) AFM image of a DNA structure in the shape of a smiley face. Adapted from [122], reprinted by permission from Springer Nature: Nature. Copyright 2006 Nature Publishing Group.

For discrete 2D structures, DNA origami is a reliable and robust technique, able to produce a high yield of well-folded structures in a one-pot reaction. (Many 3D structures have also been produced using DNA origami [5, 25, 30, 69, 76] although they can be more challenging to assemble [168].) Thus, it has become one of the most widely used approaches for building DNA nanostructures [56]. The following section explains the success of this method in terms of the folding pathway.

1.2.2.1 DNA origami folding

One of the principles underlying the success of the DNA origami method is the co-operativity of staple incorporation, which arises from using a single template strand. Incorporation of each staple reduces this entropy, bringing the template conformation closer to that of the final structure, thus a) precluding a large number of misfolded structures and b) making subsequent staple incorporation steps more favourable by reducing their entropic cost and bringing their template binding domains closer together [122]. In addition, binding of a staple provides stacking sites which favour further staple binding. The result of this high degree of co-operativity is that all staples are incorporated over a very narrow range of temperatures [135], regardless of a wider spread in their independent thermal stabilities (i.e. the melting temperatures of each domain in a short duplex with its complement). This co-operativity also results in significant hysteresis in folding and unfolding temperatures [59].

In addition, the use of a scaffold strand negates the need for perfect stoichiometry; the only requirement is that staple strands are in excess of the template (staple strands do not interact with each other and so their relative concentrations do not matter). The use of a staple excess is also thought to allow reversal of off-target interactions (e.g. where staples have bound to the wrong domain on the template strand) through strand invasion and displacement by free staple strands [122]. Both these features of DNA origami also contribute to its success. DNA origami folding is discussed in further detail in section 2.4.6.

Other methods which make use of a long template strand (and thus owe their success to the same factors as DNA origami) include grid-iron structures [49] and polyhedral meshes [13]. In addition, single-stranded origami structures have also been shown to fold successfully (although these require longer annealing protocols than standard origami) [50].

1.2.3 Single-stranded tiles

Previously, we have seen that methods which use a large number of short DNA strands (the early DNA polyhedra) have strict stoichiometric requirements which limit the size and complexity of structure which can be produced in a one-pot reaction. More recently this dogma has been overturned by the development of a reliable method for self-assembly of a large number (>300) of distinct single-stranded tiles (SSTs) in a one-pot reaction, which is robust to unpurified strands [71, 150]. These form a 2D array as shown in Figure 1.7a; adjusting the domain lengths allows formation of a 3D structure as shown in Figure 1.7b.

Simulations by Jacobs et al. have suggested that the success of this technique is due to the experimental protocol used [60]. The entropy of each SST is thought to decrease dramatically upon incorporation, which results in a significant 'nucleation barrier', i.e. binding of the first few SSTs to each other is not a very favourable process [150]. Simulations showed that the temperatures used at the start of assembly (90°C for 2D structures [150]; 80°C for 3D structures [71]) are above the window where a) the nucleation barrier is surmountable, but b) it is still significant enough to limit off-target interactions [60]. Thus, when the temperature is slowly reduced, nucleation occurs in a slow and sparse manner which suppresses competition between multiple nucleation sites and prevents aggregation of partially formed structures [60]. Monomer depletion is also not an issue due to slow nucleation, hence the stoichiometry of strands does not need to be optimised.

At these higher temperatures, partially formed structures may be more stable than the complete structure. This is a departure from classical nucleation theory

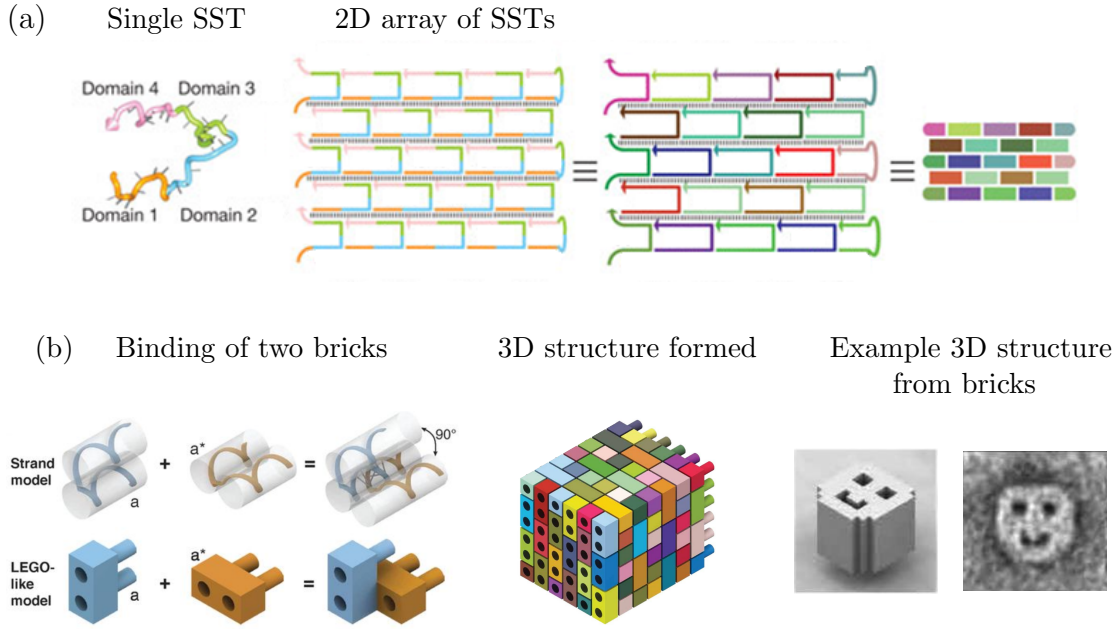


Figure 1.7: 2D and 3D structures formed from single-stranded tiles (SSTs) and bricks, respectively. (a) Formation of a 2D array from SSTs: each SST has four domains through which it can bind to four other SSTs. Adapted from [150], reprinted by permission from Springer Nature: Nature. Copyright 2012 Macmillan Publishers Limited. (b) 3D structures formed from single-stranded bricks: bricks are identical to SSTs except the length of domains results in a 90° angle between bound bricks. Right-hand two images show the design (left) and transmission electron microscope (TEM) image (right) of a 3D smiley-face structure. From [71]; reprinted with permission from AAAS.

(CNT) where, once the 'critically-sized nucleus' has been formed (i.e. the nucleation barrier has been ascended), incorporating further monomers always increases the stability of the complex. This deviation from CNT is true of all addressable, finite-sized structures [60]. Cooling is therefore required to complete formation. At lower temperatures, rapid growth occurs so that many structures are completely formed by the time strand depletion becomes an issue [71].

1.3 Applications of DNA nanostructures

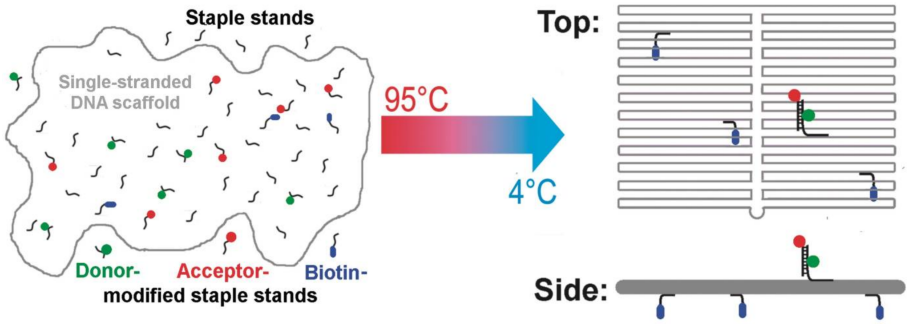
Many of the applications of DNA nanostructures make use of a) their addressability, allowing precise positioning of attached particles; and b) the interaction of the constituent DNA with other biological materials such as proteins and mRNA[1].

In terms of the former property, DNA origami has been extensively used to create nanoarrays through conjugation or aptamer binding of both organic (e.g. proteins) and inorganic (e.g. metallic nanoparticles) molecules to specific staples (Figure 1.8a). This allows placement of the particles with ∼6 nm resolution [77]. Protein arrays on origami have been used to probe protein-protein interactions [119], to build enzyme cascades [38, 47, 67, 105, 170], and to allow characterisation and imaging of proteins [29, 97]; whilst arrays of gold and other metallic nanoparticles on origami have nanophotonic and nanoelectronic applications [128]. In addition, arrays of fluorophores on origami have been used in super-resolution microscopy studies based on DNA-PAINT technology [62]. Precise positioning has also allowed DNA origami to be used as a template e.g. for the construction of liposome-type structures from lipids [26, 159] (Figure 1.8b), and to control crystal growth e.g. of gold [140].

Because of their interactions with other nucleic acids (through base-pairing or aptamer binding), DNA nanostructures have many potential biomedical applications. These include as detectors of gene expression (via interaction with mRNA)

[1]Some of these interactions mimic those found in Nature (e.g. base-pairing of complementary DNA/RNA) and some are through DNA aptamers: artificially selected oligonucleotides that have high affinity to specific targets such as proteins [154].

(a)



(b)

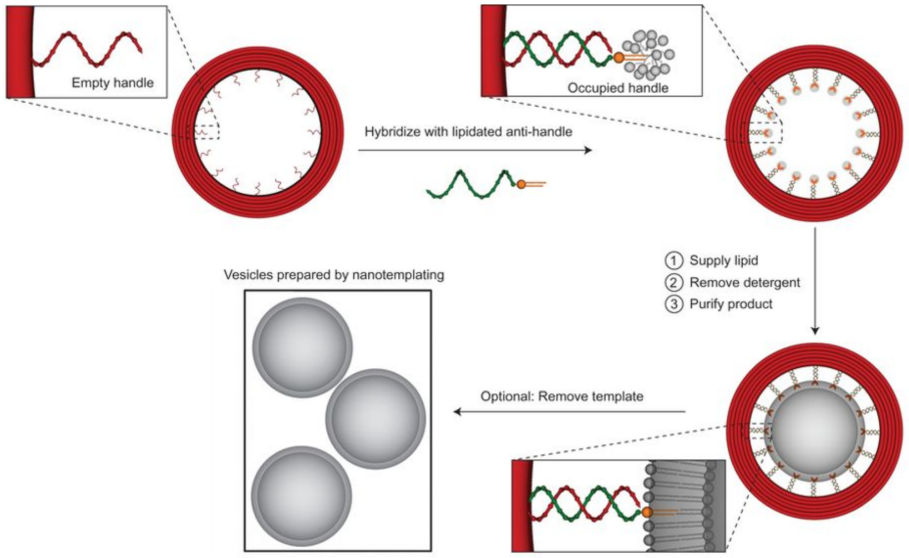


Figure 1.8: Applications of DNA origami which use its ability to precisely position other materials. (a) Functionalisation of DNA origami through conjugation of particles to staple strands. A 2D origami tile has been functionalised with donor and acceptor fluorophores on one side, and biotin on the other. Biotin allows immobilisation of the origami on a streptavidin-coated surface for total internal reflection fluorescence (TIRF) microscopy studies using the fluorophores. Reproduced from [41] by permission of Oxford University Press. (b) DNA origami (red) as a template for size-controlled liposome (vesicle) formation. Adapted from [159], reprinted by permission from Springer Nature: Nature Chemistry. Copyright 2016 Macmillan Publishers Limited.

[70], and as genotyping tools through the detection of SNPs (single-nucleotide polymorphisms; via interaction with biological DNA) [138, 165]. In addition, DNA nanostructures may be suitable for in vivo diagnostic and therapeutic uses, particularly drug delivery, which could be targeted to specific cells through the interaction of the nanostructure with cell-specific mRNAs or proteins [154, 167]. The interactions of DNA nanostructures with lipid bilayers has allowed the construction of membrane protein analogues built from DNA [22], including nanopores [9, 15, 44] and cell adhesion molecules [1]. Genetically-encoding DNA nanostructures in living cells to allow their expression in vivo may lead to the development of new research tools and biomedical applications [36]. The use of DNA nanostructures in vivo is a rapidly-developing field, but major hurdles such as ensuring cellular uptake, preventing degradation by nucleases, and minimising immunogenicity have yet to be fully overcome [14, 147].

1.4 Future perspectives and challenges

As described above, DNA nanostructures (and in particular DNA origami) have vast potential for use in a number of different fields. However, there remain obstacles which currently prevent their widespread use in industry. These include:

1. Cost: the materials for each origami design (template and staple strands) cost between several hundred and a thousand US dollars to acquire [147, 154].

2. Size: for some applications (including nanoelectronics), individual nanostructures need to be larger than the standard origami shapes produced from a single M13mp18 scaffold [128].

3. Responsive/intelligent structures. In Nature, proteins (and to some extent RNA) perform most of the cellular functions. In contrast, DNA mainly exists as a static storage material. These differences result from differences in the interactions which govern folding of these macromolecules. As a result,

DNA nanostructures (folded by high-fidelity base-pairing interactions) are also mainly static. Engineering DNA structures to have increased functionality would expand their potential applications both in vivo (as probes or molecular machines [154]) and elsewhere. This requires designing reconfigurable structures which change in response to specific stimuli.

4. Design optimisation and folding conditions: low yields of well-folded structures are still an issue for many designs [154]. In addition, most structures are assembled by thermal annealing (and can take days for complex 3D objects [61]), which is unsuitable for co-assembly with thermosensitive components such as proteins and cannot be replicated for folding in vivo.

Progress has already been made in all four of these areas. The cost of origami structures (point 1 above) has been addressed by various groups, with strategies developed for both cheaper scaffold and staple production [32, 72, 112] and reuse of staples within a structure [106]. In terms of size (point 2 above), since origami templates larger than M13 are difficult to acquire and also mechanically fragile [147], 'supramolecular' assemblies have been constructed by joining together individual origamis through sticky-end cohesion and/or base-stacking interactions [57, 87, 142] to form both 2D and 3D shapes. The complexity of these larger structures has also increased, with twisted and curved 3D origamis demonstrated by the Dietz group [25].

With this expansion in the size and complexity of structures available, the current focus of much DNA nanostructure research is reconfigurable structures as described in point 3 above [145]. Considerable work has already been done using strand displacement (see section 1.1.2) to achieve this [147], including the design of large origami structures which couple strand displacement in a particular location to dramatic changes in the overall structure [96, 171]. However, more complex systems, and those which are responsive to stimuli other than nucleic acids, are desirable. This can be accomplished by the use of environmentally-responsive motifs such as G-quadruplexes, which form in the presence of specific metal ions and low pH

[8, 16, 66, 107, 157]; i-motifs (Figure 1.9a; which also require acidic conditions to form) [27, 51, 91, 130, 141, 153]; DNA triplex formation through Hoogsteen base-pairing (Figure 1.9b; again at low pH) [4, 45]; aptamer binding (Figure 1.9c) [28, 120]; ion/temperature sensitive base stacking (Figure 1.9d) [40]; and the use of light-responsive azobenzene moieties intercalated between base pairs (Figure 1.9e) [7, 65, 78, 158]. The use of non-canonical base pairs (such as metal-ion-bridged duplexes [101, 111]) can further increase the range of stimuli to which a DNA nanostructure is able to respond to.

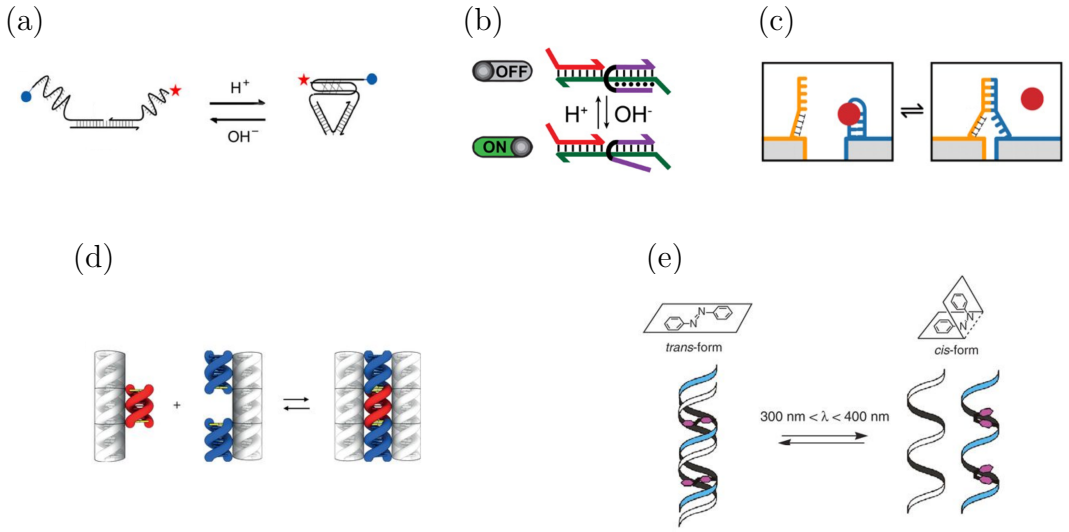


Figure 1.9: Environmentally-responsive motifs in DNA nanostructures. (a) pH-responsive i-motif, b) pH-responsive triplex formation, (c) aptamer binding, (d) ion/temperature sensitive base-stacking, (e) light-responsive azobenzene. (a) Adapted from [141], reprinted by permission from Springer Nature: Nature Communications. Copyright 2011 Macmillan Publishers Limited. (b) Reproduced from [4] with permission. (c) From [28]. Reprinted with permission from AAAS. (d) From [40]. Reprinted with permission from AAAS. (e) Adapted from [7], reprinted by permission from Springer Nature: Nature Protocols. Copyright 2007 Nature Publishing Group.

As we can see, much progress has been made in increasing the range and complexity of DNA nanostructures and their functionalities. However, the fast and robust synthesis of these structures remains an issue (especially as they become more complex) as described in point 4 above. The development of fast and isothermal assembly methods would allow co-assembly with thermosensitive components and may also enable folding in vivo. For DNA origami, various isothermal folding meth-

ods have been developed (section 2.4.6). In Chapter 2 of this thesis, I build on this work by presenting a novel folding method for an origami tile using 'helper' strands, which increases the yield of well-folded tiles obtained from isothermal folding.

Chapters 3 and 4 of this thesis detail work characterising and then using DNA 'T-motifs' (a type of junction through which two DNA duplexes can join) to design and build an 'origami-like' structure. Since the interactions which govern folding of this T-motif origami structure are weaker than those in traditional origami, this allows the structure to reconfigure at room temperature. This has implications both for the folding protocol of such structures (they can be folded quickly at low temperatures, as in point 4 above) and for their responsiveness/functionality (as in point 3 above).

Chapter 2

Folding of a DNA Origami Nanostructure using 'Helper' Strands

In this chapter, I present the results of a study into the folding of a DNA origami tile (section 2.1) using a set of 'helper' strands (section 2.2). These strands were designed to bind to the origami template and assist with folding (i.e. staple binding), but are not present in the final structure. Helper strands are a novel idea developed by myself and Dr Jon Bath; as far as we know no similar devices or methods have been used before to aid DNA nanostructure folding.

As described in sections 1.1.1 and 1.4, most DNA origami structures are assembled by thermal annealing, which is not always suitable e.g. for folding in vivo or co-assembly with thermosensitive components. The aim of using the helper strands is to allow fast and isothermal folding, which could broaden the range of applications for which origami can be used.

The helper strands were designed to coat the template strand prior to the addition of the tile staples, leaving only parts of the template single-stranded (Figure 2.1a). Staples can bind to these single-stranded regions (Figure 2.1b) and then displace the helpers via toehold-mediated strand displacement (see section 1.1.2), to

eventually form the fully-folded tile. Staple incorporation is thus nucleated at the single-stranded region of the template, and propagates out from this point.

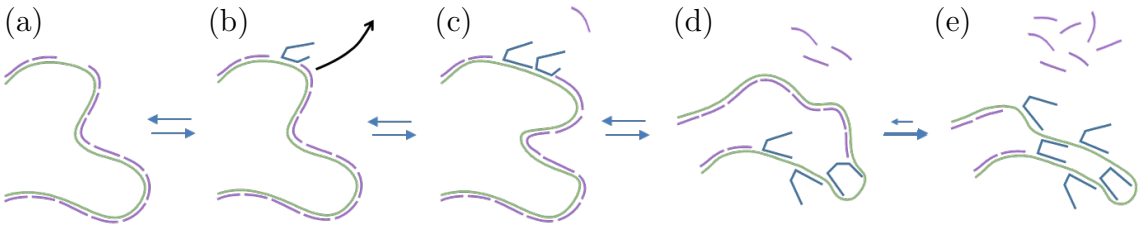


Figure 2.1: Origami folding in the 'helper-origami' system. (a) Origami template strand (green) is partially coated with helper strands (purple); (b) a staple domain is able to bind to the single-stranded region of the template; (c)-(e) strand displacement of helper strands by staples creates toeholds that allow binding of further staple domains. Binding of the second domains of staples stabilises the structure.

In addition to allowing fast isothermal folding, the system was designed to allow folding of a single origami structure from a mixture of competing staples. In such experiments, more than one staple set (i.e. corresponding to more than one origami structure) is added to a single template. When these strands are thermally annealed, the time required to reach thermodynamic equilibrium is considerably longer than for a single staple set [92]. Priming the template with helpers to favour initial binding (i.e. nucleation) and/or propagation with staples from only one staple set may reduce this time. Furthermore, it may be possible to produce two sets of templates, differing only in the positions of bound helper strands, which could allow rapid folding of two distinct structures from two different subsets of staples in a one-pot reaction.

In this chapter, I will first introduce the design of the origami tile (section 2.1) and helper strands (section 2.2), followed by the experimental methods used to study the helper-origami system (section 2.3). The results are divided into two sections: 2.4 and 2.5. Section 2.4 looks at if and how the use of helper strands improves isothermal folding of the origami tile. For these experiments, the template used was fully-coated with helper strands (i.e. no regions were left single-stranded). This is because initial experiments showed that leaving a single-stranded region (as per Figure 2.1a) produced identical results to using a fully-coated template (see Figure

2.34 and discussion in the conclusions section 2.6). Section 2.5 focuses on developing the system for use in competitive staple experiments (as described above). The findings of this chapter are summarised in section 2.6.

2.1 Tile design

The origami tile used in this chapter was designed by a former member of the Turberfield group [33]. It consists of a 2646 nt template strand derived from a double-stranded plasmid (pKD1; see Appendix A.2.2), and a staple set of 90 different strands which can bind to one or two domains on the template (see Figure 2.2a). Two-domain staples are used (in the main), as opposed to the three- or four-domain staples of conventional origami ([122]; see section 2.4.5), as this makes studying folding more straightforward. Each binding domain is 15 or 16 nts (using binding domains of the same length and therefore approximately the same melting temperature also simplifies matters) which, when bound by the complementary staple domain, corresponds to 1.5 turns of the double-helix[1]. This results in the (correctly) folded tile being a planar rectangle (shown in Figure 2.2).

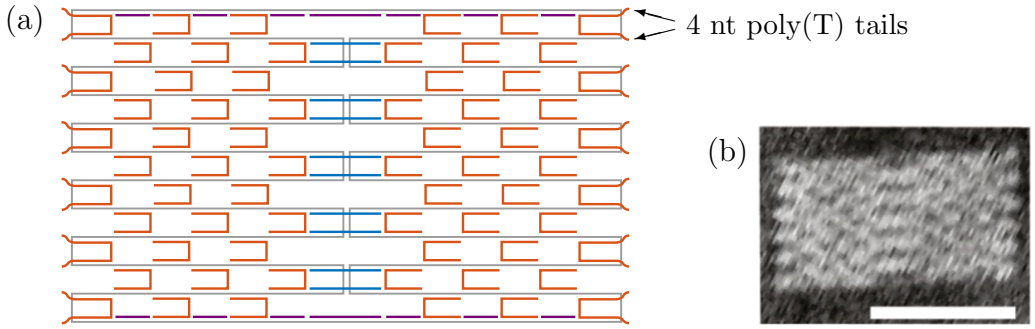


Figure 2.2: Origami tile. (a) shows the tile design, with the template strand in grey and staples in blue, purple and orange. Four-nucleotide poly(T) staple overhangs on the left and right prevent stacking between tiles. NB: all of the vertical lines in this diagram and subsequent diagrams of the tile are 0 nt in length; they are shown to make visualisation of the design easier. (b) AFM image of the folded tile (scale bar = 50 nm). (b) reprinted by permission from Springer Nature: Nature [34]. Copyright 2015 Macmillan Publishers Limited.

In common with conventional origami design (see section 1.2.2) each of the 168 binding domains of the template has a unique sequence, so that each staple domain can only bind stably to its complementary domain on the template; and the fully-

[1]Double-stranded DNA free in solution is most stable at ∼10.5 bps per turn [82], which corresponds to 15.75 bps for 1.5 turns. In the tile, every fourth domain is 15 bases (the rest are 16), which makes the average over four consecutive domains correct.

folded tile (i.e. with each binding domain hybridised to the correct staple domain as in Figure 2.2a) is the most thermodynamically stable structure possible. Thermal annealing of the template with a 2× excess of each staple strand (100 nM template: 200 nM staples) produces a high yield of the well-folded tile (as shown in Figure 2.2b).

The 90 staples of the tile can be classified as follows:

- 66 body staples (orange in Figure 2.2a) which bind to two distant (i.e. non-contiguous) domains on the template. The 12 body staples which border the short edges of the tile have a four-nucleotide poly(T) tail at both their 5' and 3' ends to prevent blunt-end stacking between tiles.

- 5 pairs of seam staples (blue in Figure 2.2a) which form the longest-range interactions (in terms of the primary sequence of the template) in the tile. They link positions where the template folds back on itself, forming a central seam as seen in other origami structures [122].

- 14 edge staples (12 one-domain and 2 two-domain; purple in Figure 2.2a) which fill in the parts of the long edges of the tile unbound by other staples.

This tile has previously been used to study origami folding during thermal annealing [33, 34]. The results showed that seam staples are incorporated at similar temperatures to body staples, despite the entropic cost of forming such long-range interactions. This is thought to be due to the high level of co-operativity between the two seam staples in each pair [34]. To test this, Dunn et al. designed a dimer version of the template (two copies joined head-to-tail) which has a number of different 'well-folded' states as shown in Figure 2.3. Changes in the distribution of these well-folded products were used to observe the effects of strengthening and weakening different interactions within the tile (i.e. making them form earlier or later in the folding pathway). The results showed that weakening the seam staple (long-range) interactions allowed the short-range (body staple-mediated) interactions to dominate the outcome, presumably because they form earlier in the folding pathway

than the now-weakened seam staple interactions. This favours the formation of the structure in which all body staples are bound to two domains that are as close as possible in terms of the template sequence (structure '**a**' in Figure 2.3).

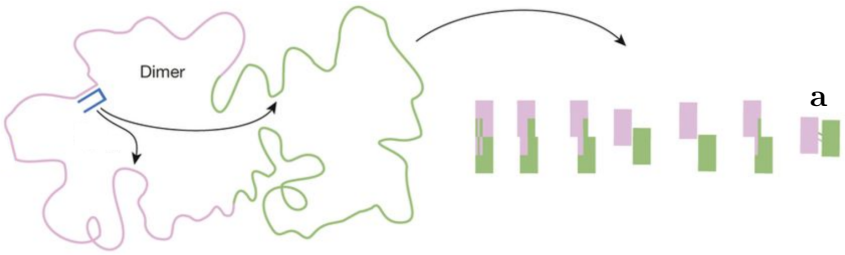


Figure 2.3: Experiments carried out by Dunn et al. using a dimer version of the tile shown in Figure 2.2. The template (left) is two copies of the monomer template sequence joined head-to-tail. This allows each two-domain staple to bind in one of two different configurations, as shown. Folding by thermal annealing produced a range of well-folded structures (right). Weakening the seam staples (by breaking one of each pair) resulted in an increase in the proportion of structure '**a**' produced. Reprinted by permission from Springer Nature: Nature [34]. Copyright 2015 Macmillan Publishers Limited.

In the helper-origami system, the tile will be folded in isothermal conditions. It is not clear how the relative binding strengths of different staples (or the distance range of the interactions which they mediate) will affect the order in which they are incorporated in this system.

2.2 Helper strand design

Each helper strand (or 'helper' as they will be referred to throughout this chapter) is the same length as a single binding domain of the tile staples (15/16 nt), but they bind to the template at an offset to the staple binding domains: either 10 nt (Figure 2.4a) or 8 nt (Figure 2.4b). The 10-offset helpers have an asymmetric offset to the staple domains, whereas the 8-offset helpers have a symmetric offset (except where the domain length is only 15 nt).

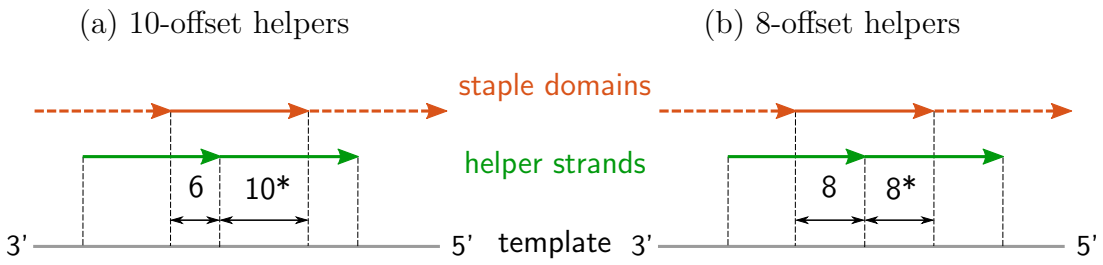


Figure 2.4: Diagram showing relative offsets of staples and helper strands in terms of their binding domains on the template strand. Each staple domain/helper strand is 15 or 16 nt. Offsets are shown in nucleotides (nt); *denotes an offset which is 1 nt shorter when the staple domain is only 15 nt.

For each of the two offsets shown above, there is a set of 168 helper strands which hybridise fully (i.e. leaving no part single-stranded) with the template strand. Figure 2.5 shows how the helper binding domains correspond to the staple binding domains in the full structure.

The system we are using relies on the origami tile being more thermodynamically stable than the helper-coated template at the temperatures used. Thermal annealing of the helper-coated template with the staples produces the folded tile (see Figure 2.18), which shows that the tile is more stable at least at the temperatures at which staple binding/rearrangement is occurring. Both structures share the same base-pair interactions and base-stacking, but differ due to the cross-linking effect of the staples. This actually reduces the stability of the origami structure in many ways: the conformational entropy of the bound template is reduced, and there is increased electrostatic repulsion due to the close packing of helices (in addition,

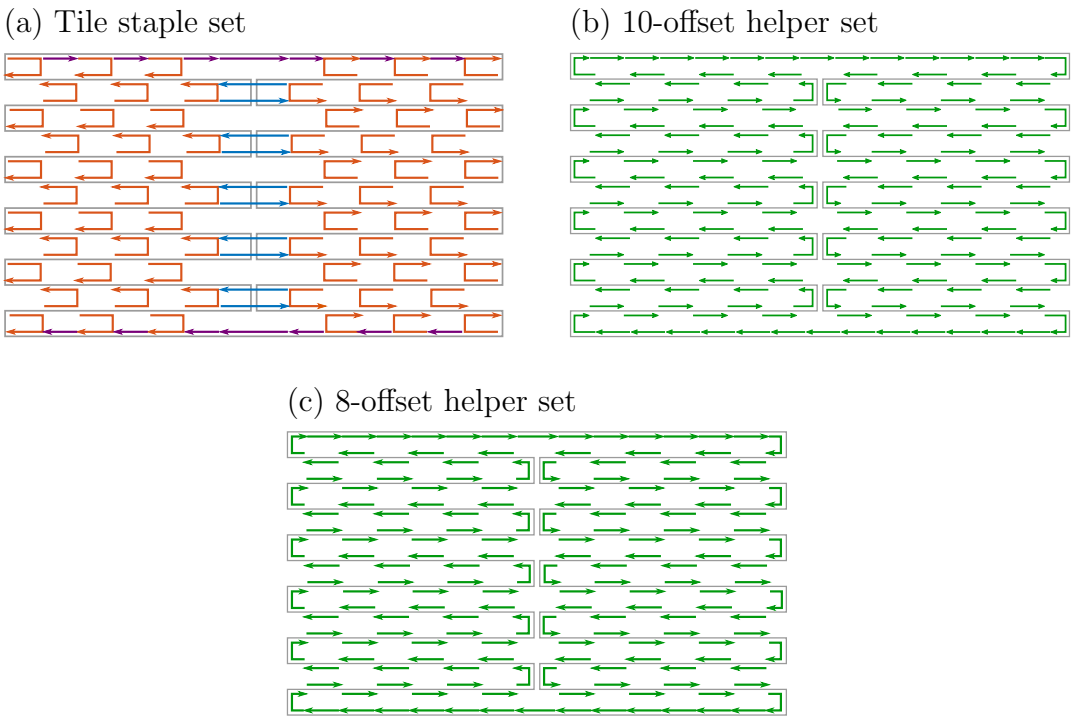


Figure 2.5: Diagrams of template binding positions of (a) the staple set, (b) the 10-offset helper set and (c) the 8-offset helper set. Arrows denote 5'-3' direction. As previously, all of the vertical lines in these diagrams, and subsequent diagrams of the tile and helper-coated template, are 0 nt in length; they are shown to make visualisation of the design easier.NB: the template is shown in the same conformation in each diagram to allow comparison, but in reality neither helper set provides connectivity between different parts of the template, so the conformation for (b)/(c) is likely to be a floppy circle similar to a nicked double-stranded plasmid.

there is a change in free energy due to changes in hydration and counter-ion binding [88]). However, this is more than compensated for by the favourable entropy gain resulting from replacing a large number (168) of bound helper strands with a smaller number (90) of bound staples. We expect that using longer (and therefore fewer) helpers would increase the stability of the helper-coated template relative to the tile. Using fewer than 90 helpers may result in the helper-coated template being more favourable than the tile; this has not been tested. (The double-stranded version of a template (in effect the use of the single, circular helper) has been shown to be more stable than an origami structure formed from one strand of the template with its staple set [54].)

Unless otherwise stated, the 10-offset helper set has been used. The only figure

corresponding to the 8-offset helper set is Figure 2.42.

2.3 Experimental design

In order to create the 'helper-coated template', template strand was thermally annealed with a 2× excess of helper strands. The tile staples (again, at a 2× excess to the template strand) were then added and the mixture incubated at a constant temperature for different periods of time (see Appendix A.2.3 for more details). The folding of the origami tile during the incubation stage was studied using three experimental techniques: AFM, gel electrophoresis and real-time fluorescence measurements. The first two techniques allowed evaluation of the end products when the conditions under which they were folded were changed (the temperature of incubation, the length of incubation time, and/or whether any of the helper-coated template was left single-stranded), while the latter allowed us to directly follow folding in real-time. In addition to this difference, these techniques differ in what structural information they are able to provide; this is summarised in Table 2.1 and detailed in the following sections.

Method	Results from method	Limits of method
AFM	• Provides direct information about the structure of the tile in a specific sample.	• Time-consuming and expensive so cannot be used for a large number of samples. • Possible sampling bias due to differences in ability to adhere to the mica of well-folded vs very poorly-folded tiles. • Can only be used for a specific time-point, not continuous measurement*.
Gel electrophoresis	• Up to 20 different samples, each representing a specific time/temperature point, each containing up to three different fluorescently-labelled strands, can be analysed at once. • Different structures present in a single sample can be observed (providing they have different electrophoretic mobilities).	• Difficult to extract a quantitative measurement to compare different folding conditions, as cannot measure a continuous range of time-points/temperatures. • A direct imaging technique (e.g. AFM) is required to confirm correct interpretation of the results. • Potential introduction of artifacts (e.g. structural changes) during running of the gels.
Real-time fluorescence measurements	• Parallel processing of many samples. • Provides quantitative data to easily compare different samples.	• Requires a direct imaging technique to confirm results (as for gels). • Depending on position of labelled strands, may only be providing information on local structural changes. • Different intermediates present in the same sample cannot be distinguished.

Table 2.1: Experimental methods used in this chapter, and what data they are able to yield about tile folding within the helper-origami system. *see reference to Song et al. [136] in section 2.3.1.

2.3.1 Atomic force microscopy (AFM)

Direct imaging of origami constructs is the most conclusive method to ascertain their structure, and (tapping-mode) AFM in liquid is the standard imaging technique of choice for 2D origami (transmission electron microscopy or TEM is commonly used for 3D structures). As we can see in Figure 2.6, which shows an AFM image of the origami tile which I am using, AFM allows us to see individual tiles and their specific defects (see Appendix A.6 for full experimental details).

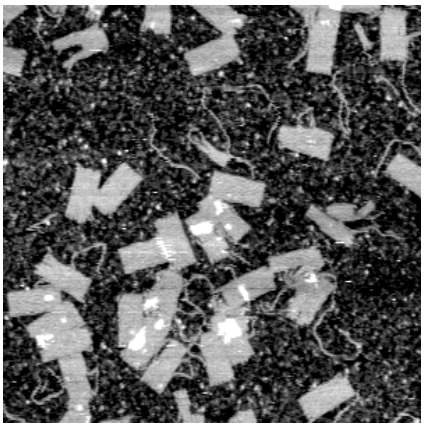


Figure 2.6: AFM image of a tile sample. Image area is 0.6 × 0.6 μm. (See Figure 2.17b for details of sample preparation.)

AFM has been used extensively to assess the products of different folding regimes, for example various isothermal methods [61, 74, 168] and competitive annealing experiments [92]. In addition, the intermediates of folding have been imaged, for example by Song et al. who used a heated-stage AFM to directly observe thermal annealing and melting of an origami rectangle [136]. Lee Tin Wah et al. also studied thermal annealing of this rectangle, using the discovery that origami folding is quenched by immobilisation on the mica surface used in AFM imaging [80]. Finally, chemical quenching using formaldehyde was used by Majikes et al. for the same rectangle, and the results correlated well with those from Lee Tin Wah et al. [93]. Although these methods are perhaps the only way of directly observing folding intermediates on a single-molecule scale, immobilising unstable intermediates on the mica surface may disrupt their native structure and thus be an inaccurate

way to determine the folding pathway of origami structures in solution. As a result (and also due to the time-consuming nature of AFM imaging), I have used AFM to characterise the end products of folding only.

2.3.1.1 Analysis of origami AFM images

In order to extract quantitative information from the AFM images, I have analysed them using Gwyddion software [103]. The process is detailed below and illustrated using the three example images shown in Figure 2.7: one of relatively well-folded tiles (Figure 2.7a), and two of poorly-folded tiles produced by different methods (Figure 2.7b). Poorly-folded origami tiles produced by incubating template and staples without helpers (Figure 2.7bi) tend to be misfolded in such a way that parts of the structure protrude upwards, producing 'double-height' regions of the tile, shown in white. This results in smaller mean areas compared to well-folded tiles. (It also results in higher mean heights, but measurement of this would require normalising the image heights to account for differences in the thickness of mica used etc., which could not be done accurately for these images due to the the presence of excess staples/helper strands in the image background). Poorly-folded tiles produced by incubating helper-coated template with staples (Figure 2.7bii) arise when displacement of the helpers by the staples is incomplete, producing tiles with lengths of template-helper duplex which appear as 'string-like' structures in the AFM images. This results in higher boundary lengths compared to well-folded tiles. The analysis process below was used to measure the mean areas and boundary lengths of tiles, and thus identify both types of poorly-folded tile.

For each sample, a total of four images (each $1 \times 1\mu$m as in Figure 2.7) were analysed. NB: the images shown in Figure 2.7 have been processed as per Appendix A.6.1 to make the tiles visible, which includes a 'line levelling' step. This involves calculating the median pixel intensity of each horizontal line and its neighbour, and adjusting the pixel intensities so that they are equal. This corrects for differences between line heights which can arise due to mechanical disturbances or the interac-

tion of the cantilever with samples. The images used for the quantitative analysis detailed below were unprocessed except for the correction of horizontal scars, as line levelling caused parts of tiles to 'disappear' into the background (an example is shown in Appendix A.6.2). See Appendices A.6.1 and A.6.2 for details.

(a) Well-folded tiles (b) Poorly-folded tiles produced:
i) without helpers ii) using helpers

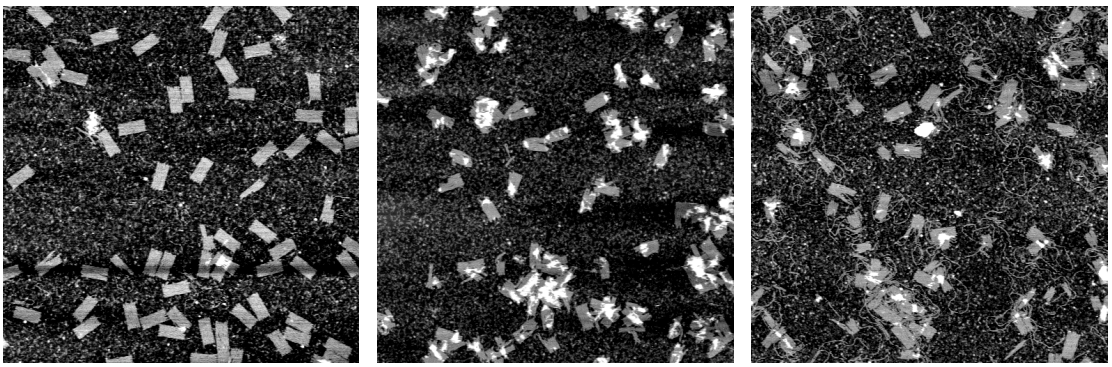


Figure 2.7: Example AFM images which will be used to illustrate the analysis process. (a) (Relatively) well-folded tiles produced by incubating helper-coated template and staples at 55°C for 4 hours. (bi) Poorly-folded tiles produced by incubating template and staples at 60°C for 1 hour. Note the appearance of 'double-height' regions of the tile, shown in white. (bii) Poorly-folded tiles produced by incubating helper-coated template and staples at 55°C for 1 hour. Note the presence of extensive 'string-like' structures. The images have been processed as per Appendix A.6.1 so that the tiles are easily visible, whereas the analysis described below was carried out on raw images (see main text). Images are 1×1 μm.

Origami AFM analysis process:

1. Individual particles in the image (henceforth referred to as 'grains') were identified using the 'Mark by Otsu's' command (see [108]) as shown in Figure 2.8.

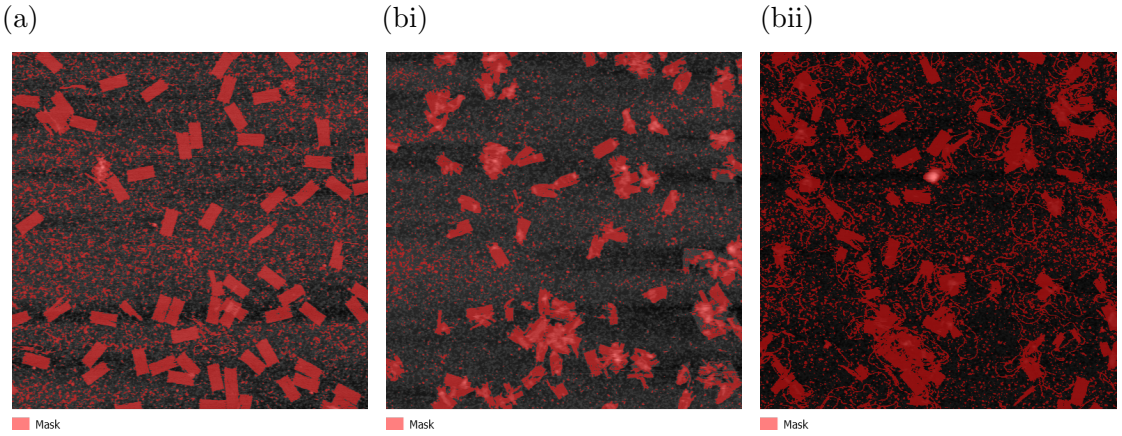


Figure 2.8: AFM images from Figure 2.7 following grain detection using the 'Mark by Otsu's' command in Gwyddion. Detected grains are in red.

2. The marked grains were filtered to remove those which were too small or too large to correspond to a single tile (see Appendix 2.3.1 for details) as shown in Figure 2.9. These include excess staples/helper strands, non-DNA particles which are presumably salt (e.g. the rounded particle seen in sample (bii)) and grains formed by two tiles touching (which was unfortunately a high proportion of the grains; with more time, repeats would have been carried out with the samples at lower concentrations to prevent this). Grains overlapping with the image edge were also removed.

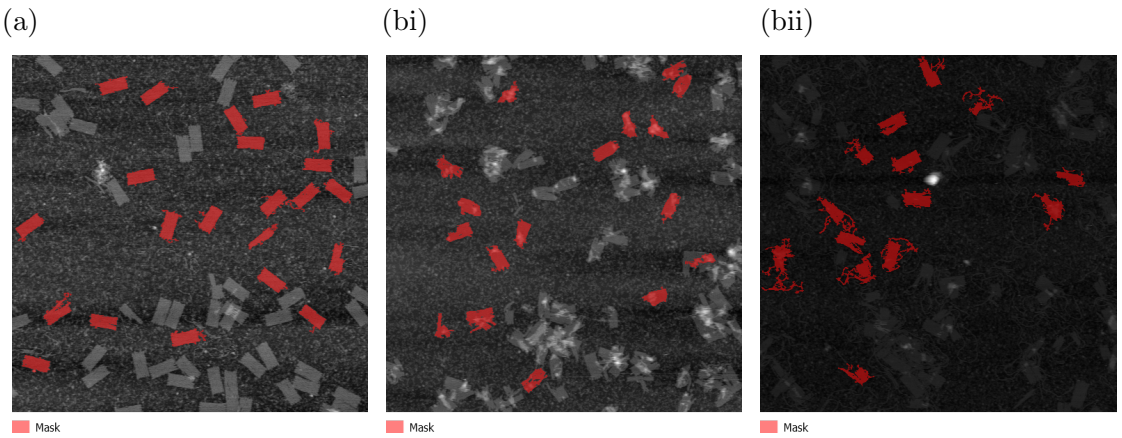


Figure 2.9: Images from Figure 2.8 following filtering to remove grains which were too small or too large to correspond to a single tile (see Appendix 2.3.1) or overlapped with the image edge.

In sample (bii), most of the 'string-like' regions of tiles have unfortunately been filtered out at this stage. However, what remains still significantly affects the

measured boundary length to allow these poorly-folded tiles to be distinguished (see Figure 2.11).

3. Remaining grains which do not correspond to single tiles were identified by visual inspection (shown in yellow in Figure 2.10) and removed. For sample (bii), this includes 'tiles' where most of the structure is string-like and thus filtered out in step 2, so that what remains does not correlate to a whole template. This does result in a bias for samples like (bii), as very poorly-folded tiles are excluded from the analysis.

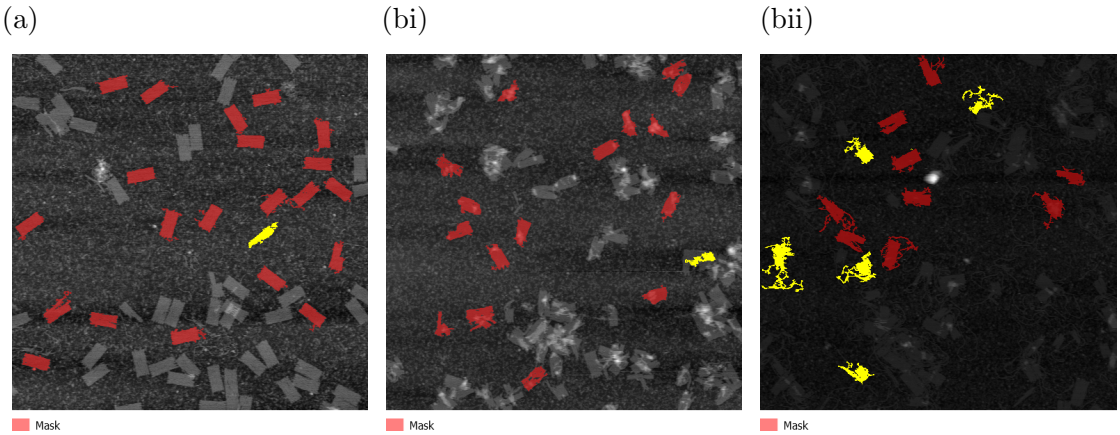


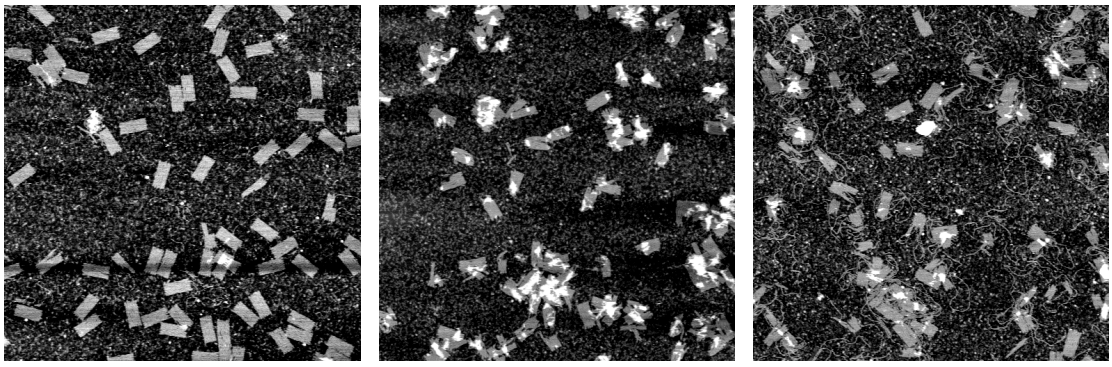
Figure 2.10: Images from Figure 2.9 were visually inspected to identify remaining grains which do not correspond to single tiles (shown in yellow), which were then removed.

4. The projected areas, mean heights, and projected boundary lengths of the grains were measured. For each sample, a total of more than 40 tiles were measured (across all four images); in sections 2.4 and 2.5 mean values for all four images are quoted.

For the three examples used here, the measurements for these specific images are shown in Figure 2.11.

(a) Well-folded tiles

(b) Poorly-folded tiles produced:
i) without helpers
ii) using helpers



grain count: 21
mean area: 2775 nm^2
mean boundary length: 345 nm

grain count: 16
mean area: 2184 nm^2
mean boundary length: 306 nm

grain count: 9
mean area: 2995 nm^2
mean boundary length: 577 nm

Figure 2.11: Mean areas and boundary lengths for the three example AFM images shown in Figure 2.7. These measurements were determined as per steps 1-4 of the 'Origami AFM analysis process' detailed above. As expected, the mean area of sample (bi) is lower than for the well-folded tiles, as misfolding has resulted in parts of the tile protruding upwards. In addition, the mean boundary length for sample (bii) is higher than for the well-folded tiles due to the presence of template-helper duplex ('string-like' structures).

Figure 2.11 shows that the two poorly-folded tile samples do deviate from the relatively well-folded tiles in Figure 2.11a as expected: poorly-folded tiles formed by incubation of template and staples (bi) have smaller areas as parts of the tile protrude upwards rather than lying flat, and poorly-folded tiles formed from the helper-coated template (bii) have longer boundary lengths due to the presence of unfolded helper-template duplex.

For comparison, the expected area of the tile (when well-folded) was calculated. Based on the design shown in Figure 2.2, the expected length is 75 nm (assuming 0.34 nm per base pair [149]) and the expected width is 35 nm (assuming each helix has a diameter of 2 nm, and there is a gap of 1 nm between helices due to electrostatic repulsion [122]). These values give an area of 2625 nm^2. Samples which were visually identified as being 'well-folded' (including sample a) had mean areas between 2750-3100 nm^2. This discrepancy may be due the resolution of the AFM images, and/or

the presence of excess staples and helpers in the background which may be wrongly included as part of a tile. It is difficult to calculate the expected boundary length as the short sides of the tile are not straight (see Figure 2.2b). Samples which were visually identified as being 'well-folded' had boundary lengths of 320-420 nm.

This analysis technique has certain limitations. A large proportion of the tiles in each image cannot be analysed, including those which:

- are in contact with another tile. Well-folded tiles produced by the conventional method of annealing template and staples adhered to the mica at such high concentrations that nearly all tiles were touching other tiles (see Figure 2.17h), and so these samples could not be analysed.

- contain a lot of 'string-like' structures as in sample bii), as the methods used are not good at identifying these structures or assigning them to the correct tile. This, combined with the reduced adherence of poorly-folded structures to the mica, results in better-folded structures being over-represented in the analysis.

The latter limitation means that small differences in how well the tile is folded cannot be identified using the measurements produced by this analysis method. These are seen more clearly by gel electrophoresis, as shown in Figure 2.12. However, the AFM images and their measurements have been a useful tool to aid interpretation of the gel experiments.

2.3.2 Gel electrophoresis

Gel electrophoresis allows us to distinguish between structures which migrate at different rates when a voltage is applied. Using 0.7% agarose gels run at 60 V for 2 hours at room temperature (see Appendix A.4 for full experimental details) allowed the differentiation of well-folded tiles from poorly-folded tiles, as demonstrated in Figure 2.12.

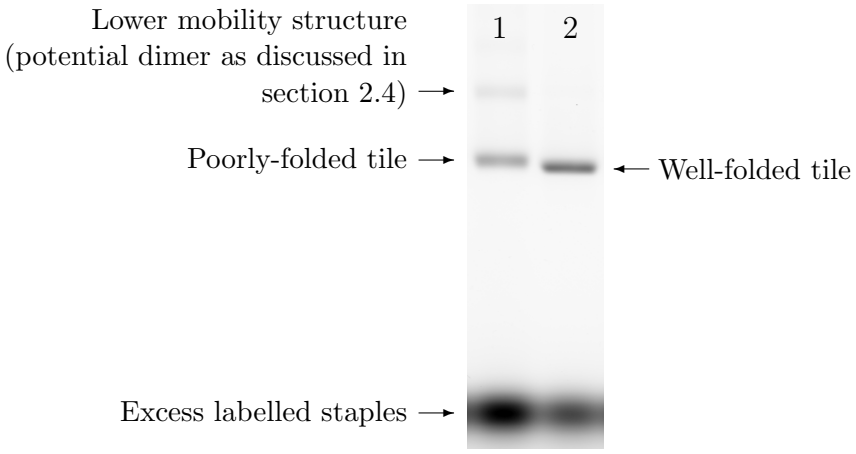


Figure 2.12: Gel electrophoresis of samples of poorly-folded tile (lane 1) and well-folded tile (lane 2). Lane 1: template incubated with the staple set at 60°C for 1 hour; lane 2: template and staples thermally annealed. One staple is fluorescently-labelled (the 'body label', see section 2.3.2.1) and visualised here. Since staples are at 2× excess to the template, excess staples (unbound to template) migrate more quickly through the gel and can be seen towards the bottom of the figure. Poorly-folded tiles were always observed to migrate more slowly than well-folded tiles, presumably due to a less compact shape. See section 2.1 for the tile and staple design, 2.4 for discussion of these results.

Gel electrophoresis is a quick and efficient way to characterise a large number of samples, and so is often used as an initial characterisation step before more detailed analysis of interesting results using AFM or TEM [98, 135]. It can be used for looking at folding intermediates, as demonstrated by Martin and Dietz [98], although it is possible that these intermediates may not be stable to the conditions used (the gel environment, temperature and time required for the gel to run). One way of reducing this potential complication is to quench intermediates in liquid nitrogen and then quickly thaw before running the gel [135].

Visualisation of the DNA structures present in the gel is made possible by either a) staining the gel with a fluorescent molecule which binds non-specifically to DNA (e.g. Sybr Gold), or b) use of fluorescently-labelled DNA strands (as in Figure 2.12, where one of the staple strands is labelled). Although this is far from a single-molecule technique like AFM, it is possible to observe multiple different structures per sample/lane (providing they have different electrophoretic mobilities, as in lane 1 of Figure 2.12), and if multiple labelled strands (with different fluorophores) are used, the distribution of these strands between different structures can also be observed. In the helper-origami experiments, each gel contained three strands labelled with different fluorophores (unless otherwise stated), as detailed below.

2.3.2.1 Fluorescently-labelled strands used for gel electrophoresis

Each gel used to characterise the helper-origami system contained at least one of the three following labelled strands (see Figure 2.18 for an example which contains all three):

1. A labelled body staple, visualised in red.
2. A labelled seam staple, visualised in blue.
3. A labelled helper strand, visualised in green. The template binding domain of this labelled helper overlaps with that of the 5' domain of the labelled body staple as shown in Figure 2.13b. The overlap region is 10 nts for the 10-offset helper set and 8 nts for the 8-offset helper set.

The template binding positions of the labelled strands are shown in Figure 2.13.

Each gel contained two 'standards' which can be seen in lanes A and B of Figure 2.18. Lane A contains helper-coated template (template annealed with the full helper strand set including the labelled helper); lane B contains folded tile (template annealed with the full staple set including both the labelled body and seam staples). These were used to normalise the pixel intensity to allow unbiased visual analysis

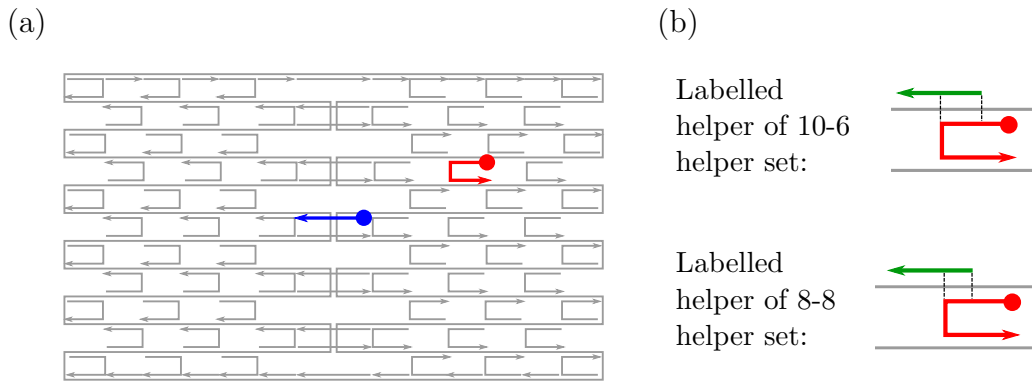


Figure 2.13: Labelled staple and helper strands used in gels. (a) Labelled staples: in red is the labelled body staple (visualised in red on the gels), in blue is the labelled seam staple (visualised in blue). (b) Template binding positions of labelled helper strands (shown and visualised in green) relative to the labelled body staple. The labelled 10-6 helpers overlaps with the labelled body staple by 10 nt, the labelled 8-8 helper by 8 nt. See Appendix B.1 for strand sequences.

(see Appendix A.4.1 for method) and also to allow comparison of electrophoretic mobilities with other samples in the gel.

2.3.2.2 Repeatability of gel results

All gel results shown in this chapter were repeated at least once. These repeats showed that the gel results from incubating helper-coated template with staples for a specific period of time at a specific temperature were not completely reproducible, in so far as the temperature at which a particular result was seen differed between repeats by up to 2°C. For example, the first lane of Figure 2.16 shows the result of an identical experiment to the 55°C lane of Figure 2.18 (the only difference is that Figure 2.16 lacks the labelled seam staple which is visualised in blue in Figure 2.18). However, the electrophoretic mobilities of the main band in each of this repeats is clearly different: in Figure 2.16 it is slower than that of the well-folded tile (lane 8), whereas in Figure 2.18 it has a similar mobility to the well-folded tile (lane 'N').

This variation between repeats is likely to be because the time between the end of incubation and loading onto the gel (during which samples were kept at 4°C) varied between repeats. Within one repeat, this length of time was identical. As previously

mentioned, this only resulted in variations of up to 2°C in the temperature at which particular results were seen, and does not affect the conclusions drawn from gel results in this chapter.

2.3.3 Temperature-controlled real-time fluorescence measurements

Fluorescence techniques can be used to study the whole folding process in solution, as it occurs, and so do not suffer from the same drawbacks as AFM and gel electrophoresis in terms of exposing sampled intermediates to conditions which might affect their structure. In addition, data from many more time-points can be easily collected.

Whole structure measurements

Both double- and single-stranded DNA absorbs UV light of wavelength 260 nm, although the absorption coefficient of double-stranded DNA is higher (meaning that it absorbs more light per unit mass) [46]. Because of this difference, it is possible to monitor the formation of staple-template interactions during origami folding (e.g. during thermal annealing), as the formation of more double-stranded DNA will result in an increase in absorption at 260 nm [6]. However, this method has a very low signal:noise ratio due to the high background signal from the single-stranded staples which must be in excess for successful origami folding [33, 90].

A better method to follow double-stranded DNA formation is to use a DNA-binding dye which is much more fluorescent when bound to dsDNA compared to ssDNA, for example Sybr Green I and EvaGreen [94, 172]. This increase in fluorescence is a result of the dyes intercalating between the stacked base pairs in the dsDNA double helix. Such dyes have been used extensively to characterise the formation and melting profiles of 2D and 3D origami structures [18, 23, 34, 80, 135], an example of which is shown in Figure 2.14 for the origami tile I am using. Here, as with most origami nanostructures, we can see the narrow temperature range over which staples bind to the template (during annealing/cooling) or dissociate from the template (during melting/heating), due to the co-operativity of these processes (see section 1.2.2.1). The small difference between melting and formation temperatures (i.e. hysteresis; as illustrated by the different positions of the red and blue peaks in

the right-hand plot of Figure 2.14) is also a result of staple co-operativity.

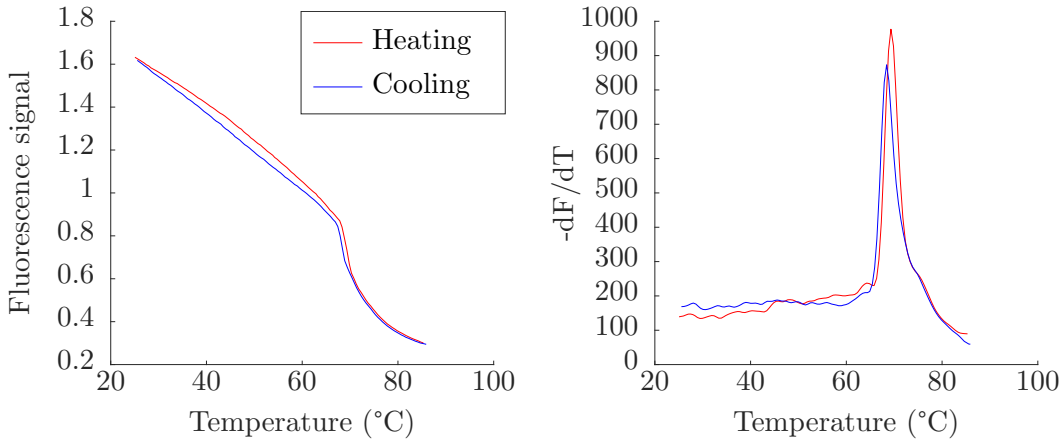


Figure 2.14: Melting and formation of the origami tile as shown by EvaGreen fluorescence. Left: EvaGreen fluorescence signal during heating and cooling. The sudden change in fluorescence at ∼69°C is due to the melting (during heating) or formation (during cooling) of the tile. Right: the negative of the differential of the fluorescence signal with respect to temperature (-dF/dT) shows the change in fluorescence at the melting temperature more clearly. See Appendix A.5 for full experimental details including data processing.

The drawbacks to such dyes are that they can show some sequence specificity, they may not be stable at higher temperatures, and their intercalation may affect nanostructure assembly and therefore the melting/formation temperatures observed (although this is thought to be minimal at the dye concentrations used [135]). EvaGreen is superior to Sybr Green I in terms of all three of these characteristics [35, 94], and was used in Figure 2.14 to characterise the formation and melting of the origami tile, and in Figure 2.19 for the the formation and melting of the helper-coated template (using the full helper set). However, since these 'whole structure' techniques measure the amount of double-stranded DNA present, they could not be used to study the transition from helper-coated template to origami tile. Instead, the following technique (using a FRET-labelled pair of staples) was used.

Probing specific parts of a structure

Using a dye which binds only to a specific part of the structure can be used to study folding/melting of a particular domain or region. One can use either a fluorescently-labelled staple strand whose fluorescence changes upon binding of the staple to the

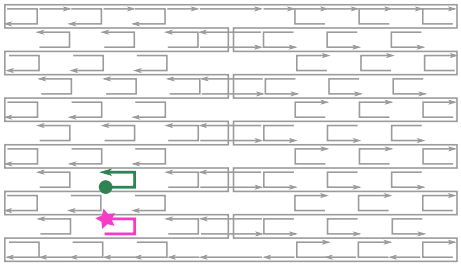
template, or (as I have done) use two staple strands which are near each other in the folded origami labelled with a pair of FRET fluorophores[2] (Figure 2.15a). For the latter experiment, when the two staples are brought into close proximity by a) both being bound to the template and b) the template being folded in such a way that the staple domains are brought close together (e.g. due to template folding during thermal annealing), excitation of the donor fluorophore (pink in Figure 2.15a) allows emission from the acceptor fluorophore (green) and so the FRET signal increases. As shown in Figure 2.15b, this leads to a similar signal to the EvaGreen fluorescence shown in Figure 2.14.

This FRET-based method has been used to follow folding and melting at different positions of 2D and 3D origami nanostructures, and to observe the effect on folding in a specific region of introducing close and distant defects [151]. In addition, it has also been used to study the assembly of non-origami DNA nanostructures [123].

All of the types of real-time measurement described in this section (2.3.3) require the ability to heat and cool samples whilst simultaneously measuring their optical absorbance/fluorescence. This can be achieved by using a real-time PCR (polymerase chain reaction) machine, which commonly allows up to 96 samples to be exposed to the same thermal cycle and monitored simultaneously. The drawback of such techniques compared to AFM and even to gel electrophoresis, is that they are bulk measurements that will not be able to distinguish between different populations in the same sample.

[2]Förster or fluorescence resonance energy transfer (FRET) is a phenomenon whereby excitation of a donor fluorophore leads to emission from a second (acceptor) fluorophore. Energy is transferred via non-radiative dipole–dipole coupling [85]. The efficiency falls steeply with increasing distance (it is related via an inverse 6th-power law) and so can be used to confirm whether the donor and acceptor fluorophore are within a short distance of each other [85].

(a)



(b)

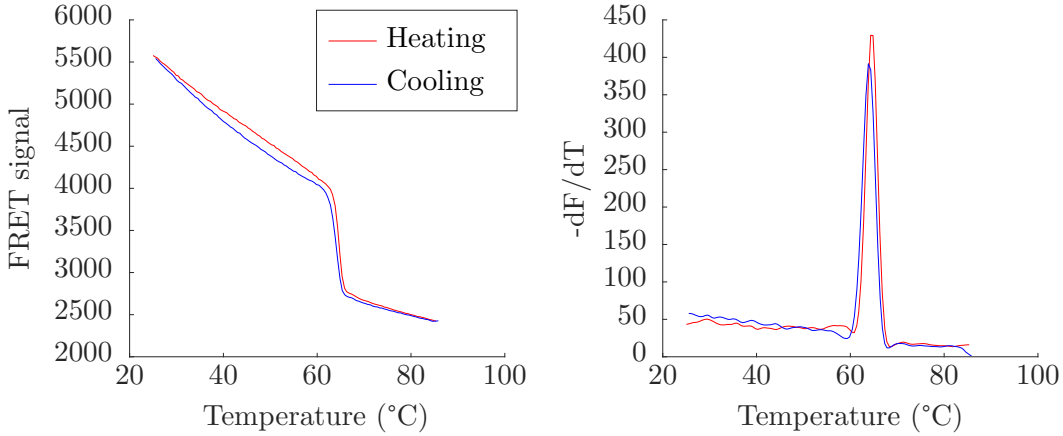


Figure 2.15: Melting and formation of (a region of) the origami tile as shown by FRET signal. (a) Position of FRET labelled staples (pink = donor, green = acceptor), (b) FRET signal (left) during heating and cooling and (right) the negative of the differential of this signal with respect to temperature. The plots look very similar to the EvaGreen fluorescence changes for the whole structure (Figure 2.14), although the peak of the differential plot here is narrower and shifted to a lower temperature (∼64°C vs 69°C for EvaGreen). These differences are due to the fact that here we are only looking at incorporation/folding of a specific pair of staples (hence the peak is narrower), which must presumably bind in a region of the tile which folds at a lower temperature to other parts of the tile (hence the temperature shift). See Appendix A.5 for full experimental details.

2.3.4 Simulations

In addition to the three experimental techniques described above, the helper-origami system was studied using simulations. These were carried out by Behnam Najafi using a domain-level model of DNA origami originally developed by Dannenberg et al. [23], which has been extended to include the effects of helpers.

Unless otherwise indicated, the simulations quoted in this chapter used the same DNA sequences as the experimental results (see Appendix B.1). Where indicated, 'average sequences' were used, where the average parameter values for base-pairing and base-stacking interactions are used for all bases.

One of the limitations of these simulations is that staple or helper 'mis-binding' (i.e. binding to a non-target domain) is not possible in the model. Incorporating mis-binding into the model would violate the domain-level abstraction used and thus require much greater computing power and/or time to complete simulations.

2.4 Helper strands improve isothermal origami tile folding

In this section, I show that the use of helper strands improves isothermal folding of the origami tile (when used to coat the template), allowing rapid folding at lower temperatures than are otherwise possible (section 2.4.1). In order to understand how the helper strands have this effect, I have characterised the folding process when helpers are used, and studied the effects of temperature (section 2.4.2) and time (section 2.4.3) upon this process. These results are discussed in section 2.4.6, in which I also compare the use of helpers to other methods which have been developed for isothermal origami folding (described in section 2.4.5).

Unless otherwise stated, all experiments in this section use a helper-coated template produced by annealing the template with (a 2× excess of) the full 10-offset helper set. As described in the introduction to this chapter, initial experiments showed that omitting helpers to leave sections of the template single-stranded had no effect on the results (see Figure 2.34 and discussion in the conclusions section 2.6), hence the use of the full set here. Staples are then added to the helper-coated template, and the mixture incubated at various temperatures and for various time periods.

2.4.1 Helper strands improve isothermal tile folding at 60°C-65°C

Helper-coated template and single-stranded template were incubated with staples at 55, 60 and 65°C for 1 hour, or thermally annealed with the staples. The gel electrophoresis results are shown in Figure 2.16.

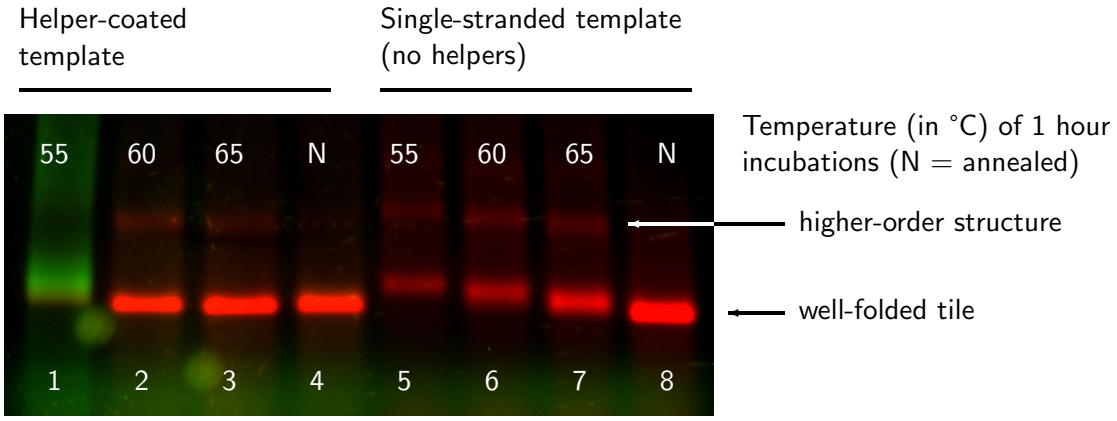


Figure 2.16: Gel electrophoresis of tile folding using helper-coated template and single-stranded template. One helper strand (green) and one body staple (red) are labelled as described in section 2.3.2.1. Lane 8 contains the well-folded tile (produced by annealing template and staples). Lanes 2 and 3 also contain this band, showing that well-folded tiles have been formed by isothermal incubation of the helper-coated template at 60 and 65°C, unlike the incubation of single-stranded template with staples at these temperatures. Lanes 1-3 and 5-7 have a secondary band of a lower-mobility structure which is not present when the samples are annealed.

Lane 8 contains the well-folded tile (produced by annealing template and staples). Lane 4 is similar in appearance to 8 (both in terms of electrophoretic mobility and fluorescence), confirming that helper-coated template annealed with staples also produces well-folded tiles. When the template is coated with helpers, folding at both 60 and 65°C produces well-folded tiles, as evidenced by the presence in lanes 2 and 3 of the same band as in lanes 4 and 8. In contrast, folding at these temperatures without helpers (lanes 6 and 7) produces structures with lower electrophoretic mobilities and lower fluorescence than the folded tile, which are presumably misfolded structures. Folding at 55°C with helper-coated template (lane 1) produces structures containing both labelled helper strand (green) and labelled body staple (red)

bound. These structures presumably represent intermediates between the helper-coated template and the folded tile. These intermediates are analysed further in section 2.4.2.

All lanes containing isothermally-folded structures (whether helpers are used or not) also contain a secondary band of a lower-mobility structure (labelled 'higher-order structure' in Figure 2.16). I propose that this is a dimeric structure, as discussed in section 2.4.2.1.

These samples were subjected to AFM imaging and analysis as per section 2.3.1.1 (Figure 2.17). The results support the observations made from the gel (Figure 2.16) in so far as:

- For helper-coated template folded at 55°C, the AFM image (Figure 2.17a) shows that much of the template is bound to helper strands (confirmed by the high boundary length), which is why the gel structure has (green) labelled helper present and a lower electrophoretic mobility than the folded tile.

- Helper-coated template folded at 60/65°C or annealed, which have similar appearances on the gel (with the exception of the lower mobility structure in the isothermally-folded samples), also appear very similar to one another (and have similar measurements) in the AFM images (Figure 2.17b-d). All 3 still appear to contain some helper-template duplex[3], but also a number of well-folded tiles.

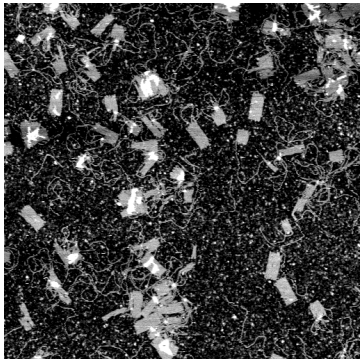
- For single-stranded template folded at 55-65°C, the AFM images (Figure 2.17e-g) show better folding with increasing incubation temperature (as seen in the increase in mean area), which correlates with the changes in electrophoretic mobility and fluorescence seen on the gel.

For the reasons outlined in section A.6.2, the AFM measurements are unable to identify small differences in folding. For example, it is difficult to clearly distinguish

[3]Some of the AFM images of tiles folded without helpers also contain small amounts of 'string-like' structure (not those shown in Figure 2.17). This is likely to be due to incomplete nucleasing during the preparation of the template strand (see Appendix A.2.2). However, this was present to such a small extent that it does not significantly affect the measurements made for these images.

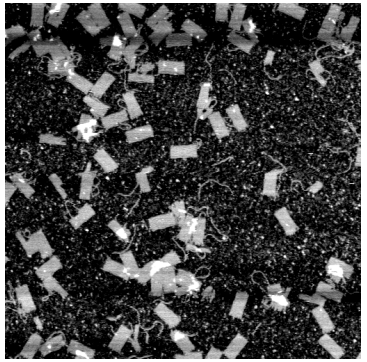
between Figure 2.17b and g (or c and g) using the AFM images or measurements, whereas the gel (Figure 2.16) shows a clear difference in both electrophoretic mobility and fluorescence for these samples (lanes 2 and 7 / 3 and 7).

(a) Lane 1: Helper-coated template, 55°C



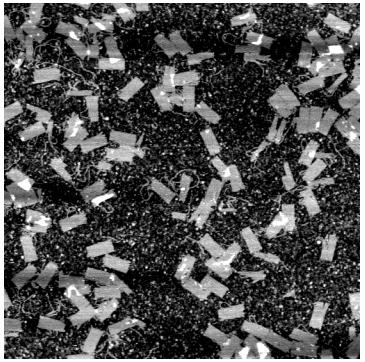
mean area: 3011 nm^2
mean boundary length: 677 nm

(b) Lane 2: Helper-coated template, 60°C



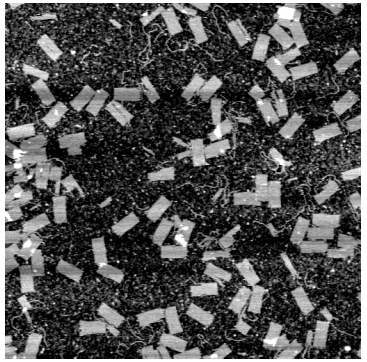
mean area: 3051 nm^2
mean boundary length: 386 nm

(c) Lane 3: Helper-coated template, 65°C



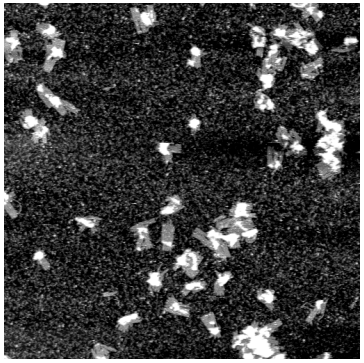
mean area: 2894 nm^2
mean boundary length: 379 nm

(d) Lane 4: Helper-coated template, annealed



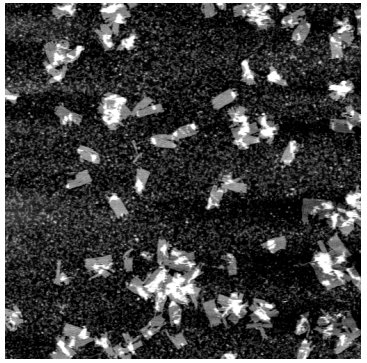
mean area: 3040 nm^2
mean boundary length: 415 nm

(e) Lane 5: No helpers, 55°C



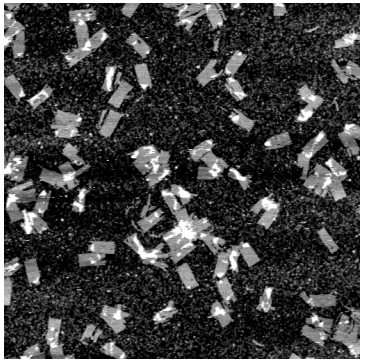
mean area: 1932 nm^2
mean boundary length: 339 nm

(f) Lane 6: No helpers, 60°C



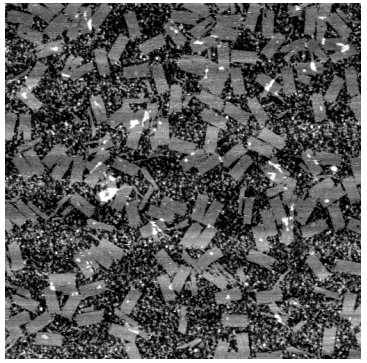
mean area: 2364 nm^2
mean boundary length: 331 nm

(g) Lane 7: No helpers, 65°C



mean area: 2547 nm^2
mean boundary length: 329 nm

(h) Lane 8: No helpers, annealed



measurements could not be made due to tiles touching

Figure 2.17: AFM images of samples from Figure 2.16, processed as per Appendix A.6.1. Measurements were determined as per section A.6.2. It is not clear why the number of tiles in h) is significantly greater than in d); this could be because the 'extra' DNA in d) (the helper strands) bind to the mica and prevent the the adhesion of tile. Images are 1×1 μm.

2.4.2 Temperature-dependent transition from helper-coated template to fully-folded tile

In the previous section, I showed that the use of helpers (to coat the template) improves tile folding when the template is incubated with staples at 60 and 65°C for 1 hour. When the helper-coated template and staples is incubated at 55°C for 1 hour, the structure formed (lane 1 of Figure 2.16a) appears to be intermediate between the helper-coated template and folded tile, containing both labelled helper (green) and labelled body staple (red). Using a higher incubation temperature allows the staples to completely displace the remaining helpers and form the tile. To study the effect of temperature further, helper-coated template was incubated with staples at a range of temperatures for 1 hour (Figure 2.18).

At temperatures of 50°C and below, the helper-coated template is unchanged by incubation with the staples (no staples are incorporated into the structure) in 1 hour incubations. With increasing incubation temperature, there is a sharp transition (over ∼4°C), and by 54°C a structure resembling the well-folded tile (which is shown in lane B for comparison) is present. NB: the bands visible in the green channel in lanes 55°C onwards are due to bleed-through from the red channel (as evidenced by the presence of this band in lane B, which contains no (green) labelled helper).

As seen in the previous section (2.4.1), isothermal incubation of helper-coated template and staples also produces a small amount of lower mobility structure, seen in Figure 2.18 as a faint band in the green channel at ∼50-54°C (the electrophoretic mobility of this band changes between 51 and 53°C), and in the red and blue channels from ∼ 53°C onwards. This band is not present in the annealed sample (N). This will be discussed in section 2.4.2.1.

As explained in section 1.2.2.1, origami structures are conventionally folded by thermal annealing: heating the template and staples up to 90/95°C (at which temperature all DNA is single-stranded) and then slowly cooling. During cooling, the staples are incorporated over a very narrow range of temperatures, due to co-operativity. For the tile used here, the melting temperature (the temperature

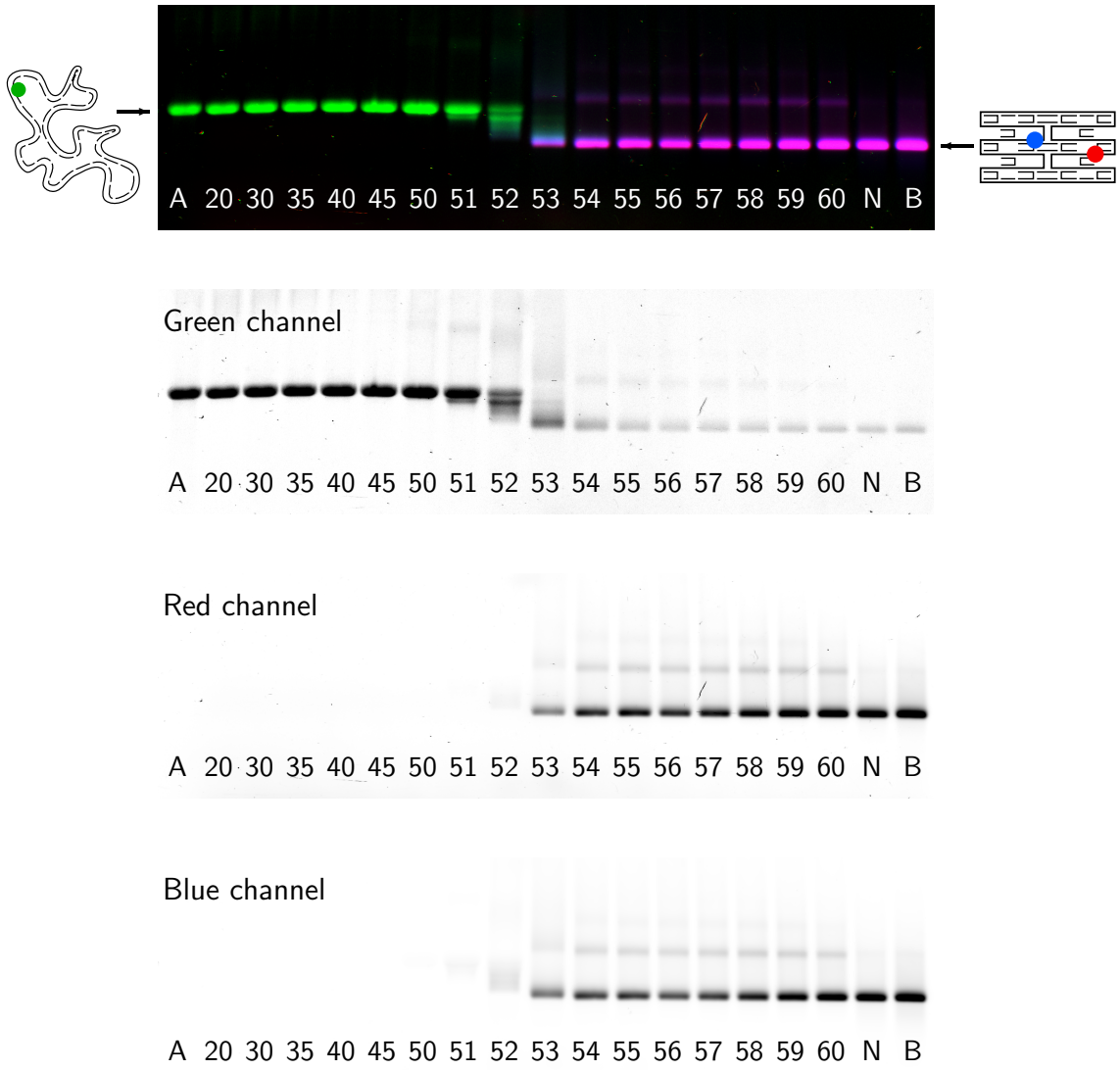


Figure 2.18: Gel electrophoresis of helper-coated template incubated with tile staples at various temperatures for 1 hour, or annealed with staples (lanes labelled N). The top image shows all three channels for the three fluorescently-labelled strands used: helper (green), body staple (red) and seam staple (blue), as described in section 2.3.2.1. The gel contains two standards represented by the schematics on either side of the gel: A = helper-coated template (no staples added); B = template annealed with staples i.e. the folded tile (no helpers). As explained in section 2.3.2.2, differences between these results and those shown in Figure 2.16 are likely to be due to differences in the time period between incubation and running the gels. NB: the bands visible in the green channel in lanes 55°C onwards are due to bleed-through from the red channel (as evidenced by the presence of this band in lane B, which contains no labelled helper). In the lower three images of individual channels, pixel intensity has been normalised using the background fluorescence and the standards (see section 2.3.2.1).

at which ∼half of the staple domain-template interactions have formed during this cooling step) is approximately 69°C (Figure 2.14). In contrast, in the helper-origami system, heating is required in order to allow the staples to displace the helpers. The temperature at which this occurs (53-54°C in 1 hour incubations) is lower than the melting temperature of the tile. It is also lower than the melting temperature of the helper-coated template (∼63°C; see Figure 2.19), a valid comparison as the concentrations of helper and template strands used in Figure 2.19 are identical to those used in the helper-origami folding experiments.

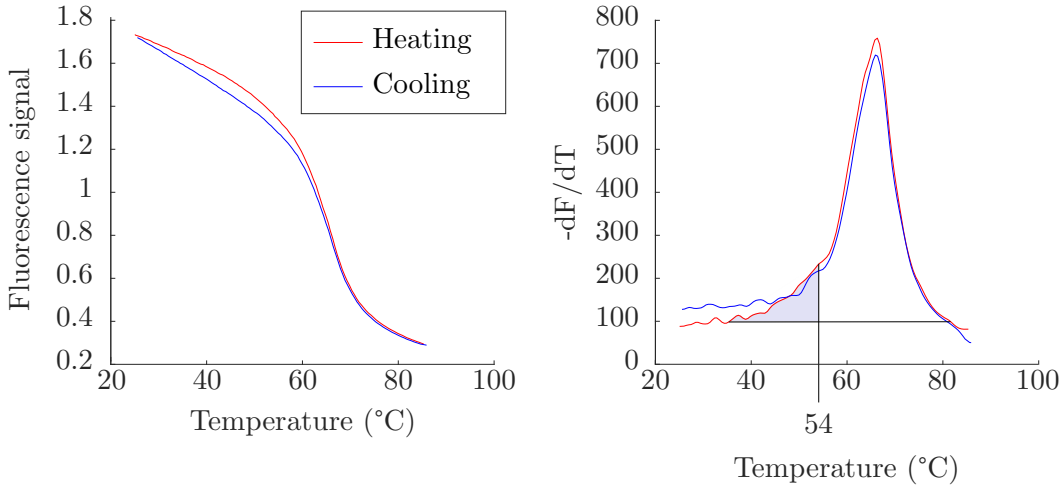


Figure 2.19: Melting and formation of the helper-coated template (using the full 10-offset helper set) as shown by EvaGreen fluorescence (see section 2.3.3 for methodology). Left: EvaGreen fluorescence signal during heating and cooling. Right: the negative of the differential of the fluorescence signal with respect to temperature (-dF/dT) shows the change in fluorescence at the melting temperature (∼63°C) more clearly. The pale purple area shows the proportion of helpers that have melted off at 54°C (∼12%). Note that the concentrations of helper and template strands were identical to those used in the incubation experiments of helper-coated template and staples.

Melting of the helper-coated template occurs over a wider range of temperatures than for the tile (due to the lower degree of co-operativity between helpers compared to staples[4]), and so a small number of helpers (∼12%, see Figure 2.19) will have melted off at 53/54°C, but not a substantial proportion. The transition from helper-

[4]Staples show co-operative binding due to a) the reduction in template entropy when a staple brings together two distant parts and b) stacking interactions between bound staples. Binding of the helper strands to the template show some co-operativity due to b), but do not benefit from a) as each binds to a single template domain only.

coated template to tile is therefore not simply due to helpers melting off and thus allowing the staples to bind.

To gain a better understanding of the transition, I have studied the intermediate structures formed during it. Figure 2.18 shows that the lowest temperature at which a structure other than the helper-coated template appears is 51°C: there is an additional band with a higher electrophoretic mobility than the helper-coated template, which presumably represents an intermediate structure in which some staples are bound to the template. By 53°C, the red and blue channels show a structure with similar (but slight lower) electrophoretic mobility than the tile (this may represent a 'nearly-folded' tile), but the green channel shows that there are also intermediates with even lower electrophoretic mobilities which have some labelled helper bound. To study these intermediates further, incubations of helper-coated template and staples were carried out at an expanded range of temperatures between 52 and 54°C (Figure 2.20; green channel only).

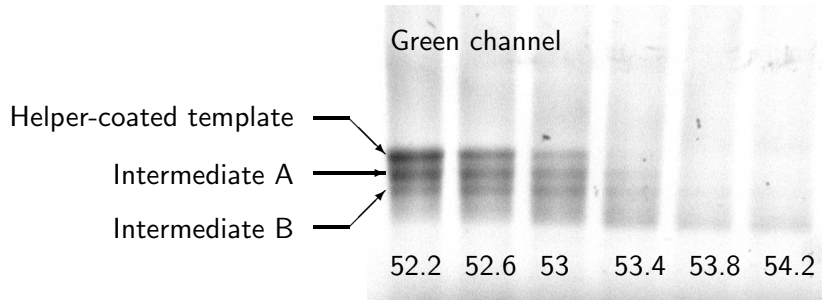


Figure 2.20: Gel electrophoresis of helper-coated template incubated with tile staples at various temperatures (shown in °C) for 1 hour. Only the green channel (labelled helper, see section 2.3.2.1) is shown. At the lower temperatures shown, a number of distinct intermediates are present which may correspond to structures with different numbers of seam cross-links formed.

At 52.2-53°C, there are at least two distinct structures present in addition to the helper-coated template (labelled 'intermediate A' and 'intermediate B' in Figure 2.20). The presence of distinct intermediates suggests that there is a defined folding pathway by which the helper-coated template folds into the tile, i.e. there is a specific order in which the staples are incorporated and different parts of the template

become folded. We hypothesised that the different intermediates observed may have different numbers of seam cross-links (see Figure 2.21a) formed, for example intermediate A may contain one seam cross-link, whilst intermediate B may contain two (Figure 2.21b).

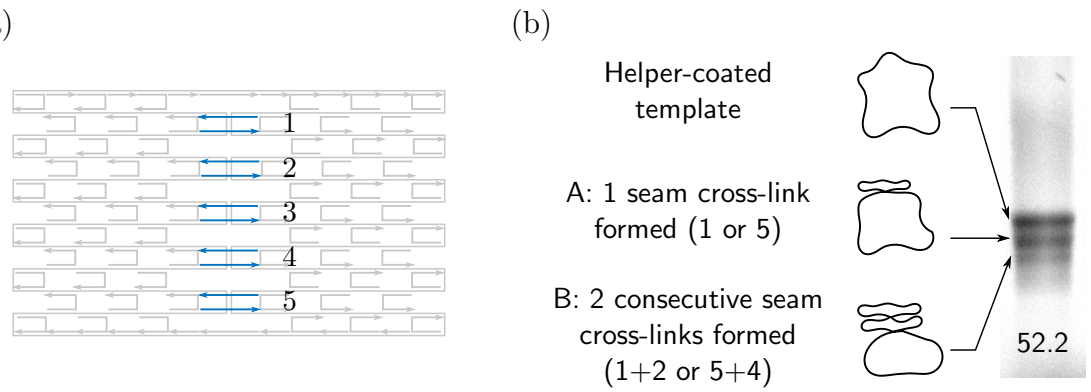


Figure 2.21: Proposed seam cross-link formation in intermediates during the transition from helper-coated template to tile. (a) The tile contains five cross-links (numbers 1-5), each mediated by a pair of seam staples (blue). (b) Gel result of incubating helper-coated template with staples at 52.2°C for 1 hour (this is lane 1 of Figure 2.20. Green channel only (labelled helper) shown. On the left are schematics of the proposed structures represented by the bands.

Each of the five seam cross-links in the tile are formed by a pair of seam staples, and binding within each pair has been shown to be highly co-operative during folding of the tile by thermal annealing [34]. This co-operativity is due to the fact that binding of one of the pair brings the template binding domains of the second of the pair into very close proximity. This co-operativity is likely to also affect seam staple binding in the helper-origami system: binding of the first of the pair will make it more favourable for the second of the pair to bind via helper displacement.

There is some evidence for this explanation of the electrophoretic mobilities of the intermediate structures. A repeat of the gel shown in Figure 2.20 was carried out, where the four seam staples involved in forming cross-links 2 and 4 (Figure 2.21) were omitted (Figure 2.22b). This change led to the apparent disappearance of the band representing intermediate B at 52.2-53°C, suggesting that the formation of B at these temperatures is dependent on the formation of cross-link 2 and/or 4.

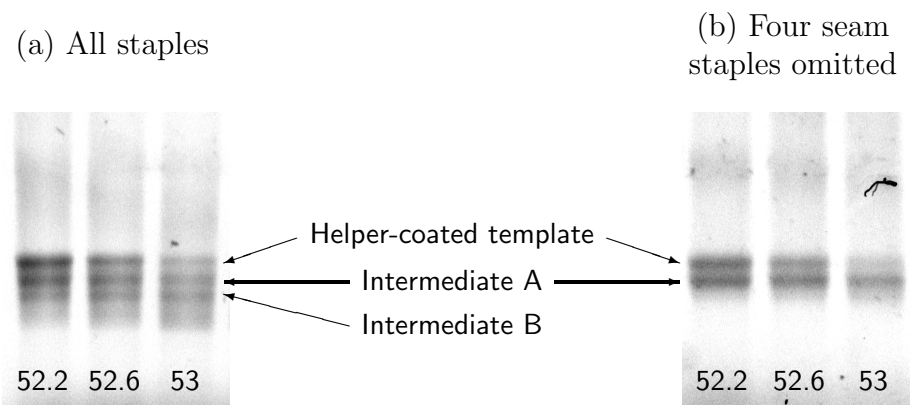


Figure 2.22: Omitting four seam staples leads to the disappearance of the intermediate B band when helper-coated template is incubated with staples (numbers show temperature in °C of 1 hour incubations). The four seam staples omitted in (b) are those which form cross-links 2 and 4 in Figure 2.21a. Bands have been identified by their relative electrophoretic mobilities: I have assumed that the lowest mobility band in both gels at 52.2°C is the helper-coated template, and the band below is intermediate A. Only the green channel (labelled helper, see section 2.3.2.1) is shown.

Simulations of thermal annealing of the tile (without helpers) carried out by Frits Dannenberg showed small differences ($\leq$2°C) in the temperature at which different body staples were incorporated [24]. To investigate whether this is also true of folding in the helper-origami system, repeats of the gel shown in Figure 2.18 were carried out using labelled helper strands and body staples with binding domains at different positions on the template[5]. However, no significant differences were observed. This may be due to the limitations associated with using gel electrophoresis to characterise the folding process: the fact that the sample is removed from the incubation and then subjected to gel electrophoresis for two hours can introduce artifacts in the samples (i.e. small changes in structure). This means that the gel results are not accurate enough to allow small differences between samples to be ascribed to differences in folding during incubation.

Further elucidation of the folding pathway by which the helper-coated template folds into the tile was aided by the results of simulations carried out by Behnam Najafi using an extended version of the model used above by Frits Dannenberg (see

[5]It should be noted that the individual sequences of the labelled strands (and those of the strands which they are competing with for template binding) will also affect the temperature at which they bind to/dissociate from the template, in addition to where on the template they bind.

section 2.3.4) and is discussed in section 2.5.

2.4.2.1 Formation of higher-order structures

We have seen in Figures 2.16 and 2.18 that isothermal incubation of either helper-coated template or single-stranded template with staples produces a secondary band of a structure with lower electrophoretic mobility than the main band. Figure 2.23 shows the same gel as Figure 2.18, with the intensities altered to better show these lower mobility structures (which are also framed by boxes). We can now see that these structures are also present, to a lesser extent, in the lanes containing annealed sample (N and B). (As previously, fluorescence in the green channel above 54°C is due to bleed-through from the red channel.)

I propose that the lower mobility structure identified by boxes in Figure 2.23 is a dimeric version of the structure in the main band. For example, in the red and blue channels at 60°C, where the main band is the well-folded tile, the 'boxed' band is a dimeric tile. The fact that the changes in electrophoretic mobility seen in the main band are mirrored by the boxed band (in all channels) supports this explanation.

These dimeric structures can arise through either a) the presence of dimer template (two copies of the template joined head-to-tail), a small amount of which is likely to be present due to the template synthesis method used[6]; or b) cross-linking of two (monomer) template copies by a two-domain staple.

Lane A of Figure 2.23 contains no staples (only helper-coated template), and so any dimeric structure in this lane must be due to dimer template (a) above). There are two very faint low mobility bands visible in this lane in the green channel (identified by lines to the left of the gel in Figure 2.23), which may correspond to helper-coated dimer template. This means that there is (at most) a very small amount of dimer template present in each lane. This means that in lanes which contain a large amount of dimeric structure (e.g. the 60°C lane, red/blue channels),

[6]The method includes plasmid amplification in DH5α E. coli cells (see Appendix A.2.2), which is likely to result in production of small amount of dimer template in addition to the monomer template [34].

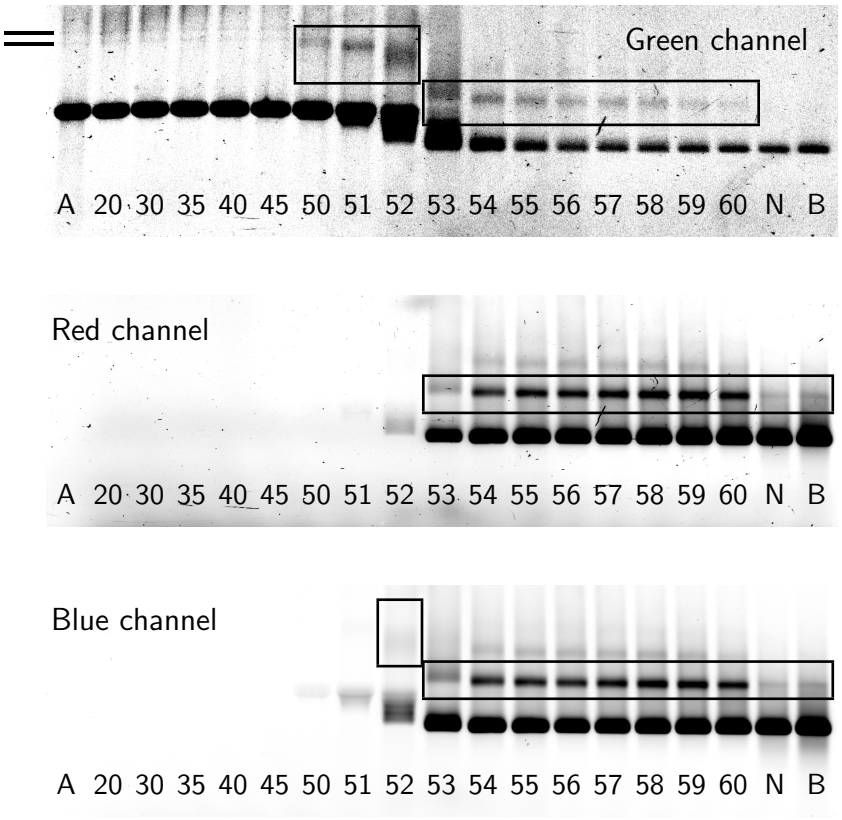


Figure 2.23: Gel electrophoresis of helper-coated template incubated with tile staples at various temperatures for 1 hour, or annealed with staples (lanes labelled N). The gel is the same as in Figure 2.18, with intensities altered to better show lower mobility structures. Boxes show lower mobility structures; lines to the left of the green channel identify structures which may contain dimer template (see main text). Green channel: labelled helper (as previously, fluorescence in this channel above 54°C is due to bleed-through from the red channel); red channel: labelled body staple; blue channel: labelled staple seam (see section 2.3.2.1). Pixel intensity has been normalised using the background fluorescence and the standards in lanes A and B (see section 2.3.2.1).

most of this must be cross-linked dimers (b) above).

Incubation of helper-coated template alone at the range of temperatures used in Figure 2.23 produced identical results to lane A at all temperatures, confirming that formation of the dimeric structures observed in the green channel at 50°C onwards requires staples. In addition, the first temperature at which these structures become easily visible in the green channel (50°C) is approximately the temperature at which staples first appear to bind to the template (51°C; i.e. when the appearance of the main band begins to deviate from that of the helper-coated template in lane A), which again suggests that staple cross-links between templates are responsible for the lower mobility bands observed.

There is a tertiary band, of lower mobility and lower fluorescence than the identified dimeric structures, visible in the red and blue channels in the 54-60°C lanes of Figure 2.23. These are thought to represent trimeric structures, as confirmed by analysing their relative electrophoretic mobilities compared to that of the dimeric and monomeric structures[7].

Figure 2.16 shows that the dimeric (and trimeric, made visible when the gel intensities are altered) structures are also produced during isothermal incubation of template and staples in the absence of helpers. At temperatures close to melting temperature of the staples from the template (69°C, see Figure 2.14) they largely disappear (see later Figure 2.31). This suggests that the cross-linking interactions are reversed by melting at these temperatures.

In general, we can see that much more of these cross-linked multimers are formed during isothermal folding compared to annealing. This is discussed further in section 2.4.5, along with other issues which can arise during isothermal folding of origami.

[7]The electrophoretic mobility of DNA structures of similar shape is inversely proportional to the log(molecular weight) [37]. Analysing the mobilities of the monomeric, 'dimeric' and 'trimeric' structures identified here suggest that their molecular weights are 1, 2 and 3, respectively.

2.4.3 Effect of incubation time

We have seen in section 2.4.2 that the transition from helper-coated template to folded tile is highly dependent on the incubation temperature. To investigate the effect of incubation time, helper-coated template was incubated with staples at various temperatures for 30 minutes, 1 hour and 4 hours (Figure 2.24).

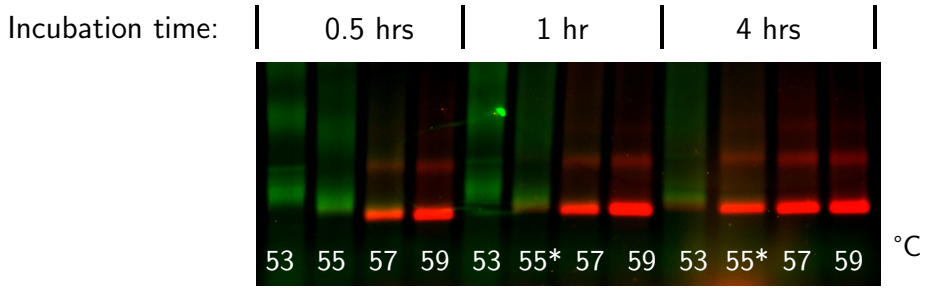
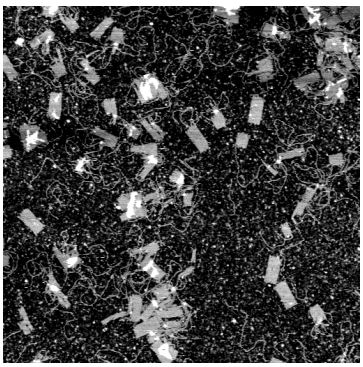


Figure 2.24: Gel electrophoresis results of helper-coated template incubated with staples at various temperatures for different lengths of time. Labelled strands are the helper (green) and body staple (red) as described in section 2.3.2.1. At temperatures below 59°C, further progression towards the folded tile occurs when the incubation time is increased (at 59°C the well-folded tile appears to be present after only 30 minutes). *samples shown in Figure 2.25.

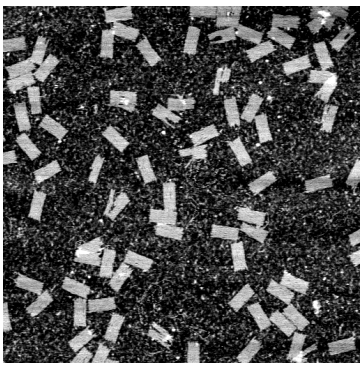
The results show that increasing the incubation time does allow further progression from the helper-coated template towards the folded tile, but clearly at temperatures below ∼60°C, it takes more than an hour for completion of this process. Figure 2.25 shows AFM images of two of the samples: helper-coated template incubated at 55°C for 1 hour (a) and 4 hours (b). The images clearly show that folding has progressed further in 4 hours than in 1 hour: there are many more well-folded tiles and much less helper-template duplex after 4 hours (as confirmed by the relative boundary lengths of the two samples).

(a) 55°C, 1 hour



mean area: 3011 nm^2
mean boundary length: 677 nm

(b) 55°C, 4 hours



mean area: 2752 nm^2
mean boundary length: 333 nm

Figure 2.25: AFM images of helper-coated template incubated at 55°C for 1 hour (a) or 4 hours (b). (a is the image in Figure 2.17a.) The reduction in boundary length between a and b correlates with the reduction in the amount of helper-template duplex which can be seen in the images as 'string-like' structures. Each image area is 1×1 μm. Images have been processed and analysed as per section 2.3.1.1 and Appendix A.6.

2.4.4 Further evidence of the role of helpers

We have seen in section 2.4.1 that the use of helpers improves tile folding at certain temperatures/timescales. An experiment was carried out to determine whether the helpers achieve this primarily by blocking the template and preventing staple 'mis-bindings' (i.e. staples binding to off-target locations on the template) and formation of template secondary structure, or by displacing already mis-bound staples.

To test this, single-stranded template was incubated with staples for 20 minutes, before the helpers were added (lane 3 Figure 2.26). The sample was then incubated for a further hour, and the results were compared to helper-coated template incubated with staples for the same length of time (lane 1), and single-stranded template where helpers were not added after 20 minutes (lane 2).

The results show that the addition of helpers after staples makes no difference to folding (lane 3 is similar in appearance to lane 2), whereas the use of helper-coated template shows better folding than both of these (as evidenced by the higher electrophoretic mobility of the structure in lane 1). A repeat of this experiment where the length of the final incubation was extended from 1 hour to 10 hours gave identical results.

These results suggest that the primary role of the helpers is in preventing unwanted interactions from forming, rather than in reversing such interactions after they have formed. These unwanted interactions may include both staple mis-binding and formation of template secondary structure. For helpers to improve folding, it must be the case that when the helpers are displaced from the template, they are mainly replaced by correctly-bound staples. This idea will be further explored in the discussion section (2.4.6).

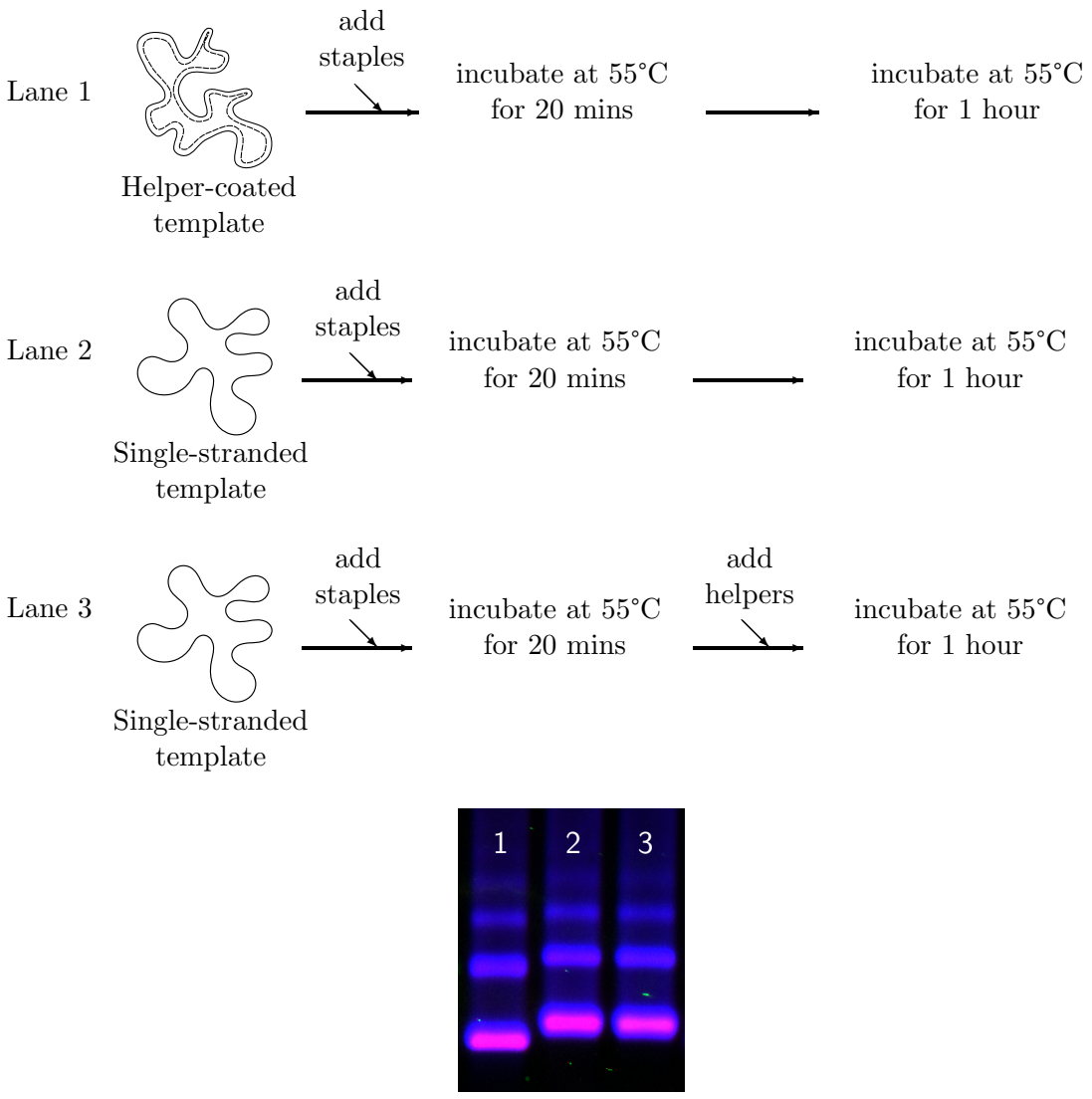


Figure 2.26: Effect of adding helper strands after template has been incubated with staples. Lane 3 is the sample of interest, lane 2 is a control where no helpers were added, and lane 1 uses helper-coated template instead of adding the helpers after the staples. Top: preparation of samples. Bottom: Gel electrophoresis of these samples. Labelled strands are the central body staple (red) and seam staple (blue) as described in section 2.3.2.1. Adding the helpers after 20 minutes made no improvement to tile folding (lane 3 is identical to lane 2), unlike the use of helper-coated template (lane 1 contains a better-folded structure than lanes 2 or 3).

2.4.5 Isothermal folding of DNA origami

In order to put the results of section 2.4.1-2.4.4 in context, I will now briefly review the alternative methods which have been developed for isothermal folding of DNA origami. The relevancy of these methods to understanding the helper-origami system is discussed in section 2.4.6.

DNA origami nanostructures are designed so that the target structure represents the minimum free energy, and as result it is the product formed at thermodynamic equilibrium. There are aspects of origami folding which favour formation of this product, in particular the co-operativity of staple incorporation (see section 1.2.2.1). Others, predominantly staple mis-binding and formation of template secondary structure, can result in misfolded structures which may represent kinetic traps if these unwanted interactions cannot be reversed (in a reasonable timescale), and so can prevent the structure reaching the target product.

Successful origami folding, whether isothermally or through thermal annealing, therefore requires limiting and/or reversing these unwanted template-staple and template-template interactions. For all origami structures, staple domains are designed so that they only hybridise fully to the target region of the template. This means that each staple usually[8] binds more strongly to the target domains than to any off-target locations. In addition, mis-bound staples are unlikely to be stabilised by co-operative binding (unlike correctly-bound staples, see section 1.2.2.1) since they have not been designed to bind simultaneously to the template. However, at room temperature this (in addition to staple co-operativity) does not provide the necessary discrimination to prevent staple mis-binding: formation of off-target template-staple interactions are still favourable at this temperature, and very slow to reverse. The use of a staple excess in origami folding is thought to improve the process of reversing these interactions (and also removal of template secondary struc-

[8]This is usually the case as any 'mis-bindings' will involve mismatches in base-pairing. However, it is possible for these interactions to be stronger than the target interactions, e.g. if a length of staple which spans more than one domain is able to bind to a region of the template with high G/C content. See section 2.4.5.4 for further discussion.

ture), through strand invasion and displacement by free staple strands [122]. Despite this, the yield of well-folded origami obtained after room-temperature incubation is generally very low, due to the formation and persistence of these interactions[9].

In thermal annealing, these issues are overcome in the following ways:

- Initially heating to 90 or 95°C removes all DNA hybridisation, including template secondary structure.
- At temperatures close to the melting temperatures of the staple-template interactions:
  (a) the difference in melting temperatures between target and off-target template-staple interactions means that formation of target interactions are significantly more favourable (i.e. this difference allows discrimination at these temperatures).
  (b) reversal of off-target interactions proceeds more rapidly than at room temperature.

Whilst the use of a thermal annealing program can overcome these problems, successful isothermal methods have also been developed for use in cases where thermal annealing is unsuitable (as described in section 1.4, e.g. when the origami is conjugated to thermosensitive hetero-molecules like proteins; with thermal annealing conjugation has to be carried out in a second step following origami assembly [74]). These are summarised below and in Table 2.2. It should be noted that most of the methods were tested on the 2D origami 'Rothemund tile' shown in Figure 2.27, which is a) larger than the tile used here (section 2.1), b) predominantly uses three-domains staples (as opposed to the two-domain staples of my tile) [122], and c) uses a larger staple excess (10-100x) [61, 74, 135] than the 2× excess which I have

[9]It should be noted that there are other potential barriers to correct folding of origami objects, in addition to the formation of unwanted interactions as described above. This is particularly true for 3D origami structures with complex scaffold routings, where the scaffold topology during the folding process must also be considered [68, 95, 98]. Since I will be using a 2D origami object I have focussed on folding of 2D origami only.

used. This point will be returned to when discussing folding of my tile in section 2.4.6.

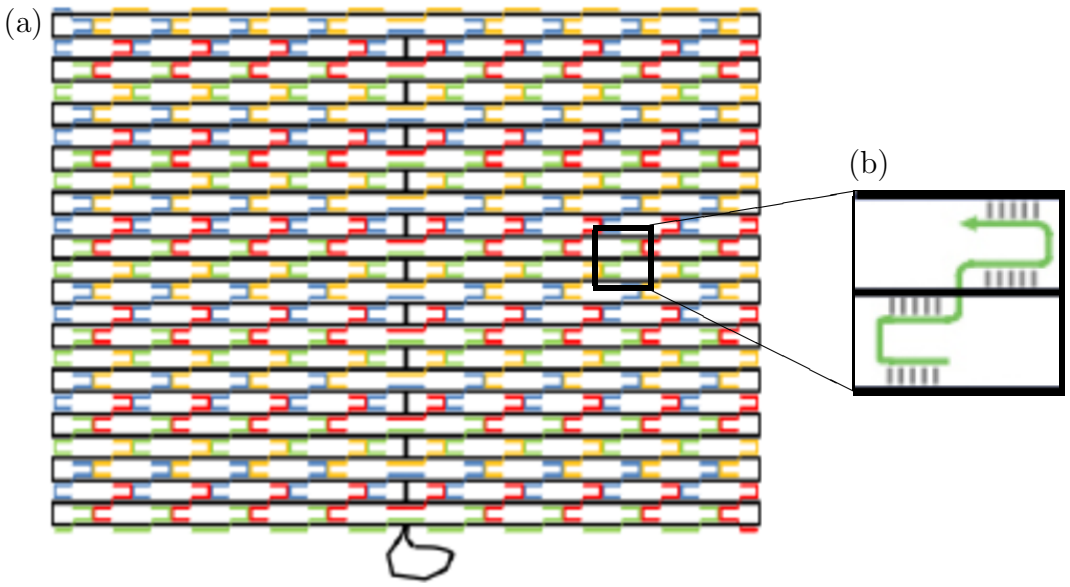


Figure 2.27: 2D origami tile designed by P. Rothemund [122]. (a) shows the full staple pattern (adapted from [3]). (b) shows the binding of an individual body staple: note the two shorter (8 nt) binding domains at either end of the staple, and the longer (16 nt) binding domain in the centre. This pattern of binding domains is repeated for all staples except those at the seam and long (top and bottom) edges. (b) adapted from [122], reprinted by permission from Springer Nature: Nature. Copyright 2006 Nature Publishing Group.

2.4.5.1 Chemical denaturants

Using a concentration ramp of a chemical denaturant has a very similar effect on the hybridisation interactions of DNA strands to a temperature ramp, and can produce similarly high yields [61]. At high concentrations of the denaturant, the DNA is completely single-stranded, and as the concentration is reduced (e.g. by dialysis [61]), base-pairing becomes more favourable. Formamide lowers the melting temperature of DNA linearly (by ∼0.6 °C per % formamide in the buffer [17]), so this would seem to be acting in an identical manner to a temperature ramp. However, further investigation has shown that a) other solvents which lower the melting temperature of DNA (e.g. ethanol and DMF) do not have a similar effect; and b) increasing the incubation temperature to simulate the melting temperature-

lowering effects of formamide produces a lower yield than when formamide is used [168]. It is therefore likely that the ability of formamide to hydrogen-bond with DNA bases is vital in allowing successful folding of the origami (and may explain why urea also works) [168]. It may be acting more as a competitive binder of both the template and staple DNA, which can be out-competed by the formation of strong template-staple interactions, disfavouring mismatches [168].

2.4.5.2 Mechanical manipulation of the template strand

In a 2014 study by Bae et al. [10], the authors posit that the most important kinetic traps encountered during origami folding are those caused by secondary and even tertiary structures formed by the template strand. They developed a method to remove these from a single (linear) scaffold molecule by mechanically stretching it using magnetic tweezers, before adding a large excess of staples. The mechanical tension was then removed, and the structure left for 5-10 minutes to allow strand displacement (to remove excess copies of multiple-copy-bound staples). This produced the correctly-folded origami structure (a 3D 10 helix bundle). Figure 2.28 shows a schematic of the process.

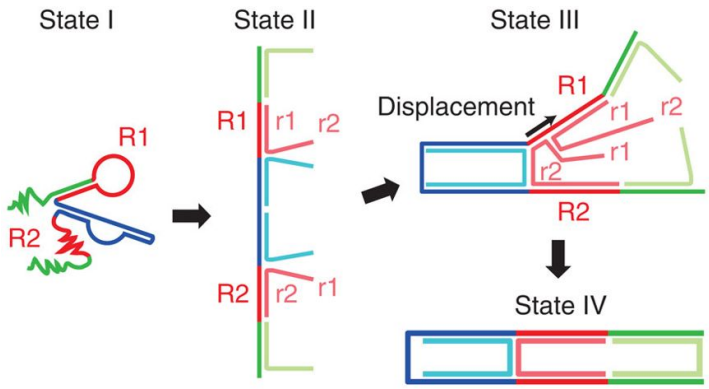


Figure 2.28: Schematic of the proposed folding process when magnetic tweezers are used to mechanically stretch the origami template strand. State I: the native secondary structure of the template; State II: secondary structure removed by the magnetic tweezers and (excess) staples added; State III: mechanical tension removed to allow strand displacement; State IV: the correctly folded structure. Figure adapted from [10]; reprinted by permission from Springer Nature: Nature Communications. Copyright 2014 Macmillan Publishers Limited.

Two-, three- and four-domain staples were used, and the schematic shows that each staple is expected to be bound via one domain only whilst the template is stretched (Figure 2.28 State II), so up to four copies of a particular staple may be bound. When the template is released, strand displacement allows one staple to replace the other copies which are bound, a process which takes less than 10 minutes at the temperature used (36°C throughout the experiment). It should be noted that when excess (unbound) staples were washed out before the mechanical tension was released, the final structure appeared to be less well-folded, suggesting that staple exchange during State III may also be an important factor.

2.4.5.3 Thermal denaturation followed by isothermal incubation

Both 2D and 3D origami structures have been shown to fold rapidly at constant temperature following initial thermal denaturation [135]. For the Rothemund tile, near 100% yield was achieved by folding at 53°C for 5 minutes following initial denaturation was at 95°C for 1 minute. The optimal folding temperature for each origami object (T_{fold}) was chosen according to the assembly temperature of the structure during thermal annealing, as shown in Figure 2.29.

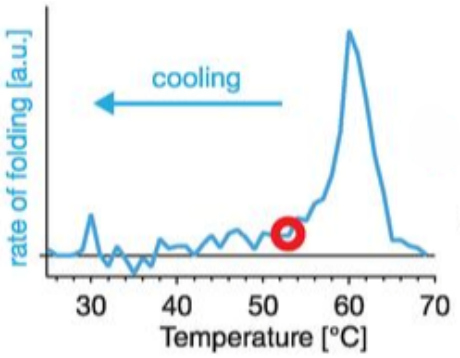


Figure 2.29: Optimal folding temperature (T_{fold}; red circle) for the Rothemund tile as determined by Sobczak et al. from the assembly temperature of the tile during thermal annealing [135]. The blue line shows the differential of the fluorescence signal, which itself is a measure of the amount of double-stranded DNA present and therefore a measure of origami assembly (see section 2.3.3). The peak in the differential plot represents the temperature range over which assembly is occurring, and T_{fold} is the lower boundary of this peak. From [135]; reprinted with permission from AAAS.

Not only is this protocol considerably quicker than thermal annealing protocols for similar objects, the yields are higher [135].

2.4.5.4 Betaine

In order to develop a method suitable for folding origami structures with components such as proteins (which none of the previous methods are: formamide and urea are damaging to proteins and other biological molecules [74], mechanical manipulation is not suitable for folding large numbers of structures, and the other method mentioned involves thermal denaturation [135]), Kopielski et al. (2015) initially tested isothermal incubations as described above without the initial denaturation step [74]. These were successful for the Rothemund tile: 80-85% yields were seen after 1 hour at 50-55°C [74]. However, at lower temperatures misfolding was observed, as evidenced by parts of the structure being two or more helices tall in AFM images [74], suggesting staple mis-binding. At higher temperatures, the structures contained holes, suggesting that some staples are not able to bind stably [74].

In order to speed-up the reversal of mis-binding interactions at lower temperatures, Kopielski et al. added betaine (trimethyl-glycine), which not only lowers the melting temperature of DNA, but equalises the melting temperature of AT and GC base pairs [116]. This reduces the chance of (staples competing with) anomalously strong-bound mis-bound staples. 1 M of betaine increased the yield from a 24 hour incubation at 37°C from $\sim$0% to $\sim$25%.

Although we do not know the exact folding process during each of these methods, from the protocols and results of each we can infer how each has managed to avoid and/or escape from the kinetic traps which can prevent correct folding. These are summarised in Table 2.2.

Method	**How is template secondary structure removed?**	**How is mis-binding of staples to template discouraged?**	**How is replacement of mis-bound staples by strand displacement accelerated?**
Thermal annealing	Heating at start	Slow reduction of temperature	N/A*
Chemical denaturation [61, 168]	High denaturant concentration at start	Slow reduction of denaturant concentration	N/A*
Mechanical manipulation of template [10]	Mechanical stretching	N/A	N/A
Isothermal with thermal denaturation [135]	Heating at start**	N/A*	N/A*
Isothermal alone [74]	N/A*	N/A*	N/A*
Isothermal with betaine [74]	N/A*	N/A*	Betaine may favour strand displacement to remove incorrectly bound staples*

Table 2.2: Various folding methods used for origami assembly, and how they avoid/overcome the unwanted interactions which can lead to the formation of misfolded structures. *NB: use of higher temperatures and/or chemically lowering the melting temperature of DNA will reduce the stability of template secondary structures; increase discrimination between correctly bound and mis-bound staples (assuming the latter are less thermodynamically stable due to mismatches) [163]; and increase the rate of strand displacement. **Template secondary structure interactions are likely to reform quickly (more quickly than staples will bind to the template) once the temperature is cooled following the initial denaturation step. However, Sobczak et al. found that for some structures, initial thermal denaturation 'presumably to resolve secondary structures in the template strand' was necessary for correct folding [135]. This suggests that the template may rearrange its secondary structure interactions over time (when it is incubated alone at e.g. room temperature) to a more relaxed state, and perhaps the fast-forming interactions are less stable than those which form later.

2.4.6 Discussion

2.4.6.1 Helpers improve isothermal folding of the origami tile

In section 2.4 I have shown that using helper strands to coat the template of an origami tile improves isothermal folding of the tile at 60 and 65°C. The results presented in section 2.4.4 suggest that helpers primarily achieve this by blocking the template and preventing unwanted interactions (which could include both staple mis-binding and formation of template secondary structure). Since folding is better than in the absence of helpers, it must be the case that the helpers are more likely to be displaced by the target staple rather than one which is mis-binding, presumably due to the difference in stability of these two interactions. In this way, the barrier imposed by the helpers is perhaps acting in a similar way to the temperature ramp used during thermal annealing or the use of formamide and other denaturants as a competitive binder of the staples and templates (see section 2.4.5.1): the proportion of incorrect staple-template interactions is reduced by increasing discrimination between target and off-target interactions.

However, section 2.23 shows that the use of helpers is not sufficient to prevent formation of all types of misfolded structures: cross-linked multimers are formed and account for 10-15% of the product formed in 1 hour incubations at 60 or 65°C[10]. These multimers may form when more than one copy of a staple is bound to the same template strand (through different domains), preventing each other from fully binding to the template. One of the partially bound staples may then bind to another template in order to satisfy its base-pairing potential. Snodin et al. (2016) suggested that binding of multiple copies of a single staple to the template could be prevented by using staples with one binding domain longer than the rest [134] as shown in Figure 2.30.

In my experiments, these cross-linked dimers persist in conditions where the majority product is well-folded monomeric tiles (e.g. when the helper-coated tem-

[10]Calculated from pixel values of lanes 2 and 3 of Figure 2.16. Only bands representing products containing a template strand were included.

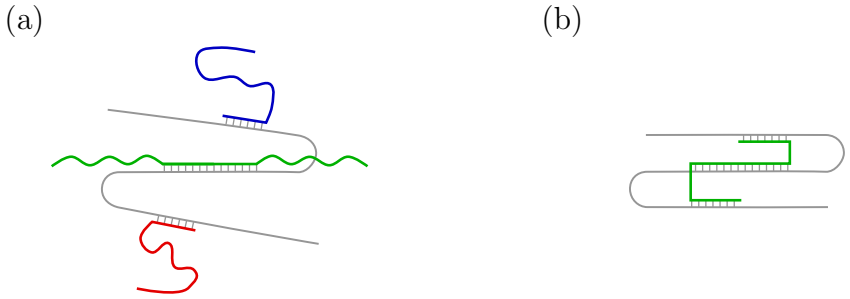


Figure 2.30: Potential binding of multiple copies of a single staple to a template strand. The staple shown here has three domains in a 8-16-8 nt pattern. This could allow up to three copies of the staple to bind to the template at once (a). However, the green copy, which can bind via the long central domain, is much more likely to be bound to the template at the temperatures used for isothermal incubations, resulting in correct binding of a single staple copy as in (b). Preventing multiple copies of the same staple binding to the template strand may prevent formation of cross-linked multimers.

plate is incubated with staples at 60°C as shown in Figure 2.23), suggesting that the staple-template interactions which form these cross-link are the most difficult type of staple mis-binding to remove. Removal of these cross-link interactions would require either melting or a 'remote-toehold'-type displacement [39].

It should be noted that in the optical tweezer experiments carried out by Bae et al. (Figure 2.28, section 2.4.5), strand displacement of extra staple copies bound to the template occurred correctly (i.e. to leave only one copy of each staple bound) in less than 10 minutes. This may be a feature of the particular domain lengths of the staples, or the very large staple excess used.

The formation of the cross-linking interactions is likely to occur more slowly than the binding of both staple domains within a single template (due to the closer proximity of a second domain on the same template compared to one on a different template), and so the formation of these cross-linked multimers is limited to 10-15% of the total product.

2.4.6.2 Isothermal folding of the tile in the absence of helpers

In section 2.4.1, I showed that one hour isothermal incubations of single-stranded template (no helpers) with staples at 60 and 65°C produce worse-folded tiles (as

shown by the gel electrophoresis results) than folding at 60°C using helpers. As described in section 2.4.5.3, Sobczak et al. [135] suggest that the optimal folding temperature for isothermal folding of an origami object (T_{fold}) is ∼7°C below the melting temperature (see Figure 2.29), which would be 62°C for the tile here. (NB: the paper referenced uses origami structures containing three-domain staples as opposed to the two-domain staples used here. In addition, Sobczak et al. suggest an initial denaturation step which I did not replicate, although this was demonstrated to be unnecessary by Kopielski et al. [74] as described in section 2.4.5.4). However, a one hour incubation of single-stranded template with staples at 62°C produced worse folding than at 65°C (as determined by gel electrophoresis results, data not shown).

Using incubation temperatures of over 65°C produced gel bands which were indistinguishable from the well-folded tile produced by thermal annealing (Figure 2.31). The secondary band (indicated by an arrow in Figure 2.31), which is thought to predominantly contain cross-linked dimers (see section 2.4.2.1) disappears at 68°C, presumably due to the fact that at this temperature these interactions are able to melt, allowing monomer tiles to fold.

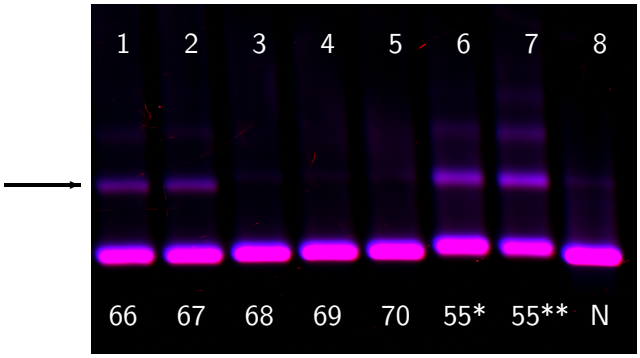


Figure 2.31: Gel electrophoresis results of tile folding without helpers. Numbers at the bottom of the gel refer to temperatures in °C for 1 hour incubations except: *1 hour 20 mins, **4 hours, N = annealed. Labelled strands are the body staple (red) and seam (blue) as described in section 2.3.2.1. Arrow to the left indicates the secondary band which is thought to predominantly contain cross-linked dimers (see section 2.4.2.1). This band is not present at incubation temperatures of 68°C and higher.

Indeed, heating to temperatures close to the melting temperature of the tile

allows the reversal of all types of unwanted interactions (staple mis-binding and template secondary structure), as in thermal annealing. The successful folding seen at 66-70°C in Figure 2.31 may predominantly be due to this ability to reverse unwanted interactions at these temperatures. (In addition, samples were cooled in a 4°C fridge between the isothermal incubation and running on a gel, which may mimic the cooling ramp used in thermal annealing.)

Folding of the tile at temperatures below 66°C (without helpers) is significantly worse than that seen for other origami structures in experiments by Sobczak et al. [135] (see section 2.4.5.3) and Kopielski et al. [74] (see section 2.4.5.4). This may be because most of the staples in the tile used here have two-domains of equal length (15/16 nt). In contrast, most conventional origami structures, including the 'Rothemund rectangle' (Figure 2.27) used by Sobczak et al. and Kopielski et al., primarily use three-domain staples with domain lengths 8-16-8 nt. This may improve isothermal folding in the following ways:

- Increasing the number of binding domains per staple may lead to increased co-operativity between staples, since the incorporation of each staple will reduce the conformational entropy of the template by a greater extent.

- The low melting temperatures of the short (8 nt) domains mean they can be rearranged (i.e. their binding to the template can be reversed and remade) at lower temperatures than 15/16 nt domains. This may allow removal of mis-binding interactions at lower temperatures.

- In designs which use three-domain staples, each link between adjacent helices of the origami structure is formed from two staple 'crossovers' (except those at the ends of helices), as opposed to one in the tile used here (see Figure 2.32). This is likely to result in greater co-operativity between the staples, in a similar manner to that seen between the seam staples in the tile design used here [34].

- Using domains of different lengths allows removal of multiple copies of one

staple, as in Figure 2.30.

(a) Three-domain staples (template in black)

(b) Two-domain staples (as used here; template in dark grey)

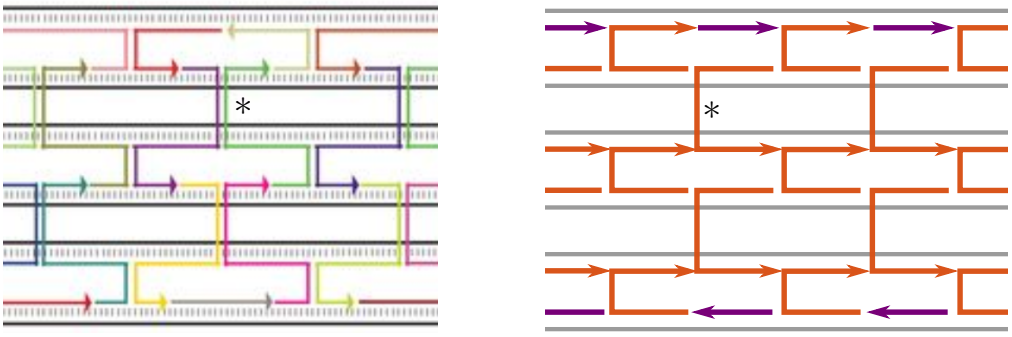


Figure 2.32: Staple crossovers in different origami designs. (a) A conventional origami which uses three-domain staples. Each link between adjacent helices is formed from two staple crossovers (e.g. at *where the purple and green staples are crossing-over). (b) The two-domain staple design used in this chapter. Each link between helices is formed from only one staple crossover (see *). a) from [50]; reprinted with permission from AAAS.

In conclusion, helpers have proved an effective method to improve isothermal folding of our tile. Without helpers, isothermal folding of the tile is poorer than for conventional origami (perhaps due to the particular design chosen). We have not tested the use of helpers in conventional origami designs (i.e. those which use three-domain staples), where they might allow folding at even lower temperatures/in shorter timescales than those seen by Sobczak et al. [135] and others.

In addition, the use of helpers may allow additional control over folding: by coating the template with helpers, but omitting helpers at specific positions, it may be possible to control where folding (i.e. staple incorporation) begins, and how quickly it propagates. This is explored in the following section (2.5).

2.5 Omitting selected helper strands further improves folding

As described at the end of section 2.4.6, in addition to improving isothermal folding of the origami tile, the use of helpers may allow kinetic control over the folding process. By coating the template with helpers, but omitting helpers at specific positions, it may be possible to control where folding begins (nucleation), and how quickly it propagates. Furthermore, it may be possible to 'prime' a template with helpers in a way that favours folding with a particular staple set, thus allowing folding of a particular target structure from a mixture of different staple sets.

We began by investigating how omission of selected helpers could favour folding with the tile staple set used previously. We focussed on favouring a) nucleation of folding, i.e. binding of the first staple domain (section 2.5.1), and b) propagation of folding, i.e. incorporation of further staple domains (section 2.5.2). In the latter section, I show that omission of helpers at or near the seams of tile allows folding of the tile at lower temperatures/shorter timescales than we have seen when the full helper set is used (an observation initially seen in simulations, and confirmed here by experimental results). We also found that the effect of omitting a particular helper is very dependent on the sequences of the staples and helper strands. This will be discussed further in section 2.6.

Unless otherwise stated, all results in this section (both experimental and from simulations) use the 10-offset helper set.

2.5.1 Nucleation

When the template is fully coated with helpers, the first staple domain can only bind via displacement of a helper. As Figure 2.33a shows, this results in a gap of single-stranded template, due to the offset between helper and staple binding domains. This step therefore reduces both the number of base pairs in the structure, and the opportunities for coaxial stacking interactions. As a result, the structure shown

on the right in 2.33a is less stable than the fully helper-coated template; this has been confirmed by simulations using average sequences. Binding of a second staple domain in the gap (Figure 2.33b) restores the full base-pairing of the template and opportunities for stacking, but this structure is still less stable than the fully helper-coated template (again, confirmed by simulations using average sequences), due at least in part to the reduction of entropy associated with increasing the number of strands bound to the template.

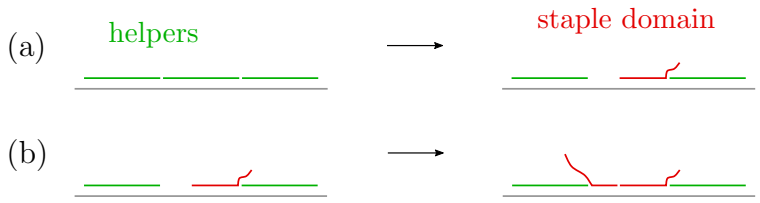


Figure 2.33: Binding of first staples to the helper-coated template. (a) Displacement of the first helper allows one staple domain to partially bind to the template, but results in a gap of single-stranded template. (b) A second staple domain can bind in the gap.

Due to the factors discussed above, binding of the first staple domain to the fully helper-coated template will only occur if the staple binding domain has a considerably higher melting temperature (i.e. binds more strongly to the template) than the helper strand with which it is competing for binding to the template. The binding domains of the staple and competing helper overlap but are offset to one another (see section 2.2), and so the difference in their melting temperatures depends on the sequence of non-overlapping region, i.e. the toehold. Simulations using the 'real' sequences of the template, staples and helper strands (i.e. those used in the experiments, see Appendix B.1) have confirmed that folding of the fully helper-coated template begins at the position where the staple binding domain is strongest relative to the helper with which it is competing for binding to the template.

The fact that nucleation is sequence-dependent means that we cannot change where it occurs in our system when the template is fully coated with helpers. (In future experiment, simulations could be used to design origami structures with sequences in which nucleation begins at a prescribed position; we have not tested this.)

As mentioned previously, omission of a single helper produced identical results to the use of the fully helper-coated template, as shown in Figure 2.34.

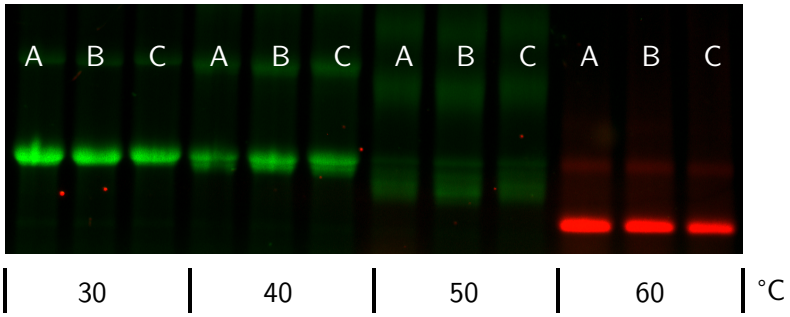


Figure 2.34: Gel electrophoresis of helper-coated template incubated with tile staples at various temperatures for 1 hour. Various subsets of helper strands were omitted: A: fully-helper coated template (no helpers omitted); B: one helper omitted; C: three consecutive helpers omitted. Green channel: labelled helper; red channel: labelled body staple (see section 2.3.2.1). As explained in section 2.3.2.2, differences between these results and those shown in Figure 2.16 are likely to be due to differences in the time period between incubation and running the gels.

Simulations at 50°C using the experimental sequences showed that omission of a single helper results in multiple initiation sites: both the site where the helper was omitted, and the nucleation site of the fully helper-coated template (i.e. the position at which there is the biggest favourable difference in binding energies between competing staple and helper domains). Since we were unable to introduce new nucleation sites through helper omission, we decided to focus on propagation of folding.

2.5.2 Propagation

We have seen in section 2.4.2 that specific intermediates are formed during folding of the helper-coated template (for example intermediates A and B in the first lane of Figure 2.35 below). This suggests that there is a specific pathway by which folding is propagated, i.e. a particular order in which staples are incorporated. The helper-origami system was designed with an offset between the helper and staple domains, so that once one staple domain is partially bound to the template, it is able to

displace the adjacent helper by toehold-mediated strand displacement (Figure 2.36). We would therefore expect helpers to be displaced sequentially, following the path of the template strand. Simulations at 50°C confirmed that the order in which staples domains are incorporated follows the path of the template strand.

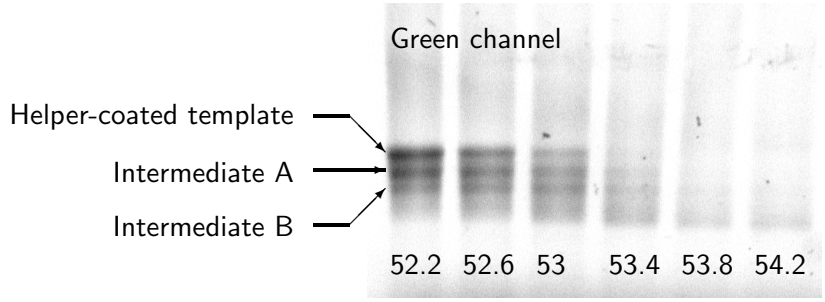
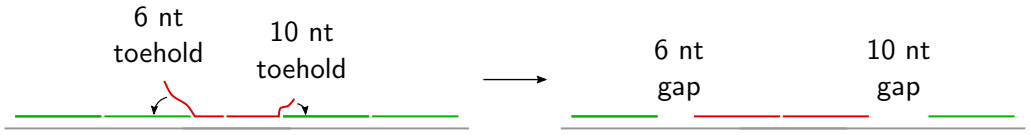


Figure 2.35: Gel electrophoresis of helper-coated template incubated with tile staples at various temperatures (shown in °C) for 1 hour. Only the green channel (labelled helper, see section 2.3.2.1) is shown. At the lower temperatures shown, a number of distinct intermediates are present. This figure is a repeat of Figure 2.20.



helpers in green, staple domains in red

Figure 2.36: Further displacement of helpers following nucleation (follows on from Figure 2.33). Partially-bound staple domains, occupying the template binding domain of a single helper, are able to displace adjacent helpers through toehold-mediated strand displacement. In the 10-offset helper set, one staple will have a 6 nt toehold, and the other a 10 nt toehold.

For the 10-offset helper set, each staple domain has a toehold of either 10 or 6 nt, as shown in Figure 2.36. We know that at 25°C, toehold-mediated strand displacement reactions using toeholds of these lengths proceed at the same rate as each other [137], but we do not know the relative rates at the temperatures which we are using for the isothermal incubations. Simulations at 50°C show that staple incorporation propagates out in both directions from the nucleation site, suggesting that there is not a significant difference in the rate of 6 nt-toehold and 10 nt-toehold mediated displacements at 50°C.

Simulations also showed that omission of helpers which have overlapping binding domains with the seam staples (see Figure 2.37) reduced the folding time. At each seam cross-link, there are six helper strands which have overlapping binding domains with the two seam staples, as shown on the right in Figure 2.37. I have divided these into three classes: 'at the seam' helpers (orange in Figure 2.37) each overlap with both of the seam staples; '5' seam-adjacent' helpers (blue) each overlap with only one seam staple and are 5' to the seam (using the direction of the template strand); '3' seam-adjacent' helpers (purple) also each overlap with only one seam staple and are 3' to the seam. Simulations showed that omission of either the 'at the seam' subset, or both of the 'seam-adjacent' subsets, reduced the folding time at 50°C.

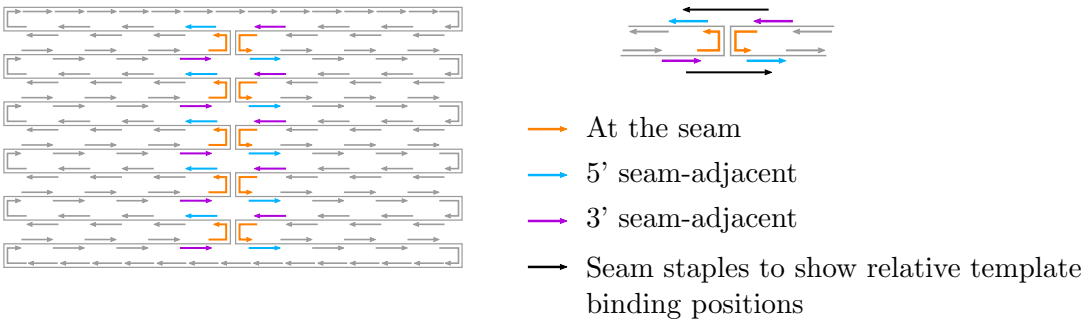


Figure 2.37: Helpers which have overlapping binding domains with the seam staples (these helpers are in blue, orange and purple). Their binding positions relative to the seam staples are shown on the right. '5" and '3" seam-adjacent labels use the direction of the template strand. In simulations, omission of either the orange set, or the blue and purple sets together, reduced folding time at 50°C.

These simulation results suggest that for the fully helper-coated template, folding is usually 'held up' at the seam[11]. At temperatures where we see partially-folded intermediates (e.g. 52.2°C in Figure 2.35), propagation of folding may have stopped when it reaches a seam. Further staples are not incorporated unless the incubation time or temperature is increased. We therefore identified the helpers which bind at or adjacent to the tile seam ('seam helpers') as suitable candidates to allow us to control propagation of folding in the helper-coated template. The following section

[11]It is not obvious why propagation at the seams should be particularly unfavourable; see section 2.5.4 for our hypothesis.

contains the experimental results when subsets of these helpers were omitted.

2.5.2.1 Real-time FRET measurements when seam helpers are omitted

Temperature-controlled real-time FRET emission (as described in section 2.3.3) was used to observe the effect of omitting helpers at or adjacent to the seam. The FRET-labelled staples do not bind close to the seam (Figure 2.38a), and so allow us to see the impact of omitting seam helpers on folding of other parts of the tile. This method is more suited to high throughput analysis than gel electrophoresis.

Helper-coated template and staples were mixed and incubated at 40°C for 10 minutes, and then heated at a rate of 0.5°C per 5 minutes to 80°C, before being cooled back to 40°C at the same rate. This temperature range was selected as we have previously seen (e.g. in Figure 2.18) that the full folding process occurs within this range.

Figure 2.38b shows the results from a control sample where the fully helper-coated template (using the 10-offset helper set) was used. During the heating step, there is an increase in FRET signal between ∼50 and 55°C which correlates to the folding of the helper-coated template into the tile. At ∼64°C, the tile melts, producing a sharp drop in FRET signal. During cooling, the tile reforms at approximately the same temperature at which it melted (there is some hysteresis due to staple co-operativity), and the tile remains present as it is cooled further. The changes in fluorescence at ∼64°C during heating and cooling are identical to those seen in the same experiment when helpers are not used (see Figure 2.15), showing that the presence of helpers does not affect the melting temperature of the tile. As previously, the melting temperature as measured by FRET (∼64°C) is lower than when Eva Green fluorescence is used (∼69°C, see Figure 2.14). This may be due to the region containing the FRET staples folding at a lower temperature to other parts of the tile.

It is not clear why the FRET signal falls during initial heating from 40 to 50°C, but this change was replicated when the staple set alone (including the FRET pair)

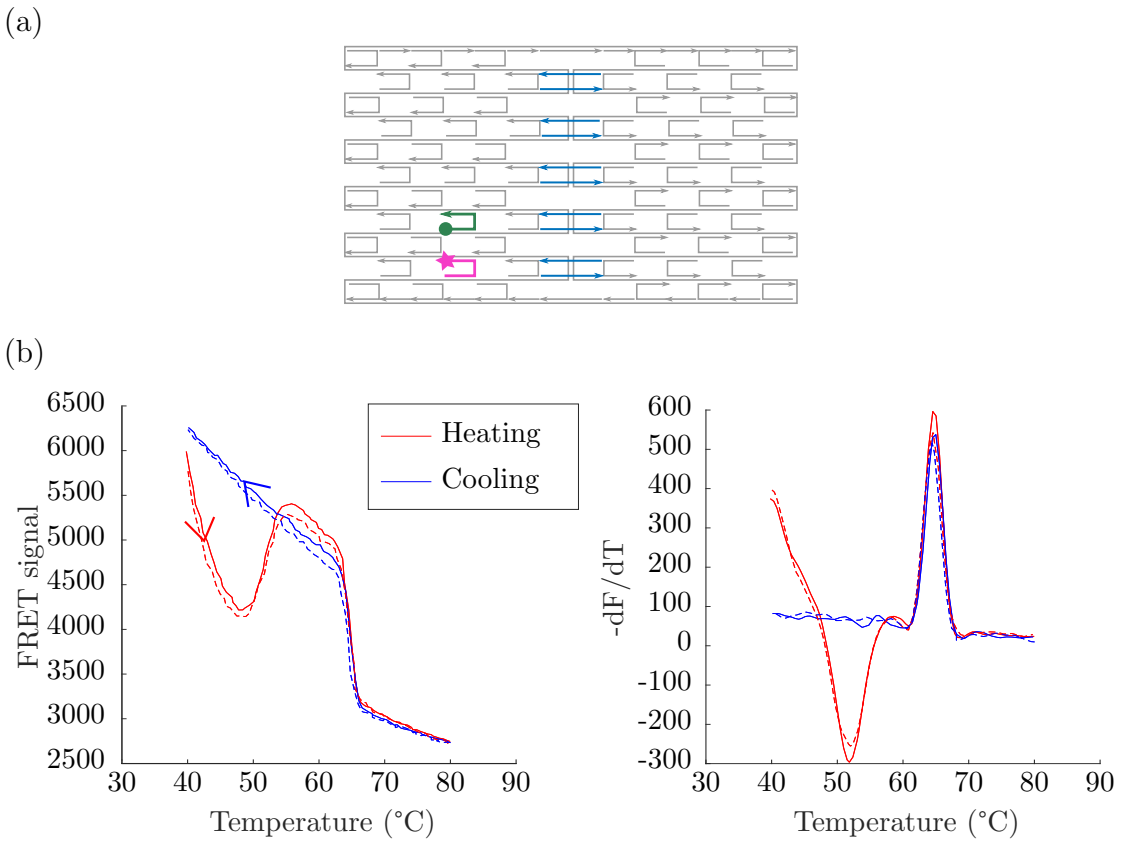


Figure 2.38: Folding of the tile from the fully helper-coated template, as observed by FRET signal. (a) Position of FRET labelled staples (pink = donor, green = acceptor). Seam staples are in blue. (b) Both the FRET signal (left) and the negative of its derivative with respect to temperature (right) are shown. Dashed lines show a second repeat. Left-hand plot: arrowheads indicate that heating step takes place prior to cooling step. The increase in fluorescence when heating from ∼50-55°C represents tile formation from the helper-coated template. The fall at ∼64°C is mirrored during cooling and represents melting and re-formation of the tile. The tile remains present as the sample is cooled further to 40°C.

were incubated at 40°C (data not shown), so does not appear to be due to interactions between the staples and template.

The negative differential of the red (heating) line gives a value for the transition temperature from helper-coated template to folded tile (the minimum value of the plot; ∼52°C in Figure 2.38). This allows easy comparison between samples, as in Figure 2.39 where subsets of helpers at or adjacent to the seam (Figure 2.39a) have been omitted.

The results show that for the 10-offset helper set, omitting helpers 3'-adjacent to the seam had a negligible effect on the transition temperature, whereas omitting either those at the seam or 5' seam-adjacent reduced the transition temperature by ∼5°C. We hypothesised that the dramatic difference between the results of omitting 5' or 3' seam-adjacent helpers could be a result of the asymmetric offset between the helper and staple domains. Further discussion of how and why this asymmetric offset could lead to the differences seen in the FRET results is not included in this thesis, because subsequent validation of the FRET results using gel electrophoresis showed that there was no significant difference between omission of 5' and 3' seam-adjacent helpers, as explained in the following section.

2.5.2.2 Gel electrophoresis results when seam helpers are omitted

Helper-coated template, with subsets of seam helpers omitted, was incubated for 1 hour with staples at a range of temperatures, and then subjected to gel electrophoresis (positions of omitted helpers are shown in Figure 2.40, results in Figure 2.41). These experiments were done in order to confirm that the results of the FRET experiments shown in the previous section a) were applicable to whole tile folding (the FRET signal measures folding in the region where the labelled FRET pair of staples bind, so may not be indicative of overall tile folding); and b) were replicated in isothermal incubations.

As a control, helper-coated template with 10 non-seam helpers omitted (the same number as in each seam helper subset) was also tested (positions of omitted

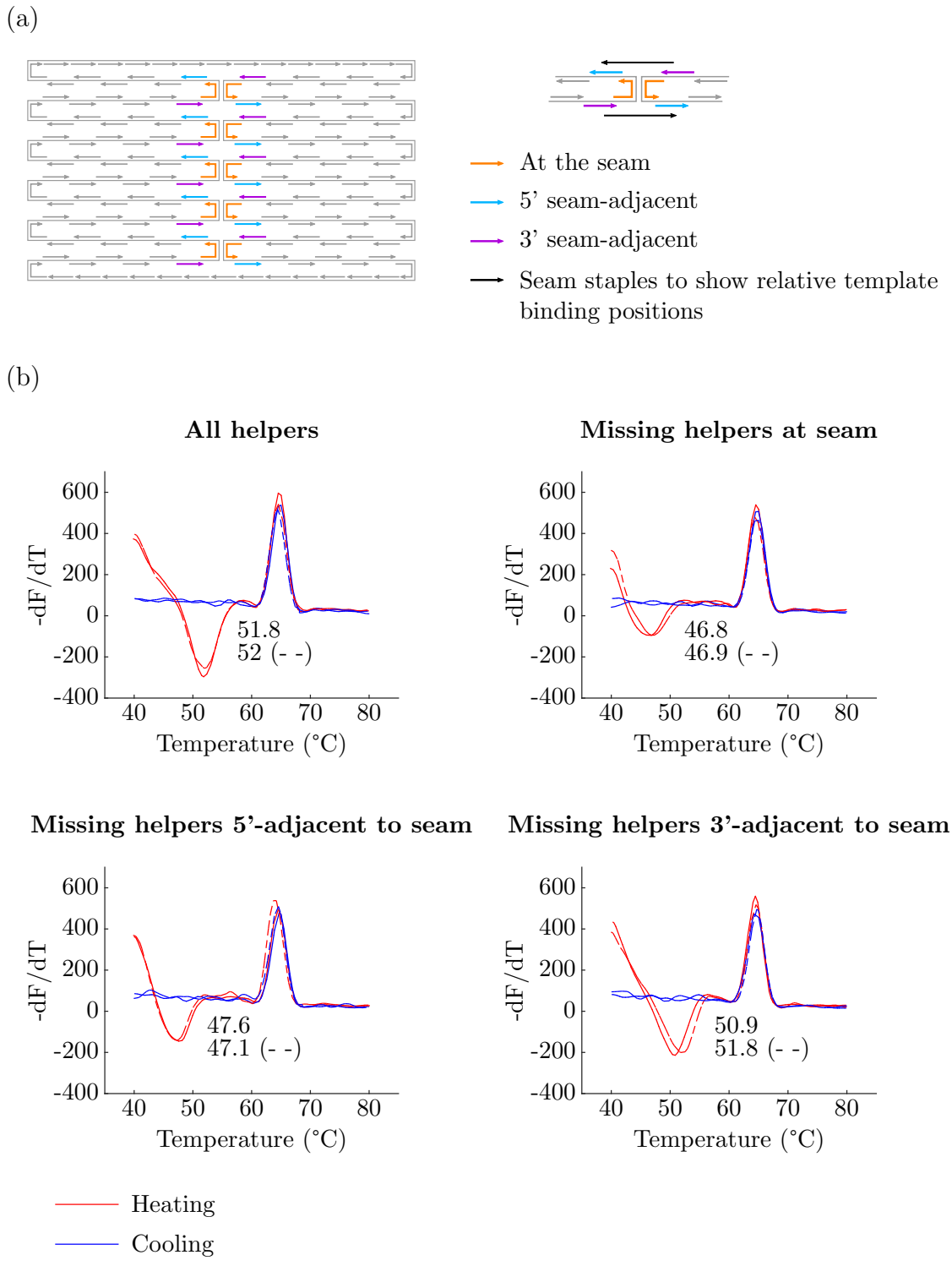


Figure 2.39: Effect on tile formation of omitting helpers at or adjacent to the seam. (a) shows the position of omitted helpers, (b) shows derivatives (with respect to temperature) of FRET traces during heating (red) and subsequent cooling (blue) of the helper-coated template with staples. Dashed lines show a second repeat. The minima of the red traces are shown; (- -) indicates minimum of dashed line.

helpers are shown in Figure 2.40, results in Figure 2.41b), and the results using the full helper set are also shown for comparison (Figure 2.41a); please note the different temperature range used for these gels compared to the gels with omitted seam helpers.

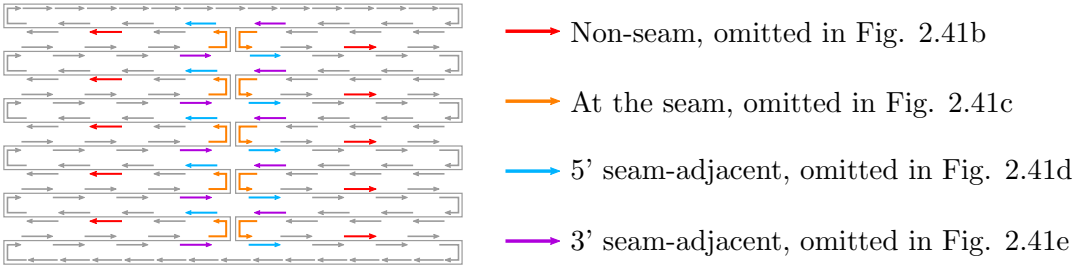


Figure 2.40: Positions of omitted helpers used in Figure 2.40.

The most striking observation is that despite the difference in transition temperature (∼5°C) observed during the temperature-ramp FRET experiments between samples where 5' seam-adjacent and 3' seam-adjacent helpers were omitted (Figure 2.39c and d), the difference in transition temperatures seen in the gels is much smaller (Figure 2.41d and e). In the gels, the first temperature at which a structure similar to the folded tile appears is ∼51°C when 5' seam-adjacent helpers are omitted vs 52°C for omission of 3' seam-adjacent helpers; a much smaller difference of 1°C (and within the margin of error due to experimental variation, see section 2.3.2.2).

This discrepancy between FRET and gel results must mean that omission of 5' seam-adjacent helpers changes folding in the region where the FRET pair binds to a much greater extent than omission of the 3' seam-adjacent helpers. It is not clear why this should be the case. The gel results show that different intermediates are produced when the 5' or 3' seam-adjacent helpers are omitted, for example at 45°C, there appears to be single type of structure present when 5' seam-adjacent helpers are omitted (Figure 2.41c), whereas there are two distinct bands (one of which contains more of the labelled staple than the other) present when 3' seam-adjacent helpers are omitted (Figure 2.41d). The formation of different intermediates shows

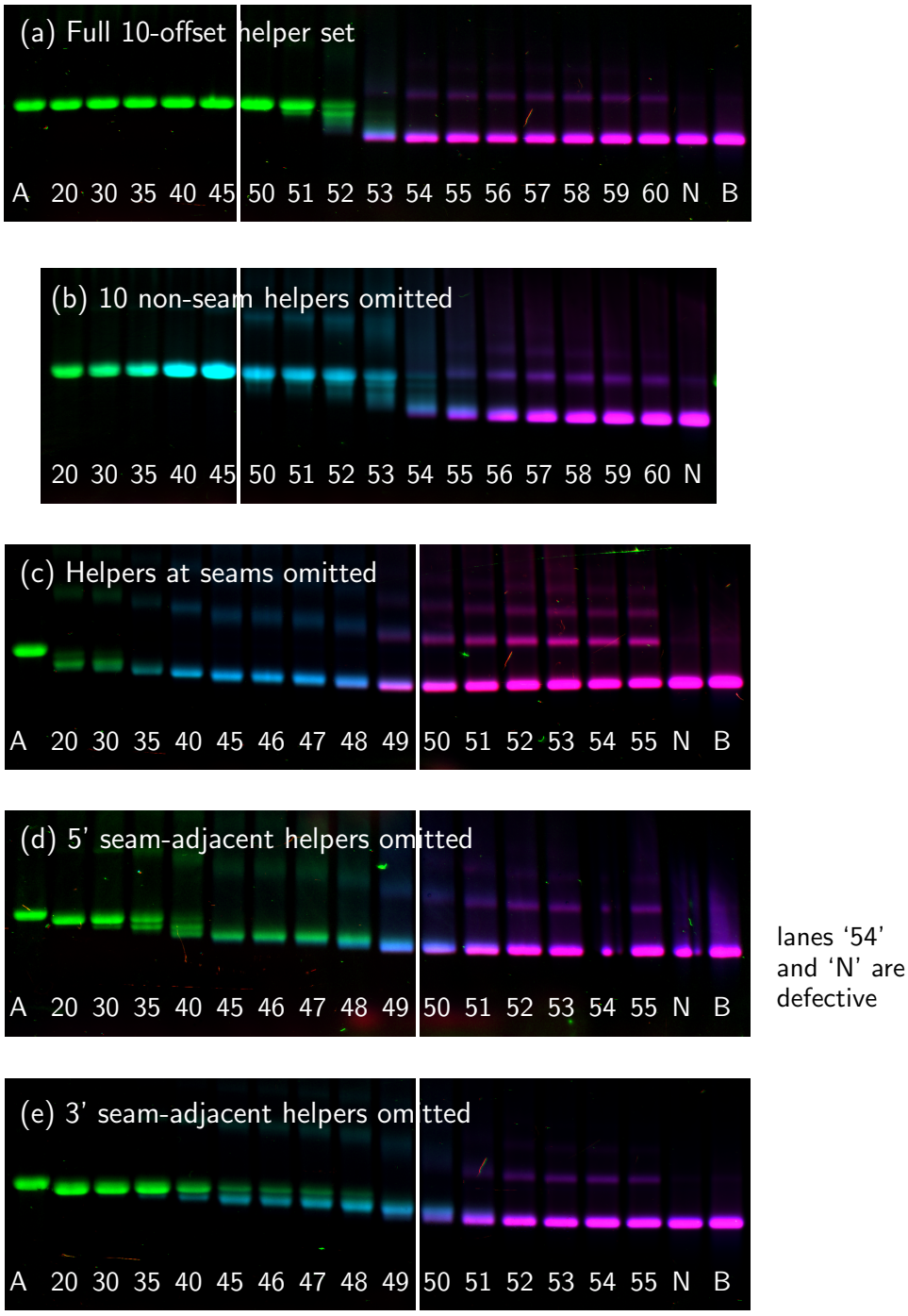


Figure 2.41: Gel electrophoresis of helper-coated template (with no helpers omitted (a), 10 non-seam helpers omitted (b), or helpers at or adjacent to the seam (c)-(e) omitted) incubated with staples at various temperatures for 1 hour, or annealed with staples (lanes labelled N). All results are for the 10-offset helper set; See Figure 2.40 for positions of omitted helpers. The white line to the left of the 50°C lane is to aid comparison between the gels. The labelled strands used are the labelled staple (red), labelled seam (blue) and labelled helper (green) as described in section 2.3.2.1. (a) was previously shown as Figure 2.18.

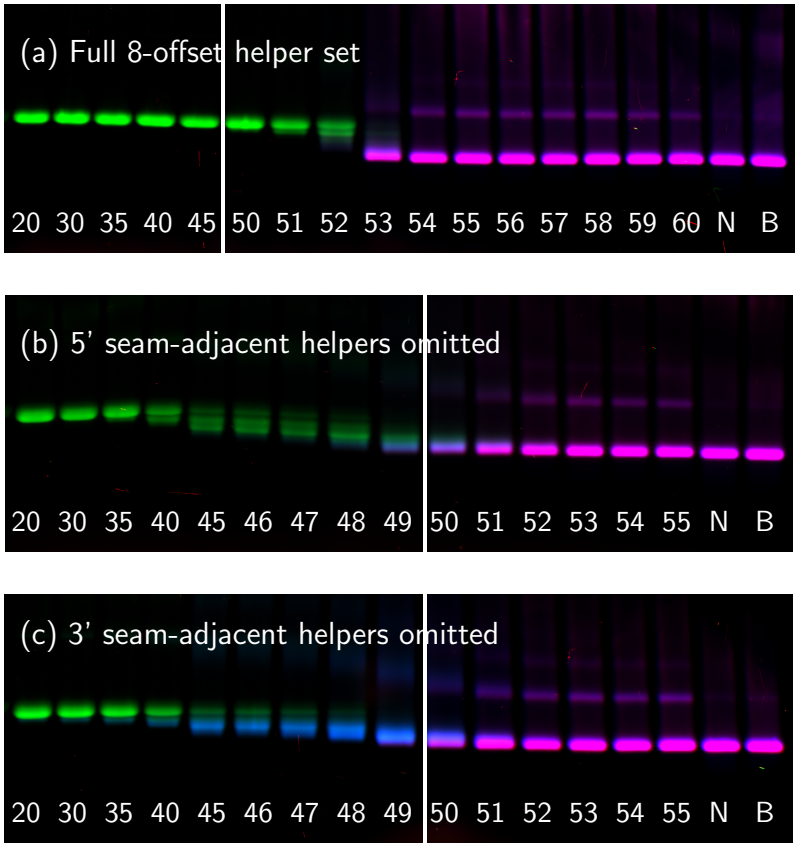
that the folding pathway differs depending on whether 5' or 3' seam-adjacent helpers are omitted. The FRET results suggest when 5' seam-adjacent helpers are omitted, folding of the area of the tile containing the FRET-labelled staples occurs earlier in the folding pathway than when 3' seam-adjacent helpers are omitted.

The folding pathways when 5' or 3' seam-adjacent helpers are omitted may differ due to either a) the asymmetric offset of helper and staple domains (which results in the 5' and 3' seam-adjacent helpers overlapping with the seam staples to different extents), or b) sequence differences between the omitted helpers and the staples with which they are competing. To test point a), the gels shown in Figure 2.41d and e were repeated using the 8-offset helper set, which has a symmetric offset with the staple domains (see Figure 2.4). These are shown in Figure 2.42; the results using the full helper set are again shown for comparison but use a different temperature range. Similarly to when the 10-offset helper set was used, the intermediates formed between 35 and 49°C differ depending on whether 5' or 3' seam-adjacent helpers are omitted. Since the 8-offset helper set have a symmetrical offset with the staple domains, this suggests that the sequence of the omitted helpers does significantly affect the folding pathway. This will be discussed further in the conclusion section of this chapter (section 2.6).

Returning to Figure 2.41 (which uses the 10-offset helper set), the results also show that omission of 10 non-seam staples does not lower the transition temperature at which the tile folds (Figure 2.41b). Despite the fact that the (blue) labelled seam is able to bind at significantly lower temperatures (more than can be accounted for experimental variation as explained in section 2.3.2.2) to when the full helper set is used, the electrophoretic mobilities of the products show that tile formation occurs at approximately the same temperature (within the error due to experimental variation). This confirms that the advantages of omitting seam/seam-adjacent helpers are not simply due to creating multiple (10) nucleation sites.

The gel results clearly show that omission of any of the three subsets of seam helpers (in the 10- or 8-offset helper set) allows folding of the tile at even lower

8-offset helpers



All numbers refer to temperatures in °C for 1 hour incubations

Figure 2.42: Gel electrophoresis of helper-coated template (with no helpers omitted (a), or helpers adjacent to the seam (b) and (c) omitted) incubated with staples at various temperatures for 1 hour, or annealed with staples (lanes labelled N). All results are for the 8-offset helper set; omitted helpers are in equivalent positions to those omitted for the 10-offset set (see Figure 2.40). The white line to the left of the 50°C lane is to aid comparison between the gels. The labelled strands used are the labelled staple (red), labelled seam (blue) and labelled helper (green) as described in section 2.3.2.1.

temperatures to when the fully helper-coated template is used (this difference is more than can be accounted for experimental variation as explained in section 2.3.2.2). These results were confirmed by AFM imaging, as shown in the following section (2.5.2.3).

2.5.2.3 AFM results when seam helpers are omitted

Figure 2.43 shows AFM images of template incubated with staples at 55°C for 1 hour using a) no helpers, b) fully helper-coated template, and helper-coated template with seam helpers omitted (c: at the seam, d: 5' seam-adjacent, e: 3' seam-adjacent).

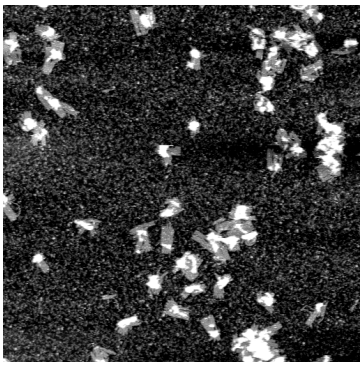
These results confirm that omission of any of the three subsets of seam helper allows folding of the tile at lower temperatures (in 1 hour incubations) than when the full helper set is used. As described in the introduction of this chapter, one of our aims was to use the helper-origami system to allow rapid folding of two distinct structures from two different subsets of staples in a one-pot reaction. The following section describes how omission of seam helpers may be used to achieve this.

2.5.3 Using helpers to fold multiple origami structures from a mixture of different staple sets

As shown in Figure 2.44, it is possible to design two origami tiles, which use the same template but different staple sets, and differ in the position of the seam in regards to the template sequence. Omitting helpers at the seam of one tile would predispose the helper-coated template to fold with one staple set over the other, and vice versa. This could allow successful competitive-annealing experiments (where both staple sets are used together) if the difference in transition temperature when seam helpers are or are not omitted is sufficient to result in large differences in yield. (This would presumably be very dependent on the incubation temperature used for the experiment.) I have not yet tested this experiment.

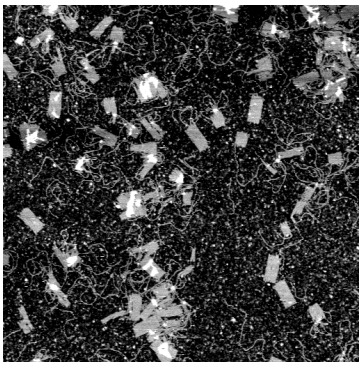
The two tile designs shown in Figure 2.44 use the same staple domains as each other, and the template regions are drawn assuming the helpers are offset by 8 nt

(a) 55°C, no helpers



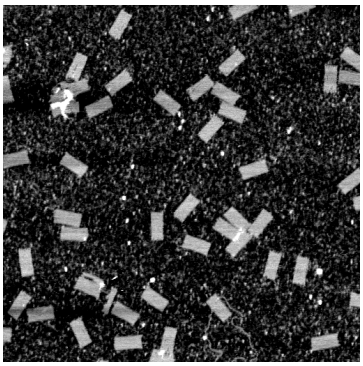
mean area: 1932 nm^2
mean boundary length: 339 nm

(b) 55°C, fully helper-coated template



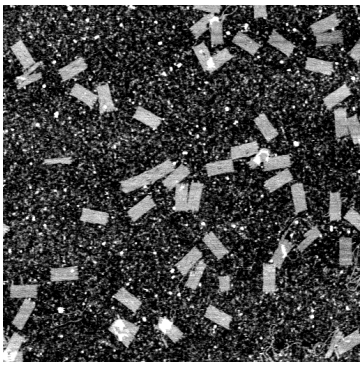
mean area: 3011 nm^2
mean boundary length: 677 nm

(c) 55°C, helpers at seams omitted



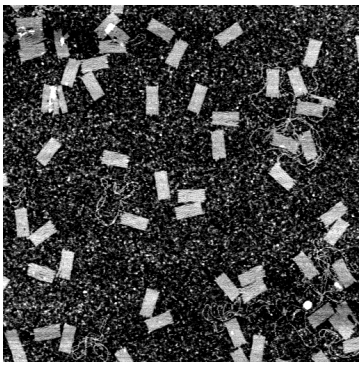
mean area: 3017 nm^2
mean boundary length: 325 nm

(d) 55°C, 5' seam-adjacent helpers omitted



mean area: 3176 nm^2
mean boundary length: 379 nm

(e) 55°C, 3' seam-adjacent helpers omitted



mean area: 2874 nm^2
mean boundary length: 362 nm

Figure 2.43: AFM images of template incubated with staples at 55°C for 1 hour using (a) no helpers, (b) fully helper-coated template, and helper-coated template with seam helpers omitted (c): at the seam, (d): 5' seam-adjacent, (e): 3' seam-adjacent). (c)-(e) show superior folding to (b) (as confirmed by the reduction in boundary length, corresponding to the reduction in helper-template duplex. Each image area is 1×1 μm. Images have been processed and analysed as per section 2.3.1.1 and Appendix A.6. (a) is the image in Figure 2.17e; (b) is the image in 2.17a and 2.25a.

Tile 1: Tile 2:

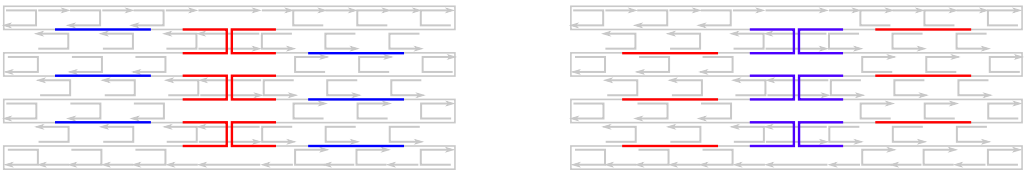


Figure 2.44: Design of two tiles which use the same template but have different staple connectivity. Template regions in red are where seam helpers bind for tile 1; the same regions for tile 2 are shown in blue on both templates. Omission of helpers in the red regions would predispose the template to fold with tile 1 staples; omission of those in the blue regions would favour tile 2.

from these domains (i.e. the same as the 8-offset set used here, see Figure 2.5). It may be preferable to use staple domains for the two tiles which are offset to each other (unlike the example shown in Figure 2.44), to reduce the formation of mixed or chimeric structures (i.e. those containing staples from both tiles) as seen in the competitive annealing experiments carried out by Majikes et al. [92].

2.5.4 Why does omitting helpers at the seam favour tile folding?

This section contains our hypothesis as to why omission of seam helpers favours tile folding more than omission of helpers at other locations on the template. The simulation results discussed below refer to folding for the fully helper-coated template using the 10-offset helper set.

Simulations using the average sequences at 45 and 50°C show that during propagation (i.e. displacement of helpers by staple domains), binding of the second domains of staples ('loop closures') are thermodynamically favourable steps. This is because when a staple is already bound to the template via one of its domains, the reduction in entropy of the template (and staple) associated with binding of the second domain is more than offset by the increase in entropy of the helper which is released. Overall, the transition from fully helper-coated template to folded tile is thermodynamically favourable because of these steps: the favourable entropy change

associated with releasing 168 helper strands more than compensates for both the binding of 90 staple strands and the reduction in template entropy due to folding of the structure (mediated by the staple connectivity).

As explained at the beginning of section 2.5.2, these simulations confirmed that incorporation of staple domains follows the path of the template strand, which is consistent with our prediction that staple incorporation is propagated by sequential toehold-mediated displacement events as shown in Figures 2.36. To clarify, 'following the path of the template strand' means that folding does not propagate via the seam cross-links; it only spreads from one half of the tile to the other where they are connected by template at the top and bottom of the tile seam.

The simulation results above suggest that when the fully helper-coated template is used, seam staples become bound to the template during folding as part of the sequential incorporation of staple domains by toehold-mediated strand displacement (Figure 2.45). However, if the other half of the tile is still coated with helpers, the other domains of these seam staples will not be able to bind until folding has propagated all the way round the template to the other half of the tile. This means that the seam staples cannot be involved in favourable loop closure steps until much later in folding. We propose that folding around the seam is particularly unfavourable because incorporation of seam staples (during folding of the first half of the tile) does not provide any favourable loop closure steps.

It may be the case that omitting helpers at/adjacent to the seam on one half the tile only (the half where folding is nucleated) may be sufficient to favour folding. This has not been tested experimentally.

Helpers in green, body staples in orange and seam staples in blue.

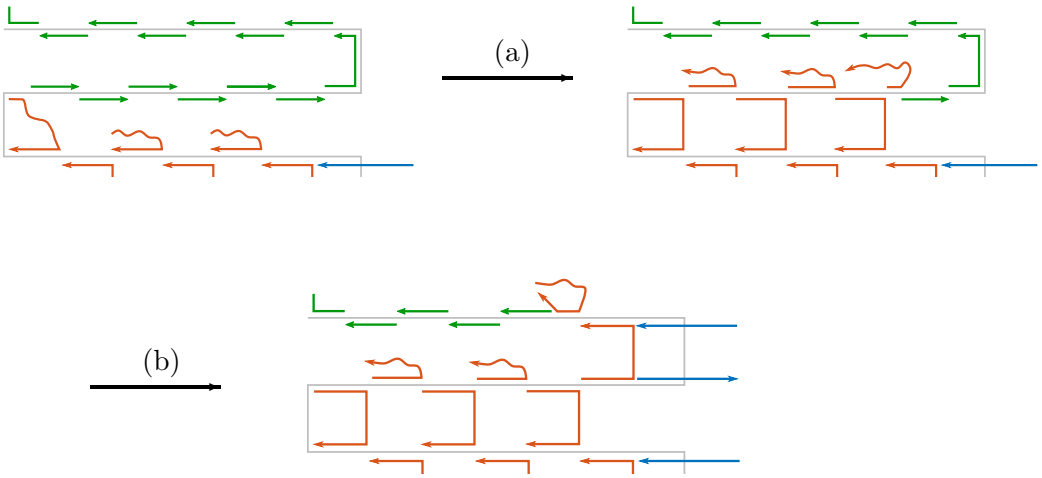


Figure 2.45: Propagation of folding around the tile seams. (a) Incorporation of staple domains along a helix of the tile. Incorporation of every other domain is a favourable loop closure step (see main text). (b) When staple incorporation reaches the seam, the seam staples (blue) are not involved in any favourable loop closure steps unless the other half of the tile is already folded. This may be why propagation is slow around the seams.

2.6 Conclusions and further work

In this chapter, I have shown that using helper strands to coat the template of a DNA origami tile allows isothermal folding of the tile at temperatures below the melting temperature of the tile, which is otherwise not possible. In addition, omitting helpers which have overlapping binding domains with the seam staples further favours folding of the tile, so that it is possible at lower temperatures and/or shorter timescales. As shown in Figure 2.44, this may allow templates to be prepared with helpers in such a way that folding with a particular set is very favourable. This could allow successful folding of a target structure in competitive folding experiments, as described in the introduction of this chapter.

In addition to competitive folding experiments, further work on the use of helpers to aid origami folding could involve:

1. Testing whether helpers can also be used in conventional origami designs i.e. those which use three-domain staples. As discussed in section 2.4.6.2, isothermal folding of the tile used here is particularly poor (in the absence of helpers), perhaps due to the use of two-domain staples. It is not clear how the use of helpers would affect folding in three-domain-staple structures.

2. Attempting to lower the temperature at which isothermal folding with helpers is successful. As explained in the introduction of this chapter and in section 2.5.2, following binding of the first staple to helper-coated template we expected staples to replace helpers through sequential toehold-mediated strand displacement reactions, following the path of the template strand. From what we know of the rates of toehold mediated strand displacement [137], it is not clear why we did not observe the process happening in one hour incubations at room temperature or 40°C. The sharp transition seen with temperature (Figure 2.18) suggests that there is some barrier to propagation which can only be overcome (in a reasonable timescale) when the system is heated to ∼53°C. Omission of helpers at or near to the seam lowers this temperature and pro-

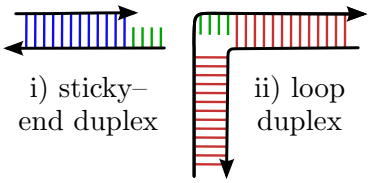
duces a broader transition with regards to temperature (Figure 2.41c-e and Figure 2.42b-c), but the system still requires heating to >50°C for full folding in one hour. Identification of the barrier(s) preventing rapid folding at lower temperatures may allow these barriers to be eliminated, allowing rapid folding at lower temperatures than have been demonstrated here.

For both of these areas of research, the use of simulations may help us to predict and understand the effect of helpers. In particular, simulations would continue to allow us to distinguish the effects of sequence and binding position (of helpers/staples), as the effect of sequence heterogeneity can be removed by using average sequences.

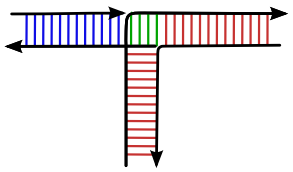
Chapter 3

Characterisation of T-motifs

DNA T-motifs were introduced in 2009 by Hamada and Murata [48] as a novel alternative to crossovers and sticky-end interactions for joining two duplexes. As shown in Figure 3.1, a T-motif is formed from two duplexes: one with a sticky-end, and the other with a single-stranded loop region which has the complementary sequence to the sticky-end. Providing that the loop/sticky-end length is not too short or too long, base-pairing between the loop and sticky-end results in the loop duplex bending through ∼90° so that the overall structure is a rigid planar 'T' shape.



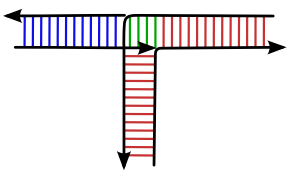
(a) Two duplexes which form a T-motif



(b) 5' T-motif



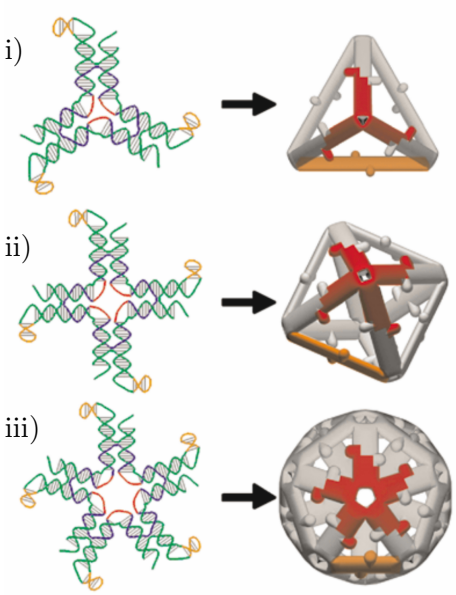
(c) 5' T-motif (helical representation)



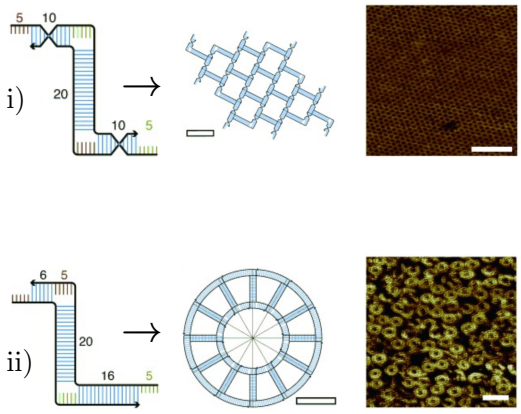
(d) 3' T-motif

Figure 3.1: Single T-motif formation from two duplexes. (a) the separate duplexes, (b) a '5' T-motif' where the sticky-end is a 5' end, (c) helical representation of a 5' T-motif, (d) T-motif with a 3' sticky-end. Arrowheads show 3' ends of strands.

The sticky-end of the T-motif can have either polarity (5' or 3'; Figure 3.1), although only 5' ends have been used in the literature so far. These 5' T-motifs have used loop/sticky-end lengths of 5-6 nts, which has allowed construction of rigid junctions ~90° as shown in Figure 3.2. (Example T-motif structures with loop/sticky-end lengths of 7 nts, from oxDNA2 simulations, are shown later in Figure 3.14.) Figure 3.2 shows the various structures which have been built using T-motifs, including discrete 3D polyhedra ([84]; Figure 3.2a) and continuous 2D lattices ([48]; Figure 3.2bi). Their use in larger but addressable DNA nanostructures (like DNA origami) has not yet been explored.



(a) Polyhedra constructed using T-motifs. Each 3D structure on the right is formed from n copies of the left-hand structure (one copy shown in red). i) n=4, ii) n=6, iii) n=12. Reprinted (adapted) with permission from [84]. Copyright (2012) American Chemical Society.



(b) Repetitive lattices formed from T-motifs: scale bars of central images are 10 nm, on the right are AFM images (scale bars = 100 nm). Reprinted with permission from [48]. Copyright 2009 WILEY-VCH Verlag GmbH & Co. KGaA.

Figure 3.2: Uses of T-motifs

In this chapter, I present a characterisation of individual T-motifs, in preparation for their use in building nanostructures as discussed in Chapter 4. The T-motifs studied cover a broader range than are present in literature: namely 3' T-motifs, and those with loop/sticky-end lengths other than 5 or 6 nts. The geometry and

stability of these T-motifs has been studied both experimentally (gel electrophoresis; performed by myself) and with simulations (performed by Behnam Najafi). The results demonstrate that T-motifs with a range of loop/sticky-end lengths, and both 5' and 3' polarity, are suitable for building nanostructures since they are able to form at μM concentrations.

3.1 Methods

Gel electrophoresis was used to measure the binding strengths (as represented by the dissociation constant K_d) of T-motifs which used different sticky-end polarities and different loop lengths. The association of the sticky-end duplex with the loop duplex (see Figure 3.1) reduces the electrophoretic mobility of the loop duplex proportionally to the amount of sticky-end duplex which is bound, as explained in section 3.1.2. Accurate measurement of the change in electrophoretic mobility (facilitated by the use of distance markers as explained in section 3.1.1) therefore allowed the amount of association between the two duplexes at different concentrations to be measured, and used to estimate K_d values.

3.1.1 Use of distance markers in gel electrophoresis

During my initial work on characterising T-motifs, it became clear that precise measurements of the electrophoretic mobilities of DNA structures within gels would be advantageous. As shown in Figure 3.3a, this was achieved by using a low-mobility and a high-mobility distance marker (LDM/HDM, respectively) along with the sample of interest (the T-motif structure or TM) in each gel lane. This allowed accurate measurement of the relative electrophoretic mobility of the TM, as shown in Figure 3.3d-f, by correcting for any differences between lanes within a gel caused by e.g. the centre of the gel becoming hotter than the edges during electrophoresis. The use of distance markers in gel electrophoresis to correct for these lane differences was introduced by Rossi-Gendron in [121]. In my work, these potential sources of error were also minimised by repeating each gel, with the lanes of the second repeat in a different order to the first. All gels described in this chapter are polyacrylamide (see Appendix A.4).

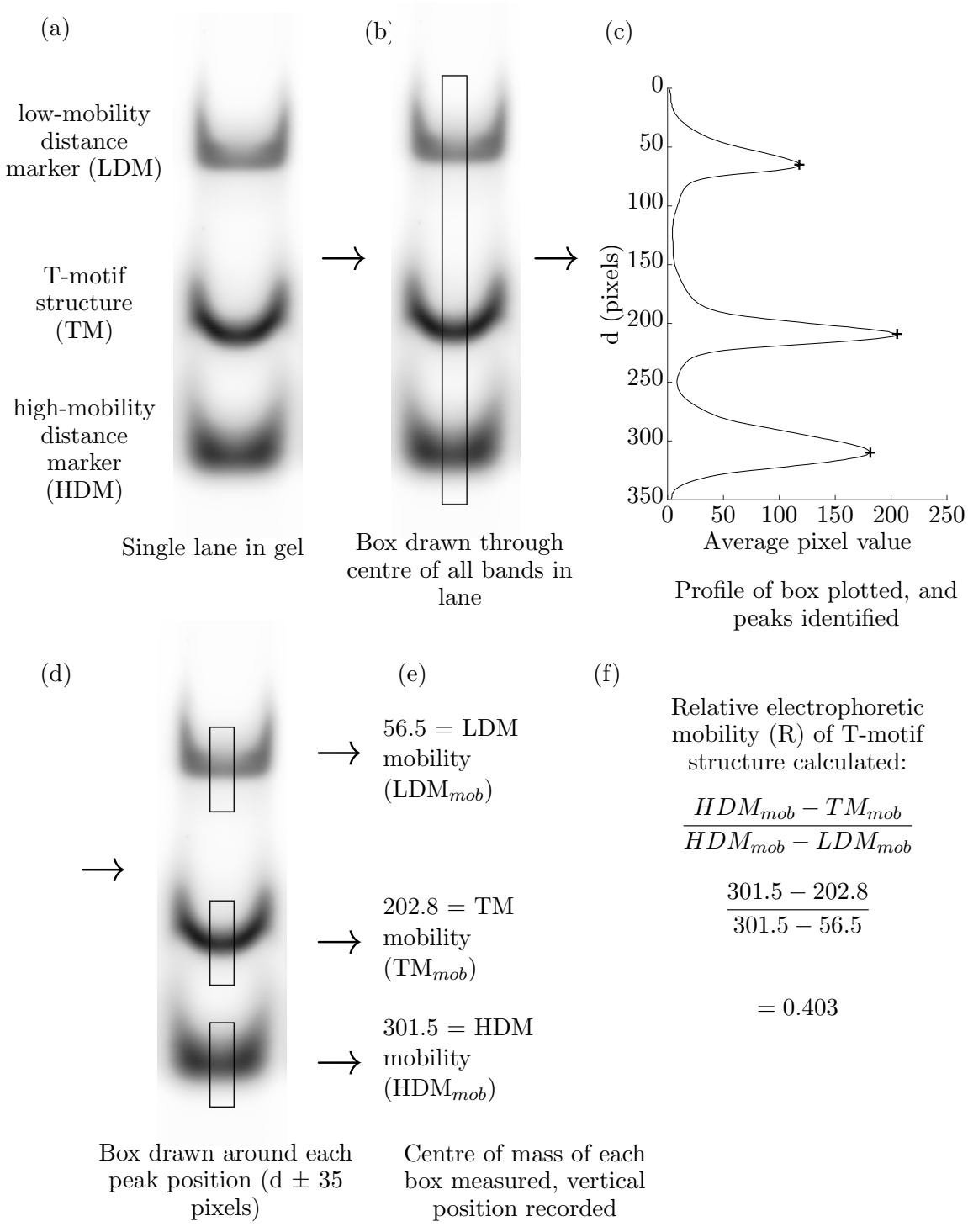


Figure 3.3: Process used to measure the relative electrophoretic mobility of T-motif structures using PAGE with distance markers. This is a semi-automated process using a Fiji [124, 125] macro where the user specifies the lane positions (b) and verifies the box positions in (d). The program automatically plots the profile of each lane (c), finds the peak positions (c), measures the centres of mass of the boxes in (d) and outputs the vertical positions of these (e). See Appendix C.1 for the full program code.

3.1.2 Construction of binding curves

Each individual T-motif is formed from four strands: two from the sticky-end duplex and two from the loop duplex. Initial gel electrophoresis experiments looking at the mobility of the loop duplex showed a shift upon addition of the sticky-end duplex (all strands were annealed together; see Appendix A.3.2.3) which was assumed to correspond to T-motif formation (see Figure 3.4c). However, the mobility of this T-motif structure altered when the concentrations of the four constituent strands were changed (Figure 3.4c). This behaviour seems to arise from the very fast on- and off-rates of the T-motif (sticky-end:loop) interaction, which means that the duplexes are binding and unbinding whilst moving through the gel, so that the mobility of the observed band is intermediate between the fully-unbound and the fully-bound loop-duplex.

I observed that reducing the length of the sticky-end duplex reduced the change in mobility of the loop duplex, as shown in Figure 3.4d. This may arise due to the size difference between the sticky-end and loop duplexes in Figure 3.4d, which results in different electrophoretic mobilities and thus separation of the two duplexes during gel electrophoresis. Alternatively, the use of a shorter sticky-end duplex may simply result in a smaller mobility shift upon binding to the loop duplex due to its smaller size. Since measurements of smaller changes are associated with larger errors, sticky-end and loop duplexes of ∼25 bp were used in this chapter to limit either possible cause of the small mobility shifts seen in Figure 3.4d.

I have used the change in mobility of the loop duplex (ΔR), when annealed with different concentrations of sticky-end duplex (Figure 3.5a), as a measure of T-motif formation:

R_{loop} = relative electrophoretic mobility (determined as per Figure 3.3) of loop duplex (no sticky-end duplex added)

R_{sample} = relative electrophoretic mobility of loop duplex annealed with a specific concentration of sticky-end duplex

ΔR = R_{sample} - R_{loop} = measure of T-motif formation i.e. amount of bound loop duplex

A binding curve of ΔR vs sticky-end duplex concentration (x) could then be constructed (Figure 3.5b) and the values fitted to the equation for a bimolecular binding interaction [2] using a non-linear regression fit in Matlab (see Appendix C.2). This equation is derived by assuming that the components are at equilibrium. The fitting parameters are the maximum value of ΔR (ΔR_{max}, also called B_{max}) and the dissociation constant K_d:

$$\Delta R = R_{sample} - R_{loop} = \Delta R_{max} * \frac{x}{K_d + x} = B_{max} * \frac{x}{K_d + x}$$

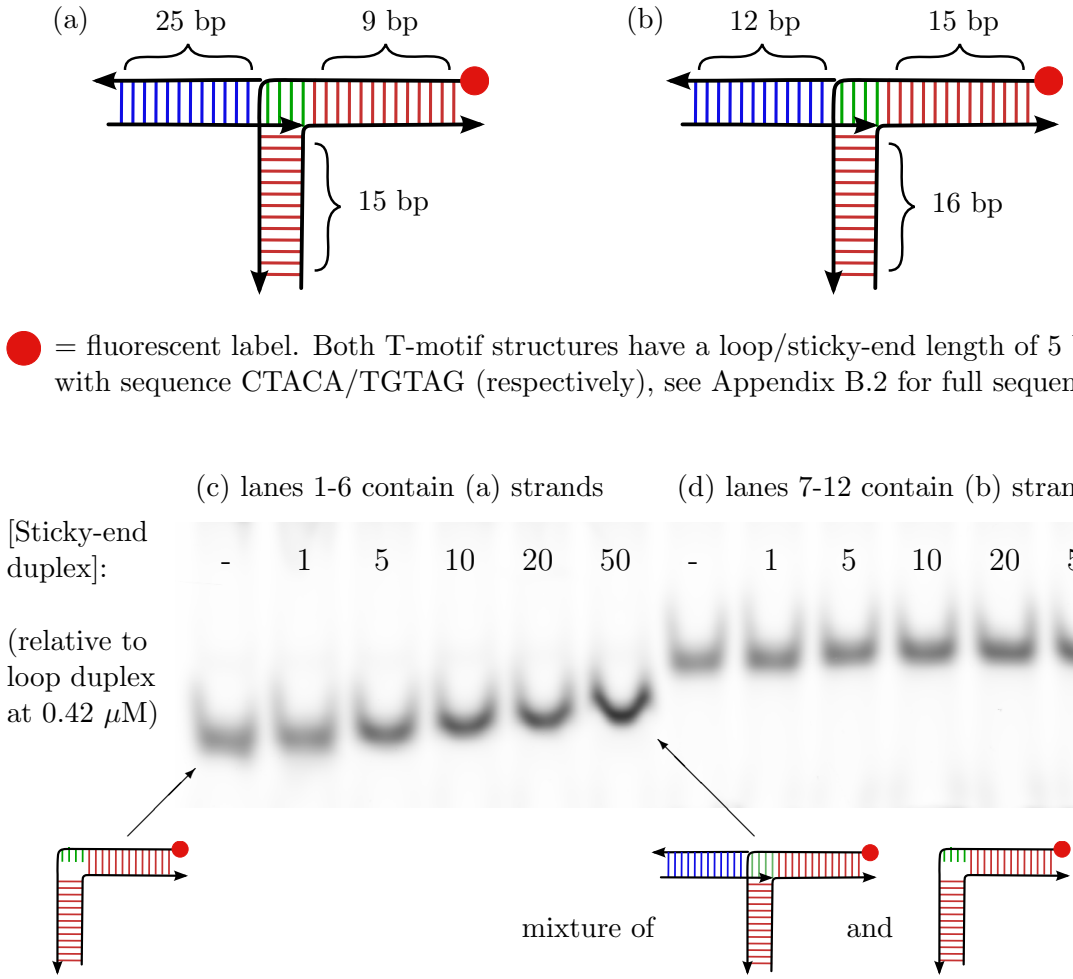
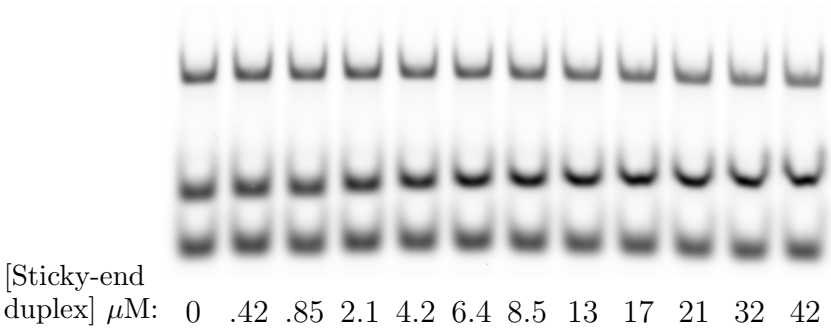
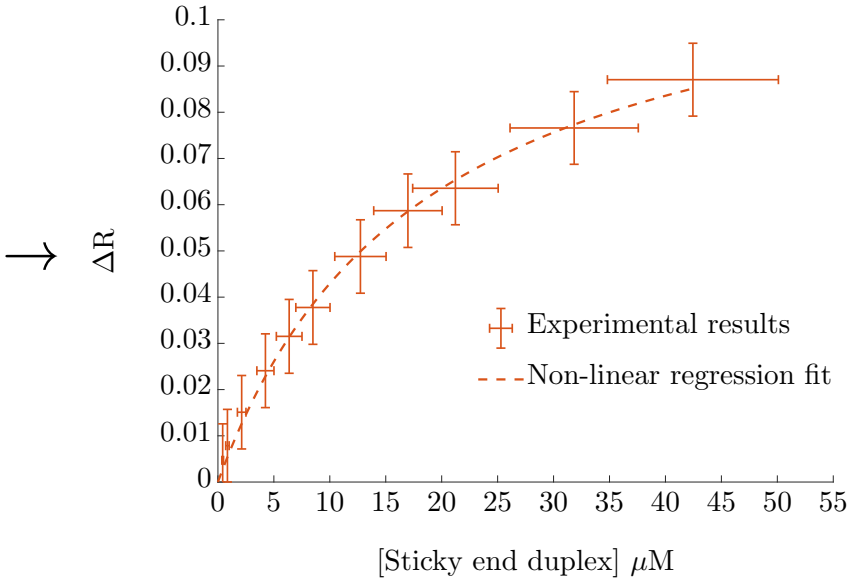


Figure 3.4: Binding behaviour of sticky-end and loop duplexes during gel electrophoresis. (a) and (b) show the strands used in gel lanes (c) and (d) respectively. Each gel band corresponds to a fluorescently-labelled strand of the loop duplex, as shown in (a) and (b). (c) Addition of the sticky-end duplex results in a mobility shift in the loop-duplex consistent with T-motif formation. This shift changes with increasing concentration of sticky-end duplex. (d) Changing the length of the three 'arms' of the T-motif to increase the size difference between the two duplexes reduces the degree of binding.



Loop duplex at 0.42 μM in all lanes

(a) Binding gel: each lane has the same concentration of loop duplex (0.42 μM), and increasing concentrations of sticky-end duplex from left to right. The relative electrophoretic mobility of the loop duplex alone (R_{loop}; first lane) was subtracted from the relative electrophoretic mobilities of the T-motif structures to give the ΔR values plotted below.



Fit to equation for one-site specific binding [2]:

$$y = B_{max} * \frac{x}{K_d + x}$$

$B_{max} = 0.122$
$K_d = 18.4$ μM

(b) Binding curve of the data from (a) (orange). Vertical error bars represent the resolution of the gel images (50 μm); horizontal errors bars represent pipetting errors which could affect strand concentrations (see section 3.2.1.1), but do not include the effect of estimating strand concentrations in the gel as discussed in the main text. A non-linear regression fit (see Appendix C.2 for details) was carried out in Matlab (using the equation for one-site specific binding shown above) and used to find the parameters B_{max} (the estimated relative electrophoretic mobility of the fully-bound loop-duplex) and K_d, as well as their standard errors.

Figure 3.5: Calculation of B_{max} and K_d from a binding gel (example here is for the same strands as used in Figure 3.4a).

In order to account for the fact that the interaction is taking place in the gel environment, the concentrations of strands within the gel were estimated using the volume of the observed bands (resulting from visualisation of the fluorescent label of one of the loop duplex strands, see section 3.1.4). The height of each band (full width at half maximum) was estimated at 1.5 mm, and the dimensions of the gel lanes (lane width = 3.35 mm; gel thickness = 0.75 mm) were used for the width and depth of the band. Since 8 μL (8 mm^3) of sample was loaded into each lane, and estimated to become concentrated into a volume of $\sim$3.77 mm^3 (1.5 mm $\times$ 0.75 mm $\times$ 3.35 mm), this meant that the strands were $\sim$2.12$\times$ more concentrated within the gel than in the original sample. The values plotted in Figures 3.5, 3.6, 3.9 and 3.10, and Tables 3.1 and 3.3, have been corrected to reflect this. This estimation of gel concentration is a source of uncertainty in the K_d estimates, as will be discussed further in section 3.2.1.

As mentioned above, in these experiments the electrophoretic mobilities of the loop and sticky-end duplexes were closely matched, so that they remained in close proximity to each other during running of the gel. This was tested experimentally as shown in Figure 3.8. Particularly for loop/sticky-end lengths of 7 nt, the duplexes did migrate at slightly different rates; the effect of this on the K_d estimates is discussed in section 3.2.1.

This method is comparable to the EMSA (electrophoretic mobility shift assay) method [52] used to characterise protein-DNA interactions. However, due to the slower off-rate of protein-DNA interactions compared to the T-motif interactions, minimal unbinding occurs during electrophoresis and so bands of only two different mobilities are present: one corresponding to the fully-unbound DNA fragment, and one corresponding to the fully-bound fragment. The distribution of DNA between these two bands at different protein concentrations can be used to calculate a K_d value in a similar way to that described here for T-motifs.

3.1.3 Controls

To ensure that the observed changes in electrophoretic mobility of the loop duplex upon addition of the sticky-end duplex were not due to either non-specific interactions at high DNA concentrations, or the higher percentage of TE buffer in the lanes with high DNA concentrations (all DNA strands were stored as 100 μM stocks in TE), controls were carried out. As shown in Figure 3.6, adding high concentrations of a sticky-end duplex with a non-complementary sticky-end to the loop of the loop-duplex, does not result in a shift of the loop-duplex (ΔR remains at $\sim$0).

Using a complementary sticky-end sequence, but normalising the TE concentration (increasing it to 100 μM in all lanes), does not give significantly different results to when [TE] is not normalised: the error bars at each reading overlap with the error bars of the standard results; and the estimated K_d is 14.3 μM, compared to 18.4 μM (see Figure 3.5) when the TE concentration is not normalised.

Both of these results show that the shifts seen in the binding gels are due to specific interactions between the complementary sticky-ends and loops of the two duplexes (i.e. T-motif formation).

3.1.4 Strand sequences, lengths and concentrations

The duplexes used for the main set of experiments are shown in Figure 3.7a; the 3' 5 nt loop structure (i.e. where n=5) is the same as that shown in Figure 3.4a and used in Figure 3.5. For each loop length, the loop duplex and sticky-end duplexes have similar mobilities in the gel. However, as shown in Figure 3.8, there is a slightly larger difference for the 7 nt loop than the 5 or 6 nt loops. This means that that the duplexes will move apart during gel electrophoresis, and will likely result in less T-motif formation (as the on- and off-rates of T-motif formation are very high, see section 3.1.2). This may lead to the values of K_d determined for the 7 nt loops being overestimates, as discussed in section 3.2.1. For future experiments it would be possible to correct for the differences between sticky-end and loop duplex mobility

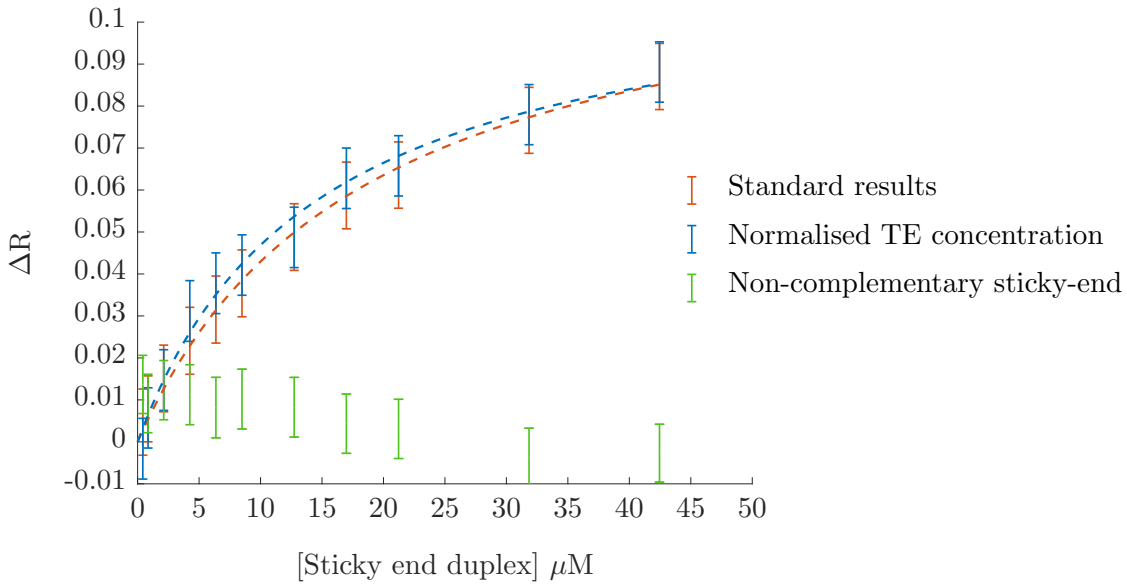
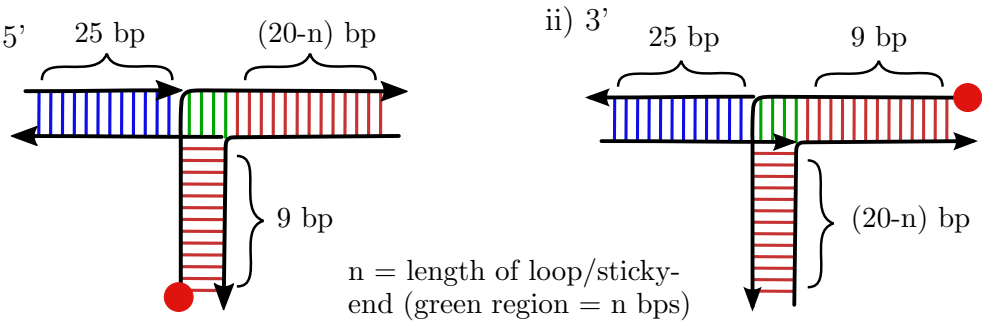


Figure 3.6: Binding curves for controls where the TE concentration was normalised in each lane (blue), and where the sticky-end duplex had a non-complementary sticky-end sequence to the loop sequence (green). Dashed lines show non-linear regression fits as per Figure 3.5. All other conditions were identical to the experiment shown in Figure 3.5 (the 'standard results' shown in orange are the experimental results shown in Figure 3.5b).

(by adding or removing base-pairs from the duplexes).

(a)



(b)

n	Loop sequence	Sticky-end sequence
5	CTACA	TGTAG
6	CTACAG	CTGTAG
7	CTACAGG	CCTGTAG

Figure 3.7: DNA strands used in experiments to measure single T-motif binding stabilities. (a) shows the duplex lengths used (and which strands were fluorescently-labelled, shown by a red circle). (b) Table 3.8 shows the loop/sticky-end sequences used.

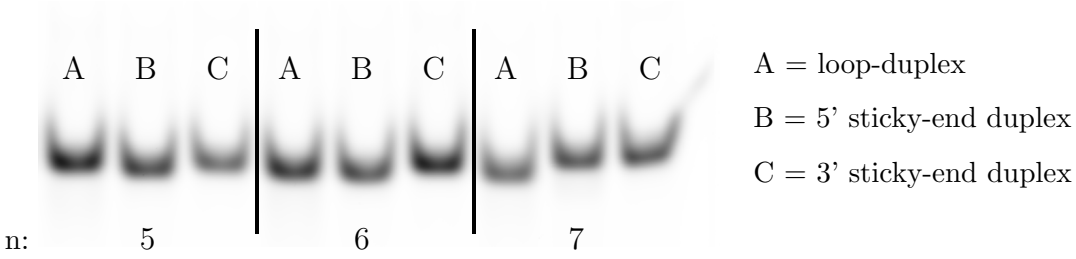


Figure 3.8: Mobility of duplexes shown in Figure 3.7. For n=5 or 6, the sticky-end (B/C) and loop duplex (A) have a similar mobility, whereas for n=7, the sticky-end duplex has a lower mobility than the loop duplex. This may lead to the K_d values determined for the n=7 T-motifs being overestimates (see main text).

For 5' and 3' T-motifs with loops of 5-7 nucleotides, the K_d and B_{max} values of their interaction was measured as described in section 3.1.2. The range of sticky-end duplex concentrations was optimised to provide a good spread of relative electrophoretic mobilities for the binding curve fitting, and was 0-42 μM for the 3' T-motifs and 0-17 μM for the 5' T-motifs.

3.1.5 T-motif simulation results

Simulations of T-motifs were carried out by Behnam Najafi using the oxDNA2 model [109, 133, 139]. This is a coarse-grained model of DNA where each nucleotide is treated as a rigid body, and includes potentials to represent sugar-phosphate backbone connectivity, excluded volume, hydrogen-bonding and base-stacking interactions (nearest-neighbour stacking, cross-stacking between adjacent base pairs in a duplex, and coaxial stacking) [109, 139]. One of the differences between oxDNA2 and the original oxDNA model is that oxDNA2 uses different widths for the major and minor grooves [133].

3.2 Results

All of the simulation results referred to in this chapter are the work of Behnam Najafi (see section 3.1.5).

Two gel repeats (as described in section 3.1.1) were carried out for each of the six structures. Anomalous results were identified by eye as those which lay outside of the binding curve, and were excluded (there were a maximum of two such readings per gel repeat). The data from each gel were used separately to determine a K_d value, which are shown in Table 3.1. The average K_d was then used in the equation $\Delta G^\circ = RTlnK_d$ to determine ΔG° at 4°C, the standard Gibbs free energy change associated with T-motif formation (at 4°C). For both repeats of the 5' n=7 T-motif, the loop duplex was nearly- or fully-saturated at all concentrations measured. This resulted in small differences between the electrophoretic mobilities at different sticky-end concentrations, which either could not be measured (in the case of one repeat) or are associated with a high measurement error (in the case of the second repeat). A K_d estimate was able to be determined from the second repeat, and used to calculate the ΔG° value, but these values are likely to be unreliable because of the high measurement errors. Nevertheless, we can conclude that this interaction is stronger than the other T-motifs measured.

Sticky-end polarity	**n**	**K_d ± SE (μM)**	**Sticky-end polarity**	**n**	**K_d ± SE (μM)**
3'	5	21.8 ± 2.2 18.4 ± 1.4	5'	5	2.58 ± 0.32 4.13 ± 0.15
	6	14.4 ± 4.2 12.2 ± 3.2		6	0.38 ± 0.048 0.268 ± 0.046
	7	16.7 ± 3.3 16.6 ± 4.4		7	0.0610 ± 0.0096

Table 3.1: Experimental K_d values and their standard errors (see Figure 3.5 for method and Figure 3.7 for strand lengths and loop sequences).

The ΔG° values are shown in Figure 3.9, alongside those calculated using K_d estimates from oxDNA2 simulations of the same strand sequences. It should be noted that the simulations were carried out at a different temperature and using different strand concentrations to the experimental results. The simulation results shown in Figure 3.9 have been scaled to correct for the concentration, but the temperature difference (25°C for simulations; 4°C for the gel electrophoresis experiments) is likely to be an important contributor to the discrepancies between the experimental and simulation results, as discussed further in section 3.2.1.

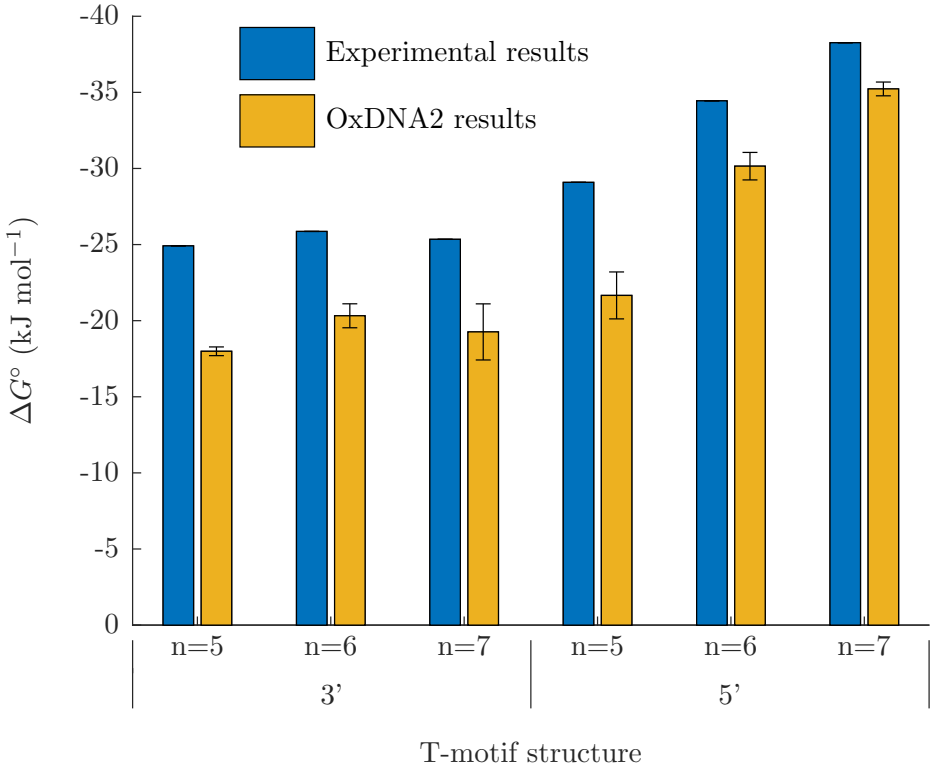


Figure 3.9: ΔG° value calculated from the average of the K_d values shown in Table 3.1 (blue). Also shown for comparison are values from oxDNA2 simulations for the same strand sequences. Error bars correspond to the standard deviation of repeated simulation results. See main text for discussion of possible sources of discrepancy between experimental and simulation results.

The experimentally-determined ΔG° values show a remarkable correlation with the simulation results. However, each experimental ΔG° is more negative (by ∼4-8 kJ mol^{-1}) than the corresponding oxDNA2 simulation result, meaning that the

experimental K_d estimates are a factor of ∼5-32 times lower than those calculated from the oxDNA2 measurements. It is likely that much of this difference is due to the temperature difference at which K_d values were measured in the simulations and experiments (25°C vs 4°C, as mentioned previously), but other factors such as the use of a gel environment when experimentally measuring these interactions, errors in the model, or errors in the experimental method may also contribute.

Possible sources of experimental error are described below and summarised in Table 3.2.

3.2.1 Possible sources of experimental error

As previously mentioned, the main contributor to the discrepancy between experimental and simulation results seen in Figure 3.9 is likely to be due to the temperature difference at which K_d values were measured (25°C for simulations vs 4°C for the experiments). However, experimental errors such as those listed here may also contribute.

3.2.1.1 a) Pipetting errors

All pipettes were from the Research 2100 series or the Research Plus range by Eppendorf, who provide the systematic and random errors of these pipettes in their technical data. From this, the accumulated errors can be calculated [99], and were a maximum of 16% (systematic) and 2% (random) for the binding gel samples. This high value for the systematic error assumes that the systematic errors are additive, e.g. that all of the pipettes over-dispense, or they all under-dispense, which is unlikely to be the case.

3.2.1.2 b) Local strand concentrations during electrophoresis

In order to produce an estimate of the equilibrium constant of the T-motif interaction, we assumed that all four strands migrated through the gel at the same speed, and so were all confined to the same (constant) volume within each gel lane during electrophoresis. Estimating this volume allowed us to calculate the concentration of each strand. However, this method has several shortcomings. Firstly, estimating the volume in which the strands were contained could not be carried out accurately (see section 3.1.2). In addition, in a gel matrix environment the DNA strands may be excluded from some of the volume. The extent to which this occurs depends on the relative size of the DNA structures compared to the pore size of the gel. For linear DNA structures, the relevant size parameter is the radius of gyration, which can easily be calculated from the number of base pairs [55]. However, for branched structures such as the T-motifs used here there is not a straightforward parameter

which can be used to describe the behaviour in a gel environment [79]. Hence, although we know that the 8% 29:1 (monoacrylamide/bisacrylamide) polyacrylamide gels used here have a median pore size of ∼90 nm [55], it is difficult to estimate to what extent the T-motif structures are able to access these pores.

We estimated that the samples are concentrated by a factor ∼2.12× within the gel compared to the original sample. Changing the strand concentrations (x values in Figure 3.5b) changes the K_d estimate by the same factor. Therefore, in order to account for the 5-32 fold difference seen between the experimental K_d estimates and those calculated from simulations, our concentrations estimates would have to be 5-32× lower than the true values. I believe such a large discrepancy is unlikely given the type of gel and method of estimation used (analysing the volume in which the fluorescent label of one of the strands is visible).

Secondly, assuming that all strands are occupying the same volume as each other during electrophoresis may not be accurate. Duplexes may migrate at different rates, as is certainly true for the structures where n=7 (see Figure 3.8). This could lead to an overestimation of the experimental K_d for n=7, and thus an underestimation of the experimental ΔG°. It is not obvious from the results that this is the case; there is not a greater difference between the experimental and simulation ΔG° results for n=7 compared to n=5 or 6.

In addition, strands may occupy differently-sized volumes to each other. In PAGE, increasing the concentration of the DNA strand or structure loaded into a lane increases the size of the band observed in that lane (in the dimension corresponding to DNA migration), i.e. the DNA occupies a larger volume. This effect may explain some of the anomalous results observed at high sticky-end concentrations, as illustrated in Figure 3.10, which shows the results of one repeat of the 3' n=7 experiment. The two results for the two highest sticky-end concentrations lie far above the binding curve predicted from the rest of the results. This may be because at these high concentrations, the sticky-end duplex occupies a larger volume of the gel, and is therefore able to bind more of the loop duplexes, perhaps by com-

pensating for the separation of the two duplexes due to different migration rates. When these results are not excluded from the non-linear regression fitting and determination of K_d values, the estimated K_d values are lower, suggesting that my K_d values may be overestimates due to duplexes moving apart during electrophoresis.

Measuring in the gel environment has the added complication that the concentration of each strand within the estimated volume is unlikely to be uniform. This effect is expected to be small, but is difficult to quantify.

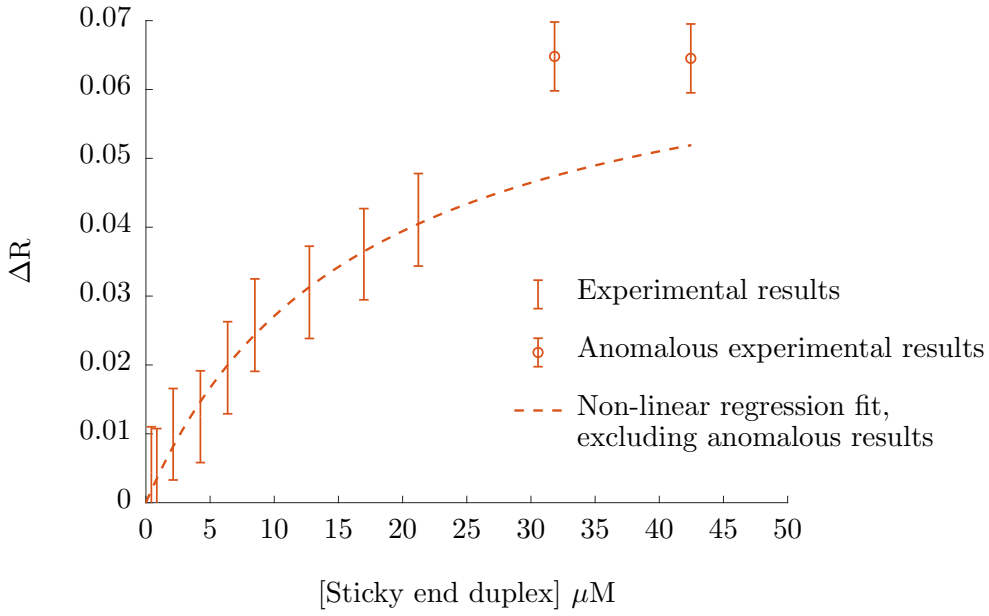


Figure 3.10: Binding curve for a 3' n=7 structure (see Figure 3.5 for method details). Experimental results are shown as markers, including two anomalous results at sticky-end concentrations of 32 and 42 μM. The non-linear regression fit (dashed line) was carried out excluding the anomalous results.

3.2.1.3 c) Errors associated with non-linear regression fitting

No error bars are shown for the experimental results shown in Figure 3.9 as it is difficult to quantify the error associated with curve-fitting using non-linear regression analysis, and the standard errors shown in Figure 3.1 are likely to be underestimates of the true uncertainty [100]. However, they do give us some idea of the confidence limits of the fitted parameters; for example we can see that there are greater standard errors (as a proportion of the K_d) associated with the 5' n=6 and n=7 measurements.

Looking at the raw data, this is because the ΔR of the T-motif structure is close to B_{max} for all concentrations of sticky-end duplex tested i.e. the interaction is strong enough that the loop-duplex appears to be nearly fully-bound even at the lowest concentration of sticky-end duplex tested. A better range of x values for the binding curve would have yielded lower standard errors, but would have increased the pipetting errors involved in preparing very dilute concentrations of sticky-end duplex. The absence of intermediate y values in the binding curves of these two structures (5' n=6 and n=7) could lead to either over- or underestimation of the experimental K_d values. Perhaps here it has been an underestimate, which could account for the smaller difference between the experimental and simulation ΔG° values compared to other structures.

Table 3.2 summarises these errors and their potential effects on the experimentally-determined K_d values.

Source of error		**Expected effect on calculated K_d values**
(a) Pipetting errors		Possible under- or overestimation of K_d; likely to be $\leq$10%
(b) Local strand concentrations during gel electrophoresis	(i) Duplexes migrating at different speeds	Possible overestimation of K_d; difficult to quantify, except results for high sticky-end suggest true K_d could be around half of the estimates here.
	(ii) Under- or overestimation of local strand concentrations (variety of causes)	Possible under- or overestimation of K_d; difficult to quantify. If strand concentrations were overestimated (strands are on average more dispersed than estimated), then calculated K_ds are overestimations, and vice versa.
(c) Errors associated with non-linear regression fitting		Possible under- or overestimation of K_d; difficult to quantify except that by eyeballing the binding curves we can see that (assuming the other errors are minimal), the true K_d must be within $\leq$50% of the estimated value.

Table 3.2: Possible sources of error in determining the experimental K_d values, and their potential effects on these values.

None of the sources of error described here could account for the large and systematic difference between the experimental and theoretical results (experimental K_d values are 5-32 times lower than the theoretical estimates, see Figure 3.9), which as previously mentioned is likely to be mainly due to the temperature difference at which K_d values were measured (25°C for simulations vs 4°C for the experiments). However, the effects of measuring the interaction in the gel (e.g. underestimating the local strand concentrations by not taking into account that the DNA structures cannot penetrate all pores in the gel matrix) may also have also contributed. That both sets of results show the same overall trends suggests that these trends are real, even if there is a discrepancy between the two sets of K_d estimates.

3.2.2 Analysis

The two main observations that we can make are, that for n (the length of loop: sticky-end hybridisation domain) equal to 5-7:

- For a specific n value, 3' T-motifs are less stable than 5' T-motifs. The n=5 structures show the smallest difference between 5' and 3' results, and even this corresponds to a 5- to 7-fold higher K_d (simulation/experimental results respectively) for the 3' structure.
- Increasing n increases the stability of 5' T-motifs, but not 3' T-motifs.

To analyse these results, we must first understand the physical processes that are contributing to the change in free energy (G°) of the structures upon T-motif formation. For all of the T-motif structures, there is an entropic cost associated with bringing the two duplexes together and the formation of hydrogen bonds between the loop and sticky-end bases. This must be outweighed by the favourable enthalpy change upon formation of these hydrogen bonds. In addition, T-motif formation can result in the generation of two additional base-stacking sites between the two duplexes (as shown in Figure 3.11), further stabilising the T-motif structures[1].

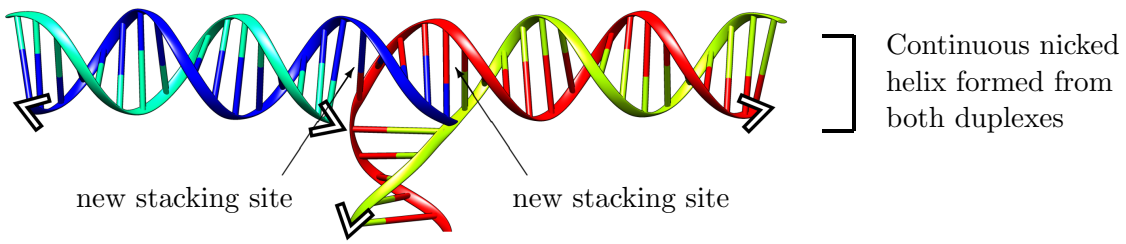


Figure 3.11: Two new stacking sites can be created during T-motif formation (a 5' T-motif is shown here). The new stacking sites are located at the beginning and end of the sticky-end (strand with sticky-end shown in blue). The formation of these stacking sites contributes to the stability of the T-motif, in addition to the hybridisation between the loop and sticky-end. Arrowheads shown 3' ends of strands.

[1]As shown by simulation results in Figure 3.13, certain loop and/or sticky-end lengths are able to hybridise (i.e. form T-motifs) without formation of these new stacking sites.

The difference in results observed for 5' and 3' T-motifs may be due to the geometry of the strands at the junction where the two duplexes meet. With a 5' sticky end, the 'branch' arm of the loop duplex spans the major groove of the continuous helix formed by the sticky end and loop duplexes (Figure 3.12a). In contrast, with a 3' sticky-end, this spans the minor groove (Figure 3.12b).

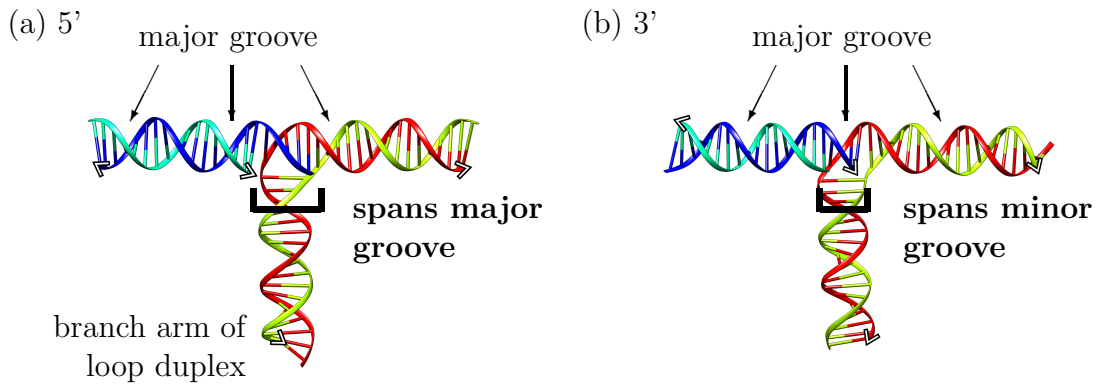


Figure 3.12: Geometry at T-motif junction for 5' (a) and 3' (b) T-motifs. In a well-formed T-motif, a continuous helix is formed from the sticky-end duplex and part of the loop duplex (Figure 3.11). In a 5' T-motif, the 'branch' arm of the loop duplex protrudes from the major groove of this continuous helix, whereas for a 3' T-motif it is joined at the minor groove. This difference may account for the lower stabilities of 3' T-motifs compared to 5'. Arrowheads shown 3' ends of strands.

This difference in geometry may account for the lower stability of 3' T-motifs: the lack of space at the junction point (due to its position in the minor groove) means that the branch arm of the loop duplex is much closer to the rest of the structure than in the 5' T-motif. This may have negative consequences for the stability of the T-motif, due to both electrostatic repulsion between the arms of the T-motif, and lack of flexibility at the junction point. It also follows that the 5' T-motif has more space and flexibility at the junction point to accommodate longer loop:sticky-end hybridisation regions, which would increase the stability of the T-motif interaction. In the 3' T-motif, longer 3' sticky-ends may be sterically-hindered when trying to thread into the minor groove of the loop-duplex, and even if they can successfully hybridise with the loop, the resulting conformation is unlikely to allow formation of the two stacking interactions between the loop and sticky-end duplex seen in Figure

3.11.

A piece of supporting evidence for this interpretation is the relative B_{max} values calculated from the binding curves as per Figure 3.5. We know that linear DNA structures migrate more quickly during gel electrophoresis than branched structures [12]. If we assume that the T-motifs are planar structures with approximately 90° geometry (i.e. a 'T' shape), the relative lengths of the three arms (as shown in Figure 3.7) means that the 3' structures, having a much longer 'branch', should move more slowly through the gel and so have higher ΔR measurements (the loop duplexes, whose relative electrophoretic mobilities are used as a baseline to calculate the ΔR values, are the same in the 5' and 3' experiments). The B_{max} values are estimates of the ΔR of the pure T-motif structures (i.e when all the loop duplex is bound), and so we would expect these values to be higher for the 3' T-motifs.

Sticky-end polarity	**n**	**B_{max} ± SE**	**Sticky-end polarity**	**n**	**B_{max} ± SE**
3'	5	0.113 ± 0.006 0.122 ± 0.004	5'	5	0.106 ± 0.004 0.126 ± 0.002
	6	0.0225 ± 0.0028 0.0178 ± 0.0019		6	0.243 ± 0.004 0.323 ± 0.006
	7	0.0723 ± 0.0080 0.158 ± 0.019		7	0.600 ± 0.004

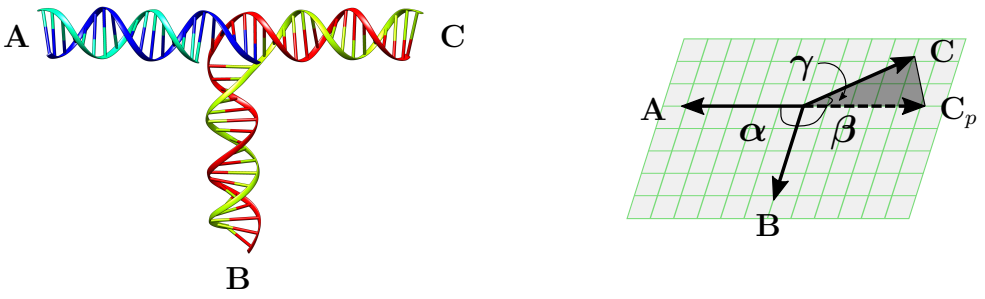
Table 3.3: Experimental B_{max} values and their standard errors (see Figure 3.5 for method and Figure 3.7 for strand lengths and loop sequences).

The fact that the 3' T-motif values shown in Table 3.3 are lower than those for the 5' structures (except for n=5) suggests that the 3' T-motifs are not forming a rigid 'T'-shaped structure. This may mean that the arms are not able to adopt the planar conformation with 90° geometry due to the space constraints at the junction; or that the junction is much more flexible than expected due to either incomplete hybridisation between the loop and sticky-end, or because the two stacking sites shown in Figure 3.11 are not properly formed. Similarly, the increasing Bmax values

seen with increasing n for 5' T-motifs may be due to changes in flexibility and geometry when longer loop lengths are used.

OxDNA2 simulations have determined the angles and degree of stacking in 'fully-formed' T-motifs (i.e. where the loop and sticky-end are fully hybridised), as shown in Figure 3.13.

(a)



Angles in the single T-motif have been labelled α, β and γ as shown above for a 5' T-motif. Arm A is the sticky-end duplex; arms B and C are parts of the loop duplex where arm C is expected to form stacking sites with the loop duplex as in Figure 3.11. C_p is the projection of arm C onto the plane formed by A and B. When the T-motif is adopting the 'ideal' planar, 90° geometry, α=90°, β=90°and γ=0°.

(b)

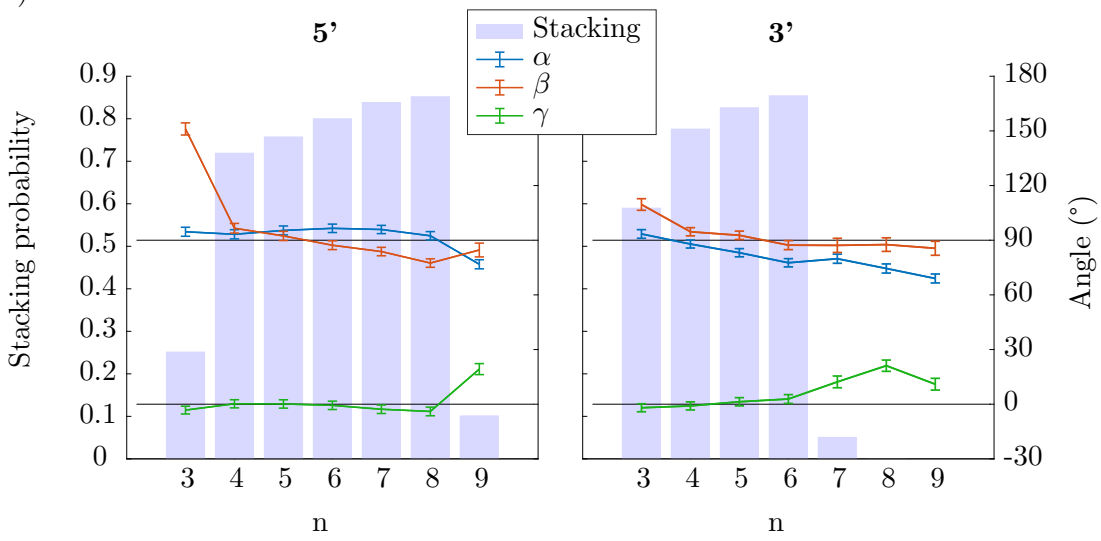


Figure 3.13: OxDNA2 simulations of stacking and angles within single T-motifs. (a) shows angle definitions. (b) Simulation results: stacking probability refers to the proportion of simulations in which both stacking sites (see Figure 3.11) were formed. Error bars for angle values represent the standard deviation from multiple simulations. Average sequences (i.e. base pairs with properties between those of a G-C and A-T base pairs) were used.

The simulation results show that 5' T-motifs can form well-stacked structures with the desired planar 90° geometry when n=4-8, unlike 3' T-motifs where this is limited to n=4-6. Figure 3.14 shows example structures of 5' and 3' n=7 T-motifs to show the difference between T-motifs which do (Figure 3.14a; 5') and do not (Figure 3.14b; 3') have this desired geometry.

(a) 5' n=7 T-motif

b) 3' n=7 T-motif

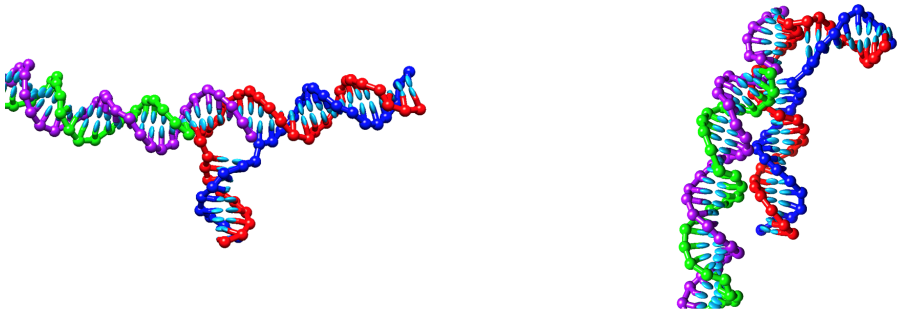


Figure 3.14: Example oxDNA2 simulations of T-motifs with loop/sticky-end lengths of 7 nt. Sticky-end duplex in green and purple (the sticky-end is on the purple strand); loop duplex in blue and red (the loop is on the red strand). a) The 5' n=7 T-motif is able to form correctly (i.e. with both stacking sites formed as per Figure 3.11, and with α and β angles of $\sim$90° and a γ angle of $\sim$0°, see Figure 3.13a). b) The 3' n=7 T-motif cannot form correctly, presumably due to the limited space in the minor groove.

The simulation results shown in Figure 3.13b may explain why the 3' n=7 T-motif has a lower B_{max} value compared to the 5' n=7 T-motif (more flexibility about the junction point), but does not explain why this was also the case for n=6 (since both 5' and 3' appear to be well-stacked in Figure 3.13b), or why the B_{max} values for 3' n=5 and 5' n=5 were similar. In addition, it is not clear why the 5' T-motif B_{max} values increase with increasing n, when there is only a modest increase in stacking, and no significant changes in the α, β or γ angles between n=5 and n=7. This inability to explain the experimental results may be a limitation of the model or may be due to the fact that average sequences were used.

3.3 Summary

Whilst various structures utilising 5' T-motifs have been cited in the literature [19, 48, 63, 83, 84, 144], this chapter presents the first thorough characterisation of T-motif stability and geometry. Through experimental and simulation studies (the latter provided by Behnam Najafi), we have shown that 5' T-motifs are more stable than 3' T-motifs, and that loop/sticky-end lengths of 5-7 can be used to form both types of T-motif at μM concentrations. Simulation results suggest that lengths of 4 would also work well (and 7-8 for 5' only). If a planar, rigid 'T' shape with $\sim$90° geometry is required, the data suggest that a 5' T-motif where n=6 is the optimum structure.

Due to the fact that the estimated experimental K_d values are in a similar range to the concentrations of DNA strands used in gel electrophoresis experiments, the difference in stability between the 5' and 3' T-motifs represents a huge difference in the concentration of strands required to see T-motif formation on a gel. Building nanostructures (which are commonly characterised by gel electrophoresis) using T-motifs is therefore much more straightforward with 5' sticky-ends, which may explain their sole use in the literature so far. However, this difference in stability also means that we can control formation of 3' T-motifs by increasing the concentration, or effective concentration, of the constituent strands. Chapter 4 looks at building nanostructures which exploit this inherent difference in stability to ensure correct assembly.

Chapter 4

Building Nanostructures with T-motifs

In Chapter 3, we have seen how estimates (both experimental and from simulations) of the relative stabilities of 5' and 3' T-motifs show that 3' T-motifs are less stable for all of the loop lengths (5-7 nt) studied. This difference is thought to arise from the fact that the 'branch arm' spans the major and minor grooves in 5' and 3' T-motifs respectively (see Figure 3.12). In addition, we have seen that when T-motif structures are characterised using gel electrophoresis, the electrophoretic mobility of the structure is intermediate between the fully-formed and fully-dissociated structure (see section 3.1.2). This characteristic needs to be considered when analysing all gel results of structures containing T-motifs.

As mentioned in the introduction of Chapter 3, the motivation for this characterisation was to allow selection of motifs for use in building DNA nanostructures. Specifically, we envisaged their use in a double-stranded 'DNA origami-like' nanostructure, an example of which is shown in Figure 4.1a. As shown in Figure 4.1b, this design uses a single template strand which hybridises to a number of short 'adapter' strands, each of which contains a single loop which can be used for T-motif formation. Binding domains (loops) can be repeated within the structure by using the same loop sequences on multiple different adapter strands. This design necessarily

means that that both 5' and 3' T-motifs must be used (see section 4.1).

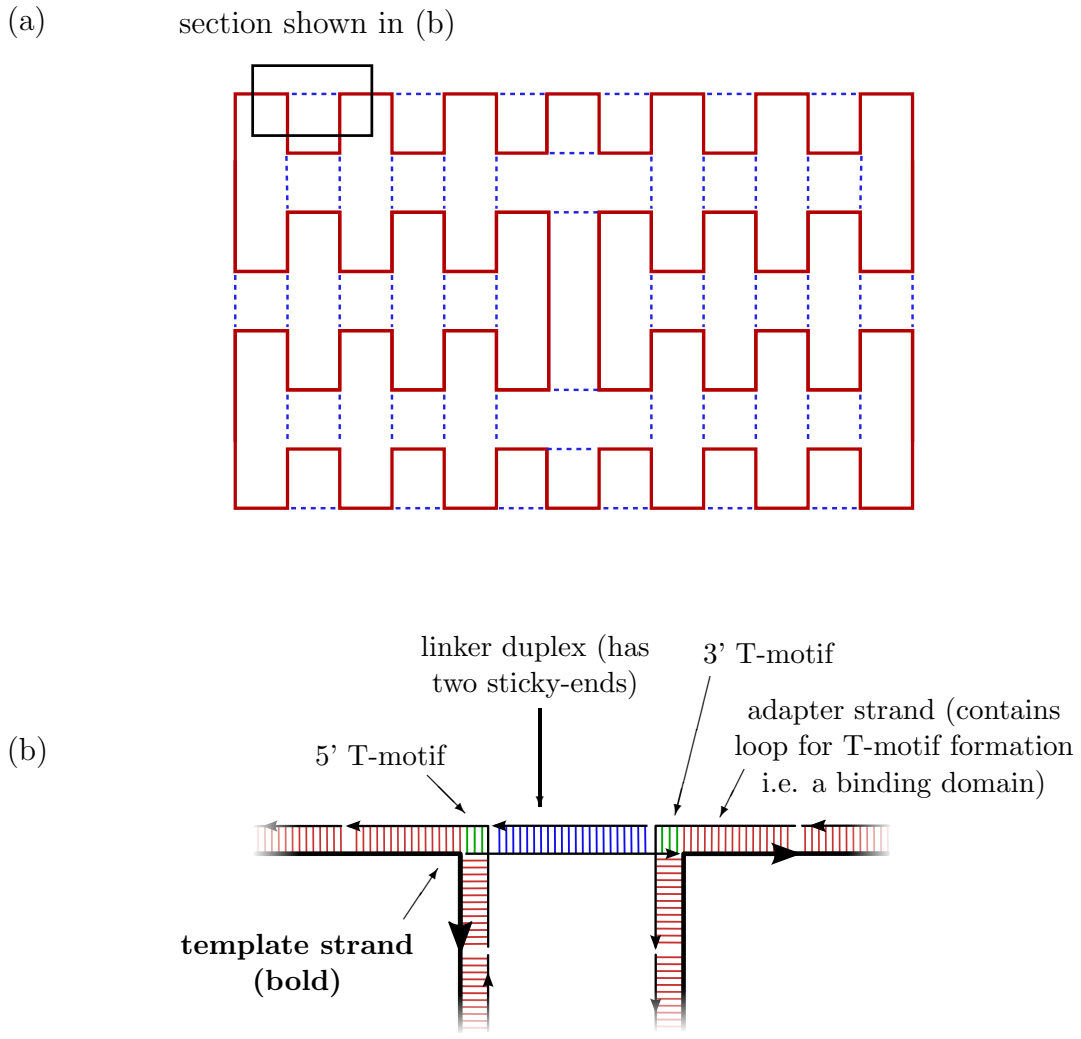


Figure 4.1: Design of a T-motif origami structure. (a) Full structure. Red lines represent the template-adapter duplex (the 'scaffold'), blue dashed lines shown linker duplexes (see b). (b) Detail of section within the structure. A single-stranded template is hybridised to a number of short 'adapter' strands (red), each of which contains a single-stranded loop for T-motif formation. Each linker has two sticky-ends (one 5' and one 3') each of which is able to form a T-motif with an adapter loop (hybridisation between loop and sticky-end shown in green).

Binding domain repetition is difficult within a conventional origami design, as it would require the use of a highly-repetitive template strand. Such templates are not readily available and are often prohibitively difficult to synthesise [102, 106]. In the 'T-motif origami' design shown in Figure 4.1, the template strand does not encode any of the binding domain information as all of the T-motif loops are on the

adapter strands, thus removing the need for it to be repetitive. The template can be designed freely, and the sequences of the non-loop regions of the adapters are then designed to hybridise in the correct positions.

The use of redundant binding domains makes T-motif origami structures ideal test systems with which to study DNA nanostructure folding. Since any linker can bind at multiple different positions, the positions where they do bind (as inferred by the final structure formed) can offer insights into the folding pathway. Such experiments are similar to those carried out by Dunn et al. using a dimer origami template [34], as described in section 2.1 of the previous chapter. In these experiments there were two copies of each binding domain, whereas the T-motif origami structures used here have a much higher degree of redundancy (and therefore many more possible outcomes in terms of the final structure; see section 4.2). The T-motif origami structures can also be used to look at how few unique binding domains are required to produce a well-folded target structure. The use of one loop (binding domain) per adapter allows the sequence of a binding domain in a particular position to be easily changed (by swapping a single adapter strand) when carrying out such experiments.

In addition, T-motif interactions have a considerably lowering melting temperature than the staple-template hybridisation interactions which determine folding in conventional origami (see Chapter 5 for further discussion), which may confer different properties to the T-motif origami structures in terms of assembly temperatures and ability to rearrange at ambient temperatures.

In this chapter, I begin with the characterisation of a small 'U-shaped structure' consisting of a short length of scaffold (i.e. template-adapter duplex) bridged by a single linker (section 4.1). Understanding the behaviour of individual linkers in this model system then allowed the design and testing of intermediate-sized T-motif origami nanostructures similar to that shown in Figure 4.1 (section 4.2).

4.1 Testing individual linker binding

The binding behaviour of individual linkers (as shown in Figure 4.1) was studied using a small 'U-shaped' structure as shown in Figure 4.2.

As shown in Figure 4.1, the design of our T-motif origami structure uses a double-stranded 'scaffold' (in red in Figure 4.1) consisting of a single template strand hybridised to many short adapter strands which contain the loops for T-motif formation (i.e. the binding domains). Locating all of the loops on adapters and not on the long template strand allows binding domains to be easily changed, and (as mentioned in the previous section) means that both 5' and 3' T-motifs must be used: we were unable to design a 2D structure with such a scaffold using solely 5' or solely 3' T-motifs.

Using linkers which form both a 5' and 3' T-motifs may offer some advantages. Given the difference in K_d values for 5' vs 3' T-motifs (see Table 3.1), it is possible to use strand concentrations where individual 5' T-motifs can form readily (and persist during gel electrophoresis) but 3' T-motifs cannot. This means that that the 3' end of each linker can only bind once the 5' end is already bound. Binding of the linker via the 5' end increases the local concentration of the binding domain at the 3' sticky-end and allows the weaker 3' T-motif interaction to form. This hierarchical assembly should prevent excess linker binding i.e. the binding of two copies of a linker, one at each of the sites which the linker was designed to bridge.

As shown in Figure 4.2, the U-shaped structure contains four loops: two can form T-motifs with the linker (on adapters A and D), and two (on adapters B and C) are left single-stranded in the 'unbraced' version (Figure 4.2a) and are used to form 5' T-motifs in the 'braced' version (Figure 4.2b). In the T-motif origami structure shown in Figure 4.1, linkers which bridge a U-shaped section of scaffold are in the 'braced' situation (although the T-motifs formed by the equivalent of adapters B and C in the full T-motif origami structure consist of one 5' and one 3' T-motif; two 5' T-motifs were used here (Figure 4.2b) for ease, as this required lower strand concentrations). The sequences of the four loops were designed to be sufficiently

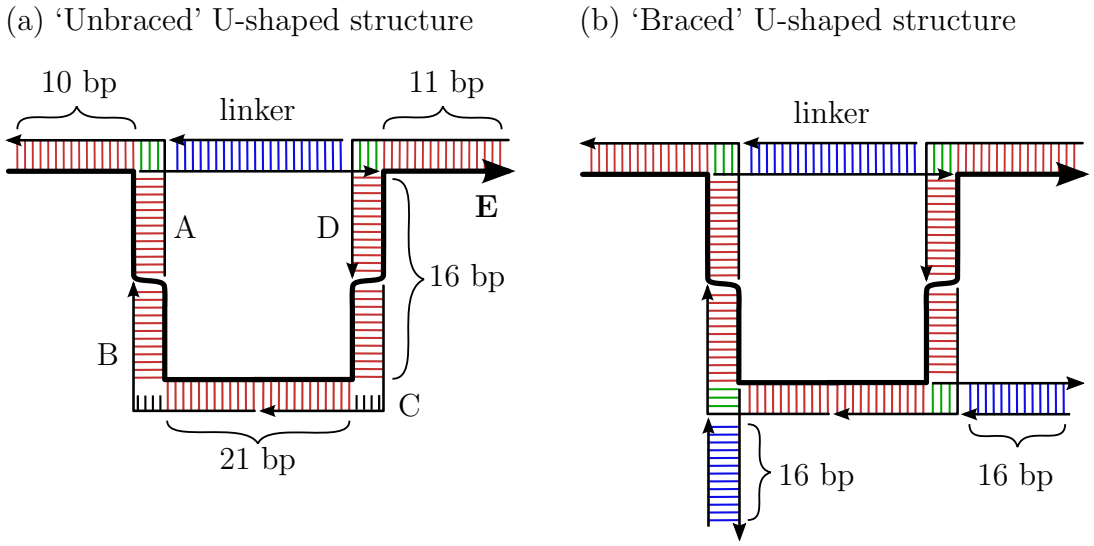


Figure 4.2: U-shaped structures used to study individual linker binding. (a) Unbraced structure: the template strand (E) is hybridised to four adapters (A-D), each of which has a loop available for T-motif formation. The loops in B and C are left-single-stranded. (b) Braced structure: identical to (a) except the loops in B and C are involved in (5') T-motif formation. This structure is predicted to be less flexible than (a). Linker length and loop/sticky-end sequences were varied between experiments.

different to each other (no runs of three or more nucleotides were repeated between loops) to prevent off-target T-motif formation (poly(T) loops were used in some of the unbraced structures).

Studies of the U-shaped structure were carried out using polyacrylamide gel electrophoresis, where all strands were annealed together prior to loading onto the gel; the same method as used for single T-motifs in Chapter 3 and detailed in Appendices A.3.2.3 and A.4.

4.1.1 Hierarchical linker binding

As described above, I expected that by using linkers with both a 5' and 3' sticky-end, the 3' end would only be able to bind to the scaffold once the 5' end was already bound (at specific concentrations). Gel electrophoresis experiments using the unbraced structure (Figure 4.2) with a number of different loop/sticky-end sequences confirmed that this is the case when the scaffold structure is at 0.1 μM and the linker duplex at 0.15 μM (see caption of Figure 4.3 for all strand concentrations). A

typical result is shown in Figure 4.3, which uses an unbraced U-shaped structure.

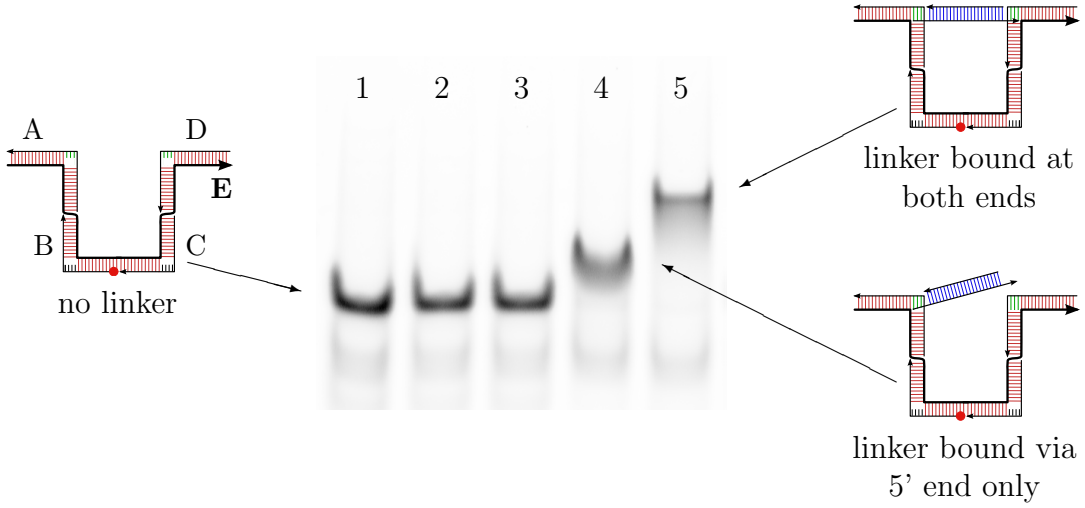


Figure 4.3: Gel electrophoresis results demonstrating hierarchical binding of a linker with both a 5' and 3' sticky-end. Adapters are labelled A-D as per Figure 4.2. Adapter B has a fluorescent label and is visualised here. All strands were annealed together (B at 0.1 μM, all other strands at 0.15 μM) prior to loading. Lane 1: no linker present (strands A-E only); lane 2: a linker is present with incompatible sticky-ends for the loops in adapters A and D, no binding is seen; lane 3: linker present with only a 3' compatible sticky-end, no binding is seen; lane 4: linker compatible at 5' end only, binding is evident from the change in mobility of the structure; lane 5: linker compatible at both ends, binding at both ends is evident from the further change in mobility. The loop sequences used are GTGTCC in adapter A and CACTCG in adapter D, (6a and 6b respectively in Appendix B.3, which contains the full strand sequences).

Lane 3 of Figure 4.3 contains the five strands of the scaffold (A-E) and a linker duplex with the correct 3' sticky-end (complementary to the loop of adapter D) but the 'wrong' 5' sticky-end (i.e. one which is not complementary to the loop in adapter A) to allow binding to the scaffold. The mobility of the band in this gel is identical to that in lane 1 (which contains only the five scaffold strands), showing that the linker is not binding to the scaffold. In contrast, in lane 4, where the linker has the correct 5' sticky-end but an incorrect 3' sticky-end, we can see that there is a shift in electrophoretic mobility indicative of linker binding. When the linker has both of the correct sticky-ends required to bind to adapters A and D, there is an even greater shift in mobility, suggesting that the linker is binding at both ends.

Eight pairs of T-motif loop/sticky-end sequences were tested: four different 5-8

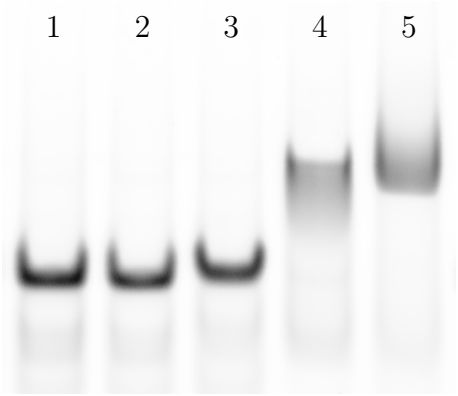


Figure 4.4: A further example of the type of gel shown in Figure 4.3. All details are the same except for the sequences used for the loops/sticky-ends of the T-motifs: here, the loop of adapter A is GTACACG, and the loop of adapter D is CAGCG (7a and 5b respectively in Appendix B.3, which contains the full strand sequences). The same pattern (i.e. minimal or no binding) is seen for the first 3 lanes (when no linker/an incompatible linker/a linker only compatible at the 3' end is used). Lanes 4 and 5 both show a shift in mobility, suggesting linker binding, but there is little difference in mobility between the structures in these two lanes. This suggests a weak interaction between the linker 3' sticky-end and the loop of adapter D.

nt sequences at the 5' T-motif (5a, 6a, 7a, and 8a in Appendix B.3) were each tested with either a 5 or 6 nt sequence at the 3' T-motif (5b and 6b in Appendix B.3). Six of these showed a similar pattern to that shown in Figure 4.3. For the remaining two, lanes 4 and 5 were more similar in appearance to each other, as shown in Figure 4.4. As mentioned previously, the rapid association and dissociation of T-motifs means that bands seen on the gel relating to T-motif structures represent a population of structures, only some of which contain the fully-bound T-motif. Bearing this in mind, I believe that the difference in appearance of the bands in lanes 4 and 5 of Figure 4.4 represent a shift in population due to the presence of structures in lane 5 which have linkers bound at both ends, which are absent from lane 4. Both of the occasions where a large shift in mobility was not seen between lanes 4 and 5 used the same 5b loop sequence at the 3' T-motif (CAGCG), which may provide particularly weak binding.

Binding of the linker reduces the electrophoretic mobility of the U-shaped structure (presumably due to an increase in the structure size). This reduction is greater when the linker is bound via both ends; presumably the square shape formed has a

larger cross-section than the flexible structure present when the linker is bound by one end only.

4.1.2 Studying geometry

In the T-motif origami structures that we envisage building (based on the structure shown in Figure 4.1), each T-motif is planar with 90° geometry (i.e. a 'T' shape, as shown in Figure 3.13). This requires each linker duplex to be of the correct length to form a planar rectangle within the structure upon incorporation. In the U-shaped structure above, the predicted optimal linker length for planar assembly is the same as the bottom section of the U-shape structure (2 turns of DNA or ∼21 bps; assuming that the sections of loop:sticky-end duplex at either end of the linker are approximately the same width as the diameter of a DNA double-helix[1]). In addition to being the correct length, the optimal linker length is an integral number of turns of DNA which positions both sticky-ends on the same side of the linker double-helix so that they are in the correct position to be incorporated into the U-shape structure (see Figure 4.2).

Simulation results (shown previously in section 3.2.2) have shown that there is flexibility at the junction point, and experimentally this has allowed the formation of 5' T-motifs with non-90° geometry (when the duplexes involved are constrained by other interactions within the structure; for example the structures shown in Figure 3.2a from [84]). Gel electrophoresis experiments were carried out to test whether 3' T-motifs are also able to form under such constrained conditions, by testing their formation when non-optimal linker lengths were used in the U-shaped structure. Using linker lengths greater or less than the optimum means that the four loops in the structure cannot (all) have perfect 90° geometry.

U-shaped structures with different linker lengths were analysed by gel electrophoresis (red structures in Figure 4.5). In order to ascertain whether the 3' sticky-end of the linker was binding, each lane also contained an identical U-shaped structure where the linker was only able to bind via the 5' sticky-end (due to the presence of an incompatible loop sequence in the D strand; green structures in Figure 4.5). Each lane also contained two distance markers (blue), which allowed accurate

[1] In B-DNA, the length of a 5 or 6 bp helical duplex of DNA (e.g. as formed between a sticky-end and loop within a T-motif) is ∼17-20Å, where the diameter of DNA double helix is 20Å [152].

measurement of the relative electrophoretic mobilities of both structures (as per section 3.1.1), shown in Figure 4.6. For each of the unbraced and braced structures, two repeats were carried out (with the lanes in a different order; as described in section 3.1.1).

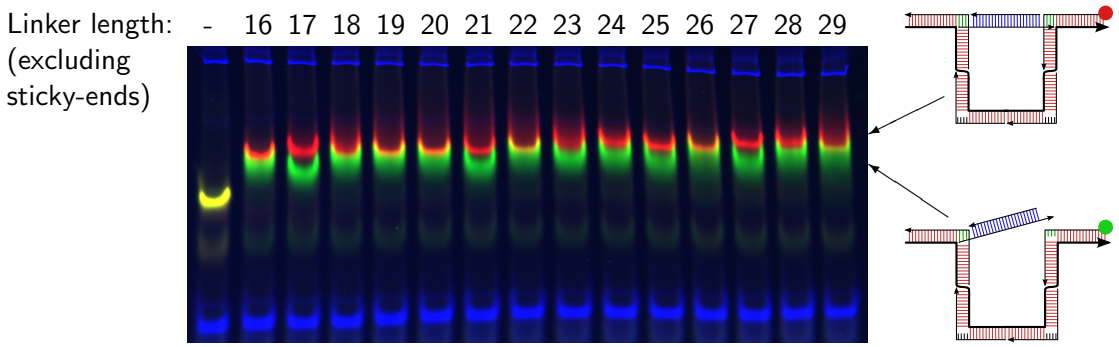


Figure 4.5: Gel electrophoresis of unbraced U-shaped structures with different linker lengths. Each lane contains two different U-shaped structures (see schematics on the right), differing only in the loop sequence and fluorescent label of adapter D. The linker used is compatible with both loops of the 'red' structure (GTGTCG in adapter A and CGACG in adapter D) but only the adapter A loop of the 'green' structure (the loop in adapter D is TCATC). All strands in each lane were annealed together prior to loading; for each of the two structures, the labelled strand was at 0.1 μM and the other strands were at 0.15 μM. Each lane also contains two distance markers (blue), which were used to accurately measure the relative electrophoretic mobilities of the red and green structures as per section 3.1.1. The results from this gel are repeat 2 of the unbraced structure shown in Figure 4.6. '-' indicates no linker.

For both braced and unbraced structures, both ends of the linker are able to bind (i.e. both 5' and 3' T-motifs are formed) at all linker lengths between 16 and 29 bps, as shown by the difference in electrophoretic mobility at each linker length of the red structure (with the compatible linker) vs the green structure (where the linker could bind via its 5' end only). This difference is not apparent in the all of the results shown in Figure 4.6, e.g. for a linker length of 26 bp the red and green plots are very close for the unbraced structure. However, gels which were run for a longer period of time did show a clear difference at all linker lengths: Figure 4.7 shows an example for the unbraced structure. These gels could not be used to obtain accurate relative electrophoretic mobility values as the high mobility distance marker ran off

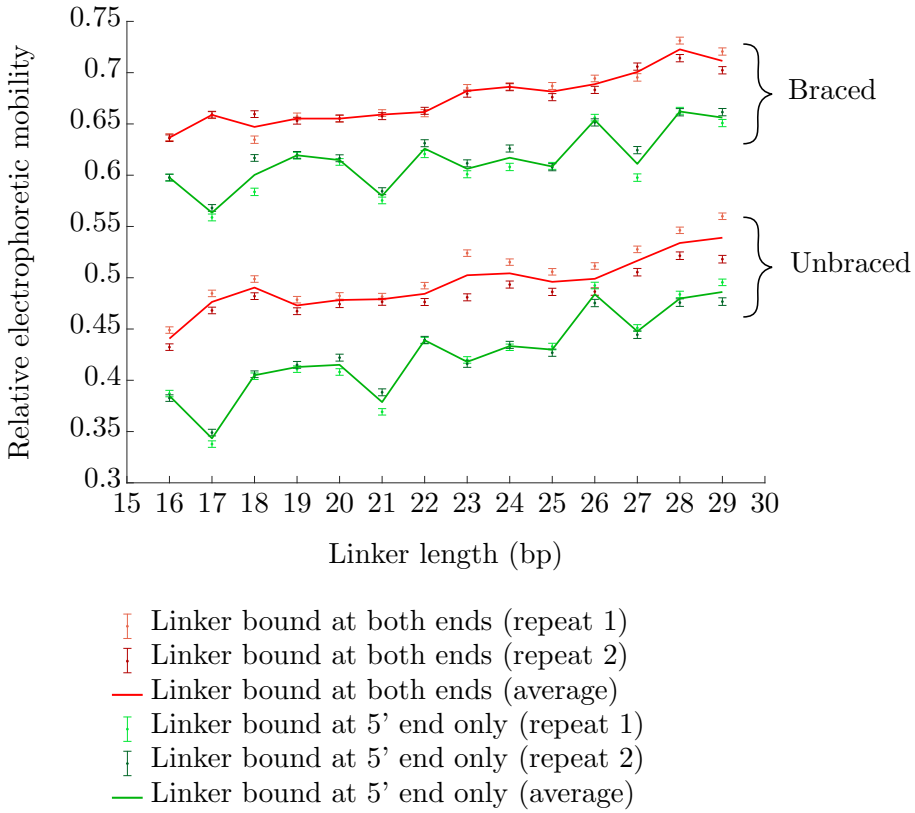


Figure 4.6: Relative electrophoretic mobilities of U-shaped structures with different linker lengths. For each of the unbraced and braced structures, two repeats were carried out (with the lanes in a different order; as described in section 3.1.1). The (blue) distance markers in each gel lane were used to measure relative electrophoretic mobilities as per section 3.1.1. Two repeats were carried out (with the lanes in a different order; as described in section 3.1.1). Red plots show the mobilities of the red structures (where the linker could bind at both ends), and green plots the green structures (linker bound at 5' end only). Error bars represent the resolution of the gel images (50 μm).

the bottom of the gels. This suggests that the flexibility around both the 5' and 3' T-motif junctions can accommodate quite significant deviations from the planar 90° geometry which is present when there are no constraints[2].

There are only very subtle changes in electrophoretic mobility when the linker length is changed, and so unfortunately the results cannot be used to confirm that 21 bp is the optimal linker length. However, simulations of the same strands have shown

[2]Figure 3.13 showed that (in simulations) both the 5' n=6 and 3' n=5 T-motifs used here have approximately planar 90° geometry on average, with both stacking sites formed, when there are no other constraints e.g. those imposed by being part of the U-shaped structure as in these experiments.

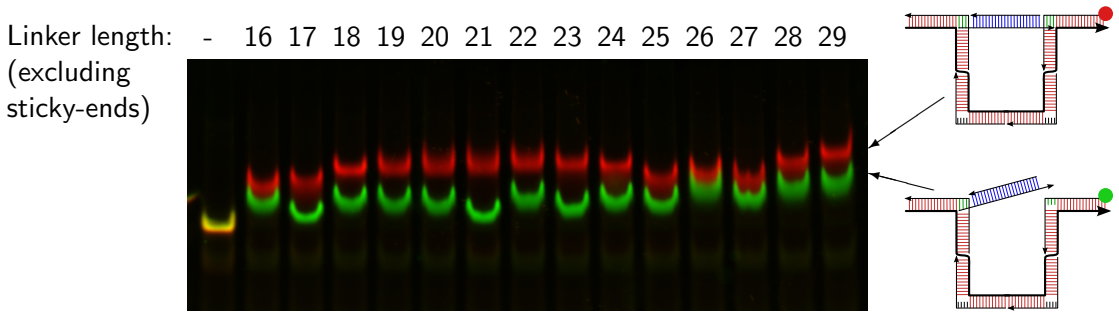


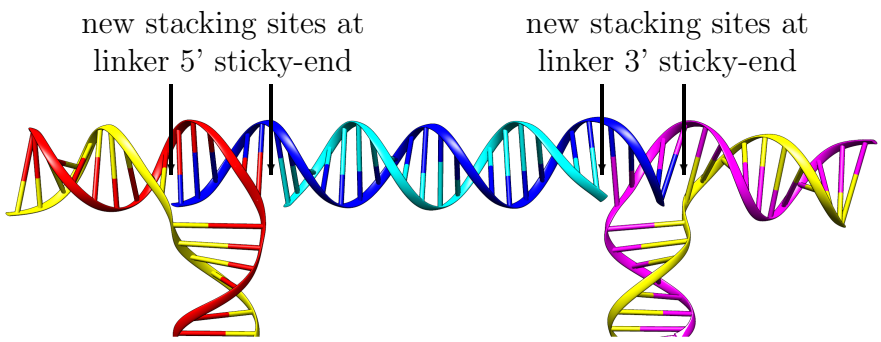
Figure 4.7: Repeat of gel shown in Figure 4.5, without distance markers and run for a longer period of time (2.5 hours compared to 2 hours). This gel shows a clear difference in the mobilities of the red and green structures in all lanes (except the first lane where no linker is present), showing that the linker is binding via both ends in the red structure at all linker lengths. All sequences and concentrations are identical to those of Figure 4.5. '-' indicates no linker.

that the highest level of stacking is seen at the two T-motif sites when linker lengths of 20-24 bps are used (Figure 4.8b). The degree of stacking gives us information about the positions of two of the arms of the T-motif junction as illustrated in Figure 4.8a: a higher level of stacking indicates that the continuous helix is forming correctly. These results therefore indicate that linker lengths of 20-24 bps are optimal for 'correct' (i.e. planar with 90° geometry) T-motif formation. Interestingly, the simulations also show relatively high stacking for 29-30 bps, i.e. around 3 turns of DNA. This makes sense, as a full number of turns will again position the two sticky-ends on the same side of the linker helix for correct formation of two T-motifs within the U-shape.

The clear pattern of variation of mobility with changing linker length seen for the 5'-compatible-only linker for both braced and unbraced structures is unexpected and unexplained; it is not apparent why increasing the linker length would have this (dramatic) effect on the electrophoretic mobility when the linker is only bound via one end.

Although our T-motif origami structures will be designed with (predicted) optimal linker lengths for planar assembly, the flexibility of both the 5' and 3' junctions means that these values do not have to be optimised to allow linker binding (e.g. we do not need to consider the potential effects on geometry of using different length

(a)



New stacking sites formed upon linker binding and formation of a 5' and 3' T-motif. 'Stacking probability' as plotted below is the probability of all four of these sites forming correctly.

(b)

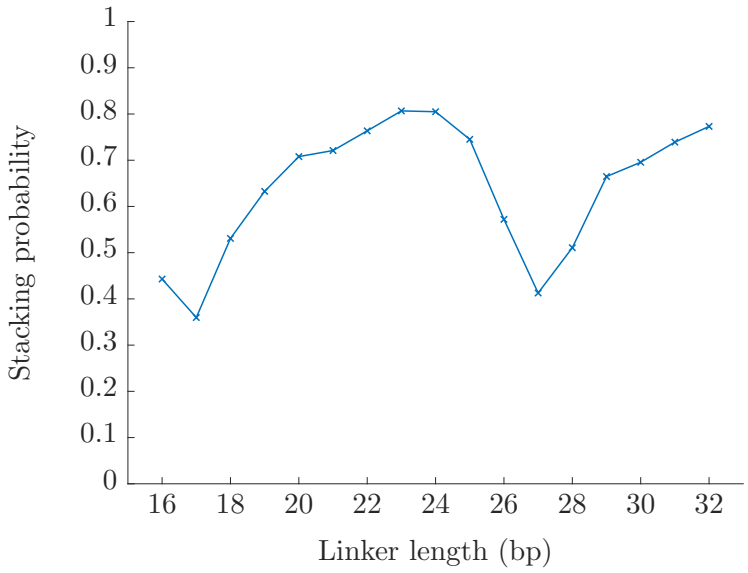


Figure 4.8: Data from oxDNA2 simulations showing how the probability of correct stacking within the U-shaped structure varies with linker length. The stacking probabilities shown in (b) refer to the probability that all four stacking sites (as shown in (a)) are formed.

loop:sticky-end duplexes). If many slightly non-planar units are combined in a much larger T-motif origami structure, the build up of stresses in the structure could result in twisting and defects, but for the size of structure tested in this chapter and the characterisation methods used (e.g. AFM, in which the adherence of DNA structures to the mica surface tends to flatten them out) this was not an issue.

4.2 T-motif origami-like structure

The characterisation of individual linker behaviour as presented in section 4.1 confirmed that using double-ended linkers with both a 5' and 3' sticky-end results in hierarchical or co-operative binding, due to the difference in binding strengths between these two ends. This should prevent excess linker binding i.e. where two copies of a linker bind, one to each of the binding domains, instead of one copy bridging both domains. In addition, it showed that the exact length of linkers does not have to be optimised based on the exact distance and geometry between its two binding domains (which can be affected by the loop lengths of these domains) in order for the linker to be successfully incorporated.

With these properties in mind, an intermediate-sized T-motif origami structure (smaller than the large structure shown in Figure 4.1), which uses redundant binding domains, was designed.

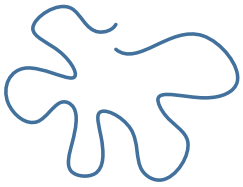
4.2.1 Design

As shown in Figure 4.9a, the structure uses a single-stranded template of ∼550 nt, a set of adapter strands (1-28) containing only four different loop sequences (so there are four different types of binding domain) shown in different colours, and three different linker types (A-C). As previously described in the introduction of this chapter, the template strands of T-motif origami structure do not have any sequence constraints and so can be freely designed: here it was based on a section of the pUC19 plasmid and artificially synthesised (see Appendices A.3.2.2 and B.4 for more information).

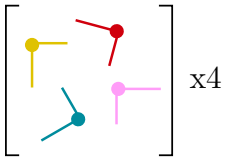
When the template is hybridised to all 28 adapters, the high degree of redundancy in the binding domains means that there are many possible positions at which each linker can bind. As Figure 4.9b shows, the first linker of type A to bind can potentially bind to any of the 7 'red' loops, and then to any of the 7 'pink', giving it 49 different potential binding positions. The next linker A can bind any of the 6 remaining 'red' loops and any of the 6 remaining 'pink' loops and so on, meaning

T-motif origami structure

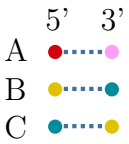
(a) Components



Single-stranded template

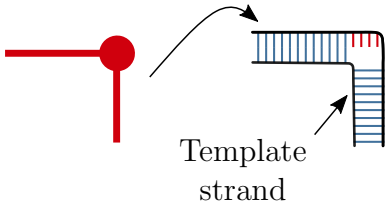


Single-stranded adapters: 28 unique strands, each with 1 of 4 different loop sequences*

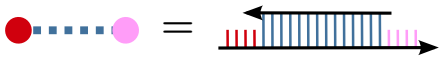


Double-stranded linkers: colours indicate which sticky-ends are complementary to the adapter loops

Key



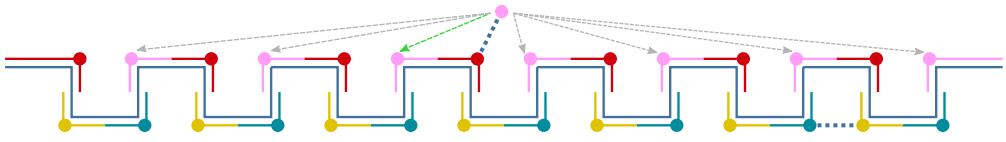
Colour represents identity of single-stranded loop which is formed upon hybridisation with the template strand



Linker has 5’ ‘red’ sticky-end (complementary to red loop) and 3’ ‘pink’ sticky-end. Each linker duplex has the same sequence for the hybridisation region and therefore the same short strand.

(b) A single copy of linker A can bind at multiple positions

The red sticky-end of the linker can bind at one of 7 different positions (red loops), following which the pink sticky-end also has a choice of 7 different positions



Template + all (28) adapters

Figure 4.9: Design of the T-motif origami structure. Four different loop:sticky-end interactions (i.e. sequences) are used, as shown by the four different colours. Of the 28 adapters, there are 7 with each loop identity; these differ in the sequences of the adapter either side of the loop which hybridise to the template. (a) shows the components, (b) shows the different positions in which a single linker could bind due to the high degree of redundancy in the binding domains. *See Figure 4.10c for loop sequences.

that there are ∼25 million different configurations in which 7 copies of linker A can bind to a single template-adapter duplex ($7^2 \times 6^2 \times 5^2 \times 4^2 \times 3^2 \times 2^2 \times 1^2$). There are the same number of configurations in which linker B or C can bind to the 'blue' and 'gold' loops, leading to the total number of configurations possible being $(25 \times 10^6)^2$ or $\sim 6 \times 10^{14}$. Some of these structures are sterically impossible, for example where parts of the structures occupy the same physical space, and many are extremely unlikely due to other constraints including the repulsion between DNA helices, and the limitations of the template-adapter duplex flexibility and the flexibility around T-motif junctions. For example, the interaction shown by a green arrow in Figure 4.9b is extremely unlikely given that it places the linker duplex and template-adapter duplex into very close proximity.

In addition, the binding of the adapters to the template (in the absence of linkers) is likely to impose some 'secondary structure' upon the template, for example bending at the locations of the single-stranded loops. This will affect the structures formed upon linker binding.

However, this still leaves a vastly higher number of potential structures than for conventional DNA origami, where each staple (the equivalent of the linkers here) has only one possible binding position and there is therefore only one target structure (this is ignoring staple mis-binding, which we assume would be minimal in a well-designed DNA origami structure).

4.2.1.1 Design of extended and compact structures

Two distinct structures were designed to fold using the T-motif origami components shown in Figure 4.9a, as shown in Figure 4.10. The 'extended form' (Figure 4.10a) consists of the template hybridised to the full set of adapters, then folded using linkers A and B. The presence of linkers bridging more distant binding domains was expected to be a minimal because a) more local interactions would form more quickly during the folding process, and b) bridging of more distant domains would reduce the entropy of the template and therefore be unfavourable.

Designing structures using T-motif origami components

(a) Extended structure

Template + all (28) adapters + linkers A and B →



Linkers

5’ 3’

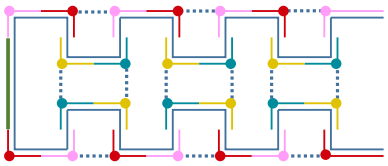
A

B

(b) Compact structure

Template + 26 adapters + ‘splint’ strand + linkers A and C →

splint strand: hybridises to same region of template as adapters 14-15, but contains no loops



Linkers

5’ 3’

A

C

(c) Loop sequences used in (a) and (b)

Adapter colour	**Loop sequence**	**Forms 5’ or 3’ T-motifs**
Red	CAGGTG	5’ only (with 5’ sticky-end linker A)
Pink	GGCTC	3’ only (with 3’ sticky-end of linker A)
Gold	CGACG	Both: 3’ with 3’ sticky-end of linker B (extended) and 5’ with 5’ sticky-end of linker C (compact)
Blue	GTGTCG	Both: 5’ with 5’ sticky-end of linker B (extended) and 3’ with 3’ sticky-end of linker C (compact)

Figure 4.10: Structures designed to fold from the T-motif origami components. The only differences in the components used to assemble (a) and (b) are: the compact structure (b) uses linker C instead of linker B, and also uses the splint strand instead of adapters 14 and 15. Both linkers B and C bridge ‘blue’ and ‘gold’ loops, but they have opposite sticky-end polarities: B has a 5’ ‘gold’ sticky-end and a 3’ ‘blue’ sticky-end, whereas C has a 5’ ‘blue’ sticky-end and a 3’ ‘gold’ sticky-end (see Figure 4.9a for detail of linker structures). (c) Loop sequences used in both structures (see Appendix B.4 for full strand sequences.)

In contrast, folding with linkers A and C was designed to produce the 'compact form' shown above in Figure 4.10b, where the template is essentially folded in half on itself. This folding in half results from the incorporation of six copies of linker C. In addition, two of the adapters (14 and 15; which provide binding sites in the middle of the template for a copy of linker A, which cannot bridge these sites when the template is folded in half as in the compact form) are omitted and instead a 'splint' strand which, when bound to the template, provides the correct geometry to allow the template to fold in half is included. This increases the relative stability of the compact structure further to make its formation favourable.

These two structures were designed to be sufficiently different in shape to allow them to be easily distinguishable by both gel electrophoresis and AFM imaging.

4.2.2 Results

For all results shown, the template strand and adapters/splint strand were thermally annealed (template at 0.2 μM, adapters at 0.3 μM). Linker strands were either present in this mixture and annealed at the same time, or added afterwards and incubated with the annealed template and adapters for 1 hour at room temperature (20°C) or 30°C. Linker duplexes were at 5 μM (i.e. at a 25× excess to the scaffold), which represents a 3-5× excess at each linker binding position given that 6-7 copies of the linker bind in each structure. All gels are agarose with magnesium in both the gel and running buffer as described in Appendix A.4. For ease of analysis, I have referred to the set of adapters present in the extended form (i.e. adapters 1-28) as the 'Ex' adapters, and the set of adapters present in the compact form (adapters 1-13 + 16-28, and the splint strand) as the 'Co' adapters.

4.2.2.1 Formation of extended and compact structures

Initial gel results showed that when the template was annealed with the adapters and linkers designed to fold it into either the extended form or the compact form, the difference in electrophoretic mobility between the structures was unexpectedly

small (Figure 4.11 lanes 2 and 4).

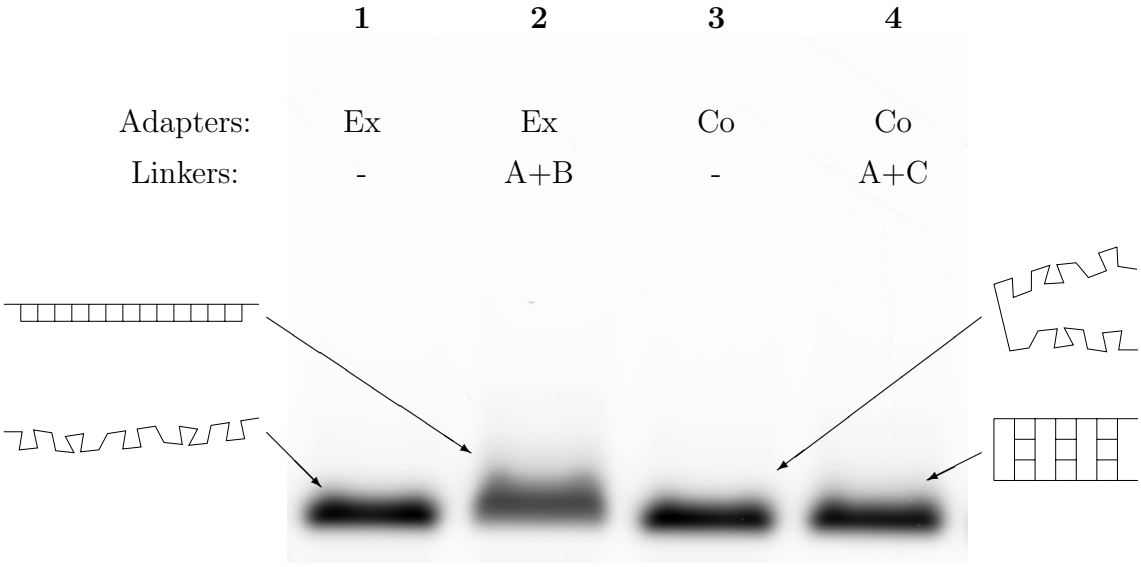


Figure 4.11: Gel results for the extended and compact structures (see Figure 4.10a and b) with and without linkers. All strands were annealed together (template at 0.2 μM, adapters at 0.3 μM and linker duplexes at 5 μM) prior to running. Lane 2 contains the template annealed with the adapters and linkers expected to fold into the extended structure (lane 1 is the same but without linkers); lane 4 contains the template annealed with the adapters and linkers expected to fold into the compact structure (lane 3 is the same but without linkers). Schematics on either side of the gel show the structures expected to be present in each lane (assuming linkers have bound correctly): black lines represent DNA duplexes. NB: structures with no linkers bound (lanes 2 and 4) are likely to be very flexible and not as rectangular as shown here. One adapter has a fluorescent label which is visualised here.

The binding of linkers in the extended structure does slightly reduce the electrophoretic mobility (Figure 4.11 lane 2 compared to lane 1), which is to be expected as the template and (either the Ex or Co) adapters alone should be quite a flexible structure which can migrate through the gel more quickly than the long rigid structure formed when the linkers are present. The compact structure (lane 4) is expected to migrate more quickly than the extended structure, and is practically indistinguishable from the structure in lane 3 (template + Co adapters, with no linkers present).

In order to confirm that the extended and compact forms are forming correctly (i.e. as they are shown in the design in Figure 4.10), AFM imaging was carried out (see Appendix A.6). Initially this was done on unpurified samples, but the majority

of images were of the excess adapter and linker strands. Imaging using a buffer with nickel omitted was also tested, in the hope that the larger structures would still be able to adhere to the mica but excess single strands would not: this did work, but unfortunately the structures were too susceptible to being moved around by the AFM tip to obtain good images. The images shown in Figure 4.12 and 4.13 were obtained using gel-purified samples and imaging buffer containing nickel.

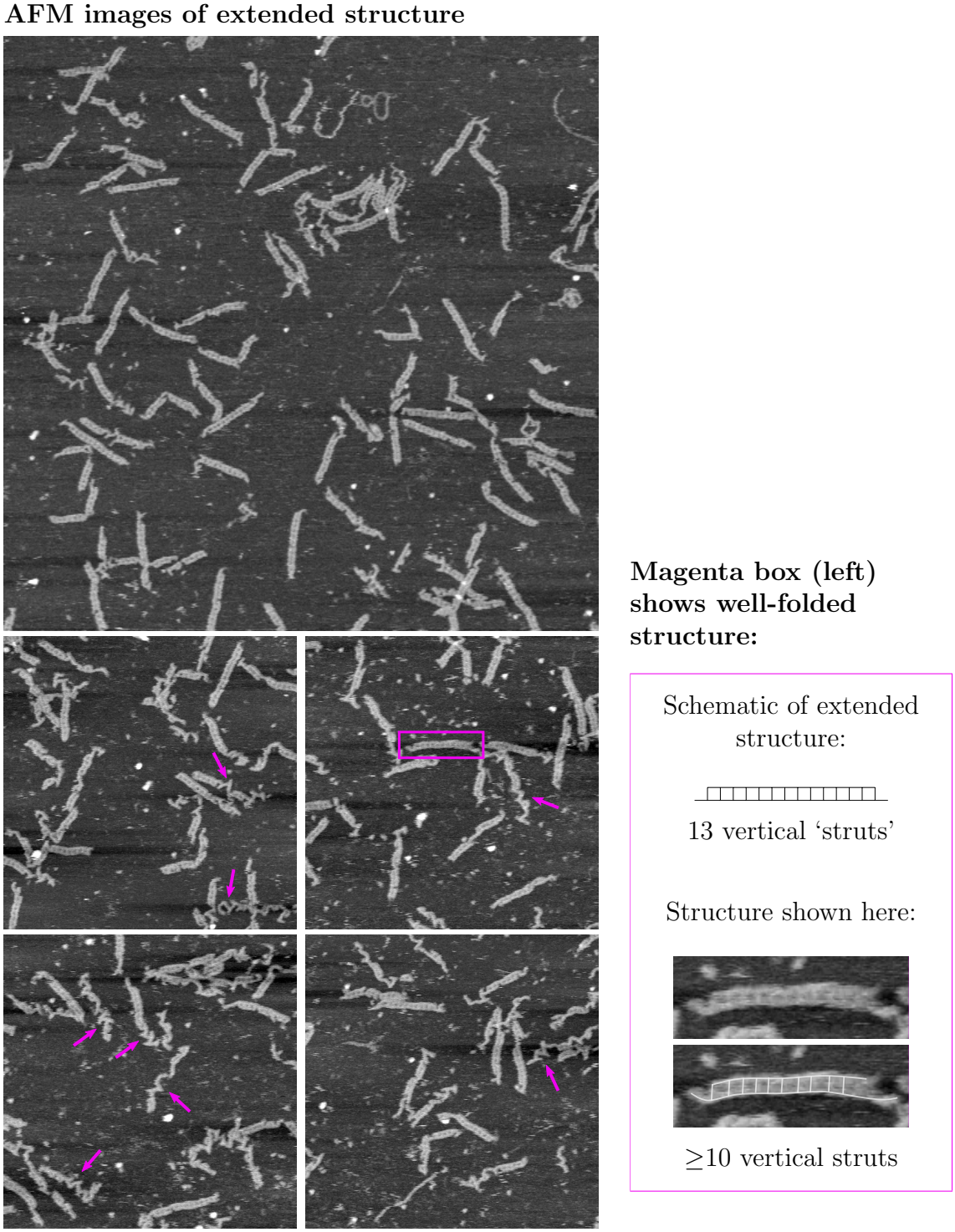
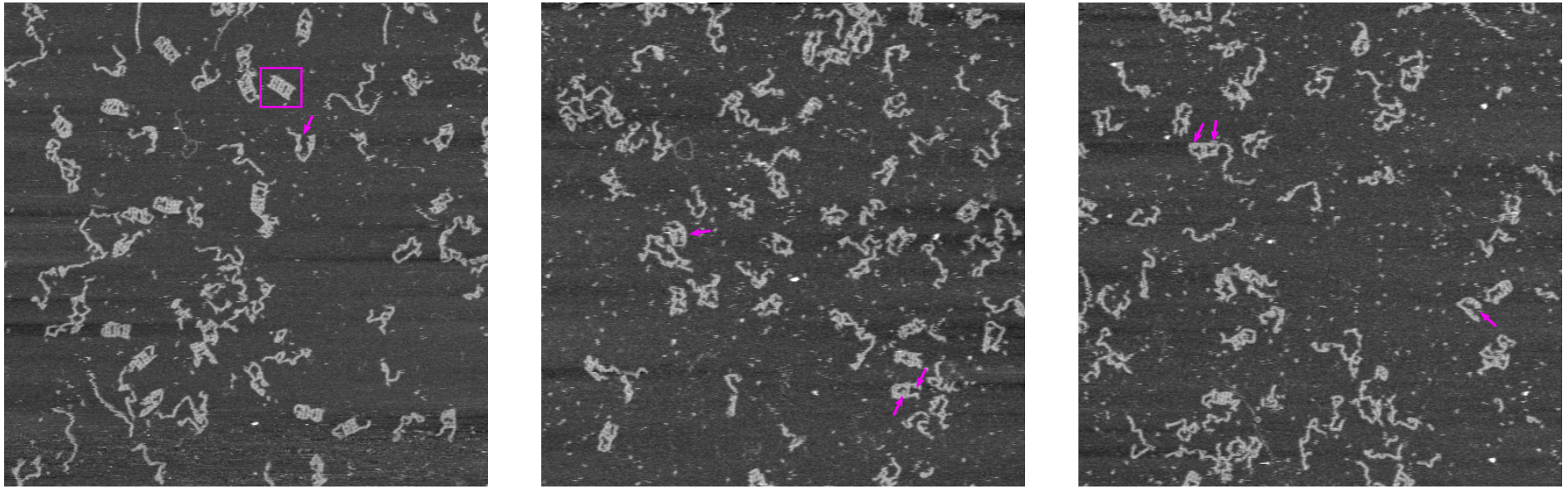


Figure 4.12: AFM images of gel-purified samples of the extended structure (all strands annealed together as per lane 2 of 4.11). Top image is 1 × 1μm, others are 0.5 × 0.5 μm. Features of interest have been annotated on the (higher resolution) 0.5 × 0.5 μm images. In the magenta box is an example of a (relatively) well-folded structure: when compared to the target structure (see right), it appears to be missing ∼3 vertical struts, although the right-hand end of the structure is difficult to resolve. Arrows point to a defect (the presence of 'zigzag' shapes) which is commonly observed for the extended structure. These appear to arise from linkers binding via their 5' sticky-end only (see Figure 4.14).

AFM images of compact structure



Magenta box (above) shows well-folded structure:

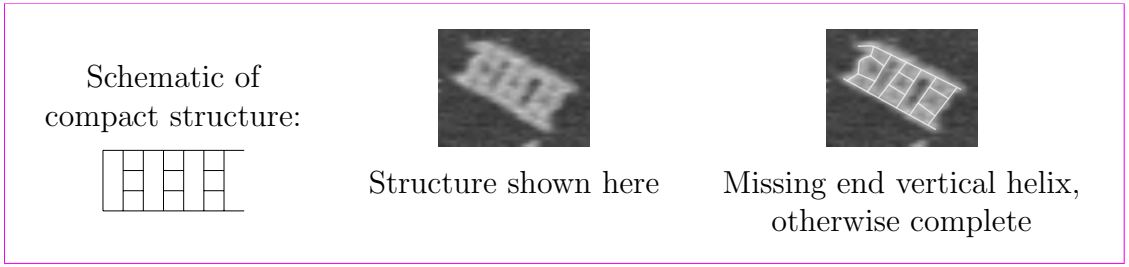
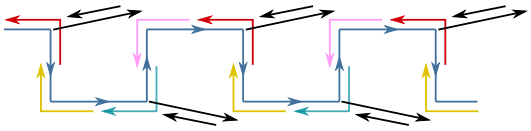


Figure 4.13: AFM images of gel-purified samples of the compact structure (all strands annealed together as per lane 4 of 4.11). All images are $1 \times 1\mu$m. Features of interest have been annotated as for Figure 4.12. In the magenta box is an example of a (relatively) well-folded structure: when compared to the target structure (see above), it appears to be missing the end helix which is formed by the template hybridised to the splint, possibly due to this being pushed around during imaging. Arrows point to a commonly-observed defect in the compact structure, which appears to arise from incomplete binding of linker C).

The AFM images show that a majority of structures are incomplete or improperly folded to some extent. For the extended structure, a commonly-observed defect is the 'zigzag' shapes which have been identified with arrows in Figure 4.12. Such structures may be due to linkers binding via their 5' sticky-end only (which is a stronger interaction than that of the 3' sticky-end), as shown in Figure 4.14.

(a) Section of extended structure with linkers (black) bound via their 5' ends only:



(b) Formation of these 5' T-motifs bends the template-adapter duplex through 90°:

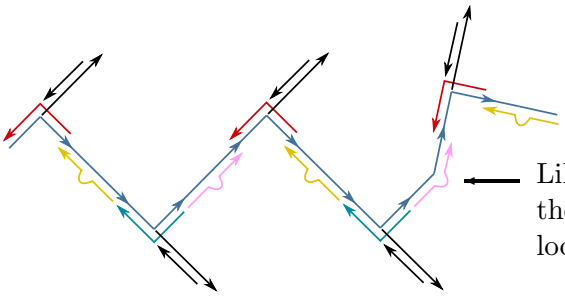


Figure 4.14: Proposed formation of 'zigzag' structures seen in AFM images of the extended structure. (a) 5' T-motifs are inherently stronger than 3' T-motifs, and so there may be structures where only the 5' sticky-ends of linkers are bound. (b) Binding (and thus formation of T-motifs) of the 5' sticky-ends will result in the template-adapter duplex bending through 90°. In contrast, template bound to adapters which have no linker bound (the yellow and pink adapters above) are likely to form a flexible but approximately linear duplex.

For the compact structure, it appears that the linkers which cross-link the two halves of the structure (linker C) are more frequently absent (or possibly half-bound, as for the extended structure, but this is not clear), as identified by the arrows in Figure 4.13. There are also many structures where copies of both types of linker are missing.

The prevalence of structures with incomplete linker binding may be due to the removal of excess linkers (by gel purification) and overall dilution of the DNA strands

prior to imaging. Reducing the strand concentrations and the linker excess lowers the melting temperatures of the T-motif interactions, and so they are more likely to dissociate at room temperature. In conventional origami, it is also common to purify out excess staples prior to AFM imaging [75, 143]. However, the higher melting temperatures of the staple-template interactions in conventional origami, as compared to the T-motifs used here, means that staples will take longer to dissociate, allowing imaging to occur before the structures fall apart.

In addition, these T-motif origami structures are more prone to being damaged during AFM imaging than more compact structures like the origami tiles shown in Chapter 2. Indeed, there was observable damage inflicted when structures were imaged more than once, including the disappearance of linkers. AFM-associated damage may therefore also contribute to the high proportion of incomplete structures observed.

In future experiments it may be possible to cross-link structures, for example with formaldehyde or glutaraldehyde [93], prior to imaging to prevent linker dissociation.

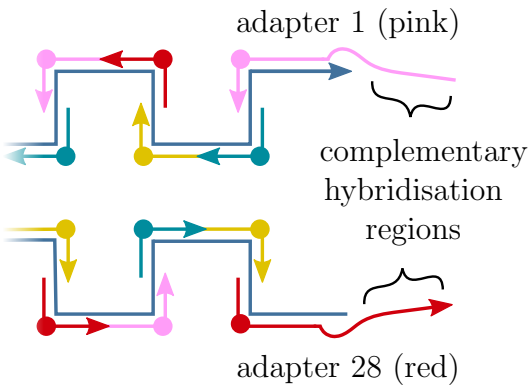
4.2.2.2 Distinguishing between structures on gels

Having confirmed that the two different structures are able to form correctly, the original designs were amended to allow them to be more easily distinguishable by gel electrophoresis. Single-stranded overhangs containing short complementary hybridisation regions were added to the adapters at either end of the template strand (as shown in Figure 4.15a). In the extended structure, these overhangs are at either end of the linear structure, so allow multiple copies of the extended structure to bind to each other and form multimers. In contrast, the overhangs are on the same side of the compact structure and are able to form an intramolecular interaction. Gel results (Figure 4.15b) confirmed this behaviour, although the compact structure does form some higher order structures.

When no linkers are present, both types of template-adapter complex (lanes 1 and 3) are predominantly present as monomeric structures. The flexibility of these

(a)

Positions of single-stranded overhangs on first and last adapters:



(b)

	1	2	3	4
Adapters:	Ex	Ex	Co	Co
Overhangs:	✓	✓	✓	✓
Linkers:	-	A+B	-	A+C

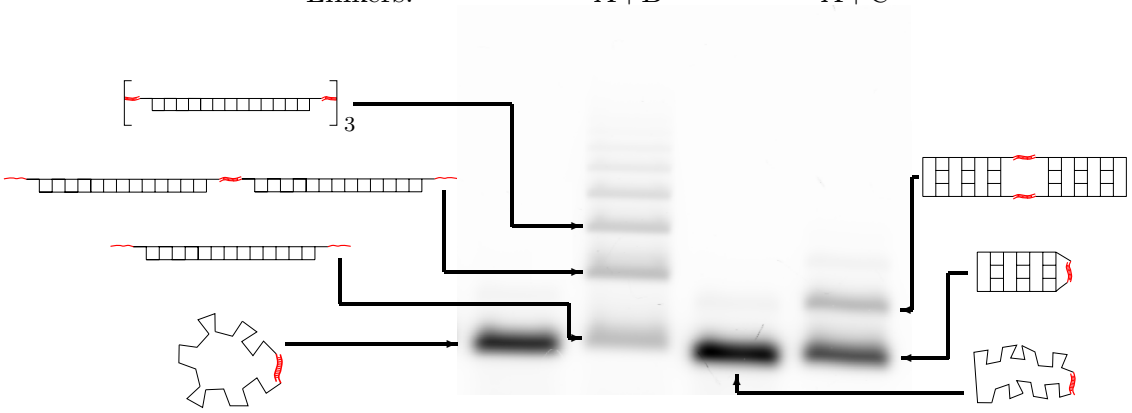


Figure 4.15: Effect of adding single-stranded overhangs to the first and last adapters of the T-motif origami structure. (a) Positions of the overhangs: overhangs were added to adapters 1 and 28 (which bind to the 3' and 5' end of the template strand, respectively), each containing a 6-10 nt hybridisation region which is complementary between the two overhangs. (b) Gel results for the extended and compact structures when adapters with overhangs containing a 10 nt hybridisation region are used, with and without linkers. Lane order and sample preparation is the same as for Figure 4.11. For both the extended and compact structures, introduction of the overhangs results in the formation of multimeric structures when linkers are present (lanes 2 and 4). Schematics on either side show the proposed structures represented by bands in the gel: overhangs are shown in red and black lines represent DNA duplexes. The proposed structures in lane 4 are shown in more detail in Figure 4.17. NB: structures with no linkers bound (lanes 2 and 4) are likely to be very flexible and not as rectangular as shown here. One adapter (adapter 2) has a fluorescent label which is visualised here.

structures presumably allows the adapter overhangs to bind intramolecularly, as shown in the schematics in Figure 4.15b. There is also a small amount of dimer present in these lanes (which is not seen in the absence of overhangs, see Figure 4.11), suggesting a small amount of intermolecular overhang hybridisation.

When linkers are present, the extended structure forms a ladder of different size structures which represent multimers containing different numbers of monomers (this was confirmed by plotting of the electrophoretic mobilities against the expected size of monomers, as described for an origami structure in point 3 of section 2.4.6). These multimers are consistent with the long rigid monomers joining end-to-end. If we assume that the probability of addition of a monomer to another monomer/existing multimer is constant (and that thermodynamic equilibrium is reached), we would expect the distribution of different size multimers to follow an exponential curve: the amount of multimer should decrease exponentially as the size of multimer increases. Figure 4.16 shows the amount (a measure of the number of molecules) of each type of multimer in lane 2 of Figure 4.15, plotted on a logarithmic scale. If the results followed the expected exponential decrease with increasing multimer size, this plot would be a straight line. The fact that the line curves downwards shows that the actual amount of each type of multimer present decreases much more dramatically with increasing size than expected. This deviation from the expected result could be due to larger structures falling apart during gel electrophoresis. Alternatively, structures containing enough multimers may be able to circularise and thus fulfil overhang binding this way, as opposed to adding in more monomers. This would also reduce the number of larger multimers seen. Unfortunately the measurements of electrophoretic mobility of the multimers here cannot distinguish between circular and linear structures.

The compact structure predominantly forms monomers (as expected) and to a lesser extent, dimers (Figure 4.15b lane 4). Unlike the extended structure, both of the overhangs are on the same side of the structure, and so it is possible to form both monomers and dimers where the hybridisation regions of the overhangs are

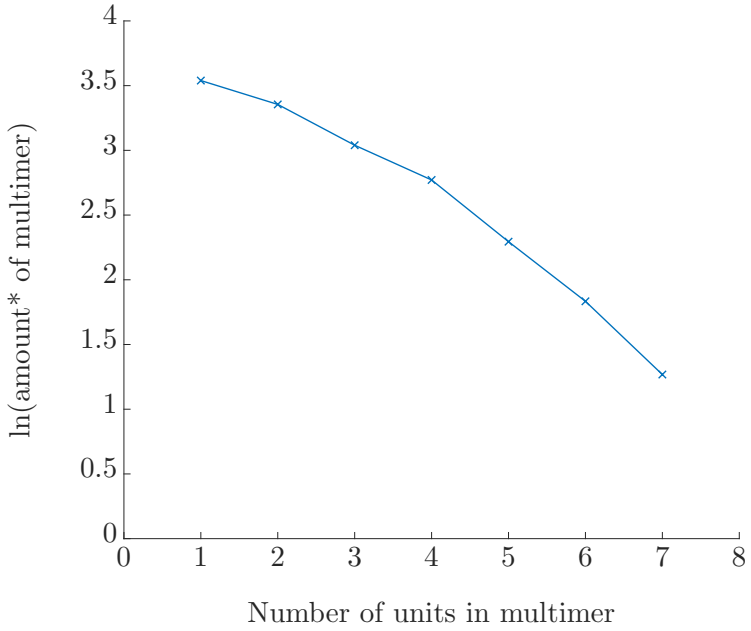


Figure 4.16: Plot of amount* of multimer vs number of units in multimer, for lane 2 of Figure 4.15.
*Amount is a measure of the number of molecules, calculated as the band intensity divided by the number of fluorescent labels per molecule.

fully satisfied (as shown in Figure 4.17). There is also a trace amount of a trimeric structure (although not when an 8 nt overhang is used as in Figure 4.18b). It not possible to infer from the electrophoretic mobility whether these structures consist of three or more monomers joined in a circle or a chain.

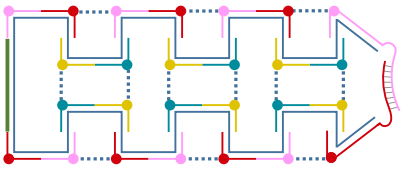
4.2.2.3 Isothermal folding

For all of the results previously shown, all strands (the template, adapters, and linkers where used) were thermally annealed in one pot.

As described in section 4.2.1, there are many different structures that could be formed using the components of the T-motif origami system. Even when only two of the three linkers are used (as is the case in the results shown so far), there are many possible positions in which each linker could bind. The fact that the gel results shown so far are consistent with the formation of well-folded extended and compact structures (we do not see a distribution of different products[3]) could be ascribed to

[3]except in Figure 4.15, where the electrophoretic mobilities of the products are consistent with them being multimers of the well-folded product.

(a) Monomeric compact structure with overhangs



This requires flexibility in the structure as the distance between the pink and red loops at the right-hand end is only 39 bp + 11 nt (some regions of the overhangs are single stranded) vs 58 bp on the left-hand side.

(b) Dimeric compact structure with overhangs

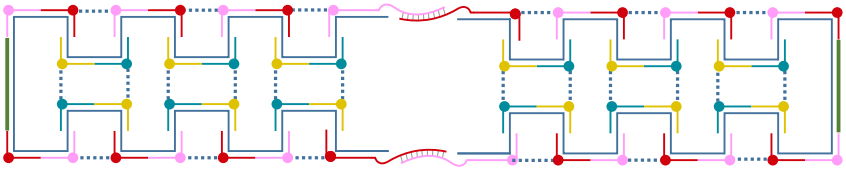


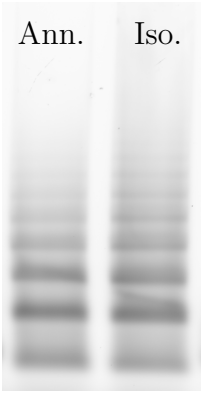
Figure 4.17: Possible hybridisation of adapter overhangs in (a) monomers and (b) dimers of the compact structure. All the hybridisation regions of the overhangs are fully satisfied in both structures, hence the compact structures is predominantly monomeric with a small amount of dimer, in contrast to the multimers seen for the extended structure.

the use of thermal annealing, which will promote formation of target interactions (since these are likely to be marginally more thermodynamically favourable, so can form at higher temperatures i.e. earlier in the annealing process) and allow linker rearrangement in order for the structure to reach a thermodynamic minimum. (We expect the well-folded extended and compact structures to be more thermodynamically stable than their misfolded counterparts, since misfolded structures will likely be able to incorporate fewer linkers and may have additional instability due to repulsion between physically-close sections of the structure and/or strain on the template strand.)

To investigate whether using isothermal conditions increases misfolding of the extended and compact structures, template and adapters were annealed together prior to the addition of linkers, and the mixture incubated for 1 hour at either 20 or 30°C. Prior to this, an overhang hybridisation length suitable for such experiments was determined: when overhangs with a 10 nt hybridisation region (as in Figure

4.15b) were used, isothermal incubation with the linkers failed to break apart the intramolecular overhang-overhang interactions which formed in the initial annealing step, and so no multimers were formed. A weaker (8 nt) hybridisation region was therefore used, the results of which are shown in Figure 4.18, alongside structures formed by thermal annealing using the same components.

(a) Extended structure



(b) Compact structure



Figure 4.18: Annealed ('Ann.') vs isothermal ('Iso.') folding of the (a) extended and (b) compact structures, with overhangs containing an 8 nt hybridisation region. In all lanes, template and adapters (the 'Ex' set for the extended structure and the 'Co' set for the compact structure) were annealed together (template at 0.2 μM, adapters at 0.3 μM). In the annealed lanes, the linkers were also present (linkers A and B for the extended structure, A and C for the compact structure; at 5 μM) during annealing. In the isothermal lanes, linkers were added after annealing (again at 5 μM), and the mixture incubated at 30°C for 1 hour.

Figure 4.18 shows that there are no significant differences between the mobilities and distribution of products observed by gel electrophoresis between thermally-annealing the linkers with the template-adapter complex (the scaffold) or incubating isothermally the linkers with the scaffold at 30°C. Structures incubated at 20°C were identical to those incubated at 30°C. At 20-30°C, it is very likely that linkers can and do form off-target interactions, and so the fact that the final result is identical to thermal annealing (which we know produces well-folded structures) suggests that linker rearrangement is possible at these temperatures. This means that it might be possible to produce structures which are reconfigurable at room temperature, an idea which is explored further in section 4.2.2.4.

When using 8 nt hybridisation regions in the overhangs (Figure 4.18a), the extended structure produces very similar results to when 10 nt is used (Figure 4.15b lane 2), whereas for the compact structure there is much more dimer relative to monomer (and no trimeric structures present) when shorter overhangs are used (Figure 4.18a vs Figure 4.15b lane 4; when folding by thermal annealing). It is likely that for both overhang lengths, the intramolecular overhang-overhang interactions present in the compact monomer structure form quickly and therefore early in the folding process. For the longer (10 nt) overhangs, this interaction is stronger and so is less likely to be reversed and replaced with the intermolecular overhang-overhang interactions which form the dimer and trimer compact structures. The results presented in the following section (4.2.2.4) use 8 nt hybridisation regions.

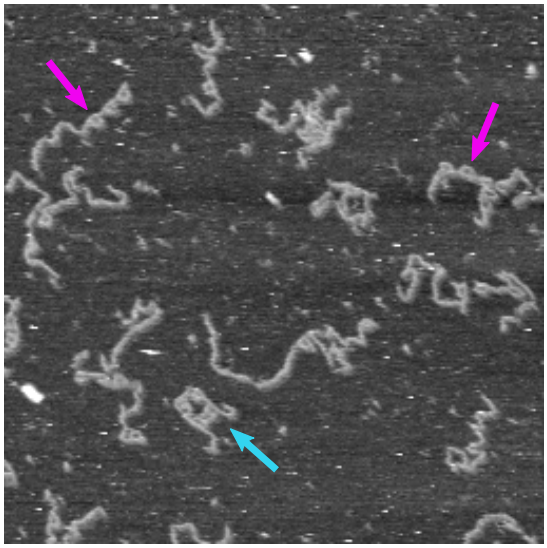
4.2.2.4 Folding preferences and reconfigurability

The template and the adapter set for either the extended structure (Ex) or the compact structure (Co) were annealed or incubated with the linkers from the other structure. The structures formed were determined by AFM and gel analysis (Figures 4.19 and 4.20). Knowledge of how these gel appearances correlate to the actual structure was important to allow interpretation of gels shown later in this section. All gel results are using overhangs with 8 nt hybridisation regions to allow multimerisation; AFM results are from gel-purified samples when no overhangs were used.

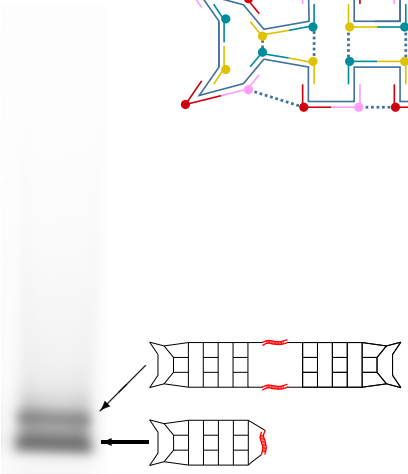
The results suggest that the extended adapters + compact linkers form a 'compact-like' structure which differs only in that adapters 14 and 15 are bound instead of the splint (Figure 4.19). For the compact adapters + extended linkers, the structure is similar to the extended structure except for the presence of a flexible region in the middle (Figure 4.20). This makes a dramatic difference to the multimerisation behaviour seen on the gel: whereas the rigid linear extended structure forms large multimers in order to satisfy the overhang hybridisation regions, the structure here can circularise with many fewer units. This is facilitated by the presence of the cen-

Extended adapters (Ex) + compact linkers (linkers A and C)

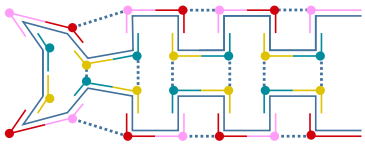
(a) AFM image (0.5 × 0.5 μm)



(b) Gel



(c) Proposed structure when all linkers are bound*:



*This could not be inferred from the AFM images, so I have based it on the compact structure, with the assumption that the binding of adapters 14 and 15 will induce 'kinks' in the template where the single-stranded loop regions are located.

Figure 4.19: AFM and gel results of template annealed with extended adapters and compact linkers (linkers A and C; strand concentrations as per Figure 4.18). Isothermal incubation of the annealed template + adapters with the linkers produced the same gel result (AFM imaging not carried out). Gel structures have overhangs to allow multimerisation; AFM structures do not.

(a) AFM imaging shows some structures where it appears that only linker A (which is common to both structures) has bound (pink arrows; although if all linker A copies are bound to the template this should produce 7 squares whereas only 3-5 are seen in the two structures indicated here). Other structures clearly have the cross-linking linker (C) bound and so appear 'compact-like' (blue arrows).

(b) Gel shows monomeric and dimeric structures, similar in appearance to when these linkers are used with the correct (compact) adapters (Figure 4.18b), suggesting formation of a 'compact-like' structure as shown in the schematics to the right of the gel. The dimer has a lower electrophoretic mobility than the true compact dimer.

(c) Proposed structure when all loops are bound. Adapters 14 and 15 are bound where the template folds in two, as opposed to the splint strand of the compact structure. This difference may explain the lower electrophoretic mobility of the dimer compared to the compact structure dimer.

Compact adapters (Co) + extended linkers (linkers A and B)

(a) AFM image (0.5 × 0.5 μm) (b) Gel (c) Proposed structure:

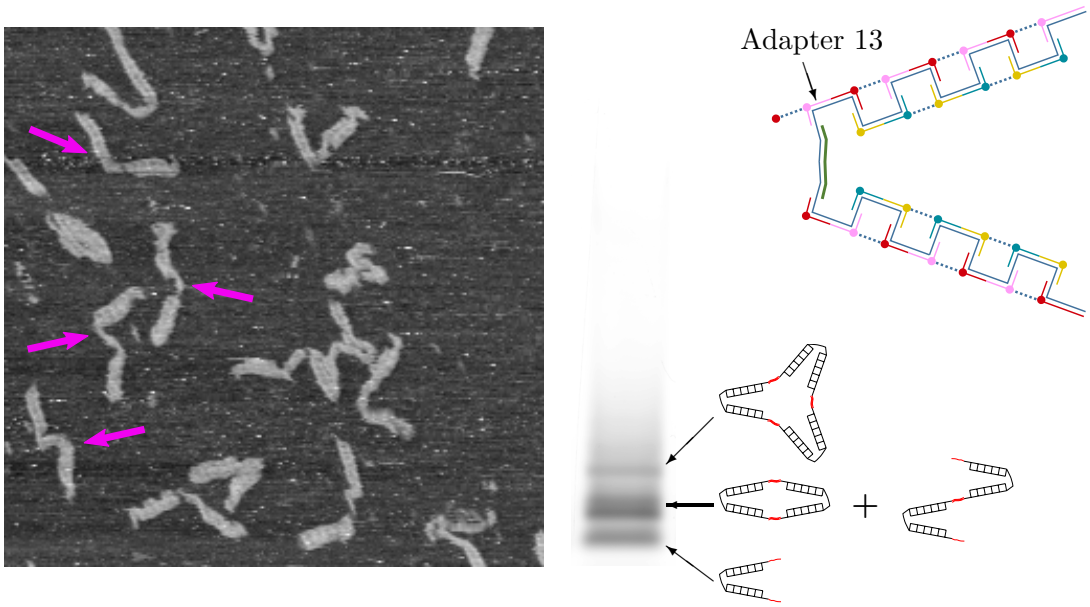


Figure 4.20: AFM and gel results of template annealed with compact adapters (including splint strand) and extended linkers (linkers A and B; strand concentrations as per Figure 4.18). Isothermal incubation of the annealed template + adapters with the linkers produced the same gel result (AFM imaging not carried out). Gel structures have overhangs to allow multimerisation; AFM structures do not.

a) AFM imaging shows structures which look similar to the extended structure, but with a thin, flexible region in the centre, which is presumably where the splint is bound to the template. The angle between the two halves of the structure suggest that there is formation of a 5' T-motif at adapter 13 (as shown in c) due to a linker binding via its 5' sticky-end, a similar effect to that seen for the extended structure (Figure 4.14).

b) Gel shows monomeric, possibly 2 different types of dimeric structure, and a higher order structure. Both dimer bands have a higher electrophoretic mobility (suggesting that they are more compact) than the true extended dimer (Figure 4.18a) and lower than the true compact dimer (Figure 4.18b). Schematics to the right show proposed structures based on these observations.

c) Proposed structure, including 5' T-motif at at adapter 13 (see a). This structure is seen clearly in the AFM results, and fits with the appearance of the gel (see schematics to the right of the gel).

tral flexible region, and may be aided by the formation of a 5' T-motif at adapter 13 as shown in Figure 4.20c. A linker may also be binding at this position in the compact structure and in the structure formed by the extended adapters + compact linkers; its effect would be less obvious in either of these since the binding of linker C also results in bending at this corner of the structure.

These results were used to aid interpretation of further gel experiments where overhangs containing 8 nt hybridisation regions were again used. The preparation of these samples is shown in Figure 4.21 and the gel results in Figure 4.22. Template was annealed with the set of adapters which are common to both the extended and compact structures (i.e. adapters 1-13 and 16-28) along with the linkers from the extended structure (linkers A and B). This was expected to form the structure shown in Figure 4.21a, which has a central section of single-stranded template. This structure has a very similar appearance on a gel (Figure 4.22 lane 1) to that of the compact adapters + extended linkers (Figure 4.20b), for which we have AFM images (Figure 4.20a). We thus assumed that the structure in Figure 4.21a is correct as it resembles these AFM images.

Various adapters were then added (Figure 4.21b-e) which bind to the template in the single-stranded region: either adapters 14 and 15 of the extended structure (b), the splint strand of the compact structure (c), or all three of these strands (d and e). The mixture was then either incubated at room temperature (20 °C) for 1 hour (b-d) or thermally annealed (e).

The structures shown on the right in Figure 4.21 are those inferred from the gel results shown in Figure 4.22. Isothermal incubation of the structure with the central single-stranded region with adapters 14 and 15 (Figure 4.21b) produces a ladder of different sized multimers (Figure 4.22 lane 2). We know that this gel appearance is characteristic of the extended structure, where the rigid linear shape means that in order to satisfy the hybridisation regions of the overhangs, more and more units must be added (see Figure 4.15b). This result suggests that the adapter strands are able to bind to the single-stranded region of the template, and incorporate a copy of

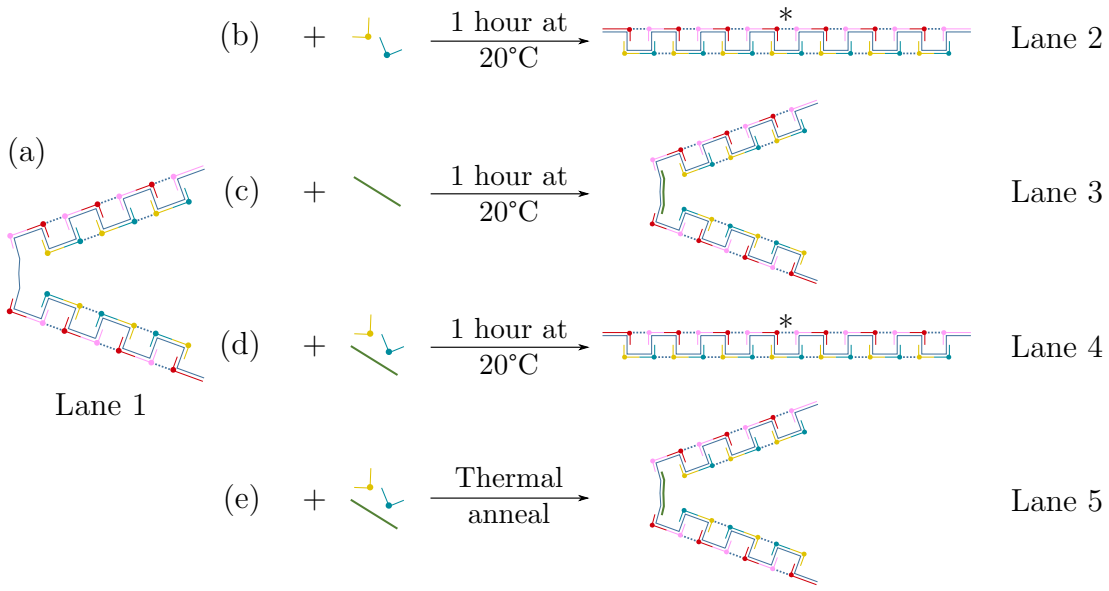


Figure 4.21: Isothermal incubation and thermal annealing with various adapters. The structure shown in (a) was prepared by annealing template with adapters 1-13 + 16-29 (those which are common to both the extended and compact structures) and linkers A and B (from the extended structure). It has a central region where the template is single-stranded. Adapters were then added which could bind to the single-stranded region of template: (b) adapters 14 + 15 (from the extended structure); (c) the splint strand (from the compact structure); or (d) and e) all three of the these structures. The mixture was incubated at 20°C for 1 hour (b)-(d) or thermally annealed (e). The structures shown on the right above are inferred from the gel results shown in Figure 4.22 (lane numbers shown here on the right). *Extra copy of linker A has been incorporated as a result of adapters 14 and 15 binding (see main text).

linker A (which bridges the two), which straightens out the structure into the rigid, linear extended form.

Isothermal incubation with the splint strand (Figure 4.21c) results in no change in the gel appearance (Figure 4.22 lane 3 is identical in appearance to lane 1). As mentioned above, we expect the structure to be very similar whether the splint is bound or not.

Finally, isothermal incubation with adapters 14 and 15 AND the splint (Figure 4.21d) results in a gel appearance similar to that of the extended structure (Figure 4.22 lane 4), suggesting adapters 14 and 15 (and a copy of linker A) are preferentially binding over the splint strand. This is in contrast to the same experiment where adapters 14 and 15 and the splint are added prior to annealing (Figure 4.22e and

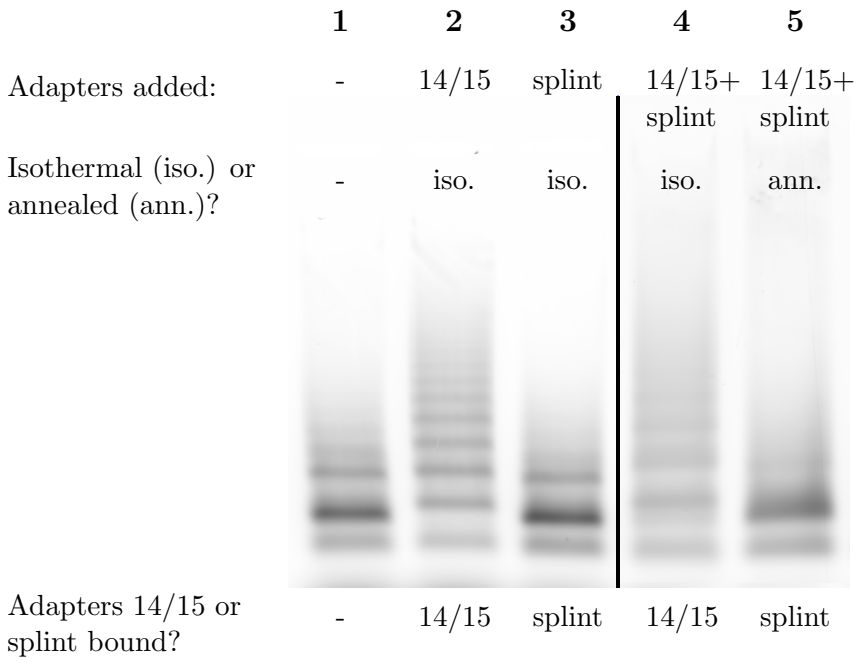


Figure 4.22: Gel results of experiments where the template annealed with adapters common to both the extended and compact structures and linkers A and B (lane 1) was incubated isothermally (lanes 2-5) or annealed with: either adapters 14 and 15 from the extended structure (lane 2), the splint strand from the compact structure (lane 3), or all three strands (lanes 4 and 5). Vertical line shows delineation between 2 separate gels. The gel results were used to infer which adapters are bound (as shown under the gel image), see main text for explanation. Sample preparation and proposed structures shown in Figure 4.21. Strand concentrations as per Figure 4.18.

Figure 4.22 lane 5), where it appears that the splint has bound.

The results show that when template is annealed with adapters 14 and 15 and the splint strand (Figure 4.22 lane 5), the splint strand preferentially binds (the results are identical to when the splint strand alone is added, Figure 4.22 lane 3). However, isothermally incubating these three strands with the template produces a different result (Figure 4.22 lane 4); the ladder of different size structures suggests binding of adapters 14 and 15 (rather than the splint strand) in at least some of the structures present. I propose that the difference between the results in Figure 4.22 lanes 4 and 5 is due to linker binding, which is able to take place at the temperatures used for the isothermal incubations but not at the high temperatures during which the adapters and splint become hybridised to the template during thermal annealing. During thermal annealing, the splint will bind to the template earlier than adapters 14 and

15 because it has a higher melting temperature. During the isothermal incubations, either adapters 14 and 15 or the splint can bind (all have melting temperatures much higher than 20°C), but the binding of a copy of linker A to adapters 14 and 15 (and thus the formation of two T-motifs) can stabilise their interaction with the template strand. It is likely that structures with the bound splint are also produced, which may explain the difference in appearance of Figure 4.22 lane 4 compared to lane 2.

4.3 Conclusions and further work

In this chapter I have presented the design and successful assembly of intermediate-sized T-motif origami structures, where each binding domain is part of a double-stranded template-adapter duplex, and folding is mediated through the interaction of linkers with these domains. Each linker has a 5' and a 3' sticky-end, which are able to form a relatively strong and a relatively weak interaction with the binding domains, respectively.

These structures have an extremely high level of redundancy in their binding domains, and thus are very different to conventional origami (see section 1.2.2), where each staple (the equivalent of the linkers here) can only bind in one position. This means that unlike conventional origami, there are many possible different structures that can be formed by the T-motif origami structures when they interact with the linkers. Despite this, the yield of 'off-target' structures appears to be low (where the target was the extended or compact structures), even when the structures are folded isothermally at room temperature. Contrast this with origami folding (Chapter 2), where isothermal folding at room temperature does not produce correctly-folded structures, despite the lack of redundancy.

The key difference is that the T-motif origami system uses relatively weak interactions (even the 5' T-motifs are a much weaker interaction that the hybridisation of a 16 nt staple domain with its binding domain), which can be re-arranged at room temperature in order to reach a more stable structure. In conventional origami, even staples which are mis-bound to the template bind too strongly for this interaction to be reversed (in a reasonable length of time) at room temperature.

This reversibility means that T-motif origami structures can be reconfigured at room temperature, as the linker connectivity can change in response to changes in the binding domains which are available. However, the corollary of this is that each possible structure is not significantly more stable than others which may be able to form. This is in contrast to conventional origami, where the target structure represents a thermodynamic minimum. These differing properties to conventional

origami may suggest future applications for T-motif origami systems where conventional origami is not suitable.

4.3.1 Further work

The next characteristic of T-motif origami systems to investigate is whether the rearrangement of linkers at room temperature is reliant on the use of (weaker) 3' T-motifs. The extended and compact structures presented in this chapter could be redesigned to use all 5' T-motifs (this would mean locating single-stranded loops on both strands of the 'template-adapter duplex'). It may be the case that only the 3' T-motifs can be broken and remade at room temperature to refine the stability of the structure, and this would likely mean that isothermal folding using a structure with only 5' T-motifs would result in a mis-folded structure (thermal annealing may still produce the target structure). In addition, the use of linkers which have both a 5' and 3' sticky-end means that the 3' sticky-end can only interact with its binding domain after the 5' sticky-end is already bound, and so the situation does not arise where you have a second copy of a linker 'blocking' the second binding site of a linker. It would be interesting to see whether this becomes a problem when only 5' T-motifs are used. (Again, I predict that this would be a problem during isothermal folding, but not folding by thermal annealing.)

We also wish to investigate whether the 3' T-motifs are more stable when part of a larger structure, closer to a conventional origami structure, such as that shown in Figure 4.1. Parts of the structure would be held together by multiple linkers, leading to increased stability, as in conventional origami. Whether or not these structures could still be reconfigured at room temperature would also be investigated.

In the introduction of this chapter I discussed investigating the minimum number of unique binding domains required to specify a particular structure. The AFM images of the compact structure shown in Figure 4.13 do appear to show some structures where linkers are bound 'incorrectly': obtaining more images could allow us to quantitatively analyse this in the compact structure and also larger structures

such as that shown in Figure 4.1.

Chapter 5

Conclusions

In this thesis, I have looked at the folding pathways of both a DNA origami tile (Chapter 2) and an 'origami-like' structure which relies on DNA T-motif interactions (Chapters 3 and 4).

For the conventional origami structure (the tile), folding is governed by hybridisation interactions between the staples and template. Such interactions are relatively strong: for the tile used here, each staple domain has a melting temperature of 38-71°C when hybridised to its complementary template domain[1]. Off-target interactions (staple 'mis-bindings') are weaker on average due to sequence mis-matches and a lack of co-operativity (which correctly-bound staples benefit from), but can have relatively high melting temperatures (up to those of the correct interactions), and thus form readily during isothermal incubations at temperatures below this. When isothermal folding of the tile was attempted at temperatures below the melting temperature of the tile, misfolded structures were produced which were not resolved on a reasonable timescale. This shows that the formation, and persistence (despite the presence of a staple excess which is thought to help remove these interactions), of these off-target interactions is a major issue when folding origami structures isothermally.

In contrast, folding of the 'origami-like' structure presented in Chapter 4 is governed by the formation of (relatively weak) 5' and 3' T-motifs. Unlike the origami tile

[1]Melting temperatures were determined using the NUPACK web application [162].

above, folding by annealing or isothermal incubation at room temperature produced identical results, presumably due to a) the extremely low melting temperatures of any off-target interactions, and b) the ability of the interactions to break and reform at room temperature. This allowed greater redundancy in the binding domains used, whilst still producing well-folded structures. Although the reconfigurability of such structures opens up potential applications where such structures could respond to the strands available, it also means that they are less stable than the conventional origami structures, as was clear from AFM imaging.

As these two examples illustrate, for a DNA nanostructure to be stable at room temperature, the interactions which hold it together must have melting temperatures significantly above room temperature so that they do not readily reverse at room temperature (as for the origami tile used in Chapter 2). However, successful assembly of such structures requires heating to above or near the interaction melting temperature, meaning that it is not possible to fold stable structures at low temperatures. The use of helper strands may provide a solution to this problem, but enabling isothermal folding (in a reasonable timescale) at temperatures below the melting temperature of the tile, apparently by increasing the discrimination between off-target and target staple:template interactions. As well as enabling isothermal folding, use of helper strands may allow folding to be controlled in other ways: e.g. by allowing folding of a specific structure from a template in a mixture with multiple staple sets.

5.1 Future directions

As explained in section 2.6, for the helper-origami system we had expected helpers to be able to be fully displaced by staples in a reasonable timescale at room temperature (once the first staple has bound). The fact that this did not occur suggests that there is some barrier to helper displacement which can only be overcome (in a reasonable time-scale) when the system is heated to $\sim$53°C. Although omission of helpers at or near to the seam lowers this temperature and produces a broader

transition with regards to temperature (Figure 2.41c-e and Figure 2.42b-c), the system still requires heating to >50°C for full folding in one hour. Identification of the barrier(s) preventing rapid folding at lower temperatures may allow these barriers to be eliminated, allowing rapid folding at lower temperatures than have been demonstrated here.

The success of isothermal folding in the absence of helpers also changes dramatically with origami design: as we saw in section 2.4.6, isothermal methods have been more successful for 'conventional' origami designs (i.e. those which use three-domain staples). It would be interesting to test the effects of helpers in three-domain-staple designs to see whether they are able to lower isothermal incubation temperatures further.

For the origami-like structure, work is on-going to address the points raised in section 4.3.1, namely building larger structures which either use both 5' and 3' T-motifs, or all 5' T-motifs.

Appendix A

Materials and Methods

All DNA sequences, including those used for PCR primers as described in this appendix, are detailed in Appendix B.

A.1 Sources of materials and definitions

All DNA strands (with the exception of the origami template strand) were purchased from Integrated DNA Technologies, Inc. (IDT). All strands were unpurified except: those with Cy5, Cy3 or TAMRA labels were HPLC purified, and strands over 60 bases in length were PAGE purified, as per IDT's recommendation. Origami staples and helpers were specified as 150 μM or 200 μM (respectively) suspensions in IDTE 7.5 pH buffer. All other strands were provided in lyophilised form, and resuspended at 100 μM in TE prior to use (or 20 ng/μL for the 537 bp template strand used for T-motif origami; section 4.2). All strands were provided single-stranded, except the T-motif origami template strand which was double-stranded.

Concentrated buffer solutions (50× TAE and 100× TE) were supplied by Pan-Reac AppliChem, SYBR Gold Stain was from Invitrogen, EvaGreen dye was from Biotium, 5× loading buffer was kindly provided by Jon Bath. Other chemicals (including magnesium chloride, LB agar, LB broth, ampicillin, agarose and 29:1 acrylamide) were supplied by Fisher Scientific. Restriction enzymes, T4 DNA ligase, exonucleases, PCR master mixes, pUC19 vector and NEB 5-alpha E. coli cells were

supplied by New England Biolabs, Inc. All thermocyclers were of the Eppendorf Mastercycler Nexus range.

In this thesis, 'TAE' refers to a 1× TAE solution (40 mM Tris, 1 mM EDTA-Na_2, 20 mM acetic acid, pH 8.5±0.2), TAE/Mg to a 1× TAE solution also containing 12.5 mM $MgCl_2$, and TE to a 1× TE solution (10 mM Tris, 1 mM EDTA-Na_2, adjusted to pH 8.0 with HCl). FAM is fluorescein amidate; TAMRA is carboxytetramethylrhodamine; Cy3/5 are cyanine-based fluorescent dyes. 'Milli-Q water' refers to ultrapure water from a Milli-Q Integral 3 Water Purification System using a 0.22 μM membrane filter (Merck).

Where 'annealing'/'thermally-annealing' has been carried out, this refers to cooling a sample from 95°C to 20°C in increments of 0.1°C, at a rate of 1°C/min (see Definitions section of this thesis for a more general definition of thermal annealing). Any LB broth and LB agar used were sterilised through the use of an autoclave, and where ampicillin was added this was to a concentration of 100 μg/ml. All materials were stored as per the manufacturer's suggestion.

A.2 DNA origami design and synthesis

A.2.1 Design

The tile origami described in section 2.1 was designed by Dr Katherine Dunn, a former member of the Turberfield group [33]. Dr Dunn synthesised a custom plasmid (pKD1) from the pUC19 vector [160], from which the origami template strand can be derived (see below). Staple sequences were designed by Katherine using caDNAno [31].

A.2.2 Template synthesis

A.2.2.1 Transformation

NEB 5-alpha competent E. coli (high efficiency) cells were transformed with pKD1 using a method based on the standard protocol from NEB [104]:

1. One 50 μL aliquot of NEB 5-alpha competent E. coli cells was thawed until fully defrosted on ice ($\sim$10 minutes).
2. The aliquot was split into two 25 μL samples on ice (one experimental and one control).
3. 2 μL of 25 ng/μL pKD1 (in TE) was added to the experimental sample, and the tube was carefully flicked to mix (vortexing is not advised).
4. Samples were placed on ice for 30 minutes.
5. Samples were heat-shocked in a 42°C water-bath for 30 seconds.
6. Samples were placed on ice for a further 5 minutes.
7. 100 μL of LB broth was added to each sample and pipetted up and down to mix.
8. 0.1× and 0.01× dilutions of each sample were made.
9. For both of the experimental and control samples, 100 μL of the 1×, 0.1× and 0.01× dilutions were spread on separate LB-agar plates containing ampicillin.
10. Plates were incubated at 37°C for 16 hours. The control samples were used to verify that the antibiotic in the plate was functional (5-alpha cells are not ampicillin resistant, so the control sample plates had no colonies). pKD1 contains an ampicillin-resistance gene which conferred resistance to the experimental sample cells.

A.2.2.2 Amplification and plasmid purification

1. From the experimental sample plates, a plate with a sufficient number of colonies (>10), but without overlapping colonies, was selected. From this plate, four colonies were selected and each transferred to a separate culture tube containing 2 ml LB broth with ampicillin (starter cultures).

2. The cultures were incubated at 37°C for 8 hours in a shaking incubator.

3. The starter cultures were each transferred to a 1 L conical flask containing 250 ml LB broth with ampicillin, and incubated at 37°C for 16 hours in a shaking incubator.

4. Plasmid from each culture was purified using a QIAGEN Maxiprep Plasmid Purification Kit as per their suggested protocol, and finally resuspended in 250 μL TE.

5. Samples were quantified (section A.7) and run on an agarose gel (A.4) to verify the identity, purity and concentration of each.

A.2.2.3 Preparation of single-stranded template strand

pKD1 contains a Nt.BspQI site on one strand (the strand which needs to be removed to leave the single-stranded template). Nicking with this enzyme followed by digestion with a exonuclease therefore yields the desired circular single-stranded template strand.

1. Samples were made up containing 350-500 μg of purified pKD1 in TE (from section A.2.2.2), 100 μL 10× NEBuffer 3.1, 25 μL Nt.BspQI, and Milli-Q water to take the total volume to 1 ml.

2. Samples were incubated in a 50°C water-bath for 16 hours.

3. 2 μL from each tube was removed for gel analysis (step 6), before adding 12 μL ExoIII.

4. Samples were incubated in a 37°C water-bath for 8 hours.

5. Enzymes were heat-killed by incubated in a 80°C water-bath for 20 minutes.

6. Samples were run on an agarose gel (including non-nucleased samples from step 2) to verify complete nicking and nucleasing. Samples were subjected to a further round of nucleasing (repeat of steps 3-6) if nucleasing was seen to be incomplete.

7. Samples were pooled.

8. The pooled sample was purified using the QIAGEN protocol for Re-Purification of Plasmid DNA Prepared by Methods other than QIAGEN Tips [114]. For the centrifugation steps, the alternative method suggested in the protocol (using disposable conical bottom centrifuge tubes) was used. The DNA was redissolved in 50-100 μL of TE during the final step.

9. The purified sample was quantified (section A.7).

A.2.3 Origami preparation

Template strand at 100 nM and staples or helper strands at 200 nM were annealed in TAE/Mg buffer in a thermocycler to produce the folded tile/template annealed with helpers respectively. For samples where both the helpers and staples were present, both were at 200 nM.

A.3 T-motif and T-motif origami design and synthesis

A.3.1 Design

For single T-motif experiments, sequences (not including the loop) were based on those used in Hamada et al. [48], which were truncated and extended as required.

Sequences of all other strands used in T-motif experiments (except loops and sticky-ends involved in T-motif formation) were designed using the NUPACK web application [162] (no sequence constraints were specified for these regions in the target structure input). For the T-motif origami template, restriction sites were also included at either end to allow amplification by cloning (see section A.3.2.2). For the loop and sticky-end sequences, GC-rich sequences (to increase the stability of these interactions) of 5-7 nucleotides in length were used, as these lengths were shown to work experimentally in the original T-motif paper [48], and simulations by Behnam Najafi (see section 3.2) confirmed that these length were optimal for both 3' and 5' T-motifs. Loop sequences in the single T-motif and U-shaped structures were also chosen to avoid 'slipping' - see Figure A.1.

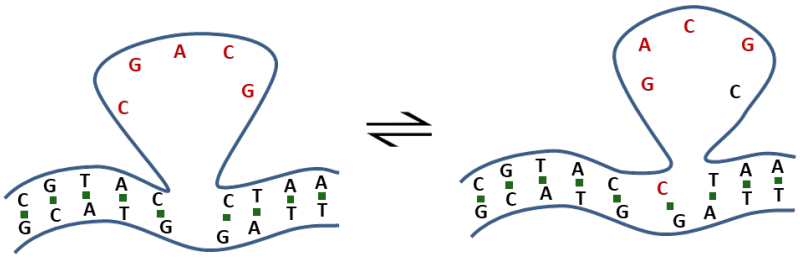


Figure A.1: Example of possible 'slipping' of the loop position. The sequences at the start and end of the loop, and adjacent to the loop on either side, can be altered to prevent this. Slipping leads to an unexpected loop sequence being exposed, which could affect sticky-end binding and therefore T-motif formation.

For both the single T-motif and U-shaped structure experiments, two DNA markers were used in each gel lane: one which had a higher mobility in the gel than the sample of interest, and one which moved more slowly (see section 3.1.1). This later allowed accurate measurements of the position of the sample bands in the gel. Marker sequences were chosen to limit interaction with the sample strands (this was checked using NUPACK), and each marker had a fluorescently-labelled strand.

A.3.2 Sample preparation

A.3.2.1 Marker synthesis (for single T-motif and U-shaped structure gels)

For the single T-motif gels (Chapter 3), the markers were two duplexes of 18 bp and 72 bp, each of which had one FAM-labelled strand. All four strands were directly ordered from IDT. For the U-shaped linker length gels (section 4.1.2), the markers were a small structure containing three short strands (ordered from IDT) and a duplex of 1456 bp (synthesised by amplifying a section of pUC19 with PCR, see below), each of which had one Cy5-labelled strand.

Synthesis (by PCR) of low mobility marker for linker-length gels of U-shaped structure

42 aliquots of 50 μL PCR reaction mix were prepared on ice. Each aliquot contained 25 μL OneTaq 2× Master Mix with G/C Buffer, 1 μL of Cy5-labelled forward primer, 1 μL reverse primer, 1 μL of 0.2 ng/μL pUC19 (see Appendix B.2 for primer sequences) and 22 μL Milli-Q water. The aliquots were placed in a preheated (94°C) thermocycler and then subjected to an initial denaturation step of 30 s at 94°C, and then 35 cycles of 20 s at 94°C, 30 s at 58°C and 40 s at 68°C (all temperatures and times as per NEB's advice, except the 58°C annealing temperature and the 40 s extension time, which were optimised for my primers and template), followed by a final extension of 5 minutes at 68°C. The samples were pooled and purified using the QIAquick PCR Purification Kit (QIAGEN): four columns were used and each was eluted in 40 μL of elution buffer. The eluates were pooled, and the pooled sample quantified (section A.7).

A.3.2.2 T-motif origami template strand preparation

The T-motif origami template was provided in double-stranded form by IDT (as a gBlock gene fragment), but required amplification to produce enough for gel experiments. Initially, this was done using molecular cloning, by cutting the double-

stranded template at either end using restriction enzymes (HindIII and SalI), and ligating the digested duplex into the multiple-cloning site of pUC19 plasmid which had been treated with the same restriction enzymes. This plasmid was sequenced to confirm that the digested fragment had been inserted correctly, followed by transformation, amplification, nicking and nucleasing as described for the pKD1 plasmid in section A.2.2. This method produced a large quantity of single-stranded template as part of a circular single-stranded plasmid. However, this construct was not ideal for observing changes in electrophoretic mobility upon formation of T-motifs within the insert, and so the template was subsequently amplified and made single-stranded using PCR followed by lambda exonuclease digestion as detailed below.

1) PCR

100 aliquots of 50 μL PCR reaction mix were prepared on ice. Each aliquot contained 25 μL Q5 2× Master Mix, 2.5 μL of forward primer, 2.5 μL reverse primer (which had a 5' phosphate group to allow later digestion with lambda exonuclease), 1 μL of 0.2 ng/μL template (the double-stranded template provided by IDT) and 19 μL Milli-Q water. The aliquots were placed in a preheated (98°C) thermocycler and then subjected to an initial denaturation step of 30 s at 98°C, and then 30 cycles of 10 s at 98°C, 20 s at 65°C and 15 s at 72°C (all temperatures and times as per NEB's advice, except the 65°C annealing temperature which was optimised for my primers and template), followed by a final extension of 2 minutes at 72°C. The samples were pooled and purified using the QIAquick PCR Purification Kit: 20 columns were used and each was eluted in 30 μL of elution buffer. The eluates were pooled, and the pooled sample quantified (section A.7).

2) Lambda exonuclease digestion

160 μL lambda exonuclease (5000 U/ml), 100 μL 10× lambda exonuclease buffer and 140 μL Milli-Q water was added to the 600 μL sample from part 1), and then flicked to mix. This was incubated in a water bath at 37°C for 45 minutes, before heat inactivation of the enzyme at 75°C for 15 minutes. This sample was cleaned up using QIAquick PCR Purification Kit, in 10 columns (each eluted in 30 μL). The

pooled eluate was quantified (section A.7) and found to be approximately 0.9 μM. To verify the identity of the product, it was annealed with a fluorescently-labelled adapter strand and subjected to gel electrophoresis.

A.3.2.3 T-motif preparation

For each sample, all strands (including marker strands when these were used) were combined in TAE/Mg buffer, and then annealed in a thermocycler, with the exception of some T-motif origami samples where linkers/adapters were added later and incubated isothermally (see section 4.2). Note that the concentrations in this section refer to the strands during annealing/isothermal incubation. Both gel electrophoresis and AFM analysis require/result in changes in the strand concentration; for gel electrophoresis these arise from both the addition of loading buffer and the process of the strands entering and moving through the gel. In Chapter 3, the concentration values given for the single T-motif experiments reflect these changes as they are an estimate of the strand concentrations within the gel environment (see section 3.1.2). AFM analysis requires dilution as described in section A.6.

For the single T-motif experiments (Chapter 3), the loop duplex was at 0.2 μM whilst the concentration of the sticky-end duplex varied (see individual gels). The two FAM-labelled marker strands (one from each duplex) were at 0.2 μM and the unlabelled marker strands at 0.3 μM.

For testing hierarchical binding within the U-shaped structures (section 4.1.1), all strands were at 0.3 μM except the labelled B strand at 0.2 μM (see Figures 4.2 and 4.3). For the linker length experiments with the U-shaped structure, each sample contained two structures with different E strands (differing in their loop sequence and fluorescent label; Figure 4.5). All strands (A-D and the linker strands) were at 0.6 μM, with each of the two E strands at 0.2 μM. For the braced structures, the extra sticky-end duplexes involved in T-motif formation with the B and C strands were at 2 μM. The linker length samples also contained the three strands of the high mobility marker at 0.015 μM, and 12.5% low mobility marker duplex prepared by

PCR as above.

For the T-motif origami structures, the template strand was at 0.2 μM, adapter/splint strands at 0.3 μM and each type of linker duplex at 5 μM (note that each template-adapter hybrid binds multiple copies of each linker, see section 4.2).

A.4 Gel electrophoresis

Prior to loading onto gels, all samples were diluted with 0.25 volumes of 5× loading buffer (containing Bromophenol Blue and glycerol). 10 μL of each sample was then loaded per gel lane.

A. Agarose

Agarose gel electrophoresis was used to analyse origami samples (Chapter 2) and T-motif origami samples (section 4.2), as well as to verify the intermediate products of origami template synthesis (section A.2.2). All agarose gels were cast using the Bio-Rad Sub-Cell GT Tray (15 × 10 cm) with a 20-well comb, and run using the Sub-Cell GT Cell, with the running buffer the same as that used to make the gel. For the origami samples and template intermediates, all gels were 0.7% agarose in TAE buffer, run at a constant voltage of 60 V for 2 hours at room temperature. For the T-motif origami, adapter strands containing single-stranded loops were unable to remain hybridised to the long strand without the addition of Mg^{2+} in the gel and running buffer and so 0.7% agarose gels in TAE/Mg buffer, run at 60 V for 2 hours at 4°C, were used.

If staining was required prior to scanning the gels, the gel was incubated with SYBR Gold Stain (at approximately 1× concentration) for 30 minutes, before destaining in water for 10 minutes. For agarose gels containing more than one fluorophore type, metal staples were inserted into the four corners of the gel to later allow alignment of the different channels.

B. Polyacrylamide

Polyacrylamide gels were used to analyse all single T-motif (Chapter 3) and U-shaped structure (section 4.1) samples. All polyacrylamide gels were 8% 29:1 (mono:bis) acrylamide gels in TAE/Mg, cast using the Bio-Rad Mini-Protean casting stand, clamps, plates with 0.75 mm spacer, and 15-well comb, and run using the Mini-Protean Tetra Cell at 150 V and 4°C. The running buffer used was the same as that in the gels. The single T-motif gels were run for 90 minutes; and the U-shaped structure gels for 90 minutes (hierarchical binding gels; section 4.1.1) or 2 hours (linker length gels; section 4.1.1) except where otherwise stated.

A.4.1 Analysis

All gels were scanned for fluorescence at (a minimum of) 50 μM resolution using an Amersham Typhoon Imager with the exception of the gels shown in Figures 2.16a and 2.24a, which were imaged at the same resolution using a Pharos FX Plus Molecular Imager.

A) For gels containing more than one fluorophore scanned using the Amersham Typhoon Imager, the images from each channel needed to be re-aligned due to movement of the gel during scanning. This was done using features of the gel plate (polyacrylamide gels) or metal staples inserted into the gel (agarose gels), and the 'Landmark Correspondences' plugin in Fiji [124, 125].

B) For the single T-motif and linker length U-shaped structure gels, distance markers were included in each lane, and these allowed the distances travelled by samples in the gel to be accurately measured (as described in section 3.1.1). Appendix C.1 shows the programs used to analyse these images.

C) Unless stated otherwise in Chapter 2, origami gels were processed to normalise the appearance (brightness and contrast) of each channel, in order to allow visual comparison between gels. This was done by adjusting the minimum and maximum pixel values in each channel to be relative to controls within the gel. The mean pixel value of the gel background (i.e. a part of the gel where no DNA was present)

was used as the minimum pixel value, and the maximum was proportional to a 'brightness' control which was loaded onto each gel. For the FAM and Cy5 channels, the control was a 10 μL sample of annealed tile (100 nM template and 200 nM staples) containing the FAM-labelled body staple and Cy5-labelled seam staple. For the TAMRA/Cy3 channel (TAMRA was used for the 10-offset helper set, Cy3 for the 8-offset set), the control was a 10 μL sample of template annealed with the full set of helpers (100 nM template and 200 nM helpers) including Cy3/TAMRA-labelled body staple. After setting the minimum pixel value to be equal to the background intensity, the sum of the pixel values of each of these brightness control bands was measured. The maximum pixel values were then scaled accordingly: for the FAM channel the maximum pixel value was set to 7 * brightness control, for Cy5 to 12 * brightness control, and for TAMRA/Cy3 to 5 * brightness control. These different scaling factors reflect differences in how different colours are perceived by the human eye.

A.5 Temperature-controlled real-time fluorescence measurements

Real-time monitoring of origami formation and melting was carried out in a Stratagene Mx3005P Real Time PCR machine (Agilent Technologies). 20 μL samples were used, contained in PCR tubes with MicroAmp optical caps (Applied Biosystems). Concentrations are those stated in section A.2.3, with the addition of 1× EvaGreen dye when required. The accompanying MxPro software was used to program different temperature profiles for the different experiments required as detailed below. A 'reading' in each experiment refers to three duplicate readings of temperature and fluorescence (of either the EvaGreen dye or the FRET signal) taken of each sample.

(a) For measuring the formation and melting of the tile and helper-coated template, the samples used were: template + all staples (all unlabelled) with 1× Eva Green (Figure 2.14); template + the full staple set including the FRET

pair (Figure 2.15); and template + a full helper set (either 10-offset or 8-offset; all unlabelled) with 1× EvaGreen (Figure 2.19). Samples were incubated at 86°C for 2 minutes, cooled to 25°C, incubated at 25°C for 2 minutes, and then heated to 86°C. This was repeated 5 times, and the rate of heating/cooling was 0.5°C per 30 seconds. Readings were taken at each 0.5°C (i.e. every 30 seconds) increment during the heating and cooling steps.

(b) For the 40-80°C incubations shown in Figures 2.38 and 2.39 the samples used were helper-coated template (with various helpers omitted) + the full staple set containing the FRET pair. Samples were held at 40°C for 10 minutes to allow thermal equilibration, before heating to 80°C and cooling back down to 40°C in increments of 0.5°C per 5 minutes. Readings were taken at each 0.5°C increment (i.e. every 5 minutes) during the heating and cooling steps.

A.5.1 Analysis

The results were analysed using Matlab. For each sample, the three duplicate temperature and fluorescence readings taken at each time-point were averaged. For the formation/melting experiments (a) above), results from repeated temperature ramps were also averaged. When the derivative (dF/dT) has been plotted, this was determined by fitting a smoothing spline (with a smoothing parameter of 0.6) to the data and differentiating this fit.

A.6 Atomic force microscopy (AFM)

AFM images of origami and T-motif origami samples were acquired using a Cypher ES AFM (Oxford Instruments). Samples were imaged in tapping mode in liquid using the blueDrive Photothermal Excitation option (with the 0.01× filter cube) and AC40 probes provided by Bruker. Origami samples prepared as per section A.2.3 were diluted 1 in 20 (to a final concentration of 5 nM) in filtered imaging buffer: 1× TAE, 12.5 mM magnesium acetate and 4 mM nickel chloride. T-motif

origami samples prepared as per section A.3.2.3 were gel purified (by running on an agarose gel as per section A.4, cutting out gel slices containing the bands of interest, and then squeezing these to remove the liquid) which resulted in a 10× dilution compared to the original sample. For all samples, 5 μL of diluted sample was pipetted onto freshly cleaved mica, left to adhere, and then covered with 60 μL imaging buffer. The tip was also covered with 50 μL of imaging buffer before beginning image acquisition.

It should be noted that sample preparation (including adhesion to a mica surface), and contact between the tip and the sample during imaging, have the potential to introduce structural changes and/or damage to the sample. However, these effects are limited for 2D origami structures such as the tile used in Chapter 2. It is possible for damage to accumulate if a tile is imaged multiple times, and so all AFM images shown in that chapter are from the first or second pass over a sample. No damage which could clearly be ascribed to the tip was observed.

A.6.1 Processing of AFM images

Gwyddion software [103] was used to process the AFM images which appear in Chapters 2 and 4 (to allow structures to be seen more easily) as detailed below:

1. Plane levelling to correct for any tilt in the mica surface.

2. Row alignment (using the 'median' option): this calculates the median of the pixel intensity of a line and its adjacent line and shifts the pixels such that the medians are equal. This corrects for differences between line heights due to the interaction of the cantilever with the sample/mechanical disturbances.

3. Removal of horizontal scars to correct for line defects.

4. Steps 2 and 3 were repeated until no further changes were seen in the image.

5. The colour range was then adjusted so that the images could easily be analysed and compared. Since all T-motif origami images were collected on the same

day, using the same mica surface and AFM tip, this range was set to be the same (-1.3 - 3 nm) for all images. For the origami samples, the range was adjusted so images were similar in appearance.

A.6.2 Quantification of origami AFM images

AFM images of origami structures (Chapter 2) were analysed to extract quantitative data as shown with examples in section 2.3.1.1. Raw images were used as opposed to those processed as per section A.6.1, because the row alignment (line levelling) step resulted in parts of some tiles disappearing into the background. This is illustrated in Figure A.2.

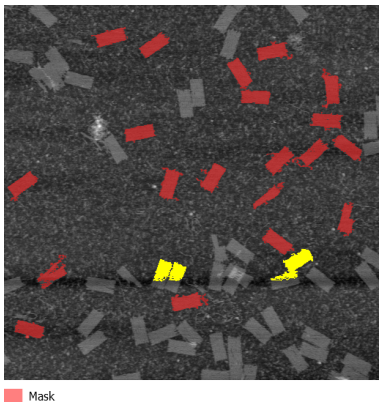


Figure A.2: Line levelling can lead to parts of some tiles (in yellow) disappearing into the background. The remaining partial tiles are sometimes then identified as 'tile-sized objects' during the analysis process, when they touch another partial or whole tile. Thus line levelling was not carried out prior to analysis. Image is identical to example a) in section 2.3.1.1, and shows a sample of helper-coated template incubated with staples at 55°C for 4 hours. Image size is 1 × 1 μm.

The analysis process is described in detail below. Gwyddion software [103] was used for all steps.

1. Horizontal scars were removed from raw images.

2. Grains were identified using the 'Mark by Otsu's' command (see [108]).

3. The marked grains were filtered to remove those not corresponding to single origami tiles. The approximate volume (the 'Laplacian background basis vol-

ume' as determined by Gwyddion) of a well-folded tile was determined from an image of well-folded tiles (produced by annealing of the helper-coated template with staples) to be $\sim 2.4 \times 10^{-24}$ m^3. Grains with a volume of less than 1.4×10^{-24} m^3 or more than 4.4×10^{-24} m^3 were removed. Using such a broad range was possible as the next step involves manual removal of grains not corresponding to single tiles.

4. Grains overlapping with the image edge were removed.

5. Remaining grains not corresponding to single tiles were identified by visual inspection and removed.

6. The properties of the grains (projected area and projected boundary length) were measured and exported.

For each origami sample, four 1 μm by 1μm images were obtained and analysed in this way. The total grain count and mean statistics for the four images were then calculated.

A.7 DNA quantification

When required, the concentration of DNA in a sample was determined by measuring the absorbance of UV light of wavelength 260 nm, using a Cary 50 UV vis spectrophotometer (Agilent) and accompanying Cary WinUV software. Samples were diluted to give an absorbance reading of 0.1-1 (since measurements outside of this range are less accurate), and 4 μL was tested. A baseline measurement was taken using the buffer in which the DNA was suspended, and subtracted from the sample measurement. An absorbance reading of 1 correlates to a DNA concentration of 50 μg/mL for double-stranded DNA and 38 μg/mL for single-stranded DNA [11].

Appendix B

DNA sequences

B.1 DNA Origami (Chapter 2)

Template strand:

GACGAAAGGGCCTCGTGATACGCCTATTTTTATAGGTTAATGTCATGATAATAATGGTTTCTTAGACGTCAGGT
GGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCT
CATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTG
TCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAA
GATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAG
TTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTA
TTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTC
ACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACAC
TGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATC
ATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATG
CCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTAGCTTCCCGGCAACAATT
AATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTG
CTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCC
CGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGG
TGCCTCACTGATTAAGCATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAAACTTC
ATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTT
TCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAAT
CTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTT
TTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCAC
CACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGG
CGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGG
GGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGA
GAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCG
CACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGC
GTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTC
CTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTAC
CGCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGG
AAGAGCGCCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTT
CCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAGGCACCCCAGGCTTT

ACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATTGTGAGCGGATAACAATTTCACACAGGAAACAGCTATG
ACCATGATTACGCCAAGCTCCTCAGCAATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGC
GTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGA
TCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGCGCCTGATGCGGTATTTTCTCCTTACGCATCTGT
GCGGTATTTCACACCGCATATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGCCCCGA
CACCCGCCAACACCCGCTGACGCGCCCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGAC
CGTCTCCGGGAGCTGCATGTGTCAGAGGTTTTCACCGTCATCACCGAAACGCGCGA

Staple and helper strands are named according to their binding positions on the tile (see Figure 2.2) in the following format (as per the output from caDNAno): a[b]-c[d]. a[b] is the position where the 5' end of the strand binds (a is the helix of the tile, counting from top to bottom (1-12); b is the nucleotide position, going from left to right across the tile (0-223). c[d] is where the 3' end of the strand binds (again c = helix, d = nucleotide). The nucleotide positions are not accurate as they assume all domains are 16 nt, whereas some are 15 (see 2.1). NB: the register of the template strand (i.e. the position in the sequence above which was taken to be 1[0]) is different for each of the two helper sets compared to the staple set. Figure 2.5 shows the relative binding positions of the staple set and two helper sets: for each of helper sets, helper 1[0]-1[16] is first helper (moving from left to right) which solely binds to the uppermost helix.

B.1.1 Staple sequences

Seam staples are in blue (except the labelled seam) and edge staples in red (see section 2.1). Labelled staples used for gels (see section 2.3.2.1) and FRET measurements (see Figure 2.15) are in green. Sequences are of the labelled version (i.e. include the fluorescent label); unlabelled versions were also used in different experiments.

Staple name	Sequence
1[0]-2[0]	TTTTTGAATACTCATACTCTGGCGACACGGAAATGTTT-TT
1[16]-1[31]	TCCTTTTTCAATATT
1[32]-2[32]	ATTGAAGCATTTATCAGAAGGCAAAATGCCG
1[48]-1[63]	GGGTTATTGTCTCATG
1[64]-2[64]	AGCGGATACATATTTGTTTCACCAGCGTTTCT
1[80]-1[95]	AATGTATTTAGAAAA
1[96]-1[127]	ATAAACAAATAGGGGTTCCGCGCACATTTCCC
1[128]-1[143]	CGAAAAGTGCCACCTG
1[160]-1[175]	TTATTATCATGACATT
1[192]-1[207]	CGTATCACGAGGCCCT
2[127]-2[96]	GGCGCGTCAGCGGGTGGTGCACCCAACTGAT
2[143]-3[143]	CAGACAAGCCCGTCAGGCTGGCTTAACTATGC
2[159]-1[159]	AAGCGGATGCCGGGAGACGTCTAAGAAACCA
2[175]-3[175]	TCACAGCTTGTCTGTGTACTGAGAGTGCACC
2[191]-1[191]	CAGCTCCCGGAGACGGAACCTATAAAAATAGG
2[207]-3[207]	AAACCTCTGACACATGTACCGCACAGATGCGT
2[223]-1[223]	TTTTTCGGTGATGACGGTGATTCGTCTCGCGCGTTTT-TT
3[0]-4[0]	TTTTCCGCGCCACATAGCAGCAATACGGGATAATATTTT

Staple name	Sequence
3[16]-2[16]	AACTTTAAAAGTGCTCCAAAAAAGGGAATAAG
3[32]-4[32]	ATCATTGGAAAACGTTTATGCGGCGACCGAGT
3[48]-2[48]	CTTCGGGGCGAAAACGGGTGAGCAAAAACAG
3[64]-4[64]	TCTCAAGGATCTTACCAGTACTCAACCAAGT
3[80]-2[80]	GCTGTTGAGATCCAGTCTTCAGCATCTTTTAC
3[96]-3[127]	TCGATGTAACCCACTCTTGGCGGGTGTCGGG
4[127]-4[96]	GTGCTGCAAGGCGATTGCCATCCGTAAGATGC
4[143]-5[143]	TGGCGAAAGGGGGATGGGTTTTCCCAGTCAC
4[159]-3[159]	TCGCTATTACGCCAGCGGCATCAGAGCAGATT
4[175]-5[175] Labelled body	/56-FAM/ATCGGTGCGGGCCTCTCGGCCAGTGAATTGCT
4[191]-3[191]	ACTGTTGGGAAGGGCGATATGCGGTGTGAAA
4[207]-5[207]	CATTCAGGCTGCGCATCATGGTCATAGCTGT
4[223]-3[223]	TTTTATCAGGCGCCATTCGCAAGGAGAAAATACCGCTT-TT
5[0]-6[0]	TTTTTGCAAAAAAGCGGTTAATGATCCCCCATGTTGTT-TT
5[16]-4[16]	GCTCCTTCGGTCCTCTGCTCTTGCCCGGCGT
5[32]-6[32]	CGATCGTTGTCAGAAGGCTCCGGTTCCCAAC
5[48]-4[48]	TAAGTTGGCCGCAGTGCATTCTGAGAATAGTG
5[64]-6[64]	TTATCACTCATGGTTAGTCACGCTCGTCGTTT
5[80]-4[80]	TGGCAGCACTGCATATTTTCTGTGACTGGTG
5[96]-5[127]	ATTCTCTTACTGTCATAAGTTGGGTAACGCCA
6[127]-6[96] Labelled seam	/Cy5/TTGCGCTCACTGCCCGACGTTGTTGCCATTG
6[143]-7[143]	ACTCACATTAATTGCGACCTGTCGTGCCAGCT
6[159]-5[159]	CCTAATGAGTGAGCTAGACGTTGTAAAACGA
6[175]-7[175]	GTAAAGCCTGGGGTGCAACGCGCGGGGAGAG
6[191]-5[191]	GCCGGAAGCATAAAGTGAGGAGCTTGGCGTAA
6[207]-7[207]	TCCACACAACATACGAGGCGCTCTTCCGCTTC
6[223]-5[223]	TTTTGTTATCCGCTCACAATTTCCTGTGTGAAATTTTTT
7[0]-8[0]	TTTTCCGGAAGGGCCGAGCGCAATAAACCAGCCAGTT-TT
7[16]-6[16]	CAGAAGTGGTCCTGCAGATCAAGGCGAGTTAC
7[32]-8[32]	ACTTTATCCGCCTCCAAGACCCACGCTCACCG
7[48]-6[48]	TCCAGTCTATTAATTGGTATGGCTTCATTCA
7[64]-8[64]	GTTGCCGGGAAGCTAGCATCTGGCCCCAGTG
7[80]-6[80]	AGTAAGTAGTTCGCCACTACAGGCATCGTGGT
7[96]-7[127]	GTTAATAGTTTGCGCACTTTCCAGTCGGGAA
8[127]-8[96]	ACATGTGAGCAAAAGGCGTGTAGATAACTACG
8[143]-9[143]	ATAACGCAGGAAAGAAACCGTAAAAAGGCCG
8[159]-7[159]	TCCACAGAATCAGGGGGCATTAATGAATCGGC
8[175]-9[175]	GGCGGTAATACGGTTATCCATAGGCTCCGCCC

Staple name	Sequence
8[191]-7[191]	TATCAGCTCACTCAAAGCGGTTTGCGTATTG
8[207]-9[207]	GGCTGCGGCGAGCGGAAAAATCGACGCTCAA
8[223]-7[223]	TTTTCTGCGCTCGGTCGTTCCTCGCTCACTGACTCGTT-TT
9[0]-10[0]	TTTTTAAAGTATATATGAGTAAGTTTTAAATCAATCTT-TT
9[16]-8[16]	AAACTTGGTCTGACAGCTCCAGATTTATCAG
9[32]-10[32]	GTTACCAATGCTTAATTCTTCACCTAGATCC
9[48]-8[48] FRET acceptor	/Cy5/CAGTGAGGCACCTATCCTGCAATGATACCGCG
9[64]-10[64]	TCAGCGATCTGTCTATAGGGATTTTGGTCATG
9[80]-8[80]	TTCGTTCATCCATAGATACGGGAGGGCTTAC
9[96]-9[127]	TTGCCTGACTCCCCGTCCAGCAAAAGGCCAGG
10[127]-10[96]	GCCTTTCTCCCTTCGGCTGACGCTCAGTGGA
10[143]-11[143]	TACCGGATACCTGTCCCTCATAGCTCACGCTG
10[159]-9[159]	TTCCGACCCTGCCGCTCGTTGCTGGCGTTTT
10[175]-11[175]	TCGTGCGCTCTCCTGGTGTAGGTCGTTCGCT
10[191]-9[191]	CCCCCTGGAAGCTCCCCCCTGACGAGCATCAC
10[207]-11[207]	AAGATACCAGGCGTTTTGCACGAACCCCCCGT
10[223]-9[223]	TTTTACCCGACAGGACTATAGTCAGAGGTGGCGAATTTT
11[0]-12[0]	TTTTGATCCGGCAAACAAACGAGTTGGTAGCTCTTTT-TT
11[16]-10[16]	CACCGCTGGTAGCGGTTTTTAAATTAAAAATG
11[32]-12[32]	GGTTTTTTTGTTTGCACGCTCTGCTGAAGCCA
11[48]-10[48] FRET donor	/Cy3/AGCAGCAGATTACGCAGATTATCAAAAAGGA
11[64]-12[64]	GCAGAAAAAAAGGATCGCTACACTAGAAGAA
11[80]-10[80]	TCAAGAAGATCCTTTGACGAAAACTCACGTTA
11[96]-11[127]	ATCTTTTCTACGGGGTGAAGCGTGGCGCTTT
12[31]-12[16]	GTTACCTTCGGAAAAA
12[63]-12[48]	CAGTATTTGGTATCTG
12[95]-12[80]	TGGTGGCCTAACTACG
12[127]-12[96]	GGTATGTAGGCGGTGCTACAGAGTTCTTGAAG
12[143]-12[128]	GGATTAGCAGAGCGA
12[159]-11[159]	GCAGCCACTGGTAACATAGGTATCTCAGTTCG
12[175]-12[160]	CTTATCGCCACTGGCA
12[191]-11[191]	ACCCGGTAAGACACGACCAAGCTGGGCTGTG
12[207]-12[192]	ATCGTCTTGAGTCCA
12[223]-11[223]	TTTTGCCTTATCCGGTAACTTCAGCCCGACCGCTGCTT-TT

B.1.2 10-offset helper set

As above, labelled strand used for gels (section 2.3.2.1) is in green. Sequence is of the labelled version. Helpers located at the seam/seam adjacent (see section 2.5) are labelled in the third column: A = at the seam, B = 5' seam-adjacent, C = 3' seam-adjacent.

Strand name	Sequence
1[0]-1[15]	CTCATACTCTTCCTTT
2[15]-2[0]	ACGGAAATGTTGAATA
3[0]-3[15]	CACATAGCAGAACTTT
4[15]-4[0]	GGGATAATACCGCGC
5[0]-5[15]	AAAGCGGTTAGCTCCT
6[15]-6[0]	CCCCATGTTGTGCAAA
7[0]-7[15]	GGGCCGAGCGCAGAAG
8[15]-8[0]	ACCAGCCAGCCGGAA
9[0]-9[15]	ATATATGAGTAAACTT
10[15]-10[0]	TAAATCAATCTAAAGT
11[0]-11[15]	GCAAACAAACCACCGC
12[15]-12[0]	GTAGCTCTTGATCCG
1[16]-1[31]	TTCAATATTATTGAA
2[31]-2[16]	AGGGAATAAGGGCGAC
3[16]-3[31]	AAAAGTGCTCATCATT
4[31]-4[16]	TGCCCGGCGTCAATAC
5[16]-5[31]	TCGGTCCTCCGATCG
6[31]-6[16]	GGCGAGTTACATGATC
7[16]-7[31]	TGGTCCTGCAACTTTA
8[31]-8[16]	GATTTATCAGCAATAA
9[16]-9[31]	GGTCTGACAGTTACC
10[31]-10[16]	ATTAAAAATGAAGTTT
11[16]-11[31]	TGGTAGCGGTGGTTTT
12[31]-12[16]	TTCGGAAAAAGAGTTG
1[32]-1[47]	GCATTTATCAGGGTTA
2[47]-2[32]	AAAATGCCGCAAAAA
3[32]-3[47]	GGAAAACGTTCTTCGG
4[47]-4[32]	GCGACCGAGTTGCTCT
5[32]-5[47]	TTGTCAGAAGTAAGTT
6[47]-6[32]	GTTCCCAACGATCAA
7[32]-7[47]	TCCGCCTCCATCCAGT
8[47]-8[32]	ACGCTCACCGGCTCCA
9[32]-9[47]	AATGCTTAATCAGTGA
10[47]-10[32]	CCTAGATCCTTTTAA
11[32]-11[47]	TTTGTTTGCAAGCAGC
12[47]-12[32]	GCTGAAGCCAGTTACC

Strand name	Sequence	
1[48]-1[63]	TTGTCTCATGAGCGGA	
2[63]-2[48]	GCAAAAACAGGAAGGC	
3[48]-3[63]	GGCGAAAACTCTCAA	
4[63]-4[48]	GAGAATAGTGTATGCG	
5[48]-5[63]	GGCCGCAGTGTTATCA	
6[63]-6[48]	GCTTCATTCAGCTCCG	
7[48]-7[63]	CTATTAATTGTTGCC	
8[63]-8[48]	TGATACCGCGAGACCC	
9[48]-9[63]	GGCACCTATCTCAGCG	
10[63]-10[48]	TCAAAAAGGATCTTCA	
11[48]-11[63]	AGATTACGCGCAGAA	
12[63]-12[48]	TTGGTATCTGCGCTCT	
1[64]-1[79]	TACATATTTGAATGTA	
2[79]-2[64]	CAGCGTTTCTGGGTGA	
3[64]-3[79]	GGATCTTACCGCTGTT	
4[79]-4[64]	CAACCAAGTCATTCT	
5[64]-5[79]	CTCATGGTTATGGCAG	
6[79]-6[64]	CTCGTCGTTTGGTATG	
7[64]-7[79]	GGGAAGCTAGAGTAAG	
8[79]-8[64]	GCCCCAGTGCTGCAA	
9[64]-9[79]	ATCTGTCTATTTCGTT	
10[79]-10[64]	TTTGGTCATGAGATTA	
11[64]-11[79]	AAAAAGGATCTCAAGA	
12[79]-12[64]	CTAGAAGAACAGTAT	
1[80]-1[95]	TTTAGAAAAATAAAC	
2[95]-2[80]	CATCTTTTACTTTCAC	
3[80]-3[95]	GAGATCCAGTTCGATG	C
4[95]-4[80]	GTGACTGGTGAGTACT	
5[80]-5[95]	CACTGCATAATTCTC	C
6[95]-6[80]	GCATCGTGGTGTCACG	
7[80]-7[95]	TAGTTCGCCAGTTAAT	C
8[95]-8[80]	GAGGGCTTACCATCTG	
9[80]-9[95]	CATCCATAGTTGCCT	C
10[95]-10[80]	ACTCACGTTAAGGGAT	
11[80]-11[95]	AGATCCTTTGATCTTT	C
12[95]-12[80]	CCTAACTACGGCTACA	
1[96]-1[111]	AAATAGGGGTTCCGCG	
2[111]-2[96]	CCAACTGATCTTCAG	B
3[96]-3[111]	TAACCCACTCGTGCAC	A
4[111]-4[96]	CGTAAGATGCTTTTCT	B
5[96]-5[111]	TTACTGTCATGCCATC	A
6[111]-6[96]	TTGCCATTGCTACAG	B

Strand name	Sequence	
7[96]-7[111]	AGTTTGCGCAACGTTG	A
8[111]-8[96]	GATAACTACGATACGG	B
9[96]-9[111]	GACTCCCCGTCGTGTA	A
10[111]-10[96]	CTCAGTGGAACGAAA	B
11[96]-11[111]	TCTACGGGGTCTGACG	A
12[111]-12[96]	GTTCTTGAAGTGGTGG	
1[112]-1[127]	CACATTTCCCCGAAAA	
2[127]-2[112]	TCAGCGGGTGTTGGCG	A
3[112]-3[127]	GGTGTCGGGGCTGGC	B
4[127]-4[112]	CAAGGCGATTAAGTTG	A
5[112]-5[127]	GGTAACGCCAGGGTTT	B
6[127]-6[112]	TCACTGCCCGCTTTCC	A
7[112]-7[127]	AGTCGGGAAACCTGT	B
8[127]-8[112]	GAGCAAAAGGCCAGCA	A
9[112]-9[127]	AAAGGCCAGGAACCGT	B
10[127]-10[112]	CTCCCTTCGGGAAGCG	A
11[112]-11[127]	TGGCGCTTTCTCATA	B
12[127]-12[112]	TAGGCGGTGCTACAGA	
1[128]-1[143]	GTGCCACCTGACGTCT	
2[143]-2[128]	AGCCCGTCAGGGCGCG	C
3[128]-3[143]	TTAACTATGCGGCATC	
4[143]-4[128]	AAGGGGGATGTGCTG	C
5[128]-5[143]	TCCCAGTCACGACGTT	
6[143]-6[128]	ATTAATTGCGTTGCGC	C
7[128]-7[143]	CGTGCCAGCTGCATTA	
8[143]-8[128]	CAGGAAAGAACATGT	C
9[128]-9[143]	AAAAAGGCCGCGTTGC	
10[143]-10[128]	ATACCTGTCCGCCTTT	C
11[128]-11[143]	GCTCACGCTGTAGGTA	
12[143]-12[128]	GCAGAGCGAGGTATG	
1[144]-1[159]	AAGAAACCATTATTA	
2[159]-2[144]	ATGCCGGGAGCAGACA	
3[144]-3[159]	AGAGCAGATTGTACTG	
4[159]-4[144]	TTACGCCAGCTGGCGA	
5[144]-5[159]	GTAAAACGACGGCCA	
6[159]-6[144]	GAGTGAGCTAACTCAC	
7[144]-7[159]	ATGAATCGGCCAACGC	
8[159]-8[144]	GAATCAGGGGATAACG	
9[144]-9[159]	TGGCGTTTTTCCATA	
10[159]-10[144]	CCCTGCCGCTTACCGG	
11[144]-11[159]	TCTCAGTTCGGTGTAG	
12[159]-12[144]	ACTGGTAACAGGATTA	

Strand name	Sequence
1[160]-1[175]	TCATGACATTAACCTA
2[175]-2[160]	CTTGTCTGTAAGCGG
3[160]-3[175]	AGAGTGCACCATATGC
4[175]-4[160]	/56-
Labelled helper	TAMN/GCGGGCCTCTTCGCTA
5[160]-5[175]	GTGAATTGCTGAGGAG
6[175]-6[160]	CCTGGGGTGCCTAAT
7[160]-7[175]	GCGGGGAGAGGCGGTT
8[175]-8[160]	AATACGGTTATCCACA
9[160]-9[175]	GGCTCCGCCCCCCTGA
10[175]-10[160]	GCTCTCCTGTTCCGA
11[160]-11[175]	GTCGTTCGCTCCAAGC
12[175]-12[160]	GCCACTGGCAGCAGCC
1[176]-1[191]	TAAAAATAGGCGTATC
2[191]-2[176]	CCGGAGACGGTCACAG
3[176]-3[191]	GGTGTGAAATACCGC
4[191]-4[176]	GGGAAGGGCGATCGGT
5[176]-5[191]	CTTGGCGTAATCATGG
6[191]-6[176]	AGCATAAAGTGTAAAG
7[176]-7[191]	TGCGTATTGGGCGCT
8[191]-8[176]	CTCACTCAAAGGCGGT
9[176]-9[191]	CGAGCATCACAAAAAT
10[191]-10[176]	GGAAGCTCCCTCGTGC
11[176]-11[191]	TGGGCTGTGTGCACG
12[191]-12[176]	TAAGACACGACTTATC
1[192]-1[207]	ACGAGGCCCTTTCGTC
2[207]-2[192]	CTGACACATGCAGCTC
3[192]-3[207]	ACAGATGCGTAAGGAG
4[207]-4[192]	GGCTGCGCAACTGTT
5[192]-5[207]	TCATAGCTGTTTCCTG
6[207]-6[192]	CAACATACGAGCCGGA
7[192]-7[207]	CTTCCGCTTCCTCGCT
8[207]-8[192]	GGCGAGCGGTATCAG
9[192]-9[207]	CGACGCTCAAGTCAGA
10[207]-10[192]	CCAGGCGTTTCCCCCT
11[192]-11[207]	AACCCCCCGTTCAGCC
12[207]-12[192]	TTGAGTCCAACCCGG
1[208]-1[223]	TCGCGCGTTTCGGTG
2[223]-2[208]	ATGACGGTGAAAACCT
3[208]-3[223]	AAAATACCGCATCAGG
4[223]-4[208]	CGCCATTCGCCATTCA

Strand name	Sequence
5[208]-5[223]	TGTGAAATTGTTATC
6[223]-6[208]	CGCTCACAATTCCACA
7[208]-7[223]	CACTGACTCGCTGCGC
8[223]-8[208]	TCGGTCGTTCGGCTGC
9[208]-9[223]	GGTGGCGAAACCCGA
10[223]-10[208]	CAGGACTATAAAGATA
11[208]-11[223]	CGACCGCTGCGCCTTA
12[223]-12[208]	TCCGGTAACTATCGTC

B.1.3 8-offset helper set

As above, labelled strand used for gels (section 2.3.2.1) is in green. Sequence is of the labelled version. Helpers located at the seam/seam adjacent (see section 2.5) are labelled in the third column: A = at the seam, B = 5' seam-adjacent, C = 3' seam-adjacent.

Strand name	Sequence
1[0]-1[15]	CATACTCTTCCTTTTT
2[15]-2[0]	GGAAATGTTGAATACT
3[0]-3[15]	CATAGCAGAACTTTAA
4[15]-4[0]	GATAATACCGCGCCA
5[0]-5[15]	AGCGGTTAGCTCCTTC
6[15]-6[0]	CCATGTTGTGCAAAAA
7[0]-7[15]	GCCGAGCGCAGAAGTG
8[15]-8[0]	CAGCCAGCCGGAAGG
9[0]-9[15]	ATATGAGTAAACTTGG
10[15]-10[0]	AATCAATCTAAAGTAT
11[0]-11[15]	AAACAAACCACCGCTG
12[15]-12[0]	AGCTCTTGATCCGGC
1[16]-1[31]	CAATATTATTGAAGC
2[31]-2[16]	GGAATAAGGGCGACAC
3[16]-3[31]	AAGTGCTCATCATTGG
4[31]-4[16]	CCCGGCGTCAATACGG
5[16]-5[31]	GGTCCTCCGATCGTT
6[31]-6[16]	CGAGTTACATGATCCC
7[16]-7[31]	GTCCTGCAACTTTATC
8[31]-8[16]	TTTATCAGCAATAAAC
9[16]-9[31]	TCTGACAGTTACCAA
10[31]-10[16]	TAAAAATGAAGTTTTA
11[16]-11[31]	GTAGCGGTGGTTTTTT
12[31]-12[16]	CGGAAAAAGAGTTGGT
1[32]-1[47]	ATTTATCAGGGTTATT

Strand name	Sequence	
2[47]-2[32]	AATGCCGCAAAAAAG	
3[32]-3[47]	AAAACGTTCTTCGGGG	
4[47]-4[32]	GACCGAGTTGCTCTTG	
5[32]-5[47]	GTCAGAAGTAAGTTGG	
6[47]-6[32]	TCCCAACGATCAAGG	
7[32]-7[47]	CGCCTCCATCCAGTCT	
8[47]-8[32]	GCTCACCGGCTCCAGA	
9[32]-9[47]	TGCTTAATCAGTGAGG	
10[47]-10[32]	TAGATCCTTTTAAAT	
11[32]-11[47]	TGTTTGCAAGCAGCAG	
12[47]-12[32]	TGAAGCCAGTTACCTT	
1[48]-1[63]	GTCTCATGAGCGGATA	
2[63]-2[48]	AAAAACAGGAAGGCAA	
3[48]-3[63]	CGAAAACTCTCAAGG	
4[63]-4[48]	GAATAGTGTATGCGGC	
5[48]-5[63]	CCGCAGTGTTATCACT	
6[63]-6[48]	TTCATTCAGCTCCGGT	
7[48]-7[63]	ATTAATTGTTGCCGG	
8[63]-8[48]	ATACCGCGAGACCCAC	
9[48]-9[63]	CACCTATCTCAGCGAT	
10[63]-10[48]	AAAAAGGATCTTCACC	
11[48]-11[63]	ATTACGCGCAGAAAA	
12[63]-12[48]	GGTATCTGCGCTCTGC	
1[64]-1[79]	CATATTTGAATGTATT	
2[79]-2[64]	GCGTTTCTGGGTGAGC	
3[64]-3[79]	ATCTTACCGCTGTTGA	
4[79]-4[64]	ACCAAGTCATTCTGA	
5[64]-5[79]	CATGGTTATGGCAGCA	
6[79]-6[64]	CGTCGTTTGGTATGGC	
7[64]-7[79]	GAAGCTAGAGTAAGTA	
8[79]-8[64]	CCCAGTGCTGCAATG	
9[64]-9[79]	CTGTCTATTTCGTTCA	
10[79]-10[64]	TGGTCATGAGATTATC	
11[64]-11[79]	AAAGGATCTCAAGAAG	
12[79]-12[64]	AGAAGAACAGTATTT	
1[80]-1[95]	TAGAAAAATAAACAA	
2[95]-2[80]	TCTTTTACTTTCACCA	
3[80]-3[95]	GATCCAGTTCGATGTA	C
4[95]-4[80]	GACTGGTGAGTACTCA	
5[80]-5[95]	CTGCATAATTCTCTT	C
6[95]-6[80]	ATCGTGGTGTCACGCT	
7[80]-7[95]	GTTCGCCAGTTAATAG	C

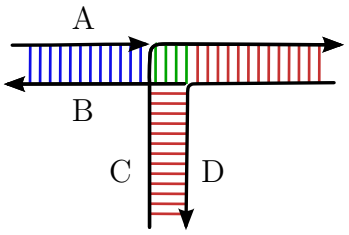
Strand name	Sequence	
8[95]-8[80]	GGGCTTACCATCTGGC	
9[80]-9[95]	TCCATAGTTGCCTGA	C
10[95]-10[80]	TCACGTTAAGGGATTT	
11[80]-11[95]	ATCCTTTGATCTTTTC	C
12[95]-12[80]	TAACTACGGCTACACT	
1[96]-1[111]	ATAGGGGTTCCGCGCA	
2[111]-2[96]	AACTGATCTTCAGCA	B
3[96]-3[111]	ACCCACTCGTGCACCC	A
4[111]-4[96]	TAAGATGCTTTTCTGT	B
5[96]-5[111]	ACTGTCATGCCATCCG	A
6[111]-6[96]	GCCATTGCTACAGGC	B
7[96]-7[111]	TTTGCGCAACGTTGTT	A
8[111]-8[96]	TAACTACGATACGGGA	B
9[96]-9[111]	CTCCCCGTCGTGTAGA	A
10[111]-10[96]	CAGTGGAACGAAAAC	B
11[96]-11[111]	TACGGGGTCTGACGCT	A
12[111]-12[96]	TCTTGAAGTGGTGGCC	
1[112]-1[127]	CATTTCCCCGAAAAGT	
2[127]-2[112]	AGCGGGTGTTGGCGGG	A
3[112]-3[127]	TGTCGGGGCTGGCTT	B
4[127]-4[112]	AGGCGATTAAGTTGGG	A
5[112]-5[127]	TAACGCCAGGGTTTTC	B
6[127]-6[112]	ACTGCCCGCTTTCCAG	A
7[112]-7[127]	TCGGGAAACCTGTCG	B
8[127]-8[112]	GCAAAAGGCCAGCAAA	A
9[112]-9[127]	AGGCCAGGAACCGTAA	B
10[127]-10[112]	CCCTTCGGGAAGCGTG	A
11[112]-11[127]	GCGCTTTCTCATAGC	B
12[127]-12[112]	GGCGGTGCTACAGAGT	
1[128]-1[143]	GCCACCTGACGTCTAA	
2[143]-2[128]	CCCGTCAGGGCGCGTC	C
3[128]-3[143]	AACTATGCGGCATCAG	
4[143]-4[128]	GGGGGATGTGCTGCA	C
5[128]-5[143]	CCAGTCACGACGTTGT	
6[143]-6[128]	TAATTGCGTTGCGCTC	C
7[128]-7[143]	TGCCAGCTGCATTAAT	
8[143]-8[128]	GGAAAGAACATGTGA	C
9[128]-9[143]	AAAGGCCGCGTTGCTG	
10[143]-10[128]	ACCTGTCCGCCTTTCT	C
11[128]-11[143]	TCACGCTGTAGGTATC	
12[143]-12[128]	AGAGCGAGGTATGTA	

Strand name	Sequence
1[144]-1[159]	GAAACCATTATTATC
2[159]-2[144]	GCCGGGAGCAGACAAG
3[144]-3[159]	AGCAGATTGTACTGAG
4[159]-4[144]	ACGCCAGCTGGCGAAA
5[144]-5[159]	AAAACGACGGCCAGT
6[159]-6[144]	GTGAGCTAACTCACAT
7[144]-7[159]	GAATCGGCCAACGCGC
8[159]-8[144]	ATCAGGGGATAACGCA
9[144]-9[159]	GCGTTTTTCCATAGG
10[159]-10[144]	CTGCCGCTTACCGGAT
11[144]-11[159]	TCAGTTCGGTGTAGGT
12[159]-12[144]	TGGTAACAGGATTAGC
1[160]-1[175]	ATGACATTAACCTATA
2[175]-2[160]	TGTCTGTAAGCGGAT
3[160]-3[175]	AGTGCACCATATGCGG
4[175]-4[160]	/56-
Labelled helper	TAMN/GGGCCTCTTCGCTATT
5[160]-5[175]	GAATTGCTGAGGAGCT
6[175]-6[160]	TGGGGTGCCTAATGA
7[160]-7[175]	GGGGAGAGGCGGTTTG
8[175]-8[160]	TACGGTTATCCACAGA
9[160]-9[175]	CTCCGCCCCCCTGACG
10[175]-10[160]	TCTCCTGTTCCGACC
11[160]-11[175]	CGTTCGCTCCAAGCTG
12[175]-12[160]	CACTGGCAGCAGCCAC
1[176]-1[191]	AAAATAGGCGTATCAC
2[191]-2[176]	GGAGACGGTCACAGCT
3[176]-3[191]	TGTGAAATACCGCAC
4[191]-4[176]	GAAGGGCGATCGGTGC
5[176]-5[191]	TGGCGTAATCATGGTC
6[191]-6[176]	CATAAAGTGTAAAGCC
7[176]-7[191]	CGTATTGGGCGCTCT
8[191]-8[176]	CACTCAAAGGCGGTAA
9[176]-9[191]	AGCATCACAAAAATCG
10[191]-10[176]	AAGCTCCCTCGTGCGC
11[176]-11[191]	GGCTGTGTGCACGAA
12[191]-12[176]	AGACACGACTTATCGC
1[192]-1[207]	GAGGCCCTTTCGTCTC
2[207]-2[192]	GACACATGCAGCTCCC
3[192]-3[207]	AGATGCGTAAGGAGAA
4[207]-4[192]	CTGCGCAACTGTTGG

Strand name	Sequence
5[192]-5[207]	ATAGCTGTTTCCTGTG
6[207]-6[192]	ACATACGAGCCGGAAG
7[192]-7[207]	TCCGCTTCCTCGCTCA
8[207]-8[192]	CGAGCGGTATCAGCT
9[192]-9[207]	ACGCTCAAGTCAGAGG
10[207]-10[192]	AGGCGTTTCCCCCTGG
11[192]-11[207]	CCCCCCGTTCAGCCCG
12[207]-12[192]	GAGTCCAACCCGGTA
1[208]-1[223]	GCGCGTTTCGGTGAT
2[223]-2[208]	GACGGTGAAAACCTCT
3[208]-3[223]	AATACCGCATCAGGCG
4[223]-4[208]	CCATTCGCCATTCAGG
5[208]-5[223]	TGAAATTGTTATCCG
6[223]-6[208]	CTCACAATTCCACACA
7[208]-7[223]	CTGACTCGCTGCGCTC
8[223]-8[208]	GGTCGTTCGGCTGCGG
9[208]-9[223]	TGGCGAAACCCGACA
10[223]-10[208]	GGACTATAAAGATACC
11[208]-11[223]	ACCGCTGCGCCTTATC
12[223]-12[208]	CGGTAACTATCGTCTT

B.2 Single T-motifs (Chapter 3)

The four strands comprising each T-motif have been labelled as follows:



A = short strand of sticky-end duplex
B = long strand of sticky-end duplex (containing sticky-end)
C = long strand of loop duplex (containing loop)
D = short strand of loop duplex

The table below shows the strands used for the main results (section 3.2; both gels and simulations). These are shown in Figure 3.7 and gels of the duplexes in Figure 3.8. Numbers after B refer to the length and polarity of the sticky-end on strand B e.g. B6 (3') means that B has a 6 nt 3' sticky-end when hybridised with strand A. To change the loop size, the length of strand D was altered to leave more of C single-stranded. The number after D refers to the resulting loop length on strand C.

Strand name	**Sequence**
A	CGTCGTTCTAGTATAGTCTCTTATT
B5 (3')	AATAAGAGACTATACTAGAACGACGTGTAG
B6 (3')	AATAAGAGACTATACTAGAACGACGCTGTAG
B7 (3')	AATAAGAGACTATACTAGAACGACGCCTGTAG
B5 (5')	TGTAGAATAAGAGACTATACTAGAACGACG
B6 (5')	CTGTAGAATAAGAGACTATACTAGAACGACG
B7 (5')	CCTGTAGAATAAGAGACTATACTAGAACGACG
C (labelled)	/56-FAM/TGGGCAGCTCTACAGGAAATCTATAGTTC
D5	GAACTATAGATTTCCAGCTGCCCA
D6	GAACTATAGATTTCAGCTGCCCA
D7	GAACTATAGATTTAGCTGCCCA

A/B5 (3')/C/D5 were used for the gels shown in Figures 3.3, 3.4a and c, and 3.5. A/B7 (3')/C/D7 were used for Figure 3.10. The strands used for Figures 3.4b and d were:

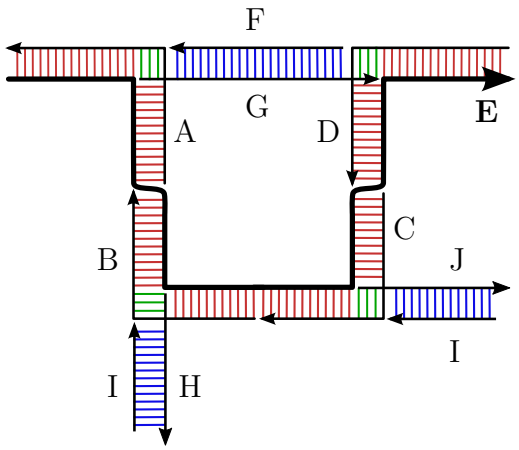
Strand name	**Sequence**
A	ACTAGAACGACG
B5 (3')	CGTCGTTCTAGTTGTAG
C (labelled)	/56-FAM/CAGTCATGGGCAGCTCTACAGGAAATCTA-TAGTTCA
D5	TGAACTATAGATTTCCAGCTGCCCATGACTG

Distance marker sequences (see section 3.1.1):

Strand name	**Sequence**
High-mobility distance marker labelled strand	/56-FAM/AGTTTTAATCGCTGCCCA
High-mobility distance marker unlabelled strand	TGGGCAGCGATTAAAACT
Low-mobility distance marker labelled strand	/56-FAM/CGGACGAGGCAGCTATGGTCGTGGGCGA-TTTCAACATGACGTCAGTCAATACGGGGAGAAGAC-CAGCGTCCG
Low-mobility distance marker unlabelled strand	CGGACGCTGGTCTTCTCCCCGTATTGACTGACGTC-ATGTTGAAATCGCCCACGACCATAGCTGCCTCGTC-CG

B.3 U-shaped T-motif structures (Chapter 4, section 4.1)

The strands for the U-shaped structure have been labelled as follows (braced structure is shown, unbraced structure is identical but lacks strands H-J):



Loop sequences were changed by using different A/B/C/D strands. Numbers and letters in round brackets () after these strand names refer to the length and identity of the loop e.g. 5a and 5b are both 5 nt loops with different sequences. A 'null' loop means that the loop was a poly(T) sequence. The corresponding sticky-ends on strands G, H and J are shown using the same labels e.g. H(7b) has the complementary sticky-end to the B(7b) loop. Both H and J have a 5' sticky-end end only. G has two sticky-ends, the 5' is shown first e.g. G(6a-5a) has a 6a 5' sticky-end and a 5a 3' sticky-end. For the linker length gels (Figure 4.5), different length F and G strands were used. The linker length refers to the hybridised region between each pair of F and G strands, this is shown after each in square brackets [ ]. Loop and sticky-end sequences are shown in red.

Strands used to study hierarchical linker binding (section 4.1.1)

Strand name	Sequence
A(5a)	TGGGACTAGCACCCCATGGCCAA
A(6a)	TGGGACTAGTGTCCCCATGGCCAA
A(7a)	TGGGACTAGTACACGCCATGGCCAA
A(8a)	TGGGACTAGCAGTAAGCCATGGCCAA
B(7b) (labelled)	/56-TAMN/CGGCAGGTTCTGTCGTCGCCGTTCGC
C(5c)	GGACCGGCGCTGGCTTGTCGGCG
D(5b)	CGGCGCGATTCCAGCGGCCGACAG
D(6b)	CGGCGCGATTCCACTCGGCCGACAG
E	TTGGCCATGGTAGTCCCAGCGAACGGAGAACCTGC-CGCGCCGACAAGGCCGGTCCCTGTCGGCGAATCGC-GCCG
F[20]	GCCAGGTGTGCCAGAGCAGG
G(5a-5b)[20]	GGTGCCCTGCTCTGGCACACCTGGCCGCTG
G(5a-6b)[20]	GGTGCCCTGCTCTGGCACACCTGGCCGAGTG
G(6a-5b)[20]	GACACCCTGCTCTGGCACACCTGGCCGCTG
G(6a-6b)[20]	GGACACCCTGCTCTGGCACACCTGGCCGAGTG
G(7a-5b)[20]	CGTGTACCCTGCTCTGGCACACCTGGCCGCTG
G(7a-6b)[20]	CGTGTACCCTGCTCTGGCACACCTGGCCGAGTG
G(8a-5b)[20]	CTTACTGCCCTGCTCTGGCACACCTGGCCGCTG
G(8a-6b)[20]	CTTACTGCCCTGCTCTGGCACACCTGGCCGAGTG

All structures used were unbraced

Strands used to study effect of linker length (section 4.1.2)

Strand name	Sequence
A(6c)	TGGGACTAGTGTCGCCATGGCCAA
B(7b)	CGGCAGGTTCTGTCGTCGCCGTTCGC
B(null)	CGGCAGGTTCTTTTTTTCCGTTCGC
C(5d)	GGACCGGCGGTCGCTTGTCGGCG
C(null)	GGACCGGCTTTTTCTTGTCGGCG
D(5e) (labelled)	/56-FAM/CGGCGCGATTCCGACGGCCGACAG
D(5f) (labelled)	/56-TAMN/CGGCGCGATTCTCATCGCCGACAG
E	TTGGCCATGGTAGTCCCAGCGAACGGAGAACCTGC-CGCGCCGACAAGGCCGGTCCCTGTCGGCGAATCGC-GCCG
F[16]	GCCAGGTGTGCCAGAG
F[17]	GCCAGGTGTGCCAGAGC
F[18]	GCCAGGTGTGCCAGAGCA
F[19]	GCCAGGTGTGCCAGAGCAG

Strand name	Sequence
F[20]	GCCAGGTGTGCCAGAGCAGG
F[21]	GCCAGGTGTGCCAGAGCAGGC
F[22]	GCCAGGTGTGCCAGAGCAGGCA
F[23]	GCCAGGTGTGCCAGAGCAGGCAT
F[24]	GCCAGGTGTGCCAGAGCAGGCATT
F[25]	GCCAGGTGTGCCAGAGCAGGCATTC
F[26]	GCCAGGTGTGCCAGAGCAGGCATTCA
F[27]	GCCAGGTGTGCCAGAGCAGGCATTCAT
F[28]	GCCAGGTGTGCCAGAGCAGGCATTCATG
F[29]	GCCAGGTGTGCCAGAGCAGGCATTCATGG
G(6c-5e)[16]	CGACACCTCTGGCACACCTGGCCGTCG
G(6c-5e)[17]	CGACACGCTCTGGCACACCTGGCCGTCG
G(6c-5e)[18]	CGACACTGCTCTGGCACACCTGGCCGTCG
G(6c-5e)[19]	CGACACCTGCTCTGGCACACCTGGCCGTCG
G(6c-5e)[20]	CGACACCCTGCTCTGGCACACCTGGCCGTCG
G(6c-5e)[21]	CGACACGCCTGCTCTGGCACACCTGGCCGTCG
G(6c-5e)[22]	CGACACTGCCTGCTCTGGCACACCTGGCCGTCG
G(6c-5e)[23]	CGACACATGCCTGCTCTGGCACACCTGGCCGTCG
G(6c-5e)[24]	CGACACAATGCCTGCTCTGGCACACCTGGCCGTCG
G(6c-5e)[25]	CGACACGAATGCCTGCTCTGGCACACCTGGCCGTCG
G(6c-5e)[26]	CGACACTGAATGCCTGCTCTGGCACACCTGGCCGTCG
G(6c-5e)[27]	CGACACATGAATGCCTGCTCTGGCACACCTGGCCGTCG
G(6c-5e)[28]	CGACACCATGAATGCCTGCTCTGGCACACCTGGCCGTC-G
G(6c-5e)[29]	CGACACCCATGAATGCCTGCTCTGGCACACCTGGCCGT-CG
H(7b)	CGACGACGACAGGTCCGACCTGC
I	GCAGGTCGGACCTGTC
J(5d)	CGACCGACAGGTCCGACCTGC

Unbraced structures used B null, C null and no H/I/J strands. Braced structures used B(7d), C(5d) and H/I/J strands.

Distance marker sequences:

Strand name	Sequence
High-mobility distance marker labelled strand	/Cy5/TTGCGCTCACTGCCCGACGTTGTTGCCATTG
High-mobility distance marker unlabelled strand	CAATGGCAACAACGTCTGCACCCGCGGCGCACGAC-GGGCCGC
High-mobility distance marker unlabelled strand	CGCCAGGCCAAACCGCGCGGGTGCAGGGCAGTGA-GCGCAA
Low-mobility distance marker labelled strand	Nucleotides 1090-2545 of pUC19, with 5' Cy5 label [160]
Low-mobility distance marker unlabelled strand	Reverse complement of low-mobility distance marker labelled strand

PCR primers used to amplify low-mobility distance marker:

Strand name	Sequence
Forward primer	/5Cy5/GTGTAGGTCGTTCGCTCCAAG
Reverse primer	ATCCGCTCATGAGACAATAACC

B.4 T-motif origami (Chapter 4, section 4.2)

Adapter names are coloured as per Figures in section 4.2 ('adapter 1' below is the rightmost adapter in Figure 4.10; adapters are numbered from right to left following the template strand). Loop and sticky-end sequences are shown in red, complementary hybridisation regions (see section 4.2.2.2) are in green.

Strand name	Sequence
Template strand	CTGCGTAAGCTTCGTATTAAATCCTTTGCCCGAACGTTATTAATTTTAAAAGTTTGAGTAACATTATCATTTTGCGGAACAAAGAAACCACCAGAAGGAGCGGAATTATCATCATATTCCTGATTATCAGATGATGGCAATTCATCAATATAATCCTGATTGTTTGGATTATACTTCTGAATAATGGAAGGGTTAGAACCTACCATATCAAAATTATTTGCACGTAAAACAGAAATAAAGAAATTGCGTAGATTTTCAGGTTTAACGTCAGATGAATATACAGTAACAGTACCTTTTACATCGGAGGAAACAATAACGGATTCGCCTGATTGCTTTGAATACCAAGTTACAAAATCGCGCAGAGGCGAATTATTCATTTCAATTACCTGAGCAAAAGAAGATGATGAAACAAACATCAAGAAAACAAAATTAATTACATTTAACAATTTCATTTGAATTACCTTTTTTAATGGAAACAGTACATAAATCAATATATGTGAGTGAATAACCTTGCTTCTGTAAATCGTCGACTGGGTA
Adapter 1	TACCCAGTCGACGATTTACAGAGGCTCAGCAAGGT
Adapter 2	TATTCACTGTGTCGCACATATATTG
Adapter 3	ATTTATGTACCGACGTGTTTCCA
Adapter 4	TTAAAAAACAGGTGGGTAATTCAAA
Adapter 5	TGAAATTGTTGGCTCAAATGTAA
Adapter 6	TTAATTTTGTGTCGGTTTTCTTGAT
Adapter 7	GTTTGTTTCACGACGTCATCTTC
Adapter 8	TTTTGCTCCAGGTGAGGTAATTGAA
Adapter 9	ATGAATAATTGGCTCCGCCTCTG
Adapter 10	CGCGATTTGTGTCGTGTAACTTGGT
Adapter 11	ATTCAAAGCACGACGATCAGGCG
Adapter 12	AATCCGTTCAGGTGATTGTTTCTCC
Adapter 13	CGATGTAAAAGGCTCGGTACTGT
Adapter 14	TACTGTATGTGTCGATTCATCTGAC
Adapter 15	GTTAAACCTGCGACGAAAATCTA
Adapter 16	CGCAATTTCAGGTGCTTTATTTCTG
Adapter 17	TTTTACGTGCGGCTCAAATAATT
Adapter 18	TTGATATGGTGTCGGTAGGTTCTAA
Adapter 19	CCCTTCCATTCGACGATTCAGAA
Adapter 20	GTATAATCCAGGTGCAAACAATCAG
Adapter 21	GATTATATTGGGCTCATGAATTG
Adapter 22	CCATCATCGTGTCGTGATAATCAGG
Adapter 23	AATATGATGACGACGTAATTCCG
Adapter 24	CTCCTTCTCAGGTGGGTGGTTTCTT

Strand name	Sequence
Adapter 25	TGTTCCGCAAGGCTCAATGATAA
Adapter 26	TGTTACTCGTGTCGAAACTTTTAAA
Adapter 27	ATTAATAACGCGACGTTCGGGCA
Adapter 28	AAGGATTTCAGGTGAATACGAAGCTTACGCAG
Splint strand	TACTGTATATTCATCTGACGTTAAACCTGAAAATCTA
Short linker strand	GCCAGGTGTGCCAGAGCAGG
Long strand of linker A	CACCTGCCTGCTCTGGCACACCTGGCGAGCC
Long strand of linker B	CGACACCCTGCTCTGGCACACCTGGCCGTCG
Long strand of linker C	CGTCGCCTGCTCTGGCACACCTGGCCGACAC
Adapter 1 + overhang with 6 nt hybridisation region	GTTCAGTAGAGTCGACGATTTACAGAGGCTCAGCAAGGT
Adapter 28 + overhang with 6 nt hybridisation region	AAGGATTTCAGGTGAATACGAAGCTTGGCGTAATCTGAAC
Adapter 1 + overhang with 8 nt hybridisation region	GTTCAGCGTAGAGTCGACGATTTACAGAGGCTCAGCAAGGT
Adapter 28 + overhang with 8 nt hybridisation region	AAGGATTTCAGGTGAATACGAAGCTTGGCGTAATCGCTGAAC
Adapter 1 + overhang with 10 nt hybridisation region	GTTCAGCGCGTAGAGTCGACGATTTACAGAGGCTCAGCAAGGT
Adapter 28 + overhang with 10 nt hybridisation region	AAGGATTTCAGGTGAATACGAAGCTTGGCGTAATCGCGCTGAAC

Adapters 1 and 28 with overhangs containing 10 nt hybridisation regions were used for Figure 4.15b, those with 8 nt hybridisation regions were used for Figures 4.18, 4.19a, 4.20a and 4.21b.

PCR primers used to amplify template strand:

Strand name	Sequence
Forward primer	CTGCGTAAGCTTCGTATTAAATCC
Reverse primer (with 5' phosphate group)	/5Phos/TACCCAGTCGACGATTTACAGAAG

Appendix C

Program codes

C.1 Measuring relative electrophoretic mobility using distance markers (Fiji macro)

Below is the macro used to determine the positions of the T-motif structure and distance markers for the single T-motif gels (see Figure 3.3). This program was adapted and used to determine the positions of the U-shaped structures and distance markers for the linker length gels (Figure 4.5); see comments (preceded by //) in program for adaptions made.

```
//User inputs number of lanes
lane_count = getNumber("How many lanes are there?", 13);
lane_heights = newArray(lane_count);

//Rotate gel
run("Rotate 90 Degrees Left");

//Define lane positions
for (i=0; i<lane_count; i++) {
waitForUser("Draw a horizontal line through the centre of lane " + i+1 + ",
through all bands");
getLine(x1, y1, x2, y2, lineWidth);
lane_heights[i] = y1;
}

if (x1<x2) {
start_box = x1;
end_box = x2;
}
else {
start_box = x2;
end_box = x1;
}

box_width = end_box - start_box;

//Input parameters
lane_width = 20; //Various lane widths were tested, all gave very similar results
```

```
peak_tolerance = 200; //This was optimised for the single T-motif gels; a peak
tolerance of 2 was used for the linker length gels.

//Get local maxima of profile
//Here, the program finds the top 3 local maxima. For the linker length gels,
the program scans through all three channels and finds one maximum for the red
and green channels (U-shaped structures) and 2 for the blue channel (distance
markers).
for (i=0; i<lane_count; i++) {
top_of_box = lane_heights[i] - (lane_width/2);
setTool("rectangle");
makeRectangle(start_box, top_of_box, box_width, lane_width);
run("Clear Results");
profile = getProfile();
maxes = Array.findMaxima(profile,peak_tolerance);
line="";
for (j=0;j<3;j++) {
new_start_of_box = start_box + maxes[j] - 20;
new_box_width = 50;
setTool("rectangle");
makeRectangle(new_start_of_box, top_of_box, new_box_width, lane_width);
waitForUser("Y");
run("Set Measurements...", "center redirect=None decimal=4");
run("Measure");
peak = getResult("XM",j);
line = line + d2s(peak,4) + "\t";
}
print(i+1 + "\t" + line);
}
```

C.2 Single T-motif binding curve fitting (Matlab)

Below is the Matlab program used to determine the K_d and B_{max} values for the single T-motifs (results shown in Table 3.1 and 3.3). Shown is an experiment for a single gel repeat (3' n=5 repeat 1). Comments are preceded by %.

```
% Input x (concentration) and y (relative mobility) values

concentrations = [0.424543946932007 0.849087893864013
2.12271973466003 4.24543946932007 6.3681592039801
8.49087893864013 12.7363184079602 16.9817578772803
21.2271973466003 31.8407960199005 42.4543946932007];

relative_mobilities = [0.00227918023755241 0.0088342367928016
0.0103272546318959 0.0210831724160957 0.0251171965802017
0.0300851986348786 0.0417941245029146 0.0491992077360604
0.054089042958321 0.066290375093928 0.0759078890931055];

% The weighting factor is the inverse square of the error
(resulting from the resolution of gel images)
weights = [231.1037382 220.8819125 213.8057062 227.469635
233.2369317 230.7997962 218.5103771 226.2646673 209.7311192
221.1792545 237.8410606];

% The model uses the equation for one-site binding
modelfun = @(b,x)((b(1).*x)./(b(2)+x));

% Initial guess at Bmax and Kd (respectively)
beta0 = [0.1 20];

% Fitting the model
mdl = fitnlm(concentrations,relative_mobilities,modelfun,beta0,
'Weights',weights);

% Output of parameters and their standard errors
Kd = mdl.Coefficients.Estimate(2)
Bmax = mdl.Coefficients.Estimate(1)
Kd_SE = mdl.Coefficients.SE(2)
Bmax_SE  = mdl.Coefficients.SE(1)
```

Bibliography

[1] E. Akbari, M. Y. Mollica, C. R. Lucas, S. M. Bushman, R. A. Patton, M. Shahhosseini, J. W. Song, and C. E. Castro. Engineering Cell Surface Function with DNA Origami. *Advanced Materials*, 29(46):1703632, oct 2017.

[2] T. Akera and V.-J. K. Cheng. A simple method for the determination of affinity and binding site concentration in receptor binding studies. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 470(3):412–423, nov 1977.

[3] B. Albrechts, D. S. Hautzinger, M. Krüger, M. C. Elwenspoek, K. M. Müller, and J. G. Korvink. Adsorption studies of DNA origami on silicon dioxide. In L. Abelmann, J. Groenland, J. van Honschoten, and H. Verputten, editors, *MME2010 Abstracts: 21st Micromechanics and Micro systems Europe Workshop*, pages 153–156, sep 2010.

[4] A. Amodio, A. F. Adedeji, M. Castronovo, E. Franco, and F. Ricci. pH-controlled assembly of DNA tiles. *Journal of the American Chemical Society*, 138(39):12735–12738, sep 2016.

[5] 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):73–76, may 2009.

[6] J.-M. Arbona, J.-P. Aimé, and J. Elezgaray. Cooperativity in the annealing of DNA origamis. *The Journal of Chemical Physics*, 138(1):015105, jan 2013.

[7] H. Asanuma, X. Liang, H. Nishioka, D. Matsunaga, M. Liu, and M. Komiyama. Synthesis of azobenzene-tethered DNA for reversible photo-regulation of DNA functions: hybridization and transcription. *Nature Protocols*, 2(1):203–212, jan 2007.

[8] H. Atsumi and A. M. Belcher. DNA Origami and G-Quadruplex Hybrid Complexes Induce Size Control of Single-Walled Carbon Nanotubes via Biological Activation. *ACS Nano*, 12(8):7986–7995, jul 2018.

[9] M. Ayub and H. Bayley. Engineered transmembrane pores. *Current Opinion in Chemical Biology*, 34:117–126, 2016.

[10] W. Bae, K. Kim, D. Min, J.-K. Ryu, C. Hyeon, and T.-Y. Yoon. Programmed folding of DNA origami structures through single-molecule force control. *Nature communications*, 5:5654, dec 2014.

[11] C. F. Barbas, D. R. Burton, J. K. Scott, and G. J. Silverman. Quantitation of DNA and RNA. *Cold Spring Harbor Protocols*, nov 2007.

[12] L. Bell and B. Byers. Separation of branched from linear DNA by two-dimensional gel electrophoresis. *Analytical Biochemistry*, 130(2):527–535, apr 1983.

[13] E. Benson, A. Mohammed, J. Gardell, S. Masich, E. Czeizler, P. Orponen, and B. Högberg. DNA rendering of polyhedral meshes at the nanoscale. *Nature*, 523(7561):441–444, jul 2015.

[14] K. E. Bujold, A. Lacroix, and H. F. Sleiman. DNA Nanostructures at the Interface with Biology. *Chem*, 4(3):495–521, mar 2018.

[15] J. R. Burns, A. Seifert, N. Fertig, and S. Howorka. A biomimetic DNA-based channel for the ligand-controlled transport of charged molecular cargo across a biological membrane. *Nature Nanotechnology*, 11(2):152–156, feb 2016.

[16] Y. Cao, X. Xiang, R. Pei, Y. Li, Y. Yan, and X. Guo. Construction of a junction DNA nanostructure and modulation of the junction switching to quadruplexes. *Royal Society Open Science*, 4(12):171337, dec 2017.

[17] J. Casey and N. Davidson. Rates of formation and thermal stabilities of RNA:DNA and DNA:DNA duplexes at high concentrations of formamide. *Nucleic Acids Research*, 4(5):1539–1552, jun 1977.

[18] C. E. Castro, F. Kilchherr, D.-N. Kim, E. L. Shiao, T. Wauer, P. Wortmann, M. Bathe, and H. Dietz. A primer to scaffolded DNA origami. *Nature Methods*, 8(3):221–229, 2011.

[19] H. Chandran, N. Gopalkrishnan, B. Yurke, and J. Reif. Meta-DNA: synthetic biology via DNA nanostructures and hybridization reactions. *Journal of the Royal Society, Interface*, 9(72):1637–1653, jul 2012.

[20] J. Chao, Y. Lin, H. Liu, L. Wang, and C. Fan. DNA-based plasmonic nanostructures. *Materials Today*, 18(6):326–335, jul 2015.

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

[22] A. Czogalla, H. G. Franquelim, and P. Schwille. DNA Nanostructures on Membranes as Tools for Synthetic Biology. *Biophysical journal*, 110(8):1698–1707, apr 2016.

[23] F. Dannenberg, K. E. Dunn, J. Bath, M. Kwiatkowska, A. J. Turberfield, and T. E. Ouldridge. Modelling DNA origami self-assembly at the domain level. *The Journal of Chemical Physics*, 143(16):165102, oct 2015.

[24] F. G. W. Dannenberg. *Modelling and verification for DNA nanotechnology*. PhD thesis, University of Oxford, 2016.

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

[26] Y. Dong, Y. R. Yang, Y. Zhang, D. Wang, X. Wei, S. Banerjee, Y. Liu, Z. Yang, H. Yan, and D. Liu. Cuboid Vesicles Formed by Frame-Guided Assembly on DNA Origami Scaffolds. *Angewandte Chemie International Edition*, 56(6):1586–1589, dec 2016.

[27] Y. Dong, Z. Yang, and D. Liu. DNA Nanotechnology Based on i-Motif Structures. *Accounts of Chemical Research*, 47(6):1853–1860, jun 2014.

[28] S. M. Douglas, I. Bachelet, and G. M. Church. A Logic-Gated Nanorobot for Targeted Transport of Molecular Payloads. *Science*, 335(6070):831–834, feb 2012.

[29] 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*, 104(16):6644–8, apr 2007.

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

[31] 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, aug 2009.

[32] C. Ducani, C. Kaul, M. Moche, W. M. Shih, and B. Högberg. Enzymatic production of 'monoclonal stoichiometric' single-stranded DNA oligonucleotides. *Nature Methods*, 10(7):647–652, jul 2013.

[33] K. E. Dunn. *DNA Origami Assembly*. PhD thesis, University of Oxford, 2014.

[34] 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, aug 2015.

[35] A. C. Eischeid. SYTO dyes and EvaGreen outperform SYBR Green in real-time PCR. *BMC research notes*, 4:263, jul 2011.

[36] J. Elbaz, P. Yin, and C. A. Voigt. Genetic encoding of DNA nanostructures and their self-assembly in living bacteria. *Nature Communications*, 7:11179, apr 2016.

[37] M. P. Fisher and C. W. Dingman. Role of molecular conformation in determining the electrophoretic properties of polynucleotides in agarose-acrylamide composite gels. *Biochemistry*, 10(10):1895–1899, may 1971.

[38] J. Fu, Y. R. Yang, A. Johnson-Buck, M. Liu, Y. Liu, N. G. Walter, N. W. Woodbury, and H. Yan. Multi-enzyme complexes on DNA scaffolds capable of substrate channelling with an artificial swinging arm. *Nature Nanotechnology*, 9(7):531–536, may 2014.

[39] A. J. Genot, D. Y. Zhang, J. Bath, and A. J. Turberfield. Remote Toehold: A Mechanism for Flexible Control of DNA Hybridization Kinetics. *Journal of the American Chemical Society*, 133(7):2177–2182, feb 2011.

[40] T. Gerling, K. F. Wagenbauer, A. M. Neuner, and H. Dietz. Dynamic DNA devices and assemblies formed by shape-complementary, non-base pairing 3D components. *Science*, 347(6229):1446–52, mar 2015.

[41] A. Gietl, P. Holzmeister, D. Grohmann, and P. Tinnefeld. DNA origami as biocompatible surface to match single-molecule and ensemble experiments. *Nucleic Acids Research*, 40(14):e110, aug 2012.

[42] R. P. Goodman. Rapid Chiral Assembly of Rigid DNA Building Blocks for Molecular Nanofabrication. *Science*, 310(5754):1661–1665, dec 2005.

[43] R. P. Goodman, R. M. Berry, and A. J. Turberfield. The single-step synthesis of a DNA tetrahedron. *Chemical Communications*, 1(12):1372–1373, jun 2004.

[44] K. Göpfrich, C.-Y. Li, M. Ricci, S. P. Bhamidimarri, J. Yoo, B. Gyenes, A. Ohmann, M. Winterhalter, A. Aksimentiev, and U. F. Keyser. Large-Conductance Transmembrane Porin Made from DNA Origami. *ACS Nano*, 10(9):8207–8214, aug 2016.

[45] L. N. Green, A. Amodio, H. K. K. Subramanian, F. Ricci, and E. Franco. pH-Driven Reversible Self-Assembly of Micron-Scale DNA Scaffolds. *Nano Letters*, 17(12):7283–7288, nov 2017.

[46] M. R. Green and J. Sambrook. Isolation and Quantification of DNA. *Cold Spring Harbor Protocols*, 2018, jun 2018.

[47] G. Grossi, A. Jaekel, E. S. Andersen, and B. Saccà. Enzyme-functionalized DNA nanostructures as tools for organizing and controlling enzymatic reactions. *MRS Bulletin*, 42(12):920–924, dec 2017.

[48] S. Hamada and S. Murata. Substrate-assisted assembly of interconnected single-duplex DNA nanostructures. *Angewandte Chemie International Edition*, 48(37):6820–6823, 2009.

[49] D. Han, S. Pal, Y. Yang, S. Jiang, J. Nangreave, Y. Liu, and H. Yan. DNA Gridiron Nanostructures Based on Four-Arm Junctions. *Science*, 339(6126):1412–1415, mar 2013.

[50] D. Han, X. Qi, C. Myhrvold, B. Wang, M. Dai, S. Jiang, M. Bates, Y. Liu, B. An, F. Zhang, H. Yan, and P. Yin. Single-stranded DNA and RNA origami. *Science*, 358(6369):eaao2648, dec 2017.

[51] L. Heinen and A. Walther. Temporal control of i-motif switch lifetimes for autonomous operation of transient DNA nanostructures. *Chemical Science*, 8(5):4100–4107, 2017.

[52] L. M. Hellman and M. G. Fried. Electrophoretic mobility shift assay (EMSA) for detecting protein-nucleic acid interactions. *Nature Protocols*, 2(8):1849–61, 2007.

[53] P. S. Ho. Structure of the Holliday junction: applications beyond recombination. *Biochemical Society Transactions*, 45(5):1149–1158, oct 2017.

[54] B. H{\"o}gberg, T. Liedl, and W. M. Shih. Folding DNA Origami from a Double-Stranded Source of Scaffold. *Journal of the American Chemical Society*, 131(26):9154–9155, jul 2009.

[55] D. L. Holmes and N. C. Stellwagen. Estimation of polyacrylamide gel pore size from Ferguson plots of linear DNA fragments. II. Comparison of gels with different crosslinker concentrations, added agarose and added linear polyacrylamide. *Electrophoresis*, 12(9):612–619, sep 1991.

[56] F. Hong, F. Zhang, Y. Liu, and H. Yan. DNA Origami: Scaffolds for Creating Higher Order Structures. *Chemical Reviews*, 117(20):12584–12640, jun 2017.

[57] R. Iinuma, Y. Ke, R. Jungmann, T. Schlichthaerle, J. B. Woehrstein, and P. Yin. Polyhedra self-assembled from DNA tripods and characterized with 3D DNA-PAINT. *Science*, 344(6179):65–69, apr 2014.

[58] H. Ijäs, S. Nummelin, B. Shen, M. Kostiainen, and V. Linko. Dynamic DNA Origami Devices: from Strand-Displacement Reactions to External-Stimuli Responsive Systems. *International Journal of Molecular Sciences*, 19(7):2114, jul 2018.

[59] W. M. Jacobs and D. Frenkel. Self-Assembly of Structures with Addressable Complexity. *Journal of the American Chemical Society*, 138(8):2457–2467, mar 2016.

[60] W. M. Jacobs, A. Reinhardt, and D. Frenkel. Rational design of self-assembly pathways for complex multicomponent structures. *Proceedings of the National Academy of Sciences*, 112(20):6313–6318, may 2015.

[61] 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):10062–10063, aug 2008.

[62] 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):4756–4761, nov 2010.

[63] R. Kageyama, I. Kawamata, K. Tanabe, Y. Suzuki, S.-i. Nomura, and S. Murata. Construction of T-motif-based DNA nanostructures using enzymatic reactions. *ChemBioChem*, 19(8):873–876, feb 2018.

[64] N. R. Kallenbach, R.-I. Ma, and N. C. Seeman. An immobile nucleic acid junction constructed from oligonucleotides. *Nature*, 305(5937):829–831, oct 1983.

[65] Y. Kamiya and H. Asanuma. Light-Driven DNA Nanomachine with a Photoresponsive Molecular Engine. *Accounts of Chemical Research*, 47(6):1663–1672, jun 2014.

[66] B. Kankia, D. Gvarjaladze, A. Rabe, L. Lomidze, N. Metreveli, and K. Musier-Forsyth. Stable Domain Assembly of a Monomolecular DNA Quadruplex: Implications for DNA-Based Nanoswitches. *Biophysical Journal*, 110(10):2169–2175, may 2016.

[67] G. Ke, M. Liu, S. Jiang, X. Qi, Y. R. Yang, S. Wootten, F. Zhang, Z. Zhu, Y. Liu, C. J. Yang, and H. Yan. Directional Regulation of Enzyme Pathways through the Control of Substrate Channeling on a DNA Origami Scaffold. *Angewandte Chemie*, 55(26):7483–7486, may 2016.

[68] Y. Ke, G. Bellot, N. V. Voigt, E. Fradkov, and W. M. Shih. Two design strategies for enhancement of multilayer-DNA-origami folding: underwinding for specific intercalator rescue and staple-break positioning. *Chemical Science*, 3(8):2587–2597, aug 2012.

[69] Y. Ke, S. M. Douglas, M. Liu, J. Sharma, A. Cheng, A. Leung, Y. Liu, W. M. Shih, and H. Yan. Multilayer DNA origami packed on a square lattice. *Journal of the American Chemical Society*, 131(43):15903–15908, nov 2009.

[70] Y. Ke, S. Lindsay, Y. Chang, Y. Liu, and H. Yan. Self-Assembled Water-Soluble Nucleic Acid Probe Tiles for Label-Free RNA Hybridization Assays. *Science*, 319(5860):180–183, jan 2008.

[71] Y. Ke, L. L. Ong, W. M. Shih, and P. Yin. Three-dimensional structures self-assembled from DNA bricks. *Science*, 338(6111):1177–1183, nov 2012.

[72] B. Kick, F. Praetorius, H. Dietz, and D. Weuster-Botz. Efficient Production of Single-Stranded Phage DNA as Scaffolds for DNA Origami. *Nano Letters*, 15(7):4672–4676, jul 2015.

[73] V. Kočar, J. S. Schreck, S. Čeru, H. Gradišar, N. Bašić, T. Pisanski, J. P. K. Doye, and R. Jerala. Design principles for rapid folding of knotted DNA nanostructures. *Nature Communications*, 1(7):10803, feb 2016.

[74] A. Kopielski, A. Schneider, A. Csáki, and W. Fritzsche. Isothermal DNA origami folding: avoiding denaturing conditions for one-pot, hybrid-component annealing. *Nanoscale*, 7(5):2102–2106, 2015.

[75] A. Krissanaprasit, M. Madsen, J. B. Knudsen, D. Gudnason, W. Surareungchai, V. Birkedal, and K. V. Gothelf. Programmed Switching of Single Polymer Conformation on DNA Origami. *ACS Nano*, 10(2):2243–2250, feb 2016.

[76] A. Kuzuya and M. Komiyama. Design and construction of a box-shaped 3D-DNA origami. *Chemical Communications*, 1(28):4182, jul 2009.

[77] A. Kuzuya and M. Komiyama. DNA origami: Fold, stick, and beyond. *Nanoscale*, 2(3):310–322, 2010.

[78] A. Kuzyk, Y. Yang, X. Duan, S. Stoll, A. O. Govorov, H. Sugiyama, M. Endo, and N. Liu. A light-driven three-dimensional plasmonic nanosystem that translates molecular motion into reversible chiroptical function. *Nature Communications*, 1(7):10591, feb 2016.

[79] H. W. Lau and L. A. Archer. Gel electrophoresis of linear and star-branched DNA. *Physical Review E*, 84(6):061916, dec 2011.

[80] J. Lee Tin Wah, C. David, S. Rudiuk, D. Baigl, and A. Estevez-Torres. Observing and Controlling the Folding Pathway of DNA Origami at the Nanoscale. *ACS Nano*, 10(2):1978–1987, feb 2016.

[81] P. A. Levene. The Structure of Yeast Nucleic Acid: IV. Ammonia Hydrolysis. *The Journal of Biological Chemistry*, 1(40):415–424, 1919.

[82] M. Levitt. How many base-pairs per turn does DNA have in solution and in chromatin? Some theoretical calculations. *Proceedings of the National Academy of Sciences*, 75(2):640–644, feb 1978.

[83] M. Li, H. Zuo, J. Yu, X. Zhao, and C. Mao. One DNA strand homo-polymerizes into defined nanostructures. *Nanoscale*, 9(30):10601–10605, 2017.

[84] X. Li, C. Zhang, C. Hao, C. Tian, G. Wang, and C. Mao. DNA polyhedra with T-linkage. *ACS Nano*, 6(6):5138–5142, jun 2012.

[85] D. M. Lilley and T. J. Wilson. Fluorescence resonance energy transfer as a structural tool for nucleic acids. *Current Opinion in Chemical Biology*, 4(5):507–517, oct 2000.

[86] D. Liu, M. Wang, Z. Deng, R. Walulu, and C. Mao. Tensegrity: construction of rigid DNA triangles with flexible four-arm DNA junctions. *Journal of the American Chemical Society*, 126(8):2324–2325, mar 2004.

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

[88] M. Lu, Q. Guo, L. A. Marky, N. C. Seeman, and N. R. Kallenbach. Thermodynamics of DNA branching. *Journal of Molecular Biology*, 223(3):781–789, feb 1992.

[89] R. R. F. Machinek, T. E. Ouldridge, N. E. C. Haley, J. Bath, and A. J. Turberfield. Programmable energy landscapes for kinetic control of DNA strand displacement. *Nature Communications*, 5(1):5324, dec 2014.

[90] J. M. Majikes. *DNA Origami Folding Pathways: Implications for Design, Thermodynamics, and Kinetics*. PhD thesis, North Carolina State University, 2016.

[91] J. M. Majikes, L. C. C. Ferraz, and T. H. LaBean. pH-driven actuation of DNA origami via parallel I-motif sequences in solution and on surfaces. *Bioconjugate Chemistry*, 28(7):1821–1825, jun 2017.

[92] J. M. Majikes, J. A. Nash, and T. H. LaBean. Competitive annealing of multiple DNA origami: formation of chimeric origami. *New Journal of Physics*, 18(11):115001, nov 2016.

[93] J. M. Majikes, J. A. Nash, T. H. LaBean, E. C. Samano, T. H. LaBean, J. Bath, J.-P. Aime, K. V. Gothelf, F. Besenbacher, and M. Dong. Search for effective chemical quenching to arrest molecular assembly and directly monitor DNA nanostructure formation. *Nanoscale*, 9(4):1637–1644, 2017.

[94] F. Mao, W.-Y. Leung, and X. Xin. Characterization of EvaGreen and the implication of its physicochemical properties for qPCR applications. *BMC Biotechnology*, 7:76, nov 2007.

[95] A. E. Marras, L. Zhou, V. Kolliopoulos, H.-J. Su, and C. E. Castro. Directing folding pathways for multi-component DNA origami nanostructures with complex topology. *New Journal of Physics*, 18(5):055005, may 2016.

[96] A. E. Marras, L. Zhou, H.-J. Su, and C. E. Castro. Programmable motion of DNA origami mechanisms. *Proceedings of the National Academy of Sciences*, 112(3):713–8, jan 2015.

[97] T. G. Martin, T. A. M. Bharat, A. C. Joerger, X.-C. Bai, F. Praetorius, A. R. Fersht, H. Dietz, and S. H. W. Scheres. Design of a molecular support for cryo-EM structure determination. *Proceedings of the National Academy of Sciences*, 113(47):E7436–E7463, nov 2016.

[98] T. G. Martin and H. Dietz. Magnesium-free self-assembly of multi-layer DNA objects. *Nature Communications*, 3(1):1103, jan 2012.

[99] D. L. Massart. *Handbook of chemometrics and qualimetrics. Part A*. Elsevier, 1997.

[100] H. J. Motulsky and L. A. Ransnas. Fitting curves to data using nonlinear regression: a practical and nonmathematical review. *FASEB Journal*, 1(5):365–74, nov 1987.

[101] J. Müller. Metal-mediated base pairs in parallel-stranded DNA. *Beilstein Journal of Organic Chemistry*, 13(1):2671–2681, dec 2017.

[102] P. M. Nafisi, T. Aksel, and S. M. Douglas. Construction of a novel phagemid to produce custom DNA origami scaffolds. *Synthetic Biology*, 3(1):ysy015, aug 2018.

[103] D. Nečas and P. Klapetek. Gwyddion: an open-source software for SPM data analysis. *Open Physics*, 10(1):181–188, jan 2012.

[104] New England Biolabs. High Efficiency Transformation Protocol, 2010.

[105] T. A. Ngo, E. Nakata, M. Saimura, and T. Morii. Spatially organized enzymes drive cofactor-coupled cascade reactions. *Journal of the American Chemical Society*, 138(9):3021–3021, 2016.

[106] S. Niekamp, K. Blumer, P. M. Nafisi, K. Tsui, J. Garbutt, and S. M. Douglas. Folding complex DNA nanostructures from limited sets of reusable sequences. *Nucleic acids research*, 44(11):e102, apr 2016.

[107] L. Olejko, P. J. Cywiński, and I. Bald. An ion-controlled four-color fluorescent telomeric switch on DNA origami structures. *Nanoscale*, 1(19):10339, may 2016.

[108] N. Otsu. A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1):62–66, jan 1979.

[109] T. E. Ouldridge, A. A. Louis, and J. P. K. Doye. Structural, mechanical, and thermodynamic properties of a coarse-grained DNA model. *The Journal of Chemical Physics*, 134(8):085101, feb 2011.

[110] S. H. Park, C. Pistol, S. J. Ahn, J. H. Reif, A. R. Lebeck, C. Dwyer, and T. H. LaBean. Finite-Size, Fully Addressable DNA Tile Lattices Formed by Hierarchical Assembly Procedures. *Angewandte Chemie International Edition*, 45(5):735–739, jan 2006.

[111] A. Porchetta, A. Vallée-Bélisle, K. W. Plaxco, and F. Ricci. Allosterically Tunable, DNA-Based Switches Triggered by Heavy Metals. *Journal of the American Chemical Society*, 135(36):13238–13241, sep 2013.

[112] F. Praetorius, B. Kick, K. L. Behler, M. N. Honemann, D. Weuster-Botz, and H. Dietz. Biotechnological mass production of DNA origami. *Nature*, 552(7683):84–87, dec 2017.

[113] L. Pray. Discovery of DNA Double Helix: Watson and Crick — Learn Science at Scitable, 2008.

[114] Qiagen. Re-Purification of Plasmid DNA Prepared by Methods other than QIAGEN Tips.

[115] E.-A. Raiber, R. Hardisty, P. van Delft, and S. Balasubramanian. Mapping and elucidating the function of modified bases in DNA. *Nature Reviews Chemistry*, 1(9):0069, sep 2017.

[116] W. A. Rees, T. D. Yager, J. Korte, and P. H. Von Hippel. Betaine can eliminate the base pair composition dependence of DNA melting. *Biochemistry*, 32(1):137–144, jan 1993.

[117] N. J. Reiter, L. J. Maher, and S. E. Butcher. DNA mimicry by a high-affinity anti-NF-κB RNA aptamer. *Nucleic Acids Research*, 36(4):1227–1236, mar 2008.

[118] L. P. Reynaldo, A. V. Vologodskii, B. P. Neri, and V. I. Lyamichev. The kinetics of oligonucleotide replacements. *Journal of Molecular Biology*, 297(2):511–520, mar 2000.

[119] S. Rinker, Y. 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, jul 2008.

[120] M. Rossetti and A. Porchetta. Allosterically regulated DNA-based switches: From design to bioanalytical applications. *Analytica Chimica Acta*, 1012:30–41, jul 2018.

[121] C. Rossi-Gendron. DNA origami melting: precise measurement and photocontrol. In *Poster session at DNATEC17*, Dresden, 2017.

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

[123] B. SaccÃ , R. Meyer, and C. M. Niemeyer. Analysis of the Self-Assembly of 4Ã—4 DNA Tiles by Temperature-Dependent FRET Spectroscopy. *ChemPhysChem*, 10(18):3239–3248, dec 2009.

[124] J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J.-Y. Tinevez, D. J. White, V. Hartenstein, K. Eliceiri, P. Tomancak, and A. Cardona. Fiji: an open-source platform for biological-image analysis. *Nature Methods*, 9(7):676–682, jul 2012.

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

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

[127] N. C. Seeman. DNA in a material world. *Nature*, 421(6921):427–31, jan 2003.

[128] N. C. Seeman. Nanomaterials based on DNA. *Annual Review of Biochemistry*, 79:65–87, jan 2010.

[129] N. C. Seeman and N. R. Kallenbach. Design of immobile nucleic acid junctions. *Biophysical journal*, 44(2):201–209, nov 1983.

[130] L. Shi, P. Peng, Y. Du, and T. Li. Programmable i-motif DNA folding topology for a pH-switched reversible molecular sensing device. *Nucleic Acids Research*, 130:2380–2381, mar 2017.

[131] W. Shih. Biomolecular assembly: Dynamic DNA. *Nature Materials*, 7(2):98–100, feb 2008.

[132] W. M. Shih, J. D. Quispe, and G. F. Joyce. A 1.7-kilobase single-stranded DNA that folds into a nanoscale octahedron. *Nature*, 427(6975):618–21, feb 2004.

[133] B. E. K. Snodin, F. Randisi, M. Mosayebi, P. Šulc, J. S. Schreck, F. Romano, T. E. Ouldridge, R. Tsukanov, E. Nir, A. A. Louis, and J. P. K. Doye. Introducing improved structural properties and salt dependence into a coarse-grained model of DNA. *The Journal of Chemical Physics*, 142(23):234901, jun 2015.

[134] B. E. K. Snodin, F. Romano, L. Rovigatti, T. E. Ouldridge, A. A. Louis, and J. P. K. Doye. Direct Simulation of the Self-Assembly of a Small DNA Origami. *ACS Nano*, 10(2):1724–1737, jan 2016.

[135] J.-P. 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, dec 2012.

[136] J. Song, J.-M. Arbona, Z. Zhang, L. Liu, E. Xie, J. Elezgaray, J.-P. Aime, K. V. Gothelf, F. Besenbacher, and M. Dong. Direct Visualization of Transient Thermal Response of a DNA Origami. *Journal of the American Chemical Society*, 134(24):9844–9847, jun 2012.

[137] N. Srinivas, T. E. Ouldridge, P. Sulc, J. M. Schaeffer, B. Yurke, A. A. Louis, J. P. K. Doye, and E. Winfree. On the biophysics and kinetics of toehold-mediated DNA strand displacement. *Nucleic Acids Research*, 41(22):10641–10658, dec 2013.

[138] H. K. K. Subramanian, B. Chakraborty, R. Sha, and N. C. Seeman. The Label-Free Unambiguous Detection and Symbolic Display of Single Nucleotide Polymorphisms on DNA Origami. *Nano Letters*, 11(2):910–913, feb 2011.

[139] P. Šulc, F. Romano, T. E. Ouldridge, L. Rovigatti, J. P. K. Doye, and A. A. Louis. Sequence-dependent thermodynamics of a coarse-grained DNA model. *The Journal of Chemical Physics*, 137(13):135101, jul 2012.

[140] W. Sun, E. Boulais, Y. Hakobyan, W. L. Wang, A. Guan, M. Bathe, and P. Yin. Casting inorganic structures with DNA molds. *Science*, 346(6210):1258361, nov 2014.

[141] S. Surana, J. M. Bhat, S. P. Koushika, and Y. Krishnan. An autonomous DNA nanomachine maps spatiotemporal pH changes in a multicellular living organism. *Nature Communications*, 2(1):340, sep 2011.

[142] Y. Suzuki, M. Endo, and H. Sugiyama. Complexing DNA origami frameworks through sequential self-assembly based on directed docking. *Angewandte Chemie International Edition*, 57(24):7061–7065, apr 2018.

[143] A. J. Thubagere, W. Li, R. F. Johnson, Z. Chen, S. Doroudi, Y. L. Lee, G. Izatt, S. Wittman, N. Srinivas, D. Woods, E. Winfree, and L. Qian. A cargo-sorting DNA robot. *Science (New York, N.Y.)*, 357(6356):eaan6558, sep 2017.

[144] C. Tian, C. Zhang, X. Li, C. Hao, S. Ye, and C. Mao. Approaching the Limit: Can One DNA Strand Assemble into Defined Nanostructures? *Langmuir*, 30(20):5859–5862, may 2014.

[145] T. Tørring, N. V. Voigt, J. Nangreave, H. Yan, and K. V. Gothelf. DNA origami: a quantum leap for self-assembly of complex structures. *Chemical Society Reviews*, 40(12):5636–46, dec 2011.

[146] J. C. Wang. Helical repeat of DNA in solution. *Proceedings of the National Academy of Sciences*, 76(1):200–3, jan 1979.

[147] P. Wang, T. A. Meyer, V. Pan, P. K. Dutta, and Y. Ke. The Beauty and Utility of DNA Origami. *Chem*, 2(3):359–382, 2017.

[148] X. Wang, W. Lai, T. Man, X. Qu, L. Li, A. R. Chandrasekaran, and H. Pei. Bio-surface engineering with DNA scaffolds for theranostic applications. *Nanofabrication*, 4(1):1–16, may 2018.

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

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

[151] X. Wei, J. Nangreave, S. Jiang, H. Yan, and Y. Liu. Mapping the Thermal Behavior of DNA Origami Nanostructures. *Journal of the American Chemical Society*, 135(16):6165–6176, apr 2013.

[152] R. Wing, H. Drew, T. Takano, C. Broka, S. Tanaka, K. Itakura, and R. E. Dickerson. Crystal structure analysis of a complete turn of B-DNA. *Nature*, 287(5784):755–758, oct 1980.

[153] N. Wu and I. Willner. pH-Stimulated Reconfiguration and Structural Isomerization of Origami Dimer and Trimer Systems. *Nano Letters*, 16(10):6650–6655, oct 2016.

[154] X.-r. Wu, C.-w. Wu, and C. Zhang. Discrete DNA three-dimensional nanostructures: the synthesis and applications. *Chinese Journal of Polymer Science*, 35(1):1–24, jan 2017.

[155] P. Yakovchuk, E. Protozanova, and M. D. Frank-Kamenetskii. Base-stacking and base-pairing contributions into thermal stability of the DNA double helix. *Nucleic Acids Research*, 34(2):564–574, jan 2006.

[156] 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*, 301(5641):1882–4, sep 2003.

[157] Y. Yan, Y. Cao, C. Xiao, Y. Li, X. Xiang, and X. Guo. 5'-(CGA)n sequence-assisted pH-controlled assembly of supramolecular DNA nanostructure. *Royal Society Open Science*, 5(8):180123, aug 2018.

[158] Y. Yang, M. Endo, K. Hidaka, and H. Sugiyama. Photo-Controllable DNA Origami Nanostructures Assembling into Predesigned Multiorientational Patterns. *Journal of the American Chemical Society*, 134(51):20645–20653, dec 2012.

[159] Y. Yang, J. Wang, H. Shigematsu, W. Xu, W. M. Shih, J. E. Rothman, and C. Lin. Self-assembly of size-controlled liposomes on DNA nanotemplates. *Nature Chemistry*, 8(5):476–483, mar 2016.

[160] C. Yanisch-Perron, J. Vieira, and J. Messing. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mpl8 and pUC19 vectors. *Gene*, 33(1):103–119, jan 1985.

[161] R. M. Zadegan and M. L. Norton. Structural DNA nanotechnology: from design to applications. *International Journal of Molecular Sciences*, 13(6):7149–7162, jan 2012.

[162] 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(1):170–173, jan 2011.

[163] D. Y. Zhang, S. X. Chen, and P. Yin. Optimizing the specificity of nucleic acid hybridization. *Nature chemistry*, 4:208–214, 2012.

[164] D. Y. Zhang and E. Winfree. Control of DNA Strand Displacement Kinetics Using Toehold Exchange. *Journal of the American Chemical Society*, 131(47):17303–17314, dec 2009.

[165] H. Zhang, J. Chao, D. Pan, H. Liu, Y. Qiang, K. Liu, C. Cui, J. Chen, Q. Huang, J. Hu, L. Wang, W. Huang, Y. Shi, and C. Fan. DNA origami-based shape IDs for single-molecule nanomechanical genotyping. *Nature Communications*, 8:14738, apr 2017.

[166] Y. Zhang and N. C. Seeman. Construction of a DNA-Truncated Octahedron. *Journal of the American Chemical Society*, 116(5):1661–1669, mar 1994.

[167] Y. Zhang, J. Tu, D. Wang, H. Zhu, S. K. Maity, X. Qu, B. Bogaert, H. Pei, and H. Zhang. Programmable and Multifunctional DNA-Based Materials for Biomedical Applications. *Advanced Materials*, 30:1703658, feb 2018.

[168] Z. Zhang, J. Song, F. Besenbacher, M. Dong, and K. V. Gothelf. Self-Assembly of DNA Origami and Single-Stranded Tile Structures at Room Temperature. *Angewandte Chemie*, 125(35):9389–9393, aug 2013.

[169] J. Zheng, J. J. Birktoft, Y. Chen, T. Wang, R. Sha, P. E. Constantinou, S. L. Ginell, C. Mao, and N. C. Seeman. From molecular to macroscopic via the rational design of a self-assembled 3D DNA crystal. *Nature*, 461(7260):74–77, sep 2009.

[170] L. Zhou, Y. Liu, H. Shi, X. Yang, J. Huang, S. Liu, Q. Chen, J. Liu, and K. Wang. Flexible assembly of cascade enzyme on DNA triangle prism nanostructure for controlled biomimetic generation of nitric oxide. *ChemBioChem*, 19:2099–2106, jul 2018.

[171] L. Zhou, A. E. Marras, H.-J. Su, and C. E. Castro. Direct Design of an Energy Landscape with Bistable DNA Origami Mechanisms. *Nano Letters*, 15(3):1815–1821, mar 2015.

[172] H. Zipper, H. Brunner, J. Bernhagen, and F. Vitzthum. Investigations on DNA intercalation and surface binding by SYBR Green I, its structure determination and methodological implications. *Nucleic Acids Research*, 32(12):e103, jul 2004.