Structures and Mechanisms for Synthetic DNA Motors

Natalie Emma Charnell Haley
St. Catherine's College
University of Oxford

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

Trinity Term, 2017

Abstract

DNA provides an ideal substrate for nanoscale construction and programmable dynamic mechanisms. DNA mechanisms can be used to produce DNA motors which do mechanical work, e.g. transportation of a substrate along a track.

I explore a method for control of a DNA mechanism ubiquitous in DNA motor designs, toehold-mediated strand displacement, by which one strand in a duplex can be swapped for another. My method uses a mismatch between a pair of nucleotides in the duplex, which is repaired by displacement. I find that displacement rate can be fine-tuned by adjusting the position of the mismatch in the duplex, enabling the design of complex kinetic behaviours.

A bipedal motor [1,2] is designed to walk along a single-stranded DNA track. Previously the motor has only taken a single step, due to a lack of designs to extend the single-stranded track. I present a novel design for track held under tension using a 3D DNA origami tightrope, and verify its assembly. The bipedal motor design is adapted and a method to specifically place motors on tightropes is demonstrated. Motor operation is investigated on truncated tracks and tightrope tracks by electrophoresis and spectrofluorometry. The motor does not accumulate appreciably at the track end; this is tentatively attributed to rearrangement of the motor between track sites without interaction with fuel.

Tightrope origami can hold single-stranded DNA under pN tension. I use tightropes to study hybridization kinetics under tension and find dramatic, non-monotonic changes in hybridization rate constants and dissociation constants with tension in the range ∼0-15 pN.

Extended tracks for a 'burnt-bridges' motor which destroys its track as it moves [3] are created on the inside of DNA nanotubes, which can be polymerised to create tracks up to a few μm in length, and on tiles which I attempt to join in a specific order. Crossing of the motor between tubes is verified, and microscopy experiments provide some evidence that track is being cleaved by the motor, a requirement for movement along the track. Tile based tracks are imaged by super-resolution DNA PAINT [4], providing proof-of-principle for track observation to infer motor movement.

For Tom.

Acknowledgements

I wish to thank my supervisor, Professor Andrew Turberfield, for believing that I was able to undertake a DPhil in the first place, for his supervision and guidance scientifically, and also for his incredible ability to be (almost) always right about things. Dr Jonathan Bath has acted as a sort of unofficial second supervisor to me, with his vast knowledge of DNA, and general life advice. The group really would not be the same without him.

I was highly fortunate to be able to work with Dr Tom Ouldridge, who has been a brilliant theoretical collaborator on the mismatch project, and Dr Robert Machinek, a former lab member with whom I worked on an earlier mismatch project during my Masters year. Robert taught me much, from how to pipette, to how to use a fluorometer, and our work together inspired the project on mismatch repair in this thesis.

Elsewhere in the lab I had the joy of working with Flo Benn and Alex Lucas on the DNA nanotube motor project. Sam Tusk taught me pretty much everything I know about optical microscopes and has been a great friend throughout my time here. Antonio Garcia Guerra provided thesis-writing companionship over the past months. Thank you to Alex Lucas, Antonio Garcia Guerra, Céline Journot, and Robbie Oppenheimer for proofreading some sections of this thesis, which has undoubtedly improved it. Other lab members: Benham Najafi, Emma Silvester, Ibon Santiago, Joel Spratt, Juan Jin, Le Liang, Seham Helmi, Wenjing Meng, Mike Boemo, Aiman Entwistle, Robert Schreiber and Katie Young, have all helped to sustain and grow the friendly, supportive atmosphere in the group, making it a wonderful place to work.

I am grateful for the fantastic opportunities presented to me as a student at the Doctoral Training Centre. Thanks to Dr Omer Dushek and the Van der Merwe and Omer Dushek groups for welcoming me during my first DTC project, where I learnt much which has

influenced and aided the projects in this thesis. Thank you to the friends that I have made at the DTC (and constructed many a rushed PowerPoint presentation alongside) in particular: Eleanor Law, Ali Cigari, Farah Colchester, Ferdinando Randisi, Gareth Langley, Joe Donlan, Joe Page and Ross Johnstone. As is traditional, I have hidden something for you.

Thanks also to Professor Fritz Simmel, and his group in TUM, for kindly hosting me for a secondment, letting me use their lab, and showing me the wonderful city of Munich.

The biophysics labs really do fall apart without the assistance of first, Sarah Matthews, and later, Sanna Pippo-Henderson. Thanks to both of them for their help keeping everything running smoothly. Thank you also to Alan Francis and Dan Franks in stores, for dealing with my many envelopes of DNA and other eclectic science stuff.

I would also like to thank the Engineering an Physical Sciences Research Council, St. Catherine's College and the Worshipful Company of Leathersellers for financial support in the form of scholarships.

Thank you to my lovely parents, for encouragement throughout my life and for nurturing my love of science. Last, but certainly not least, I wish to thank my partner Tom Hocking. His love and support over the last years have been invaluable to me, and it is to him that I dedicate this thesis.

Contents

List of Figures

List of Tables

Glossary

A Adenine.

Abasic Without nitrogenous base.

AFM Atomic force microscopy.

AGE Agarose gel electrophoresis.

Alexa Fluorophore family by Molecular Probes Inc.

APE-1 Apurinic/apyrimidinic endonuclease, specifically cuts to 5' end of abasic nucleotide in a duplex.

ATP Adenosine triphosphate: Biological energy store.

β-gal Lactase enzyme.

Biotin Vitamin which binds to streptavidin.

bp base pair.

C Cytosine.

caDNAno Computer aided DNA origami design software.

cando DNA origami structure prediction software.

CCD Charge coupled device: Used to convert signal electrons in the camera to a digital signal

Crossover DNA backbone link between a pair of double helices.

Cy Cyanine flurophore.

DNA Deoxyribonucleic acid.

DNAzyme DNA enzyme.

dsDNA Double stranded DNA.

DX Double crossover.

EDTA Ethylenediaminetetraacetic acid: Hexadentate ligand able to chelate metal ions such as magnesium.

EM Electron multiplying

EMCCD Electron multiplying CCD: Produces multiple secondary electrons from each signal electron

FAM Carboxyfluorescein fluorophore.

FJC Freely jointed chain.

FOV Field of view.

FRET Förster resonance energy transfer.

FWHM Full width half maximum.

G Guanine.

HCR Hairpin chain reaction.

HF High fidelity.

HindIII Restriction endonuclease.

IEL Intuitive energy landscape.

Iowa Black Dark quenchers by Integrated DNA Technologies.

IPTG Isopropyl β-D-1-thiogalactopyranoside, lacZ inducer.

lacZ Gene encoding β-gal.

λ phage Bacteriophage with ∼50 kbp double stranded DNA genome.

LB Lysogeny broth (rich medium).

LHS Left hand side.

LNA Locked nucleic acid: Conformationally restricted synthetic ribonucleotides.

LoG Laplacian of Gaussian.

LUT Look up table.

M13 Bacteriophage with single stranded DNA genome.

M9 Minimal medium.

MCS Multiple cloning site.

MFE Minimum free energy.

mFJC Modified freely jointed chain model.

Mismatch Any base pair other than A-T or C-G.

mRNA Messenger RNA.

MWCO Molecular weight cut-off.

nt Nucleotide.

Nt.Bbv CI Nicking enzyme with recognition sequence `CCTCAGC`.

NUPACK Nucleic acid package: DNA thermodynamics software.

ODE Ordinary differential equation.

oxDNA Coarse grained model of DNA.

PAGE Polyacrylamide gel electrophoresis.

PAINT Points acquisition for imaging nanoscale topography: Super-resolution localisation technique.

PCR Polymerase chain reaction.

PEG Polyethylene glycol.

pfu Plaque forming units (measure of phage).

PMT Photon multiplyer tube

poly-T Polymer of thymine nucleotides.

primer Short length of DNA complementary to one end of a DNA target for PCR.

RF Replicative form.

RHS Right hand side.

RMS Route mean squared.

RNA Ribonucleic acid.

RNAse H RNAse enzyme specifically cuts RNA/DNA hybrids.

RNAse HII Thermostable RNAse enzyme specifically cuts RNA-DNA chimera when hybridized to DNA.

ROI Region of interest

S-300 Sephacryl size exclusion chromatography resin with 300kDa MWCO.

SDS Sodium dodecyl sulphate: Anionic detergent used to denature protein.

SNP Single nucleotide polymorphism.

SphI Restriction endonuclease.

ssDNA Single stranded DNA.

Stator Stationary part of motor system.

Streptavidin Tetrameric protein with a high affinity for biotin.

Sybr Gold Nucleic acid stain.

Sybr Green II Nucleic acid stain with higher relative affintity for ssDNA and RNA than Sybr Gold.

T Thymine.

TAE Tris acetate EDTA buffer.

TE Tris EDTA buffer.

TEM Transmission electron microscopy.

TET Tetrachlorofluorescein fluorophore.

TIRFM Total internal reflection fluorescence microscopy.

TMSD Toehold-mediated strand-displacement.

U Units (measure of enzyme activity).

U Uracil.

UV Ultra violet: Electomagnetic radiation with wavelength 10–400 nm.

X-gal Substrate for blue-white screen for lactase activity.

Chapter 1

Introduction

I present work on synthetic DNA motors, powered mobile devices constructed from DNA. I investigate both dynamic DNA processes which are involved in mechanisms of DNA motors, and DNA nanostructures for use as tracks for DNA motors. In this chapter I briefly present some of the fundamental features of DNA which make it such a useful substrate for nanotechnology, and some early examples of DNA nanostructures. I then discuss dynamic processes involving DNA, especially the fundamental mechanisms of hybridization and strand displacement as these will also be the focus of chapters 2 and 4. I explore how dynamic DNA mechanisms have been applied by nanotechnologists to create dynamic DNA devices, in particular DNA motors. Next I will jump an order of magnitude up the length scale and into the world of larger DNA structures formed by the DNA origami method, focusing in particular on how DNA origami structures have been used in relation to DNA motors, and on DNA origami as a tool to stretch single-stranded DNA, both of which are themes to be studied in chapters 3, 5 and 6. Finally I will set the scene for the remainder of the thesis, giving a brief précis of what is to be expected in each of the coming chapters.

1.1 DNA for Nanotechnology

The pioneering X-ray crystallographic work of Rosalind Frankin and Maurice Wilkins [5, 6], alongside theoretical work of Francis Crick and James Watson [7], resulted in the discovery of the 3D structure of DNA. Structure is intimately related to function. A core principle of structural biology is that we can learn more about the function of something, for example a protein, if we know its shape and structure, as morphology will prescribe what the

molecule is capable of, e.g. flexible regions can act as hinges and allow constrained motion of one part about another, or a specifically shaped pocket might act as an active site for a protein catalyst or enzyme. When Crick and Watson proposed the double helical form of a pair of complementary strands of DNA, they hypothesised that the structure of DNA suggests a function in DNA replication [7, 8], which was indeed later confirmed. Knowledge of structure can not only be used to infer existing functions, it can also be used to engineer new functions, made possible by the structure. DNA nanotechnology, does exactly that, uses the now well known structure of DNA, as a substrate for nanoscale fabrication from which all manner of structures and devices have already been constructed.

1.1.1 DNA as a building material

DNA is a hetero-polymer formed of sub-units called nucleotides (nt). A DNA nucleotide has a sugar with a phosphate group attached to the 5' carbon atom of the deoxyribose sugar, and a nitrogenous base group attached to the 1' carbon atom. In a DNA strand, the 5' phosphate of one nucleotide is covalently attached to the 3' carbon of the next nucleotide. The most common form of DNA is B-form DNA where the two anti-parallel complementary strands composing a DNA duplex wrap around each other in a double helix, with the bases in the core of the helix. Each base is held in a planar arrangement with its complementary base partner by hydrogen-bonding interactions. These planar, aromatic, DNA bases also stack on each other forming π-π stacks due to delocalised electrons in the π orbitals [9]. The combination of π-π stacking, and hydrogen-bonding creates the overall stabilising effect of base pairing of consecutive nucleotides [10]. The deoxyribose sugar-phosphate backbone linking nucleotides along each strand of DNA is on the outside of the helix. Base-backbone vectors of nucleotides in a base pair are not parallel, this creates major and minor 'grooves' in the double helix. The helical repeat of DNA is ∼10.4-10.6 base pairs, indicating a rise per base pair of approximately 34° [11, 12]. This allows the choice of angle at which two parts of a structure join, by the number of base pairs along a linking helix.

A base pair (bp) is a pair of two complementary bases, one purine and one pyrimidine. There are two native 'Watson-Crick' base pairs in DNA, Adenine (A) and Thymine (T), and

Guanine (G) and Cytosine (C), which form either two, or three inter-base hydrogen bonds respectively. Runs of consecutive base pairs form regions of interaction called domains, where domain A will stick to its complementary domain A', and the DNA sequence of A' is the reverse complement of A. Many orthogonal domains can be used to create specific contacts between different parts of a DNA structure.

The DNA duplex is a rigid building block on short length scales. The rigidity and tensile strength of helices on length scales much shorter than the persistence length (50 nm [13–15]) allows the precise positioning of structural elements on this length scale. Flexiblilty of certain areas of a DNA structure can be increased by adding single-stranded domains which have a much shorter persistence length ~ 0.75 nm [16, 17].

1.1.2 Holliday junctions and DX tiles

Holliday junctions are structures with four duplex arms formed from DNA first proposed by Robin Holliday as intermediates in the process of homologous recombination [18,19]. They are a key motif in DNA nanotechnology for linking multiple DNA strands. The structure of a Holliday junction is highly salt-dependent [20]. At the central point at which the four arms join, the central nucleo-bases can be either stacked or unstacked. Divalent cations such as magnesium can stabilise the stacked conformations.

Nadrian Seeman, who would go on to found the field of DNA nanotechnology, constructed Holliday junctions from a set of 4 synthetic oligonucleotides [21]. Seeman later used Holliday junction structures to create DNA double crossover tiles (denoted DX tiles), units with pairs of double crossovers [22] (figure 1.1a) which can be linked to form arrays [23].

1.1.3 3D structures

DNA nanostructures are not restricted to a plane, three dimensional structures were first hypothesised in 1991 with Ned Seeman's 'cube' where sequential hybridization of complementary domains and ligation of nicks created a nano-object with the connectivity (edges and vertices) of a cube [24]. It was not possible to demonstrate that a three dimensional,

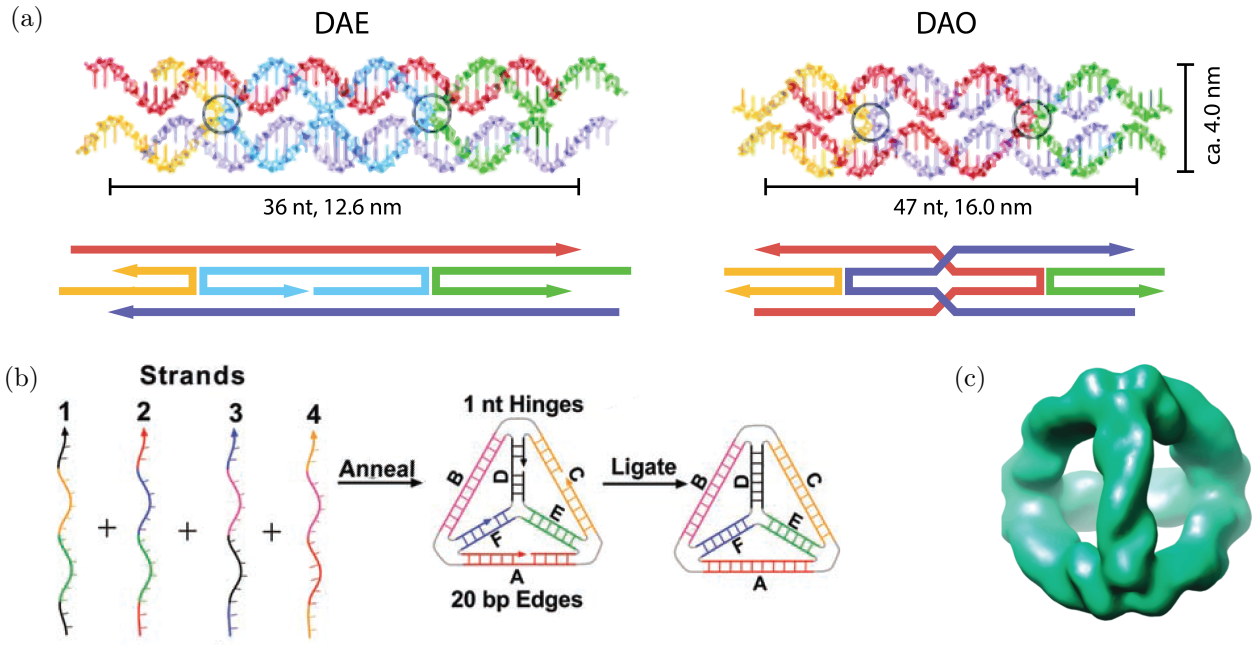

Figure 1.1: DNA nanostructures from branched junctions. In DNA line diagrams, 3' ends of strands are indicated by terminal arrowheads. **a** Helical representation and line schematic of DX tiles. Left: Antiparallel DX molecule with even number of half-turns between crossovers. Right: Antiparallel DX molecule with odd number of half turns between crossovers. Holliday-junction-like crossovers are indicated by circles. Each component strand has a unique colour. **b** Scheme for the construction of a DNA tetrahedron from 4 strands. Colour in this schematic indicates an individual domain and its complement. **c** Cryo electron microscopy reconstruction of DNA tetrahedron. Figure **a** adapted by permission from Macmillan Publishers Ltd: Nature, reference 23, copyright (1998). Figure **b** from reference 25. Reprinted with permission from AAAS. Figure **c** reprinted with permission from reference 26 Copyright (2009) American Chemical Society.

cubic structure had been produced as the vertices of the cube were flexible.

The successful design and synthesis of a similar wire-frame DNA nanostructure, a tetrahedron, was later demonstrated from the one pot annealing, and subsequent ligation of 4 strands of DNA [25] (figure 1.1b). Annealing, in the context of this thesis, is where DNA strands are heated to high temperature to break any existing base pairing interactions within or between the strands which might interfere with the folding, and then cooled quasi-statically to reach their free energy minimum, which, in most cases, will be the state which maximises the number of base pairs formed. By the nature of the design, the tetrahedron, if formed, must project into a third dimension. The three-dimensional structure was confirmed by atomic force microscopy (AFM) and by cryo-electron microscopy [26] (figure 1.1c).

1.2 Dynamic DNA

DNA has been demonstrated to be capable of producing two and three-dimensional nanostructures, however these structures described so far are static and immobile post assembly. DNA can also be used to produce systems that evolve with time. I will present here some of key processes which are used in generating such time evolving DNA systems and then examine their use in a selection of published examples.

1.2.1 Hybridization

Hybridization is the coming together and binding of complementary, single-stranded domains of DNA to form runs of base pairs. This process is used in the self-assembly of nanostructures such as those detailed above.

Hybridization is temperature dependent. Above the melting temperature, fewer than 50% of native base pairs are formed. Melting temperature is dependent upon complementary domain length, salt type, and concentration. The thermodynamics of hybridization are of great importance for the design of static nanostructures. A widely used model for DNA thermodynamics is the nearest-neighbour model of DNA base-pairing. This model assumes that interactions between a base pair and the neighbouring pair of base pairs are sufficient to describe binding thermodynamics. The inclusion of nearest-neighbours, instead of considering only the base pair in isolation, is required because base-stacking is a significant contributor to the free energy, and varies between one pair of base pairs and another [27]. John SantaLucia created a unified set of thermodynamic parameters [28] from data collected by himself and others [29–35], giving nearest-neighbour free energy contributions from each of the 10 possible nearest-neighbour pairs, and additional parameters for helix initiation [34], including penalties for terminal A-T base pairs, and symmetry correction for self-complementary sequences. The set of parameters was later expanded to include terms for the 24 possible internal single-base mismatches, helix terminal mismatches, coaxial-stacking, dangles, loops and bulges [36]. The kinetics of hybridization have also been studied, and will be discussed in chapter 4, where I also investigate hybridization to my own DNA structure.

Hybridization kinetics have been harnessed in the single molecule localisation microscopy technique DNA PAINT (points acquisition for imaging nanoscale topography) [4, 37]. By using a fluorescently labelled DNA oligonucleotide of length 8-10 nt with melting temperature around room temperature, each oligomer will bind to a complementary target transiently, and, once bound, can be localised by the fitting of a 2D Gaussian approximating the point spread function to the diffraction-limited spot produced. Bulk concentration can be altered to fine-tune hybridization kinetics. DNA PAINT has inherent resistance to photobleaching, the process through which fluorophores are damaged by light to produce a permanently dark state, because though an individual fluorophore might bleach during the time it is bound and within the illuminated region, the oligonucleotide with the bleached fluorophore will dissociate on observable timescales and be replenished from bulk solution where the fluorophores are not illuminated and unbleached. PAINT relies upon dynamics to achieve 'blinking' of the fluorophores and therefore allow super-resolution localisation.

1.2.2 Branch migration

Branch migration is a dynamic process, whereby a junction between multiple DNA strands can move by inter-strand exchange of base pairs. Early studies of branch migration were in natural Holliday junctions. The Holliday junctions, which form during homologous recombination, bear high levels of sequence similarity: this allows them to be motile, and swap base pairs from one strand to another (figure 1.2a) [41].

1.2.3 Toehold-mediated strand displacement

Initiating a branch migration process can be slow. Take, for example, a duplex and an invading single strand of DNA which has an identical sequence to one of the strands of the duplex, and therefore, in principle, should be able to form a duplex with the complementary invading strand in place of the strand already bound. Such a system is depicted in figure 1.2b. These processes are commonly referred to as blunt-ended strand displacement. Reynaldo and colleagues performed kinetic investigations into blunt-ended strand displacement [42]. They considered two pathways by which one strand can replace an identical strand in a duplex

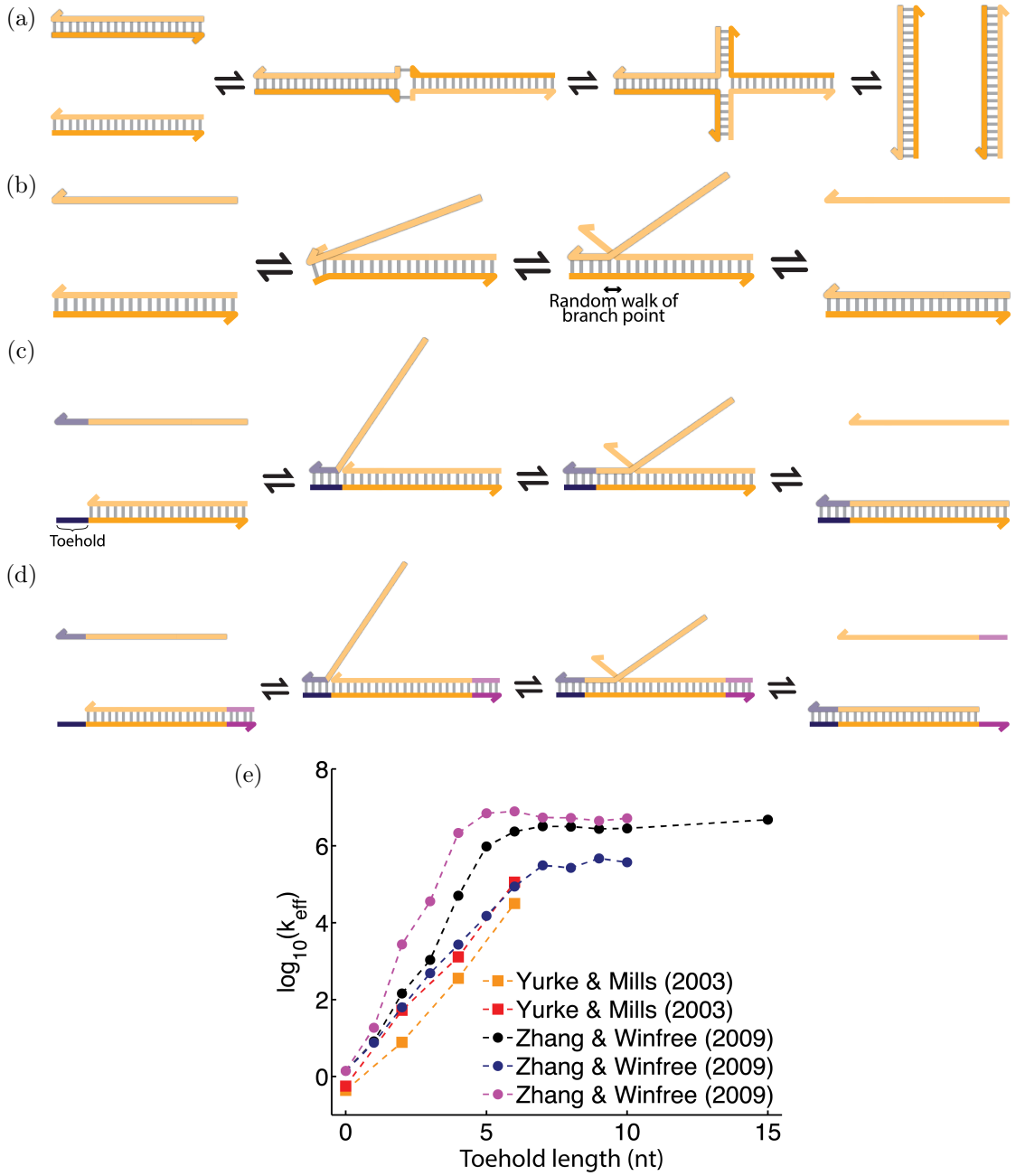

Figure 1.2: Strand displacement mechanisms. **a** Holliday junction branch migration. Two homologous duplexes can associate and swap base pairs through a series of Holliday junction intermediates. **b** Blunt strand displacement. A duplex and a single strand identical to one of the duplex strands can associate and form a three-stranded branched junction intermediate which may lead to dissociation of the incumbent strand from the duplex. **c** Toehold-mediated strand displacement. Similar to blunt strand displacement but the association step is aided by the single-stranded toehold domain overhang on the duplex which allows the invader strand to bind. Displacement is also thermodynamically driven due to the additional base pairs which can be formed in a duplex with the invader. **d** Toehold exchange. Toehold-mediated strand displacement with toeholds on either side to allow the toehold-mediated reverse reaction. **e** Plot showing the dependence of effective second-order rate constant for toehold-mediated strand displacement on toehold length using data from reference 38 and 39. Figure **e** reproduced from reference 40, Nucleic Acids Research (2013), by permission of Oxford University Press.

with a common complement. The first pathway involves complete dissociation of one of the strands from the duplex and the subsequent hybridization of the invading strand, and the second pathway replaces the original strand with the invading strand by branch migration. They measured displacement kinetics using a gel-shift timecourse, and found the dominant pathway to be the dissociative pathway, close to the melting temperature of the duplex, but the branch migration pathway at lower temperatures.

To initiate branch migration, the duplex and the invading single strand must first associate. This requires the spontaneous melting of bases in the duplex before association can occur as, whilst the duplex is complete, it is unreactive. To overcome the association barrier, a toehold can be added to a duplex. The toehold is a domain which remains single-stranded and acts as a binding site for the invading strand, from which it can progress into branch migration (see figure 1.2c). This process is called toehold-mediated strand displacement or TMSD. Toeholds have been demonstrated to increase the rate of strand displacement processes compared to blunt-ended strand displacement [38]. The displacement rate constant for TMSD has been shown to increase exponentially with increasing toehold length up to a saturating toehold length of 7 nt (see figure 1.2e) [39, 40]. I will elaborate on methods used to influence TMSD kinetics in chapter 2 where another method for control of displacement kinetics is presented.

1.2.4 Toehold exchange

Toehold exchange is an extension of toehold-mediated strand displacement where the invading strand binds at a toehold on one end of the duplex but does not bear the nucleotides required to base pair fully at the opposite end of the duplex. Consequently, once the invading strand has displaced the victim strand initially present by spontaneous melting of the last few base pairs, it has created a new toehold wherby the victim, or another invading strand, is able to rebind and displace the invading strand (figure 1.2d). The toehold exchange method was used in the construction of 'see-saw' gates by Lulu Qian and colleagues [43], which have been demonstrated in complex DNA cascades capable of performing 4 bit square root calculations [44] and simple 'neural network' computation [43].

1.2.5 Hairpins

Displacement processes can proceed intramolecularly. A duplex where both domains are linked at one end forms a motif called a hairpin. The duplex forms the 'stem' of the hairpin, and additional linking ssDNA is termed the 'loop'. TMSD can be used to open hairpins either via a single-stranded toehold on the end of the stem, opposite the loop, or by toehold domains in the loop of the hairpin itself, with the latter occurring at a lower rate than the former for similar toeholds [45]. Hairpins can be harnessed as kinetic traps, which sequester domains, preventing reactions until the directed opening of the hairpin has occurred.

Strand displacement by a catalyst strand can be used to trigger a reaction between two metastable hairpins [46], which can be used for linear signal amplification [47], or detection of specific sequences [48]. A similar design can be used to trigger a polymerisation chain reaction of metastable hairpins, where an input strand opens a hairpin, which then opens another, and so on, forming linear [49] or branched [50] hairpin polymers. Such reactions, know as hairpin chain reactions or HCR, have recently been used to create a 'ribosome-like' ordered chemical assembly line [51], and to amplify signals from fluorescence *in situ* hybridization (FISH) probes for microscopy of nucleic acids such as expressed mRNAs in organisms [52].

1.3 DNA Machines

Using the techniques detailed above, DNA based machines and motors can be constructed which produce a mechanical output in response to a stimulus [53, 54]. The earliest example was DNA tweezers, where a strand was cycled between opened and closed states by addition of closing strands which hybridize to tweezer arms, and opening strands which strip away the closing strands by TMSD [53].

1.3.1 Clocked DNA motors

DNA motors are a class of dynamic DNA machine which couple some form of energy input to motion, often along a track. The earliest examples of DNA motors were from the

groups of Niles Pierce and Nadrian Seeman in 2004. Both developed bipedal DNA walkers with similar locomotive principles. The motors had two single-stranded feet, linked to each other either by a duplex in the case of the Shin-Pierce walker [55] (see figure 1.3a), or by a triple crossover molecule in the Sherman-Seeman walker [56]. The tracks were some form of rigid structure with periodic single-stranded protrusions, this was a duplex with ssDNA branches for Shin and Pierce, and a DX tile array for Sherman and Seeman. The motor was held in place at a specified position on the track by coupling strands, one for each foot, which bore the complement to a motor foot, and a ssDNA track site. The coupling strands had a further domain, which remained single-stranded when holding the motor in place, which could act as a toehold for the removal of the coupling strand from the foot and track to release the foot. A released foot could then be coupled to the next track site along by a new coupling strand. This method requires the sequential addition of coupling strands and removal strands which leads to the accumulation of waste products with the potential to stall the motor [57]. The energy input for this type of mechanism comes from the free energy of hybridization [38, 53] of various domains during the motor operation. After each strand addition more base pairs are produced in the system than were previously present.

1.3.2 Autonomous devices

These early motors were unable to make progress without the clocked addition of reagents which allow them to step. One desirable property of a motor is for it to be able to make progress autonomously without the requirement for timed addition of control strands, or other external perturbation. Its locomotion must therefore be conditional upon the previous step having been made for the next step to be initiated and so on. The first example of an autonomous motor was the work of Peng Yin and colleagues who developed a system to transport six nucleotides along a track by cycles of restriction with alternating endonucleases followed by ligation [58] (see figure 1.3b). Asymmetric cleavage by the endonuclease pair ensures directionality of transport. The energy source for this transporter comes from the free energy of hydrolysis of ATP by the ligase as new covalent bonds are formed [59]. This motor has strict limitations on the cargo it is able to transport. The cargo length and

sequence is prescribed by the restriction endonucleases used. The nucleotides of the cargo are not fully-connected, as they are three nucleotides each on each strand of the track duplex and are staggered (not complementary pairs).

1.3.3 Biased tracks

Another approach to producing a free-running motor which does not require clocked addition of strands uses a energy gradient along the track sites. The examples described so far have track binding sites with approximately isoenergetic binding sites, therefore there is not a thermodynamic bias towards binding at one site over another. By controlling the track energy landscape, the direction of motion, towards lower free energy, can be achieved. One example of such a motor is a rolling motor [60] where a ligated single-stranded circle of DNA rolls along a single-stranded track where binding sites increase in LNA content and length. LNA are locked nucleic acids and produce base pairs lower in free energy than the DNA equivalents [61]. The thermodynamic gradient causes the roller to move towards the track end when brake oligonucleotides blocking track sites are removed.

1.3.4 Burnt-bridges motors

In 2005 two groups simultaneously developed their own variants of a single-legged DNA motor which performs a series of toehold-mediated strand displacement reactions to move across a series of single-stranded stators [3, 62] (see figure 1.3c). The mechanism relies upon the creation of a free toehold domain on the 'motor' strand by cleaving away the stator strand close to the top if and only if the motor strand is bound, such that the top stator fragment is able to thermally dissociate. The exposed motor top is now a toehold, allowing the cut stator to be displaced from the motor strand by the next stator along by TMSD, moving the motor on to this stator and along the track. This mechanism is named burnt-bridges, as the process by which the motor moves destroys the track behind it, biasing the motor to step forward rather than diffuse backwards where fewer base pairs can be formed. The main difference between the designs of Bath et. al. and Tian et. al. was the method of stator-top cleavage. Tian et. al. used a DNAzyme motor strand [62], able to cleave DNA/RNA chimera stators,

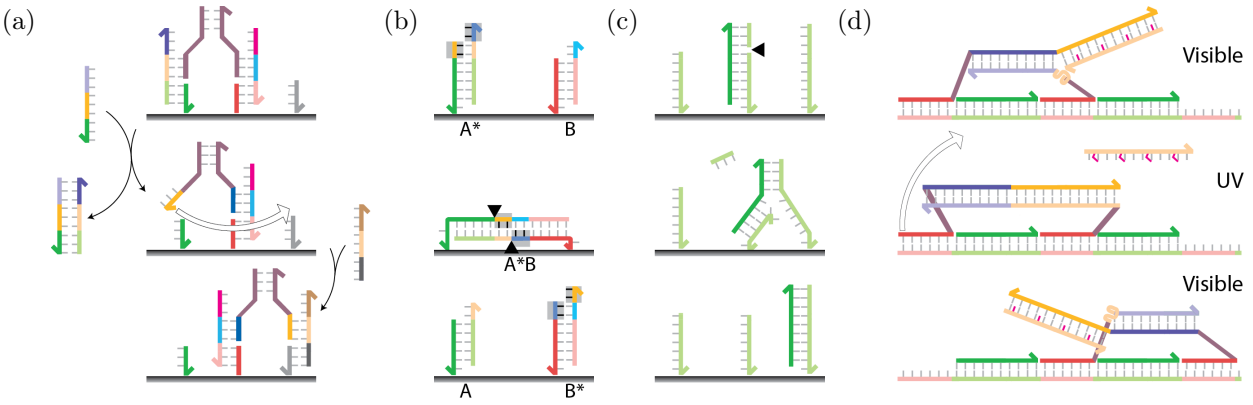

Figure 1.3: Example DNA motors. **a** Shin-Pierce design for clocked motor [55]. **b** Autonomous nucleotide transporter of Yin and colleagues [58]. **c** Autonomous burnt-bridges motor of Bath and colleagues [3]. **d** Azobenzene-ratcheted bipedal motor of Yeo and colleagues [64]. **a-c** Reprinted by permission from Macmillan Publishers Ltd: Nature Nanotechnology reference 65, copyright (2007).

wheras Bath et. al. opted for an external nicking enzyme Nt.BbvCI [3], which cuts a single strand of a duplex at a particular recognition sequence. Another form of the burnt-bridges motor uses pyrene-catalysed photolysis of disulphide-modified stator backbone linkages [63].

1.3.5 Ratchets

Ratchet-like motors have been developed using the reversible cis-trans isomerisation of azobenzene under different wavelengths of illumination [64, 66, 67]. This conformational change in the azobenzene is able to destabilise existing duplexes, or prevent new duplex formation. Resulting changes to the free energy landscape upon isomerisation create new energy minima upon switching illumination and can be used to direct motion.

One example mechanism for an azobenzene ratchet motor extends the motor body linking the two feet with azobenzene isomerisation [64] (see figure 1.3d). The motor has two identical feet joined by a body duplex. The body has adjustable length depending on the azobenzene isomerisation state. Feet bind to identical ssDNA segments along a single strand of DNA with alternating ssDNA foot binding sites and duplex spacer domains. Extending the body preferentially breaks the duplex between the back foot and track as this is in peeling configuration, and the front foot is in shearing configuration. Peeling apart bp occurs at lower force than shearing apart bp [68–73]. Cycling illumination allows the raised foot to rebind where it came from or the next site along, making eventual forward progress.

1.3.6 Autonomous bipedal motor

A bipedal autonomous DNA motor was developed by Green and colleagues [1], the motor was composed of identical feet linked by a duplex, and walked on a single strand of DNA with repeating foot binding domains. This motor will be presented in greater detail in chapter 5, where it is further developed and studied. The mechanism relies upon asymmetric lifting of feet, due to competition between feet making it easier to lift a particular foot with fuel. This motor has been demonstrated to lift feet with both DNA haripin-based fuels and anti-fuels [1], and enzyme-cleaved fuels [2]. Unlike burnt-bridges designs, this motor does not destroy its track as it goes, and is reversible by design of the fuel added. To date it has not been able to make multiple steps due to limitations in producing long tracks of ssDNA. A method for producing an extended track for this motor is also explored in chapter 3.

1.4 DNA Origami

Just over ten years ago, Paul Rothemund presented a new technique for the design of DNA nanostructures [74]. Insipired by William Shih's octahedron formed from a folded single strand of DNA [75], and Nadrian Seeman's double crossover tiles [22], he postulated that a single strand of DNA, with a non-synthetic sequence, could be folded into a designed shape with the aid of many short strands of DNA to hold it together. These short strands are called staples, and pin multiple domains of the long single-stranded DNA, known as the scaffold or template, together, creating multiple double helices linked periodically by Holliday-junction-like crossovers. Rothemund named the technique DNA origami, after its Japanese paper-art namesake, which proved a fitting name for the method as both enable the formation of a plethora of shapes by folding a substrate (paper or single-stranded DNA) in different ways. The DNA origami technique has since been widely adopted, and extended to produce three-dimensional [76] and curved structures [77] of various shapes. Computational tools such as caDNAno [78] and cando [79, 80] have also been developed to help researchers design their own origami. RNA origami has also been produced from single transcripts of RNA [81]. Such structures are able to fold on themselves *in vitro* or co-transcriptionally

without the need for staples. In a move from the standard origami scaffold routing, which sends the scaffold up and down the structure in antiparallel sheet-like pattern, wireframe scaffold routing enables a new class of structures to be produced [82–84]. By extending staple sequences that protrude from the surface of a DNA origami structure, it can be functionalised with other DNA or DNA-coupled units of choice.

DNA origami is usually produced by annealing scaffold DNA with a 2 to 20 × excess of each staple strand. Folding with 1:1 stoichiometry of scaffold to staples has been shown to produce poorly folded structures [74]. Annealing protocols vary from structure to structure but temperature ramps used are typically steeper for single-layer or pseudo-2D structures [74] than for 3D origami [76], although origami can be folded at constant temperature [85]. If desired, origami can subsequently be purified away from the assembly excess of staples for example by polyethylene glycol precipitation [86], size exclusion chromatography [87] or excision of bands from agarose gel (see appendix A). This is particularly important if some of the staples have added functional domains.

The maximum size of a DNA origami structure is determined by its scaffold length. For a space-filling origami structure (as opposed to wireframe origami) with a ∼7 kbp M13 scaffold this leads to typical dimensions of smaller than 100 nm for a single origami object. Larger origami have been produced using longer scaffolds such as ∼50 kbp λ DNA [88], or λ phage-M13 hybrid DNA as a scaffold [89]. An alternative method for producing larger DNA origami constructs is to create structures from multiple DNA origami each with their own scaffold, joined together. There are two main interactions through which such origami polymerisation has been attempted, through base pairing interactions [90–92] and through coaxial stacking interactions [91, 93–96], or some combination of the two [97–99]. A hybridization-based approach is considered in chapter 6 as a method to retro-fit existing origami with the ability to form polymers of two or more origami types.

1.4.1 Transport on origami

All of the motors described so far have been limited to tracks constructed from short DNA oligonucleotides, often forming a single duplex with branches [3,55,58,62,63,66,67,100–103], or even just a single strand of DNA [1, 2, 60, 64]. Such methods for track construction introduce limits on how far the motor might be able to travel due to the persistence length of the track. Mechanical rigidity is also important for certain motor mechanisms [67,103,104]. Some motors have utilised multi-stranded DNA tiles [56, 105] which present a stiffer track substrate. DNA origami is an ideal design scheme for the development of motor tracks. Even simple DNA origami tiles [74], produce a nanoscale breadboard with evenly spaced addressable points spaced by ∼6 nm on each side of the tile.

Origami-based motor tracks appeared for clocked [92, 106, 107] and autonomous burnt-bridges motors [87, 108, 109]. Wickham and colleagues added the burnt-bridges motor of reference 3 to a DNA origami tile enabling the direct observation of motor stepping using high speed atomic force microscopy [87]. The burnt-bridges motor was also sent down a branched track with 4 end stations, and could be identified reaching each station by bulk fluorometry [108]. The end station reached could be influenced by external control strands or by control sequences appended to the motor. Lund and colleagues created a 'DNA spider', a streptavidin molecule with three DNAzyme legs, able to move along a track on DNA origami by a burnt-bridges mechanism [109]. A fourth leg was used to hold the spider at a start site.

Lieber and colleagues demonstrated that a Shin-Pierce design of clocked motor was able to move repeatedly between a pair of linked tiles. This was the first demonstration of a motor being able to bridge between two origami [92].

Taking a different approach, in order to achieve more rapid transport across origami, a tethered arm was positioned on an origami tile by Kopperger and colleagues [110]. By constraining the diffusion of the arm end to a hemispherical shell or volume (depending on the arm geometry), super-diffusive transport of a cargo from one station to another station on the opposite side of the tile could be achieved. This process was faster[1] than the reported

[1]The authors find a lower-bound velocity of 2.5 nm s^{-1}, but true velocity is likely to be much higher.

rates of burnt-bridges type motors of $\sim$1 nm s^{-1} [3].

1.4.2 Long distance DNA motor motion off origami

Transport of a nanoparticle cargo along DNA stators adsorbed onto a carbon nanotube have also been demonstrated [111]. The particles were observed to move distances of $\sim$1 μm by two channel (IR and visible) light microscopy observation. Motors which move longer distances are easier to observe with light microscopy if the distance travelled exceeds the diffraction limit for the light used for observation. A burnt-bridges-inspired rolling spherical nanoparticle was produced [112], where the nanoparticle was coated with single-stranded DNA anchors, and would bind multiply to a surface coated with complementary single-stranded RNA. RNAse H, an enzyme which will hydrolyse RNA only when it is part of an RNA-DNA duplex, is able to cut the duplexes between the sphere and the surface causing the sphere to roll along in a 2D random walk on a uniformly coated surface.

1.4.3 Stressed DNA origami structures

DNA has been shown to fold against tension in single-stranded regions of its scaffold. Liedl and colleagues produced three dimensional 'tensegrity'[2] structures, such as kites and prisms, from regions of multi-helix origami bundles, linked by single-stranded scaffold sections [113]. The global 3D shape requires the opposing tensions of single-stranded links, otherwise the bundles would not be held in place. A later example where tension in ssDNA was used to bend an origami hinge structure was demonstrated by Castro and colleagues [114,115]. The structure had rigid arms made from multi-helix bundle DNA origami, linked by a compliant row of single layer origami which could be bent by multiple ssDNA links opposite. Changing the length of the ssDNA links changes the equilibrium distribution of hinge angles. Helix bundles can be used as springs [116]: Using the insertion and deletion of base pairs to introduce curvature [77], two helix bundles were wound into a coiled spring which was used to observe tethered myosin VI move along actin and change stepping mode from hand-over-hand to shorter inchworm steps in high force regimes.

[2] A portmanteau of 'tensional' and 'integrity'.

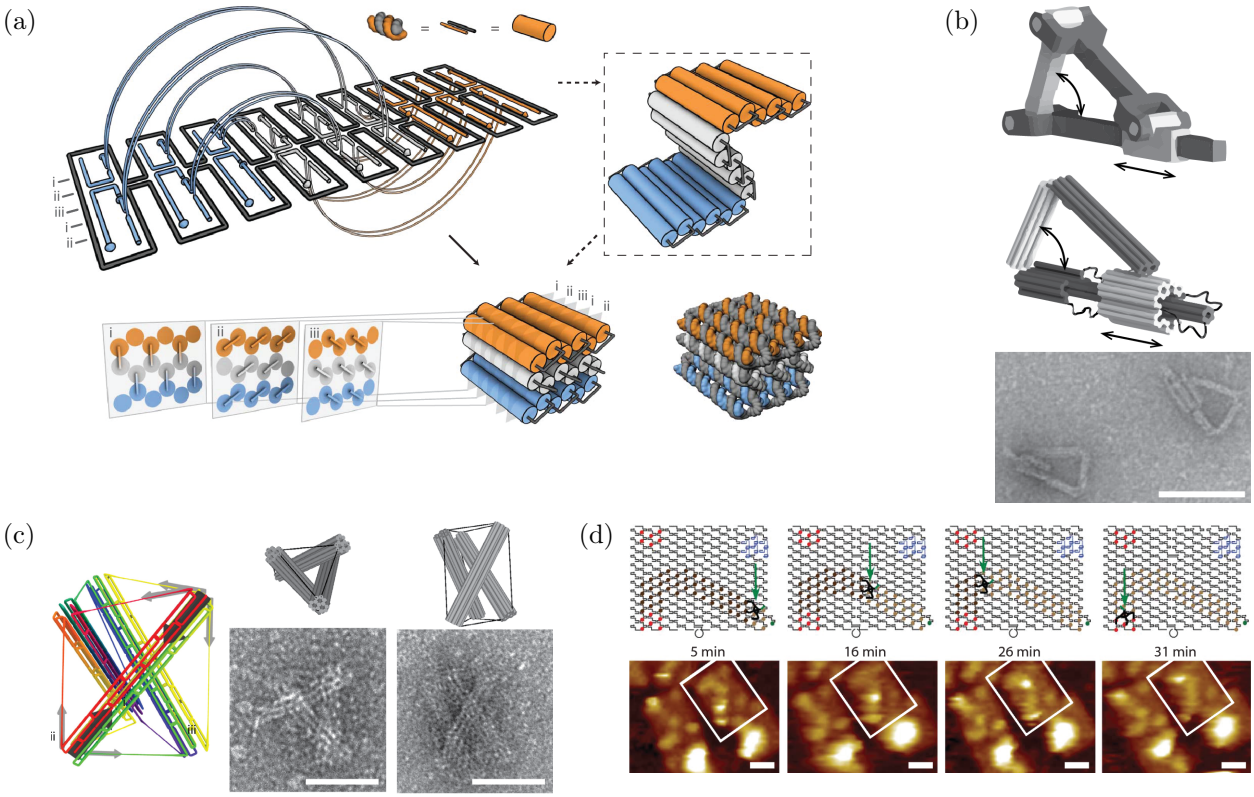

Figure 1.4: DNA origami structures. **a** Scheme for three-dimensional DNA origami design [76]. **b** Crank slider from rigid DNA origami with flexible elements [121]. **c** Tensegrity prism formed of helix bundles held in place by spring-like ssDNA [113]. **d** Burnt-bridges spider moving on DNA origami tile track [109]. **a** Reprinted by permission from Macmillan Publishers Ltd: Nature reference 76, copyright (2009). **b** Adapted from reference 121. **c** Reprinted by permission from Macmillan Publishers Ltd: Nature Nanotechnology reference 113, copyright (2010). **d** Reprinted by permission from Macmillan Publishers Ltd: Nature reference 109, copyright (2010). All scale bars 50 nm.

1.4.4 Towards whole-origami motors

What if origami were not just part of a track for a DNA motor, but formed the motor itself? A number of recent advancements have been made in this direction, for example hinges [114,117], rotaxanes [118,119], and controllable rotors [120] allow dynamic assemblies of one or more origamis [93, 121]. Constructions such as the DNA origami crank slider of Marras and colleagues where a slider rotaxane is coupled to a flexible hinge, can be used to convert linear motions into rotational motions [121]. Dietz and colleagues demonstrated that cross-like origami could form stacked closed conformations or unstacked open formations, acting as macromolecular Holliday junctions, controllable by external modulation of temperature or salt concentration [93]. These designs pave the way for clocked, or even autonomous motors, made from multiple DNA origami.

1.5 Thesis Aims

This thesis aims to develop structures and mechanisms for synthetic DNA motors, ultimately to enable observation of the autonomous bipedal DNA motor of Bath et al [2] moving along an extended track (chapter 5). To design DNA motors, knowledge of the mechanisms which dictate their motion is required. The processes of DNA hybridization, and DNA strand displacement introduced in the chapter are used ubiquitously in DNA motors as we have seen so far. I aim to understand how key design elements in the bipedal motor design (mismatches between the motor foot and track, and tension in the track) can affect the kinetics of DNA strand displacent or hybridization (chapters 2 and 4).

Novel structures are required to extend the range of DNA motors. We have seen in this chapter the current state-of-the-art DNA motors are limited in their range by factors such as the genome size of M13mp18 used as origami scaffolds in the case of DNA burnt bridges motors [87, 108], or by a lack of current designs for single stranded DNA tracks which hold the track under tension in the case of the bipedal motor [1, 2]. In this thesis I aim to design and construct longer range tracks for both of these motor varieties (chapters 3 and 6).

1.6 Thesis Structure

In chapter 2 I investigate a method to tune the rate of toehold-mediated strand displacement. I use TMSD to repair a mismatch with a Watson-Crick base pair, and explain the dramatic effect that the position of this mismatch has on displacement kinetics. The lessons learnt from these investigations are used to demonstrate a design for a DNA pulse generator based on informed choice of mismatch repair position and the competition between a kinetically favoured and a thermodynamically favoured product. Additionally, mismatch repair processes are involved in the operation of the bipedal DNA motor introduced in section 1.3.6 which will be described in more detail at the beginning of chapter 5.

Tension of nucleic acids can have dramatic and interesting effects on dynamic processes, both in nature [122] and in synthetic fields such as DNA nanotechnology [1, 64, 66]. I present the design and structural verification of a DNA origami 'tightrope' in chapter 3. This structure is designed to produce a defined tension across a single strand of DNA by stretching

a defined number of nucleotides across a gap of fixed distance. The tightrope can be used to study some of the effects of tension on dynamic DNA processes.

In chapter 4 I use the tightrope nanostructure to investigate the effect of single-stranded DNA tension on the kinetics and thermodynamics of DNA hybridization. Anomalous hybridization kinetics can be seen when one attempts to hybridize DNA strands to the stretched single strand of the tightrope. By hybridizing strands of different lengths to the tightrope I determine the rate constants for hybridization k_{on} and dissociation k_{off} whose ratio provide equilibrium dissociation constant $K_D = k_{\text{off}}/k_{\text{on}}$. I discover that changing the tension can change both the hybridization kinetics and equilibrium binding of strands.

The tightrope was conceived originally as a track for the bipedal DNA motor from section 1.3.6, as simulations have shown the application of tension to the track to improve motor forward stepping bias [123]. In chapter 5 I use the tightrope origami structure to attempt to move the bipedal motor along multiple steps, a feat which has never previously been achieved. I present a method to place the motors at a specific position on a tightrope track with multiple sites where the motor could choose to bind. By comparison with oligonucleotide-based control tracks I am able to decipher some of the aspects of bipedal motor operation on tightrope origami tracks with bulk fluorescence experiments. Unfortunately I am unable to observe significant quantities of bipedal motor reaching the end of a 6 site tightrope track.

In chapter 6, I make preliminary investigations into methods to extend the track for another DNA motor, the burnt-bridges motor introduced in section 1.3.4. Two methods are considered, placing a repeating track inside polymerised DNA origami nanotubes, and placing a track over a series of DNA origami tiles linked in a defined order, where each tile can bear a track with a different pattern.

Methods used in this thesis are usually standard protocols in the field. Methods may be presented in the chapters in which they are used, however complete protocols and experimental conditions are otherwise presented in appendices. In particular, the matrix, buffer, running, and staining conditions for any gel electrophoresis experiments are listed in appendix B table B.2, and experimental buffer conditions in appendix G.

Chapter 2

Mismatch Repair for Enhanced Control of Displacement Kinetics

In this chapter we investigate the effect on displacement rate of the position of a mismatch present in the substrate duplex which will be repaired by toehold-mediated strand displacement with a fully complementary invading strand. We show that, through choice of mismatch repair position, thermodynamically decoupled tuning of displacement kinetics can be achieved, and the relationship between displacement rate and mismatch position is non-monotonic. We use oxDNA to independently estimate reaction rate constants determined by experiment and find good agreement with experimental data. Aided by oxDNA simulation results we propose an explanation for this behaviour. We then use what we have learnt in the study of mismatch repair position, to design kinetic behaviour in a competitive displacement reaction. By choice of mismatch repair position, we can preferentially accelerate the reaction with one of two competing invading strands, even when the other invading strand is given the thermodynamic advantage.

Author contributions

This project was a collaboration with Thomas Ouldridge (Imperial College London UK) who provided simulation data to support my experimental results and was also involved in the interpretation of our combined results.

2.1 Motivation

DNA strand displacement underpins the operation of the majority of dynamic synthetic DNA devices including, but not limited to, DNA motors. Strand displacement can be

accelerated by adding a toehold to the system, increasing displacement rates by up to 6 orders of magnitude [38, 39]. Controlling displacement with toehold length, however, has the disadvantage that any change in displacement kinetics is coupled to a change in the relative free energies of reactants and products. Our earlier work [124] produced a facile method for fine-tuning strand displacement kinetics by changing the position of a mismatch to be introduced into a product duplex by strand displacement, without changing the overall free energy change of reaction. In this chapter we explore the converse scenario, where the position of a mismatch initially present in a duplex is varied, and the effect of its repair position by a perfectly matching invading strand on displacement kinetics is investigated.

Understanding the precise impact of mismatches on displacement processes is of great interest. Dynamic DNA systems such as enzyme-driven motors [2,3], may require the inclusion of mismatches in certain domains to prevent cleavage by nickases. Mismatches may also be used in design to destabilise certain domain interactions [2, 125, 126]. They are also used in DNA sensors for the identification of SNPs [47, 127–132]. Often there is no consideration during design of the kinetic impact of mismatch position choice. Our method will provide a new design tool to engineer devices with tuned kinetic behaviour.

2.2 Previous Work

2.2.1 OxDNA for coarse-grained simulations of DNA

Molecular dynamics (MD) simulations of macromolecular structures such as DNA have the potential to uncover details or intermediate states which cannot be probed experimentally. 'All-atom' MD simulations use detailed physical force fields [133–135] and calculate, for each timestep, the effect on each atom in the simulation, which is computationally costly due to the large volume of state space. Computational cost of simulation can be reduced by coarse-graining the system to one with far fewer degrees of freedom, which still encompasses the important behaviour. A tool developed by Ouldridge and colleagues [136, 137] provided one such coarse-grained model of DNA, by reducing a DNA nucleotide to a rigid unit with a number of interaction sites. Each nucleotide has a hydrogen bonding site, a stacking site, a backbone repulsive site, and a base repulsive site to represent excluded volumes. Nucleotides

in a strand of DNA are connected by a spring potential. Interactions between nucleotides occur pairwise according to a potential including hydrogen bonding, cross-stacking, and excluded volume terms. Nearest-neighbour nucleotides also incur contributions to the potential from stacking (with a minimum at a separation of 0.34 nm between bases), and backbone connections [136]. The model originally neglected sequence effects [136], but has since been extended to include sequence-dependent effects on hydrogen bonding and stacking [138], salt dependence, and major and minor grooving [139]. OxDNA has repeatedly proved successful in reproducing experimental behaviour [40, 115, 124, 137, 140–142], and can also be used to inform design [103, 123]. OxDNA is used in this chapter to independently produce information on the dependence of TMSD displacement kinetics on mismatch position and to provide a probabilistic explanation of the results.

2.2.2 Studies of toehold-mediated strand displacement kinetics

By adjusting toehold binding strength it is possible to change the rate of strand displacement. We learnt in chapter 1 that the rate of toehold-mediated strand displacement increases exponentially with toehold length as demonstrated first by Yurke and Mills [38], and then by Zhang and Winfree who also showed that the exponential dependence on toehold length saturated for toeholds longer than 7 nt. Other methods of altering toehold binding affinity include introducing mismatches into the toehold domain, and metallo-toeholds, where non-native base-pairing interactions can be favoured in the presence of metal ions such as Mercury (II) ions. The mismatched base-pair T–T is stabilised by Hg^{2+}, therefore the binding affinity of a toehold containing these mismatches can be tuned by varying Hg^{2+} concentration [143].

An alternative method for controlling toehold-mediated strand displacement moved the toehold from its conventional location, immediately prior to the displacement domain, by introducing a spacer. Remote toeholds were investigated by Genot and colleagues [144] where the authors showed that displacement rate can be reduced by at least three orders of magnitude by the addition of a spacer between the toehold and the displacement domain in both the invading strand and substrate duplex. The flexibility of the linker also influences displacement rate with remote toeholds (e.g. double stranded DNA spacers exhibit slower

displacement rates than single stranded spacers with the same number of nucleotides). The addition of a spacer can increase kinetic differentiation between toeholds with lengths beyond 7 nt, as the spacer reduces the probability of successful displacement from a toehold-bound state. Related work by Li and colleagues [145] investigated a single-sided remote toehold, where the spacer was only present between the toehold domain and the displacement domain of the substrate duplex, with no spacer in the invading strand. They show that longer ss-DNA spacers have greater temperature dependence than shorter spacers, and that increasing temperature increases displacement rate. For the single-sided remote toehold, displacement rates were also found to decrease with increasing spacer length. Yang and colleagues used an 'allosteric' toehold strand to expose a toehold on a reactant duplex by partially displacing one end of a duplex with a regulator strand and a separate regulator toehold for the regulator TMSD. The size of the newly exposed toehold was determined by the penetration depth of the regulator strand into the duplex, therefore reaction rates could be tuned with the concentration of regulator strand, the length of penetration and the length of the regulator toehold domain. Longer regulator toeholds, higher regulator concentrations, and greater penetration lengths all increase displacement rates. The idea of using another DNA strand to partially expose a domain and produce a toehold was not novel however, it had already been used implicitly in the design of bipedal motors with inter-foot competition [1, 2].

The toehold need not be formed from DNA. A non-DNA toehold was demonstrated [146], where biotinylated substrate duplex, and biotinylated invading strand, are brought together with a streptavidin molecule (strepatvidin is a tetramer with 4 high affinity binding sites for biotin). The, otherwise blunt, displacement is enhanced by increases in local concentration due to co-binding a single streptavidin. This design is sensitive to properties of the spacer between the biotin and the displacement domain on both biotinylated reactants, as they must span the distance between streptavidin binding sites to intra-molecularly react.

As an alternative to modifying the toehold, the displacement domain can be changed to alter displacement rates. The branch migration phase of TMSD can be manipulated by changing the energy landscape for branch migration. This can be achieved by introducing mismatches into the displacement domain as the reaction proceeds, which have been shown

to slow branch migration [147]. Intriguingly we discovered in our previous work that the displacement rate varies with the position of the mismatch within the displacement domain [124]. This effect cannot be explained using thermodynamics alone, as the overall free energy change of reaction does not change significantly with mismatch position. We found that the rate of TMSD increased as the mismatch was moved further from the toehold, eventually reaching the rates of TMSD where no mismatch was introduced, as the mismatch was introduced towards the end of the displacement domain opposite the toehold both in experiment, and in oxDNA coarse grained simulation. Similar results were later reproduced for surface immobilised components [148, 149].

To understand the underlying physics which dictates these effects, rather than simply characterising with an empirically fitted effective rate constant, we must move beyond models of branch migration as a single, reversible first-order transition. Increased resolution of the branch migration process is required. Srinivas and colleagues presented a model for strand displacement called the intuitive energy landscape model or IEL [40]. This model describes toehold binding and branch-migration-intermediates with a free energy of association, where the invader and substrate duplex form their initial contact, followed by zippering steps of toehold base pair formation, and then saw-tooth oscillating free energies of branch-migration-intermediates, with additional energetic penalty for formation of a three-stranded branch-point when branch migration begins. They found that to fit the experimental data, the combined ΔG from the saw-tooth amplitude, and the penalty for branch-point formation had to be larger than expected from nearest-neighbour thermodynamics. Melting experiments of complexes designed to represent branch migration intermediates provided evidence for the free energy penalty for branch-point formation, also confirmed by oxDNA simulation. Additionally, oxDNA demonstrated that the progression of the branch-point produced metastable intermediates with high free energy, where more stacking interactions were broken than in simple fraying of the toehold, justifying a saw-tooth amplitude greater than the energy difference between base pairs in the toehold. OxDNA was able to reproduce the exponential dependence of TMSD kinetics on toehold length, and the saturation for long toeholds, which was not possible for models including only nearest-neighbour thermodynamics.

Branch-migration-intermediates were found to be more likely to be coaxially stacked when close to the ends of the displacement domain, and unstacked further from the ends when the steric hindrance of the branches is greater.

2.3 Method

2.3.1 Experimental system

The system used to study strand displacement rate consists of two modules. The first is the mismatch repair strand displacement reaction we wish to investigate. The second is a 'reporter' displacement reaction which results in a fluorescence change [39, 124].

In module 1, a substrate duplex VT is composed of a victim strand V hybridized to a target strand T by a 20 bp displacement domain. The displacement domain is complementary except for the presence of a single mismatching C-C base pair between the two strands. 16 nt at the 5' end of V form a displacement domain for the reporter displacement reaction in module 2, and produce a dangling end on duplex VT (green) opposite the toehold (dark blue) which is on the 5' end of target strand T. Invading strand I is able to bind at the toehold and compete with V for binding within the displacement domain, creating a three-

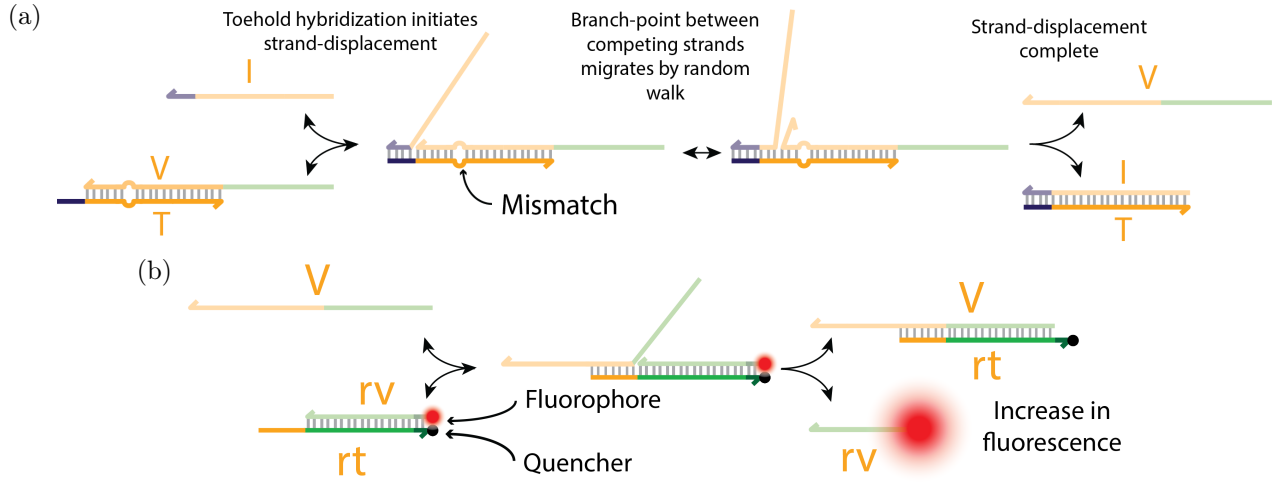

Figure 2.1: Mismatch repair experiment schematic. **a** Module 1: The mismatch repair reaction where a duplex VT containing a single mismatching pair of bases, is attacked by an invading strand I, fully complementary to the target strand T of this duplex, including the complementary toehold domain (dark blue). I competes with V for binding to T and the reaction progresses as the branch-point between the two strands diffuses along the displacement domain in a random walk. If the branch-point reaches the end of the displacement domain, strand V is displaced from the duplex. **b** Module 2: Free V, the output of module 1, is able to displace the reporter victim strand rv from the reporter complex liberating the Cy5 fluorophore and increasing observed fluorescence.

strand branch-point between V, T, and I which can move along the displacement domain by a random walk. Migration of the branch-point beyond the mismatch position will 'repair' the mismatch with a G-C base pair. As the branch-point reaches the end of the displacement domain the victim strand V is liberated from the duplex.

The output from module 1 is a free, single-stranded, victim strand V, which acts as the input for module 2. Module 2, an indirect reporter module [39, 124], uses V as an invading strand for another displacement reaction with the reporter complex. The toehold for the reporter complex displacement reaction is the 7 5'-most nucleotides of the mismatch repair displacement domain. Displacement of the victim strand (rv) of the reporter complex by V results in the separation of a Cy5 fluorophore and an IowaBlack RQ quencher which are initially held close together in the reporter complex, in a state where Cy5 is quenched. After displacement, the Cy5 is unquenched, and the recorded fluorescence at the excitation and emission wavelengths for Cy5 increases. This signal is used to determine the effective second-order displacement rate for the mismatch repair in module 1. Reporter complexes are prepared with an excess[1] of reporter victim strand (6:5 rv:rt) to ensure that there is no free reporter target strand (rt), which could hybridize to liberated V without causing a change in observed fluorescence.

2.3.2 Toehold and displacement domain design

The toehold was located at the 5' end of target strand T. The toehold length was 4 nt, chosen to tune the sensitivity of the system to changes in mismatch repair position. For long ($\geq$7 nt) toeholds, the effective second-order rate constant of TMSD saturates at $\sim 10^6$ $M^{-1}s^{-1}$ [39] as toehold binding is effectively irreversible and therefore guarantees successful displacement once the toehold has associated. The system is then insensitive to mismatch repair position as further increases in displacement probability after the mismatch is repaired will cause little change to the overall success probability which is already high [148]. Use of a short toehold also simplifies analysis of results downstream as the reaction rates with shorter toeholds are slowed relative to those with longer toeholds such as the

[1] Verified by titrating rv with target strand rt until no further quenching was observed.

reporter reaction. In the limit that the reporter reaction rates are much greater than the rate of the mismatch repair displacement under investigation we can assume that observed fluorescence changes are proportional to the rate of release of victim strand V. As a mismatch is being repaired, we hypothesise that the no-mismatch reaction will be one of the slowest investigated, and introducing a mismatch in the substrate duplex will increase the rate beyond this, potentially by up to 3 orders of magnitude [124]. Reactions with no mismatches and a 4 nt toehold reach half-completion in $\sim$ days. This would allow all measurements to be taken within the measurable range of the equipment used (where data acquisition interval is 60 sec) while avoiding as far as possible unduely long experiments.

The mismatch used for all cases was a cytosine-cytosine (C-C) mismatch which is the most destabilising mismatch [36]. The base sequence for the strands used and the reporter complex were adapted from sequences used in reference 124. These sequences were designed to minimise secondary structure using NUPACK software [150]. The length of the displacement domain was . To create a mismatch, compared to the equivalent perfect match, the target strand and the invading strand which form the product duplex are co-mutated to form the mismatch between the target and victim when part of the initial duplex, but to maintain a match between target strand and invading strand, thus allowing the same victim strand to be used for the matching and mismatching cases. This minimises the number of different victim strands required, and is desirable as changes to the victim strand can change the response of the reporter reaction to which the victim strand is the input.

A set of multiple displacement domains were designed in order to produce multiple different mismatch positions whilst maintaining most of the nearest-neighbour interactions within the domain and hence maintaining a similar duplex free energy [28] over the range of mismatch positions. To achieve this, the mismatch neighbourhood was defined as the mismatch itself and its two nearest-neighbour base pairs (see figure 2.2a). This set of three base pairs was shifted along the displacement domain by moving the base pair immediately to one side of the mismatch neighbourhood to the opposite side of the mismatch neighbourhood as depicted in figure 2.2b. At most two nearest-neighbours out of the total 19 will be changed by moving the mismatch position in this way, and the base content of the strand will be

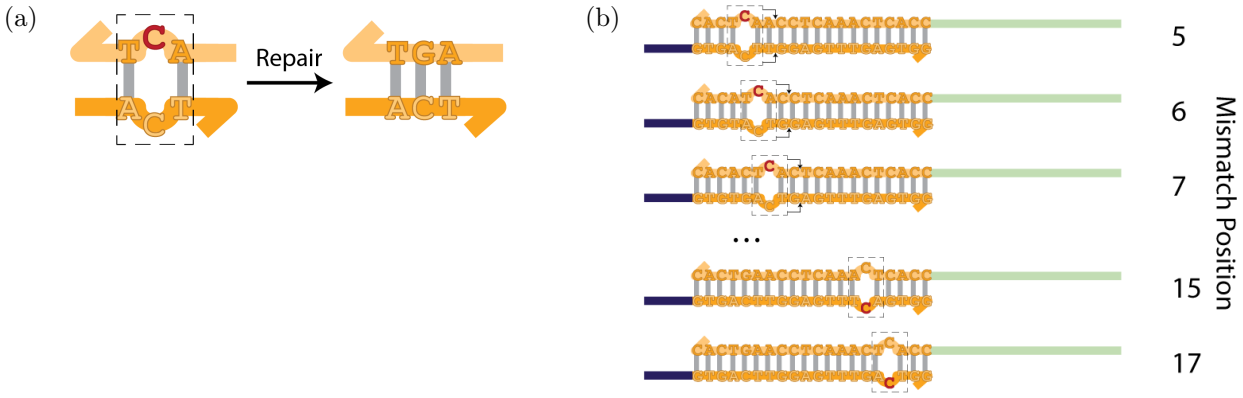

Figure 2.2: Mismatch repair sequence design. **a** The C-C mismatch and its neighbouring base pairs constitute the mismatch neighbourhood depicted within the grey dashed line. The mismatch is repaired to a fully complementary 'match' by displacement with a strand which has a guanine base in place of the mismatching cytosine. **b** Scheme for generating displacement domains for different mismatch repair positions. To move a mismatch position by one base pair, the base pair to one side of the mismatch neighbourhood is moved to the opposite side of the mismatch neighbourhood. This method allows mismatch positions 2 to 13 to be generated. For late mismatch positions 15 and 17, to avoid mutating the reporter toehold domain, a guanine base in the the target strand (dark orange) found in the same mismatch neighbourhood can be mutated to a cytosine.

identical. This method allowed mismatches between 2 bp and 13 bp from the toehold to be produced. Mismatches further along the displacement domain were not achievable by this method because they occur in the domain which forms the toehold for the reaction with the reporter complex. Any change in sequence in this domain would not allow the use of a common reporter for different mismatch positions and is therefore undesirable. Within the reporter toehold domain there are, however, two occurrences of the mismatch neighbourhood. By the switching of one base on the target strand from a guanine to a cytosine, mismatch positions 15 and 17 could be produced without any changes to the victim sequence in the reporter toehold domain. The free energy changes of TMSD predicted using NUPACK software [150] are given in table 2.1. The free energy change of reaction is least negative for displacement with no mismatch because repairing a mismatch causes a drop in free energy. The free energy change of reaction for mismatch positions 2 and 3 was also less negative than the majority of other positions. This is because NUPACK predicts incomplete mismatch enclosure, i.e. the duplex end close to the mismatch may be frayed, rather than base-paired causing the VT duplex to have a more negative free energy[2] and therefore a smaller free energy drop upon TMSD. All other mismatch positions show a near-identical free energy

[2]Frayed duplex ends are less destabilising than enclosed mismatches for cases close to the duplex end.

change of TMSD reaction, therefore any differences in the rates of these processes are not due to changes in relative stabilities of reactants and products.

Mismatch position (bp from toehold)	$\Delta\Delta G‡$ kcal mol^{-1}
No mismatch	-5.58
2	-7.86
3	-10
4	-11.08
5	-11.27
6	-11.27
7	-11.27
8	-11.27
9	-11.27
10	-11.26
11	-11.27
12	-11.27
13	-11.27
15	-11.21
17	-11.27

Table 2.1: Free energy change of toehold-mediated strand displacement reaction for different mismatch positions from the toehold $\Delta\Delta G$. ‡Calculated using NUPACK software [150].

2.3.3 Reporter complexes

Despite considerations in the original sequence design attempting to prevent the requirement for multiple reporter complexes, ultimately two different reporter complexes were needed to probe the full range of mismatch positions. The reporter complex used for the majority of experiments, Rep A, was an 18 bp duplex with a 7 nt toehold. 16 bp of the duplex were complementary to the 5' reporter displacement domain on V. The final 2 bp of the duplex were anticipated to melt rapidly when the branch point reaches the end of the 16 bp displacement domain.[3] A schematic of this reporter complex is shown in figure 2.3a.

Leak reactions

Reacting components of module 1 are designed to remain in a kinetically trapped state where they cannot trigger module 2 without the mismatch repair reaction taking place. To achieve this, the toehold for reaction with the reporter complex (orange) is sequestered

[3]The reverse reaction has a 2 nt toehold, but is thermodynamically disfavoured as it forms 5 fewer bp.

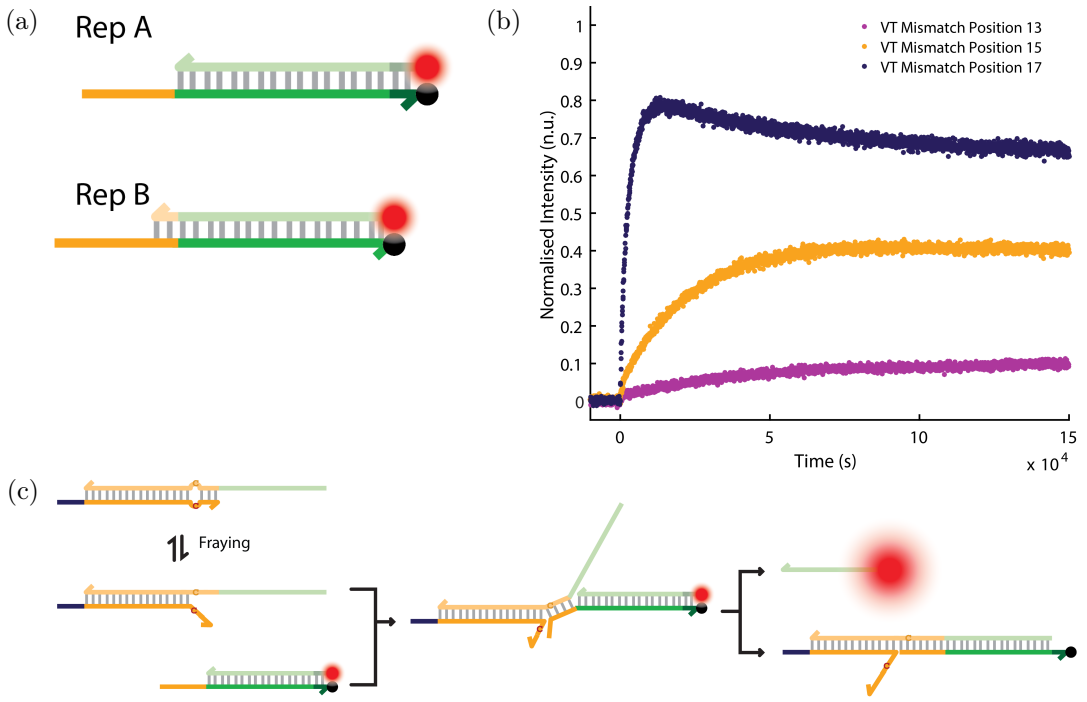

Figure 2.3: Leak reactions and reporter complexes. **a** Reporter complexes used in experiments. Rep A produces leak reactions with mismatches at positions 15 and 17. Rep B does not cause leak reactions but is incompatible with mismatch positions 12 and 13. Data from both reporters is required to probe the full range of mismatch positions. **b** Data showing leak reactions between solution of 25 nM reporter complex after the addition of 12.5 nM VT with a mismatch at position 13 (purple), 15 (orange), or 17 (dark blue). Data is normalised to 0 prior to addition of VT and 1 at the fluorescence level after the displacement of 12.5 nM Rep A. **c** Schematic showing possible leak pathway for a position 17 mismatch with Rep A. As the position 17 mismatch is close to the end of the displacement domain, the duplex is able to fray at the end, transiently exposing bases at positions $\geq$ the mismatch. This exposes a partial toehold for reaction with the reporter complex via a 4-strand intermediate, displacing the reporter victim strand rv without displacing V from VT.

in the displacement domain of the duplex VT. The reporter displacement domain on V (green in figure 2.1a), however, is single-stranded. Interactions between this domain and the reporter complex are by blunt-ended strand displacement which occurs with second-order rate constant $\sim 1\ \mathrm{M}^{-1}\mathrm{s}^{-1}$. The system was recorded after addition of VT complexes to the reporter in solution for every experiment, and no appreciable leak reaction was observed for the majority of mismatch positions. Mismatches at positions 15 and 17 however, caused leak reactions when substrate duplex VT was incubated with reporter complex but without invading strand. Exemplary data is shown in figure 2.3b. Data has been normalised to the maximum achievable fluorescence if all VT had reacted with the reporter complex.

A potential mechanism for the reaction is shown in figure 2.3c. As the mismatch is close to the end of the displacement domain it may not be fully enclosed all of the time.

The post-mismatch base pairs may spend some proportion of their time unbound, leaving a partial toehold for the reporter reaction exposed, and able to initiate TMSD with the reporter complex. The reporter victim strand rv can then be displaced from the duplex without V ever being released from VT.

To remove the leak reaction, the reporter complex displacement domain was extended by 2 bp into the displacement domain of the mismatch repair reaction. This is anticipated to prevent leak because the fraying of terminal base pairs will then produce a shorter effective toehold for the leak reaction. The toehold was also extended by 2 nt to maintain a reporter toehold length of 7 nt. The last 2 bp at the opposite end of the reporter duplex from the toehold were also omitted such that the reporter complex duplex length remained 18 bp. This reporter is referred to as Rep B, and is depicted in figure 2.3a. Using Rep B in place of the original reporter Rep A eliminated leak reactions for mismatch positions 15 and 17. Due to the further penetration of the toehold domain of Rep B into the displacement domain of VT, this reporter is incompatible with mismatch positions 12 and 13[4], however it can be used with all other mismatch positions.

2.3.4 Spectrofluorometry measurements

In a fluorometer experiment the reporter complex was initially present at 25 nM. Data was collected until the fluorescence level settled. 12.5 nM substrate duplex VT was then added to the reporter complex solution and a further baseline was collected. Finally 12.5 nM invading strand was added, marking the mismatch repair reaction start.

2.3.5 Fitting

Experimental data was zeroed by setting t=60 at the first data point collected after mixing (where 60 s is the measured typical mixing time) and subtracting from all fluorescence counts the fluorescence of the last data point collected before addition of invader. The zeroed data f was then fitted using MATLAB function '`fit`' between $f(t = 0)$ and the first point at which the fluorescence exceeds 99% of the maximum recorded value.

Data was fitted assuming an irreversible displacement reaction $VT + I \xrightarrow{k} IT + V$

[4]Because the mismatch neighborhood overlaps with the reporter toehold domain.

with stoichiometric quantities of reacting species VT and I[5]. The fitted function was `a(1-1/(1+r*t))` where `a` is the plateau height and `r` is the rate of TMSD. Second-order rate constant k is related to rate `r` by the relation $k = \mathtt{r}/c$, where $c = 12.5$ nM is the initial concentration of reacting species. Fit parameters `a` and `r` were initiated at `a`= f_{max} a.u., `r`= 10^{-3} s^{-1} and bounded $0.9 \times f_{max}$ a.u.≤`a`≤ $1.5 \times f_{max}$ a.u., 0 s^{-1} ≤`r`≤1000 s^{-1}. The plateau height `a` varies between experiments due to a bug in the instrument control software where changes in emission slit widths were reported without being applied (see appendix C). Wider emission slits produce higher signals but do not affect the sample. Displacement rate `r` and derived second-order rate constant k should be independent of plateau height `a`.

2.4 Results

Derived second-order rate constants k were divided by the second-order rate constant of an equivalent TMSD reaction with no mismatch to be repaired k_0. These values are shown in

[5]The assumption of irreversibility should be valid in the limit $k_{\mathrm{rev}} \ll k$. This condition is satisfied for all $\Delta\Delta G$ in table 2.1.

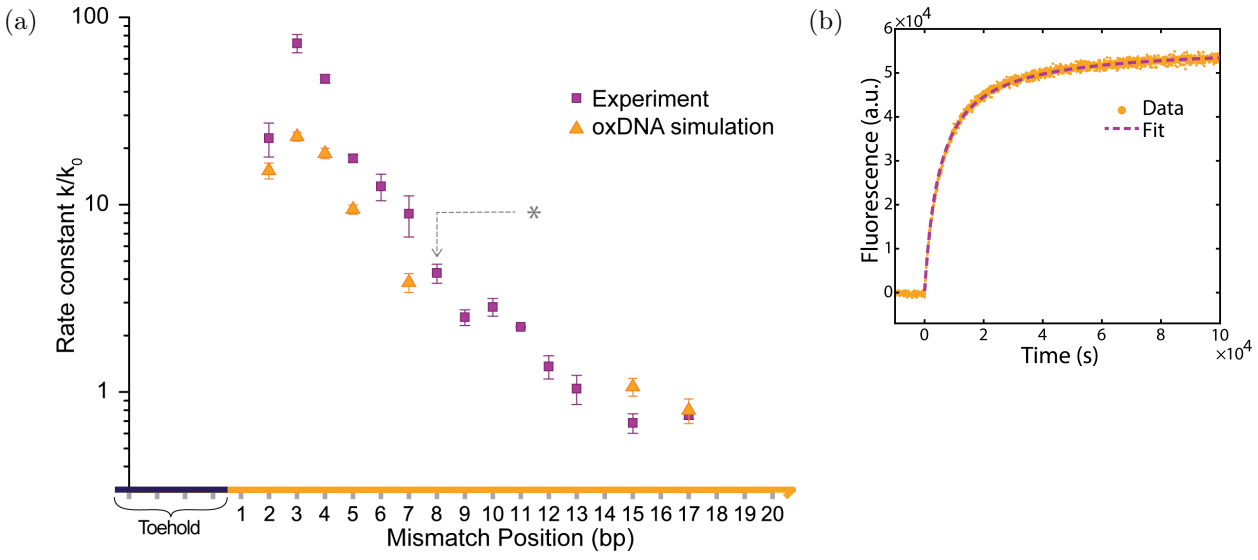

Figure 2.4: Relative second-order rate constants for TMSD for a range of mismatch repair positon. **a** Plot of relative rate constants k obtained from either second-order fits of data up to 99% maximum recorded fluorescence (purple), or oxDNA simulation (orange), divided by the rate of displacement where no mismatch was present k_0 ($= 2.4 \times 10^3$ M^{-1}s^{-1}). Data points show mean values and error bars show SEM determined from repeat measurements. Experimental data points have additional systematic shift of ±0.14×, and simulation points ±0.11× due to error in k_0. **b** Exemplary data (orange) and second-order fitted curve (purple) for position 7 repair which contributes to the data point marked * in **a**.

figure 2.4a. Relative rate constants k/k_0 were found to vary by two orders of magnitude. The relationship between mismatch repair position and relative rate constant was non-monotonic, first increasing, then decreasing, as the mismatch position moves further from the toehold. The maximum observed rate was for mismatch position 3 bp from the toehold.

2.4.1 OxDNA

Relative rate constants were independently determined from oxDNA simulation using the same sequences of duplex experimentally tested. These data are plotted in orange alongside experimentally determined relative rate constants in figure 2.4a. In toehold-mediated strand displacement, toehold binding is concentration dependent, however branch migration is independent of the bulk concentrations of VT and I during a single encounter between invader and target toeholds.

The exact details of the simulations are beyond the scope of this thesis, however some key details are included here. The simulation concentration (set by the dimensions of the simulated 'box' containing reacting species) was $\sim$ 100 μM which is 8000 fold larger than the experimental concentration. This high concentration allowed toehold association and branch migration processes to proceed on similar timescales. Initial association timescale of the reacting species VT and I should scale linearly with inverse concentration, whereas branch migration timescale (or time spent in a three-stranded intermediate state VTI) will be independent of concentration. Consequently, were the simulation concentration similar to the experimental concentration, the time spent undergoing branch migration would be negligible compared to the association timescale. The time spent in branch-migration-intermediate states is therefore neglected in the calculation of rates from simulation. Instead, to determine reaction rates, toehold association rates from simulation are multiplied by the probability that the displacement will complete during an individual toehold association event (i.e. the victim strand V will dissociate before the invading strand I dissociates). The reporter complex was not included in the simulation. For further details of the simulation method see reference 124, where similar simulations were performed.

Additionally, though oxDNA simulations have an implicit associated timescale, this timescale is subject to many assumptions and approximations. It is for this reason that relative rates instead of absolute rates are compared as all rates should be affected by these approximations and assumptions in the same way, and therefore comparisons between relative rates of similar processes should be more reliable [40, 123, 124].

Relative reaction rate constants determined from oxDNA and from experiment show remarkable agreement. Both methods show the maximum reaction rate for mismatches 3 bp from the toehold and show the same exponential relationship between rate and mismatch position for intermediate mismatch positions. Both show that the late mismatch position 17 has a near-identical rate to that of non-mismatching strand displacement. Simulated relative rates are systematically lower than the experimental results for the same mismatch position which may be due to an underestimation of the destabilising effect of the mismatch inherent in the oxDNA model according to Thomas Ouldridge.

2.4.2 Comparison of fitting methods

Three different fitting methods were used to find the second-order rate constant. The motivation for attempting to fit the data via different methods was due to artefacts in the raw data which cause it to deviate from simple second-order displacement kinetics. The main issue was with photobleaching of the reporter complex. Photobleaching was not uniform between samples. Variations in photobleaching could be due to differing amounts of air sealed in the cuvette with the sample solution, and variation in the excitation intensity as the fluorometer bulb ages. As a result of photobleaching, some traces with rapid kinetics reach a peak and then decay. Second-order displacement kinetics cannot describe this non-monotonic behaviour.

Second-order fit with exponential decay envelope

$$ae^{-bt}\left(1-\frac{1}{1+ckt}\right) \tag{2.1}$$

The three parameters, a, k, and b were fitted, and results for fitted k are shown as green diamonds in figure 2.5a. Example fits are shown with solid green lines in figure 2.5b and 2.5c.

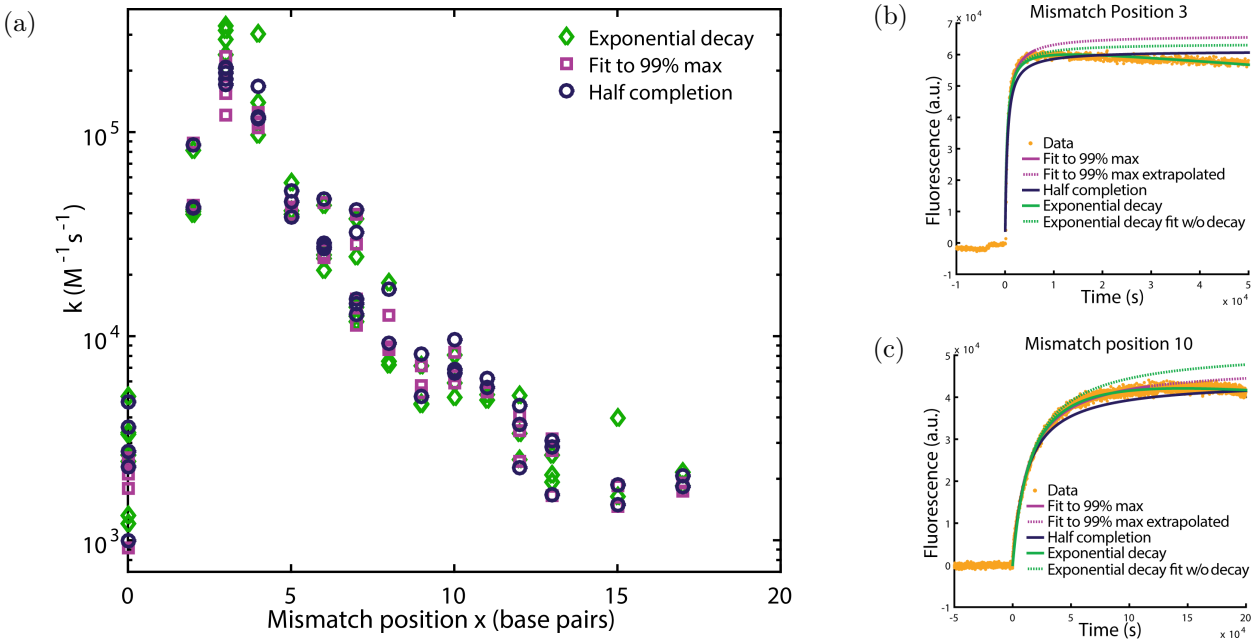

Figure 2.5: Robustness of rate constant to fitting method. **a** Plot of second-order rate constant for all experimental data determined by three different fitting methods: Second-order displacement with exponential decay envelope (green diamonds), Second-order fit up to 99% maximum fluorescence (purple squares) and second-order rate constant from half-completion time. Individual fits to experimental data (orange) are shown for positions 3 **b** and 10 **c**. Curves in plots **b** and **c** are: second-order displacement with exponential decay envelope (solid green), second-order displacement with parameters from exponential envelope fit but shown without the envelope (dotted green), second-order displacement fit up to 99% maximum fluorescence (solid purple), same fit extrapolated beyond fitted range (dotted purple), and second-order displacement from half-completion time (solid blue).

Dotted green lines in figure 2.5b and 2.5c show the second-order function without exponential decay envelope ($b = 0$) for parameters a and k obtained from fits with the envelope.

Second-order fit data up to 99% peak intensity

This is the method used to produce the fits used in figure 2.4a. As the second-order function is monotonically increasing, fitting values after a maximum has been reached (i.e. when net fluorescence decreases due to photobleaching) fit quality is poor. We therefore fit data points up to the first data point exceeding 99% maximum recorded fluorescence value and ignore later data points. The fitted function is given by equation 2.2:

$$a\left(1 - \frac{1}{1+rt}\right) \tag{2.2}$$

Results are shown as purple squares in figure 2.5a. Example fitted curves are shown with solid purple lines in figure 2.5b and 2.5c, and dotted purple lines outside the fitted range.

Half-completion time

The relationship between second-order rate constant k and time until half completion $t_{1/2}$ for a second-order displacement reaction with stoichiometric initial concentrations c is given by equation 2.3:

$$k = \frac{1}{c \times t_{1/2}} \tag{2.3}$$

Using this relationship, values of k were determined by finding the time at which data exceeds half the maximum value. Results are shown in dark blue circles in figure 2.5a, and second-order curves as dark blue lines in figure 2.5b and 2.5c.

By comparing the fitted curves and rate constants determined by each method (figure 2.5) we can see that the use of a different fitting procedure produces qualitatively and quantitatively similar results and trends in fitted second-order rate constant, demonstrating that any conclusions drawn from fitted rate constants are robust to fitting method.

2.4.3 Comparison of reporter complexes

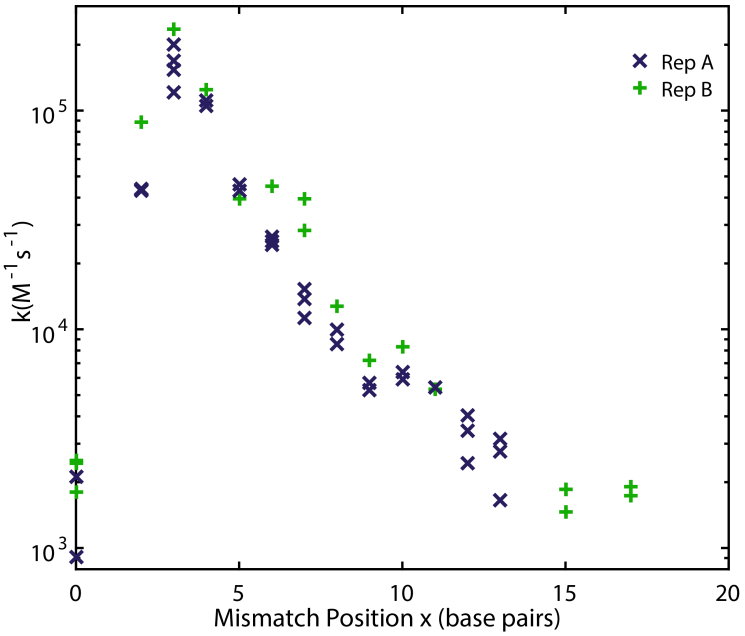

Figure 2.6: Comparison of rate constants determined using reporter complexes Rep A (blue crosses) and Rep B (green plusses).

Where possible (i.e. for mismatch positions 0-11) measurements of TMSD were performed with both reporter complexes Rep A and Rep B. Results from both are combined to produce figure 2.4a, however, second-order rate constants determined from each reporter complex

separately are shown in figure 2.6. Both reporter complexes show a qualitatively similar relationship between mismatch position and displacement rate. Both show derived rate constants which are quantitatively similar, allowing the two to be combined in the analysis.

2.5 Insights from oxDNA Simulations

OxDNA simulations can provide information about the probability of success of a strand displacement encounter with the toehold (success is defined as the displacement of V whilst I remains bound), which cannot be determined from the experimental data. This information, as previously noted, is used for the calculation of reaction rates, but can also be used to decipher the reasons behind the variation in rate with mismatch repair position.

Systems with and without mismatches were simulated. Those with mismatches are shown by orange points in figure 2.7, and those without are shown by purple points. Figure 2.7 shows displacement probabilities. Coordinate x is the position along the displacement domain. In systems with mismatches, position x is also the mismatch position.

We consider the probability of mismatch repair, P(Repair at x), defined as forming a

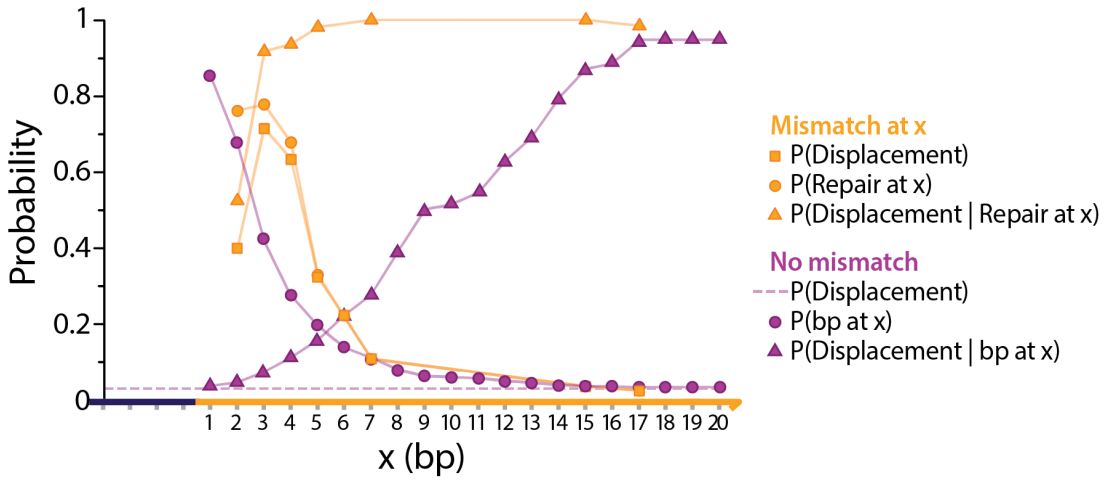

Figure 2.7: Displacement and repair probabilities during a single encounter with the toehold from oxDNA simulation for a system with a mismatch located at position x (orange), or without a mismatch (purple). Probability of successfully displacing V in a single encounter with the toehold is given by squares for a system with a mismatch and by a purple dashed line for the no-mismatch system. Probabilities of forming a base pair between the invader I and target T at position x during an encounter is given by circles. In the system with a mismatch at position x this will involve the 'repair' of the mismatch. Probability of completing displacement given that a base pair between I and T at x has been formed is indicated by triangles.

correct, Watson-Crick, base pair at the site where formerly there was a mismatch at position x (orange circles). This can happen once or more during the encounter, (i.e. the mismatch can be repaired and then later re-introduced, and repaired again). In the no-mismatch system the analogous quantity considered is the probability of forming a base pair at least as far as position x between I and T, P(bp at x) (purple circles).

We also consider the probability of displacement given that a mismatch has been repaired P(Displacement|Repair), or base pair formed P(Displacement|bp at x), at position x. This tells us how likely complete displacement is when strand I has made it as far as position x by branch migration.

From Bayes' theorem $P(A|B)=\frac{P(B|A)P(A)}{P(B)}$, and given that P(Repair at x|Displacement)=1, as displacement cannot occur without repair, it can be shown that the product of quantities P(Displacement|Repair at x)$\times$ P(Repair at x) will give the probability of displacement P(Displacement) (orange squares). For the system with no mismatch P(Displacement)= P(Displacement|bp at x)$\times$ P(bp at x) is independent of x (dashed purple line in figure 2.7).

The non-monotonicity of the relationship between P(Displacement) and mismatch position can be explained by the probability of displacement given repair (orange triangles). Mismatch positions 2 and 3 have similar probabilities of mismatch repair ($\sim$80%) (orange circles), however P(Displacement|Repair at x) is low, only $\sim$50% for a mismatch repair at position 2, indicating that even if the mismatch is repaired, dissociation of I prior to displacement of V will happen half of the time. For mismatch positions 3 or later the success probability given mismatch repair is over 90%, indicating that repair almost guarantees that displacement will be successful. By contrast, P(Displacement|bp at x) in a non-mismatching displacement reaction (purple triangles) increases more gradually with x, over the full range of x, and only exceeds 90% for $x \geq 15$ bp. P(Displacement|bp at x) is matched by a decreasing probability of reaching x, P(bp at x), in a single encounter (purple circles). The reason for this difference in success probability between cases where mismatches are repaired, and cases where there is no mismatch, is the free energy gain from mismatch repair. This gain is lower for the position 2 mismatch than other positions (see table 2.1).

The phenomenon of decreasing displacement rate with increasing mismatch position for intermediate mismatch positions follows from the probability of repairing a mismatch at x, P(Repair at x). Mismatch repair is less probable the further from the toehold the mismatch is, but for late mismatch positions, once the mismatch has been repaired, successful displacement is almost guaranteed as P(Displacement|Repair at x)$\approx$1. Note that the P(Repair at x) is greater than P(bp at x) for x <7 bp. This is because a mismatch destabilises nearby base pairs making it easier to displace as far as x when a mismatch is present. Beyond x=7 bp, there is negligible difference between P(Repair at x) and P(bp at x).

2.6 Interpretation

Mismatch repair generally increases the rate of toehold-mediated strand displacement compared to the equivalent TMSD reaction where no mismatch is present, as the mismatch repair process biases branch migration by introducing an energy barrier to disfavour backwards steps which reintroduce the mismatch. The position of the mismatch in the displacement domain determines how much the reaction rate will be increased relative to the no mismatch case. For late mismatch positions (> 13 bp in a 20 bp displacement domain) there is no significant difference in the rate of displacement compared to rates without a mismatch. For the repair of early mismatches, displacement rates can be accelerated by two orders of magnitude compared to the no-mismatch displacement rate. This is because reaching an early mismatch by branch migration is likely once bound to the toehold, and once the repair has occured, the probability of completing displacement is high. The further the mismatch is moved from the toehold, the less likely it is that branch migration will reach the mismatch in order to repair it before the invading strand dissociates, and hence the lower the overall rate of displacement as the mismatch moves further from the toehold. If the branch-point reaches as far as 15 bp along the 20 bp displacement domain, success probability is high, regardless of whether there is a mismatch or not, and this explains why there is no difference between rates with and without mismatches when mismatches are late.

2.6.1 Comparison with mismatch introduction results

Mismatch repair and mismatch introduction can be considered as inverse processes. For a 0 nt toehold they are exactly reverse processes. We address this by comparing our results from the previous study of mismatch introduction [124], and our findings for mismatch repair, to ensure that they are consistent.

In the limit of a 0 nt toehold, for a system with an n bp displacement domain, mismatch repair at position x will be the reverse of mismatch introduction at position $n-x$. By detailed balance the ratio of rate constants for these processes is specified by the free energy difference between matched and mismatched states, the free energy cost of mismatch formation.

$$\frac{k_{\text{repair at } x}}{k_{\text{introduction at } n-x}} = e^{\frac{\Delta G_{\text{mismatch}}}{k_B T}} = e^{\frac{\sim 5.7 \text{kcal mol}^{-1}}{k_B T}} \tag{2.4}$$

Consequently, because $\Delta G_{\text{mismatch}}$ is largely independent of mismatch position, we would expect positions x where mismatch repair has a high rate, to indicate that mismatch introduction will have a high rate at position $n - x$. We observe maximum rates at mismatch repair position $x = 3$, and mismatch introduction position $n - x = 14$, where $n = 17$ $\implies x = 3$ consistent with this. In mismatch repair we observe a general decrease in repair rate with increasing x. This should translate to a decrease in mismatch introduction rate with increasing $n - x$, which is equivalent to an increase in mismatch introduction rate with increasing x, as is observed in reference 124. We therefore conclude that the findings of the current study, and our previous work are consistent.

2.7 Informed Design of a DNA Pulse Generator

Equipped with the information about the effect of mismatch repair position on displacement rate we are able to design dynamic devices with kinetic bias. A simple example is a duplex with two toeholds able to be attacked from either end by two invaders. Both toeholds are the same, but displacement is designed to be kinetically asymmetric due to the position of a mismatch in the substrate duplex to be repaired. The choice of the position of a mismatch initially present between the victim and the two-toeholded target will determine which invader is kinetically preferred, and will displace the victim strand more rapidly from

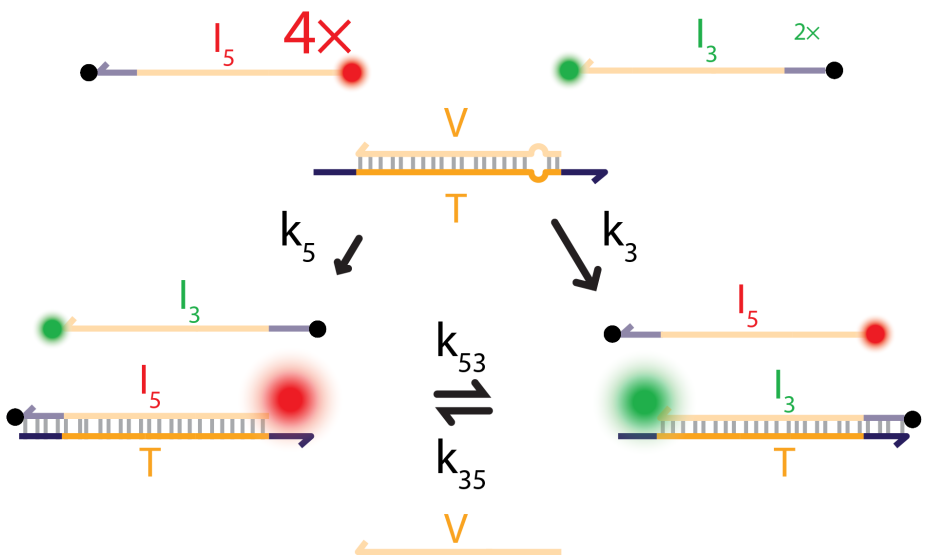

Figure 2.8: Two-toehold mismatch repair schematic. A substrate duplex VT with an asymmetrically positioned mismatch has two toeholds at either end of strand T, one at the 5' end and one at the 3' end (a mismatch position 18 bp from 5' toehold, and 3 bp from 3' toehold is depicted). 5' and 3' toeholds are attacked by 5' and 3' invaders I_3 and I_5 respectively, each of which is able to displace V from VT repairing the mismatch. Product duplexes I_3T and I_5T can equilibrate via toehold exchange. A thermodynamic bias is implemented by adding one of the invaders in greater excess. Both invaders have a fluorophore at the end opposite their toehold and a quencher at the toehold end which should cause an increase in fluorescence when an invader is bound to the target stand.

the duplex. As both invaders are able to repair the mismatch, they will both produce the same number and type of base pairs when bound to the two-toeholded target, and therefore near-identical free energy of product duplex. Equilibration between the two product duplexes I_5T and I_3T can be achieved by toehold exchange [39]. To favour a particular invader at equilibrium, the excess of that invader is increased relative to the other.

Instead of using an external reporter for these experiments, the two invading strands were labelled with a fluorophore at the end opposite their toehold domain, and a quencher at the toehold end. The fluorophore quencher pair for I_5 was 5' Cy5 and 3' Iowa Black RQ, and the pair for I_3 was 3' Cy3 and 5' Iowa Black FQ. Each invading strand is partially quenched in solution [151] due to the shorter end-to-end distance of ssDNA compared to dsDNA. Once an invading strand binds to the target strand T, the fluorophore and quencher separation increases and intensity of the fluorophore associated with that invader increases.

Mismatches of different types were either repaired, or introduced at positions 3 bp from one toehold and 18 bp from the other. In an experiment, the two invading strands were mixed, with the invader intended to be kinetically disfavoured by mismatch position at

50 nM, and the invader intended to be kinetically favoured by mismatch position at 25 nM. The higher concentration of the kinetically disfavoured invader provides a thermodynamic bias for T binding this invader at equilibrium. The reaction was initiated by the addition of the two-toeholded substrate duplex VT at 12.5 nM, to the solution containing invaders I_3 and I_5. A schematic for the components and displacement reactions is shown in figure 2.8.

Fitting

The set of reactions are modelled by the ODEs:

$$\begin{aligned}\frac{d[I_5T]}{dt} &= -k_{-5}\left([I_5T]+[I_3T]\right)[I_5T] + k_5\left([I_5]_0-[I_5T]\right)\left([VT]_0-[I_5T]-[I_3T]\right) \\ &\quad -k_{53}\left([I_3]_0-[I_3T]\right)[I_5T] + k_{35}\left([I_5]_0-[I_5T]\right)[I_3T] \qquad (2.5)\\ \frac{d[I_3T]}{dt} &= -k_{-3}\left([I_5T]+[I_3T]\right)[I_3T] + k_3\left([I_3]_0-[I_3T]\right)\left([VT]_0-[I_5T]-[I_3T]\right) \\ &\quad +k_{53}\left([I_3]_0-[I_3T]\right)[I_5T] - k_{35}\left([I_5]_0-[I_5T]\right)[I_3T] \qquad (2.6)\end{aligned}$$

$[I_5]_0$, $[I_3]_0$ and $[VT]_0$ are experiment dependent parameters. We assume no reverse reaction due to output displacing invader from I_XT therefore set $k_{-5} = k_{-3} = 0$. The remaining 4 parameters k_5, k_3, k_{53}, and k_{35} are parameters to be fit by the model.

Prior to fitting, data was zeroed and normalised, and corrected for linear fluorescence drift. Full details of the processing and normalisation are given in appendix C.

ODEs are solved numerically using MATLAB function `ode15s` to find the time evolution of normalised product duplex concentrations $[\hat{I_3T}]$ and $[\hat{I_5T}]$ given normalised initial concentrations $[\hat{XT}]_0 = [XT]_0/[T]_{tot}$ and normalised rate constants $\hat{k}_X = k_X/[T]_{tot}$. A score is given to the solution with this parameter set by calculating the difference between the numerical solution and the normalised data points for $0 < t < 15000$ seconds, weighted logarithmically between 10 and 1, to give greater weight to early time points. The Euclidean lengths of the weighted difference vector for both wavelengths are summed to give the score. An extra penalty of 10^{51} is added if any parameter is negative, or if $\left|\log(\hat{k}_{53}/\hat{k}_{35})\right| > 3$ i.e. rates cannot be negative and equilibration rates forwards and backwards toehold exchange cannot differ by more than three orders of magnitude. The score is minimised by varying the normalised rate constants using MATLAB function `fminsearch`. Initial values for normalised rate constants were $\hat{k}_5 = \hat{k}_3 = \hat{k}_{53} = \hat{k}_{35} = 2 \times 10^5$ s^{-1}. Absolute rate constants are

recovered by multiplication by $[OT]_0 = [T]_{tot}$, to give k values in $M^{-1}s^{-1}$.

2.7.1 Mismatch Repair vs Mismatch Introduction

By placing a mismatch asymmetrically, at position 18 from the 5' toehold ($\equiv$ position 3 from the 3' toehold), we would expect displacement by attack of the 3' toehold by incoming strand I_3 to occur far more rapidly than displacement from the 5' toehold by incoming strand I_5. The two product duplexes (I_3T, I_5T) would eventually equilibrate, but on a longer time scale than the 3' mismatch repair because mismatch repair increases reaction rate. In figure 2.9a, the substrate duplex is challenged with a twofold excess of I_3, the kinetically favoured invader, and a fourfold excess of I_5, to shift the equilibrium towards I_5T.

We observe the expected transient behaviour for mismatch repair T–C $\rightarrow$ T–A at this position. There is an initial sharp rise in Cy3 fluorescence as I_3 rapidly displaces output strand V from the substrate duplex VT repairing a position 3 mismatch. Cy3 fluorescence reaches a peak before falling again as I_3 is itself displaced fom product duplex I_3T by I_5 to reach equilibrium. This equilibration process is no-mismatch toehold exchange. We would expect it to occur at a much slower rate than position 3 mismatch repair with the same toehold [39]. The Cy5 signal rises more slowly than the Cy3 signal. The rise in Cy5 signal is a combination of two processes, first the displacement of V from VT, and secondly the displacement of I_3 from I_3T. Both processes have similar rate constants but the displacement of V is completed by early times, primarily by the other incoming strand I_3. After the majority of V is displaced, changes in Cy5 intensity will be due to the equilibration process. The duration of the pulse is determined by the equilibration rate.

Fit results confirm that the observed pulse morphology is due to increase in k_3 by more than tenfold compared to the rate without a mismatch repair (see figure 2.9c and table 2.2), whilst k_5 was not significantly changed. Equilibration toehold exchange second-order rate constants k_{53} and k_{35} were both of similar order of magnitude $\sim 10^3$ $M^{-1}s^{-1}$. This is approximately five fold slower than the rate constant predicted by reference 39 of $\sim 5 \times 10^3$ $M^{-1}s^{-1}$ for toehold exchange with two 4 nt toeholds, however it is of similar magnitude to the no mismatch TMSD rate recorded in our earlier experiments of $\sim 2 \times 10^3$ $M^{-1}s^{-1}$.

When the kinetic asymmetry favouring I_3 was attempted by G–C → G–G mismatch introduction at position 3 from the 5' toehold (c.f. reference 124) in figure 2.9b we observe a dramatic slow down to effectively zero for the displacement of V from VT with a position 3 mismatch introduction by the incoming strand I_5. All displacement of V from VT occurs via the incoming strand I_3 which introduces a position 18 mismatch. We find mismatch repair two-toehold competition produces a sharper, more rapid peak than mismatch introduction. Peak lifetime is determined by equilibration processes which have the same rates in both cases, but peak height is determined by how quickly the kinetically favoured incomer can displace V from VT, and mismatch repair is faster than mismatch introduction.

Similar results were obtained when attempting to bias the kinetics in favour of I_5 while thermodynamically favouring I_3T at equilibrium with a greater excess of I_3. Results for a mismatch repair C–C → C–G, 3 bp from the 5' toehold are shown in figure 2.9d. We do achieve the tenfold speed up of the 5' invader relative to the no mismatch case (see figure 2.9f and table 2.2). We also observe a twentyfold slowdown of k_3 when a mismatch G–C → G–A is introduced 18 bp from the 5' toehold and 2 bp from the 3' toehold (figure 2.9e). Despite the expected changes in rate constants consistent with our earlier findings, the shape of the fluorescence traces does not show a peak similar to those observed in figures 2.9a and 2.9b because the general rate of 5' invasion without a mismatch is slower than 3' invasion.

Mismatch		Position (bp)		Fitted rate constant (M^{-1} s^{-1})				Figure
		5'	3'	k_5	k_3	k_{53}	k_{35}	
T–C → T–A	Repair	18	3	5.6×10^3	1.3×10^5	1.0×10^3	8.9×10^2	2.9a
G–C → G–G	Intro	3	18	9.2×10^{-4}	8.2×10^3	1.1×10^3	7.0×10^2	2.9b
	None	N/A	N/A	2.7×10^3	7.6×10^3	7.4×10^2	8.5×10^2	2.9c
C–C → C–G	Repair	3	18	2.4×10^4	1.4×10^4	7.8×10^2	9.5×10^2	2.9d
G–C → G–A	Intro	18	3	1.9×10^3	3.8×10^2	6.5×10^2	8.4×10^2	2.9e
	None	N/A	N/A	2.6×10^3	7.9×10^3	9.2×10^2	6.5×10^2	2.9f

Table 2.2: Fitted rate constants for two-toehold displacement experiments from figure 2.9.

5' vs 3' toeholds

It is apparent from the raw data and the fitted rate constants in figure 2.9 and table 2.2 that even in the absence of any mismatch there is a kinetically favoured invader $k_3 > k_5$

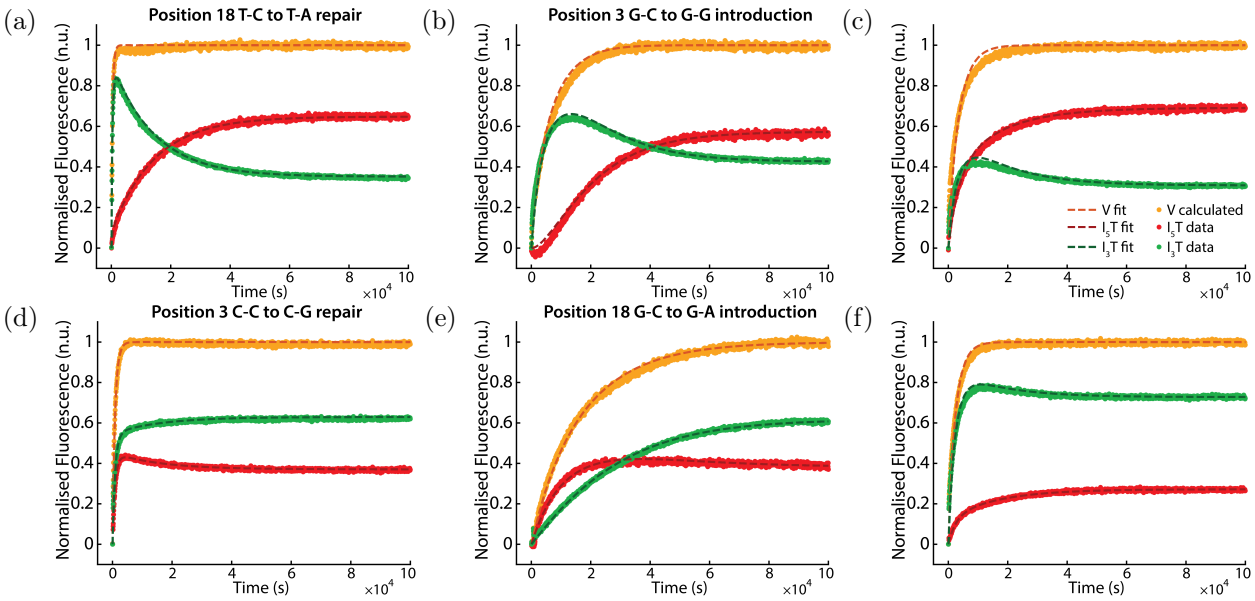

Figure 2.9: Two-toehold kinetic data and fits. Raw data is shown in figure C.2. Relative component concentrations $I_3 : I_5 : VT$ are 2:4:1 in **a-c** and 4:2:1 in **d-f** where [VT]=12.5 nM.

and the ratio of rate constants $k_3/k_5 \sim 3$. Two-toehold displacement experiments were attempted with invaders where the displacement domain had been reversed to determine if this effect was sequence dependent. If the result were sequence dependent then reversing the sequence should switch the invader which is kinetically preferred in the absence of a mismatch. Reversing the displacement domain sequence did not, however, change the invader which was kinetically preferred, though the ratio of rate constants for TMSD with no mismatch was reduced to $k_3/k_5 \sim 1.5$ indicating a smaller rate discrepancy between the invaders. We conclude that the effect is partially sequence dependent.

An explanation for the difference in rate of 3' and 5' invasion is that the two invaders have different terminal modifications. The differences are both in the fluorophore/quencher chemical structure, the part of the fluorophore which is proximal to the DNA, and the additional linker required when the modification is to the 3' end could each have an impact on stabilities of hybridized or unhybridized states, or of intermediate states during displacement. It is possible that these differences are sufficient to change the relative affinities of each invader for its toehold and therefore change the observed rates. Despite the 5'/3' asymmetry, rates can be controlled by altering mismatch position as demonstrated in figure 2.9.

2.8 Summary

We have demonstrated that the position of a mismatch repaired by toehold-mediated strand displacement is crucial in determining the resulting reaction rate, and that reaction rates vary with mismatch position by two orders of magnitude for the 4 nt toehold tested. We note that the change in reaction rate with mismatch repair position is non-monotonic and shows a maximum at a position 3 bp from the toehold. Beyond this position, further from the toehold, displacement rates decrease exponentially with increasing distance of the mismatch from the toehold, reaching rates similar to displacement rates without a mismatch beyond position 13. We can explain these observed changes in terms of displacement and repair probabilities, aided by oxDNA simulations which enable us to compute such probabilities. We interpret the overall decrease in mismatch repair rate with increasing distance from the toehold as a result of the fact that mismatch repair is less likely to occur if the repair site is further away, but once the mismatch has occurred branch migration is almost guaranteed to ultimately displace the victim strand before the invader dissociates.

Because mismatch repair processes generally increase reaction rates compared to those without a mismatch, we are able to produce sharper, more rapid peaks of a kinetically favoured product than by using the introduction of a mismatch to create kinetic asymmetry.

These results may be useful in the design of other dynamic DNA devices, such as the bipedal walker investigated in chapter 5, where mismatch introduction and repair processes occur during the motor operational cycle.

Chapter 3

Tightropes

In this chapter I explore the design and synthesis of an DNA origami structure capable of holding single-stranded DNA under tension as a prerequisite for the extended operation of a DNA biped studied in chapter 5. The structure created is an origami tightrope, named for its macroscopic circus-residing counterpart also designed to bear a biped walker. This structure also forms the basis for the studies of hybridization kinetics under tension in chapter 4. I present the design for the tightrope origami, I then refine a cloning and characterisation pipeline to identify bacteriophage which contain the DNA sequence I wish to hold under tension. Finally I attempt to verify the folding of the origami and the tension across the DNA of interest by imaging techniques and bulk FRET.

Author contributions

The work in this chapter was carried out exclusively by myself except for TEM micrographs, which were acquired by myself or Robert Schreiber (University of Oxford UK).

3.1 Motivation

Unlike many DNA motor designs [3, 55, 56, 62, 92, 105, 106], the design for the bipedal motor studied in this thesis requires a continuous single-stranded DNA track of repeating foot-binding sites [1, 2] (see figure 3.1a and 5.2 for illustration). This is distinct from a track of many single-stranded anchorages which can be added to a DNA origami structure for rigidity and spacing with ease by simply appending staple sequences with anchorage domains [87,106,109]. To allow the motor to operate as designed, walking stepwise along the

track, the foot-binding sites must be well separated to prevent feet sticking to non-adjacent sites and stalling the motor [123] (see chapter 5). A single strand of DNA is a flexible hetero-polymer well-approximated as a freely jointed chain (FJC) [16]. The expectation for the RMS distance between polymer ends using this model is $\sqrt{Nlb} \approx 0.87\sqrt{N}$ nm where N is the number of nucleotides in the chain, $l \approx 0.5$ nm is the length of one nucleotide, and $b \approx 1.5$ nm is the Kuhn length or length of a chain segment. This gives an average separation between foot binding sites (which occur at intervals of 16 nt) on an unstructured track of 3.5 nm. The feet of the motor are separated by a 16 bp duplex ≈5.4 nm in length, so a motor would be expected to easily bind non-adjacent sites on a free ssDNA track. To date this has limited the studies of the bipedal motor to, at most, a single step [1, 2, 152].

The end-to-end length of an FJC can be increased by applying tension, up to a maximum of the contour length of the polymer Nl. I propose to use a DNA origami structure [74, 76] to extend the single-stranded DNA track.

3.2 Previous Work

3.2.1 DNA under tension

Tension can be applied to ssDNA in different ways. The simplest is to apply tension directly to the single strand. In order for tension-induced failure there would have to be a break in the ssDNA backbone. Tension can also be applied to the strand indirectly by pulling on a coupling strand, hybridized to the ssDNA at one end. In these cases failure would occur when the base pairs between the coupling strand and the stretched strand break. The failure force depends on the hybridization geometry, for example, if the coupling helix was aligned with the applied force all of the base pairs would need to shear, requiring ∼50 pN [68–72]. If the helix was anti-parallel to, or perpendicular to, the direction of force application then base pairs could unzip one-by-one, requiring only ∼15 pN tension [70, 71, 73].

Tension imposed by the presence of a rigid rod opposing the tension (e.g. a DNA helix or helix bundle) is subject to failure due to the phenomenon of Euler buckling [153, 154]. This is the threshold force beyond which any additional force will lead to buckling rather than

compression of the structure. It is a function of the persistence length of the rod l_p, and the contour length of the rod L given by equation 3.1.

$$F_c = k_B T l_p \left(\frac{\pi}{L}\right)^2 \tag{3.1}$$

DNA origami has been shown to fold against tension in single-stranded sections [113]. As we discovered in chapter 1, Liedl and colleagues constructed tensegrity origami structures with multiple rigid helix bundles held in a specific 3D arrangement because of entropic string ssDNA segments pulling against each other, and demonstrated that the structures could fold against up to 14 pN of tension, beyond which buckling was observed. The energy required to stretch the ssDNA segments comes from the free energy of hybridization from base pairs holding the scaffold in place. This demonstrates that DNA origami can fold against tensions in the range which should influence bipedal motor operation, as simulations suggest tensions of 7.3-14.6 pN increase motor forward stepping bias [123].

3.2.2 Tracks for a DNA biped

The longest track so far used to study this type of motor was made of three sites with the central site binding an immobile foot. This track was circularised by ligation and braced by hybridizing a bracing oligonucleotide to make a double-stranded region opposite the track [152] (see figure 3.1b). This method is only effective for tracks with a short number of binding sites as it is limited by the rigidity of the double-stranded segment, a duplex with persistence length $l_p = 50$ nm [16, 155]. For a three-site track, Euler buckling will prevent tension application of more than 7 pN using a single duplex, which is already below the tension required to stretch ssDNA to the length of dsDNA (see appendix E). For a 6-site track with the same constraint, Euler buckling would occur beyond 4 pN. The production of ligated circles to use as tracks is also low in yield, reportedly 5% of reactants form the desired product [152], though yield would be expected to vary with the size of circle. To produce longer tracks, a different method is required, and higher track yield would be desirable.

An attempt was made to construct a track for the biped piece-wise along a 2D origami tile [156]. Each repeating unit of the track was attached to the tile by appending both

Figure 3.1: Previous tracks for a bipedal motor. **a** Generic track is a single strand of DNA with repeating domains. **b** Braced three site track [152]. **c** Piece-wise track on 2D origami tile [156]. Each track segment has two (grey) staple ends (see zoom view in red).

5' and 3' ends with staple sequences such that each 16 nt track segment was stretched a distance of 6.2 nm between the two staples producing ∼12 pN track tension (see figure 3.1c). 3 bp were hoped to form between neighbouring track segments to form a continuous track. Unfortunately the correct integration of such staples into the origami could not be confirmed. Additionally, the motor was found not to operate as intended on such tracks [156]. It was hypothesised complications were to do with topological issues associated the piecewise nature of the track as the motor foot would straddle two track segments when bound fully, which might impede foot lifting with fuel.

3.3 DNA Origami Tightrope Design

The DNA origami tightrope was designed with the aid of caDNAno software [78]. I opted for a honeycomb cross-section where each helix has up to three adjacent helices and possible crossovers between helices occur at intervals of 7 bp.

The tightrope is a pseudo-prism, 85 nm long with an approximately triangular cross section (see figure 3.2d) with the 35 nm single-stranded domain along one edge. The face opposite the edge with the single-stranded domain is flat with a large surface area to enable immobilisation to a surface for microscopy. It is depicted by orange helices in figure 3.2b-d. This face of the structure must also oppose any Euler buckling due to the tension of the single-stranded domain. It is composed of a row of 6 helix bundles which individually have been shown to exhibit persistence lengths of 1-2 μm [113, 157]. Over the length scale 35 nm, the length of the single-stranded domain, a single six helix bundle would be expected to buckle at more than 50 pN [153]. This is greater than the force required to shear apart

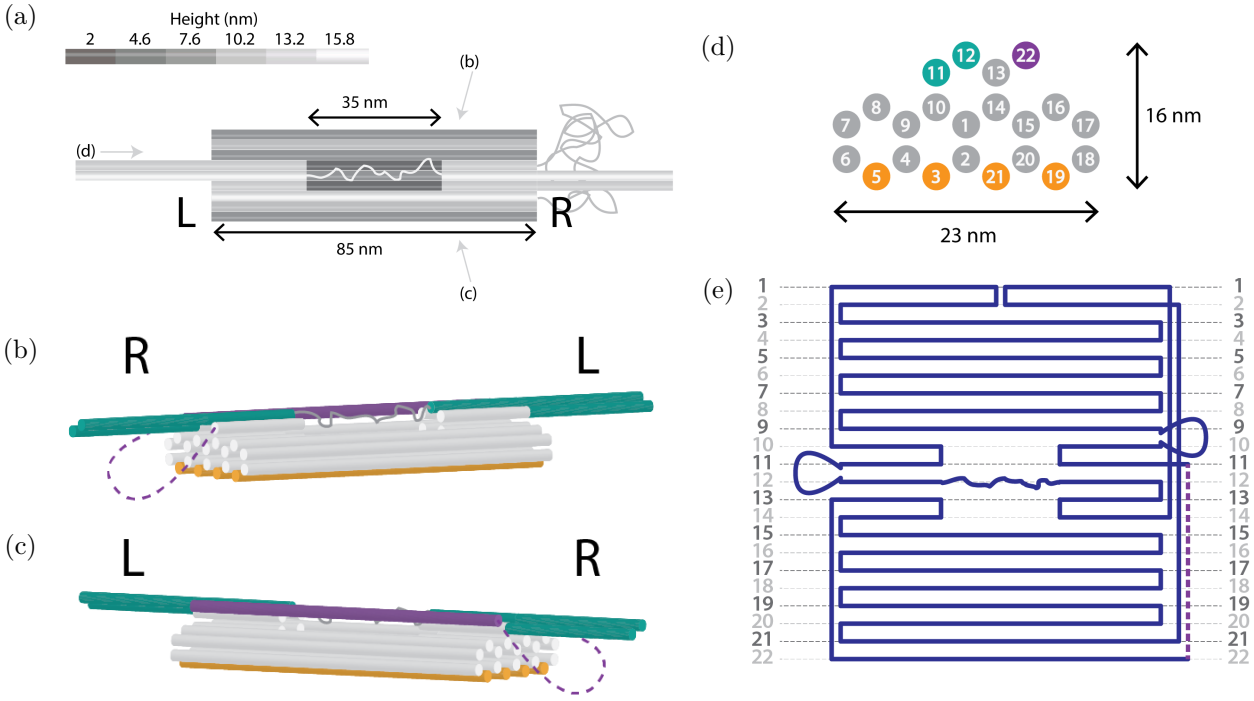

Figure 3.2: Design of "Tightrope" DNA origami. **a-d** show the structure viewed from different directions. The left hand side of the structure is labelled L and the right hand side of the structure is labelled R. **a** Top view of the structure coloured from dark to light grey according to vertical height. The single-stranded DNA track spans a crater in the origami and has raised regions holding it in place at either side. **b** Back view of the structure. From this orientation the single-stranded track region across the gap is not obscured by the scaffold routing helix (purple cylinder). **c** Front view of the structure. The single-stranded track region is obscured from view by the scaffold routing helix. **d** End view of the structure from the left hand side. Cross-section dimensions are 16×23 nm. **e** Scaffold routing pattern. Helices are numbered as depicted in 3.2d. Tension setting loops are between helices 11 and 12 on the left hand side and between helices 9 and 10 on the right hand side. 1992 nt ssDNA link non-adjacent helices 11 and 22 represented by the dashed purple line.

DNA base pairs. Using the persistence length of a single 6 helix bundle as a lower bound for the persistence length of the row of 6 helix bundles, the structure would not be expected to buckle under the tension of the single-stranded region (throughout this thesis tensions are never greater than 20 pN except in deliberate misfolds).

The single-stranded domain is part of the scaffold strand, not a part of one or more staple strands. This circumvents issues of incorporation of a stressed staple into a structure. With an excess of stressed staple to scaffold, it would be possible to hybridize two such staples each by only one end and neither staple would be under stress. This phenomenon would be more likely in cases where the intended tension is higher due to the entropic cost of stretching the staple. Even if stoichiometry was correct, with only one staple per origami,

the staple could be bound at one end only. In my design, as the single-stranded domain is part of the scaffold, the only way the structure could fold fully without tension across the ssDNA domain would be if there were a break in the ssDNA backbone in this domain.

The mode of force application will dictate the maximum force applicable to the ssDNA before the structure fails. In this origami the tension will be applied parallel to the helical axis, in shear conformation, rather than perpendicular to the helical axis peeling conformation, as the failure force is ∼50 pN rather than ∼15 pN [68–73].

In the vicinity of the helix bearing the single-stranded track (helix 12 in figure 3.2d), gaps are left lengthwise along neighbouring helices (helices 11 and 13 in figure 3.2d) to reduce potential undesired interactions between these and the single-stranded DNA. To route the scaffold through the structure with these helices missing the central region it was necessary to also remove the central section of helices 10 and 14 (see figure 3.2e).

It is desirable to be able to change the tension of the single-stranded DNA across the gap without the need for an entirely new set of staples. In order to fold tightropes at different tensions with minimal staple substitutions, loops were introduced between helix 12 bearing the track and the next helix the helix routes on to (helix 11 on the left hand side (LHS) and 13 on the right hand side (RHS)). The loops contain extra DNA which can be spooled into the gap to reduce the tension; alternatively, excess ssDNA in the gap can be brought into the structure by increasing loop size to increase the tension. The loops can be either single-stranded DNA or double-stranded two-helix bundles. The latter is used in the majority of this thesis but for simplicity the loops may not be depicted as two-helix bundles in figures. The only staple sequences which must be adjusted to change tension are those of staples with any domains on track bearing helix 12. Tension setting domains are cyan in figure 3.2.

An origami scaffold is commonly a single-stranded circle of DNA. The scaffold routing pattern for a tightrope is shown in figure 3.2e. It begins at the RHS of the structure on helix 11 and terminates on helix 22 following the 5'-3' direction of scaffold DNA. The role of helix 22, named the scaffold routing helix, is to return the scaffold to the same end of the origami on which it initiated, otherwise the scaffold would need to stretch across the entire length of

the structure to reach its initiation point. The starting and ending helices are not adjacent. They are separated by approximately 8 nm, therefore, in order to link the helices, 1992 nt single-stranded nucleotides of scaffold are left at the RHS of the structure. This is shown as a purple dashed line in figure 3.2.

Tightropes will be distance clamps and not force clamps. In practice this should provide approximately constant tension on the single-stranded DNA during motor operation where the number of hybridized base pairs on the track oscillates by 16 bp as feet bind and lift (see section 4.3). Indeed, on a track where the number of single-stranded (ss) nucleotides across the gap is equal to the number of duplex base pairs (ds) required to span the gap, the average expectation value of tension should be approximately constant as base pairs hybridize/unhybridize between motor and track. I call this ss=ds length. When DNA hybridizes to the ssDNA spanning the gap there can be a change in ssDNA tension. In cases of higher tension than at ss=ds length, hybridizing to a complementary strand will shorten the end-to-end length of a domain. This means that the remaining ssDNA will need to stretch even further, thus tension is increased by hybridization. In cases of lower tension than at ss=ds length the opposite is true. Hybridizing a complement will increase the end-to-end length of a domain, and tension will decrease.

3.4 Inserting Biped Track into the Origami Scaffold

3.4.1 M13 bacteriophage

In most cases the sequence of the scaffold strand for a DNA origami nanostructure is not user-designed, but fixed by its biological source. In my case, as I require the biped track with a repeating designed sequence[1] to be part of the scaffold strand, this is not sufficient. Traditionally phage or plasmid DNA is used as a scaffold for origami. The DNA must be made single-stranded, if it is not already, for example by an exonuclease. Both phage and plasmid vectors can be cloned to contain a region of user-specified sequence. I opted to use M13 bacteriophage cloning to produce the scaffold for my origami due to its genome size

[1]Further details of the track design and sequence will be presented in chapter 5.

and inherent single-strandedness.

M13 is a filamentous bacteriophage. It has a single-stranded genome, ideal for use as an origami scaffold, encapsulated within a protein coat. The protein coat has a a variable number of copies of coat protein pVIII which allows the size of the genome to be modified by using more or fewer pVIII copies. This is not the case in phage where the genome is compacted within a protein capsid of fixed size such as lambda.

Back in 1974 Joachim Messing set out to introduce restriction sites for cloning in M13 [158]. He saw the appeal of the single-stranded genome of M13 for sequencing applications but this required the ability to clone in sequences of interest. Wild type M13, first discovered in the sewers of Munich (the 'M' of M13 is for Munich) [159], did not naïvely posses these cloning sites. Messing succeeded in adding, first Eco RI [160], and later many other restriction sites [161, 162], to M13 creating the multiple cloning site (MCS) in a series of phage M13mp (where mp strands for Max Planck, the institute where Messing developed the series). He also introduced the gene coding for the α fragment (146 amino acids at the N terminus) of the enzyme β-galactosidase with a lac promoter which allows phage containing this truncated gene to metabolise lactose, if and only if, the bacterial strain infected contains the remainder of the lacZ gene and can produce incomplete 'complementary' β-galactosidase. This process is called α-complementation [163]. When the gene is induced by IPTG, the reconstituted protein is produced which can digest white X-gal, a lactose analogue, to form a blue precipitate. The gene in M13mpX is in the MCS so it will be disrupted by the insertion of additional DNA by cloning. Loss of blue colour can be used to identify M13 clones.

M13 can infect strains of *E.coli* possessing F-type pili for example DH5αF' (New England Biolabs) or JM109: a strain developed by Messing for this purpose [164]. The phage has two forms. When inside a host cell, phage DNA is in a double-stranded replicative form (RF) containing both + and − genome strands. The RF phage can then amplify its DNA by a rolling circle mechanism to produce single-stranded copies of the + strand. The + strand is packaged into progeny bacteriophage and extruded from the cell non-lytically.

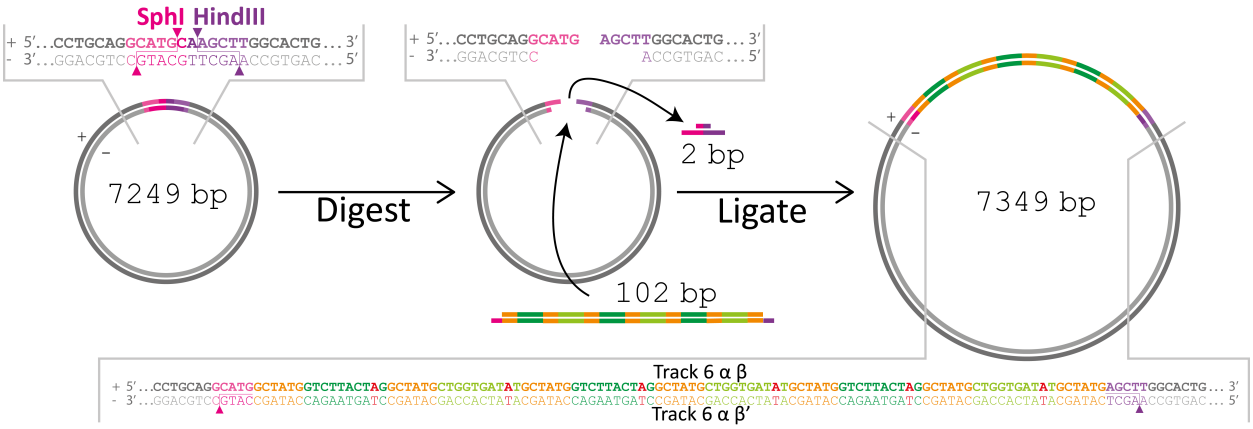

Figure 3.3: Cloning a track insert into M13mp18. Replicative form bacteriophage is cut with restriction enzymes HindIII and SphI. Initially RF DNA is double-stranded, circular and supercoiled. Cleavage produces a long linear fragment with two sticky ends, both on the + strand. 2 nt are lost from the + strand. After dephosphorylation, the linear RF DNA is ligated with a duplex containing the 102 nt track sequence on the + strand, and the track complement sequence with appropriate sticky ends on the - strand, with 5' phosphates on both strands. The ligated product DNA should be able to re-circularise only when the track insert duplex is present.

3.4.2 Cloning

My track insert duplex was formed from two 5' phosphorylated oligonucleotides (Integrated DNA Technologies) hybridized together. The duplex has 2× 4 nt sticky ends on the strand complementary to the track, but the strand with the track sequence has no additional nucleotides. The sticky ends are designed to hybridize to the sticky ends left after M13mp18 RF DNA has been cleaved by two restriction enzymes: HindIII and SphI. Once the insert duplex is ligated into the M13mp18 RF, the product should no longer be cleavable by HindIII or SphI because the restriction sites for these enzymes would both be incomplete.

M13mp18 RF was digested sequentially with HindIII HF[2] and SphI HF (see figure 3.3)[3]. SphI has greater activity close to the end of a DNA duplex than HindIII [165], and as the restriction sites for these two enzymes are adjacent, cutting with one enzyme will leave the restriction site for the other close to an end, with one base pair before the end of the duplex. I therefore cut first with HindIII, and then with SphI. Antarctic phosphatase was used to dephosphorylate 5' termini of restriction products and prevent the restriction products from re-circularising without insert DNA. Dephosphorylated linearised vector DNA was then ligated with insert DNA using T4 DNA ligase (all enzymes from New England Biolabs).

[2]High fidelity.
[3]Cloning attempts with a simultaneous digest failed to produce any M13 with the desired insert.

3.4.3 Phage plaques

5 ng or 500 pg ligated DNA in a volume of 1 µl was transformed into 50 µl competent DH5α *E. coli* (New England Biolabs) by heat shock at 42°C for 30 seconds. A mixture of 40 µl transformed cells and 200 µl lawn bacteria was plated in 3 ml top agar containing 1×LB, 15 mM $MgCl_2$, 6.4 mM X-gal and 2.4 mM IPTG over an M9 agar base. This is 10× the X-gal concentration and 2× the IPTG concentration of published protocols [166, 167]. As M13 is non-lytic, phage plaques are observed as areas of slower growing bacteria in a lawn of denser bacteria, rather than areas of no bacterial growth. Plaques containing unmodified M13 should appear blue whereas plaques where the LacZ gene fragment has been disrupted (for example by the insertion of my track DNA) should be colourless (figure 3.4a).

A variety of phage plaques were picked and grown on a small scale by infecting cultures. Plaques from a positive control of DH5α transformed with M13mp18 RF all appeared a deep blue colour. No plaques (and no blue colour) were observed in negative control plate transformed with no DNA. All plaques from transformations with track ligated DNA were colourless. I picked 45 of these plaques as candidates for insert incorporation.

3.4.4 Phage PCR length tests

I attempted to estimate the length of DNA between two positions on M13 either side of the HindIII and SphI sites with PCR. I used two standard primers, M13 Forward(-41) and M13 Reverse (-48), which bind to the + and – strands respectively, either side of the MCS. In M13mp18 RF there are 149 bp between the binding positions for 5' ends of this pair of primers. After a successful ligation with track insert, the expected length between 5' ends of primers would increase by 100 bp to 249 bp.

PCR can be achieved directly from a picked colony of bacteria e.g. containing a plasmid to be amplified. I hypothesised that PCR could also be achieved directly from phage plaques, or from small cultures of bacteria infected with phage from a plaque. I therefore attempted each of these alongside PCR of ss and ds M13mp18 DNA. I performed PCR with 1 µl plaque infected culture per 50 µl PCR reaction for my screen using the method shown in figure 3.4.

The 45 phage plaques from the track ligated DNA transformation were grown up to

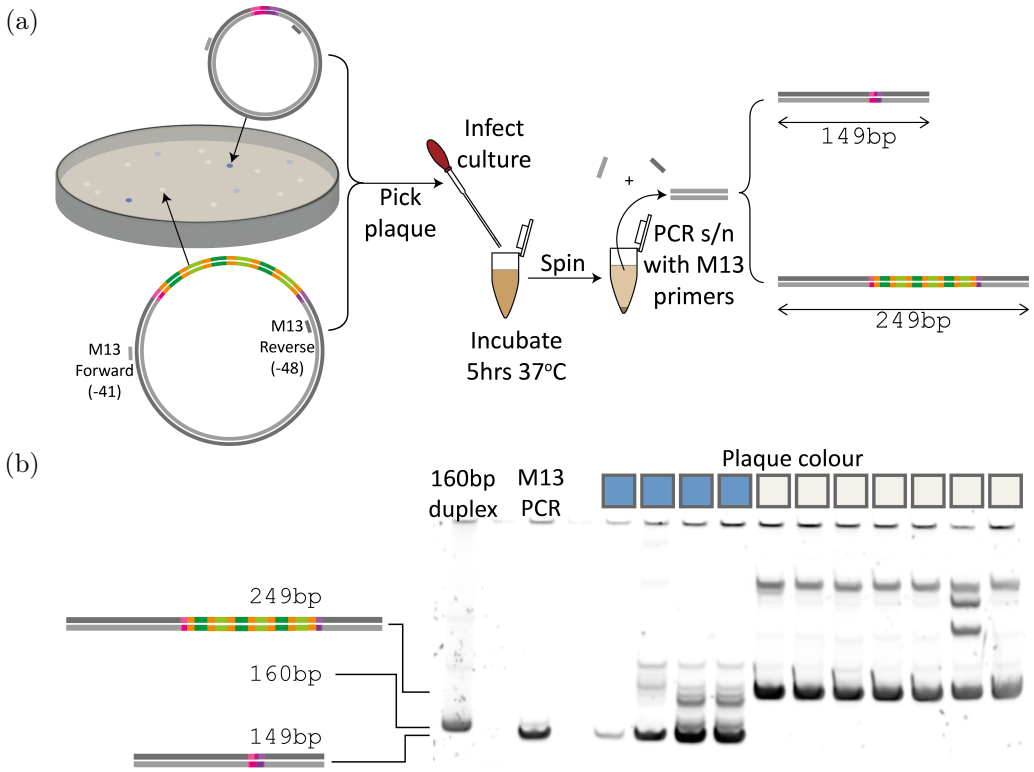

Figure 3.4: Diagnosing cloning success from X-gal/IPTG phage plaques using PCR. **a** Plaques are areas of lower bacterial density in an otherwise uniform lawn of bacteria. Plaques are blue (intact M13mp18) or colourless (M13mp18 with disrupted MCS). Individual plaques are picked and used to infect cultures of bacteria which are then grown for 5 hrs at 37°C. Cultures are spun down and bacterial pellets discarded. 1 µl phage supernatant is used in PCR with M13 forward and reverse primers. **b** PCR product length measured by native PAGE. A 160 bp duplex is used as a length marker (lane 1). A control reaction using 1 ng commercial M13mp18 ssDNA as a template for PCR (lane 3) produces a product duplex slightly shorter than 160 bp. Lanes 5-15 show PCR products from reactions containing 1 µl phage supernatant from blue plaques (lanes 5-8) produced from M13mp18 RF transformed bacteria or colourless plaques (lanes 9-15) produced from track ligation transformed bacteria. Blue plaque PCR product runs to the same position as the M13mp18 PCR. Colourless plaque PCR product runs higher than the 160 bp band, and to the same position as the 250 bp marker on a 50 bp ladder (not shown).

small scale cultures. I then performed the plaque PCR length screen on each. Results for a representative sample of 7 such plaques are shown in figure 3.4b lanes 9-15. Positive control blue M13mp18 plaques were also picked and PCR screened (figure 3.4b lanes 5-8). Polyacrylamide gel electrophoresis confirmed that 42/45 plaques produced PCR fragments of the expected length to show probable insert incorporation.

3.4.5 Phage stock characterisation and sequencing

Phage from 0.5 ml of the phage supernatant (host bacteria had been pelleted and phage containing supernatant retained) from small scale 1.1 ml plaque-infected cultures was used to

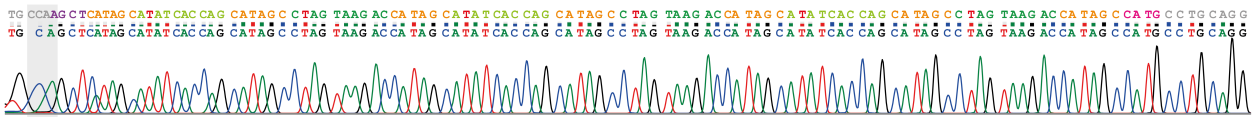

Figure 3.5: Chromatogram except for M13mp18 with attempted 6 site track DNA inserted. Upper sequence shows the expected sequence, and lower sequence, the sequencing result. The only discrepancy between the two is in the early nucleotides, shown in grey, and can be explained by insufficient resolution of pairs of identical consecutive bases.

infect 2.5 ml DH5αF' and grown in 250 ml LB + 5 mM $MgCl_2$. 15 ml phage supernatant from this preparation was preserved as a generation 1 phage stock. Due to the high mutation rate in M13, and the tendency of the phage to excise inserted DNA, phage should not be grown for long periods to reduce the probability of these events having occurred [166]. Sambrook and Russell recommend not growing M13 infected cultures for more than 5 hours, and not growing cultures past the second generation of phage i.e. phage produced from the first 5 hr infected culture can be used to infect and grow a further 5 hr infected culture which would be the second generation, but phage from this second generation culture should not be used to infect more cultures of phage without first reforming plaques and reverifying sequence.

I estimated the number of plaque forming units (pfu) in my generation 1 phage stock by plating in top agar 10 μl of either 10^{-6}, 10^{-7}, 10^{-8} or 10^{-9} diluted phage in 200 μl JM109 lawn bacteria. Plaques were observed only in the 10^{-6} dilution plate. The inferred concentration of phage was $1.6 \pm 1.1 \times 10^{10}$ pfu ml^{-1} in the undiluted stock.

I attempted a second generation growth by infecting 2.5 ml JM109 cultures with 0.1 ml, 0.5 ml and 2 ml generation 1 phage supernatant or 1 ml LB in which a picked plaque from the 10^{-6} generation 1 dilution plate had been soaked. The infected cultures were then grown in 250 ml LB + 5 mM $MgCl_2$. The recommended quantity of phage for infection is 5×10^{11} pfu [166], larger than any of the quantities tested. The total mass of ssDNA extracted from phage grown from these preparations was between 510 μg and 560 μg inferred by 260 nm absorbance. Infecting with more pfu did not produce more ssDNA.

Sequencing results for each of these samples showed track DNA inserted in all cases. The sequence reported contained the exact sequence of the reverse complement of the track . The sequence after the SphI restriction site matched exactly the sequence of M13mp18 expected in this region suggesting that this side has been correctly incorporated into M13mp18. The

sequence after the HindIII restriction site did not match (see figure 3.5), however sequencing data in this region is of poor quality due to the proximity to the binding domain for the primer used (M13 Forward (-41)). Comparing the desired sequence `TGCCAAGCT` to the sequencing result (the same for all 4 sequencing attempts) `TGCAGCT` it seems the duplicated nucleotides C and A have each been read as single nucleotides. Looking at the chromatogram this seems plausible as the C and A peaks are both broad. Furthermore in a sequencing attempt with commercial M13mp18 ssDNA (Affymetrics) and the same primer, the same two nucleotides were omitted in the sequence. From the sequencing results I conclude that the six-site track insert is correctly incorporated into the M13mp18 in the first and second generation phage preparations from phage grown from a particular plaque have the anticipated sequence. This phage DNA will be referred to as M13NH1 throughout the thesis.

3.5 Tightrope Assembly Verification

3.5.1 Agarose gel electrophoresis

The initial attempt to assess folding of tightropes was made using agarose gel electrophoresis. Tightropes were folded with 102 nt ssDNA across the gap, with a 10 or 20 fold excess of staples to M13NH1, in a buffer of 1× TE + 15 mM $MgCl_2$, using a temperature ramp cooling non-linearly from 65°C to 25°C over a period of 16 hrs (see appendix A). When annealed with this excess of staples, the mobility of the band appeared similar the M13NH1 band. It was difficult to conclude good folding from a gel because the origami bands all migrate close to the M13NH1 scaffold band, and a well defined band does not confirm that origami has folded into the desired shape.

3.5.2 Transmission electron microscopy (TEM)

Transmission Electron Microscopy (TEM) was used for high resolution imaging of structures in bands of interest extracted from agarose gel (see appendix A). Samples were deposited on carbon coated copper grids (Agar scientific) and stained in uranyl acetate (see appendix D). The staining process can produce regions of negative staining and positive staining on the same grid.

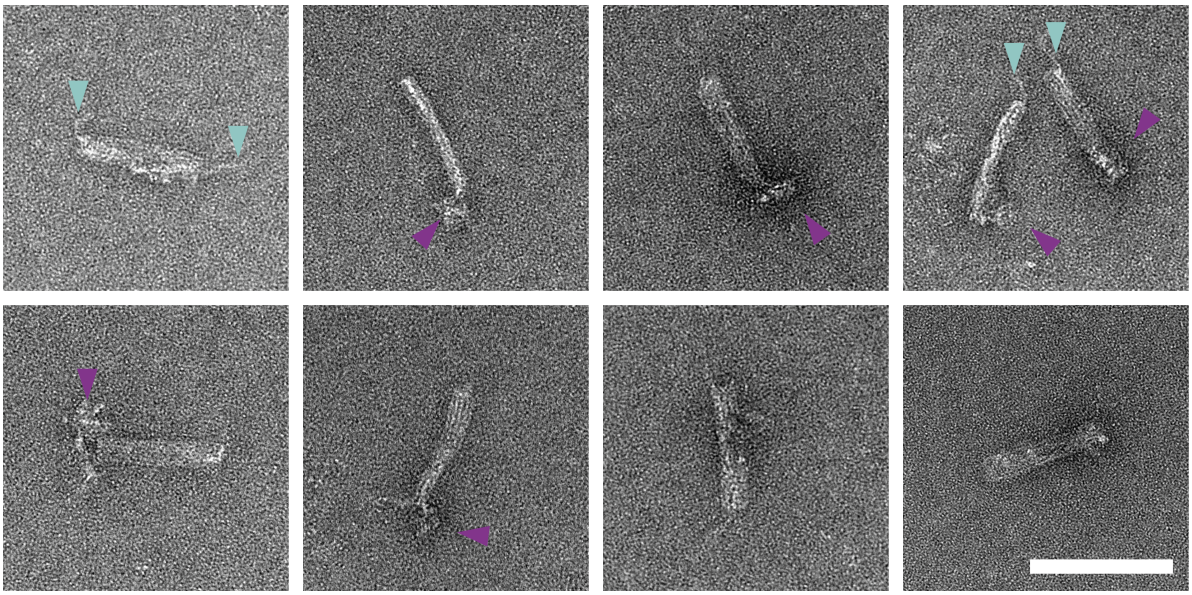

Figure 3.6: Negative stain TEM micrographs of tightrope origami folded with 102 nt ssDNA across a gap of 35 nm. Scaffold routing loop of single-stranded DNA can be observed as disordered pale features and or a dark halo at one end of the structure as indicated by purple markers. Tension setting loops (2× staple linked dsDNA double helices at each side of the structure) can be observed in some of the micrographs as indicated by cyan markers. The majority of structures appear straight suggesting that they are not buckled as a result of the tension. The central gap over which the ssDNA is stretched cannot be clearly observed in these micrographs. Scale bar 100 nm.

TEM micrographs for tightropes folded with 102 nt across the gap purified from agarose gel are shown in figure 3.6. Most observed structures look like straight, multi-helix bundles and have a length of 90±5 nm similar to the length expected from the original design of 85 nm. In many cases the double-stranded tension setting loops can be seen as flexible filaments either end of the body of the structure (cyan markers). The excess single-stranded DNA of the design can be seen in some images as a dark blob on one end of the structure (purple markers). From these images is is hard to see the gap in the centre of the structure over which the single-stranded DNA should be held. This may be because it is partially obscured by the scaffold routing helix, or because there are limited orientations of tightropes where this would be visible anyway (recall figure 3.2b and c). As TEM is a transmission technique, features in the image on top of the structure and on the bottom of the structure are indistinguishable, further restricting the orientations where the gap could be seen.

Tightropes were folded with either 144 nt (figure 3.7b), 102 nt (figure 3.7c), or 2 nt (by using a M13mp18 scaffold instead of M13NH1) (figure 3.7d) across the 35 nm gap. The latter shows the effect of attempting to fold against too high a tension on the morphology of

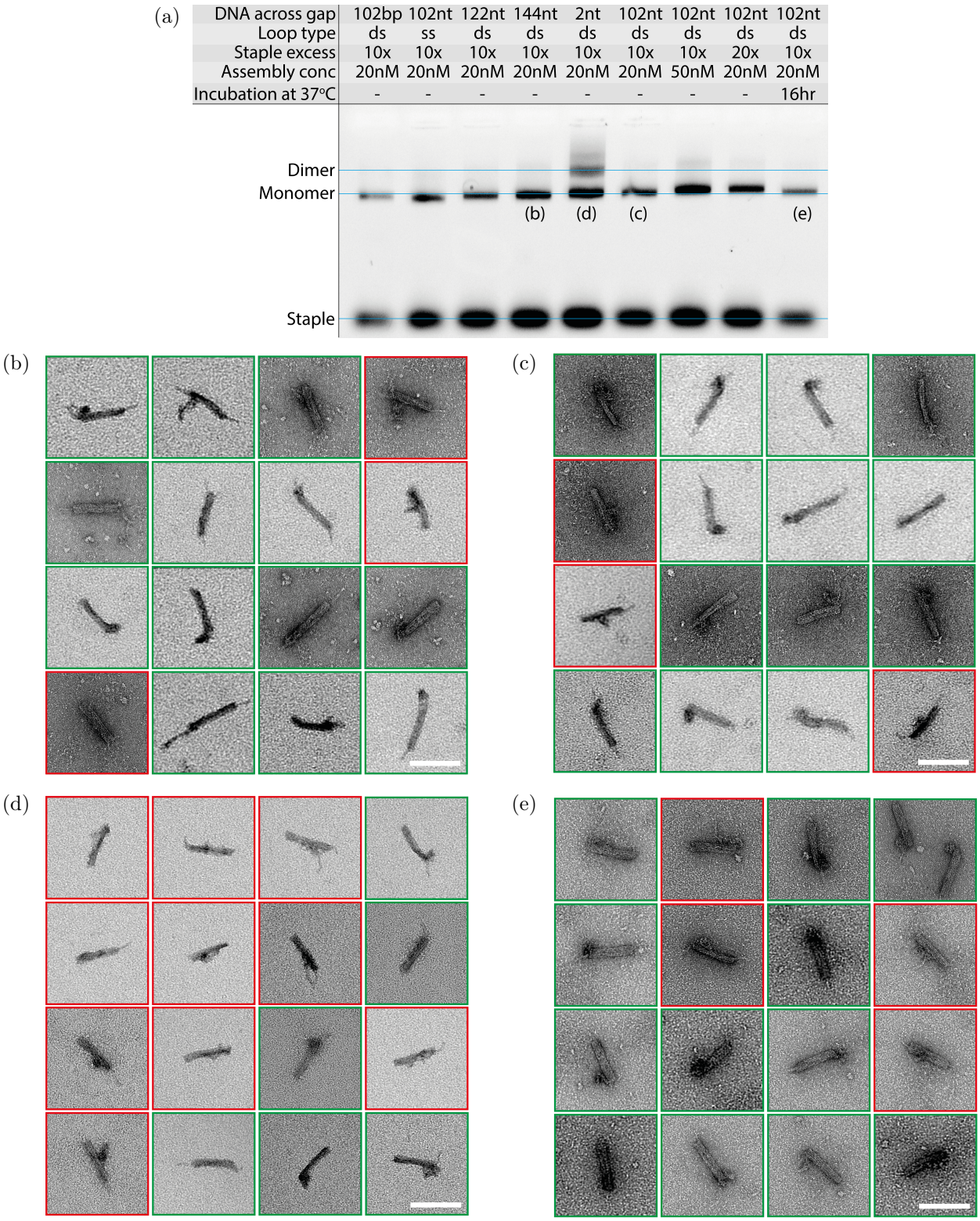

Figure 3.7: Representitive TEM micrographs of origami tightropes. **a** Agarose gel of a variety of Cy5-labelled tightropes. Blue lines give guides for positions of monomer, dimer, and staple bands. Annotations below bands excised from gel for TEM micrographs in **b-e**. Structures assessed by eye for features centrally along the length which may indicate misfolding. Disordered features at the edge are assumed to be excess scaffold loop and are not classified as misfolds. Misfolded structures are outlined in red and well folded structures are outlined in green. Structures folded with **b** 144 nt, **c** 102 nt, **d** 2 nt across the 35 nm gap. **e** Tightropes folded with 102 nt across the 35 nm gap incubated for 16 hrs at 37°C. Scale bars 100 nm.

the structure as it is impossible to fold fully with only 2 nt ssDNA to stretch across the gap. Tightropes were purified from agarose gel (figure 3.7a), and imaged by TEM. Tightropes folded with 102 nt across the gap which had been incubated for 16 hrs at 37°C were also imaged to determine if long incubations at elevated temperatures used for enzymes would cause changes to the structure. Structures were assessed by eye for evidence of misfolding. The presence of the excess ssDNA loop at one end of the structure complicates assessments of folding quality because it is flexible and disordered and may appear similar to a misfolded section of the nanostructure. The position of excess ssDNA loops should be distributed about one end of the structure. In my assessment of good folding I therefore chose not to discount structures with an asymmetrically positioned disordered region as badly folded, but to mark any structures with a centrally positioned disordered region, or multiple disordered regions, as misfolded. These analyses, though they may not tell me with certainty if an individual structure is well or badly folded, should give me a metric by which to compare samples. In figure 3.7 possible well folded structures are outlined in green, and possible misfolds in red. There are more structures which appear misfolded in the sample with 2 nt to span the 35 nm gap (figure 3.7d) than in other tested samples. This is expected as spanning the gap with so little ssDNA is impossible. There is no TEM visible deterioration of structures after incubation for a week at 37°C (figure 3.7e). This is useful as the structures should be able to withstand these conditions when attempts are made to operate motors on the tightropes.

Folding quality can only be crudely assessed by TEM due to issues in the ability to observe certain features of the design, such as the gap across which the ssDNA is stretched, and the ambiguity of the position of the excess ssDNA scaffold.

3.5.3 Atomic force microscopy (AFM)

Atomic force microscopy is often used to image two-dimensional DNA nanostructures such as tiles, and shapes of only single helical thickness. In some cases it can also be used to look at three-dimensional DNA nanostructures particularly if the structures are reasonably flat, though usually image quality is lower than for the two dimensional counterparts.

Atomic force microscopy in tapping mode in fluid was attempted with a variety of

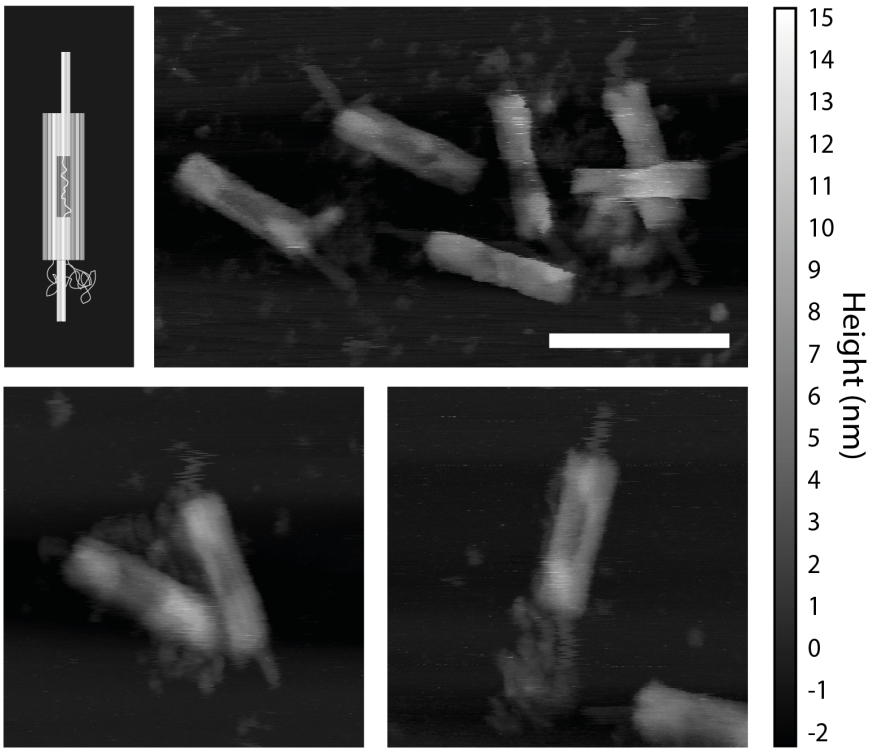

Figure 3.8: AFM images acquired of tightrope origami with 102 nt across the 35 nm gap with 200 bp dsDNA loops either side. Top left shows design schematic for comparison. Scale bars 100 nm.

tightrope origami purified by PEG precipitation (see appendix A). Samples were deposited on a mica substrate and imaged in fluid in AFM buffer. All samples show the tightropes are generally mono-disperse and tend to stick to mica by the flat base of the structure. The structure was designed such that it would stick by its base to surfaces for fluorescence microscopy but it is serendipitous that the large flat base also interacts preferentially with the mica. In the AFM images, for tightropes stuck base down, the height of the structure could show the central crater over which the ssDNA is stretched surrounded by higher, tension bearing helices on either side. In the highest resolution images obtained, shown in figure 3.8, individual double helices, including the scaffold routing helix, could be seen. The double-stranded tension setting loops can also be seen at either end of the structure.

AFM was performed on tightrope origami with a range of different tensions. Images of samples prepared with 94 nt, 102 nt and 144 nt across the gap are shown in figure 3.9. Morphological features were picked out by markers in these images. Cyan markers point to tension setting helices extending beyond the main structure. The variable lengths of these helices (required to set different tensions across the gap) can be seen. All samples have a

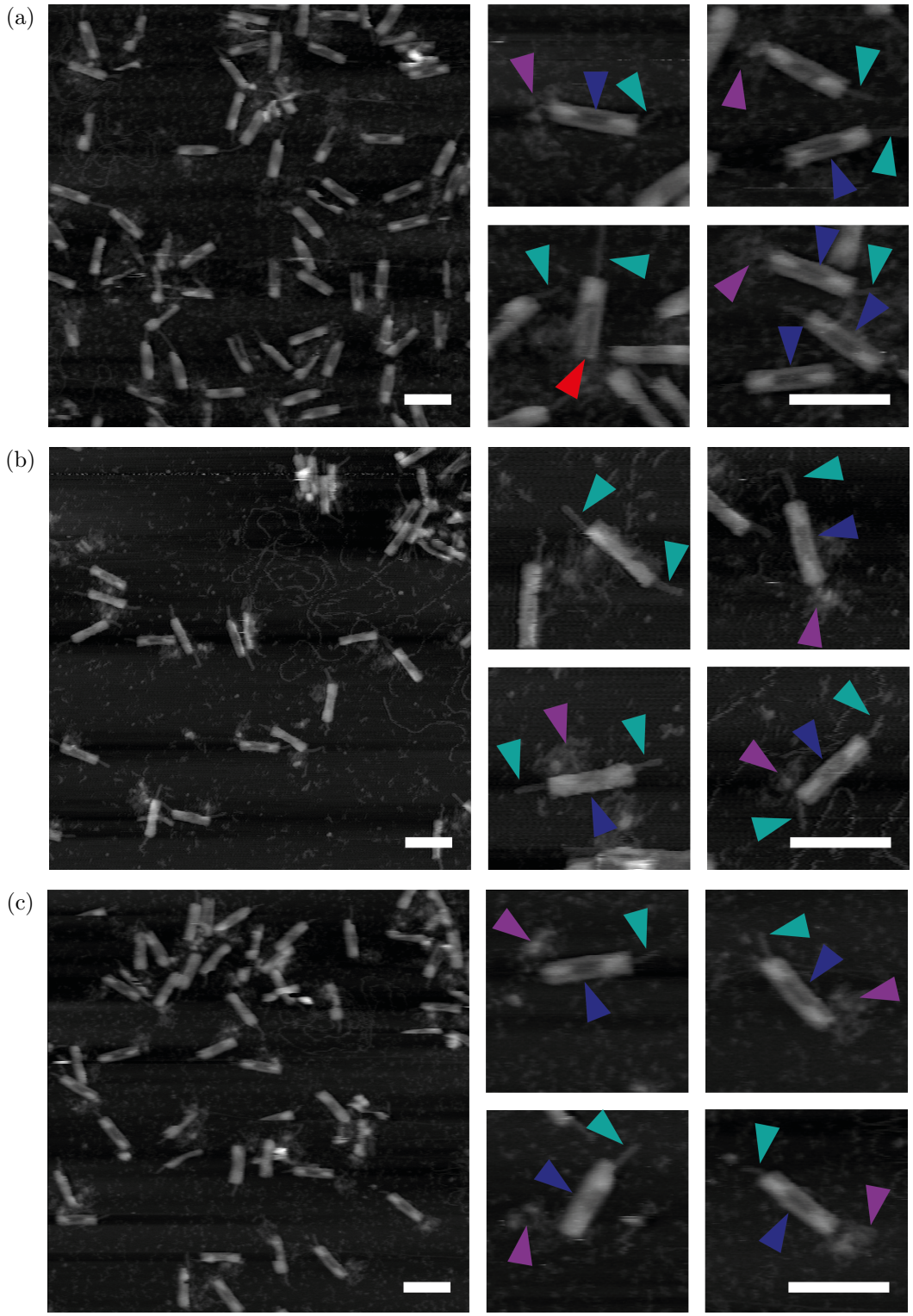

Figure 3.9: AFM images of tightrope origami folded with **a** 94 nt, **b** 102 nt, and **c** 144 nt spanning the 35 nm gap. Right hand panes have features labeled with markers; Red = misfold motif where the top of one side of the structure does not form, Blue = central gap over which the ssDNA is stretched, Cyan = tension setting loops, Purple = excess ssDNA scaffold. Scale bars 100 nm.

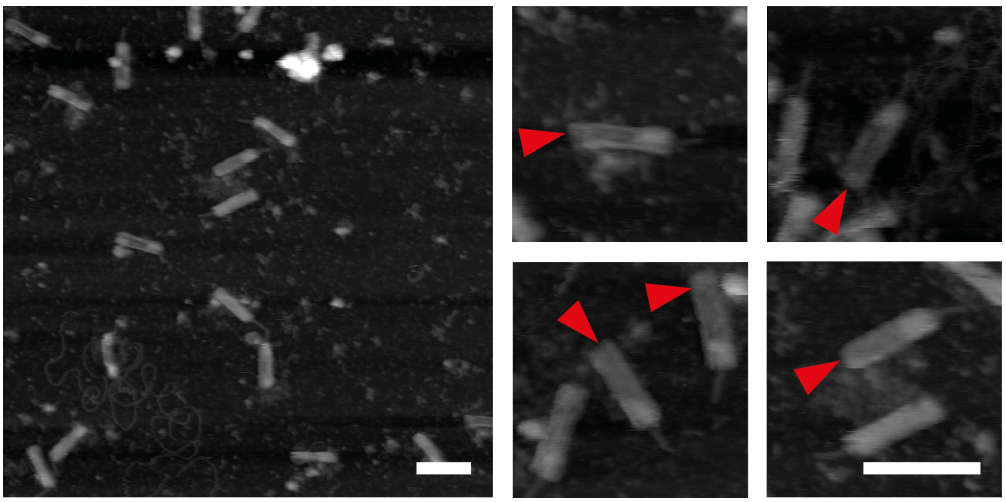

Figure 3.10: AFM images of tightrope origami with RHS tension bearing helix staples omitted. Red markers indicate misfold where the top RHS of the structure does not form. Scale bars 100 nm.

LHS loop of 188 bp but the 94 nt sample in figure 3.9a has a RHS loop of 104 bp, the 102 nt sample in figure 3.9b has a RHS loop of 196 bp, and the 144 nt sample in figure 3.9c has a RHS loop of 54 bp. Purple markers show excess single-stranded scaffold on the RHS and blue markers mark the central indentation over which the ssDNA should be stretched. A misfolded motif is shown with a red marker. This misfold is where one side of the tension-bearing top of the structure is not present, and can be produced deliberately by omitting all of the staples which hold the scaffold in place on one of the sides. AFM images of such a sample are shown in figure 3.10 and the majority of tightropes appear similar to the tightrope marked red in figure 3.9a. This misfold is only observed in the 94 nt case which is the highest tension assembly. It is likely that the misfold is due to the high tension.

3.5.4 Bulk Förster resonance energy transfer (FRET)

Imaging methods alone cannot confirm the correct folding of the structure, particularly as none of the imaging methods attempted were able to show the single-stranded DNA across the gap. Even if the bulk of the structure appeared unambiguously well-folded, this would not preclude the possibility that the ssDNA is discontinuous. I needed a method to test the integrity of the single-stranded DNA in the gap itself to demonstrate that the structure was functioning as designed and holding a length of ssDNA under tension.

In this thesis I use a freely jointed chain model with extensible segments (mFJC) from reference [16] to calculate estimated tensions (see appendix E). This model, and its short-

comings, will be discussed in chapter 4. In this section, I assume that the extension per single-stranded nucleotide is equal, which may not be the case due to secondary structures.

Fluorescence resonance energy transfer (FRET) techniques allow quantitative measurements of distances on the scale of the Förster radius R_0 between a donor and acceptor fluorophores. Förster radii for most pairs of organic fluorophores are on the scale of 6 nm, a convenient length scale for DNA constructions as 6 nm corresponds to a separation of approximately one 18 bp duplex. Folding my structure under different tensions should change the separation of a given pair of nucleotides along the ssDNA region.

Experimental method

For FRET experiments, I use an M13mp18 scaffold rather than M13NH1 scaffold with a repeating track, as the M13mp18 sequence stretched across the gap is uniquely addressable.

Tightropes with different ssDNA lengths across the gap were annealed with two 'bookend' strands which bind to domains on the stretched ssDNA section and provide domains for hybridization of complementary probe strands appended with FRET donor or acceptor fluorophores (figures 3.11b and 3.11c). Between the bookends there is a single-stranded domain. A domain length d=24 nt was chosen to give high FRET in the lowest tension case, an unfolded structure, and to give low (but observable) FRET for high tightrope tensions.

Experiments were performed in bulk with probes initially single-stranded in solution, each at 15 nM concentration. The donor fluorophore was Alexa546 and the acceptor fluorophore was Alexa647. Fluorescence intensity at Alexa647 excitation and emission (I_{AA}), Alexa546 excitation and emission (I_{DD}), and FRET at Alexa546 excitation / Alexa647 emission (I_{DA}) were recorded (see table C.2).

After recording baseline unhybridized probe fluorescence, tightrope origami with bookends was titrated into the probe solution in steps of 5 nM (figure 3.11b). This was to encourage the hybridization of both probes to each origami (figure 3.11c). At the point at which the addition of origami caused no further change in FRET, additions were stopped, and a spacer strand complementary to the ssDNA between the bookends was added in excess (figure 3.11d). Hybridization of this strand to any tightrope should separate bookends by the

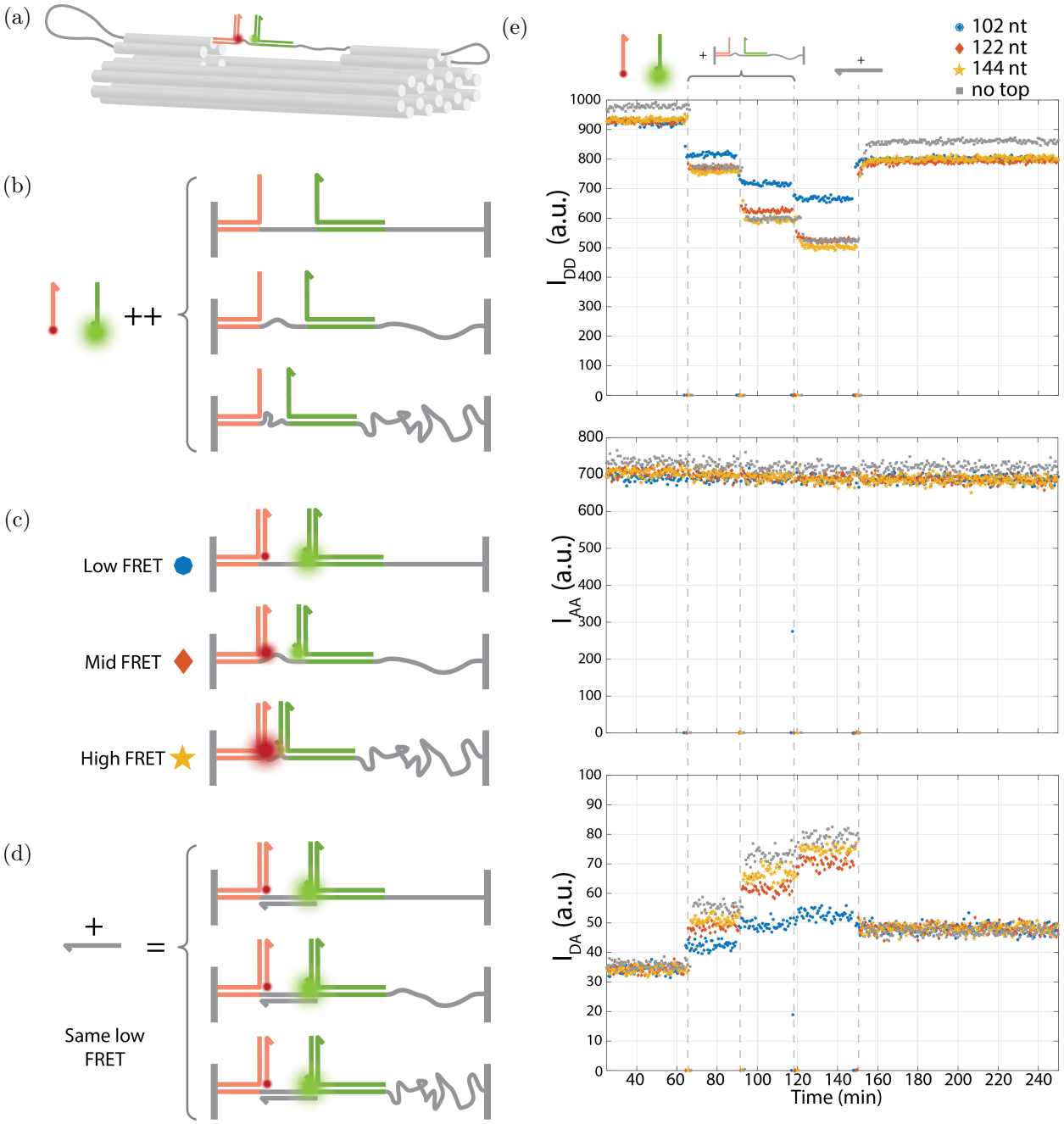

Figure 3.11: Bulk FRET test of track tension. **a** Diagram of 'bookend' coupling strands, used for bulk FRET, on a tightrope. **b-d** Schematic of experimental procedure. **b** Tightrope region for three different tightrope origami annealed with bookends. Tightropes were slowly titrated into a solution containing bookend-complementary donor and acceptor probes which **c** hybridize free domains on the bookends. **d** After tightrope addition causes no further fluorescence change, a 24 nt spacer duplex was added in excess to hybridize between the probes and transforming all tightrope types to the same FRET state. **e** Experimental traces for 102 nt (blue circles), 122 nt, (red diamonds) and 144 nt (yellow stars) across the 35 nm gap. Grey squares show the same experiment for an origami folded without any tension bearing helices and hence with ∼530 nt spanning the 90 nm structure (∼3 pN). Both donor quenching (top) and acceptor FRET (bottom) show the expected trend of increasing FRET as number of nucleotides spanning the gap is increased. Acceptor fluorescence (centre) is unaffected by hybridization of probes or spacer DNA.

length of a duplex of length d. Change in fluorescence between unhybridized baseline and fluorescence after spacer addition should only depend on the concentration of origami with both probes attached. Change in fluorescence before and after spacer addition is a function of both concentration of origami with both probes attached *and* the change in distance between the fluorophores as the spacer hybridizes. Using this information I can estimate the ratio of FRET efficiencies E_{ds}/E_{ss}.

Extracting a FRET efficiency ratio from the data

FRET efficiency was determined from donor quenching signal[4] using the relationship:

$$E = 1 - \frac{I_{DA}}{I_{DD}} = 1 - \frac{F^*}{F_0^*} \tag{3.2}$$

F^* is the contribution to fluorescence from origami with two fluorophores bound. There is also a contribution to fluorescence from origami with only one fluorophore bound, and from fluorophores not bound to an origami (e.g. those bound to a detached bookend), F_u. The relationship between observed fluorescence F, and the fluorescence contribution F^*, is given by the relationship:

$$F^* = F - F_u \tag{3.3}$$

F_0 is the contribution to fluorescence from unhybridized probes; before the addition of any origami. Note that F_0^* here must be defined differently because there are no fluorophores bound to origami before the addition of any origami. It is instead defined as the contribution to fluorescence from unhybridized probes which *will eventually* bind a tightrope as a pair, given by the equation $F_0^* = F_0 - F_u$.

I assume that fluorophores not bound to an origami, or bound singly to origami, will have constant fluorescence (i.e. F_u =constant). This is consistent with control experiments which show hybridization of a neighbouring duplex did not affect the fluorescence of either fluorophore when the partner fluorophore was not present, and from the I_{AA} trace which shows no change in acceptor-only fluorescence after spacer addition (figure 3.11e centre). FRET efficiency can now be defined mostly in terms of observable quantities:

$$E = \frac{F_0}{F_0^*}\left(1 - \frac{F}{F_0}\right) \tag{3.4}$$

[4]Donor signal had higher signal to noise ratio than acceptor FRET signal.

Hence ratios of FRET efficiencies for a sample are calculable from measured values:

$$\frac{E_{ds}}{E_{ss}} = \frac{F_0 - F_{ds}}{F_0 - F_{ss}} \tag{3.5}$$

Results

Figure 3.11e shows the raw data from the bulk FRET tightrope titration experiment and table 3.1 shows values extracted from the data for the median fluorescence values at titration plateaus using MATLAB via equation 3.5.

Tightrope	Fluorescence (a.u.)			
	F_0	F_{ss}	F_{ds}	$\frac{E_{ds}}{E_{ss}}$
102 nt	920	665	804	0.46
122 nt	928	525	796	0.33
144 nt	935	499	804	0.30
no top	975	524	864	0.25

Table 3.1: Donor fluorescence values I_{DD} for unhybridized (F_0), single-stranded 24 nt spaced (F_{ss}), and double-stranded 24 nt spaced (F_{ss}) FRET probes with a range of tightrope origami. Values were extracted from the median fluorescence value at saturation using MATLAB 2014b.

Experimental results were compared with predicted values of the FRET efficiency from experted separation r according to $E = \frac{1}{1+(r/R_0)^6}$. For calculations I used Förster radius R_0=7.4 nm [168] and assumed the separation of fluorophores when the spacer is double-stranded[5] to be 8.3 nm giving $E_{ds} = 0.34$. A crude calculation assuming ssDNA length r is a delta function at its expectation value $\langle r \rangle$ for a given tension gives values of 0.36, 0.74 and 0.92 for E_{ss} with 102 nt, 122 nt and 144 nt across the gap respectively, and ratios $\frac{E_{ds}}{E_{ss}}$ of 0.94, 0.46 and 0.37. This is a poor assumption because there will be a distribution of end-to-end lengths r for ssDNA under a force, contributions from $r = \langle r \rangle + \delta r$ to E will not cancel $r = \langle r \rangle - \delta r$. The distribution of lengths r will also be asymmetric about $\langle r \rangle$ (see appendix E). $\langle E \rangle$ can be calculated from the Boltzmann distribution $z(x)$ for displacement x from equilibrium using equation 3.6 (see section E.4.1).

$$\langle E \rangle = \frac{\int E(x) z(x) dx}{\int z(x) dx} \tag{3.6}$$

The resulting values of E_{ss} are 0.38, 0.72 and 0.87, and hence ratios $\frac{E_{ds}}{E_{ss}}$ of 0.88, 0.47 and 0.39 for 102 nt, 122 nt and 144 nt across the gap respectively. Comparing these ratios to table

[5]Calculated 0.34 nm per bp along the helix and $\sqrt{2}$ nm perpendicular to the helical axis because 24 bp is approximately $2\frac{1}{4}$ helical repeats.

3.1, we see that in both cases the ratio decreases monotonically as the number of nucleotides across the gap increases from 102 to 144. This is promising as it indicates the designed order of tensions is correct. Experimental ratios are systematically lower than calculated ratios. This may be due to contributions from misfolded tightropes. Misfolds will typically have lower tension in the single-stranded region under test and therefore tend to inflate observed E_{ss}, reducing ratio $\frac{E_{ds}}{E_{ss}}$. There is also a significant uncertainty in the fluorophore separations r used to calculate ratios. These uncertainties result in large errors in calculated E because the calculation of E involves the sixth power of r.

To obtain quantitative results this experiment should be repeated as a single molecule FRET experiment. This would allow individual fold populations to be deconvolved from the bulk data and the single and double-stranded separations to be measured. The bulk data are sufficient to demonstrate that the expected trend in tension has been achieved by design, but not to confirm exact tensions, or give an idea of proportion of well folded origami.

3.6 Other Work

After my work presented in this chapter had been carried out, a similar structure addressing the same issue was published [169]. The 'force-clamp' [6] design of Nickels and colleagues also held DNA under prescribed tension with an origami structure. Their design has many similar features to my DNA tightrope, for example some helices close to the stretched DNA are not present, Loops hold excess DNA at either side of the structure which can be added across the gap to achieve different tensions, and only a subset of staples (those spanning the helices which are contiguous with the stretched DNA) determine the tension. Nickels and colleagues also choose a similar M13 sub-sequence across their gap as they also use the MCS to generate custom stretched DNA sequences. Their work provides independent verification of the design principles used in this thesis to develop the tightrope.

3.7 Summary

I have presented a design for a DNA origami tightrope which holds a single-stranded section of DNA under tension. The tension applied to this section can be changed by

[6] This is the name given by the authors to their structure, though it is technically a distance clamp.

swapping only a small subset of staples used to construct the tightrope prior to assembly, to stretch a different number of nucleotides over a distance defined by the rest of the structure.

To enable the stretched region to have a specific sequence, for example a motor track sequence, the DNA origami scaffold strand sequence was changed by cloning M13mp18 bacteriophage to contain an inserted subsequence in the MCS. The scaffold sequence was aligned in the origami structure such that the single-stranded section is in the MCS. I presented a screen for inserted DNA of the correct length by PCR across the region containing the insert, directly from small scale phage supernatant without requiring separate DNA isolation and purification steps prior to the PCR reaction. The correct insertion of the desired track sequence was verified by sequencing candidate phage.

I attempted to fold the structure and verify assembly by electrophoresis and imaging. Folding quality was hard to assess from agarose gel electrophoresis due to poor separation between scaffold-only and folded origami bands. TEM of gel purified samples verified that the components had assembled into multi-helical bundle structures, however details of the structure, such as the stretched ssDNA section and gap, could not be well resolved. AFM provided clearer images of the gap as structures stuck preferentially base-down on mica.

Imaging showed the formation of the desired 3D tightrope shape, but could not confirm that tightropes contained intact ssDNA under tension. To verify application of tension, FRET probes were added to the single-stranded section spaced by 24 ssDNA nucleotides. FRET was measured in bulk solution for tightropes assembled under three different tensions and the highest tension design, produced the lowest FRET efficiency, and the lowest tension produced the highest FRET efficiency, with the intermediate tension producing intermediate FRET efficiency, as expected due to the increase in probe separation with increasing tension. Additional experimental uncertainties prevented a quantitative conclusion on the separation of fluorophores, possibly due to heterogeneity of sample and uncertainties about the fluorophore spacing geometry. Single-molecule FRET techniques would possibly be able to overcome these issues, however from the bulk data it appears that a significant population of tightropes are well formed and hold the ssDNA under perscribable tension.

Chapter 4

Thermodynamics and Kinetics of Hybridization to Tightropes

In this chapter I consider the effects of tension on the kinetics and thermodynamics of hybridization of a free strand of DNA to a complementary single-stranded domain held under tension with the tightrope DNA origami. The association rate changes non-monotonically with tension, with a maximum rate observed for 6 pN tightrope tension. The relationship between dissociation rate and tension is also non-monotonic with minimum at 9-12 pN. Free energy change of hybridization for a 10 bp duplex varies slightly with tension in the range 6-12 pN, with a minimum at 6 pN, and duplexes with these tightropes are more stable than the equivalent duplex on M13. At lower, 4.5 pN, tightrope tension duplex stability is reduced and similar to that of the duplex on M13.

4.1 Motivation

It is vital to understand the effect of tension on dynamic processes such as hybridization when working with dynamic, out-of-equilibrium systems where tension is involved. An example of such a system is the bipedal DNA motor on the tightrope origami track to which I will return in chapter 5.

Although the thermodynamics and kinetics of DNA hybridization are well studied [170–175], less is known about the effect of tension on hybridization kinetics. Specific effects of the tightrope origami, such as electrostatic interactions with the multi-helix bundles of a DNA origami structure [37, 176, 177], in combination with the tension, are also unknown.

4.1.1 Why use tightropes to study DNA dynamics?

The tightrope holds the ends of DNA at fixed positions. It is topologically different from a oligonucleotide where ends are free, and different from a circle of DNA, such as M13, where ends are linked. Dynamic DNA processes may be affected by free ends e.g. loop-mediated opening of a hairpin [45] by a strand of DNA is likely to be suppressed if the ends of this strand are fixed.

Most studies of DNA under tension to date [16,178] have been performed using optical or magnetic tweezers to apply force. Both are single-molecule techniques, enabling identification of multiple individual states which contribute to the ensemble. However, throughput is low and specialised experimental apparatus is required. By using the tightrope origami structure from chapter 3, I am able to hold a single strand of DNA under tension defined by the number of nucleotides stretched across the gap, without the need for optical or magnetic tweezers. The tightrope provides an ideal system for bulk solution studies of DNA tension.

Tightropes can be produced in high concentrations $\sim$ 50 nM, and decorated with fluorophores conjugated to DNA oligonucleotides. Tightropes are also compatible with a range of single-molecule methods such as confocal microscopy and wide field microscopy.

4.2 Previous Work

4.2.1 DNA hybridization kinetics

Morrison and Stols used fluorescence spectroscopy to observe hybridization in solution between complementary oligonucleotides bearing fluorophores which form a FRET pair [170]. Dissociation was observed by adding a large excess of a non-fluorescent version of one strand, which replaces labelled strands when they dissociate, causing a reduction in FRET[1]. A similar method is used in this chapter to study hybridization on tightropes.

Secondary structure, determined by sequence, has been demonstrated to have significant effects on hybridization kinetics [172] with three-base-pair intramolecular secondary structure slowing hybridization of a 20mer twofold. Stronger secondary structures were reported to

[1]This requires strand displacement processes to be slow relative to dissociation.

cause the hybridization kinetics to deviate significantly from a two-state association model.

Simulation with oxDNA also found sequence to affect hybridization kinetics [173]. Initial contacts formed during association by the first couple of base pairs have a small probability of generating the correctly hybridized final state (33% at 27°C). Fully hybridized structures can be formed from misaligned initial contacts by internal-displacement, which is sequence dependent. Stabilising misaligned states creates higher observed displacement rates as success probability given initial association is increased. Zhang and colleagues used machine learning to build a model to predict hybridization rate from sequence [174]. Their model uses experimental conditions, including hybridizing probe sequence, salinity, and temperature, to calculate features, such as free energies of hybridization of subsequences, or free energies of misalignment, which give each hybridization experiment a position in feature space which can then be used to predict hybridization rate, from experimental data close in feature space.

Hybridization to protruding sticky ends of a DNA nanostructure formed from packed DNA double helices is slowed compared to hybridization in solution [176] and hybridization is further slowed when there is additional steric hindrance from surrounding single-stranded or double-stranded DNA [177]. Similarly, adhesion of DNA strands to other surfaces has also been shown to slow hybridization rate approximately 20 fold [172].

4.2.2 ssDNA under tension

Smith, Cui and Bustamante investigated the elastic response of single-stranded lambda DNA using optical tweezers [16]. Each end of the DNA was fixed to a latex bead: one bead was held by pipette, the other in a dual-beam trap. The position of the trapped bead, and hence end-to-end length of the molecule, could be manipulated by moving the trap. Force was inferred from bead displacement within the trap, and force-extension curves were obtained for stretching the DNA. The force-extension curve for single-stranded lambda DNA in 150 mM NaCl was observed to fit a freely jointed chain (FJC) model with Kuhn length 1.5 nm for forces <15 pN. Beyond this, an extension to the model allowing elastic Kuhn segments, with stretch modulus 800 pN, was required to fit the high-force data well. Data collected in low salt and 20% formaldehyde was not well described by a FJC model in the low-force

regime <10 pN. This can be explained by lack of secondary structure and reduced screening of backbone phosphate charges increasing the observed end-to-end extension at low forces. Similarly, force-extension curves for ssDNA molecules with differing GC content (therefore differing strengths of secondary structures) only converged for high forces (>10 pN), demonstrating that differences in low-force behaviour may be due to secondary structure [122]. For low force, ssDNA cannot be described as a FJC [179–181], due to repulsion at low extension, however in higher salt conditions, hairpin formation and other base-pairing in ssDNA are possible, decreasing the end-to-end distance and enabling a FJC model to be fit, though not for the reasons, such as no self-avoidance, implied by the model.

OxDNA simulations (see 2.2.1) found the relationship between force and extension for 100 nt ssDNA with random sequence, where secondary structures can form, and when secondary structure formation is forbidden [140]. Consistent with other work [16, 122], curves for random sequences, and curves with no secondary structure converged for forces > 10 pN. Below this tension, extension at low forces was enhanced when secondary structure was forbidden, resulting in lower tension at a given extension without secondary structure.

4.2.3 Hybridization kinetics under tension

A recent study investigated hybridization kinetics with varying tension [178]. Whitley and colleagues used a confocal microscope with two optical traps to apply the tension to a strand of ssDNA. Binding of complementary, Cy3-labelled, oligonucleotides to the stretched DNA was observed. Lifetimes of unbound and bound states were integrated and fitted to exponential functions to obtain rate constants for association and dissociation respectively.

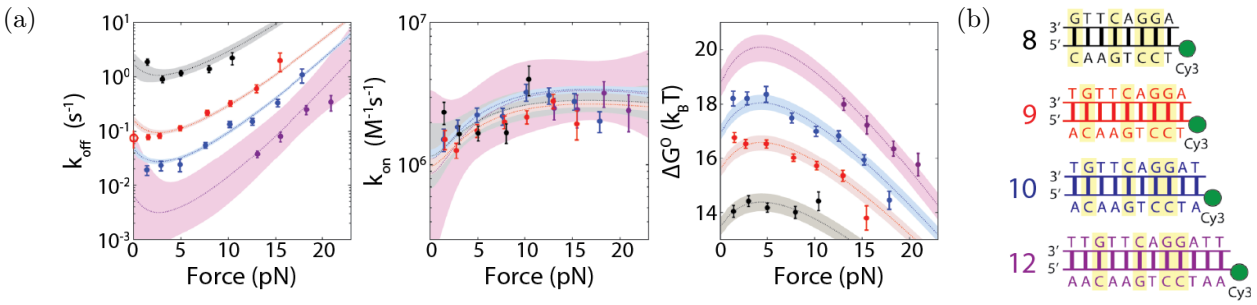

Figure 4.1: Results of Whitley et al. Hybridization under tension. **a** k_{off} (left), k_{on} (centre) and ΔG° (right). Dashed lines show fits and shaded regions 95% C.I. for fits. For details see main text. **b** Sequences for probes length 8-12 nt. Colour matches data in **a**. Figure reproduced from reference 178 by permission of Oxford University Press.

Results shown in figure 4.1a show an increase in k_{off} with increasing tension, little change in k_{on} with tension, and a corresponding decrease in stability with increasing tension, over the tension range investigated. This system was a true force clamp, and was not subject to steric and electrostatic complications of proximity to multi-helix bundles of DNA origami.

Also plotted in figure 4.1a are dashed lines showing fits to the force-varying rate constant data of a modified Arrhenius equation [182] with a force-dependent activation energy. Figure 4.1a left and centre plots are globally fit by functions $k_{\text{off}}(F) = k_0 e^{-\Delta G^*(F)/k_B T}$ and $k_{\text{on}}(F) = k_0 e^{(-\Delta G^\circ(F)+\Delta G^*(F))/k_B T}$, where k_0 is a collision rate, $\Delta G^\circ(F) = \Delta G^\circ(0) - l\int_0^F (x_u(f) - x_b(f))\,df$ is the force dependent standard-state free energy change of hybridization for an l nt probe, and $\Delta G^*(F) = \Delta G^*(0) - l\int_0^F (x^*(f) - x_b(f))\,df$ is the force dependent activation energy for dissociation for an l nt probe. $x_u(F)$, $x_b(F)$, and $x^*(F)$ are the force-dependent extensions of unhybridized ssDNA, hybridized dsDNA, and transition state intermediate respectively. $x_u(F)$ and $x_b(F)$ were calculated from literature, and empirical values using worm like chain (WLC) models. The transition state properties determining $x^*(F)$ (contour length per nt, and persistence length) were fitted parameters, and fitted values appeared sensible, where both contour length per nt and persistence length of the transition state were between those of dsDNA and ssDNA, but closer to those for ssDNA. Free energy of hybridization $\Delta G^\circ(F)$, which is independent of the transition state, was also fit to data in figure 4.1a right, with $\Delta G^\circ(0)$ for each of the probes in 4.1b as a free parameter. High tension points were better fitted than low tension points, where the fitted model function showed non-monotonicity in $\text{k}_{\text{off}}(F)$, and $\Delta G^\circ(F)$ not demonstrated by the data.

4.3 Modelling the Tightrope

Single-stranded DNA is an example of an entropic spring. Pulling on the ends of the spring will align chain segments with the applied force. This alignment reduces the number of microstates corresponding to the observed macrostate or end-to-end length, and a force opposing the applied force. In this thesis I model the tightrope using the Smith mFJC (FJC with extensible Kuhn segments) model, and parameters in appendix E, to estimate tensions for different lengths of DNA stretched across the gap. As noted, the Smith model

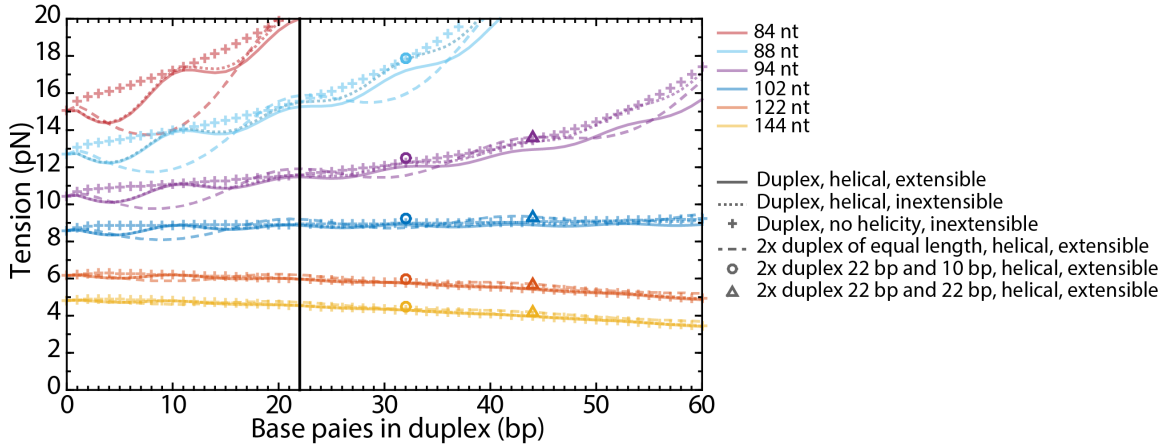

Figure 4.2: Tension calculations for tightropes with a duplex region. Tensions calculated for tightropes with a total number of nucleotides across the gap (both ss and ds) of 84 nt (maroon), 88 nt (cyan), 94 nt (purple), 102 nt (blue), 122 nt (orange), and 144 nt (yellow). Duplex region is modelled as: an extensible, helical duplex FJC segment (solid lines); an inextensible, helical duplex segment (dotted lines); an inextensible duplex segment with no helicity (pluses); or two extensible, helical, duplexes with independent orientations each with half the total bp (dashed lines). Vertical line at 22 bp marks the length of quenching strand R used in later experiments. Points for two independent duplex segments (no stacking) with the first duplex of 22 bp for R, and the second duplex of 10 bp for V_{10} (circles) or 22 bp for V_{20} (triangles). Where used in calculations the dsDNA stretch modulus is 1000 pN, helical radius is 2 nm, and repeat is 10.5 bp, length per bp is 0.34 nm.

will overestimate tensions below ∼10 pN, in cases when secondary structure differs from that of lambda DNA. The implication of this will be discussed later in the chapter.

The tightrope is not a force clamp

The fixed quantity in the tightrope design is the gap across which ssDNA is stretched. Hybridization of an oligonucleotide to the ssDNA will change the tension by converting x flexible ssDNA nucleotides to a rigid duplex of $0.34x$ nm. The stretched DNA has a different contour length and elasticity after hybridization, hence the tension will be different[2].

Introducing non-homogeneous FJC segments to model double-stranded regions

A FJC with uniform segments is generalisable to a chain with heterogeneous, independent, segments (see appendix E). I can include a single chain segment of different Kuhn length to represent the duplex portion of the chain. Similarly two duplex segments can be included assuming their orientations are independent[3]. As duplexes formed on the tightrope are far shorter than the persistence length of double-stranded DNA, (∼50 nm [14,183,184]) I treat them as rigid rods with length D. I consider cases where D= axial duplex length, and

[2]Except at the tension of ds=ss length which occurs at around 8.6 pN.
[3]This is a poor assumption when duplexes are adjacent due to helical end stacking between the duplexes.

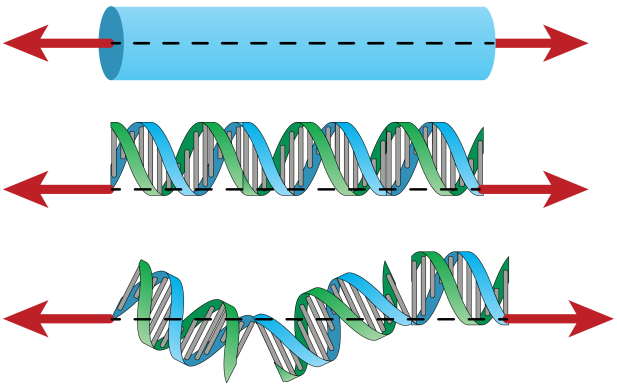

Figure 4.3: Helicity of independent duplex segments affects contour length. Force (red arrows) applied to the centre of a helix face (top) representing the no helicity assumption, or to the blue strand at either end of a continuous DNA duplex with 8 half-turns (centre) or three discontinuous, independent duplexes with 3, 3, and 2 half-turns each (bottom) where the blue strand is continuous but the green strand is discontinuous. Only in a continuous duplex with an even number of half-turns (eg centre), will the length will be the same as the cylinder (top).

where D= duplex length including a correction for helicity. I also consider duplex segments with elasticity defined by stretch modulus S_{ds} ∼1000 pN [14] (see appendix E).

Tightrope tensions for each model are plotted in figure 4.2. There is smaller deviation between the models for the tightrope with 144 nt across the gap. There is negligible difference between models with and without extensible duplex segments for 144 nt-102 nt tightropes, but higher tension 94 nt-84 nt tightropes show small differences in tension when extensibility is included. In figure 4.2 solid lines (extensible duplexes) are lower than dotted lines (inextensible duplexes). Models including duplex helicity display peak tension at an integer number of helical repeats (10.5 bp) because at these points, the contour length of the DNA across the gap is a minimum. Splitting the duplex into two independent duplexes of half the length (dashed lines) emphasises the effect of helicity, as two duplexes with an odd integer of half-turns will have longer total contour length than a single duplex with an even integer number of half-turns but the same total number of base pairs (see figure 4.3).

From these results it appears that the inclusion of helicity and extensibility in the model have a large effect on calculated tension for tightrope DNA with a duplex section, especially when number of nucleotides across the gap is close to or above the ss=ds length of 102 nt. I therefore choose to use the model with both helicity and extensible duplexes for tension calculations in the remainder of this chapter. Furthermore the independence of duplex segments, when multiple duplexes are present on the tightrope, will have an impact on the

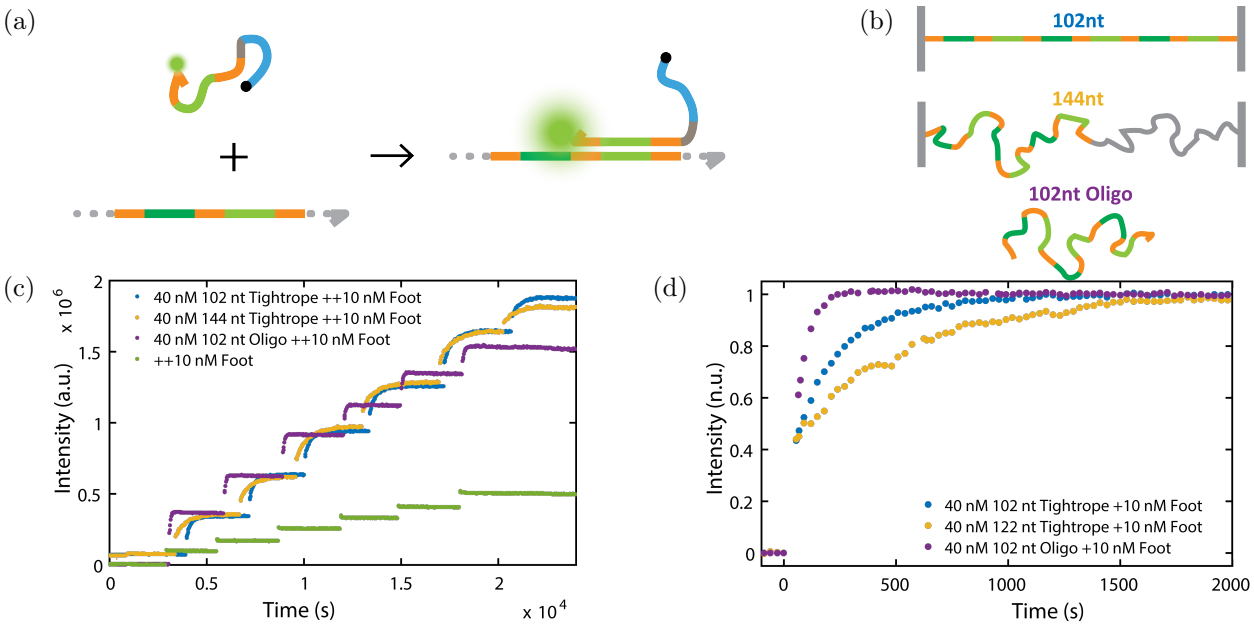

Figure 4.4: Titrating tracks with dual-labelled foot. **a** Schematic of dual-labelled fluorescent foot binding to a track binding site. Foot fluorescence increases upon hybridization as of fluorophore-quencher separation increases. **b** Schematic for tracks used in the titrations. The top two tracks are stretched across a 35 nm gap using tightrope origami. The third is the same ssDNA sequence as an oligonucleotide. All tracks contain three binding sites for the foot. **c** Raw fluorescence data for the stepwise addition of 10 nM foot to 40 nM track either: on tightrope origami with 102 nt (blue), or 144 nt (yellow) across the gap, or on a 102 nt long oligonucleotide. A control where 10 nM foot is added to buffer is shown in green. **d** The first 10 nM addition to 40 nM track. Fluorescence normalised to 0 before addition and 1 after equilibration. As the foot added is fluorescent, some of the fluorescence increase (∼0.4 n.u.) is an increase in background signal from foot addition.

calculated force. I choose to consider multiple consecutive duplexes as a single 'effective duplex' in the remainder of this chapter due to stacking interactions between consecutive duplexes [27, 28]. The true value will lie between this and the value for independent duplex segments (for example for a 22 bp duplex and consecutive 10 bp duplex, used in experiments later in the chapter, the value will be between the solid line at 32 bp and the open circle point in figure 4.2). When helices are not adjacent I will treat them as independent segments.

4.4 Investigating Hybridization Kinetics Using Tightropes

4.4.1 Hybridization of motor feet to tracks

Proof-of-concept hybridization experiments were performed using one of the bipedal motor feet (see chapter 5) with a Cy3 fluorophore at the 3' end and a quencher Iowa Black FQ at the 5' end. The 22 3'-most nucleotides can hybridize to the track[4] (see figure 4.4a). Tightropes have a 6 foot-site track for a motor with non-identical feet (i.e. three foot-binding

[4]There will be a single AG mismatch in this 22 bp duplex.

sites for each type of foot). M13NH1 Tightrope samples were folded at different tensions, purified from excess staple by PEG precipitation (see appendix A), and resuspended to 40 nM. A sample of the same track as a linear oligonucleotide at 40 nM was also prepared. To each sample 10 nM foot was added stepwise and fluorescence change recorded. As the foot has three binding sites per track, the concentration of binding sites in solution is 120 nM.

Results are shown in figure 4.4. In figure 4.4d I show only the first addition, as binding sites are in the largest excess and it is less likely for multiple feet to bind to the same track[5]. Hybridization to tightropes is slower than the hybridization to an oligonucleotide track for all six additions of 10 nM foot. Hybridization to a tightrope track with 102 nt across the gap is faster than to a tightrope track with 144 nt across the gap. The difference between hybridization to oligonucleotide and tightrope origami tracks may be due to factors other than ssDNA tension, such as electrostatic/steric effects, and topology, but differences between the 102 nt and 144 nt ssDNA tracks could be due to the tension. A system without the complications of multiple ssDNA binding sites, and inter-probe quenching will be used in the remainder of this chapter to investigate hybridization kinetics on tightropes.

4.4.2 A low secondary structure, uniquely-addressable, tightrope

The sequence of M13 across the gap has secondary structure. Predictions of this secondary structure in a single-stranded section in the absence of the rest of the M13 are shown in figure 4.5. To reduce the effects of the secondary structure of the single-stranded section on hybridization experiments, a new insert[6] was cloned into M13mp18 as described in section 3.4.2. The new M13 construct is named M13NH2.

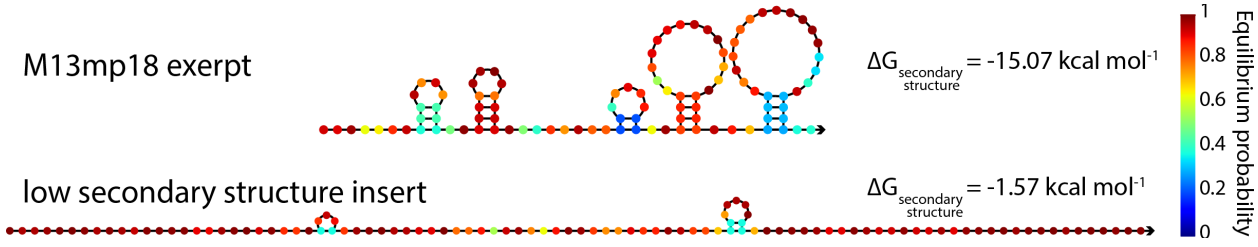

Figure 4.5: Minimum free energy secondary structures of M13 in the single-stranded section of the tightrope (top) and the designed low secondary structure insert (bottom) at 25°C calculated using NUPACK software [150]. Figure adapted from NUPACK software output.

[5]Multiple feet binding the same track is likely to lead to some inter-foot quenching, and the presence of one foot bound to a tightrope will change its tension (see section 4.3).

[6]Also of length 102 nt to use the same tension setting staples as M13NH1.

Candidate sequences for the insert were generated using MATLAB 2014b. A random 102 nt sequence was produced and all possible 4 nt sub-strings and their reverse complement sub-strings were collated. Repeated sub-strings or reverse complements of sub-strings were mutated iteratively until a sequence without any repeated or complementary 4 nt domains was obtained. Secondary structures of a number of these candidates were compared using NUPACK software [150] to predict MFE secondary structures, and a structure with low MFE secondary structure selected. The MFE secondary structure of the selected insert sequence is shown in figure 4.5, and shows reduced secondary structure compared to M13mp18.

4.4.3 Experimental scheme

Tightropes were assembled with a quencher strand R bearing 5' Iowa Black FQ hybridized to the DNA spanning the tightrope gap. Hybridization probe strand V is complementary to the ssDNA domain on the tightrope adjacent to the 5' end of R (see figure 4.6). A 3' Cy3 label on V will quench upon hybridization to the tightrope next to R.

R-labelled tightropes are added to a solution containing free V, as this allows the measurement of unhybridized baseline fluorescence for V, required for normalisation. To achieve 10 nM tightropes in the sample cuvette, 30 µl of 50 nM tightropes are added to 120 µl 25 nM V solution. This dilutes V to 20 nM. By fitting the fluorescence change after tightrope addition, the kinetic parameters k_{on} and k_{off} for the hybridization process can be obtained.

Dissociation rate k_{off} was also investigated using a probe dilution method [170] where, after equilibrium has been established, an excess of DNA with identical sequence to the hybridization probe V, but no fluorophore label, is added. Assuming second-order blunt-ended displacement processes occur at a much slower rate than first-order dissociation, the

Figure 4.6: Hybridization experiment schematic. **a** Schematic of the single-stranded tightrope segment bearing quencher probe R. Hybridizing probe strand V is able to hybridize to its complement domain immediately adjacent to R which holds the Cy3 fluorophore close to the Iowa Black FQ quencher on R, reducing fluorescence intensity upon hybridization. **b** Normalisation complex is able to form 22 bp with V to achieve a 'fully quenched' fluorescence level.

presence of the excess unlabelled probes will mean that once a tightrope bound fluorescent probe melts from the tightrope, it is statistically unlikely to rebind a tightrope in the limit $V \ll V_{\mathrm{u}}$. At the new equilibrium, bound strands will be predominantly unlabelled. The exponential decay of VT will be observed as an increase in fluorescence to a new plateau.

At the end of the experiment, an excess of normalisation complex is added to hybridize any V in solution and provide a fluorescence measurement for 20 nM fully hybridized V. The normalisation complex consists of a 'tightrope-like' strand T^* with domains complementary to R and V (see figure 4.6b). The complex is constructed with a 3:1 excess of R to T^*.

Hybridizing domains of different lengths x were tested. All V were 22 nt long, but of the 22 nt only the 3'-most x nt were complementary to the domain adjacent to R on the tightrope. A 'long' strand V_{22} where $x = 22$, was expected to hybridize to V effectively irreversibly ($K_d = k_{\mathrm{off}}/k_{\mathrm{on}} \approx 0$). A 'short' strand V_{10} where $x = 10$ is expected to have lower melting temperature, and greater k_{off} than V_{22}. Regardless of the length of the tightrope hybridizing domain, all strands V can form 22 bp with their normalisation complex. The hybridizing domain length of 10 nt was chosen as its melting temperature was close to the 25°C used for observation, therefore its stability is similar in hybridized and unhybridized states. This should produce an observable change in fluorescence in the hybridization phase, but bound state lifetimes should be short for dissociation to occur on an observable timescale.

4.4.4 Data processing and normalisation

Addition of DNA increases the fluorescence of Cy3-labelled probes V in solution. Details of this non-specific effect are presented in appendix C. This has the effect of reducing the apparent dilution factor due to tightrope addition. Adding 30 μl non fluorescent solution to 120 μl fluorescent solution, would naïvely be expected to reduce observed fluorescence to $120/150 = 80\%$ of its initial value. Control experiments show that instead, the fluorescence after dilution is 85±3% of its initial value. An 'effective dilution factor' is used to calculate the unhybridized fluorescence of 20 nM V from the baseline fluorescence at 25 nM $\times 0.85$.

Example raw fluorescence data is shown in figure 4.7a with 'dilution factor' scaled data in orange. Association phase data is normalised by scaling the 20 nM baseline to zero, and the

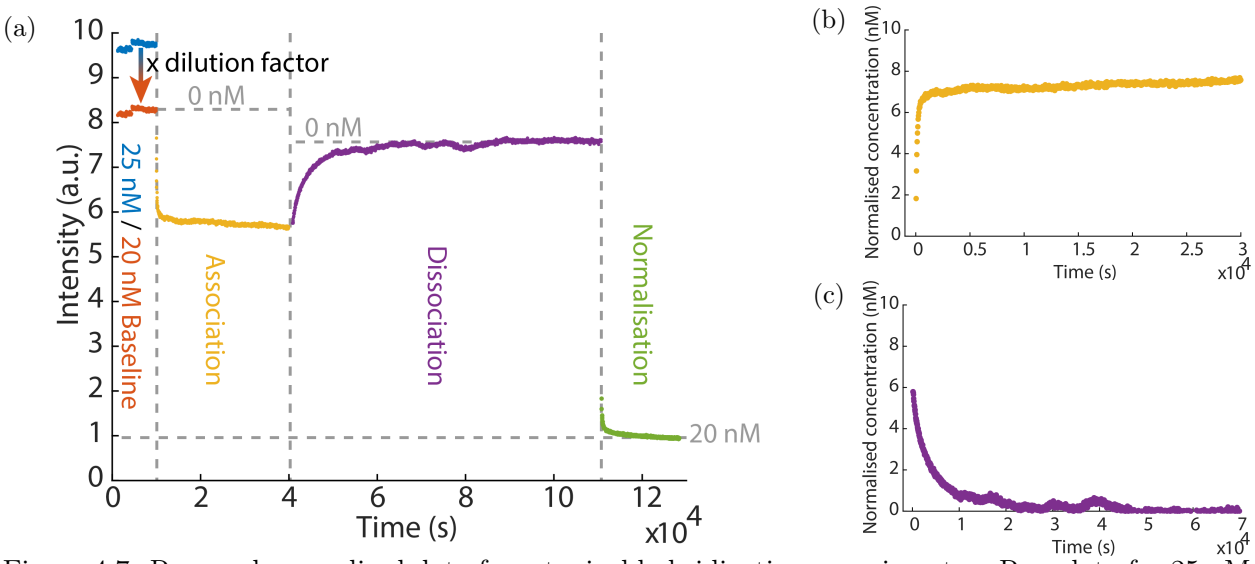

Figure 4.7: Raw and normalised data for a typical hybridization experiment. **a** Raw data for 25 nM V_{10} diluted to 20 nM by addition of 10 nM T with 122 nt DNA across the gap. Prior to addition of T raw fluorescence counts (blue) are scaled by an effective dilution factor of 0.85 to approximate the equivalent fluorescence of 20 nM (rather than 25 nM) free V_{10} in solution (orange). This provides the zero for the association phase (yellow). After the association reaction has plateaued, 400 nM unlabelled V_{10} are added to observe recovery of fluorescence as bound V_{10} dissociate from T (purple). Finally 600 nM normalisation complex is added to provide a 'fully hybridized' fluorescence level where all V_{10} are bound adjacent to a quencher. **b** Association data normalised to zero at the 20 nM baseline level (orange) and 20 nM after normalisation complex addition (green). **c** Dissociation data normalised to zero at its maximum and 20 nM after normalisation complex addition (green).

level after normalisation complex addition to 20 nM (total concentration of V). Dissociation phase data is normalised by scaling the plateau after dissociation to zero and the level after normalisation complex addition to 20 nM. Time is also zeroed for each phase.

4.4.5 Kinetic model

Hybridization can be modelled as a reversible bimolecular association reaction.

$$V + T \underset{k_{\text{off}}}{\overset{k_{\text{on}}}{\rightleftharpoons}} VT \tag{4.1}$$

The resulting ODE for the rate of change of $[VT]$ is analytically soluble. The concentration of hybridized VT as a function of time is given by equation 4.2.

$$[VT] = \frac{[V]_0[T]_0(1 - e^{-k_{\text{on}}\beta t})}{\beta + [VT]_{eq}(1 - e^{-k_{\text{on}}\beta t})} \tag{4.2}$$

Where $\beta = \sqrt{\left([V]_0 + [T]_0 + \frac{k_{\text{off}}}{k_{\text{on}}}\right)^2 - 4[V]_0[T]_0}$ and $[VT]_{eq} = \frac{[V]_0 + [T]_0 + \frac{k_{\text{off}}}{k_{\text{on}}} - \beta}{2}$.

During dissociation, assuming no displacement and $[V_u] \gg [V]$, $[VT]$ is given by equation 4.3.

$$[VT] = [VT]_0 e^{-k_{\text{off}} t} \tag{4.3}$$

Where $[VT]_0 \sim [VT]_{eq}$ is the starting concentration of VT immediately prior to addition of excess V_u. For derivations see appendix F.

4.5 Results

4.5.1 Dissociation

Normalised dissociation data (see figure 4.7c) was fitted for V_{10} by taking the natural logarithm of normalised data and fitting a straight line of gradient $-k_{\text{off}}$. Data was well fit by an exponential decay as shown in figure 4.8a. Longer V_{22} were not tested because the decay timescale was too long to measure (reactions are effectively irreversible). Even the dissociation of V_{11} probes was too slow to measure accurately with the apparatus used.

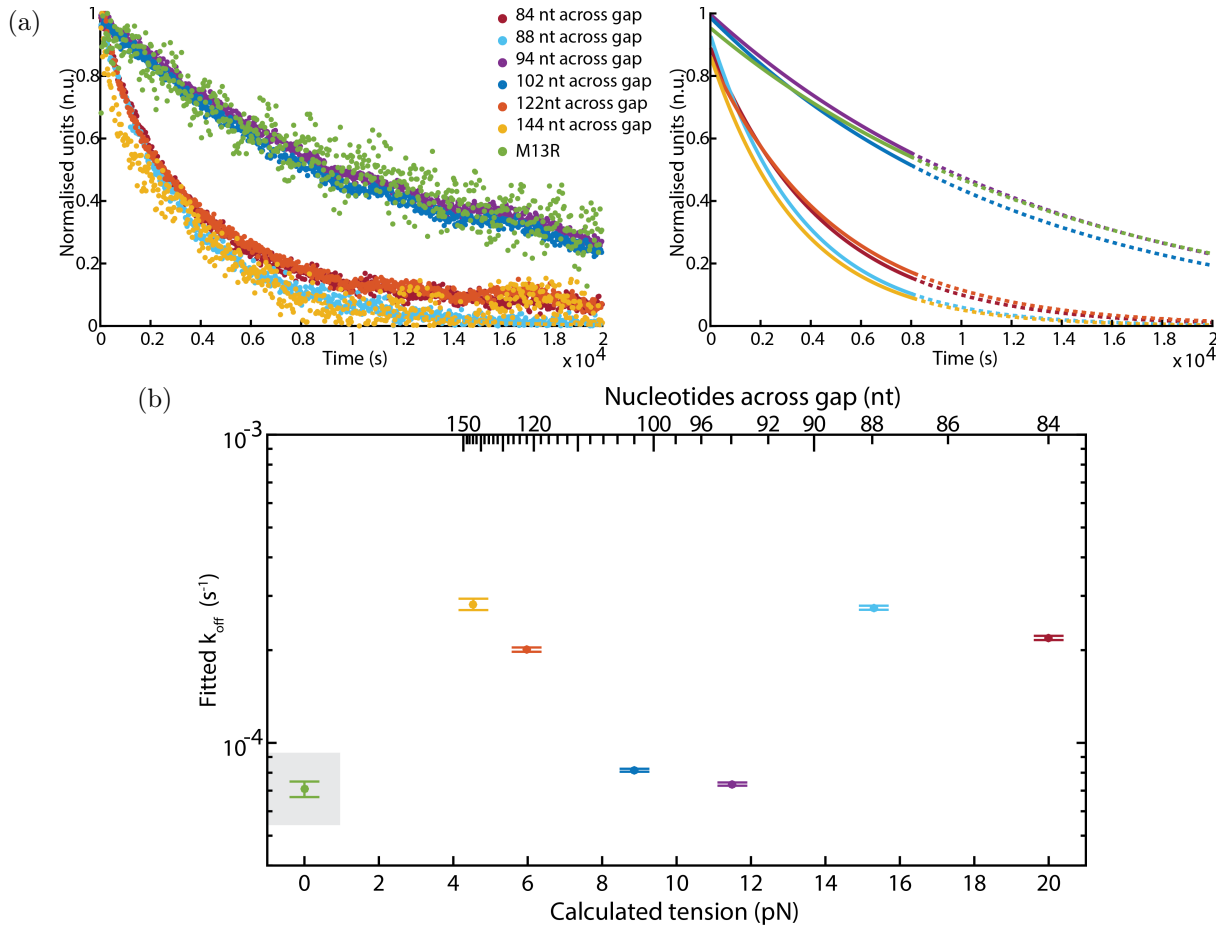

Figure 4.8: First-order fits of V_{10} dissociation from tightropes. **a** Normalised fluorescence dissociation data rescaled to lie between 0 and 1 for visual comparison (left) and fits to the data between 0 s and 8000 s (right). Dotted lines show extrapolations of fit to later times. **b** Fitted k_{off} with calculated tension using a mFJC model including an extensible duplex segment. The 0 tension case is unfolded M13 scaffold bearing quencher strand R. Error bars show 95% confidence intervals for fitted values. Grey box indicates M13 data point which is non-tightrope based.

Decay data for all constructs was rescaled to lie between 0 and 1 for visualisation on the same plot in figure 4.8a. First-order decay fits to the data are shown in the right hand plot. There are two clusters of decay rate, indicating two different decay rate regimes: 'fast' where $k_{\text{off}} \sim 2\times10^{-4}$ s^{-1} and 'slow' where $k_{\text{off}} \sim 7\times10^{-5}$ s^{-1}, threefold slower. The off-rate is slow for M13 which has not been folded into a tightrope. Tightropes with 144 nt or 122 nt of DNA stretched across the gap (estimated tensions of 4.5 pN and 6.0 pN respectively) both show fast off-rates. The tightropes with fewest nucleotides, 88 nt or 84 nt stretched across the gap (estimated tensions of 15 pN and 20 pN respectively) also produced fast off-rates, whereas intermediate numbers of nucleotides, 102 nt or 94 nt across the gap (estimated tensions of 8.9 pN and 12 pN respectively) produced slow off-rates similar to those observed with M13. The variation of off-rate with tension is non-monotonic with minimum rate close to the ss=ds length. Tightropes folded with more, or fewer nucleotides across the gap, have increased k_{off}.

4.5.2 Association

The hybridization of V_{22} is assumed to be effectively irreversible due to the large free energy of hybridization of 22 base pairs. Both reversible bimolecular association (equation 4.2), and irreversible bimolecular association (where k_{off} is fixed at zero) were fitted to the data. For the irreversible bimolecular association fit, an additional free scaling parameter was used to account for differences between the limit of the fitted function $\lim_{t\to\infty}[VT] = \min([V]_0, [T]_0)$ and the normalised data at late times. Rates for reversible bimolecular association were also derived form half-completion times (see section F.1.3).

For hybridization of V_{10}, irreversible binding cannot be assumed as the melting temperature of this 10 bp duplex is 22±1°C [150], below the experimental temperature of 25°C. I therefore only find the fitted and half-completion-derived bimolecular association rate constant for these experiments.

Bimolecular association rate constant k_{on} determined by these methods for both V_{22} and V_{10} are plotted in figure 4.9. Results from the two different length hybridizing probes show a similar trend of k_{on} with tension however association rates with V_{22} are consistently $\sim 3\times$ higher than equivalent rates with V_{10}. For both lengths of hybridizing probe the lowest

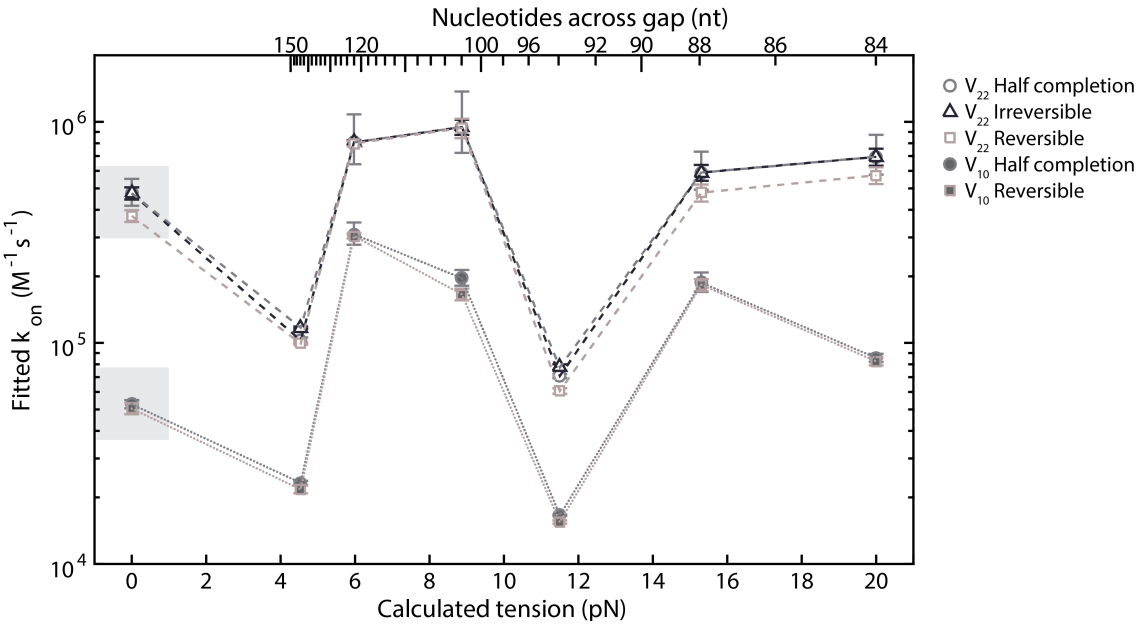

Figure 4.9: Fitted second-order association rate constants k_{on} for hybridization of V_{22} (light filled symbols and dashed lines) or V_{10} (dark filled symbols and dotted lines) to tightropes. Multiple methods were used to determine second-order rate constant from scaled data. For the hybridization of the 10mer, A reversible bimolecular association fit with free parameters k_{on} and k_{off} is shown by squares, and k_{off} determined from half-completion time and plateau concentration according to equation F.29 is shown by circles. Both methods are also used to produce rate constants k_{on} for hybridization of V_{22}. A fit of irreversible bimolecular association was also performed with two free parameters: k_{on} and a scaling factor $0.7 \leq a \leq 1.1$ to allow for some deviation of the normalised data plateau from $\min([V]_0, [T]_0)$ shown by triangles. Error bars show 95% confidence intervals for fitted values for reversible and irreversible fit methods and show k_{on} determined from $t_{1/2} \pm 10$ s for the half-completion method, the estimated error in determining half-completion time. Grey boxes indicate M13 data points which are non-tightrope based.

association rates occur for 94 nt or 144 nt across the gap (calculated tensions of 12 pN and 4.5 pN respectively). In-between these tensions the hybridization rate is significantly higher. The highest k_{on} for hybridizing probe V_{10} occurs for 122 nt across the gap. For V_{22} both 102 nt and 122 nt across the gap exhibit similar k_{on}, fitted values are within error of each other. Unfolded M13 , and the two highest attempted tensions with 88 nt and 84 nt stretched across the gap have higher k_{on} than the tightropes with 94 nt or 144 nt across the gap.

Quality of second-order fits to association data

Bimolecular association fits to data provide a useful proxy for comparison of association rates in figure 4.9, however, from fit shape and residuals it is clear that a single second-order process cannot fully describe the observed behaviour. Fits systematically underestimate the initial rise, and plateau prematurely, compensating by under-fitting the plateau value. With only two fit parameters k_{on} and k_{off} it is not possible for the data to be well fit.

Second-order bimolecular association fits to the data are shown in figure 4.10. The fitting algorithm weights data points in the fit logarithmically between 10^3 and 10^1. Removing this weighting increases the extent of underestimation of the gradient at early times in favour of fitting the plateau. Second-order association fits also provide a value for k_{off}. This value is less reliable than the value of k_{off} determined from dissociation fits as it is very sensitive to the plateau level, and as noted the fitted second-order association curve does not describe the data well, whereas the first order dissociation fit does. Nonetheless values for k_{off} determined from association and dissociation can be compared.

Table 4.1 shows rate constants k_{off} determined by both methods for strand V_{10}. Results show poor agreement between the two methods for deriving k_{off}. k_{off} values from dissociation fits are lower than those from bimolecular association fits. Errors indicated are 95% confidence intervals for the fitted parameters. They do not show the quality of fit, merely the parameter confidence given the fit function and weighting vector, which may not be an appropriate model for the behaviour. Consistent with the observation that the k_{off} deter-

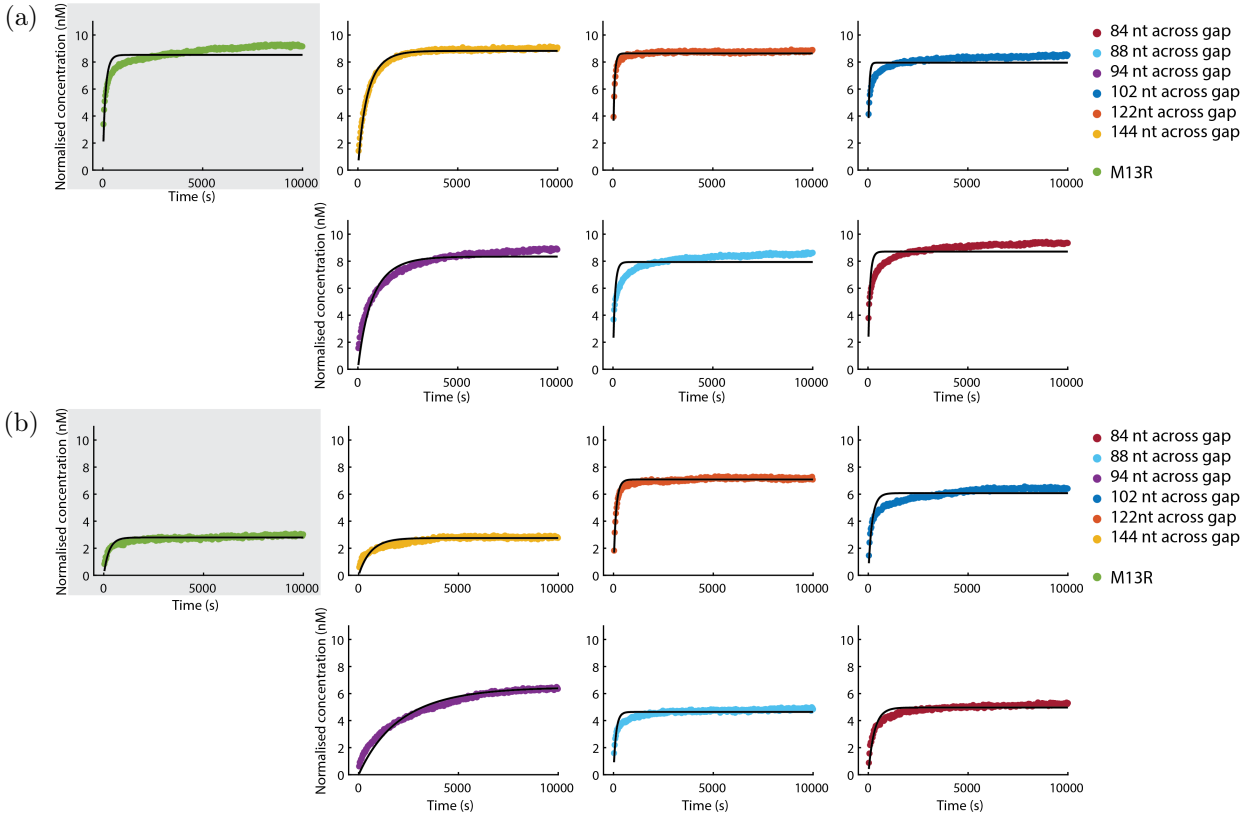

Figure 4.10: Second-order bimolecular association fitted curves for **a** V_{22} hybridizing probe and **b** V_{10} hybridizing probe. Fitted curves are plotted in black and data points are coloured according to the legend. These fits are the source for k_{on} values in figure 4.9 depicted with square points. Grey boxes indicate M13 data points which are non-tightrope based.

Structure	k_{off} Dissociation (s^{-1})	k_{off} Bimolecular association (s^{-1})
94 nt	$7.35\pm0.09 \times 10^{-5}$	$1.24\pm0.04 \times 10^{-4}$
102 nt	$8.15\pm0.09 \times 10^{-5}$	$1.63\pm0.11 \times 10^{-3}$
122 nt	$2.01\pm0.03 \times 10^{-4}$	$1.77\pm0.05 \times 10^{-3}$
144 nt	$2.82\pm0.12 \times 10^{-4}$	$1.06\pm0.05 \times 10^{-3}$
84 nt	$2.19\pm0.03 \times 10^{-4}$	$1.36\pm0.06 \times 10^{-3}$
88 nt	$2.75\pm0.04 \times 10^{-4}$	$3.51\pm0.23 \times 10^{-3}$
M13	$7.08\pm0.41 \times 10^{-5}$	$2.40\pm0.15 \times 10^{-3}$

Table 4.1: Comparison of dissociation rate k_{off} from dissociation and bimolecular association fits. Errors are 95% C.I. in fitted parameters.

mined from fits to the association phase tend to be faster than those determined from fits to the dissociation phase, if bimolecular association fits are attempted with only k_{on} as a free parameter, and k_{off} fixed to be equal to the value determined from the dissociation phase, the fit quality is poor due to a low fitted value of k_{on}, in order to fit the plateau level, which is a function of the free energy and therefore ratio $K_D = k_{\text{off}}/k_{\text{on}}$.

Two-population fits

The systematic deviation of fitted curves from experimental data, and the disagreement between fitted off-rates determined from association and dissociation phases, suggests that the process is not a single-population second-order bimolecular association reaction. A model with two different populations, each of which undergo a second-order bimolecular association reaction with different on-rates, but a shared off-rate equal to the dissociation fitted off-rate, is able to fit the data well at early and late times. The rationale for sharing k_{off} determined from the dissociation is that the shape of the dissociation appears to be a single exponential decay, rather than the sum of exponentials with different decay times, indicating a single k_{off} rate. This fitting function has three free parameters, two of these are the two fitted on-rates k_{on1} and k_{on2}, and the third, p, describes the proportion of structures in each of the two populations where p is the probability that a structure is in population 1 with on-rate k_{on1} and $1-p$ is the probability that a structure is in population 2 with rate k_{on2}.

Fit results to the V_{10} data are shown in figure 4.11. The fitted curves are far closer to the data points than those fitted with a single bimolecular association, though this is with the caveat that there are three rather than two fit parameters in the two-population model.

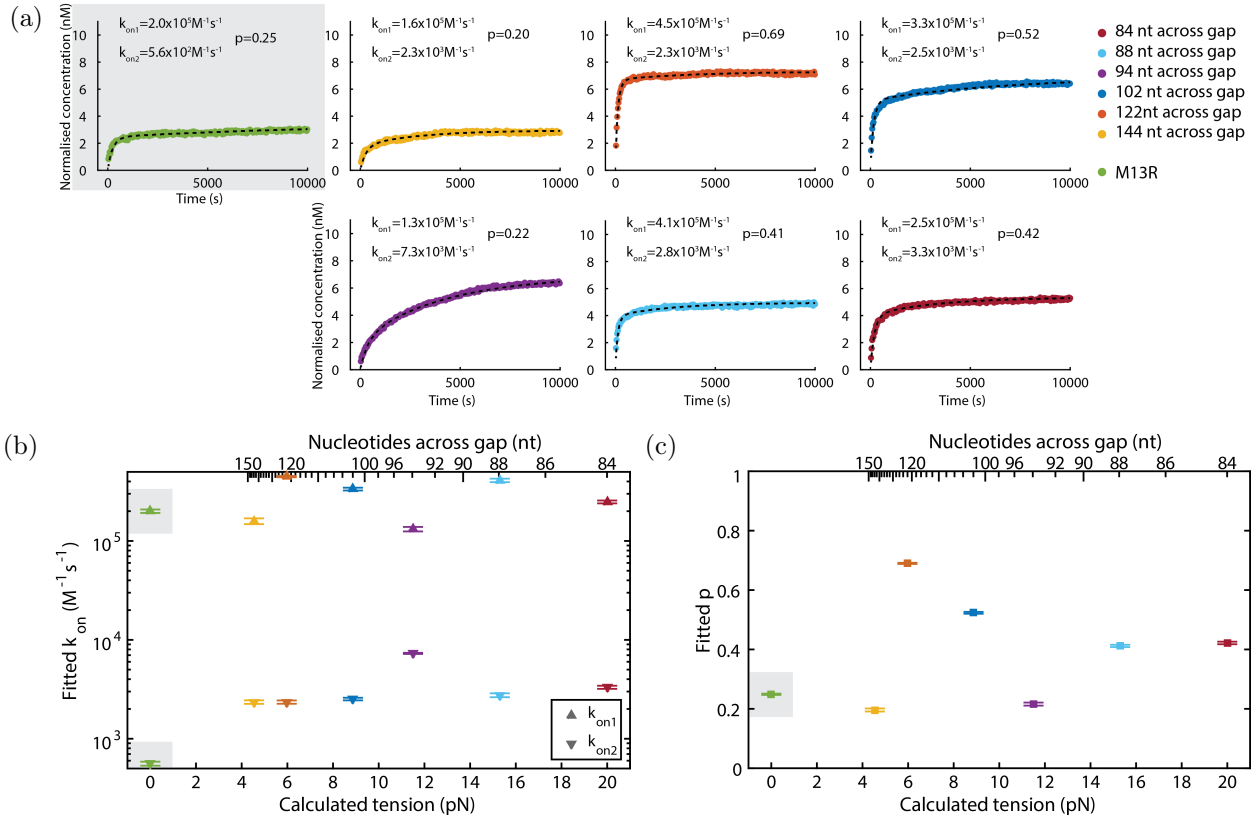

Figure 4.11: **a** Two-population second-order bimolecular association fit with k_{off} set from the dissociation phase, and 3 free parameters, k_{on} for each population and proportion p of structures in population 1. **b** Fitted k_{on1} (triangles) and k_{on2} (inverted triangles). **c** Fitted p. Error bars show 95% C.I. for fitted parameters. Grey boxes indicate M13 data points which are non-tightrope based.

Two of the parameters, k_{on1} and p, appear to covary. The initial rapid rise is accounted for by population 1, and a greater initial gradient can be achieved by either increasing k_{on1} or by increasing p, the proportion of structures in population 1. Data was also fitted by fixing each of these parameters in turn, setting the fitted parameter equal to its mean determined with all three free parameters. Fitting with fixed $k_{\text{on1}} = 2.8 \times 10^5$ $\text{M}^{-1}\text{s}^{-1}$ still matched the data well, however fitting with fixed p , or k_{on2} yielded poor results. The association rate constant for the faster population k_{on1}, and association rate constant k_{on} from single population bimolecular association (figure 4.9) appear strongly correlated, however k_{on2} and k_{on} are not correlated.

4.5.3 Free energy change of hybridization

The plateau height after equilibration in the association phase can be used to estimate a free energy change of hybridization of V_{10} as the plateau gives the concentration of VT at equilibrium. From the Arrhenius equation $K_D = \frac{k_{\text{off}}}{k_{\text{on}}} = \frac{[V]_{eq}[T]_{eq}}{[VT]_{eq}} = e^{\Delta\Delta G/k_B T}$ [182]. All

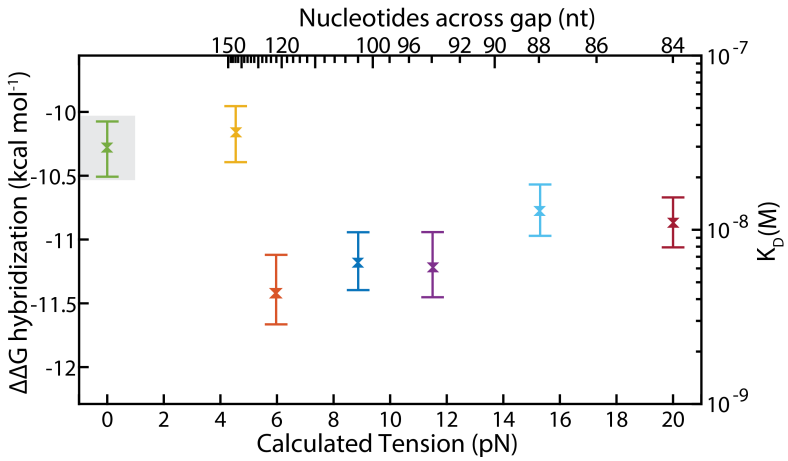

Figure 4.12: Plot of free energy change upon hybridization $\Delta\Delta G$ with calculated tension. Free energy change is calculated from a K_D value by the Arrhenius equation. Plateau height after association is used to extract a K_D (filled crosses). Error bars show values obtained assuming maximum and minimum dilution factors of 0.82 and 0.88 respectively, the greatest source of uncertainty in data analysis. Grey box indicates M13 data point which is non-tightrope based.

of these quantities can therefore be calculated given knowledge of the plateau level and the initial concentrations of V_{10} and tightrope origami T.

Free energy change for hybridization $\Delta\Delta G$ and dissociation constant K_D are given in figure 4.12 for the calculated tensions trialled with a V_{10} hybridizing strand. Results show there is little change in $\Delta\Delta G$ with tension over the experimental range, only a variation of ~ 1.5 kcal mol^{-1} is observed. To put these values in context, the free energy of a single base pair is ~ 1.42 kcal mol^{-1} at 37°C in 1M NaCl [28]. Values of K_D vary by at most 10 fold. The most stable duplex is formed between V_{10} and tightropes with 122 nt, 102 nt or 94 nt across the gap, and duplexes between V_{10} and M13, or tightropes with 144 nt across the gap are the least stable resulting in the least negative free energy change of hybridization.

4.6 Discussion

4.6.1 Misfolded tightropes

The relationships between rate constants k_{on}, k_{off} and calculated tension have multiple maxima. This may be because true tensions differ from those calculated. Indeed, for the tightropes folded against the highest tension, structures may not be well folded, with all staples in place, necessary to produce tensions close to those predicted by the mFJC model. A schematic depicting how partial folding of tightropes could produce lower tension in the

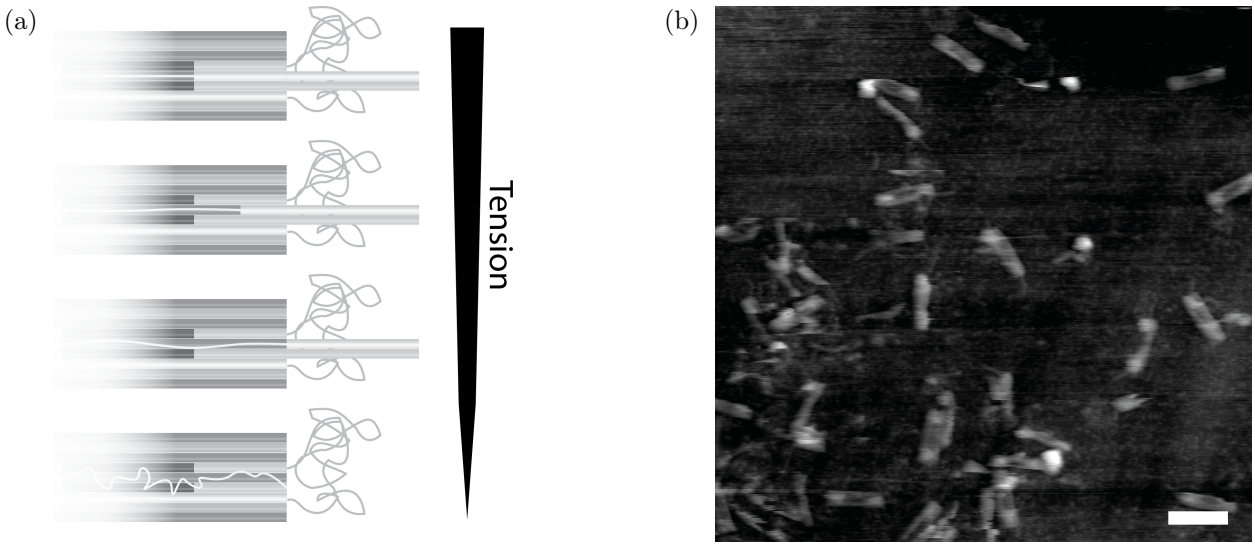

Figure 4.13: **a** Schematic showing possible misfolding motifs for the RHS of tightropes when the number of nucleotides across the gap is shorter than the number of basepairs in a duplex spanning the gap. The fully-folded structure is depicted at the top, but by omitting staples close to the gap going outwards, the tension in the structure can be reduced as both effective gap size and number of single-stranded nucleotides are increased. **b** AFM micrograph of tightropes folded with 84 nt across the 35 nm gap. Some structures are malformed. Scale bar 100 nm.

ssDNA region is shown in figure 4.13a. By omitting some staple domains on the tension-bearing helix during folding, more nucleotides are stretched across larger gaps producing lower tensions. I believe 94 nt tightropes to be well folded because high resolution AFM images show mostly intact structures with the expected features (see figure 3.9a). An AFM image for a tightrope with 84 nt across the gap is shown in figure 4.13b. Some extended central gaps can be seen (consistent with partial folding at the outside edges of the structure).

4.6.2 Behaviour of tightropes with 84 nt - 144 nt across the gap

The fitted association rate k_{on} is highest at 122 nt and lower either side, with minimum association rate for 94 nt tightropes, but higher association rate again for the two highest tension tightropes (which may not be well folded). k_{off} is also non-monotonic. It is highest for tightropes with 144 nt, 88 nt, or 84 nt across the gap, and lower for intermediate numbers of nucleotides across the gap. K_D is highest for 144 nt tightropes suggesting the duplex on the 144 nt tightrope is less stable. This is consistent with low k_{on} and high k_{off} for the 144 nt tightrope. Increasing tension past 144 nt appears to stabilise the duplex. There is a minimum $\Delta\Delta G$ around 122 nt - 94 nt across the gap, after which the free energy change increases. Stretching more than 94 nt across the gap destabilises the duplex.

During hybridization a disordered length of ssDNA must be converted to an ordered double helix with fewer degrees of freedom as rotation about covalent bonds in the DNA backbone is restricted by base stacking and hydrogen bonding interactions. The decrease in entropy is compensated for by an enthalpic gain as hydrogen bonds and $\pi - \pi$ stacking interactions are formed in consecutive base pairs. Applying tension to the ssDNA may pre-order the inital state making it more similar to the transition state, partially aligned in such a way as to make it easier to form a duplex. Tension should also remove secondary structure in the ssDNA, though this effect is reduced in my system by design of a low secondary structure sequence. This could explain the increase in association rate observed between 144 nt and 122 nt stretched across the gap. Some of the entropic cost of duplex formation has already been paid by stretching the ssDNA into the tightrope, increasing the free energy released by hybridization, hence $\Delta\Delta G$ decreases between 144 nt and 122 nt.

Once a critical tension is exceeded, further stretching of the ssDNA moves the initial state further from the transition state, for example tensions too high might stretch the DNA beyond the equivalent duplex length so that it must contract to form a helix. This could explain increases in k_{off}, and $\Delta\Delta G$, when only 88 nt or 84 nt are stretched across the gap, compared to 102 nt or 94 nt across the gap. In the high force limit, duplexes melt [16, 185, 186], therefore eventually $\Delta\Delta G$ for duplex formation will become positive.

4.6.3 Comparison with other studies of hybridization under tension

Results initially appear contradictory to those of reference 178. Whitely *et al.* found decreasing dissociation rate with increasing tension in the range 2 pN to 18 pN whereas my results show non-monotonicity. They note that the association rate varies little with tension which, again, is contrary to my findings, where on-rate k_{on} varies by an order of magnitude. The authors of reference 178 define free energy change of hybridization as $-k_B T \ln K_D$ which has the opposite sign from the definition presented here, however when the sign is corrected, results from both my own study and reference 178 find decreasing stability (increasing $\Delta\Delta G$) of duplexes with increasing tension beyond 5 pN, however they do not observe destabilisation below this tension seen in my data for tightropes with 144 nt across the gap.

4.6.4 Systematic overestimation of tension with mFJC model

Above 10 pN the mFJC model should provide a reasonable estimate for ssDNA tensions, however below 10 pN the tension will be overestimated, particularly because the insert used was designed to minimise secondary structure, therefore hairpin formation at low force will be suppressed in the first 102 nt of the gap spanning ssDNA (see figure 4.5). Consequently any values for tension derived with the mFJC model will be higher than true values. As the force-extension relation for a FJC is monotonically increasing, the order of tensions should not be changed by the mFJC interpretation (i.e. a tightrope with 122 nt across the gap will have higher tension than a tightrope with 144 nt across the gap). This means that results may not be inconsistent with the findings of Whitley et. al. [178], who do not cover a range of tensions below 2 pN. Indeed their model predicts non-monotonicity in k_{off} and K_D but their data is not sufficient to show this. My results on the other hand show the non-monotonicity.

4.6.5 Interpretation of two-population fit

The model with two-populations appears to better describe the data than simple bi-molecular association with one population. This might indicate that there are indeed two-populations of species undergoing the hybridization reaction. Could the two-populations be fully folded and misfolded/unfolded origami? If this was the case what would be the expected result, and why would only a single off-rate be observed? The observation of a single off-rate suggests that either all dissociation processes occur with a single off-rate; or that there are two or more off-rates but that these vary by so many orders of magnitude that they are not observable on the same timescale (effectively all of the 'faster' dissociating population has dissociated before any appreciable dissociation of the 'slower' decaying population). A single off-rate is implausible as the observed off-rate varies with tension, and therefore it seems improbable that both folded and misfolded tightropes would always respond to tension with identical off-rates, regardless of the nominal folding tension. Two off-rates with such different magnitudes that only a single off-rate could be observed in the timescale of the experiment would require the reaction with the slower decay to be effectively irreversible, which, again, appears unlikely for two reasons: Firstly non-zero dissociation is observed in the M13 case (and in experiments with origami folded without RHS tension

setting staples - not shown), suggesting that misfolds which are M13-like or missing one side will still not display irreversible association. Secondly the fitted proportion p would indicate the final level to which the decay would reach, or the proportion of V which dissociate from the origami during the dissociation phase, which can be determined from the raw data (such as figure 4.7) by finding the raw fluorescence change during the dissociation phase divided by the raw fluorescence change during the association phase (using the dilution factor corrected baseline). This estimates the proportion of V which dissociates from the origami during the dissociation phase. The fitted p values do not correlate with the calculated proportions.

If the two-populations are not folded and misfolded tightropes, then could they be correct and faulty V strands? During synthesis some V strands may not be correctly synthesised, either during the DNA synthesis or fluorophore conjugation. Different 'types' of V strand may interact with the tightropes in a different way producing different on-rates [187]. If this were the case, I would expect the proportion p to remain the same across all experiments which it does not (see figure 4.11c). Fixing the proportion p produces poor quality fits.

Instead of considering the two-populations as two distinct reacting species, the populations could be considered as two different pathways to reach the same final state. An example of this would be a V strand forming its initial base pairs with the tightrope either near the quencher, or forming the initial base pairs at the opposite end of the domain from the quencher probe. If internal 'zippering' of the DNA from the first base pairs into a complete duplex is slow compared to the initial base pair formation, the two-populations could simply be duplexes which initiate from the quencher proximal end of the duplex and duplexes which form from the quencher distal end of the duplex. Quenching would be rapid for the former, as even before complete duplex formation the fluorophore and quencher are close together, and for the latter there would be some rapid quenching contribution (as the fluorophore and quencher, though not adjacent, will be closer than they were in solution) followed by the slower recovery of complete quenching after the helix has fully formed. Both methods form the same final state and should therefore display a single 'off' rate. However, as the experimental concentrations of hybridizing species are low (tens of nM) it is unlikely that internal rearrangement would be slow compared to initial association [173].

4.7 Further Work

Though these investigations have produced some initial insights, more investigations are required to better understand the processes involved in hybridization to tightrope origami. Development of a method for direct measurement of tension, or an improved model for estimating tension such as oxDNA simulation, would better elucidate true tightrope tensions.

A greater number of intermediate tensions could be tested particularly those with between 122 nt and 102 nt across the gap, to determine the tension of the peak association rate. This tension of maximum rate might prove a better tension for operation of the bipedal motor, as it allows rapid hybridization and setting down of feet.

An improved estimate for K_D could be determined by varying concentrations of tightropes for constant total $[V]$, and measuring the equilibrium fluorescence of each to find $[VT]_{eq}$.

The dissociation phase could be attempted with a smaller excess of unlabelled V_{10} to confirm that the dissociation process was due to spontaneous melting rather than blunt-ended strand displacement by unlabelled V_{10}. Spontaneous dissociation rate would be concentration independent, whereas blunt-displacement rate would be concentration dependent.

Other dynamic processes could also be studied on tightropes. For example, the effect of changing tension on TMSD rate could be investigated. Tightropes with V and R pre-hybridized could form the TMSD substrate, and V could be displaced by an invading strand

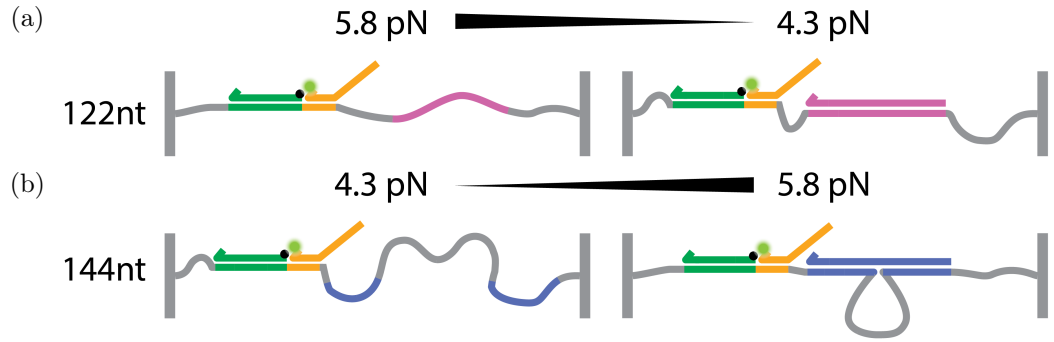

Figure 4.14: Adjusting tightrope tension after assembly to encourage allosteric hybridization or dissociation. **a** Tension can be decreased by hybridizing a domain of ssDNA to form a duplex. **b** Tension can be increased by gathering up additional ssDNA into a loop held in place by a pair of duplexes. Tensions calculated assuming two independent extensible helical duplex segments of 32 bp for the first duplex and either 47 bp or 42 bp for the second duplex in **a** and **b** respectively, with total nucleotides across the gap of 122 nt in **a** and 144 nt or 144-30 nt in **b** when the 30 nt loop length of ssDNA is gathered away.

as in chapter 2, resulting in an increase in fluorescence.

Given the enhanced stability of duplexes of tightropes with 5.8 pN tension, compared to duplexes of tightropes with 4.3 pN tension, dynamic adjustment of tightrope tension, could allosterically control equilibrium binding fraction. An example design is shown in figure 4.14.

4.8 Summary

Preliminary results suggest that tension in a single strand of DNA has a significant effect on hybridization kinetics which causes small variations in the stability of a duplex where one strand is under tension in a tightrope origami nanostructure. Some of the highest tension tightropes may be misfolded, however, there does appear to be a clear effect of changing tension. The relationship between on-rate and tension is non-monotonic, first increasing, then decreasing as tension increases. The maximum on-rate occurs with 122-102 nt across the gap, corresponding to tensions of 6-9 pN. The off-rate is also non-monotonic with tension, first decreasing with increasing tension up to 12 pN, when 94 nt are stretched across the gap, then increasing again. Stability, as inferred from equilibrium concentration of hybridized 10mer strand V_{10}, is not significantly changed by tension for tightropes in the range 6-12 pN with 122 nt - 94 nt across the gap, however the stability of the duplex on a 4.5 pN, 144 nt tightrope is reduced as shown by an increase in $\Delta\Delta G$ of $\sim$1 kcal mol^{-1}. This may be the free energy required to pull apart weak secondary structures present in the ssDNA at low tensions. The use of a distance clamp rather than a force clamp to determine hybridization rates is likely to have some influence on results as hybridization can change the tension of the ssDNA and hence free energy of the nanostructure, stabilising duplexes at tensions below the tension of ss=ds length and destabilising duplexes at tensions above this.

The process of hybridization to tightropes is not well described by a single reversible bimolecular association, and is far better described, by the sum of two reversible bimolecular reactions for two sub-populations. This may not be the only function to better describe the data. Other models might also fit the observations well, but if the two-population model is physically relevant it could suggest two different varieties of hybridizing species such as well-folded and misfolded origami, or two different pathways to the same final state.

Chapter 5

The Bipedal Motor

In this chapter I adapt the design of a bipedal motor [2] to disfavour a configuration which can inhibit the motor mechanism, and demonstrate a scheme for specific attachment of the motor to the start of a tightrope-based ssDNA track presented in chapter 3. I test the operation of the motor using truncated tracks with only two or three foot-binding sites by gel electrophoresis and bulk fluorometry, and perform similar observations on tightrope origami. I am able to observe changes of FRET between fluorophores on the motor consistent with foot lifting after fuel addition, and foot replacement after fuel cleavage, both on truncated tracks and tightropes. I am unable to detect significant quantities of motor reaching the end of the track. At the end of the chapter, possible explanations for the lack of motor progress are explored, and directions for future work considered.

5.1 Motivation

A bipedal autonomous DNA motor with a single-stranded DNA track has not previously been able to take more than a single step [1, 2, 152, 156]. I wished to enable the motor to make multiple steps and walk along a track. Informed by previous simulation results [123] I hoped to achieve the extended motion of the walker by applying tension to its track.

5.2 Challenges for Nanoscale Motors

An effective nanoscale stepping motor should outperform diffusive translocation. The motor must be able to move along its track faster than it would via purely diffusive mechanisms. If it cannot do this, then any directed progress will be dissipated away. Ideally

the turnover of the fuel or energy source should be tightly coupled to stepping. The fewer fuel molecules required to make a step, the greater the efficiency of the motor. In kinesin, a biological stepping motor, there is a 1:1 relation between ATP hydrolysis and forward stepping [188]. The motor should be processive. It should not dissociate from its track, but should remain bound. The motor should be directed. Directed motion can be achieved by a break in symmetry, energy dissipation and (for a two-footed motor) coordination.

5.3 Previous Design for a Bipedal, 'Kinesin-Like', Motor

5.3.1 Mechanism

The bipedal motor walks along a ssDNA track of alternating domains [1, 2] as depicted in figure 5.1. The two domains are a competition domain (orange) and a binding domain (green). A motor is composed of two identical feet, joined by a 16 bp duplex, the motor body (cyan), at their 5' ends with a flexible 4T spacer between body and foot (brown). Each foot has three domains which can bind to the track. From 5' to 3' these are a complement to the competition domain closest to the motor body (heel), followed by a complement to the binding domain (foot arch), and another complement to the competition domain (toe). If one were to imagine the motor movement akin to a person walking, it would move heel first i.e. forward motor stepping would be backwards human stepping.

Fundamental to its mechanism of operation, the motor must bind two consecutive foot-binding sites on a track. It is not possible for both feet to be fully hybridized when bound to consecutive sites, because only one foot can occupy the central competition domain. For a ssDNA track heading from 5' - 3', the leading foot toe will compete with the trailing foot heel to hybridize the competition domain, transiently exposing one or other of these domains.

Foot lifting

The exposure of toes and heels allows their use as toeholds[1] for toehold-mediated displacement of foot from track given an appropriate fuel strand. In previous studies [1, 2] two different fuel designs have been demonstrated. The first was a DNA hairpin with competition domain toehold after the 3' end of the hairpin stem. This fuel required an additional

[1]Not to be confused with toes.

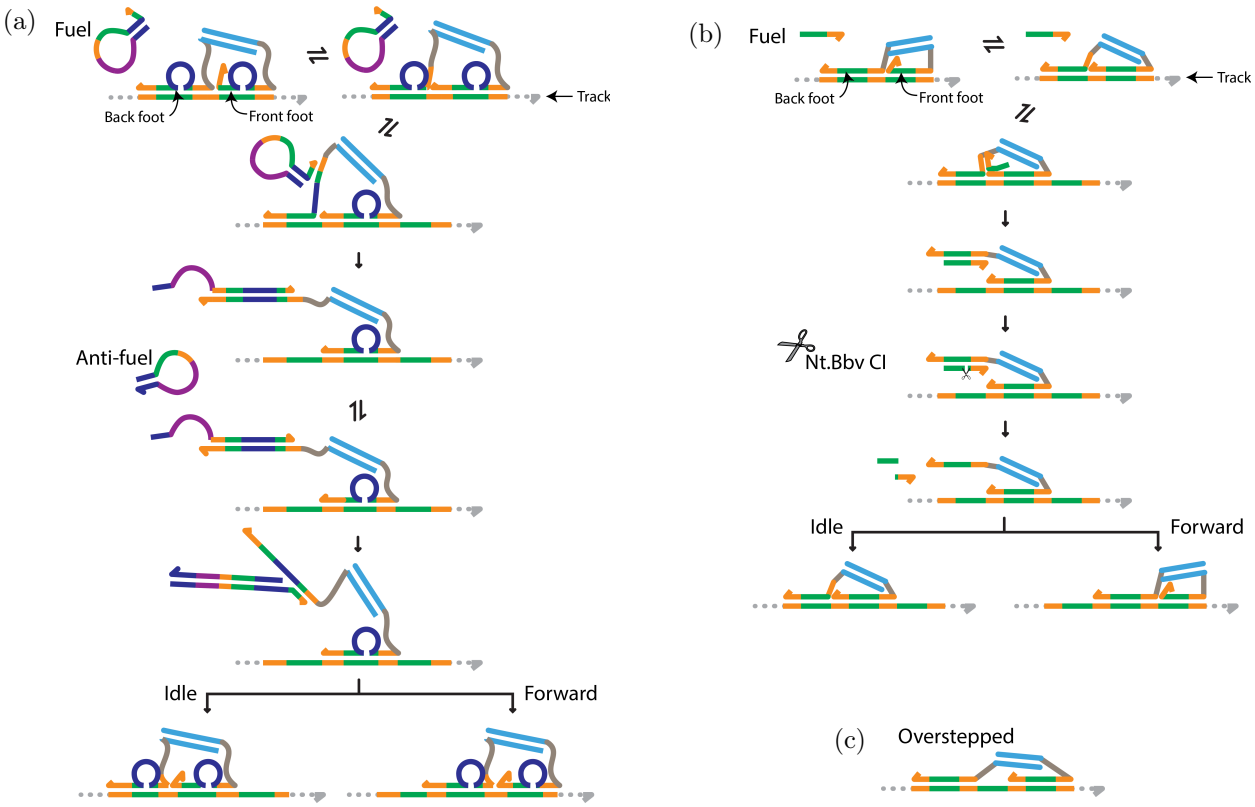

Figure 5.1: Bipedal motor mechanisms. Hybridization [1] **a** and enzyme [2] **b** driven variants of an autonomous bipedal DNA walker. Two identical feet are bound to adjacent foot-binding sites on a single-stranded DNA track. Feet compete over a central domain (orange), transiently exposing either front foot heel or back foot toe. When back foot heel is lifted this provides a toehold for a forward fuel. In **a** fuel is a DNA hairpin where the hairpin stem can be opened by displacement with a loop domain (dark blue) in the foot. In **b** fuel is ssDNA which can lift the foot by toehold exchange [39]. Fuel is removed to allow foot to rebind the track. In **a** this is by displacement of fuel with anti-fuel hairpins. The fuel:anti-fuel interaction is foot-catalysed. In **b** fuel is cleaved by nicking enzyme Nt.BbvCI. Short fragments dissociate from the foot. Lifted foot can rebind track behind or in front of the stationary foot, or can overstep - missing out a foot-binding site as in **c**. Figure **a** adapted with permission from reference 1. Copyrighted by the American Physical Society.

loop domain within the foot arch domain to thermodynamically favour the displacement of foot from track with fuel. The hairpin fuel ssDNA overhang has a toehold for lifting exposed heel (orange), and half of a foot arch domain (green), as far as the loop on the foot (dark blue). The stem of the hairpin fuel (also dark blue) is complementary to the loop domain in the foot, allowing the interaction between foot and fuel to proceed as follows:

1. Hairpin fuel toehold binds exposed back foot heel on track-bound motor.
2. TMSD between the fuel and the foot-track duplex then proceeds as far as the loop domain on the foot. When displacement has reached this stage, the loop will be open.
3. The opened ss foot loop domain (blue) can open the stem domain of the fuel hairpin.

4. The opened loop of the hairpin fuel contains the complement to the remaining track-bound domains of the foot, and can displace them by TMSD, lifting foot from track.
5. The opened loop of the hairpin has a domain (purple) for its later removal by anti-fuel.

In order to free the foot to rebind track an anti-fuel hairpin is required. The anti-fuel is the complement of the fuel hairpin but without the single-stranded stem overhang. It can displace the fuel from the foot by TMSD using a toehold in its loop [45].

The second fuel type uses a similar mechanism to lift foot from track and remove fuel from lifted feet. The feet in this variant do not have central single-stranded loops when bound to track. Fuel strands are simply a 5' binding domain followed by a 3' competition domain, able to hybridize either exposed front foot toe or back foot heel but only able to continue toehold-mediated strand displacement with the back foot after binding back foot heel. The remaining 6 nt back foot toe domain is expected to dissociate from the track after the displacement of the rest of the foot with fuel, lifting the foot. Instead of using an anti-fuel to remove the fuel strand from the lifted foot, a nicking enzyme Nt.BbvCI (New England Biolabs) is used to cleave the 16mer fuel centrally, producing two 8mer fragments. These fragments should melt and leave the foot single-stranded, able to rebind track. The analogous cut to track by Nt.BbvCI is prevented by including a mismatch between the foot and fuel at the position of the restriction site for Nt.BbvCI. The mismatch also shifts the thermodynamics in favour of binding fuel over binding track.

The two fuels described will preferentially interact with, and lift, the back foot and are therefore forward fuels. Thermodynamically, as the feet are identical, it is possible to lift either foot with the fuel but the kinetics of lifting the back foot should be more rapid because of the fuel directionality, and therefore the back foot lifted isomer is the favoured kinetic product. Fuels which preferentially lift the front foot (reverse fuels) can also be produced. In the case of the enzyme driven fuel variant this is as simple as reversing the domain order on the fuel strand so that it has a 5' competition domain and 3' binding domain.

Foot replacement

The free foot has various options for rebinding the track. The null outcome is that the foot rebinds the track exactly where it has just come from i.e. for the forward fuel case the

foot rebinds track immediately behind the stationary foot. This is referred to as an idle step, as the motor has not made any progress either forward or backward along the track. Alternatively the foot can rebind the track on the opposite side of the stationary foot i.e. immediately in front of the stationary foot in a forward step. A motor which makes only forward and idle steps will make net progress forwards.

A third option is for the motor to overstep and bind a non-adjacent foot-binding site, either in front or behind the stationary foot (see figure 5.1c). When bound to non-adjacent site there will be no competition between front foot toe and back foot heel to expose toe-holds for interaction with fuels. The motor is said to stall when in this state. To restore an operational motor configuration one foot must dissociate from track via an alternative mechanism such as blunt-ended strand displacement with fuel which is much slower [38].

Simulations using oxDNA have shown that overstepping by more than one site can be reduced by increasing tension applied to the ends of the track from 7.3-14.6 pN [189], however a singly overstepped state is still the modal result for foot replacement under 14.6 pN tension [123]. One reason for this is that greater applied forces extend the single strand of DNA which forms the track, an entropic spring, and by doing so move the position of an overstepped state further to the extrema of the probability distribution of positions for a free foot, reducing the likelihood of the foot binding such a site. Tension can also increase the dissociation rate for feet which do overstep, as the overstepped motor would be strained and prone to unpeeling of the front foot heel [123, 189].

Simulations have also shown that the direction of rebinding is influenced by the order of fuel fragment dissociation [123]. For both forward and reverse fuel types with the enzyme driven scheme, motor efficiency (defined as the bias of stepping over idling) is improved if the fuel fragment closest to the motor body dissociates first.

5.4 Design Modifications

I consider changes to the enzyme driven design to address three main points: Reduce overstepping, Improve motor efficiency and Introduce track tension.

5.4.1 Non-identical feet to reduce overstepping

In an effort to reduce overstepping, particuarly as even under applied tension of 14.6 pN a singly overstepped state is still the most common result of foot replacement [123], I broke the symmetry between feet by giving each foot a unique arch domain, but leaving the toe and heel the same in both preserving competition between the feet. This prevents overstepping by an odd number of foot-binding sites, because the arch domain no longer matches. Instead of overstepping by one site, the motor must overstep by at least two foot-binding sites in order to fully hybridize. Use of two different foot types requires two different fuel types for movement in a single direction, one per foot.

Two feet with Nt.BbvCI

I considered sequences for pairs of feet which allow the use of Nt.BbvCI to cleave fuel. Nt.BbvCI recognises the sequence 5'-`CCTCAGC`-3' and cleaves after the first two nucleotides, between `C` and `T`. Possible foot sequences will be highly constrained by the sequence requirements for Nt.BbvCI. Using domain lengths of 10 nt for the binding domain and 6 nt for the competition domain there are 16 possible positions for the enzyme recognition site, however, all designs either produce fuel fragments longer than 11 nt which may not melt from the foot at opeating temperature, or only cleave one direction of fuel. If using two copies of the enzyme recognition site which do not overlap, there are 23 permutations, and as the start and end nucleotide of the recognition site are both C, positions with two overlapping recognition sites `CCTCAGCCTCAGC` can also be considered, of which there are a further 16. Restricting these 39' possible sequences to those which produce fragments no longer than 9 nt, and can cut both forward and reverse fuel at least once reduces the number of choices down to 10. From this set of 10 restriction site positions, we require two sequences which are able to share the competition domain, but have sufficiently different sequences in the binding domain. 6 pairs of sequences have compatible competition domains. The binding domain of each of these pairs were compared by sequence alignment [190] which produced scores for pairs of sequences based on sequence similarity. All compatible pairs of binding domain sequences were found to all be highly similar scoring between 29.5 and 38 where a score of 0 is given to orthogonal strands and a score of 50 to identical strands. I conclude that, although there

are many appropriate designs using Nt.BbvCI for a single foot type, there are no solutions for two different foot sequences which fulfil the sequence constraints of the enzyme, the requirement for forward and reverse movement, the requirement for orthogonality, and the requirement that fuel fragments produced dissociate at operating temperature.

Two feet with alternative enzyme and fuel combinations

To remove enzyme-imposed sequence constaints I consider enzymes which are sequence independent. Two candidates are APE-1, a DNA repair enzyme which cleaves the phosphodiester DNA backbone to the 5' end of an abasic nucleotide in a duplex, and RNAse H, a RNAse which specifically digests RNA when hybridized to DNA.

Mismatches and sequences

To use APE-1 for fuel cleavage, the fuel strand requires an abasic nucleotide modification in place of one of the fuel nucleotides. At least one mismatch between the foot and track is required to maintain the thermodynamic preference of binding fuel over track, especially as the presence of the abasic site will reduce the number of base pairs between fuel and foot. The position of this mismatch between the foot and the track is now decoupled from the enzyme cleavage position, as it is not required to prevent the cleavage of track. We can choose the mismatch position to kinetically favour particular displacement reactions (see chapter 2 and reference 124). The same is true if using an RNA fuel or DNA/RNA chimera fuel strand and RNAse H enzyme, no mismatch is required to prevent cleavage, however a mismatch is desirable to favour fuel binding, and therefore the position of the mismatch is not constrained by the position of cleavage by RNAse H. In a fully RNA fuel, enzymatic digestion can occur between all ribonucleotides.

I chose the mismatch to be asymmetrically positioned in the binding domain, but in the same relative position for both foot types. The mismatch is placed at a position two bases from the heel competition domain (this is 9 bases from the toe competition domain). This position will provide a mismatch repair site close to the toehold for back foot lifting with forward fuels which was shown to provide near maximum displacement rate in chapter 2 so

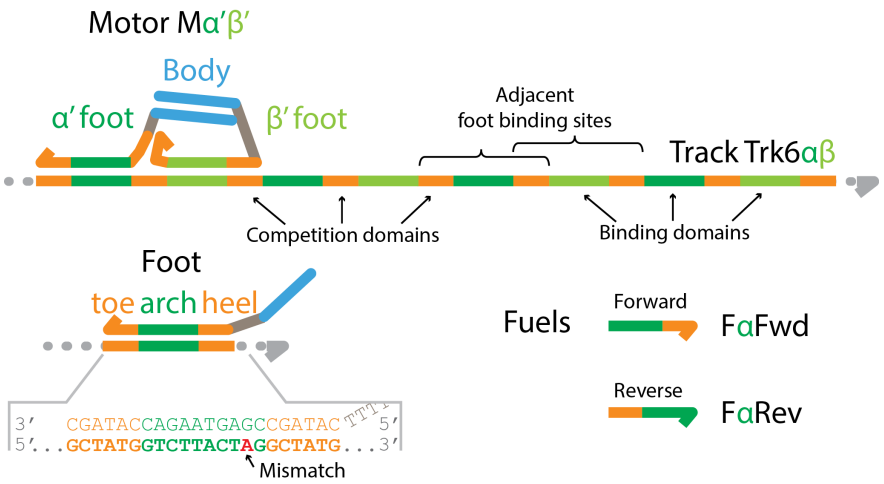

Figure 5.2: Nomenclature for the bipedal motor. The track is a single strand comprising of alternating binding and competition domains. The motor has two different foot types α'-foot and β'-foot, and there are two corresponding types of binding domain. Both feet share a common competition domain. A foot-binding site is a binding domain with a competition domain either side. Foot binding sites overlap as indicated. A motor is formed of one α'-foot and one β'-foot joined by a duplex called the body. A foot has a toe domain at its 3' end, a central arch domain, and a heel domain closest to the motor body. The foot-track duplex contains a mismatch close to the heel. There is no mismatch between foot and fuel. Fuels can drive the motor either forward or reverse depending on the domain order.

forward foot lifting kinetics should be rapid. Displacement of foot from track by reverse fuels on the other hand involve a late mismatch repair and will therefore experience less of a 'speed up' effect from the mismatch than if the mismatch was closer to the heel. Consequently I expect the motor to operate better in the forward than the reverse direction.

I chose to place the abasic site on an APE-1 substrate fuel such that the fuel would be cleaved centrally with a 3' 8 nt fragment which can form 8 bp with foot, and a 5' 8 nt fragment which can form 7 bp with foot. After cleavage it is anticipated that the 7 bp fragment should dissociate more rapidly, and this will be the fragment closest to the walker body which simulations suggest should improve foot replacement bias in the desired direction of motion.

The sequences for both binding domains, the competition domain, and the motor body were designed using NUPACK software [150].[2] An annotated schematic for the redesigned motor is shown in figure 5.2, along with nomenclature which will be used in the remainder of this chapter to describe components or domains of the motor system. Components such

[2]At the time of writing, the software would not permit the specification of pseudoknotted structures, of which the motor is one, so feet and body were designed as separate structures then joined together by hand.

as feet, fuels and track binding sites are referred to by Greek characters, or primed Greek characters. A component labelled with a Greek character should have domains complementary to the same primed Greek character so that they can hybridize. Fuels are abbreviated F, followed by the Greek letter specifying their binding domain, and directionality (Fwd or Rev for forward and reverse respectively). Motors are abbreviated M, followed by primed Greek letter corresponding to the domains for the feet. Tracks are abbreviated Trk, followed by the number and/or order of foot-binding sites.

5.4.2 Tension

Tension to the track is applied using the tightrope DNA origami nanostucture presented in chapter 3. In this chapter I use the tightrope with a 102 nt track of 6 foot-binding sites (three complete motor cycles) stretched across the 35 nm gap. The estimated tension across the track is ∼9 pN (see appendix E). In chapter 4 I found this tension to be a good compromise tension for high on rates and low off rates for hybridization.

5.5 Loading the Motor onto Tightropes

For observations of the motor on tightrope origami I require a single motor per tightrope, and a specific start position on the track. Loading the motor on the track by simply incubating motor and track can produce tightropes with multiple motors bound, and motors could bind at any position on the track. I implement a method to slow the interaction of motor feet with track, and to introduce a specific interaction of motor with one end of the track to reduce multiple motor binding and synchronise loaded motors at the track start.

5.5.1 Mechanism

Foot blockers and loading domain

To reduce interaction between the foot-binding sites on the track and the motor feet, fuel-like 'load block' strands were used. Both motor feet can be pre-lifted with load block strands before adding to the track. A toehold sequence was appended to one end of each load block strand which can be used to remove the strands from the feet by TMSD. By

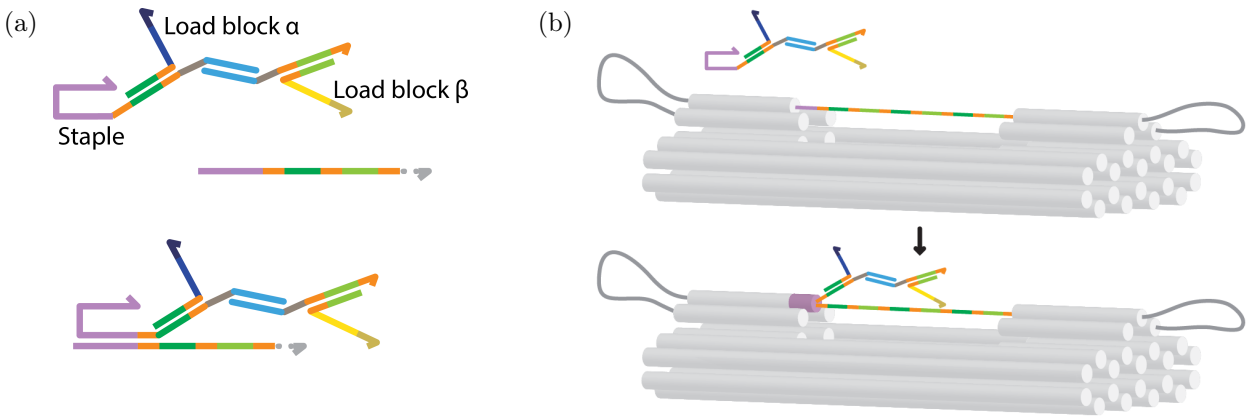

Figure 5.3: Loading biped with staple extension and blocked feet. The motor is constructed with load block strands for each foot which are forward fuels with toeholds for their removal (blue and yellow). The load block strands reduce the rate of interaction of motor with track, as the binding domains are sequestered in duplexes. α'-foot is extended by a the sequence of a staple in the tightrope origami found to the 5' end of the scaffold track. By appending the motor with this staple, and omitting the staple from tightrope origami assembly, the staple domain can guide the motor to the track start.

using orthogonal toehold sequences for the two load block strands, they can be removed independently and sequentially.

If the motor is blocked it should not interact with track, at least not rapidly. To introduce an interaction between motor and track which would favour the motor binding to a specific binding site, an additional domain was added to one of the motor feet. This design is shown in figure 5.3. The foot which can bind the first foot-binding site on the track is extended by a loading domain which is complementary to the scaffold sequence immediately adjacent to the 5' end of the inserted track DNA. This loading domain forms part of a staple used to assemble the tightrope origami, named the loading staple. In order to use the domain for loading, this staple must be omitted from the origami assembly. Motors appended with loading domains or loading staples will preferentially bind to the foot-binding sites close to the loading staple position (at the track start).

Release and cleavage

Once the motor has been loaded into the origami it can be released by the ordered addition of two release hairpins as shown in figure 5.4. The first hairpin to be added is the release for α-foot, the foot extended with the loading staple. This hairpin has a toehold which can bind to the toehold of load block α, and then open the hairpin stem by TMSD.

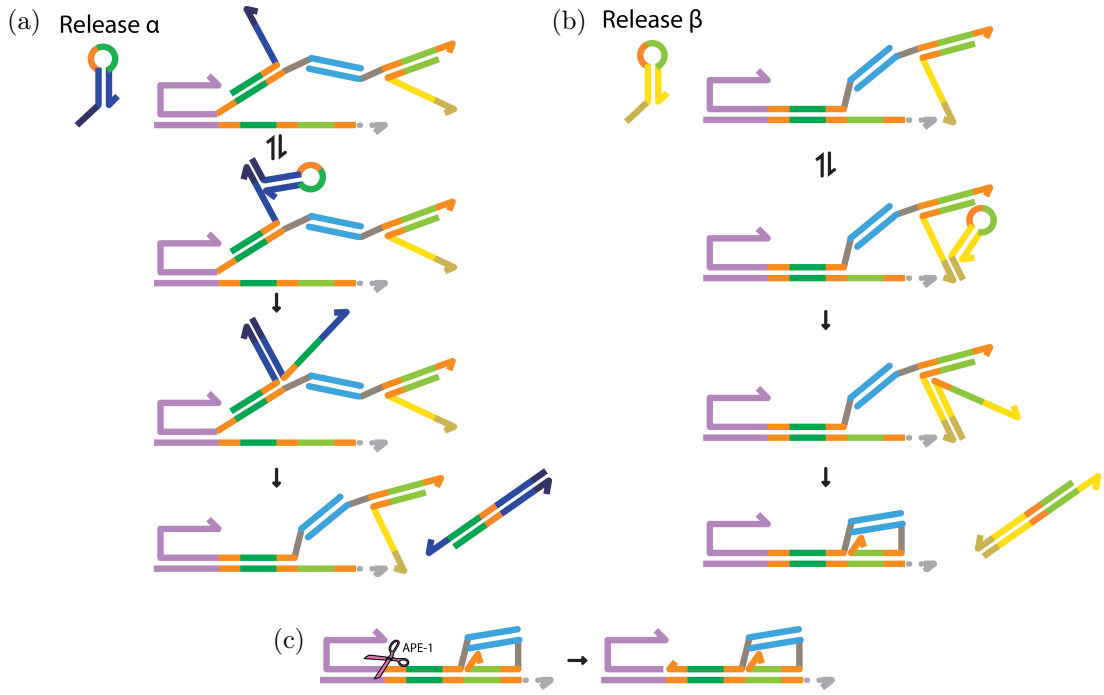

Figure 5.4: Removing blocking strands to sit the motor on the track with release hairpins. Release α is added first, **a**, then release β, **b**. Both hairpins strip the load blocks from feet by first binding via a single-stranded toehold domain on the hairpin, then opening the hairpin stem by toehold-mediated strand displacement with the load block, then displacing the fuel-like blocking domains from the foot using foot-like domains previously in the hairpin loop. The product should be an inert duplex as no foot-like or fuel-like domains are exposed. **c** Motor is cleaved from loading domain by APE-1.

The product duplex hides all domains which are fuel-like (e.g. load block strands) and could interfere with motor feet, and also hides all domains which are foot-like and could bind track. Release hairpins should be topologically restricted from binding to track with their foot-like domains without opening the stem. Spontaneous stem melting is unlikely at operating temperatures.[3]

After release, the motor must be cleaved from the loading domain to free its back foot. This is achieved by introducing an abasic nucleotide immediately after the toe domain of the back foot. APE-1 will cleave to the 5' end of this nucleotide when it is enclosed in a duplex (i.e. after loading), removing the loading domain and freeing the motor to walk (figure 5.4c).

5.5.2 Loading tests

The principle of loading with a staple or staple domain was tested with oligonucleotide test tracks extended by the sequence complementary to the track start staple. In figure

[3] Stem melting temperatures of release α and release β are 73°C and 79°C respectively [150].

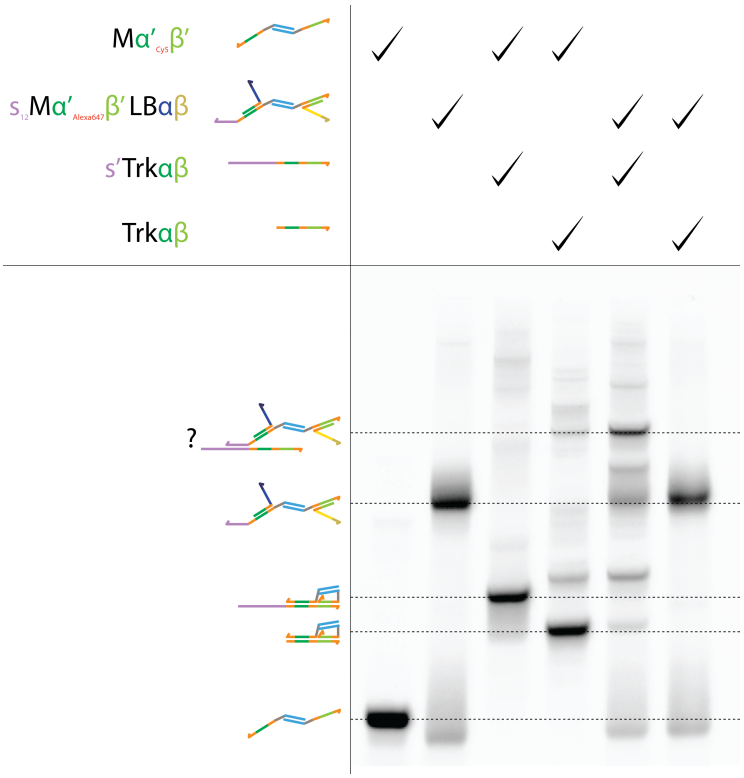

Figure 5.5: Testing the role of loading domain in adhesion of blocked motor to track. Gel shows Cy5 fluorescence from Cy5 labelled α'-foot. Unblocked motor without loading domain (lane 1) and blocked motor with loading domain (lane 2) were mixed with two-site tracks either with or without the loading staple complement domain and incubated for 1 hr. Unblocked motor interacts with tracks with and without the loading staple complement (lanes 3 and 4 respectively). Blocked motor with loading domain interacts with track with loading staple complement (lane 5) but does not interact with track without loading staple complement (lane 6 which appears similar to lane 2). Bands indicated with dashed lines are identified as the complexes depicted on the left hand side. The uppermost band indicated may be loadable blocked motor bound to track by its loading domain although the gel does not contain the appropriate control species to prove this.

5.5 for the loading domain (pink) I use not a full staple, but the domain of the track start staple which is on the track bearing helix, 12 nt long. Two forms of the motor are used, the first is an unblocked motor without a loading domain, the second is a blocked motor with loading domain. Two test tracks are used, one with the staple complement at its 5' end, the other without. All combinations of the two motors and tracks were mixed together with a motor:track stoichiometry of 1:1.5 and incubated for 1 hr. The products formed were run on a PAGE gel and scanned for red fluorophore emission. The α'-foot of each motor has a red fluorophore (Cy5/Alexa647) so any bands produced must contain α'-foot. Unblocked motor can hybridize either track type producing a majority band corresponding to a single fully-bound motor. Higher order bands are likely to be motors with one track-bound to each foot, or topologically-linked track-bound motors. Mobility is reduced when the motor is bound to

a track with the staple complement domain, rather than a track without, as expected. When blocked loadable motor is incubated with track with loading staple complement domain, a number of new bands form, indicating that the motor is able to interact with the track. The exact configuration and composition of species in the bands cannot be determined without additional control lanes not included in this gel, however, the band with greatest intensity might be blocked motor bound to track by its loading domain, as it has mobility lower than blocked motor alone. Critically, when loadable blocked motor is incubated with track *without* staple complement, there is no change in motor bands produced from a sample of blocked motor alone. This indicates that, as designed, the mechanism for blocked motor interaction with the track is primarily via the loading domains and not, for example, via the exposed toe domains. The load block strands prevent interaction with track via the feet.

Proving the loading domain works as a method to place blocked motor to an oligonucleotide test track does not prove that this method is capable of loading the blocked motor into origami. Origami differs from oligonucleotide tracks in many ways. The electrostatic environment is different in the origami, and sterically it may be more difficult to hybridize to a domain in origami due to many bulky DNA helices nearby. Topologically the end of the loading domain is fixed in the origami whereas it is free on the oligo. All of these effects may cause the loading domain mechanism to fail in origami. To test loading onto origami, blocked and unblocked motors were incubated with origami lacking the staple used for loading in their assembly. The motors were added in 3× excess to origami which has a six-site track, and therefore is capable of hybridizing three motors by both of their feet. Results are shown in figure 5.6. In all cases when motor is not blocked, more motors bind to the tightrope. There may be multiple motors binding per tightrope. Gels show that when the motor is blocked and has no loading domain, very little motor sticks to the tightrope, however more blocked motor binds when the loading domain is present, than blocked motor binds when there is no loading domain. These results indicate that the loading domain enables blocked motors to bind to origami, and that blocking of feet may reduce the probability multiple motors binding to origami. It is not possible to be certain about the number of motors per origami with this method, however, the relative integrated intensities of gel bands from

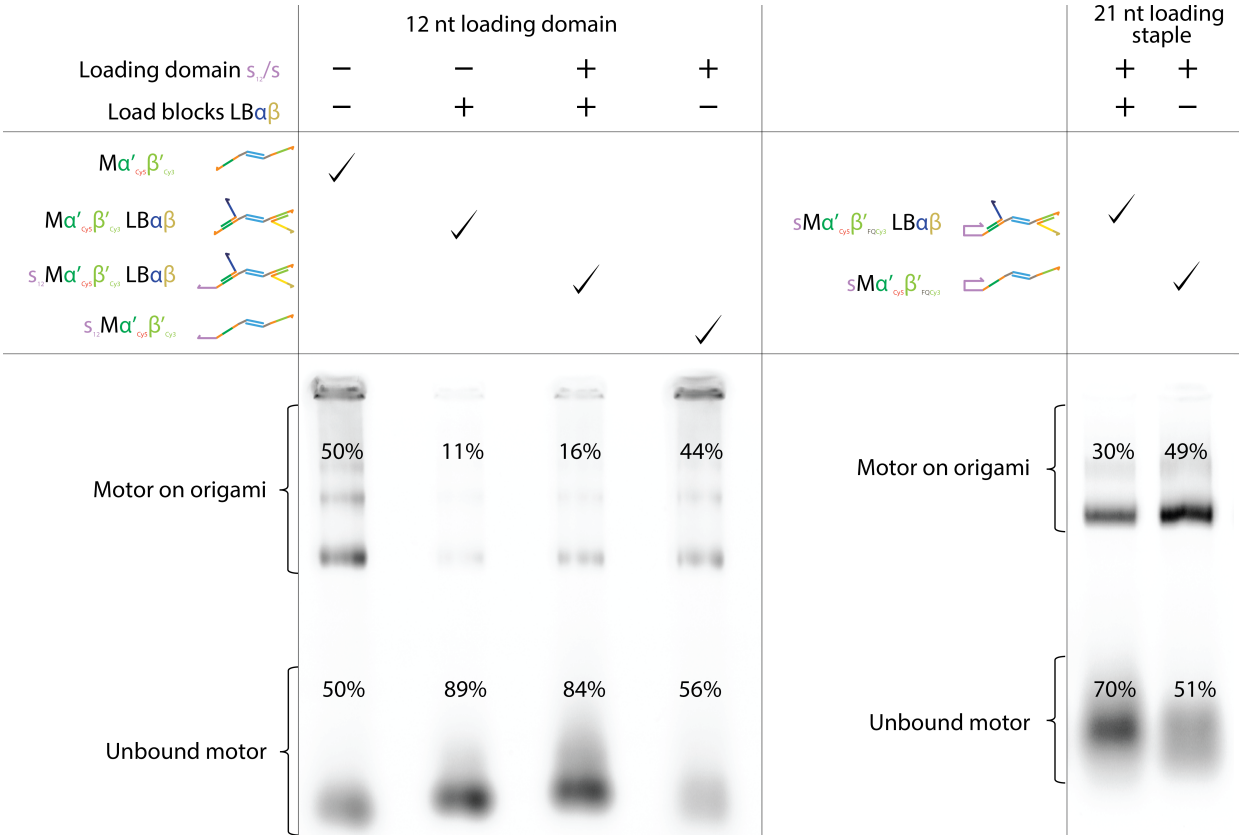

Figure 5.6: Loading tests with blocked and unblocked motors on tightropes. Gels show Cy5 fluorescence from Cy5 labelled α'-foot for tightrope origami incubated with 3× excess motor for 2 hrs at room temperature (left) or 1 hr at 37°C (right). Left gel shows motors with and without a 12 nt loading domain and with and without load block strands. The largest quantity of motor on origami is produced when there are no load blocks (lanes 1 and 4). With load blocks, more motor binds to origami with the loading domain (lane 3) than without loading domain (lane 2). The right gel shows motors with a 21 nt loading staple . With load blocks there are fewer motors bound to origami, and more unbound motors, than without load blocks. Percentage of integrated intensities for origami-bound and unbound motors are marked on the lanes. For quantization well intensity was discounted as it is optically different from the rest of the gel.

bound and unbound origami would suggest greater than 33% of motors are bound in all unblocked cases. As motors are in 3× excess to origami this would indicate >1 motor per origami. Stoichiometry could be confirmed by single molecule measurements counting Cy5 fluorophores per origami (e.g. TIRF or confocal microscopy).

5.5.3 Testing release mechanism

Is the release mechanism necessary to allow the motor to sit on the track?

Although the kinetics of blocked motor interacting with track have been demonstrated to be slow in the absence of the loading domain, the loading staple itself can be considered a toehold for a strand displacement reaction which will displace load block α from α'-foot.

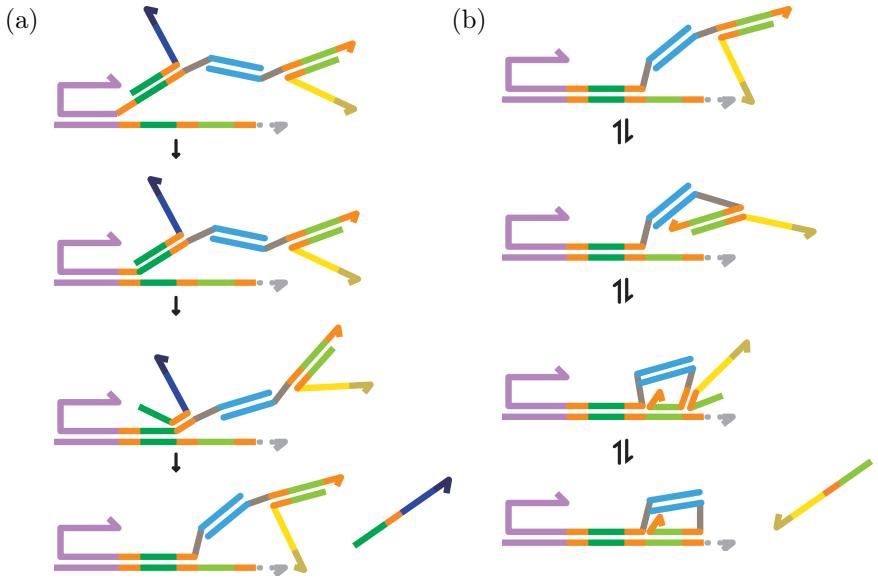

Figure 5.7: Possible mechanism for spontaneous removal of load blocks from loaded motors. **a** The loading staple acts as a toehold for displacement of load block α from α'-foot by track. **b** The loading staple is a remote toehold [144], increasing the local concentration of blocked β'-foot at the track binding site where a displacement reaction can occur.

The reaction would lead to the eventual placement of α'-foot at the track start without the use of release hairpins. The staple loading domain on α'-foot cannot be considered as a proximal toehold for the reaction between blocked β'-foot and track, but the local concentration of blocked β'-foot at the foot-binding site is increased by the loading process, so it can be considered a remote toehold for this reaction [144]. A possible mechanism for the spontaneous release of load block strands from motor feet are shown in figure 5.7.

Spontaneous and release-hairpin-triggered setting of motors on tracks were tested using PAGE. Results are shown in figure 5.8. Blocked or unblocked motors with loading staple domains were incubated with a two foot-binding site track Trk2αβ with loading staple complement domains. Unblocked motor and track ran as a single band of motor sat fully on track (lane 3). Blocked motor and track produced a number of bands (lane 4), however the majority band appears to be the fully track-bound motor indicating a significant amount of spontaneous release of blocked feet. Control lanes verify that the other 'motor' bands (i.e. those with both Cy3 and Cy5 signal) in lanes 4-6 correspond to blocked motor, and motor with α'-foot sat fully on the track, but β'-foot not sat on track. After addition of both release strands (lane 7) all of the motor is fully sat on the track as in the unblocked case. These results show that the release mechanism can remove load block strands from motor

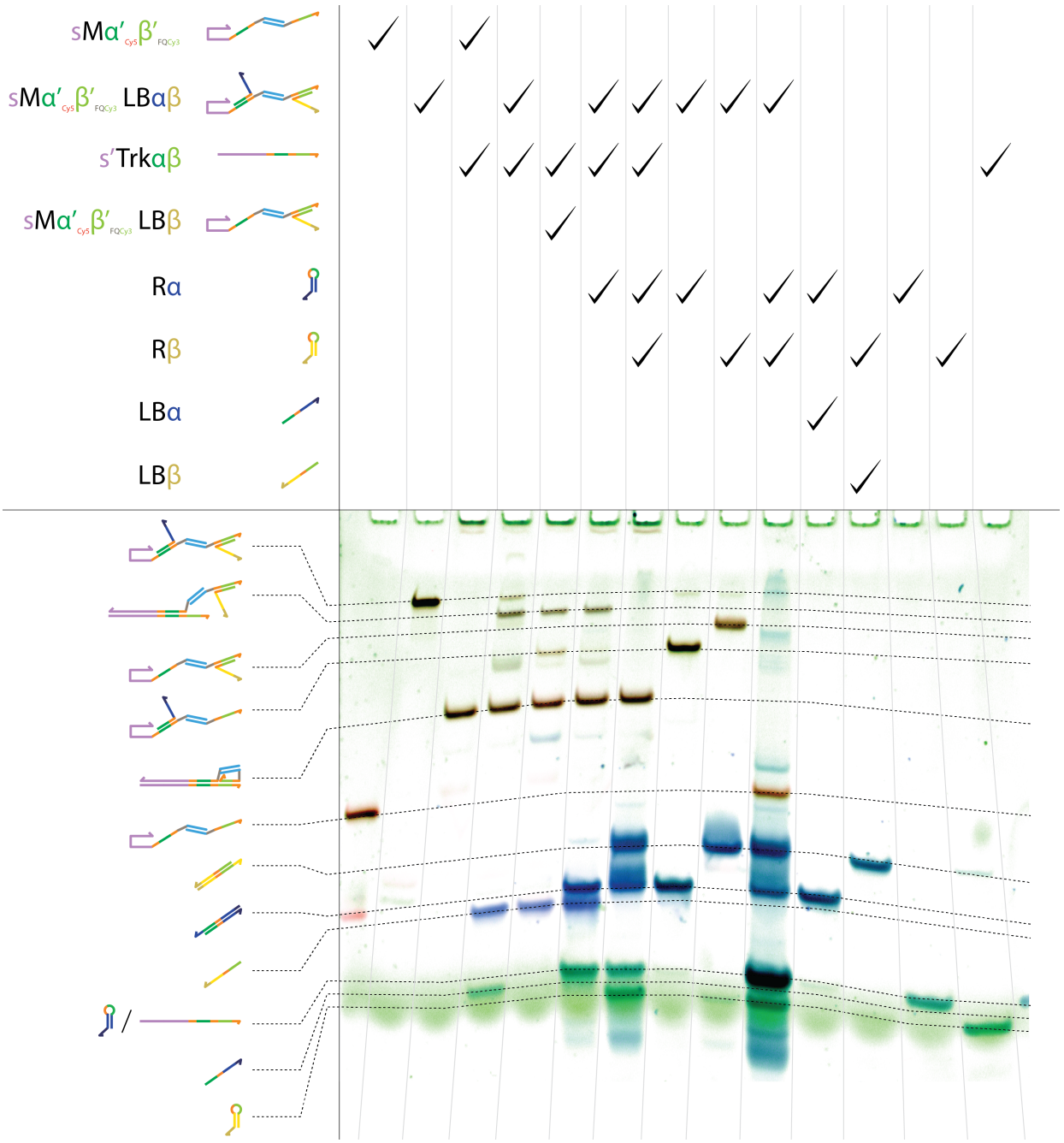

Figure 5.8: Testing the release mechanism with PAGE. Gel shows RGB overlay of Cy5 on α'-foot (red), Cy3 on β'-foot (green), and Sybr Gold stain (blue) fluorescence. Lane 1 shows the migration of unblocked motor, lane 2 shows the migration of blocked motor. In lane 3 the unblocked motor is incubated with a two-site oligonucleotide track with loading domain s'. A single band formed shows the position of motor sat fully on track with no blocking strands. Lane 4 shows the motor with both feet blocked incubated with the track. The majority of motor sits fully on the track despite bearing blocking strands. Another band in lane 4 corresponding to α'-foot sat on track but β'-foot bound can be identified by comparison to lane 5 where motor with only β'-foot blocked is incubated with track. Lane 6 is blocked motor and track incubated with release strand Rα. Lane 7 is blocked motor and track incubated with both release strands sequentially. Lane 8 is blocked motor, incubated with release Rα. Lane 9 is blocked motor incubated with release Rβ. Lane 10 is blocked motor incubated with both release hairpins sequentially. Lane 11 is load block α incubated with Rα. Lane 12 is load block β incubated with Rβ. Lanes 13, 14, and 15 are Rα, Rβand track respectively. Complex stoichiometry for Motor:Track:Release is 1:1:1.5. Each incubation was for 2 hrs at 37°C.

feet, however spontaneous release of load block strands by track is possible.

5.5.4 Cleavage of loading staple from motor

Loading staple cleavage on oligonucleotide origami mimics

To test loading domain cleavage from foot on oligonucleotide test tracks, blocked and unblocked motors were incubated with two-site tracks with the full loading-staple-complement domain of 21 nt. The motors both had a loading domain of 12 nt (the first of which was abasic) such that when fully hybridized to track there will be a 9 nt single-stranded domain (part of the loading staple complement) left single-stranded. 0.5 $U\mu l^{-1}$ APE-1 was added to some of the samples, with others left as no-enzyme controls. The reaction buffer was RNAseH reaction buffer + Mg^{2+} (see appendix G) and all incubations were 1 hr at room temperature. After incubation with APE-1, the full loading staple was added to some of the samples. The full staple can displace cleaved 12 nt loading domains by TMSD. If the loading domain is not cleaved, it will remain attached to the motor-track complex after the displacing staple is added.

Results are presented in figure 5.9. The last 4 lanes corresponding to blocked motor confirm the evidence of figure 5.8, that load block strands can be spontaneously displaced as bands corresponding to load blocks can be seen. Both for motors originally blocked, and motors originally unblocked, when samples have been incubated with APE-1 and loading staple (lanes 8 and 12), a high mobility band appears, likely to be the cleaved loading domain. When the motor-track samples are incubated with APE-1 and no loading staple is added (lanes 9 and 13), this loading domain fragment band is also visible but more feint. This confirms APE-1 is able to cut loading domain away from foot when hybridized to test track in 1 hr at room temperature in RNAseH reaction buffer + Mg^{2+}. Although there is not a control lane containing the cleavage product, I expect that most of the substrate has been cleaved because in the Cy5 channel when lanes 8 and 10 are compared the lowest band in each has significantly altered mobility. As the contents of the two lanes are identical except for the presence of APE-1 in lane 8 but not in lane 10, this difference might be due to cleavage. If this is the case then the complete absence of the lowest band in lane 8 from

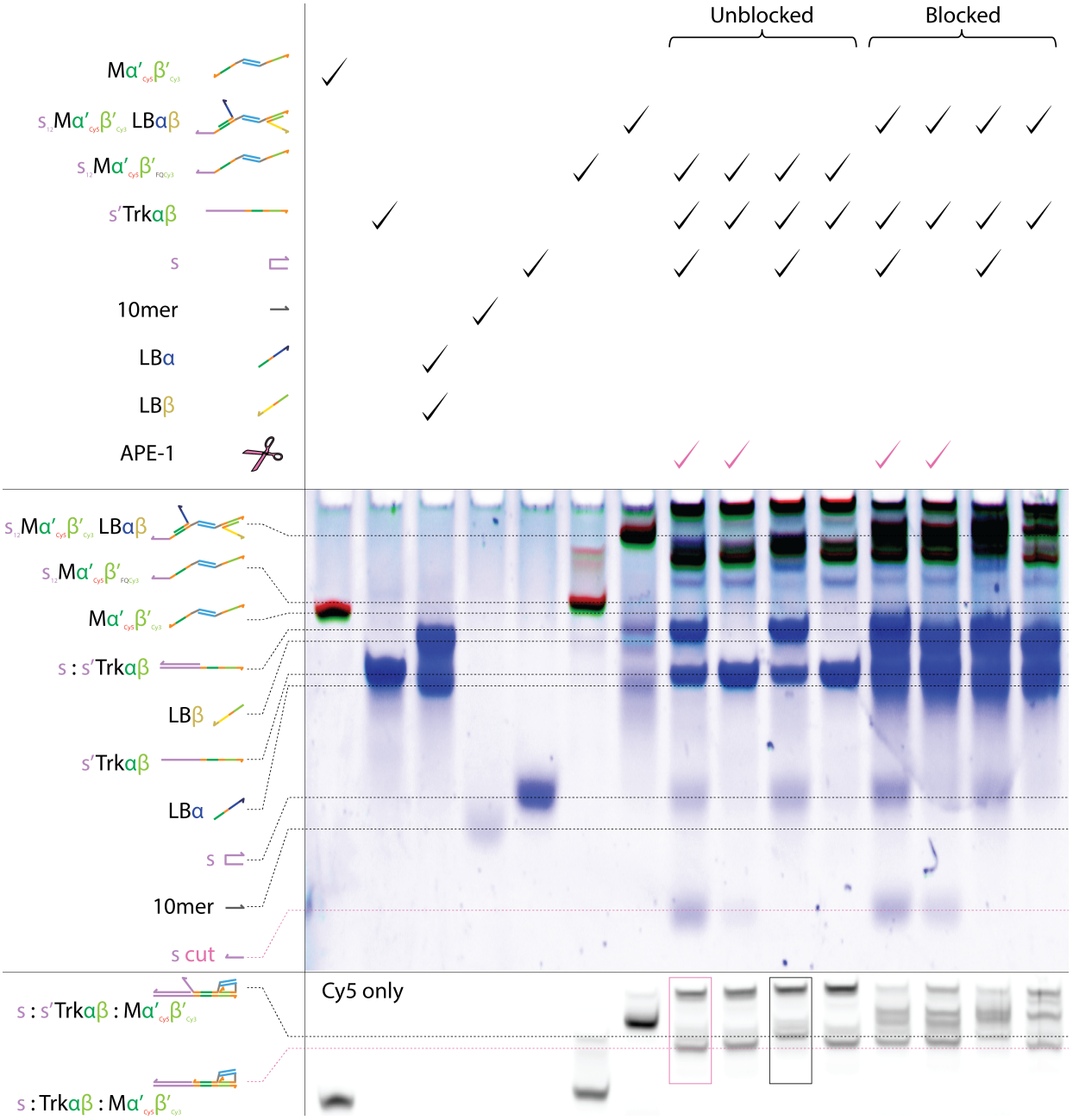

Figure 5.9: Test of staple cleavage from motor foot on oligonucleotide test tracks. Gel shows RGB overlay of Cy5 on α'-foot (red), Cy3 on β'-foot (green), and Sybr Gold stain (blue) fluorescence. Either blocked (lane 12) or unblocked (lane 8) motors with 12 nt loading domains were incubated with a 1.5× excess of two-site track with staple complement domain for 1 hr at room T in RNAse H reaction buffer + Mg^{2+}. APE-1 is added to achieve a final concentration of 0.5 $U\mu l^{-1}$ and the sample is incubated for another hour at room T. Finally a 2× excess of loading staple is added. This should be able to remove the cleaved loading domain from the track oligo by TMSD. Other lanes show control samples where one or more of these components are omitted. Evidence of a cleaved fragment (indicated by pink dashed line) can be seen in both lane 8 and lane 12, and in lanes 9 and 13 where APE-1 has been added but no displacer has been used. When no APE-1 is present the fragment is not observed (lanes 10, 11, 14, 15). All samples were treated with 0.2% SDS before running on the gel. Bottom panel shows same gel but only the Cy5 channel. Here differences in band mobility between samples with (pink box) and without (black box) APE-1 can be seen, which may be attributed to cut and uncut complexes depicted.

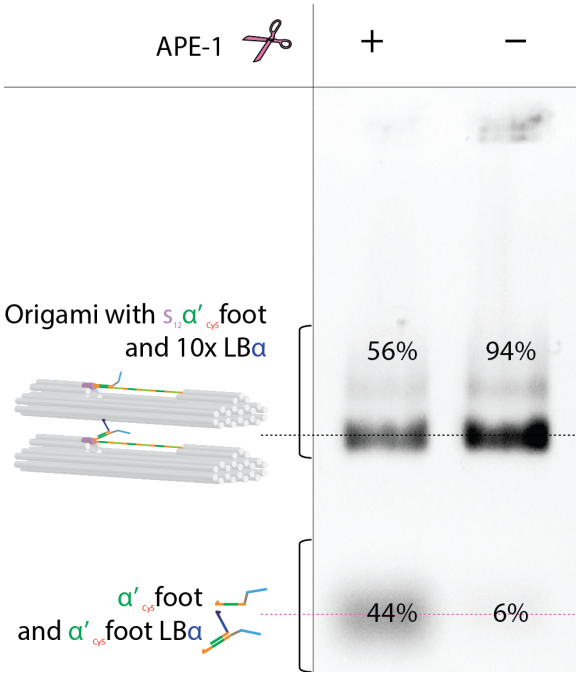

Figure 5.10: Test of staple cleavage from motor foot on tightrope origami. Gel shows Cy5 fluorescence from Cy5 labelled α'-foot. 20 nM unpurified origami annealed with 1× α'-foot, and 10× excess of load block α was incubated with (left) and without (right) 0.5 Uμl^{-1} APE-1 for 1 hr at room temperature. Successful cutting of the loading domain from α'-foot enables dissociation of α'-foot from the origami (possibly aided by load block α) creating a higher mobility foot band marked with pink dashed line. A higher proportion of free α'-foot is observed when APE-1 is present.

lane 10 indicates that undetectable levels of substrate remain.

Loading staple cleavage on tightrope origami

To demonstrate that cleavage of the loading domain from foot is effective on origami, a tightrope origami sample was annealed with 1× α'-foot and 10× load block α. The sample was not purified after assembly so there was still load block α in solution. The sample was diluted to 20 nM in a final buffer of Tightrope buffer supplemented with 50 mM NaCl. The tightrope sample was then split into two halves, and to one half 0.5 Uμl^{-1} APE-1 was added, but not to the other. After incubation at room temperature for 1 hr both samples were run on an agarose gel presented in figure 5.10.

The left hand lane containing APE-1 shows an origami band of lower intensity than the right hand lane origami band (band position is indicated by a black dashed line). A lower band containing small complexes, such as feet, in the left lane is of greater intensity than the equivalent band on the right hand lane. This band is produced by cleavage by the APE-1.

After cleavage of the loading domain, the foot is less strongly bound to the origami and more likely to dissociate from the origami. This dissociation may be aided by the load block strand which is fuel-like and can bind foot-binding and competition domains instead of the track. These results show that cleavage is possible on origami with APE-1, however, it does not provide information about the proportion of loaded motors which have been successfully cut, as to be in the higher mobility band a foot must have both been cleaved from its loading domain, and have dissociated from the origami. Feet which have been successfully cleaved, but have remained bound to origami, migrate with the mobility of the origami.

5.6 Foot Discrimination with Fuel

On an oligonucleotide test track with two foot-binding sites each foot and each track site is unique, so there is only one orientation in which the motor can bind. When two-site tracks with α and β foot-binding sites are incubated with motor Mα'β', the track-bound motor band formed has mobility dependent on the order of the domains. This is shown in figure 5.11. The first pair of lanes are controls for lifted front foot and lifted back foot respectively formed by incubating Mα'β' with a dummy foot-binding site γ, which is unable to bind either foot, either at the front or at the back. The last two lanes are Mα'β' incubated with track Trk2αβ, or Trk2βα. In these lanes many bands are seen, the uppermost bands are likely to be multimers[4], complexes containing multiple motors and/or multiple tracks. The lowest band is the fully track-bound motor band. In lane 3 and 4 the lowest band has different mobility. This could be due to differences in the centre of the structure due to uneven competition over the competition domain causing a different track-bound motor morphology, or it could be due to the positions of fluorophore labels on the structure which can alter the overall shape and therefore mobility of the structure.

5.6.1 Lifting of a specific foot with fuel

Using oligonucleotide tracks the motor operation mechanism was tested. The specificity of foot lifting with fuel was first investigated. A forward-type fuel is designed to preferentially

[4]Evidence from gels where track:motor stoichiometry was varied, and deviation from 1:1 stoichiometry increased the intensity of low mobility bands.

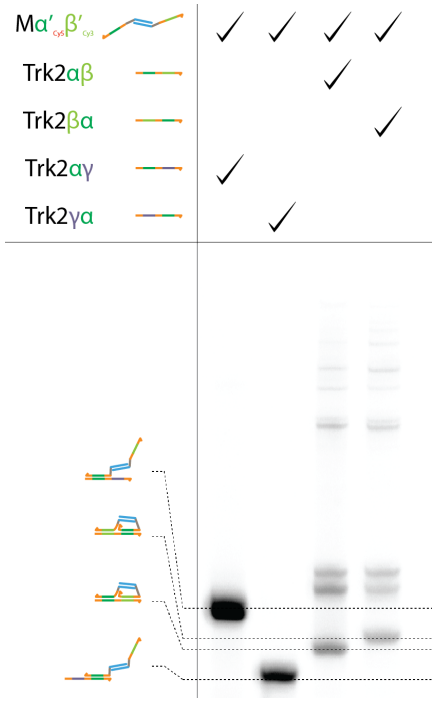

Figure 5.11: α' β' Motors incubated with 1.5× excess tracks. Gel shows Cy5 fluorescence from Cy5 labelled α'-foot. Lifted foot controls for lifted back and front foot in lanes 1 and 2 respectively are prepared by incubating Motor α' β' with tracks with a dummy front or back foot-binding site γ in place of the β foot-binding site. β'-foot can bind track in both cases. Lane 3 shows motor incubated with Trk2 αβ. The lowest band is assumed to be the fully bound motor band where the motor is attached by both feet. The highest bands are likely to be multimers due to their low mobility. Multimers could be multiple tracks bound to a single motor, multiple motors bound to a single track, or topologically interlinked motors bound to tracks. Lane 4 shows motor incubated with Trk2βα. The lower band in this lane has lower mobility than the lower band in lane 3.

lift back (trailing) foot, and interact with a front foot on a much longer timescale, allowing back foot lifting to be kinetically favoured. This means that only back feet should be rapidly lifted by fuel, and only when the fuel-binding domain matches the back foot.

Foot lifting was attempted with a variety of forward fuels, with either α or β binding domains, and composed of either DNA[5], DNA with a single central abasic site, DNA with a single central ribonucleotide substitution, or RNA.

Back foot is only lifted when the matching forward fuel is used

Results are shown in figure 5.12 for tracks with foot-binding sites ordered α β, and figure 5.13 for binding sites in the opposite order i.e. β α. Results show that back foot is only lifted when the matching forward fuel is used. In figure 5.12 this is in lanes 2-5 where FαFwd fuels

[5]A DNA fuel is not a true fuel as it does not contain a nicking site sequence and there is therefore no candidate enzyme with which to cleave it. It should however lift feet in the same way as a true fuel.

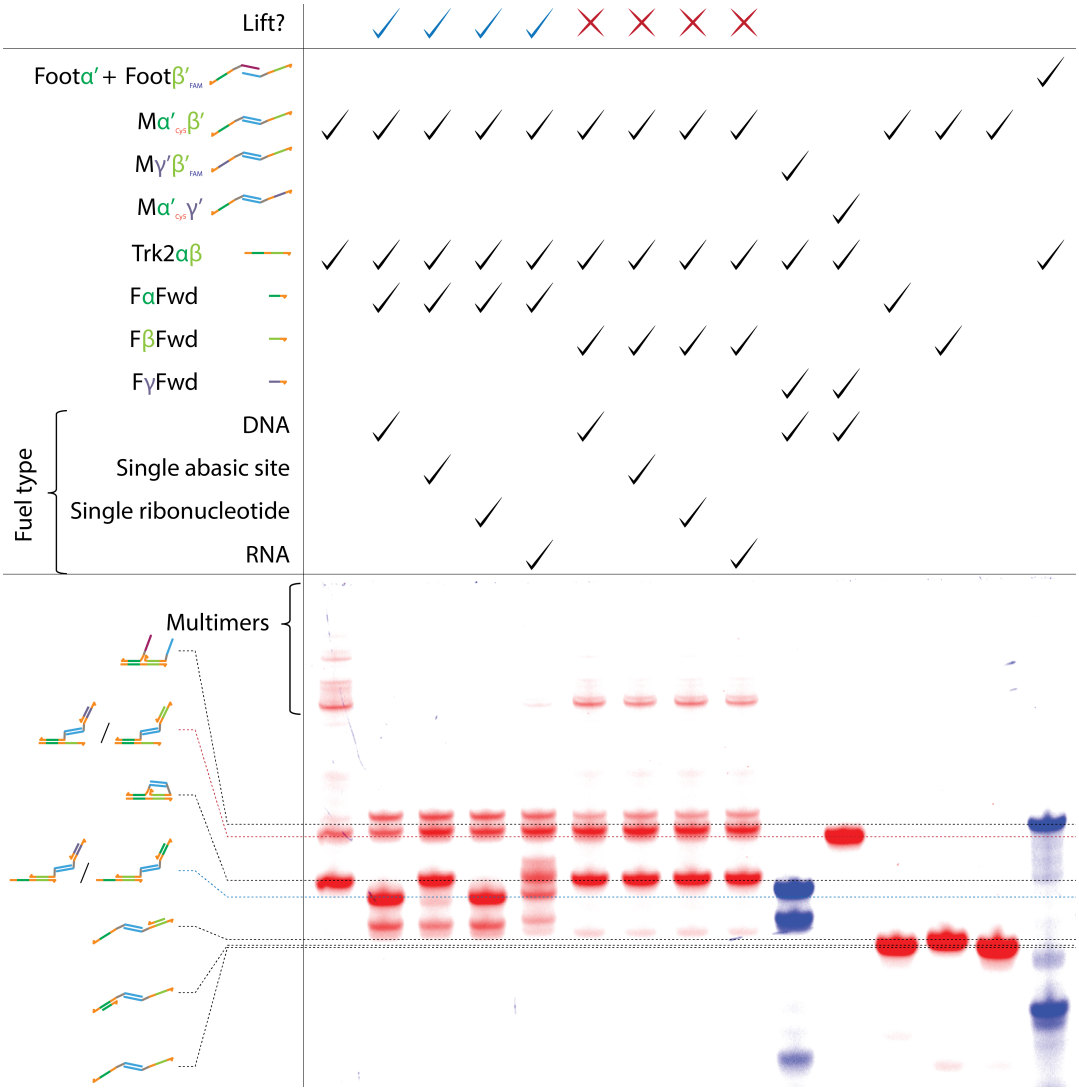

Figure 5.12: Foot lifting with forward fuels on two-site αβ tracks. Gel shows Cy5 fluorescence on α'-foot (red) and FAM fluorescence on α'-foot (blue). Forward fuel types were incubated with Motor Mα' β' bound track Trk2αβ to to determine if feet could be lifted by each fuel. Lane 1 shows motor bound to track with no added fuel and the majority band is annotated as the position of fully bound motor. Lanes 2-5 show the effect of addition of various forward α fuels. A fully foot complementary DNA FαFwd (lane 2) removes all track-bound motor band and a new lifted back foot band appears (marked by blue dashed line). A DNA FαFwd with a central abasic site (lane 3) lifts very little back α'-foot, the band corresponding to track-bound motor remains and only a hint of lifted foot band is seen. A DNA FαFwd with single central ribonucleotide (lane 4) lifts the back α'-foot correctly. RNA FαFwd (lane 5) lifts some of the α'-back foot. Lanes 6-9 attempt foot lifting with forward β fuels. None of the FβFwd fuel types show significant lifting of front (blue dashed line) or back (red dashed line) feet as intended. Lanes 10-15 are control lanes. Lane 10 shows the position of fuel lifted back foot using dummy feet and fuels γ. Lane 11 shows the position of lifted front foot. Lane 12-14 show free motor with FαFwd, FβFwd or no fuel respectively as controls for motors which have fallen off tracks. Lane 15 is a control lane for motors which have split body domains bound to track.

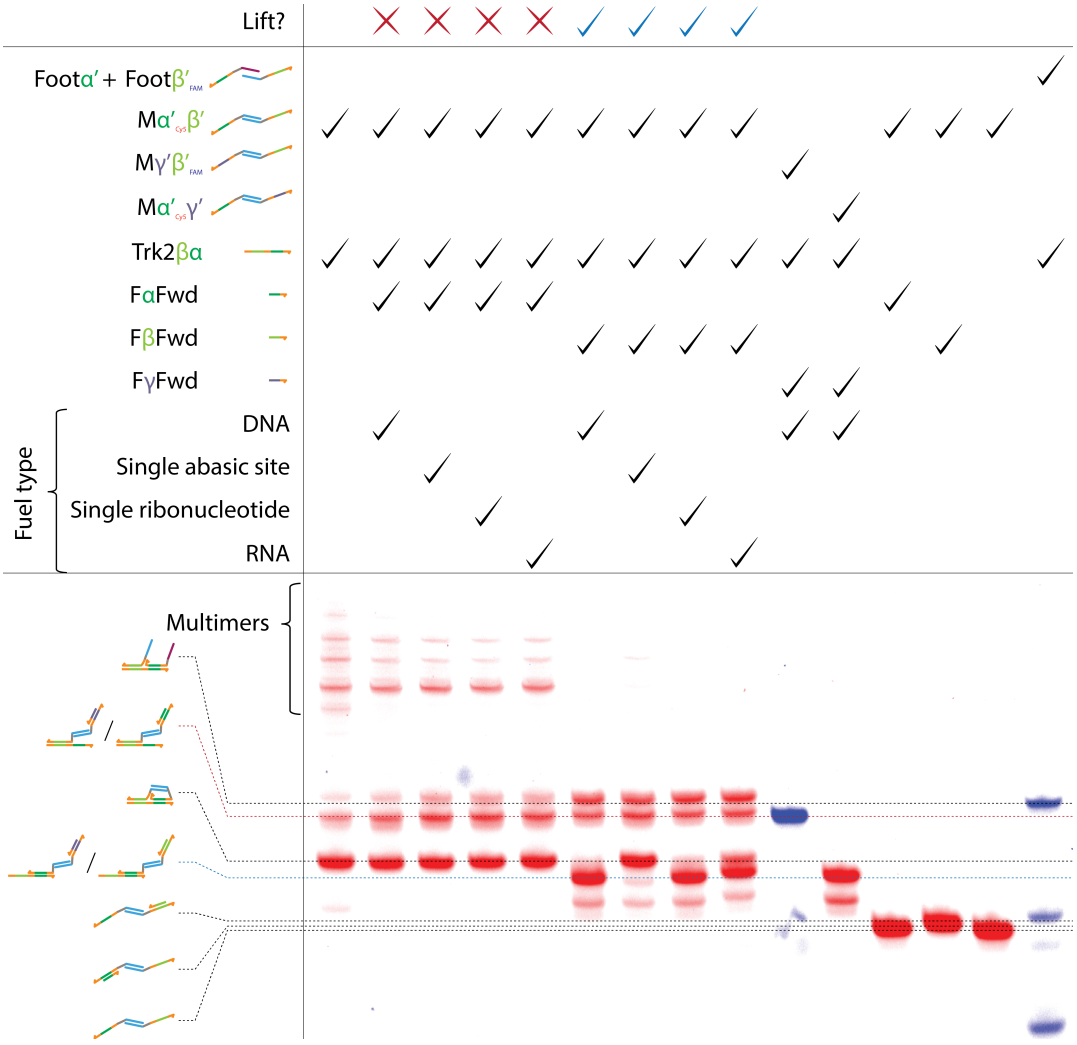

Figure 5.13: Foot lifting with forward fuels on two-site βα tracks. Identical to figure 5.12 but with Trk2βα in place of Trk2αβ. As in figure 5.12, back foot (β'-foot) is lifted best by its appropriate forward fuel (FβFwd) when the fuel is DNA, or DNA with a single ribonucleotide, followed by the appropriate RNA fuel, and only slightly lifted by DNA fuel with an abasic site.

are used to lift a trailing α'-foot, and in figure 5.13 this is in lanes 6-9 where FβFwd fuels are used to lift a trailing β'-foot. All of these correct fuels produce a band indicated by the blue dashed line which corresponds to a control sample of fuel lifted back dummy foot (lane 10 in 5.12 and lane 11 in 5.13). This band is not observed in any samples incubated with the incorrect forward fuel type, where the binding domain does not match the back foot.

When the back foot and forward fuel type match and the back foot can be lifted, the higher order bands also disappear from the lanes of the gel. There are no higher order bands in lanes 2-5 for figure 5.12 or lanes 5-9 of figure 5.13. This could be achieved via the opening of loops by foot lifting which could separate topologically interlinked motors bound to tracks, or in cases of inter-motor rather than intra-motor competition over the competition domain,

where the lifting of a motor foot is possible and could break a multimer into monomers.

Front foot is not lifted with forward fuels

By design, no forward fuel should rapidly lift a front foot, even if the binding domain matches that of the front foot. This is because forward fuels use lifted heels as toeholds, and the heel domain of the front foot is not exposed as it is not subject to competition. The rate constant for front foot lifting should be similar to that for blunt displacement $\sim 1\ M^{-1}s^{-1}$, far slower than the toehold-mediated rate constant of $\sim 10^5\ M^{-1}s^{-1}$ for a 6 nt toehold [38, 39].

Lifted-front-foot band position is indicated by red dashed lines in both figure 5.12 and 5.13 inferred from control lanes for fuel lifted dummy front foot (lane 11 in 5.12 and lane 10 in 5.13). There is some lifted foot even in the absence of any fuel (lane 1). Evidence from various gels show that lifted feet with and without 16mer fuel bound to the lifted foot migrate with similar mobilities on 10% PAGE gel. Despite the confounding factor that there is already some lifted front foot without fuel, or at least something present with a very similar electrophoteric mobility to lifted front foot, the intensity of the band at the position of lifted front foot appears roughly uniform across lanes 1-9 in both figure 5.12 and 5.13 suggesting that no fuel is significantly lifting front foot.

Fuel types show different amounts of foot lifting which can be explained by thermodynamic considerations

The composition of a fuel strand can affect its ability to lift foot from track. For both α and β fuels lifting their respective back feet, more foot was lifted by a DNA fuel, or a DNA fuel with a single ribonucleotide, than RNA or abasic-site-containing fuels (figure 5.12 lanes 2 and 4, figure 5.13 lanes 6 and 8). These fuels showed the greatest reduction in track-bound motor band, and the greatest intensity of fuel-lifted back foot band. The RNA fuel provided an intermediate level of foot lifting (figure 5.12 lane 4, figure 5.13 lane 9), some track-bound foot remained, but some fuel lifted back foot band also appeared. The DNA fuel with an abasic site provided only low levels of foot lifting (figure 5.12 lane 3, figure 5.13 lane 7). Most of the track-bound motor band remained after fuel incubation and only a small quantity of lifted back foot band appeared. These results can be understood

by considering the relative free energies of foot-fuel duplexes and foot-track duplexes. A foot-track duplex has a single mismatch, and 15-21 base pairs depending on the competition state. For simplicity I consider the liftable foot-track duplex as 15 bp with one mismatch, as the total number of base pairs between both feet and track will only reduce by 15 when the foot is lifted as then the remaining track-bound foot can permanently occupy the former competition domain. Nearest neighbour free energies for duplexes are given in table 5.1.

	$\Delta G_{\alpha' \text{foot:Trk}}$		$\Delta G_{\alpha' \text{foot:fuel}}$		$\Delta\Delta G$	
Fuel type	†	‡	†	‡	†	‡
DNA[a]	-14.11	-15.8	-18.17	-20.06	-4.06	-4.26
Single ribonucleotide[b]	-14.11	-15.8	-18.02	-20.14	-3.91	-4.34
RNA[c]	-14.11	-15.8	-16.93	-19.26	-2.82	-3.46
Abasic site[d]	-14.11	-15.8	-16.43	-18.32	-2.32	-2.52

Table 5.1: Nearest neighbour free energies in kcal mol^{-1} of foot-track and foot-fuel duplexes [28, 36, 191] at 37°C in †1 M NaCl or ‡RNAse H buffer + Mg^{2+} using salt correction of references 36, 192.

[a]The DNA fuel can form 16 bp with the foot with zero mismatches.

[b]The DNA single ribonucleotide fuel can form 16 bp with the foot. I have used an average of the RNA/DNA duplex and DNA/DNA duplex nearest neighbour free energies around the ribonucleotide and DNA/DNA duplex nearest neighbour free energies elsewhere.

[c]The RNA fuel can form 16 bp with the foot. For a RNA/DNA hybrid, the duplex free energy had a higher initiation penalty despite individual base pairs on average having lower free energy than their DNA/DNA equivalent. For short duplexes such as this foot/fuel duplex the result will be that the duplex has higher free energy than the DNA/DNA equivalent duplex.

[d]The DNA fuel with the single abasic site can form 15 (non-continuous) bp but there are no mismatching base pairs. I have assumed that the abasic 'base-pair'[e] makes no negative or positive contribution to the duplex free energy and that the free energy. The abasic site may be more destabilising due to the disruption of base stacking when there is no base present on one strand, increasing $\Delta\Delta G$.

[e]Use of the term base pair here is technically incorrect because there are not a pair of bases.

At equilibrium the greatest foot lifted fraction will be from the foot lifting process with the most negative free energy change of reaction $\Delta\Delta G$. The contribution to $\Delta\Delta G$ which varies between the different fuel types (ignoring fuel secondary structure) suggests that DNA or single ribonucleotide fuel will have the most favourable free energy change, followed by RNA fuel, with the least favourable free energy change from lifting with abasic fuel. This reflects the results of gels in figures 5.12 and 5.13.

5.6.2 Orientation dependent fuel turnover

Once fuel strands have lifted feet from track they must be able to be cleaved by an enzyme to remove them from the foot and allow the foot to rebind track. On a two-site track, there

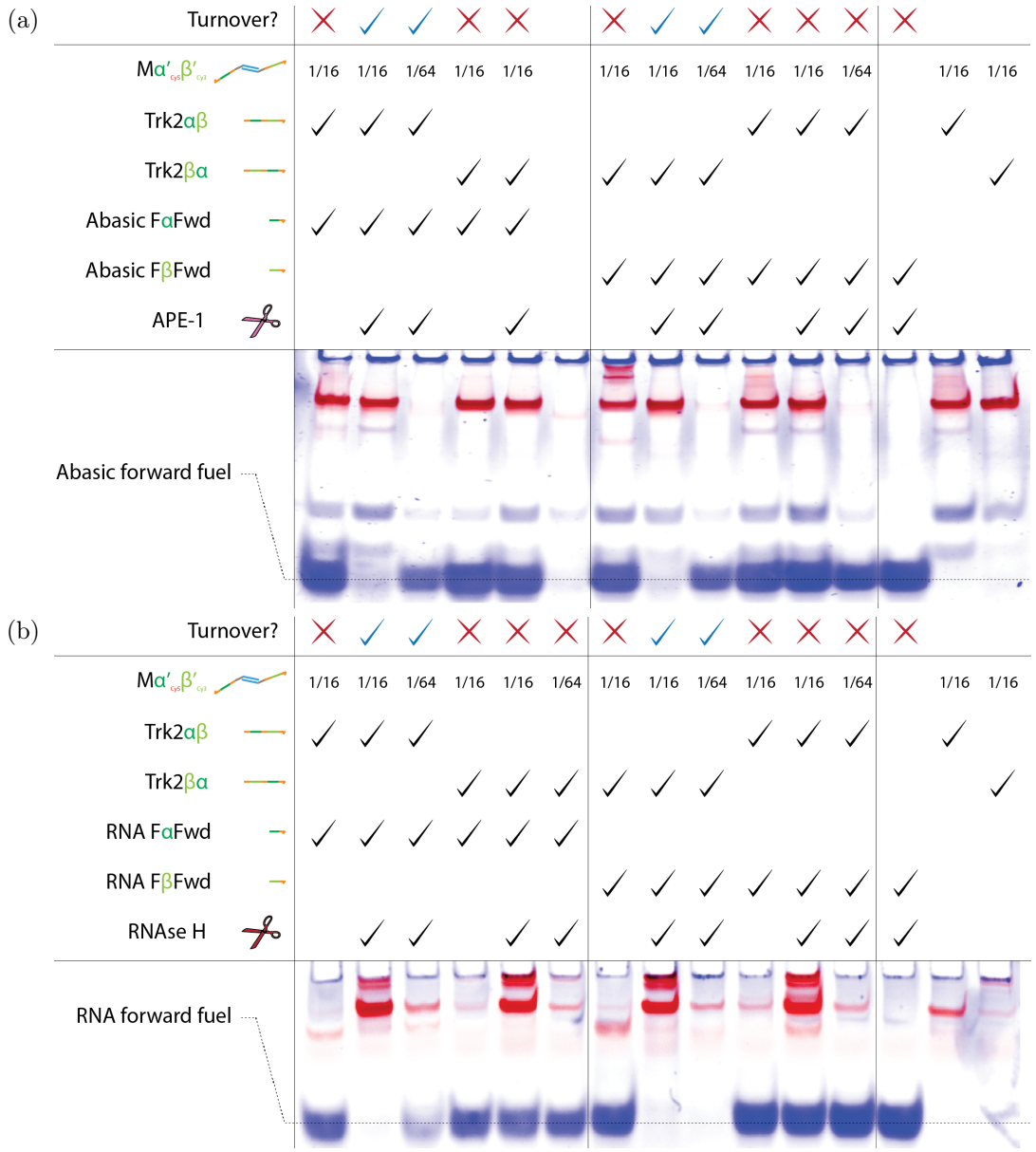

Figure 5.14: Orientation dependent forward fuel turnover. The value in the motor row gives the ratio of motor to fuel. Gel shows Cy5 fluorescence on α'-foot (red) and Sybr stain fluorescence (blue).

is only one foot-binding site available for lifted foot to bind, the same site it came from, so all steps taken on a two-site track must be idle steps. The outcome of this is that the track-bound motor orientation is fixed on a two-site track system. If this orientation allows fuel-mediated foot lifting and enzymatic fuel cleavage, it will do so repeatedly, and cleave multiple copies of the fuel strand per motor.

>16 forward fuels can be turned over by correctly orientated motors

Orientation dependent fuel turnover was assessed by the reduction of fuel band in a PAGE gel. Motors where the back foot and forward fuel type match should be able to turnover fuel in the presence of appropriate enzyme, and motors where the back foot type and forward fuel type do not match should not digest fuel.

Results are shown in figure 5.14a for forward fuels with an abasic site using enzyme APE-1 and reaction buffer NEB4, and in figure 5.14b for forward RNA fuels using enzyme RNAse H in reaction buffer RNAse H reaction buffer + Mg^{2+}. Each 10 μl reaction contained 1 μM fuel strand, either 1/16 μM or 1/64 μM motor with 1.5× excess track, and 0.5 $U\mu l^{-1}$ RNAse H or 0.5 $U\mu l^{-1}$ APE-1. Samples were left for 16 hrs at either 37°C for figure 5.14a or room temperature for figure 5.14b. If a fuel band disappears in the gel, either 16× or 64× fuels have been turned over by a track-bound motor depending on the initial motor:fuel ratio.

For both gels 5.14a and 5.14b, where the motor orientation was correct (lanes with blue ticks) the fuel band was fully digested with a 16× excess of fuel to motor. In 5.14a where an abasic fuel/APE-1 combination was used the fuel band was partially digested when there is 64× excess of fuel to motor, however in 5.14b where an RNA fuel/RNAse H combination was used, the 64× fuel excess was fully digested for FβFwd, but partially digested for FαFwd. For both gels where there was either no enzyme, or the orientation of motor did not match the fuel present (lanes with red crosses), little or no fuel was digested. These data suggest that in 16 hrs a correctly orientated motor can turn over 16-64× abasic forward fuel strands (of either α or β type) at 37°C, 16-64× RNA FαFwd and >64× RNA FβFwd at room temperature[6].

5.7 Foot Replacement on Three-Site Tracks

As noted in the previous section where I explored foot lifting and fuel turnover on two-site tracks, a track with only two foot-binding sites will only allow idle stepping as there is no 'third' track binding site to step forwards (or backwards) on to, and a lifted foot can therefore

[6]Incubation with half of the concentration of RNAse H (0.25 $U\mu l^{-1}$) did not change the amount of fuel turned over (additional data not shown)

only rebind its original site. Tracks with only three-sites should not allow overstepping for a motor with non-identical feet, as one foot type only has a single, central, binding site, the other foot type has two foot-binding sites either side, but whichever site the second foot binds, the two feet will occupy adjacent sites. Three-site tracks allow a single complete step of one of the foot types to invert the orientation of the motor (half a full motor cycle).

Moving to tracks with a greater number of binding sites presents even greater challenges in band identification. The more repeated domains there are, and the more component strands there are, the more possible binding configurations there can be. The system can therefore produce a large collection of bands. It would be very difficult to identify all of the bands in the system, and would require a large number of controls. One method to reduce the number of bands produced would be to strictly control component stoichiometry for example by extracting the band corresponding to a desired complex from a polyacrylamide gel. This process is laborious and time consuming, and also oxidising conditions used in the preparation of the polyacrylamide matrix can damage fluorophores [193]. Due to these undesirable factors, the gels shown in this section do not use purified components.

5.7.1 Preferred orientation of motors

On a three-site track I can define a 'start' and an 'end' of the track, to match the definitions of forwards and reverse fuels. The 'start' of the track should be the edge site of a track from which a forward fuel can move the motor forwards, this is the 5' end of the track. Conversely the 'end' of the track, is the 3' end of the track. On a three-site track[7] the orientation of the motor at the start and at the end will be the opposite way round (i.e. on a three-site track Trk3αβα, the start position will have the motor orientated α'β', and the end position will have the motor orientated β'α' along the track.

If all orientations and track positions have identical free energies, incubation of motor with track would produce motors at the start and the end of the track, at a 1:1 ratio. Three-site tracks with 'poly-T tails' (10T nucleotides at both ends of track) produced two

[7]Or any track with an odd number of foot-binding sites.

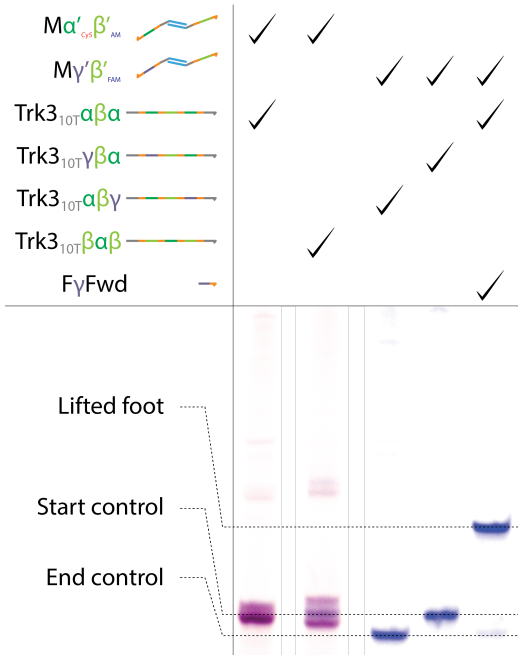

Figure 5.15: Motor can be in two track positions at equilibrium. Motors incubated were with 1.5× excess three-site tracks. Gel shows a Cy5 (red)/FAM (blue) overlay for motor labelled with Cy5 α'-foot and FAM β'-foot. Gel shows selected lanes for comparison, vertical lines indicate where the gel has been truncated for clarity. Two bands are observed when motor is incubated with three-site tracks either Trk3$_{10T}$αβα (lane 1) or Trk3$_{10T}$βαβ (lane 2). A third band is observed in lane 2, between two purple bands, and appears bluer. The difference in colour indicates different foot stoichiometry.

distinct bands in a gel[8] (see figure 5.15), for both αβα (lane 1) and βαβ(lane 2) tracks[9]. These bands may correspond to start and end motor positions. Control samples where the motor was forced to bind at either the start (lane 3) or the end (lane 4) of the track were used to aid band identification. Subtle differences in domains required to produce the control samples mean that control sample mobility may differ from the mobility of the configuration of working motor it is a control for, especially if the dummy domains γ in the control, are left single-stranded in control complexes. Using the relative mobilities of start and end control structures as a guide, I infer that the highest mobility band is likely to be the end position, and the lower mobility band the start position. In both cases the highest mobility band is higher intensity than the lower mobility band. This would suggest that the track positions do not have equal binding energies, and preferentially bind at the track end.

[8]Only after adding poly-T tails to three-site tracks are two separate bands for the motor positions observed.

[9]There is also a third band in-between the two in lane 2 but this has incorrect foot stoichiometry.

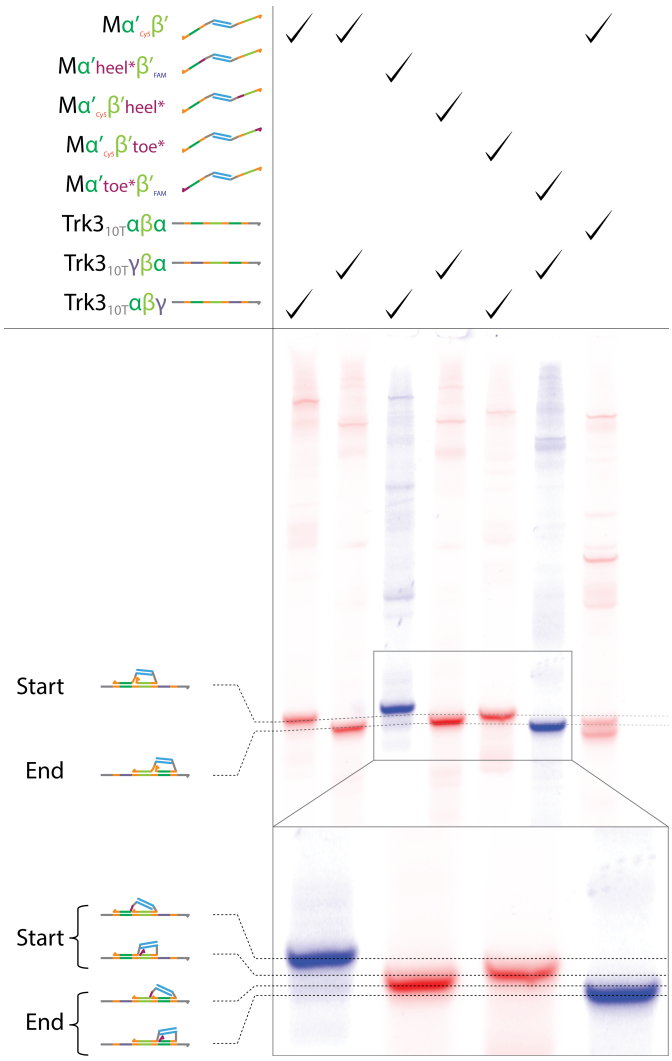

Figure 5.16: Preferred orientations and competition state of motors on extended tracks. Control complexes with dummy toes (lanes 5 and 6) or heels (lanes 3 and 4) in the central competition domain bound to the start αβ sites (lanes 1 and 3) or end βα (lanes 2 and 4) of three-site tracks.

Central competition state alters gel mobility

It is not just the position of the motor on the track which can affect electrophoteric mobility. As previously identified in gels of motors on two-site test tracks such as figure 5.11, motors in the same track position can have differing mobilities. One hypothesis is that this difference in mobility is due to the central competition state, or which foot is winning the battle to hybridize to the central domain. This may explain the difference in band separation between the two-site track motor bound bands in figure 5.15 lane 1 compared with lane 2. If differences in competition state are visible on a gel then electrophoresis could be used to decipher competition state and determine its effect on motor operation.

To test this hypothesis, a gel was run with control structures on three-site tracks with

heel or foot domains of one of the feet replaced with poly-T. This should prevent the mutated heel or toe from binding to the competition domain, and allow the other foot to 'win' the competition and hybridize the central domain. The control samples are shown in lanes 3-6 of figure 5.16. The two samples with the highest mobility were those with motors at the track end. This is consistent with previous gels. Of the two samples with motors at the end, the sample where front-foot-toe was mutated and back-foot-heel track-bound was higher in mobility than the sample where the back-foot-heel was mutated and front-foot-toe track-bound. The same was true for relative mobilities of the competition states of the two motors at the start position, despite the fact that the motor orientation was the opposite in these cases, suggesting that the difference in mobility is mostly due to central competition state rather than motor orientation.

What does this mean in terms of separation of observed 'track-bound' bands for other samples? Assuming that in each case the top band is motor at start and the bottom band is motor at end; in the absence of fuel, the difference in separation of these two bands can be explained by a particular foot type being better at winning the inter-foot competition. Recall the results from figure 5.15 lane 1 and 2 where the pair of bands were closer for motor bound to $Trk3_{10T}\alpha\beta\alpha$ and further apart for motor bound to $Trk3_{10T}\beta\alpha\beta$. This might suggest that for $Trk3_{10T}\alpha\beta\alpha$, when motor is at the end, β'-foot heel is lifted and α'-foot wins the competition (slower end band), and when the motor is at the start, β'-foot toe is lifted and α'-foot wins the competition (faster start band). Overall this would reduce the separation between start and end bands. Conversely for $Trk3_{10T}\beta\alpha\beta$, when motor is at the end, β'-foot toe is lifted and α'-foot wins the competition (faster end band), and when motor is at the start β'-foot heel is lifted and α'-foot wins the competition (slower start band) increasing the separation between start and end bands.

If one foot type always won the inter-foot competition, only the losing foot type would ever have an exposed toehold for foot-lifting with fuel, and the motor could only make a single step. Our case cannot be this extreme because two-site track gels show lifting of both foot types when motor orientation matches the forward fuel. The fuel turnover assay, however, shows greater turnover of FβFwd, consistent with α'-foot winning the competition.

5.7.2 Timecourse

Now that a test system has been created in which both start and end positions are observed at equilibrium, the next step is to attempt drive the system out of equilibrium and to move motors at the start of the track to the end of the track.

Figure 5.17 shows the pattern of bands produced after the addition of RNA forward fuel and or RNAse H to track-bound motor Mα'β' labelled with Cy5. Figure 5.17a shows the timecourse for an αβα track and α fuel. The fuel appears to lift foot in less than 10 min, as a clear lifted foot band appears after 10 min in lane 4, and there is even evidence of some lifting after 0 mins (where the fuel was added immediately before the lane was loaded and the gel was run). For samples incubated between 10-240 mins, all fuel-lifted foot appears to have been 'removed' by RNAse H enzyme (lane 5,7,9,11,13). The remaining bands appear identical to the no enzyme, no fuel control sample in lane 1. The sample incubated for 360 mins with fuel and enzyme (lane 18) shows some lifted foot band. This may indicate that the enzyme has degraded and is no longer able to turnover fuel. Strangely, in lane 20 where the sample has been incubated for 480 min, there is no lifted foot band.

Similar results are seen in figure 5.17b except for the 480 min sample with enzyme showed more lifted foot remaining than the 360 min sample rather than less. Neither gel shows a change in the ratio of motor occupation of the two positions. This might indicate an issue with foot replacement where feet are prone to rebind the site they have come from, or an issue with rapid equilibration between the track-bound configurations.

5.7.3 Equilibration of motor between track start and track end

To determine if the two configurations can interchange at room temperature, gels were run in two dimensions. The first dimension was run in a gel as normal then incubated at room temperature. Lanes of interest[10] from the gel were recast rotated by 90°, and run again. Total time spent at room temperature was ∼1 hr. As well defined bands can be observed in gels run at 5°C, it is assumed that species are not exchanging significantly at this temperature. The motor is usually operated at room temperature, however, and at

[10]Lanes 1 and 2 from figure 5.15.

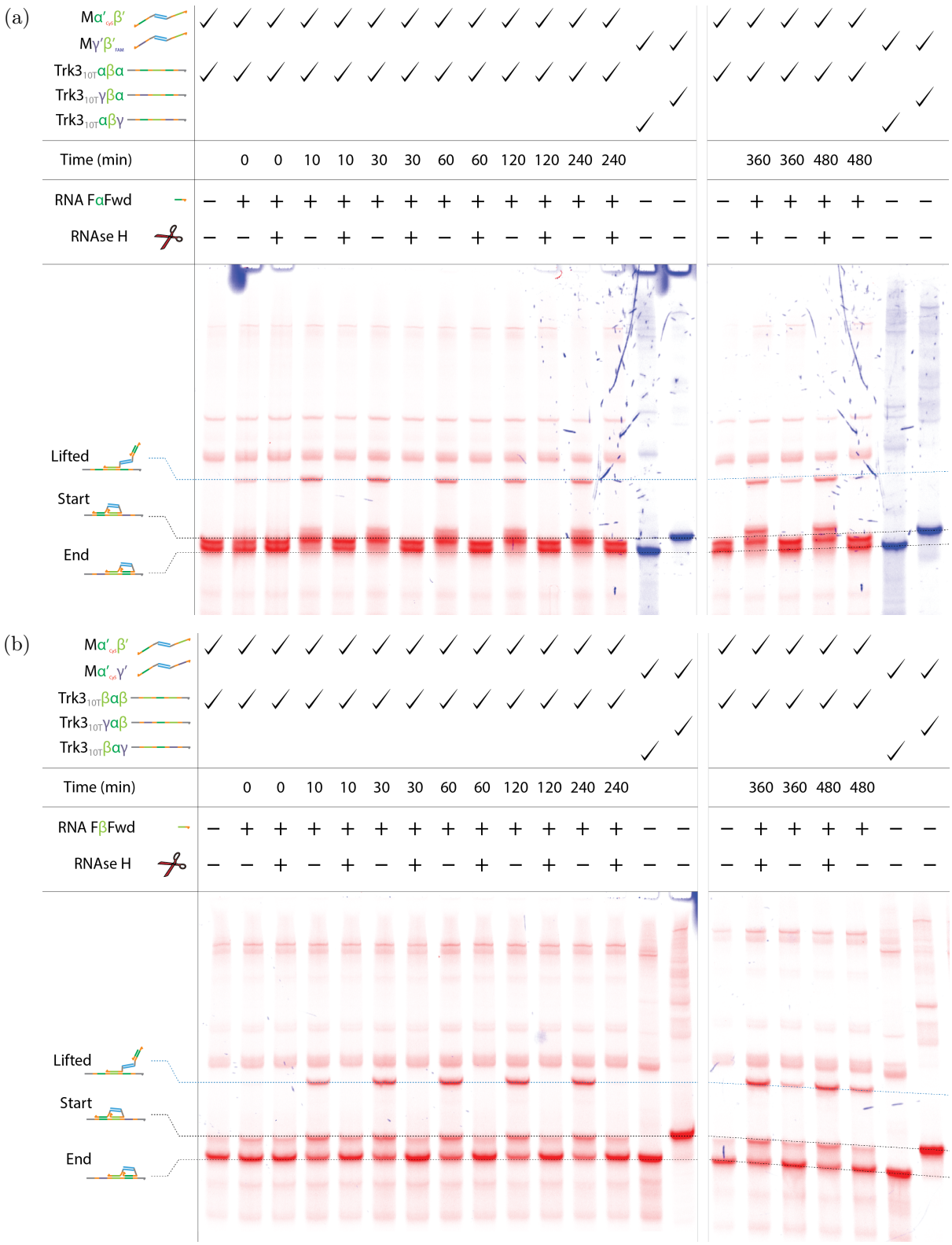

Figure 5.17: Timecourse of biped on three-site tracks. Mα'β' on three-site tracks **a** Trk3αβα, or **b** Trk3βαβ. Lanes 1 and 16 show motor on track without fuel or enzyme (−, −), subsequent pairs of lanes (2-13, 17-20) show track-bound motors with (+) added fuel with (+) and without (−) RNAse H lanes where both are added a number of minutes prior to running on gel given above the respective lane. Control lanes 14 and 21 show end bound motor, and 15 and 22 show start bound motor. Lanes 1-15 and 16-22 are from two different gels. The gap between gels is indicated with vertical grey lines.

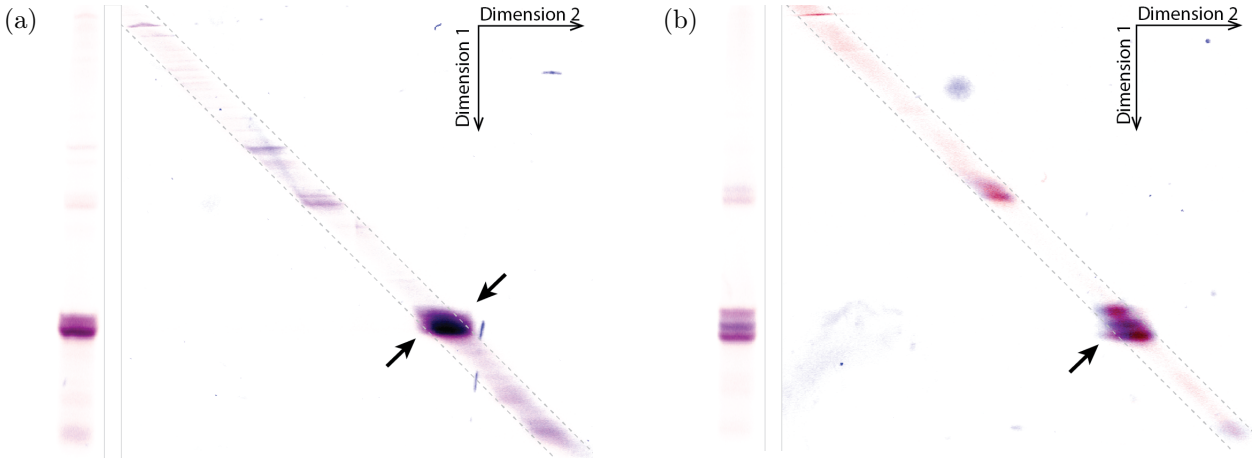

Figure 5.18: 2D PAGE electrophoresis Cy3 (red)/ FAM (blue) overlay of **a** figure 5.15 lane 1, Motor $\alpha'_{Cy5}\beta'_{FAM}$ on $Trk3_{10T}\alpha\beta\alpha$, and **b** figure 5.15 lane 2, Motor $\alpha'_{Cy5}\beta'_{FAM}$ on $Trk3_{10T}\beta\alpha\beta$. Gel was run along dimension one (shown left), scanned, lane extracted, repolymerised, then run along dimension 2 after 1 hour at room temperature. Off diagonal bands (diagonal is marked with grey dashed lines) indicate conversion of one species to another. There are off diagonal bands in both **a** and **b** (indicated with arrows) suggesting exchange between start and end track positions.

this elevated temperature exchange may be possible. Any species which has not changed its shape/bonding will run with the same mobility in the second dimension creating a diagonal line of bands. If exchange is occurring, this will appear as off-diagonal 'bands'.

Results are shown in figure 5.18. Results are inconclusive because band separation was smaller than band width, which makes any off diagonal shift difficult to see, but it does appear that there may be some off diagonal bands around the track-bound motor bands in both figure 5.18a and figure 5.18b, supporting the hypothesis that start and end configurations are exchanging at room temperature on the time scale of 1 hour or less. It is possible that the exchange is sufficient to re-equilibrate any motors which have stepped forward to an 'end' position on a track. Equilibration was previously observed for the motor with two identical feet [1, 2, 152] on two-site tracks in <40 mins.

5.8 Observing Conformations with FRET

Fluorescent label positions for motors used in this chapter are presented in figure 5.19. The Cy3 labelled β'-foot was designed to be able to report information about the motor orientation. This foot has both a Cy3 fluorophore on its 3' toe, and an Iowa Black FQ quencher at its 5' end. The separation between this fluorophore quencher pair in α' β' orientation is smaller (more quenching - lower observed Cy3 signal) than the separation

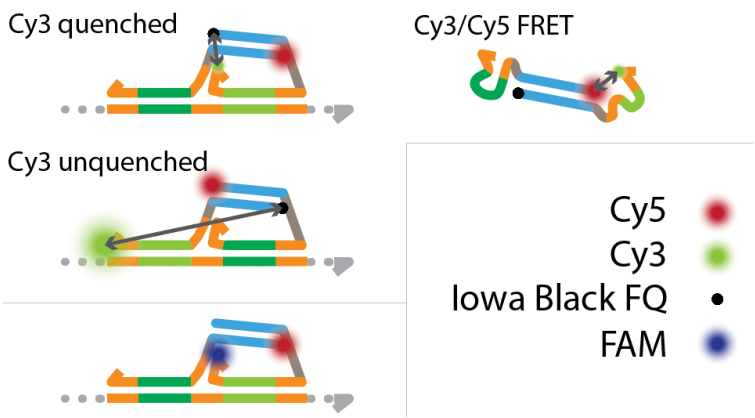

Figure 5.19: Schematic of the positions of fluorophores and quenchers on the bipedal motor. Cy5 α'-foot has the fluorophore on the motor body at its 5' end. Cy3 β'-foot has the fluorophore on its toe on the 3' end and a Iowa Black FQ dark quencher at its 5' end. The combination of this fluorophore and quencher allows the foot to report its conformation in the motor. Top schematic shows the Cy3 fluorophore quenched by Iowa Black FQ because it is a short distance away in the α'β' motor orientation. Middle schematic shows the motor unquenched because Cy3 and Iowa Black FQ are well separated in the β'α' motor orientation. FAM labelled β'-foot has fluorophore on its toe at the 3' end. FRET is also possible between Cy5 and Cy3 fluorophores, or Cy5 and FAM fluorophores but the distance between the two fluorophores does not change dramatically between its various configurations except when β'-foot is single-stranded (top right).

between the fluorophore quencher pair in β' α' orientation (less quenching - higher observed Cy3 signal). A motor where a foot is lifted should also appear unquenched because Cy3 and Iowa Black FQ are not held in close proximity.

5.8.1 Band identity may be better established by mobility and Cy3 quenching than by mobility alone

Bands in three-site track gels are often difficult to interpret due to imperfect controls, and subtle differences such as in the labels used on motor feet are enough to cause changes in band mobility. Using the Cy3/quencher β'-foot, I can attempt to identify complexes in gel by both band position and Cy3 quenching. Figure 5.20 shows an example of this type of band interpretation. The gel image is formed by an overlay of a sum image of Cy5+Cy3 scans in green and a difference image of Cy5-Cy3 scans in red. Two motor types are used, usually the motor type alternates from one lane to the next but the samples in pairs of lanes (1&2, 4&5, 6&7, 8&9, 10&11) should be otherwise identical. Differences in mobilities between these pairs of lanes demonstrate the potential for fluorophore labels to alter mobility. Lanes 12-15 are control samples for motors with Cy3/Iowa Black FQ β'-feet at the start (12, 14) or the end (13, 15) of the track in either α'β' orientation (12, 15) or β'α' orientation (13, 14).

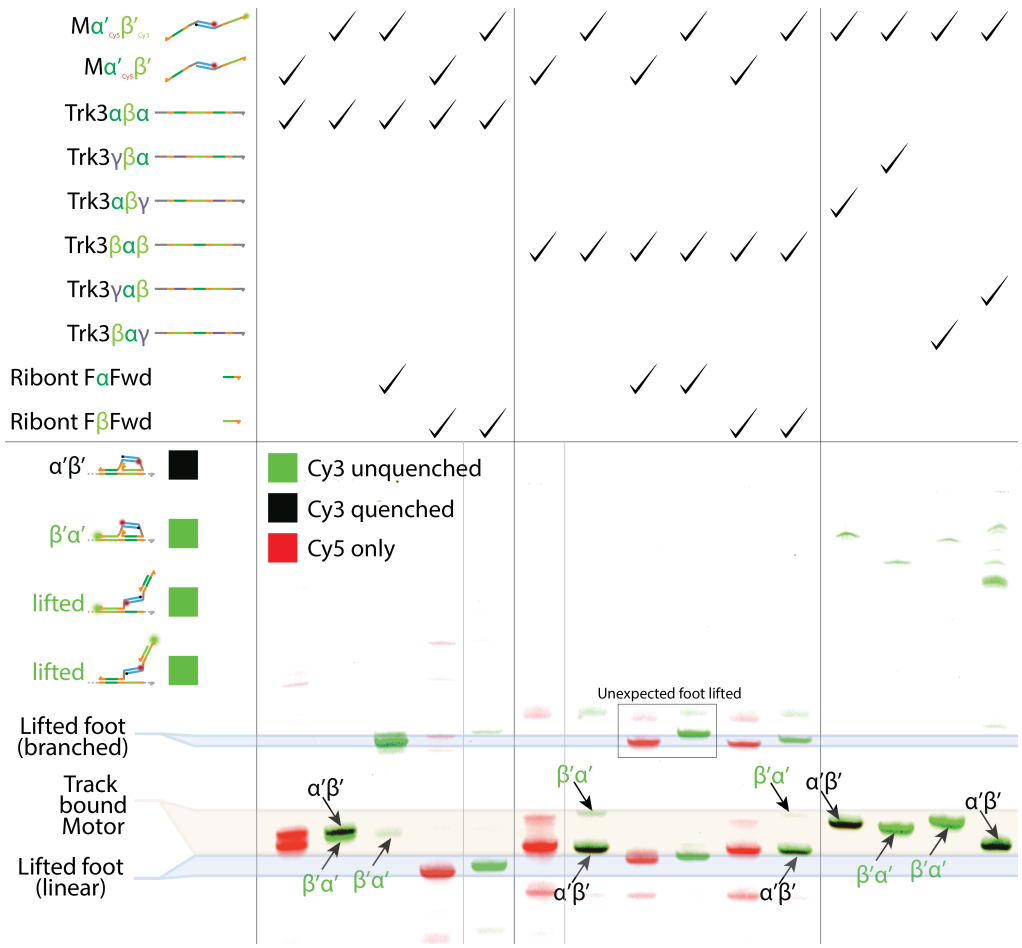

Figure 5.20: Two versions of Mα'β' were incubated with $Trk3_{10T}$αβα or $Trk3_{10T}$βαβ and challenged with either α or β forward fuel (DNA fuel with single central ribonucleotide). Motors all have Cy5 α'-foot, and either unlabelled β'-foot, or 3' Cy3 and 5' quencher labelled β'-foot. Overlay shows sum of Cy3 + Cy5 in green and difference Cy5 - Cy3 in red to emphasise differences in Cy3 quenching. Key enables interpretation of this overlay. Possible orientations and configurations of motor are depicted on the left along with the colour of band expected in a gel. Blue regions of the gel contain lifted complexes. The upper blue region corresponds to more branched structures e.g. those with the central foot remaining bound, and track either side of it, and the lower blue region corresponds to more linear structures where the foot still stuck to the track is at one end of the track. Yellow region of the gel contains track-bound motor complexes where both feet are bound to the track.

Band colour in the overlay confirms the predicted effect of orientation on Cy3 fluorescence: configuration α'β' is more quenched than β'α'. This can be used to interpret track-bound complexes (found in the yellow region of the gel) as either α'β' or β'α' (see annotations).

We also learn about foot lifting from this gel. Lifted-foot bands (identified from control lanes in similar gels, and by observation of which type of band appears when different fuels are added) can appear in regions coloured in blue, either below or above the position of the track-bound complexes depending on the position of the foot which remains bound to the track. A central track-bound foot produces the low-mobility band as it is a highly-branched

structure. A track-bound foot at the track end produces a more linear structure with higher mobility than the track-bound complexes. A track-bound foot at the start produces another low-mobility branched structure. All lifted feet have unquenched Cy3 inferred from band colour in the overlay of figure 5.20. The above explanation for lifted foot band positions can explain almost all of the bands produced assuming forward fuels only lift back foot of the correct type. The only surprising bands are those produced by βαβ track and α fuel, in which case the expected foot remaining after lifting would be the β'-foot at the track end (high mobility band). This band does appear, but the β'-foot at track start band also appears, which is not expected as it would require a front foot to be lifted or equilibration between the two lifted α'-foot complexes.

5.8.2 Conclusions from three-site track gel experiments

Though three-site gels are complex to interpret, some conclusions can be made. Results from various attempts to lift feet from track do show clear evidence of lifted-foot with various fuel types, and RNAse H has been shown to remove lifted-foot bands in figure 5.17 and allow foot to rebind track. Results so far have shown that after lifting and digestion the ratio of motors at the start and end track positions is unchanged. The most strongly binding fuels are also able to lift almost all track-bound motor, it is not just the band corresponding to the correct orientation which is lifted, in contrast to the two-site track results in figure 5.14. This might indicate, not that the fuel is lifting the incorrect orientation of motor, but that the track-bound motor is rearranging back towards its equilibrium distribution of start and end positions and therefore all track-bound motor eventually depletes. Preliminary evidence supporting the hypothesis of motor rearrangement comes from 2D electrophoresis of track-bound motor incubated at room temperature for 1 hr in figure 5.18, and has been observed in earlier studies [1, 2].

Moving the motor will be a constant battle against equilibration mechanisms. Some of these mechanisms may be hindered by using a track which is stretched, and anchored at both ends such as the tightrope track. It is possible that the tightrope track would improve the motor locomotion by slowing equilibration. I therefore continue this chapter by attempting

to operate the motor on the tightrope rather than an oligo test track, and use FRET to decipher its conformation.

5.9 Real-Time Observation of Biped Operation

In this section I attempt to observe fluorescence changes during motor operation on test tracks and origami. In most experiments I use a 'triple-labelled' motor shown in figure 5.19 with Iowa Black FQ and Cy3 at the 5' and 3' ends of β'-foot, and a red fluorophore such as Cy5 or Alexa 647 at the 5' end of α'-foot. For experiments involving loading, the 11 nt domain + 1 abasic nucleotide is appended directly after the 3' toe of α'-foot.

In bulk fluorescence experiments, three sets of excitation and emission wavelength data are collected at 60 second intervals: Cy5 (Cy5 excitation / Cy5 emission), Cy3 (Cy3 excitation / Cy3 emission) and FRET (Cy3 excitation / Cy5 emission). I expect some conformational changes of the motor to be reported by Cy3 quenching by Iowa Black FQ, but it is possible for Cy3 to be quenched by FRET with Cy5. By comparing Cy3 and FRET signals I can infer when changes are due primarily to Iowa Black FQ quenching, or FRET with Cy5. If a change in the Cy3 intensity is coupled to an opposite change the FRET channel then the change is primarily FRET with Cy5 (e.g. a reduction in Cy3 donor fluorescence and increase in FRET), otherwise, if the change in Cy3 intensity is coupled to a change in Cy5 FRET in the same direction, then the change is primarily Iowa Black FQ quenching (a reduction in Cy3 fluorescence due to Iowa Black quenching would also reduce FRET signal as less energy can be transferred to the acceptor).

5.9.1 Loading and release

I attempted to load a blocked, triple-labelled, loadable motor onto oligonucleotide test tracks in a fluorometer. During loading, the motor should change from a blocked configuration where Cy3 and Iowa Black FQ are well separated, to a configuration where the motor is sat oriented α'β' on the track and Cy3 and Iowa Black FQ are held in close proximity, reducing Cy3 signal. Cy5 fluorescence change during loading should be minimal, as it should

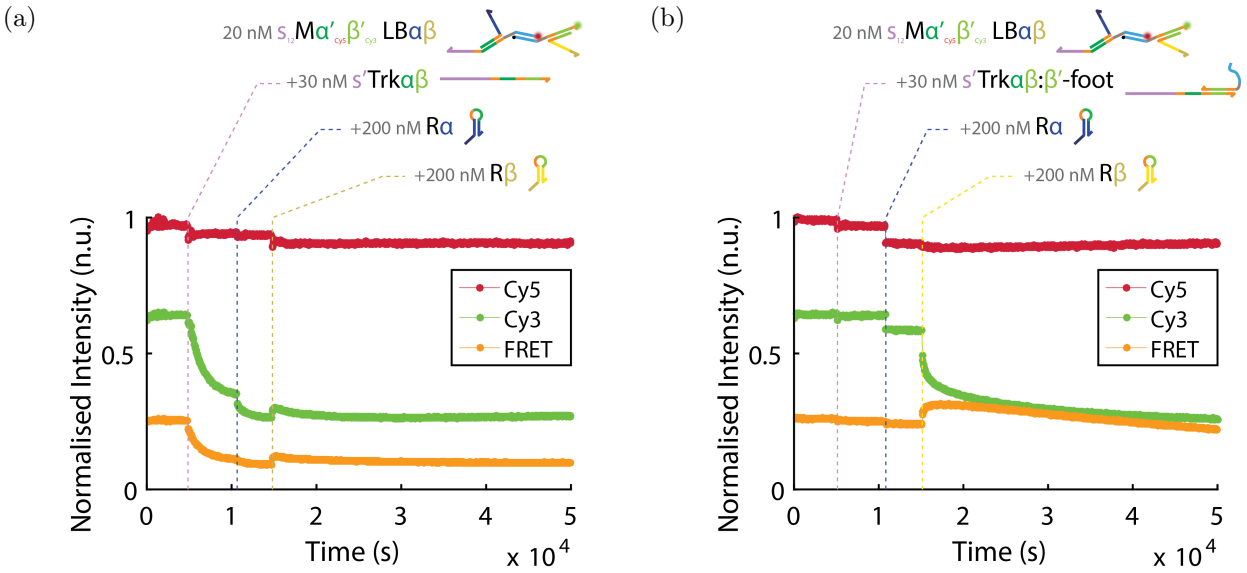

Figure 5.21: Fluorescence changes as motor is loaded onto test tracks. **a** 20 nM blocked motor with 12 nt loading domain in solution is mixed with 1.5× excess track with loading domain complement (addition time marked with pink dashed line), then 10× excess of release α (blue dashed line), then 10× excess of release β (yellow dashed line). Cy3 fluorescence drops both after track addition, and after the addition of Rα. FRET signal also decreases indicating quenching by Iowa Black FQ. Addition of Rβ causes a temporary increase in Cy3 fluorescence followed by a decrease. **b** Same protocol as **a** except the added track was pre-annealed with unlabelled β'-foot to block the β track site. Here the addition of track and Rα do not cause Cy3 quenching because β'-foot is prevented from sitting down on the track. The addition of Rβ causes a large decrease in Cy3 fluorescence and an initial rise in FRET indicating Cy5 FRET as the β'-foot to is released, becomes single-stranded, and closer to Cy5 on the motor body. FRET signal then decreases and the Cy3 signal continues to decrease, as β'-foot displaces the incumbent foot from the track and sits down on the track. Data was normalised by dividing each signal by the maximum Cy5 signal.

remain a constant distance (the length of the motor body) from the Iowa Black FQ quencher.

Fluorescence changes during loading are due to β'-foot sitting on the track

Fluorescence changes during loading were observed on test tracks. Results are shown in figure 5.21. Figure 5.21a shows a blocked motor with loading domain in solution to which first 1.5× excess track with loading domain complement, then 10× excess of each release hairpin is added. The addition of track causes a decrease in Cy3 fluorescence, and a decrease in FRET signal. The decrease in FRET signal is presumably due to reduction in Cy3 fluorescence by energy transfer to Iowa Black FQ reducing the energy transfer to Cy5, hence reducing Cy5 FRET. Addition of release Rα causes further decrease in Cy3 fluorescence and Cy5 FRET. Addition of release Rβ causes a transient rise in Cy3 fluorescence followed by a decrease. This may be due to an intermediate of displacement of LBβ by hairpin Rβ with increased

Cy3 quencher separation.

I attribute decreases in Cy3 fluorescence to the placement of β'-foot on the track site in front of α'foot which holds fluorophore and quencher in close proximity. To test this hypothesis the added track was pre-annealed with unlabelled β'-foot which blocks the available binding site for β'-foot. This sample is shown in figure 5.21b and, as predicted, the addition of track:β'-foot complex, or Rα no longer causes a decrease in Cy3 fluorescence because β'-foot of the motor is prevented from binding. Fluorescence changes in figure 5.21a can be attributed to placement of β'-foot on track, and occur before addition of Rα due to spontaneous release (figure 5.7). In figure 5.21b addition of the second foot release hairpin Rβ does cause a rapid decrease in Cy3 fluorescence coupled with an increase in FRET signal. The increase in FRET suggests that the Cy3 is primarily quenched by FRET rather than Iowa Black FQ. This can be understood because the removal of LBβ from the blocked motor by Rβ makes β'-foot single-stranded. When the foot is single-stranded, the ensemble averaged distance between the Cy3 and Cy5 fluorophores is reduced, as depicted in the top left schematic in figure 5.19. The foot cannot rapidly bind track because the track binding site is blocked by the pre-annealed foot. Eventually, however, the foot will displace the incumbent foot from the track (or the incumbent foot will displace labelled β'-foot from the motor) as this is entropically favourable. The gradual decrease in FRET signal is thought to be due to the displacement of incumbent foot from track as the motor foot sits down.

Fluorescence changes were also observed during release of a tightrope origami sample pre-incubated with load blocked motor with 12 nt loading domain for 1 hr and then purified by PEG precipitation. Release hairpins were added to these loaded tightropes in the fluorometer. Recall the blocked motor is able to spontaneously release itself from load blocks on truncated tracks as depicted in figure 5.7. This may also be possible on tightropes, therefore some of the motors may have released spontaneously prior to fluorometer observation.

Loading onto tightrope origami (figure 5.22) also shows a decrease in Cy3 fluorescence after addition of Rα similar to figure 5.21a. There is also a further decrease after Rβ addition which is not observed in figure 5.21a. Cy3 quenching could be from blocked motors sitting

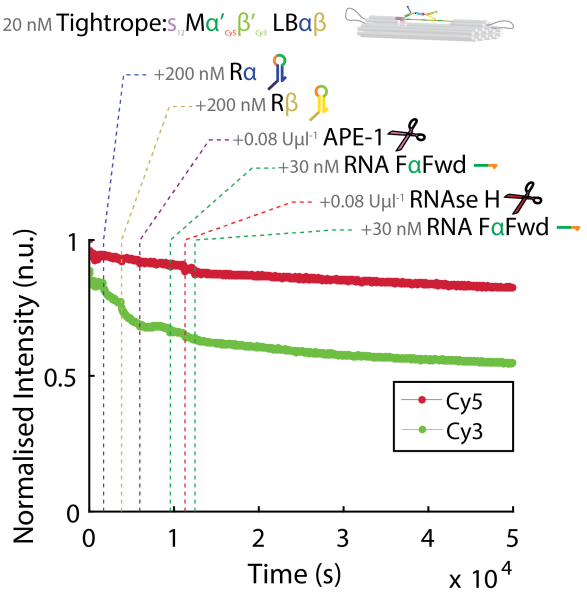

Figure 5.22: Effect of strand addition to tightropes purified after 1 hr incubation with loadable, blocked, triple-labelled motor. Release strands Rα and Rβ both cause Cy3 signal to decrease as β'-foot sits on the track. Data was normalised by dividing each signal by the maximum Cy5 signal.

on the tightrope. It may not have been observed after addition of Rβ on oligo tracks because spontaneous release was more rapid on these tracks and therefore Rβ was not required to enable the motor to sit on track but if the process is slower on origami, the release of β'-foot may help the motor sit down. Further additions of enzyme and 1.5× excess of fuel do not produce significant fluorescence changes. Addition of APE-1 was not expected to cause a fluorescence change, and the fuel excess was small, so it may not have been sufficient to produce a large enough foot-lifted population to observe. It is also possible that release hairpins interfere with the motor operation. The design held potentially promiscuous 'foot-like' domains in the hairpin loop, but some hairpins may be truncated due to incomplete synthesis, dimerised/multimerised, or otherwise open, and able to bind and block track binding sites or fuels. For this reason, other experiments on tightropes in this chapter purify away excess release strands before observation in the fluorometer.

5.9.2 Foot lifting and fuel cleavage

Two-site tracks were used to observe foot lifting with RNA forward fuel FαFwd. Tracks in the two possible orientations Trk2αβ and Trk2βα were used, along with a motor with no track as a control for fluorescence changes in unbound motor. Figure 5.23 shows the results of these tests. Figure 5.23a shows the fluorescence response of unblocked motor in solution

mixed with 1.5× Trk2αβ, to orientate the motor α'β', at time marked with pink dashed line. Cy3 fluorescence increases after the addition of track, but FRET signal decreases, as high FRET between Cy5 on the motor body and Cy3 on the toe of a single-stranded foot is reduced when feet bind track. The orientation α'β' should quench Cy3, however, quenching by Iowa Black FQ when the motor is sat on the track appears lower than the quenching by Cy5 FRET when the foot is single-stranded.

A 10× excess of RNA FαFwd is then added to the sample (dark green dashed line). This fuel addition causes an increase in Cy3 fluorescence and a small increase in FRET. These results indicate that the interaction with α fuel has reduced the quenching of Cy3 by Iowa Black FQ which is consistent with the expected change for lifting the foot. Subsequent addition of RNAse H decreases Cy3 fluorescence from the level achieved after fuel addition, but to a level higher than the track-bound level. The reduction in fluorescence may indicate that some of the motors have reset their α'-feet on the track after enzyme cleavage, and the only available α track site will restore the motor to α'β' orientation where Cy3 is quenched by Iowa Black FQ. The difference between fluorescence level after track binding, and fluorescence level after digestion with RNAse H indicates that not all α'-feet have rebound track. There is excess fuel present so the discrepancy may be due to feet rebinding fuel, or being re-lifted by fuel, after the digestion of the first fuel and so on, producing steady-state fuel digestion. The Cy3 fluorescence appears to increase on a longer timescale (see right hand side of plot) which might indicate the degradation of the enzyme over time shifting the steady state Cy3 fluorescence upwards in the excess fuel limit. A near-identically composed sample is shown in figure 5.23d. Here all components and concentrations are identical except for the RNAse H which was at 50% concentration, and from a newer batch of enzyme. In 5.23d behaviour is the same as 5.23d until the addition of RNAse H, which reaches a similar Cy3 fluorescence immediately after RNAse H addition, but continues to decrease, rather than increase, approaching the fluorescence level prior to fuel addition. An explanation for the differing behaviour could be batch-to-batch or batch-age-related variation in RNAse H activity. In 5.23a I postulated that the enzyme activity reduced before the fuel was depleted. In the sample of figure 5.23d, higher enzyme activity in the sample could allow the

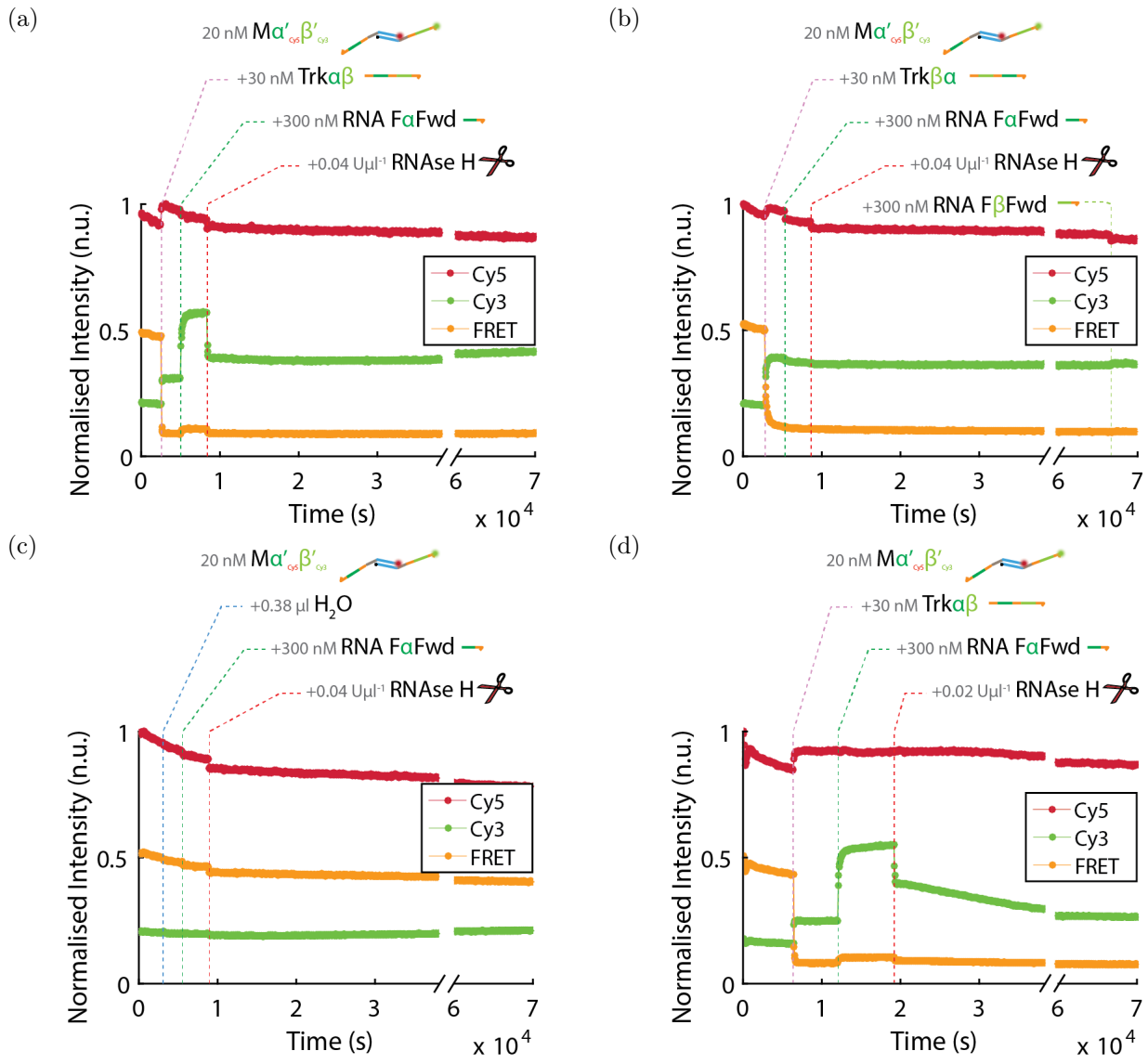

Figure 5.23: Foot lifting fluorescence changes with α forward fuel on two-site tracks. To unblocked motor in solution either **a**&**d** Trk2αβ, **b** Trk2βα, or **c** H_2O was added to achieve a 1.5× track excess. 15 × RNA FαFwd, and RNAse H were then added to all samples. All samples showed a reduction in FRET and increase in Cy3 fluorescence after track addition. Only samples with track Trk2αβ (**a**&**d**) showed a change in fluorescence after fuel addition, which caused an increase in Cy3 fluorescence in these samples as feet were lifted by fuel from a quenched, track-bound state. RNAse H addition to these samples caused a rapid decrease to a lower Cy3 fluorescence state as foot-bound fuel was digested and feet rebind track. In **a**, Cy3 fluorescence eventually rose however in sample **d**, identical to **a** except RNAse H from a newer batch was used, and concentration of RNAse H was halved, there was an eventual decrease in fluorescence approaching the track-bound motor fluorescence level. No other samples showed significant fluorescence change after RNAse H addition. Data was normalised by dividing each signal by the maximum Cy5 signal.

complete consumption of fuel before the enzyme degrades, which would ultimately restore fluorescence to the level prior to fuel addition. To test this hypothesis, deliberately 'damaged' RNAse H (achieved by short incubations at elevated temperatures) could be used in identical experiments alongside fresh RNAse H from the same batch. Differences in the behaviour after RNAse addition could then be attributed to differences in the enzyme activity.

Motor where track Trk2βα is added is shown in figure 5.23b. The binding of motor to track forming motors orientated β'α', causes a greater increase in Cy3 fluorescence than motors bound in α'β' orientation in figure 5.23a, despite a similar reduction in Cy5 FRET signal. This indicates that, as expected , the α'β' orientation has lower Cy3 fluorescence intensity than the β'α' orientation. Addition of α forward RNA fuel (dark green dashed line) does not cause any further change in Cy3 fluorescence. This could be because the motor is not able to interact with the fuel rapidly whilst in this orientation. It could also be explained if the foot is able to lift with FαFwd but the lifted foot state has similar average Cy3 fluorescence so no overall change is observed. I predict that the former case is the true explanation, because I expect a lifted foot to produce similar Cy3 fluorescence independent of the orientation it is lifted from. Comparing fluorescence level after addition of FαFwd in figures 5.23a and 5.23b, the level in figure 5.23b is lower, which suggests less foot has been lifted when the motor is in the β'α' orientation. This also agrees better with gel results in figures 5.12 and 5.13.

To ensure the fluorescence changes observed were not due to fuel interactions with non-track-bound motors, a no-track control sample was prepared in figure 5.23c. To this sample, first an equivalent volume of water (to control for dilution and mixing effects) was added, to which no significant fluorescence change was observed. Then a 10× excess of RNA FαFwd fuel was added, to which there was still no significant change in fluorescence intensity. In this case the lack of fluorescence change is not because motor is unable to hybridize fuel, gel controls show a shift in position of a motor band after incubation with FαFwd (see figure 5.12), but because the interaction between motor and α fuel does not cause a change in Cy3 fluorescence. Digestion of any foot-bound α fuel after the addition of RNAse H also leaves fluorescence levels unchanged. I conclude that that the fluorescence changes after fuel

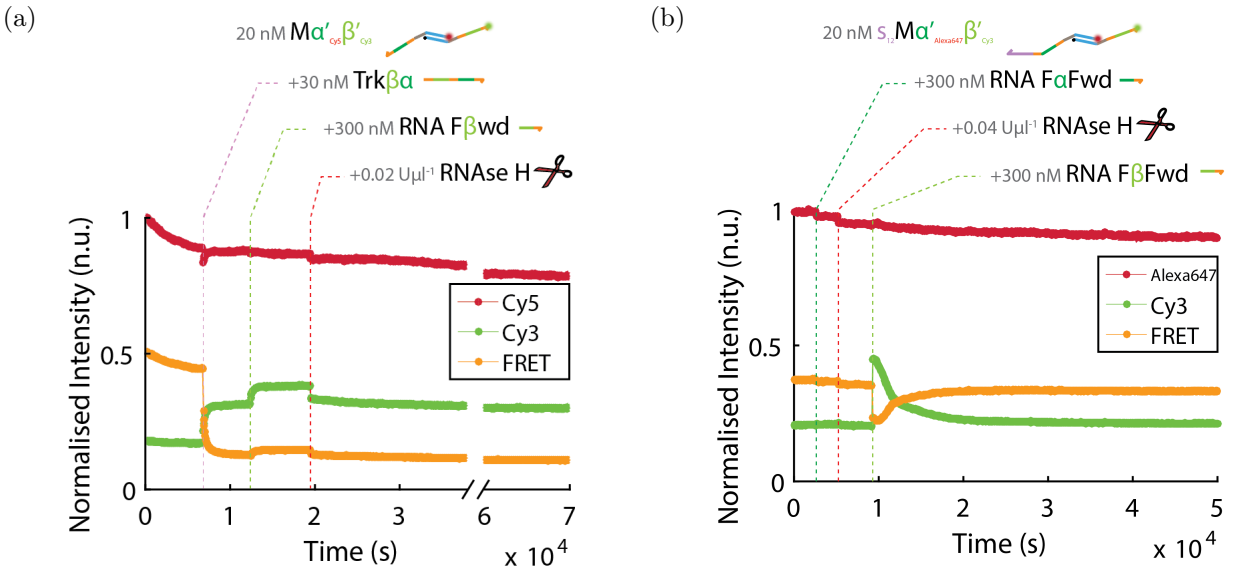

Figure 5.24: Fluorescence changes with β forward fuel. **a** Foot lifting fluorescence changes with β'α' orientated motor and β forward fuel. Unblocked motor in solution was mixed with 1.5× excess two-site oligonucleotide track Trk2αβ, then 15× excess of RNA FβFwd, then RNAse H. After addition of track, FRET signal decreases and Cy3 signal increases as motor hybridizes track. Addition of fuel causes a small further increase in Cy3 fluorescence as fuel lifts foot from track. RNAse H addition restores the fluorescence level to the level observed before fuel addition, where motor was bound to track, indicating that feet have rebound track after fuel cleavage. **b** Fluorescence changes when RNA FβFwd is added to motor in solution (pale green dashed line) in the presence of RNAse H. Addition of this fuel causes an increase in Cy3 fluorescence coupled with a decrease in Cy3/Alexa647 FRET as fuel hybridization separates Cy3 and Alexa647 fluorophores decreasing FRET. This change is only transient as it is restored to levels before fuel addition due to RNAse H in solution. Data was normalised by dividing each signal by the maximum Cy5/Alexa647 signal.

addition are due to foot lifting from an Iowa Black FQ quenched initial configuration, to an unquenched foot-lifted configuration.

The effect of β forward RNA fuel was also investigated, with results in figure 5.24. On a βα track designed to allow foot lifting with FβFwd, addition of FβFwd after motor had bound track caused an increase in Cy3 fluorescence. This increase was smaller than the increase in fluorescence observed by α' back foot lifting in figure 5.23a, which may be because the initial state was less quenched to begin with, therefore foot lifting caused a smaller decrease in quenching. The effect of FβFwd on fluorescence of an unbound motor sample was also tested, similar to tests with α fuel in figure 5.23c. Unbound motor was mixed with 15× FαFwd, then RNAse H, then 15× FβFwd. The first two additions caused no significant fluorescence changes (figure 5.23c). The addition of FβFwd however did cause a large fluorescence change. The Cy3 fluorescence increased dramatically and the Alexa647

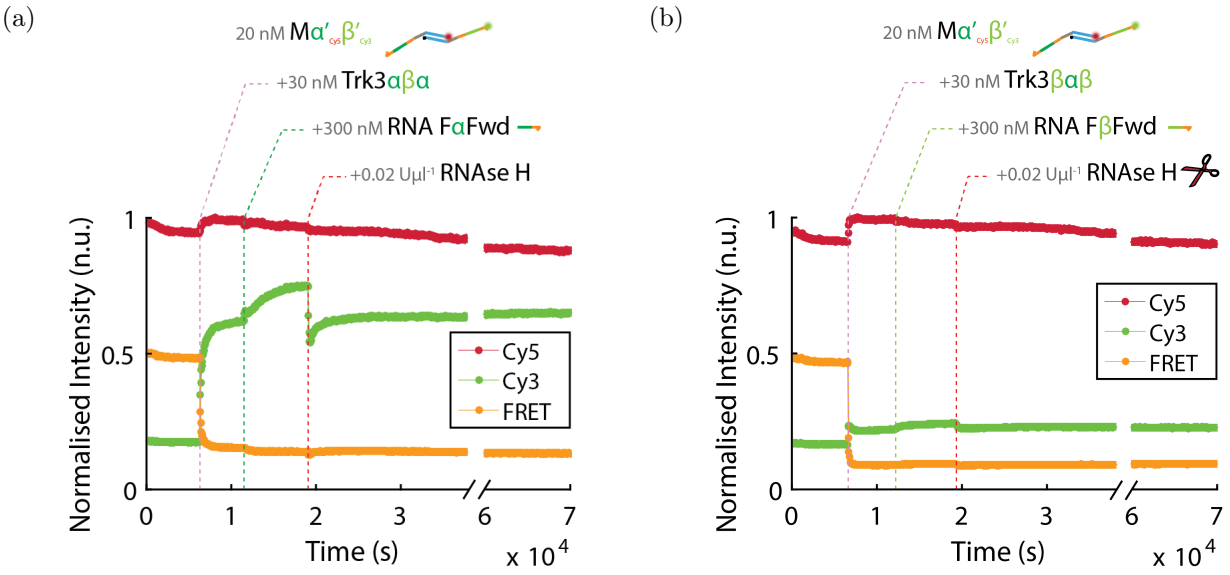

Figure 5.25: Fluorescence changes on three-site tracks with start foot fuel. Data was normalised by dividing each signal by the maximum Cy5 signal. **a** Unblocked motor sample was mixed with 1.5× excess of Trk3αβα causing a large increase in Cy3 fluorescence and large decrease in FRET. The majority of motors are likely in unquenched orientation β'α', at the track end. Addition of RNA FαFwd caused a further increase in Cy3 fluorescence with slower kinetics than was observed for lifting α'-foot on Trk2αβ with the same fuel type and concentration. Addition of RNAse H caused a large early decrease in Cy3 fluorescence which later increased restoring Cy3 fluorescence to the level prior to fuel addition. **b** Unblocked motor sample was mixed with 1.5× excess of Trk3βαβ causing a small increase in Cy3 fluorescence and large decrease in FRET. The majority of motors are likely in quenched orientation α'β' at the track end. The fluorescence peaked and then began to decrease again which may indicate initial binding of motors on track followed by rearrangement of feet. Addition of 15× excess RNA FβFwd caused a small increase in Cy3 fluorescence. Fluorescence levels were restored to levels prior to fuel addition after RNAse H addition.

FRET signal decreased, indicating that the loss of Cy3 quenching was because of a reduction in FRET with Alexa647. When unbound to track, the link between Cy3 and Alexa647 is 26 nt of single-stranded DNA (see figure 5.19 top right). When FβFwd is hybridized to β'-foot, the link is 4 nt ssDNA, then a 16 bp duplex, then 6 nt ssDNA. This link has greater RMS end-to-end distance when fuel is hybridized, therefore lower FRET.

Rearrangement on a three-site track

As the motor is unable to step forward on a two-site track, three-site tracks were tested in the fluorometer to investigate stepping/foot replacement. On three-site tracks non-idle stepping would produce the alternative fluorescence orientation.

Figure 5.25 shows unblocked motors incubated with a three-site track (Trk3αβα in figure 5.25a and Trk3βαβ in figure 5.25b) then mixed with excess RNA forward fuel designed to

lift the back foot of motors at the start position on the three-site track (FαFwd in figure 5.25a and FαFwd in figure 5.25b). Instead of monotonic changes in fluorescence after the addition of track similar to those observed on two-site tracks, the change in Cy3 fluorescence when Trk3βαβ (figure 5.25b) was added to unblocked motor increased, but then decreased forming a maximum and then lower steady-state Cy3 fluorescence. Both samples showed some increase in Cy3 fluorescence after the addition of the forward fuel. This increase in figure 5.25b was small, either because the orientation to be lifted was already unquenched or because most motors were not lifted by FβFwd. The increase in Cy3 fluorescence after FαFwd addition in figure 5.25a was larger despite the larger initial Cy3 fluorescence which points to the majority of motors orientated β'α' . The increase is formed of a small rapid rise, and a larger rise which occurs with slow kinetics, far slower than the observations of α'-foot lifting on two-site test tracks in figure 5.23a which may indicate an additional reaction step in the foot lifting process which slows the overall rate of foot lifting. One possible explanation of this is, again, rearrangement of motors on track . Foot lifting by α fuel depletes the population of motors at the track start. The system then re-equilibrates as motors rearrange on the track to restore the equilibrium track site occupation, producing more motors orientated to allow foot lifting, which then lift, and so on.

After RNAse addition, Cy3 fluorescence decreases for both samples. In figure 5.25a, the fluorescence change is non-monotonic. This could also be explained by the rearrangement hypothesis. Feet are replaced after fuel is cleaved either in front of or behind the central track-bound foot but the motors then equilibrate restoring the steady-state Cy3 fluorescence observed prior to fuel addition.

Lifting motor feet on tightrope origami

Using insights about the effect of fuel and enzyme addition on fluorescence gained from experiments on oligonucleotide test tracks, motor operation on the tightrope was investigated. Motors were loaded onto tightropes and treated with release hairpins and enzyme, and purified by PEG precipitation prior to observation in the fluorometer. Once loaded, released, cleaved, and purified; tightrope origami was added to the fluorometer at 20 nM

tightrope concentration. To the tightrope sample, first 15× RNA FαFwd was added, then RNAse H, then 15× RNA FβFwd. Results are shown in figure 5.26. Addition of 15× RNA FαFwd caused both Cy3 and FRET signals to rise. I attribute this to foot lifting, the change is primarily in quenching of Cy3 by Iowa black FQ[11] . This is comparable to two-site track samples where α'-foot lifting caused fluorescence to increase when motors were oriented α'β'. Similarly addition of RNAse H caused fluorescence of Cy3 to decrease, rapidly at first then slowly approaching its steady state prior to the addition of fuel. The addition of RNA FβFwd caused a small rise in Cy3 fluorescence, which then decreased due to the presence of RNAse H in solution. None of these fluorescence changes are observed for free motors (figure 5.24b) so I conclude they are due to conformational changes of tightrope bound-motors.

The restoration to the same steady-state at late times may suggest that the ensemble orientations of motors are the same as observed prior to fuel addition. This could be because motors have not moved anywhere, and remained at their original position. It could be because motors have on average taken an even number of steps, and therefore maintained their orientation. It could be because both before fuel addition, and at late times after fuel addition, the motors, able to diffuse along track positions, have equilibrated.

Samples where the motor was loaded, released and not treated with APE-1 (therefore should have been stuck at the start) in figure 5.26c and samples where the motor was not blocked and loaded but instead unblocked motor was incubated with tightropes (therefore non-specifically positioned on track) in figure 5.26d showed similar changes in fluorescence when RNA FαFwd, RNAse H, and RNA FβFwd were added.

As the track has six sites, start and end positions have the same motor orientation, so I cannot tell from these data that the motor has moved along the origami, nor conclude with confidence that it has not moved. To confirm occupation at the track end I require a second reporting method, which will cause a fluorescence change when motors reach the track end.

[11]Also verified by a control sample where no Cy5 fluorophore was present, where similar changes in Cy3 intensity were observed (figure 5.26b).

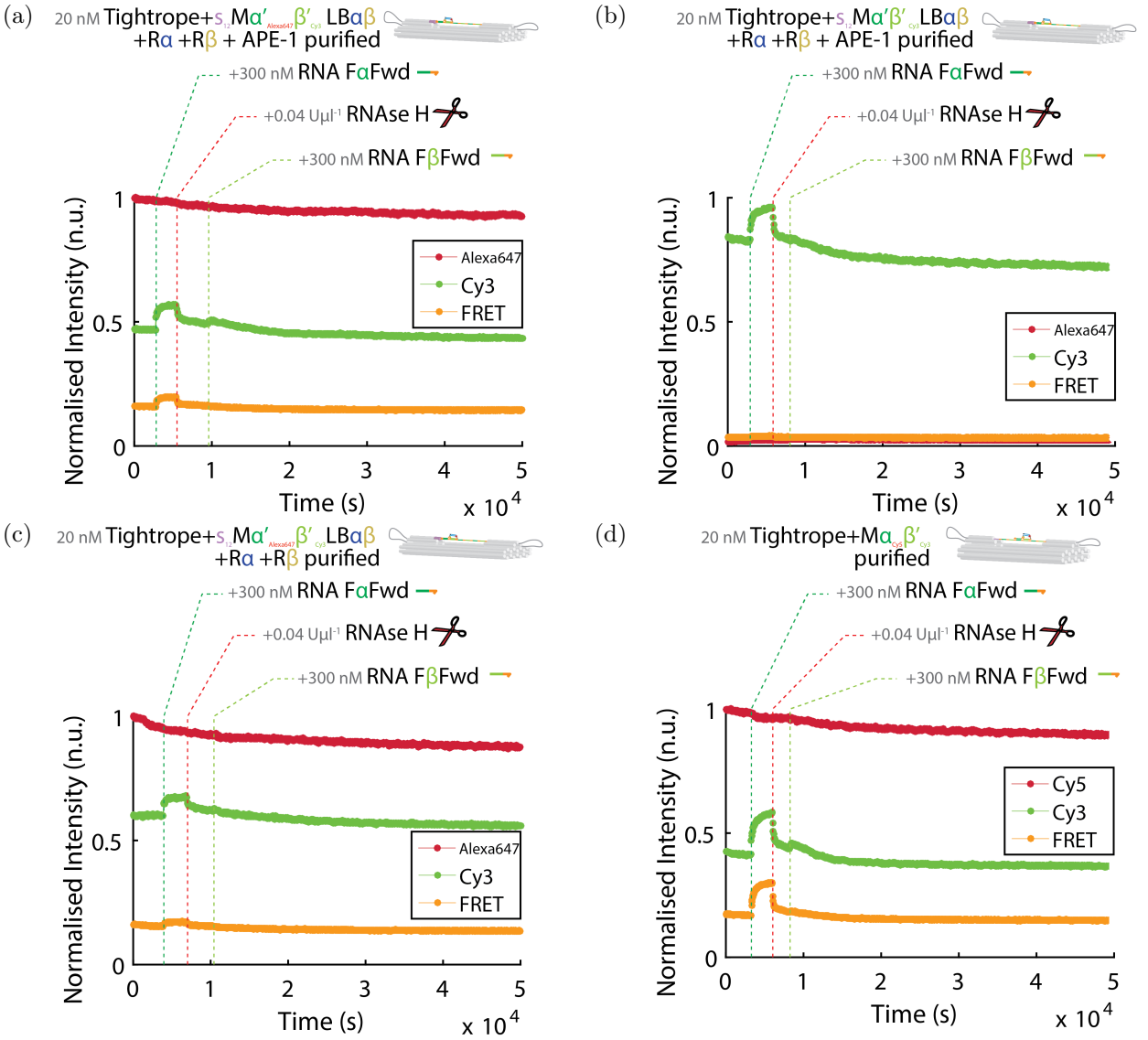

Figure 5.26: Foot lifting fluorescence changes on tightropes. Data was normalised by dividing all fluorescence intensities by the maximum recorded Cy5/Alexa647 fluorescence, or the maximum Alexa647 fluorescence of an identical concurrent experiment including an Alexa647 label for **b**. All samples use motors on tightropes folded with 102 nt ssDNA across the gap at 20 nM. Tightropes were treated with release hairpins and APE-1 before PEG purification and fluorometer observation unless specified otherwise. Additions whilst in the fluorometer are marked with dashed lines representing 300 nM RNA FαFwd (dark green), RNAse H (red), 300 nM RNA FβFwd (pale green), or H_2O. **a** Tightropes loaded with blocked motor with Alexa647, Cy3 and Iowa Black FQ fluorophores. **b** Tightropes loaded with blocked motor with no red flurophore, only Cy3 and Iowa Black FQ fluorophores. **c** Tightropes loaded with blocked motor with Cy5, Cy3 and Iowa Black FQ fluorophores which has not been treated with APE-1 to cleave away loading domain. **d** Tightropes incubated with 1× unblocked motor with Cy5, Cy3 and Iowa Black FQ fluorophores, not released or APE-1 treated. Motor is non-specifically placed on the track.

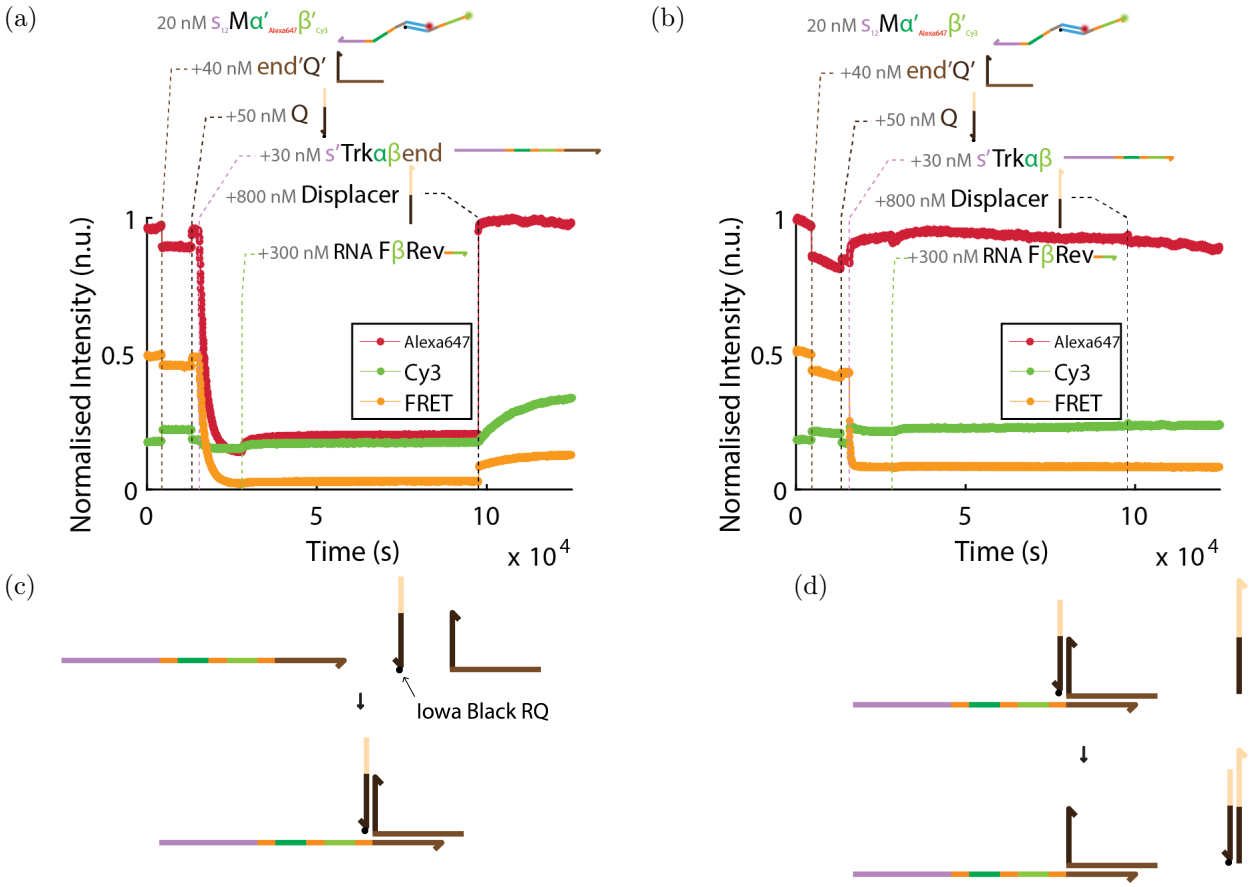

Figure 5.27: Testing the track-end marker on an oligonucleotide track. **a** Construction of oligonucleotide track end with Iowa Black RQ quencher by addition of components from schematic **c** to unblocked motor in solution. After assembly Alexa647 fluorescence is quenched to less than 15% its initial fluorescence level. Addition of a reverse RNA β fuel to lift the front β'-foot increases the Alexa647 fluorescence slightly as the fluorophore is further from the end quencher. A displacer strand can be used to remove the end quencher from the track as shown in schematic **d**. After displacing the end quencher Alexa647 fluorescence recovers to its initial high value. **b** Similar additions as in **a** except the oligonucleotide track is truncated and does not have the domain required to bind the track-end staple which hosts the quenching strand. As the Alexa647 fluorophore and Iowa Black RQ quencher are no longer co-localised, the Alexa647 fluorescence does not decrease as it hybridizes to the track.

5.9.3 The end of the track

The staple at the end of the track was extended with the complement to a 16 nt domain of a strand bearing an Iowa Black RQ fluorophore. This fluorophore should be close to the red fluorophore (Cy5/Alexa647) on the motor body when it is in the track-end position, but not in any of the other track positions.

Oligonucleotide tests of the track-end marker are shown in figure 5.27 The test track is assembled with unblocked motor already in solution to form the structure as shown in schematic 5.27c. To the 20 nM unblocked motor, 40 nM of the track-end-complement staple

is added, which also bears a domain complementary to quencher strand Q. 50 nM Q is then added which can hybridize to the track-end-complement staple. Little change in fluorescence is expected at this stage as there is no track present to change the environment of fluorophores on the motor, and this is what is observed in figure 5.27a. Finally 30 nM track with track end domain is added in figure 5.27a. The stoichiometries of component strands are chosen so that no motors should be free in solution, and that all tracks will have a staple and quencher strand on their track end. Some excess Q, and end'Q':Q complex will be free in solution but the lack of any major change in fluorescence when these components were added indicates that they will not have a major effect on fluorescence levels. After the track has been added, Alexa647 fluorescence drops by over 85%. This suggests that the track-end quencher is able to quench the red fluorophore on the motor when the motor sits at the end of the track. For the sample in figure 5.27b, the track added does not have a track end domain, therefore the track-end quencher cannot assemble on the track. In this sample there is not a decrease of Alexa647 fluorescence as the track is added, indicating that all Alexa647 quenching was due to the motor sitting on tracks where there was a quencher at the track end.

No increase in motor population at the track end.

As the track-end quenching scheme had been shown to work in principle on an oligonucleotide track, I applied the scheme to origami. Tightrope origami is assembled without the track-end complement staple, and after assembly 2× excess of the missing track-end staple and 4× quencher strand Q were added and incubated for 2 hrs at room temperature. This was done alongside motor loading. Tightropes were PEG precipitated to remove excess unbound staple or quencher and excess motor loading components.

Figure 5.28 shows only the Alexa647 fluorophore signal for a number of tightrope origami experiments where the tightrope bears the track-end quencher. The track-end quencher was displaced either before or after the addition of fuel and RNAse H in the fluorometer in different samples. All 4 samples were prepared from the same assembled origami, of which one quarter of the sample was removed prior to treatment with APE-1 (pink). The remaining 3/4 of sample had APE-1 added to cleave the loading domain. One of the three

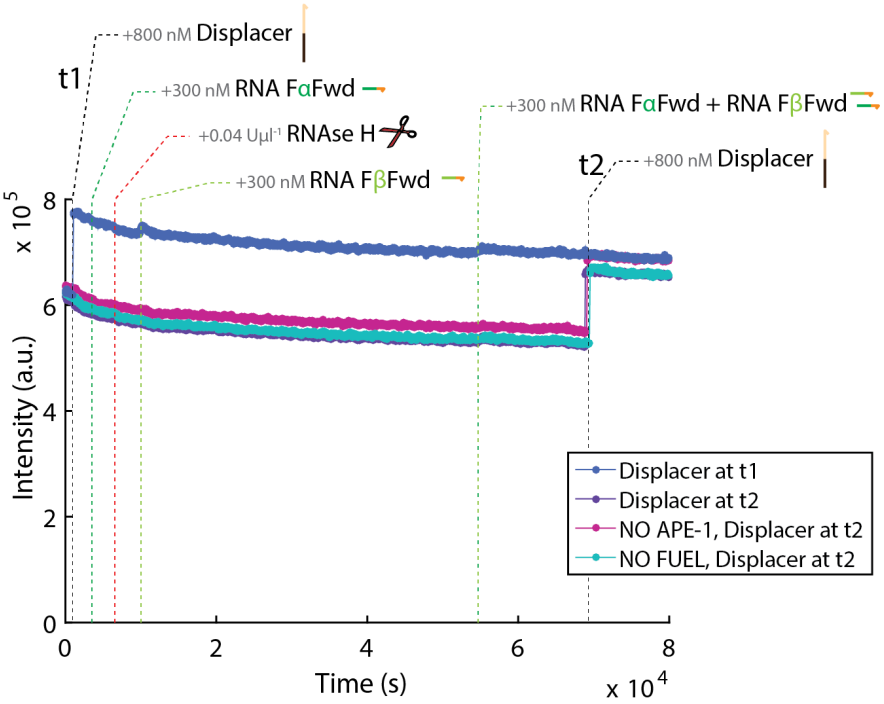

Figure 5.28: Plots of Motor Alexa647 fluorescence on tightropes with quencher Iowa Black RQ at the track end. To each sample displacer is added at either time t1 or t2 which strips the quencher from the tightrope. Blue trace shows blocked motor loaded on tightrope, released and digested with APE-1 before purification and observation. The displacer is added at t1, the experiment start before addition of any fuel or RNAse H. Purple trace is identical to the blue trace except displacer is added at t2, after addition of other components. Pink trace shows blocked motor loaded on tightrope, released but not treated with APE-1 so loading domain remains uncleaved. Cyan trace has buffer added instead of fuel. All samples show similar fluorescence increase after displacer addition.

APE-1 treated samples had RNAse H added to it whilst in the fluorometer, but no fuel strands (cyan). Three of the samples had the track-end quencher stripped at the end of the experiment (purple, pink, cyan), and one sample had the track-end quencher stripped at the beginning (blue). Regardless of APE-1 treatment during loading, presence of RNA fuels in solution, or time of stripping of the track-end quencher from tightropes, all samples showed similar increases in Alexa647 fluorescence as the end quencher was removed of around 20%. Because there was a change in fluorescence when the quencher was displaced (rather than no change as observed in figure 5.27b) it is likely that some motors are near the track end[12].

A crude estimate for the number of motors at the track end, assuming binding at any track position is equally probable[13], would be 20% at the track end as there are 5 possible

[12]This could be from a population of misfolded origami where the track is not stretched so separation between all track sites and the track-end domain is reduced.

[13]I have already demonstrated that on three-site oligo test tracks the β'α' orientation appears to be favoured at equilibrium in section 5.7.1, which may also apply on origami. This would contradict the assumption of equal binding probabilities.

positions. On oligo test tracks when 100% motors must be positioned at track ends, Alexa647 fluorescence is reduced to 85% of its unquenched intensity (see figure 5.27a). If I further assume that the fluorescence drop would be the same on origami as on the oligo tracks, and that quenching would be zero for motors on any track site other than the track end, 20% of motors at the track end would cause a 17% drop in fluorescence from the level after the track end quencher is displaced. This is close to the observed fluorescence of ∼20% below the unquenched level after displacement. Could this indicate that the motor is just positioned with an equilibrium distribution along the track, and that I have not managed to drive the system out of equilibrium even when fuel and RNAse H have been added? Even the sample where the loading domain could not have been cleaved showed a similar result, despite the expectation that uncleaved loading domains would bias motor binding at the track start.

A further difficulty arises from photobleaching. Both the gradual arrival of motors at the track end, and photobleaching of the red motor fluorophore would produce a decrease in red fluorescence. The ability to displace the track-end quencher provides a measure of the number of motors at the end at a single timepoint by comparison of fluorescence shortly before and after displacement, but this can only be done once for each sample.

To overcome this issue, the track-end quencher could be replaced with a track-end fluorophore able to produce a FRET signal with the red fluorophore Cy5/Alexa647. Using a green track-end fluorophore as a FRET donor, green fluorescence would decrease as the motor reaches the track end, however red/green FRET signal would increase.

5.10 Motor Model Including Rearrangement

Can the experimental result of no motor accumulation at the end of the tightrope track be explained by a simple kinetic model? A biped moving along a track of n positions can be modelled by a series of $3n + 2$ ODEs. There are three ODEs for motors at each track position corresponding to the three motor states in the model: Fully track-bound T_i, Back foot lifted with fuel, FxL_i, and single-stranded back foot lifted L_i. These states are depicted in figure 5.29.

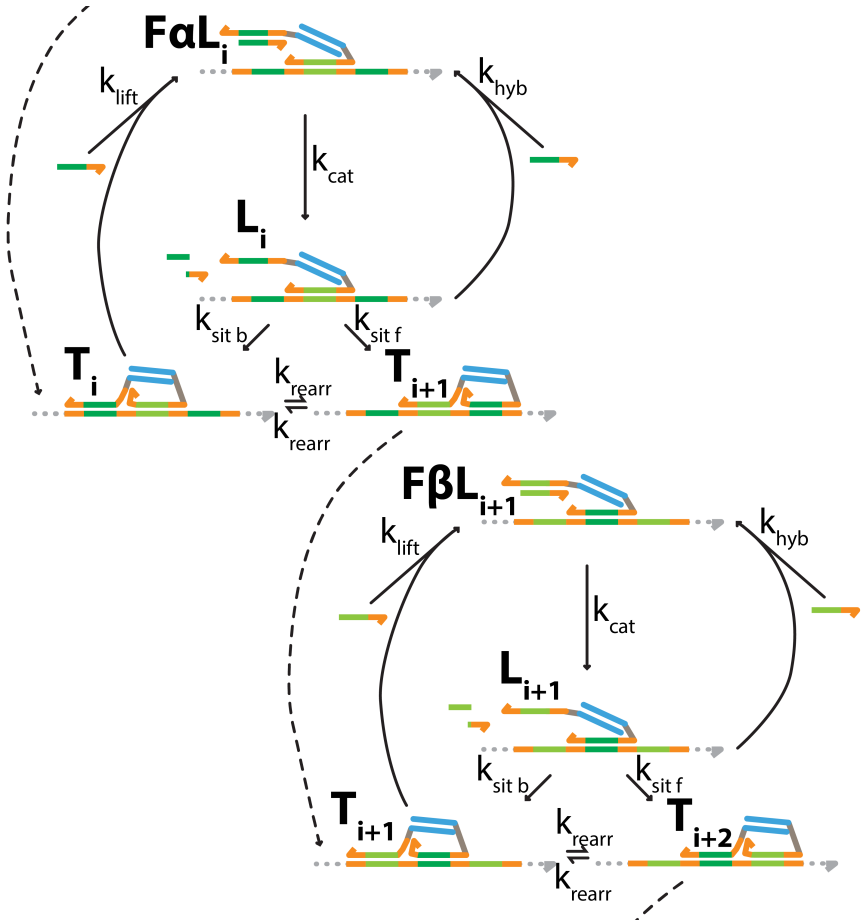

Figure 5.29: Schematic for a toy model of the biped motor including rearrangement.

For central track positions i=2 to 4, where $x = \alpha$ for odd i and $x = \beta$ for even i:

$$\frac{d[T_i]}{dt} = k_{\text{sit f}}[L_{i-1}] + k_{\text{sit b}}[L_i] + k_{\text{rearr}}\left([T_{i-1}] - 2[T_i] + [T_{i+1}]\right) - k_{\text{lift}}[T_i][Fx] \quad (5.1)$$

$$\frac{d[FxL_i]}{dt} = k_{\text{lift}}[T_i][Fx] - k_{\text{cat}}[FxL_i] + k_{\text{hyb}}[Fx][L_i] \quad (5.2)$$

$$\frac{d[L_i]}{dt} = k_{\text{cat}}[FxL_i] - k_{\text{hyb}}[Fx][L_i] - k_{\text{sit f}}[L_i] - k_{\text{sit b}}[L_i] \quad (5.3)$$

For start position $i = 1$:

$$\frac{d[T_1]}{dt} = k_{\text{sit f}}[L_0] + k_{\text{sit b}}[L_1] - k_{\text{rearr}}\left([T_1] + [T_2]\right) - k_{\text{lift}}[T_1][Fa] \quad (5.4)$$

$$\frac{d[F\alpha L_1]}{dt} = k_{\text{lift}}[T_1][F\alpha] - k_{\text{cat}}[F\alpha L_1] + k_{\text{hyb}}[F\alpha][L_1] \quad (5.5)$$

$$\frac{d[L_1]}{dt} = k_{\text{cat}}[F\alpha L_1] - k_{\text{hyb}}[F\alpha][L_1] - k_{\text{sit f}}[L_1] - k_{\text{sit b}}[L_1] \quad (5.6)$$

For end position $i = 5$:

$$\frac{d[T_5]}{dt} = k_{\text{sit f}}[L_4] + k_{\text{sit b}}[L_5] + k_{\text{rearr}}\left([T_4] - [T_5]\right) - k_{\text{lift}}[T_5][F\alpha] \quad (5.7)$$

$$\frac{d[F\alpha L_5]}{dt} = k_{\text{lift}}[T_5][F\alpha] - k_{\text{cat}}[F\alpha L_5] + k_{\text{hyb}}[F\alpha][L_5] \quad (5.8)$$

$$\frac{d[L_5]}{dt} = k_{\text{cat}}[F\alpha L_5] - k_{\text{hyb}}[Fx][L_5] - k_{\text{sit f}}[L_5] - k_{\text{sit b}}[L_5] \quad (5.9)$$

Note $[T_0](t) = 0$, as this state does not exist. The states $[L_0](t)$ and $[F\beta L_0](t)$ do exist, but cannot be entered because the model does not allow front foot lifting.

Assumptions of the model:

- All fuels are forward fuels and only lift the intended foot. There is no front-foot lifting, or lifting with mismatched fuel.
- Rate constants for similar processes involving alpha or beta domains are the same.
- Foot lifting with fuel is irreversible.
- No foot replacement bias. Sitting down lifted foot on either side of the bound foot is equally likely ($k_{\text{sit f}} = k_{\text{sit b}}$) [1]. Note that simulations show this is not the case for a track under tension [123].
- Sitting down is irreversible.
- Rearrangement only occurs between adjacent track-bound positions T.

Model parameters can either be empirically estimated, or estimated from literature values. Default values used in the model for each parameter are given in table 5.2.

Parameter	k_{sit} (s^{-1})	k_{lift} ($\text{M}^{-1}\text{s}^{-1}$)	k_{cat} (s^{-1})	k_{hyb} ($\text{M}^{-1}\text{s}^{-1}$)	k_{rearr} (s^{-1})
Value	[a] 2×10^4	[b] 7×10^4	[c] 7	[d] 10^7	[e] 3×10^{-4}

Table 5.2: Motor model parameters estimated from experiment and literature.

[a] Hybridization rate for a 16mer [170] at local concentration 1 per 6 nm radius.
[b] Second order time to half completion from foot lifting on two-site tracks in figures 5.23a, 5.23d and 5.24a. This is similar to foot lifting rate $k_{\text{lift}} \approx 6 \times 10^4$ $\text{M}^{-1}\text{s}^{-1}$ reported in reference 1.
[c] [194, 195].
[d] [170].
[e] Assumes rearrangement occurs on a timescale of ∼1 hour. This may underestimate rearrangement rate.

Most motors reach the track end position $i = 5$ after 10^4 s, regardless of whether the motors were initially at the start position 1 (figure 5.30a) or evenly distributed across all 5 positions (figure 5.30b). Changing the rate constants k_{sit}, k_{hyb}, and k_{cat} by ± 3 orders of magnitude had little or no effect on the observed results. k_{lift} appears to be the rate limiting step in the turnover of fuel in this model, and the motor appears to spend negligible time in a foot-lifted state. A simpler model could therefore be used where all the processes involved in stepping are further coarse grained into a single effective rate constant.

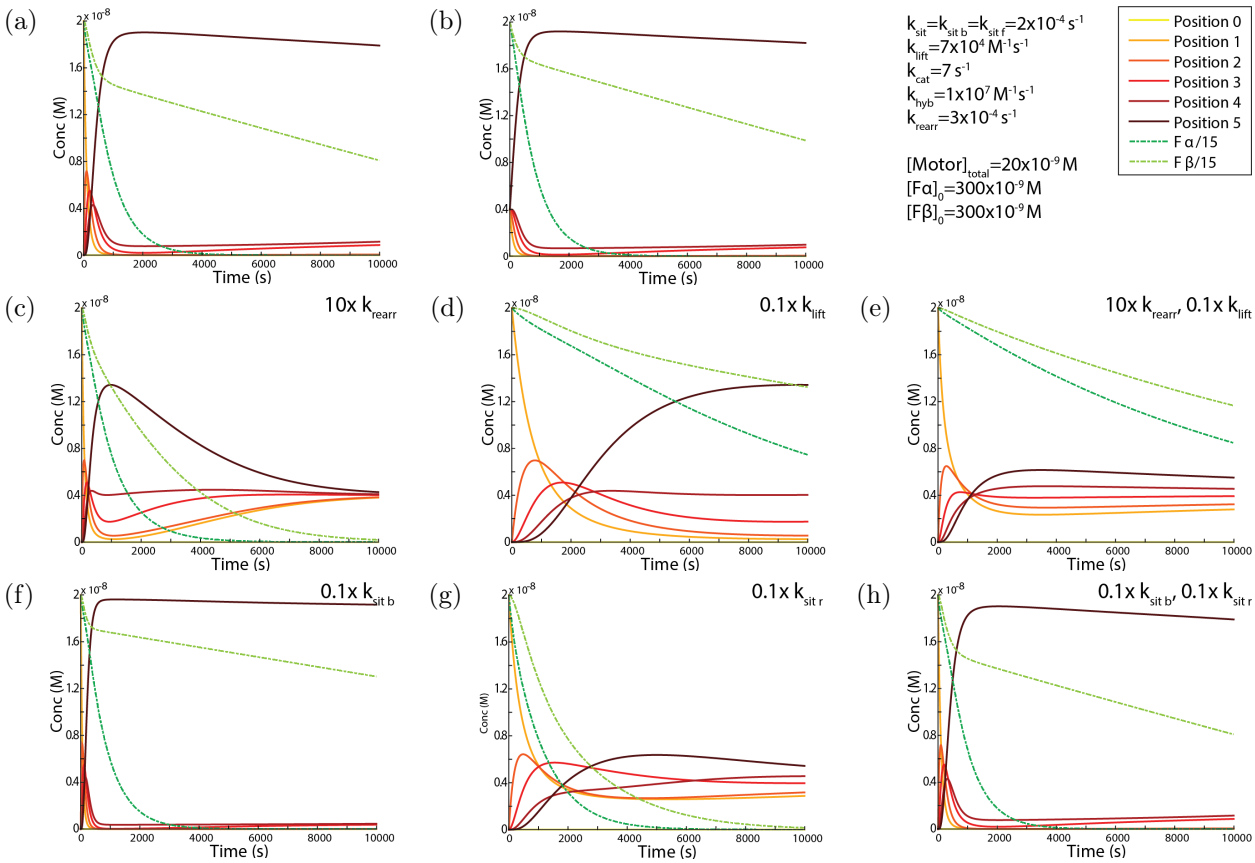

Figure 5.30: The effect of altering model parameters on motor position at time t. Motor concentrations at positions 0 to 5 are given by solid lines from yellow to dark red. Fuel concentrations (divided by 15, the fuel to motor excess) are shown by green dashed lines. **a** Default parameters and motors at position 1 at t=0 s. **b** Motors evenly distributed across the track positions at t=0 s. **c** k_{rearr} increased tenfold. **d** k_{lift} decreased tenfold. **e** k_{rearr} increased tenfold and k_{sit} decreased tenfold, **f** $k_{\mathrm{sit\ b}}$ decreased tenfold. **g** $k_{\mathrm{sit\ f}}$ decreased tenfold, **h** $k_{\mathrm{sit\ b}}$ and $k_{\mathrm{sit\ f}}$ decreased tenfold.

The model was sensitive to perturbing foot-lifting rate constant (see figure 5.30d), where a tenfold decrease was able to both decrease the maximum concentration of motor reaching the track-end position 5, and to increase the amount of time for this peak to be reached. Foot lifting rate is a model parameter with which I have reasonable confidence on oligonucleotide tracks, as it has been empirically determined, and the value agrees well with previous work [1]. The effect of the tightrope on foot lifting rate, however, has not been well characterised. It is conceivable that tension, and proximity to DNA origami would affect the rate of TMSD. Studies of TMSD on strands tethered at one end to a DNA origami tile have shown slowed kinetics [149], so the figure 5.30d may be physically relevant.

Increasing the rearrangement rate constant tenfold was also shown to decrease the maximum concentration of motor reaching the track end as shown in figure 5.30c. The distribution

of motor positions after 10^4 s in this case was even closer to the equilibrium motor position distribution. The initial value of $k_{\text{rearr}} \approx 3 \times 10^{-4}$ s^{-1} may be an underestimate, as it was obtained by assuming a rearrangement time constant of 1 hour, as rearrangement has occurred within an hour (see figures 5.18, 5.25a). The track tension, and fixing of track ends is anticipated to reduce rearrangement, as one class of rearrangement mechanisms would need to pull track against a restoring force to bring necessary domains close together. Rearrangement could also occur by transient melting and hence (non-fuel mediated) lifting of foot from track, followed by replacement. This mechanism may be less hindered by tension. It is possible that the rearrangement rate constant used in figure 5.30c is physically relevant.

Combining both of these effects as shown in figure 5.30e may be sufficient to never allow the motor to escape from a position distribution close to its equilibrium position distribution.

Another way to influence the model behaviour is to introduce a foot replacement bias, such that $k_{\text{sit b}} \neq k_{\text{sit f}}$. Simulations using oxDNA suggest that forward foot replacement bias can be introduced by increasing track tension to 14.6 pN [123], where the ratio of binding the site in front of the track-bound foot to binding the site behind the track-bound foot was 34:7, but in the absence of tension, foot replacement showed no considerable bias [1, 2]. By slowing the replacement rate $k_{\text{sit b}}$ tenfold, more motors reach the end of the track, and do so more rapidly (figure 5.30f). The restoration to equilibrium by rearrangement is also slowed. This is because a greater foot replacement bias tightens the coupling between fuel turnover and forward stepping. The opposite case, where the rate of replacement in front of the track-bound foot $k_{\text{sit f}}$ is slowed tenfold (figure 5.30f), shows a decrease in motors reaching the end of the track, due to the weakened coupling between fuel turnover and forward stepping. If the rate $k_{\text{sit}} = k_{\text{sit b}} = k_{\text{sit f}}$ is decreased tenfold but the ratio of forward and backward replacement is not changed, the model behaviour in figure 5.30h is unaffected.

The results from the simple model including rearrangement demonstrate that there are physically relevant regimes where the motor may not be able to accumulate appreciably at the track end position due to rearrangement, however it seems unlikely that some transient pulse of motors reaching the track end before fuel depletion would not be observed in experiments

such as those shown in figure 5.28. The assumptions used to build the model may be too crude to capture aspects of the motor behaviour. Sequence dependent effects may cause equivalent processes involving α or β domains to occur at different rates. I have already suggested in section 5.7.1 that there may be preferred motor orientations, and this is not captured in the model. Effects of the tightrope origami on dynamic processes are still unknown. It remains ambiguous whether observed experimental results can be explained by rearrangement.

5.11 Future Directions

It would be interesting to clone new two and three-site tracks into the origami to examine the effects of shorter tracks under tension, and to control for the effect on red fluorophore fluorescence when all motors can only bind to the track-end binding site.

Fuel turnover could be observed in bulk using a doubly-labelled fuel strand and watching the fluorescence change as it is broken down. Depletion of β-fuel is indicative of rearrangement, a lack of stepping of motors, or undesired interaction between motors and β-fuel, and might provide a useful diagnostic for the lack of motors reaching the track end.

Single molecule imaging experiments such as high speed AFM or super-resolution localisation microscopy might enable the observation of motor diffusion along the tightrope track if it is occurring. Single molecule FRET could also be used to examine conformations.

As the RNA fuel does not lift as effectively as DNA fuel it may be worth considering a chimera fuel with two or more consecutive ribonucleotides in an fuel primarily formed of DNA. This should be cleavable by RNAse H[14].

The motor would not be without use if it transpired that it does not out-perform diffusion. Applications of a diffusing 'shuttle' on a 1D stretched track could be considered. Such a shuttle may be able to facilitate the diffusion of some cargo from one side of the track to the other. Indeed diffusive motion (rather than ATP driven directed motion) is harnessed by the kinesin MCAK as approach to depolymerising microtubules [196, 197]. Remaining associated to the microtubule reduces dimensionality of its diffusive search.

[14]Rather than the single ribonucleotide version which requires RNAse H II.

5.12 Summary

I have adapted a bipedal motor to have non-identical feet and I am able to specifically load this motor onto tightrope origami using a mechanism where motor feet are blocked to reduce non-specific track binding, and a specific loading domain is added to guide motors to a specific site. I have demonstrated that foot lifting works in an orientation dependent manner on two-site test tracks, and that forward fuels are only turned over when the orientation of the motor has the foot matching the fuel used positioned at the back.

Results become more ambiguous on three-site tracks, where a motor can occupy one of two possible positions on the track: start or end. Track-bound motor orientation on gel can be inferred from Cy3 quenching by Iowa Black FQ fluorophore when the dye and quencher are at 3' and 5' ends of one of the foot strands. Foot lifting can be observed on three-site tracks however, it appears that the ratio of track-bound states is not significantly changed by foot lifting, and once the fuel has been digested by enzyme and is free to rebind track, there is no significant change in the ratio of track-bound states either. This behaviour can be explained by rearrangement of motors on the track. Further evidence for rearrangement was found from 2D gels where the mobility of a species from a single track position changed after incubation at room temperature producing off-diagonal bands in a 2D gel.

Fluorescence changes can be observed using bulk spectrofluorometry as the motor is released from its blockers, and as it interacts with fuel, lifting feet. Foot lifting can be observed both on oligonucleotide-based test tracks and on tightrope origami. A method to report motor arrival at the track end was demonstrated on test tracks, however when used on tightropes, it appeared that the population of motors at the end of the track was unchanged by the addition of fuel and enzyme. This is consistent with rearrangement working against any effect the motor has of stepping forward.

Chapter 6

Long Distance Locomotion of a Burnt-Bridges Motor

In this chapter we investigate two methods to create extended tracks for a DNA burnt-bridges motor to walk along >100 nm distances. The first method places a track on the inside face of DNA origami nanotubes. We show that the motor can move between two polymerised DNA nanotubes connected in register. We attempt to observe motors in DNA nanotubes with TIRF microscopy. We can infer the presence of a fluorophore-labelled motor in an nanotube by stepwise photobleaching. We also attempt to observe the cutting of fluorophores from a track by an operating motor.

The second method attempts the construction of more complex track designs by polymerising multiple, distinguishable, DNA tiles in a defined order. Constructing a track from separately-addressable tiles would be useful for DNA motor-based computation, currently limited by the dimensions of single tiles. The method succeeds in producing some polymers of the desired pattern, however there are still undesired interactions between monomers of the same type. I finally propose a method by which it may be possible to perform single-molecule imaging of locomotion down complex tracks with DNA PAINT. I have imaged a zig-zag track with a resolution of at least 50 nm using DNA PAINT. Future work could use DNA PAINT to watch the burning of a track as the motor moves along it.

Author contributions

The work on motor operation in DNA nanotubes was a collaborative effort between myself, Florence Benn and Alexandra Lucas (University of Oxford UK). The majority of experimental design and execution was performed jointly except for specific areas where one

person performed the experiment/design, detailed below.

- DNA nanotube design was performed solely by Florence Benn. A brief summary of her design is included in this thesis. Detailed designs can be found in reference 198.
- All microscopy (AFM, TIRF) and electrophoresis included in figures was performed by myself. AFM studies of nanotubes performed by Florence Benn are not presented in this thesis.
- The original polymerisation scheme was designed by Florence Benn, but the asymmetric polymerisation scheme was designed by myself.
- Fluorometer experiments were carried out by myself using protocols developed jointly.
- Tube assembly was carried out by either Florence Benn, or myself.
- Alex Lucas performed the analysis and processing of microscopy images in figure 6.9 under advice from both myself and Florence Benn.
- Figure 6.1 depicting the DNA nanotube design was produced by Jonathan Bath.

The work on burnt-bridge tracks spanning multiple tiles was carried out by myself, however, I benefited from a collaboration with Anna Kostina (Technische Universität München DE), during which I visited the lab of Prof. Friedrich Simmel (Technische Universität München DE) . AFM images in figures 6.12 and 6.13a were acquired by Ali Aghebat Rafat (Technische Universität München DE) during this secondment.

6.1 Motivation

Burnt-bridges motors have been demonstrated to walk along 2D origami tiles [87,108,109], to perform logical computation [108, 199], and sequential, programmable, chemical assembly [101]. Motor-based computation could be improved by placing the burnt-bridges tracks across multiple tiles to both reduce error [199], and allow more logic gates to be incorporated, increasing the complexity of possible calculations. Additional spatial compartmentalisation of a burnt-bridges motor, for example by placing it on the interior of a tube or wire, could enable the assembly of rotaxanes, as a cargo is threaded through the tube. Cargo transported on the interior of a hollow DNA structure would be in a different, and tunable, local environment, which may prove useful in chemical assembly lines.

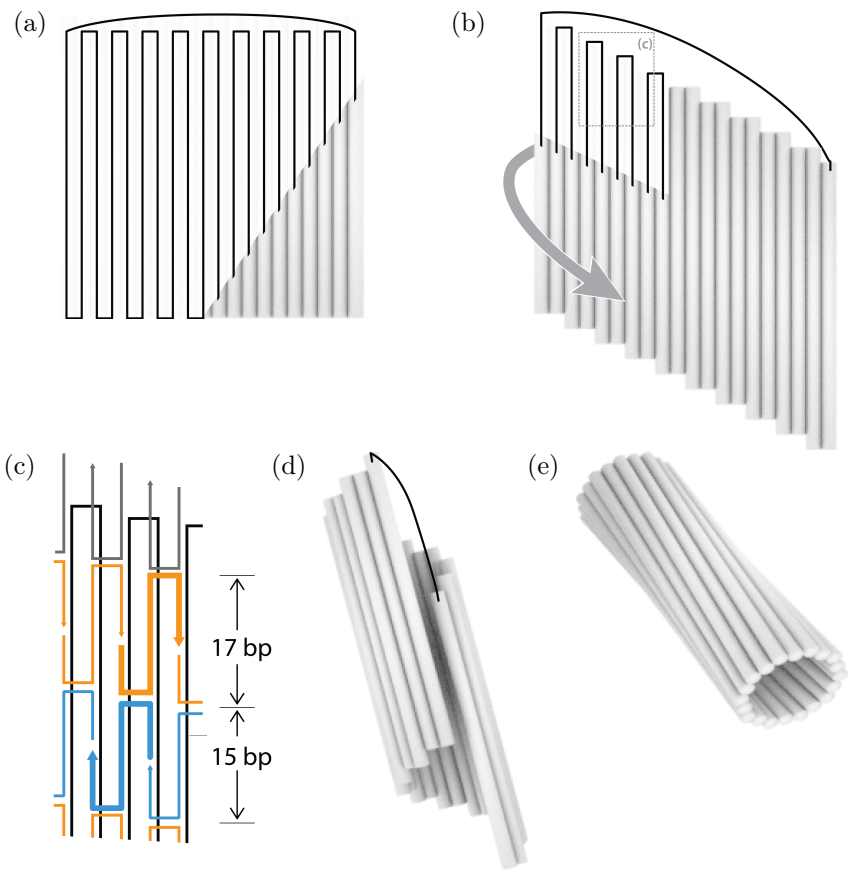

Figure 6.1: Design of a DNA nanotube. Helices are depicted by grey cylinders, staples by grey, orange or blue lines with 3' arrowheads, and scaffold by black lines. The loop of scaffold linking helices at the top of **a**, **b** and **d** is of length 0 nt. **a** A 2D tile-like DNA origami with scaffold connection between helices on opposite edges can form a DNA nanotube. **b** Alternate helices are staggered by changing the spacing of crossovers. **c** Insertion and deletion of base pairs either side of a crossover introduces global curvature. **d** The curvature curls the tile into a tube with the inner face defined by the curvature. **e** Staple linkages hold pairs of adjacent helices together creating a tube with slight right-handed chirality. Figure by Jonathan Bath reproduced here with permission.

6.2 Motors Through DNA Nanotubes

6.2.1 DNA nanotubes with addressable interior and exterior

DNA nanotubes can be formed by connecting opposite edges of planar DNA structures such as tile arrays [200,201], or two-dimensional DNA origami tiles. DNA nanotubes formed in this way can have a range of diameters, as multiple monomers can fold together to form tubes with larger radii. There is also little control over which face of the tile becomes the inside or the outside of the DNA nanotube. To address these issues, our nanotube design takes a tile-like DNA origami with scaffold routing with periodic boundary conditions (figure 6.1a). The tile can be made to curl in a defined direction [77] (figure 6.1d), by shifting the position of a crossover between pairs of helices by one bp relative to the one between the next

pair of helices. This is accomplished by altering the staples of the tile such that there is an addition of a nucleotide one side of the crossover, and a deletion of a nucleotide on the other side (figure 6.1c). The face which becomes the inside of the tube is defined by the introduced twist. This also hinders the formation of tubes with larger diameters. Helices at opposite edges of the tile were linked by staple crossovers such that the tube forms a flat circular face at one end as shown in figure 6.1e, also giving the tube a slight right-handed chirality. The adressability of the inside and the outside faces of the DNA nanotube, and its designed chirality were confirmed by circular dichroism experiments [198, 202] on tubes assembled with gold nanoparticles attached to positions either inside or outside the nanotube. The approximate tube dimensions are 120 nm length and 10 nm radius. Such a nanotube forms the platform for a burnt-bridges DNA motor presented in this chapter.

6.2.2 Polymerisation of DNA nanotubes

The nanotubes can be polymerised to form long strings of multiple tubes in series. The simplest way to do this is to leave helix ends blunt, then DNA tubes meta-stack due to base-stacking at the ends of the 20 component helices. This can produce long filaments, but the filaments are only held together by stacking interactions; there are no hydrogen-bonding interactions holding neighbouring tubes together to strengthen the links. Additionally the rotational offset of tube stacking about the tube axis is not controlled. To create a continuous track for a DNA motor along a series of nanotubes we require the connection between tubes to occur in register so that the end of the track on one tube lines up with the start of the track on the next tube. Directed stacking of tubes can be achieved with two-domain linker staples which have one domain on a helix on one face of the tube and the other domain on a corresponding helix on the opposite face (figure 6.2b), whilst undirected stacking can be reduced by adding poly-T edge staples (figure 6.2a) [74].

Forming multimers after annealing is difficult as formation of a kinetically-trapped state is possible where all domains on the ends of a single tube are hybridized to a single domain of a linker staple, with the second domain left single-stranded. Two such tubes would have to exchange domains by blunt strand displacement processes to successfully form a polymer.

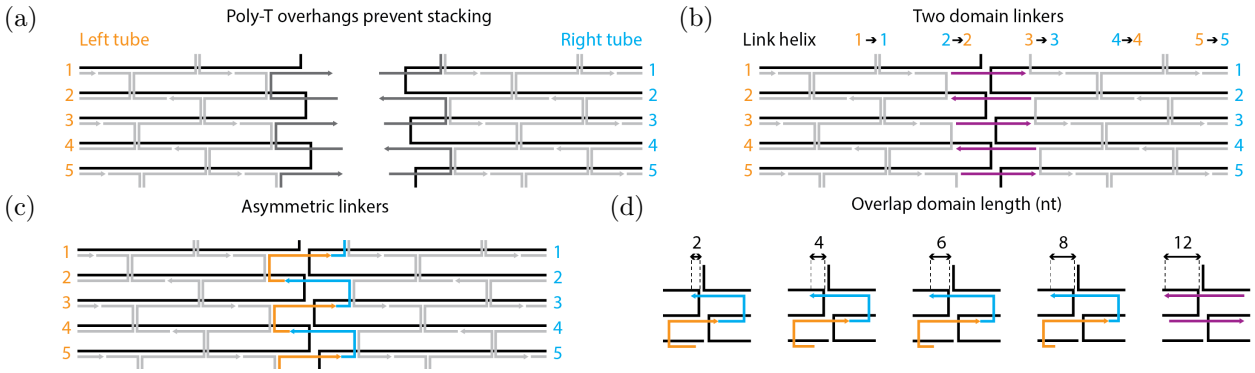

Figure 6.2: Tube edge schematics for linkable or passivated faces. Scaffold strand is shown in black. **a** Stacking between tubes is prevented by edge staples with poly-T overhangs. **b** Tubes can be joined by linking pairs of odd or even helices together. **c** Introducing asymmetry into the staples linking the helices gives each linking staple a preferred tube edge which may improve polymerisation post annealing. **d** Length of the domain which overlaps onto the opposite tube edge. This length is used to name asymmetric staple types.

They are effectively passivated against doing this, therefore two-domain linkers are not ideal for forming multimer structures after annealing.

I introduced asymmetric linkers, in place of the two-domain linkers, to improve assembly of polymers. In the asymmetric linker scheme each helix link between two edges is mediated by two staples. Staples hybridize a long domain on one helix on one edge, a short domain on a second neighbouring helix, and a third domain on the same helix of the opposite edge (see figure 6.2c). It is expected that the long domain will hybridize first, followed by the short domain of the neighbouring helix, leaving edges with sticky ends and sticky recesses complementary to the opposite edge, which associate forming the polymer.

Optimal domain lengths were determined by agarose gel electrophoresis (figure 6.3 left). One type of tube (A) was assembled with poly-T staples at one edge, and no edge staples on the opposite edge, and with 21 Cy3 fluorophore labels. The other tube (B) was assembled without edge staples on both ends, and no fluorophore labels. After assembly both tube types were PEG purified to remove excess staples (see appendix A). A tubes were mixed with a 10× excess of linker staples, with and without a 5× excess of B type tubes. By comparing fluorescence in the well (from polymerised nanotubes too long to enter the agarose matrix) to fluorescence in the monomer band (from A tubes which are not part of a polymer) across samples, polymerisation can be compared for the different domain lengths of asymmetric linkers. Ideally A tubes should not be able to polymerise with each other due to the poly-T

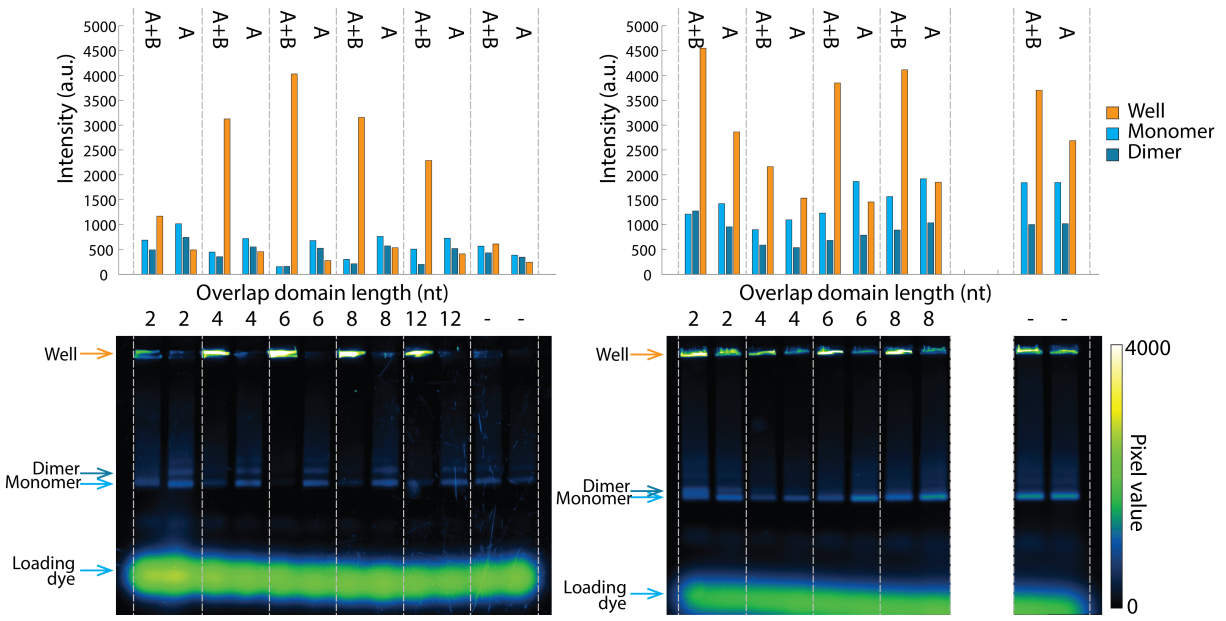

Figure 6.3: Determining optimal tube linkers (see figure 6.2). Agarose gels showing polymerisation of A, or A+5×B nanotubes for different linker staple domain lengths. Left shows polymerisation of purified nanotubes assembled without linker staples, then incubated with a 10× excess of linker staples. Right shows polymerisation of nanotubes assembled with 10× excess linker staples and purified before mixing together. Type A tubes are labelled with 21× Cy3 and have only one linkable edge, the other edge is passivated with poly-T tailed staples. Type B tubes have both edges available for polymerisation and no fluorophore. Integrated band intensities for the monomer band, dimer band and well (containing long polymers) are shown in the bar graphs (top).

staple extensions on one of the edges, and should only form filaments in the presence of B tubes. We seek a linker able to form long fluorescent structures (remaining in the well) when A, B and linkers are present, but not when B tubes are omitted. The optimal linker for this assembly was an asymmetric linker with short domain length of 6 nt, because it displayed maximum fluorescence in the well, and minimum fluorescence in the monomer band when B tubes were present, but one of the lowest well fluorescence values in the absence of B tubes.

We also tried adding linker staples during assembly (figure 6.3 right), in the hope that polymers with sticky ends and recesses would be produced, which would stick together when mixed. Again, A tubes were the only tube type with fluorophore labels, and they were assembled with one poly-T edge and the other edge containing half of the linker staples with their long domain on that edge. B tubes were assembled with all of the linker staples. Both A and B tubes were purified by PEG precipitation before mixing them together with a 5× excess of B tube to A tube. This assembly method also suggested the overlap domain length of 6 nt was optimal, as it gave one of the highest well to monomer ratios when A and B

were present, but one of the lowest well to monomer ratios with A alone. There is also no dimer band evident, in contrast to the case where the short domain is 2 nt where presumably stacked dimers can form in both A:B mixtures and A alone.

These results suggest that adding linkers after tube assembly enables more successful polymerisation than adding linkers during assembly as it gives a higher well to monomer ratio in the A+B case, and lower well to monomer ratio in the A only case. When tubes are polymerised by addition of linkers after tube assembly, most A tubes become part of tube polymers, so if an A tube can be identified, it is likely to be part of a polymer. This does not suggest that most tube polymers will include an A tube, so if a polymer can be identified, it does not indicate that the polymer contains an A tube.

To confirm that these structures are polymerised tubes, and not an aggregated clump of DNA, similar samples were imaged by AFM and TIRF microscopy. Results are shown in figure 6.4. AFM of a sample of purified B tubes purified with S-300 (see appendix A) then polymerised with 10× excess asymmetric linker staples with 6 nt overlap domains, confirms that tubes do form long polymers when polymerised in this way, however there is evidence of aggregation of tubes observed on the mica. I cannot be sure if this is due to deposition of the tubes on a surface or whether aggregation would be seen in solution.

The dynamics of the polymerisation process was observed with TIRF microscopy. Surface immobilised A type tubes (able to polymerise from only the right hand edge) were prepared, then asymmetric linkers with 6 nt overlap domains and B type tubes (able to polymerise from both ends) were added to the flow chamber in the microscope. All tubes were labelled with multiple Cy3 fluorophores during assembly and purified before adding to the flow chamber. Images at the surface of the coverslip show the gradual formation of long thin structures, consistent with nanotubes and some large bright spots which could be aggregates.

6.2.3 Adaptations to the burnt-bridges motor and stators for observation inside nanotubes

To the polymerised nanotubes, we wished to add a motor able to transport cargo through the tube interior. We used a burnt-bridges motor [3, 87, 108] to achieve this. Please see

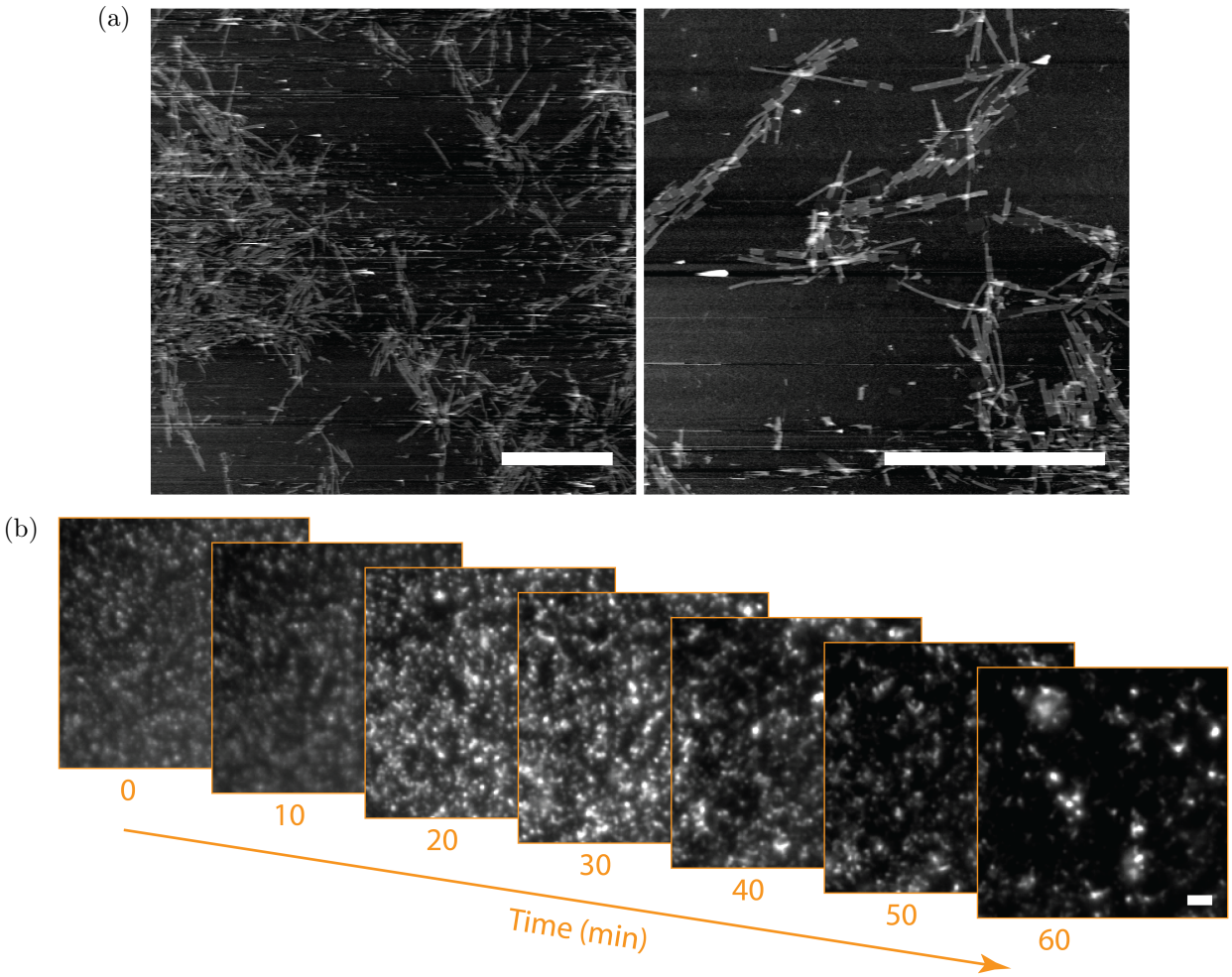

Figure 6.4: Observation of nanotube polymerisation by AFM and TIRFM. Each tube monomer has length ∼120 nm and radius ∼10 nm **a** AFM images of nanotubes polymerised with 10× excess asymmetric linker staples with 6 nt overlap domains deposited on mica. **b** TIRFM images showing nanotube polymerisation over time. Each frame shows a median projection from 500 frame movies (individual frame 100 ms exposure, 200 EM gain) collected in 10 minute intervals after the addition of asymmetric linker staples with 6 nt overlap domains, and nanotubes able to polymerise from both ends labelled with Cy3 fluorophores to a surface functionalised with seed tubes. Scale bars 2 μm.

chapter 1 section 1.3.4 and figure 1.3c for a description of the burnt-bridges motor mechanism. Burnt-bridges motors were observed using either bulk fluorescence or TIRF microscopy. The former required a fluorescence change when motors move along a track and was achieved by monitoring fluorescence at different stator positions along the track. The latter required marking either the motor itself with a fluorophore, or marking the track with fluorophores removed when the motor passes.

The burnt-bridges motor used had the same sequence as reference 87, with 2 nt to thermodynamically bias binding to a start type stator at the 5' end, followed by the 22 nt displacement domain, and a 3' 6 nt toehold domain (immediately above the nicking site for

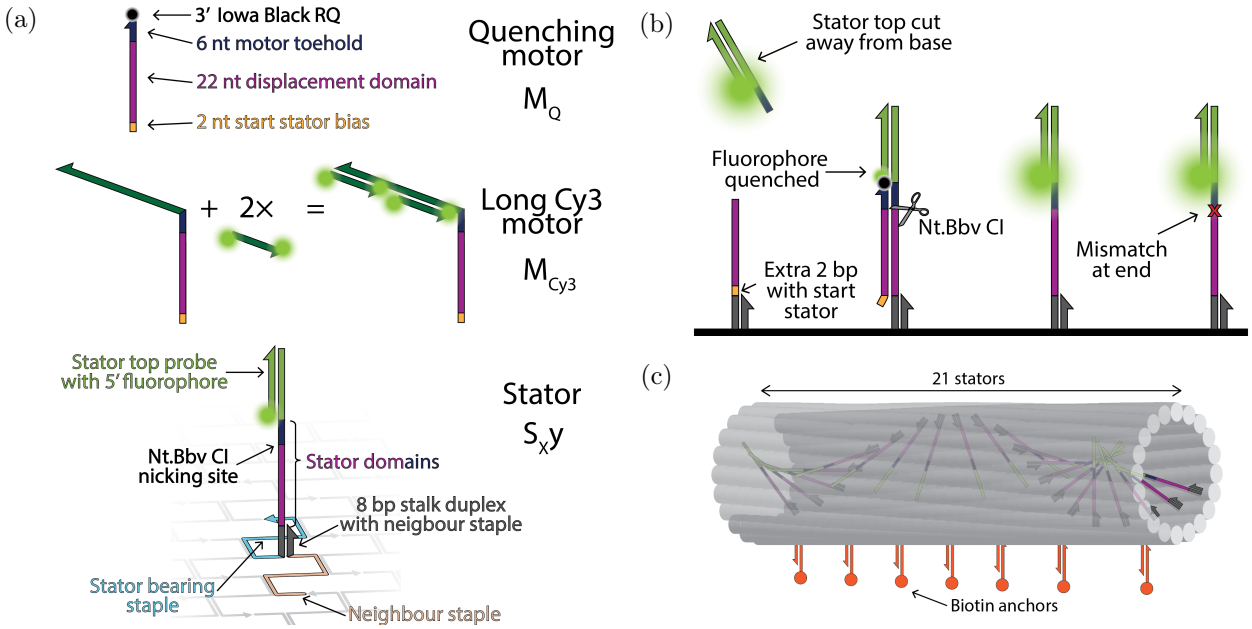

Figure 6.5: Adding a burnt-bridges motor to a DNA nanotube. **a** Motor and stator design. Quenching motor M_Q (top) with 3' Iowa Black RQ quencher will quench fluorophores on the stator it occupies. Long motor (middle) M_{Cy3} extended to hybridize two dual-fluorophore-labelled strands. Stator (bottom) with motor complement domains at the centre appended by an 8 bp stalk domain at the base, and a stator-top domain with a complementary stator-top probe. Stator top probe and or stalk neighbor staple are not integral to motor movement. **b** Schematic of burnt-bridges motor with stator-top probes. As the motor passes a stator, Nt.BbvCI enzyme cleaves the stator toehold and stator-top from the base, allowing it to dissociate. In the fluorometer, with M_Q, this will increase fluorescence of the fluorophore on the stator-top as it moves away from the quencher. In TIRF microscopy, the motor should leave a dark wake behind it as fluorescent probes cut from the top of the used stators diffuse out of the tube, while unused stators in front remain bright. **c** Schematic of stators inside the nanotube. Stators point into the nanotube and are spaced by 6.2 nm and one helix such that stator positions have greater right handed chirality then the tube. Each tube has a maximum of 21 stators and the first and 21st stator are on the same helix. Seven biotin anchors are added to two helices along the tube axis for specific adhesion to a microscope coverslip.

Nt.BbvCI). For observation in the fluorometer, an Iowa Black RQ quencher is appended to the 3' end of the motor to form quenching motor M_Q. We also use a form of the burnt-bridge motor where, instead of a 3' quencher, the 3' end is extended by two copies of a domain complementary to a strand labelled with Cy3 at both ends named M_{Cy3} (see figure 6.5a). This motor is used for microscopy, and has multiple fluorophores to increase its overall fluorescence lifetime and to increase the number of emitted photons which would allow higher localisation precision if used for super-resolution tracking[1].

Stator sequences are complementary to the 22 nt displacement domain and 6 nt toehold domain of the motor. Stators were extended with an additional 16 nt domain to hybridize a

[1]Localisation precision scales as the square root of the number of emitted photons [203].

'stator-top' probe at a position above the motor binding site as depicted in figure 6.5a. The fluorescent probe will be proximity-quenched when the motor, bearing a quencher, occupies that stator. As the stator-top with probe fluorophore is cut during motor operation, it will be liberated into solution increasing its fluorescence as depicted in figure 6.5b. The base of the stator was also extended by an 8 nt stalk domain designed to reduce the threading of the stators through the nanotube origami walls. If a stator threads through the walls, it will be on the wrong (outer) face of the tube and will not be accessible for branch migration of the motor, effectively creating a gap. Threading has been reported to reduce when a stalk is added to a strand to reduce the flexibility of the DNA close to the origami [204]. This was done by creating a duplex between the stator-bearing staple and the neighbouring staple (see figure 6.5a) which should form pointing perpendicularly inwards during the annealing process. In order to thread through the tube, the single-stranded domain of the stator would have to reach further. Stators are identical except for the start stator, which can form an additional 2 bp with the motor, and the end stator, which forms a mismatch in the Nt.BbvCI nicking site to prevent the cleavage of the final stator and trap the motor at the track end [3, 87, 108]. If end-stator cleavage was not prevented, the motor would diffuse along the burnt track as all sites would be isoenergetic. Stators are named $S_X y$ where X is the tube on which the stator is located, and y is the stator position (1-21) along that tube.

To stick down tubes on a coverslip surface for TIRF microscopy we introduced 7 staples with 16 nt extensions complementary to a biotinylated oligonucleotide containing a further 30 nt spacer between the binding domain and the 5' biotin modification (see figure 6.5c). This method should enable the specific attachment of tubes onto a streptavidin functionalised microscope coverslip (for full protocol see appendix D).

6.2.4 Design issues

Undesired stator secondary structure

The original stator-top sequence produced a hairpin secondary structure. This interferes with the operational portion of the stator by partially blocking the motor toehold domain when the stator-top is single-stranded. When the stator-top is double-stranded this should

not be an issue, but if stator-tops are not annealed into the structure to begin with, but incubated with it afterwards, it is possible that some hairpins would remain and slow the migration of the motor down the track. Some of the experiments presented were carried out with this version of the stators, and where this is the case it is noted.

Misplaced stator S2

The position on origami of stator S2 was mistakenly two helices from its correct position in the majority of experiments, as the mistake only became apparent later on. This would act to slow the motor down as to get from S1 or S2 to S3 and onwards would require a double length step, shown to slow the motor [87, 156]. All experiments except the bulk fluorometer tube crossing experiment in section 6.2.6 were carried out with S2 in the wrong position.

Direction of anchor staple extensions

The anchor staples were intended to point radially outwards from the outer face of the tube as shown in figure 6.5c. Nicks between staples point radially inwards from the internal face of the tube. To obtain radially outward pointing anchor domains the position between staple breaks must be moved along by 16 bp (∼1.5 helical turns of the DNA double helix i.e. facing the opposite direction), however the anchor domains were actually appended to the current nick positions so they will point inwards. This may reduce the ability of the biotinylated tubes to stick to the streptavidinated surface. Some specific sticking of nanotubes to a streptavidinated surface was observed, compared to a negative control of non-biotinylated tubes. This might be because some of the anchors have threaded through the walls of the tube, and are therefore on the outside, and once on the outside and stuck to the streptavidin, they can no longer rethread back inside the tube [205]. All microscopy experiments used this form of the anchor staples.

6.2.5 Tests of enzyme cutting

For microscopy experiments we use a buffer containing an anti-bleaching cocktail [206, 207] (see appendix G). This buffer contains an oxygen scavenging system [208] in combination with a triplet-state quenching system [209, 210] which increases fluorophore lifetime

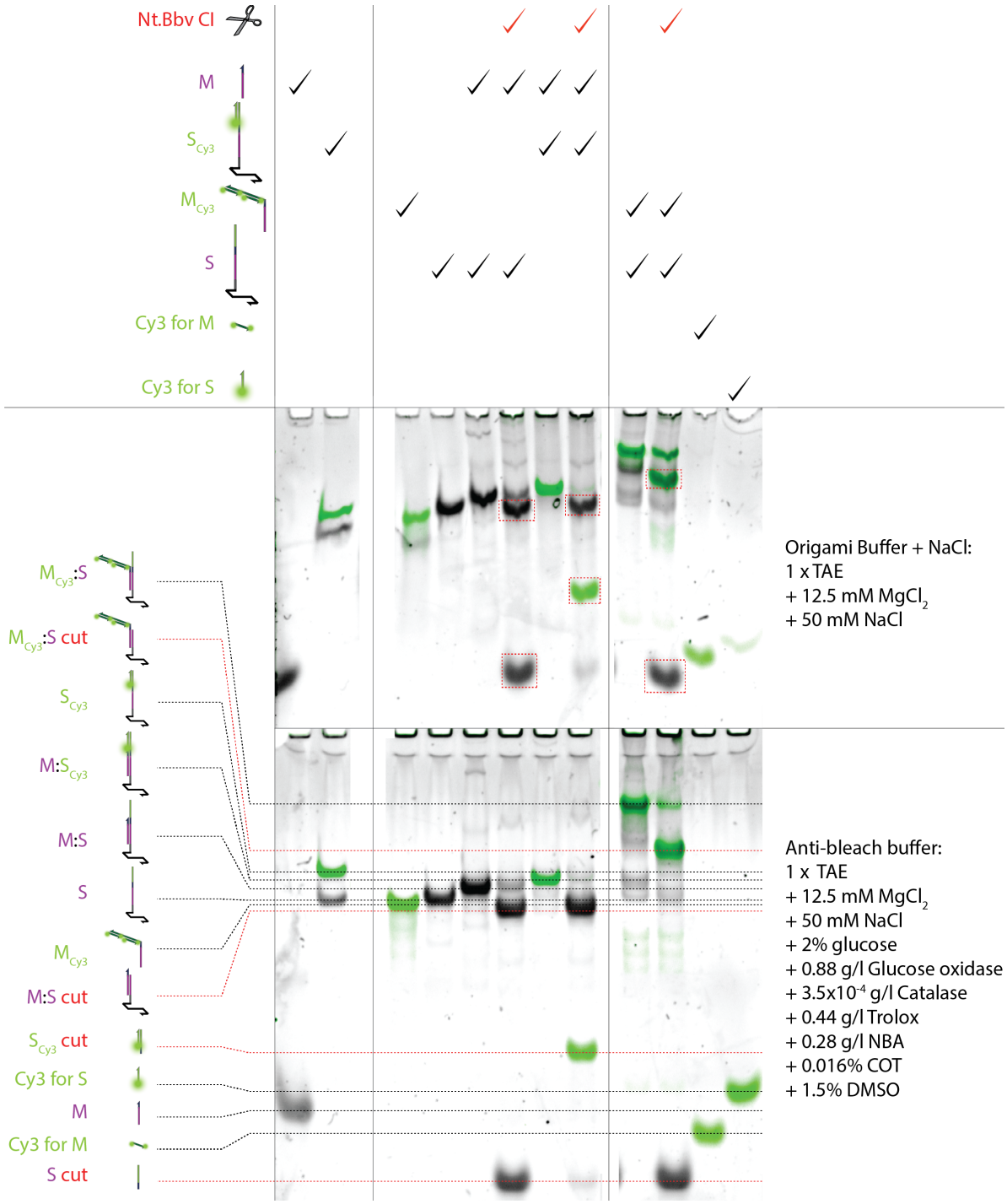

Figure 6.6: PAGE verification of stator cleavage with different motor types, stator types, and buffers. Black=Sybr gold stain, Green=Cy3 fluorescence. Complexes were annealed, mixed and then incubated at 37°C for 1 hr before addition of 0.01% SDS and running on gel. Two gels contain the same samples in different buffers: Origami buffer + NaCl (top), and Anti-bleach buffer (bottom). Gels show the same pattern of bands. All bands can be identified and species are depicted as annotations to the bottom gel. Bands which are only possible after successful nicking with Nt.BbvCI are indicated by red dashed lines or boxes.

by removing reactive oxygen species able to irreversibly damage excited fluorophores, and to rapidly convert 'dark' triplet-states back to ground state for further excitation/emission cycles. The anti-bleach buffer has a different composition from the recommended buffer for Nt.BbvCI. We needed to ensure that the enzyme is active in this buffer, and that the enzyme is able to cleave away the stator tops when they have additional bulk in the vicinity of cutting. Cutting tests were performed in two buffers: Origami buffer + 50 mM NaCl, and Anti-bleach buffer. Complexes were assembled with a motor:stator stoichiometry of 1.2:1 and 0.2 U/μl of Nt.BbvCI where present. After mixing, all samples were incubated for 1 hr at 37°C, the enzyme was then denatured by addition of 0.01% SDS prior to electrophoresis.

Results shown in figure 6.6 confirmed Nt.BbvCI activity in both buffers. Nt.BbvCI in Anti-bleach buffer showed at least as much activity as in Origami buffer + NaCl, with the majority of the substrate band converted into cleaved product bands after 1 hr at 37°C.

All combinations of motor and stator tested were cleaved in both Origami buffer + 50 mM NaCl and Anti-bleach buffer. This includes M:S, M:S_{Cy3}, and M_{Cy3}:S complexes.

6.2.6 Bulk fluorescence tube-crossing

To determine whether the motor was able to pass between two tube linkages, two tube types A and B were purified then mixed together 1:1. A monomeric tube mixture was prepared with poly-T staples on both edges of both A and B tubes. AB dimer tubes were prepared by adding a 10× excess of asymmetric linker staples with 6 nt overlap domain to A tubes with poly-T staples on the left hand edge, and B tubes with poly-T staples on the right hand edge. Linker staples could then stick the right hand edge of A to the left hand edge of B as shown in figure 6.2. The complete track of 26 stators for the burnt-bridges motor is split across tubes A and B with the first 21 stators on tube A (S_A1-S_A21) and the final 5 stators on tube B (S_B1-S_B5)[2].

[2] S_A21 and S_B1 will be separated by 13.6 nm , longer than the usual stator separation of 6.2 nm.

Loading the motor at the start of an A tube

The motor was loaded onto A tubes using a similar method to reference 108 whereby the A tubes were folded without the staple bearing S_A1, then incubated at 37°C for 1 hour with a complex containing S_A1 pre-annealed to a marginally sub-stoichiometric quantity of Quenching motor (ratio S_A1 to M_Q = 1:0.95). The concentration of stator-motor duplex S_A1:M_Q incubated with incomplete A tubes was also sub-stoichiometric, S_A1:M_Q to tube is 0.9:1. This was to reduce the probability of motor sticking in the wrong place, as excess motor in solution could hybridize to arbitrary stators, and excess of stator-motor duplex could lead to non-specific addition of motors to arbitrary stators by strand displacement.

Stator occupation on two separate tubes was monitored

Two stators were labelled with fluorophores to be observed during the experiment. S_A1 had a TET flurophore stator-top probe to show the cutting of the first stator on the A tube, and S_B5 had a Cy5 flurophore stator top probe to show motors arriving at the last stator on the B tube. All other stators had an unlabelled single-stranded stator-top and the stator-tops did not have the undesired secondary structure described in section 6.2.4. In order for the motor to reach S_B5 it had to cross between two tubes.

At the experiment end, a 10× excess of quencher motor M_Q to stators[3] is added to each sample. This should fill any remaining stators (cut or uncut) with motors and quench and stator-top probes which have not been cut away. The end stator S_B5 cannot be cut because of its restriction site mismatch, therefore Cy5 probes on S_B5 will all be quenched after excess M_Q. This gives the fluorescence of fully occupied S_B5.

Data normalisation

Stator state is 0 when a motor is bound to that stator, and 1 when the stator is empty. Data for S_A1 were normalised to zero before the addition of enzyme, where we assume all motors occupy the start stator S_A1[4]; and to 1 at the maximum fluorescence value reached when S_A1 stator tops have been cut away. Data for S_B5 were normalised to zero after

[3]This is a 26× excess of M_Q to each tube type.

[4]Duplex M:S_A1 has at least two extra base pairs over a duplex between motor and another stator.

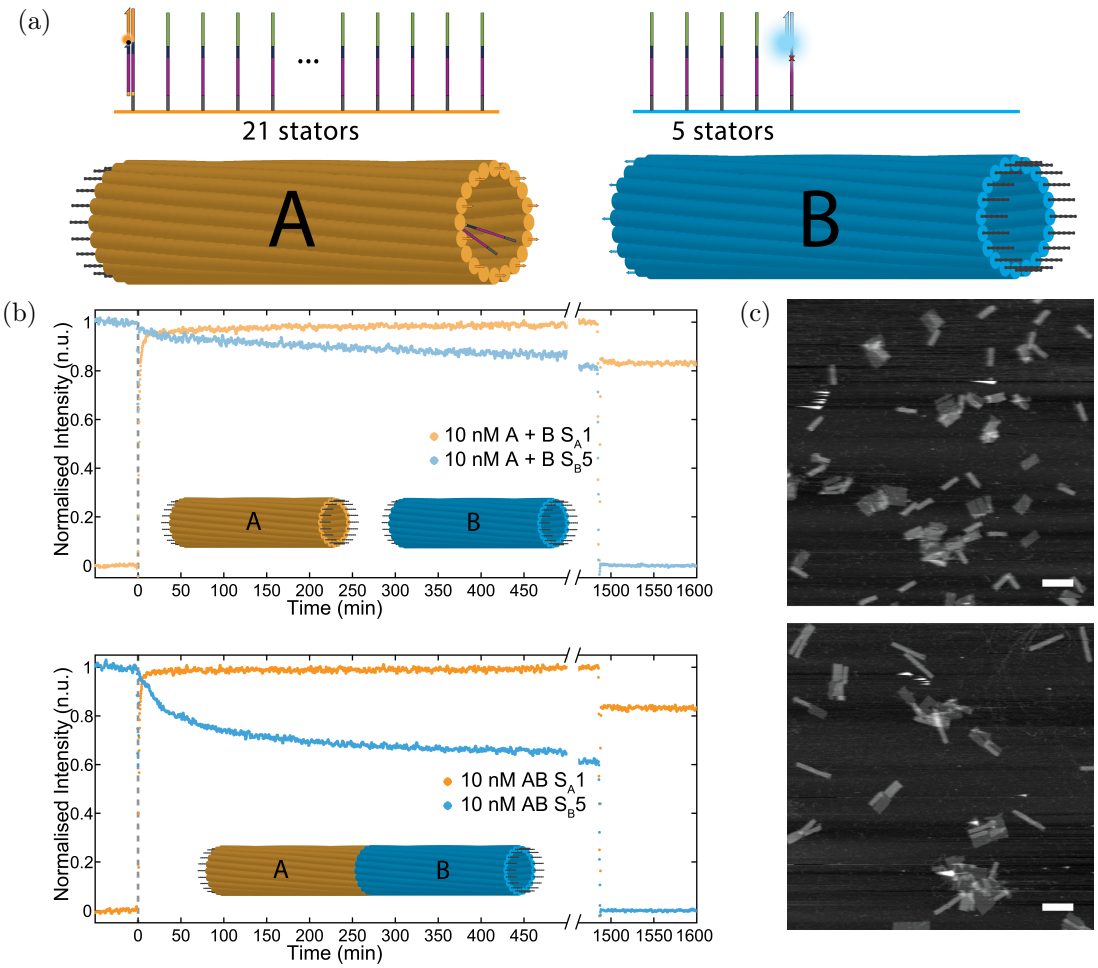

Figure 6.7: Motor crossing tube junctions. **a** Schematic of stator layout on A and B tubes. A tubes have 21 stators along their full length. The first stator S_A1 is loaded with quencher motor M_Q and has a TET stator-top probe. B tubes have only 5 stators on the left hand side of the tube. The final stator S_B5 has a Cy5 stator-top probe and a mismatch in the Nt.BbvCI site. **b** Normalised fluorescence data (0 = no motor occupation, 1 = 100% motor occupation) showing cleavage of stator top of S_A1 (orange) and motor occupation of end stator S_B5 after Nt.BbvCI addition (grey dashed line) for mixture of A+B monomers (top) and AB dimers (bottom) **c** AFM images of mixture of A+B monomers (top) or AB dimers (bottom) after experiment in **b**. Scale bars 200 nm.

saturation of stators with excess M_Q, where all stators are occupied and cannot be cut due to the mismatch in end type stator S_B5; and to 1 before the addition of enzyme where it is assumed all such stators are uncut and unoccupied.

Results

Results are shown in figure 6.7b and confirm more motor reaches end stator S_B5 when the tubes are polymerised, with an estimated 35% of S_B5 stators occupied with a motor after 500 minutes, than when the two tube types are mixed together and not polymerised, where less than 15% of S_B5 stators are occupied with a motor over the same time period.

In both cases the stator top on S_A1 is cut away with similar kinetics.

Unlike in reference 156, we did not have direct experimental access to the kinetics of stepping off the first stator from the fluorophore on S_A1 because the fluorescence signal only reports part of the stepping process, the cleavage and dissociation of the stator top, and not the kinetics of TMSD onto the next stator. We shall call this rate constant k_{cut}. The analytic solution for this irreversible cutting and dissociation is a first-order exponential decay of the concentration of uncut S_A1 with timescale $1/k_{\text{cut}}$, hence the normalised fluorescence response is described by Normalised Intensity$= 1 - e^{-k_{\text{cut}}t}$ and is found to be $k_{\text{cut}} \sim 0.02$ s^{-1}, roughly double the reported stepping rate $k_{\text{step}} \sim 0.009$ s^{-1} [3].

The low, 35%, yield of motor reaching S_B5 in the polymerised case is due to a number of contributing factors. First of all, the sub-stoichiometric concentrations of motor used during loading, which would lead to an upper bound of 85% A tubes with a motor initially. Secondly the polymerisation yield of dimers: as can be seen from AFM images of the dimer tube sample, figure 6.7c, not all tubes form dimers. Aggregated or opened structures, and the occasional monomer are visible in addition to well-formed dimer tubes.

There is also a non-zero occupation of S_B5 at the experiment end when the tubes are not dimerised. One possible explanation for this observation is inter-tube crossing of motor between A and B monomers. This could be from either edge of the A tube, from S_A1 prior to enzyme addition, or from S_A21 after the motor has moved the length of the A tube to the opposite end. When a motor-occupied A tube edge transiently meets a B tube edge bearing S_B1, motor could diffuse between these stators and onto the B tube, and migrate towards end stator S_B5 on that tube. This effect would be reduced in dimerised tubes because B tube stators are no longer exposed to receive a motor as they are at the centre of the dimer.

The monomeric inter-tube crossing rate was similar to the rate observed previously for tiles where 15% of motors crossed between two separate tiles after 500 min [156] with a transfer rate of $k_{\text{inter-tube}} \sim 10^{-6}\text{s}^{-1}$.

To test some of these hypotheses and produce cleaner results, a method to purify away malformed structures such as aggregates and tile-like structures would be desirable. Agarose

gel purification followed by sample concentration would be one such method. The start stator should also be positioned closer to the centre of the tube away from the exposed face, though this may hamper loading of S_A1:M_Q.

6.2.7 TIRF observations of the burnt-bridges motor

Fluorescent motors can be identified by stepwise photobleaching

Long Cy3-labelled motors (see figure 6.5a) were loaded into A tubes as described in section 6.2.6. These A tubes were polymerised with B tubes in a 1:20 A:B ratio before immobilising on a coverslip surface (see appendix D for immobilisation protocol). Neither A nor B tubes had stator top fluorophores, therefore the only source of fluorescence from the tubes should come from the motors, which can carry up to 4 Cy3 fluorophores. These experiments involved the stators with the undesired hairpin secondary structure described in section 6.2.4. TIRF observation of the sample using 532 nm laser excitation showed a number of bright spots (figure 6.8 top right). To confirm that these spots behaved as motors, we captured movies of 2000 frames at a frame rate of 7.4 Hz with 100 ms exposure and 200×EM gain. Despite imaging in an Anti-bleach buffer, these bright spots bleached on the order of seconds. Their bleaching profiles showed discrete stepwise bleaching, in up to three steps of similar size (see plots in figure 6.8). This suggests that the spots do come from motors as the photobleaching profile of a motor would be stepwise bleaching with uniform sized steps, and the number of steps equal to the number of fluorophores on the motor. A fluorescent contaminant would not show this type of bleaching behaviour, and therefore we infer that the spots are from motors. Four steps were not observed despite this being the maximum number of fluorophores on the motor. This may have been because the first bleaching step was rapid and occurred before acquisition began. Gel studies indicate that the 11mer motor top strands are not firmly attached to the motor, for example origami with many such motors leaves a trail in front of it on agarose gel where the tops dissociate from the motor during electrophoresis and subsequently migrate faster (this can be seen in figures 6.14b and 6.15b). Some motors may consequently have 2 fluorophores (or 0), instead of 4.

The short lifetimes of these fluorophores complicates observations over long timescales. A

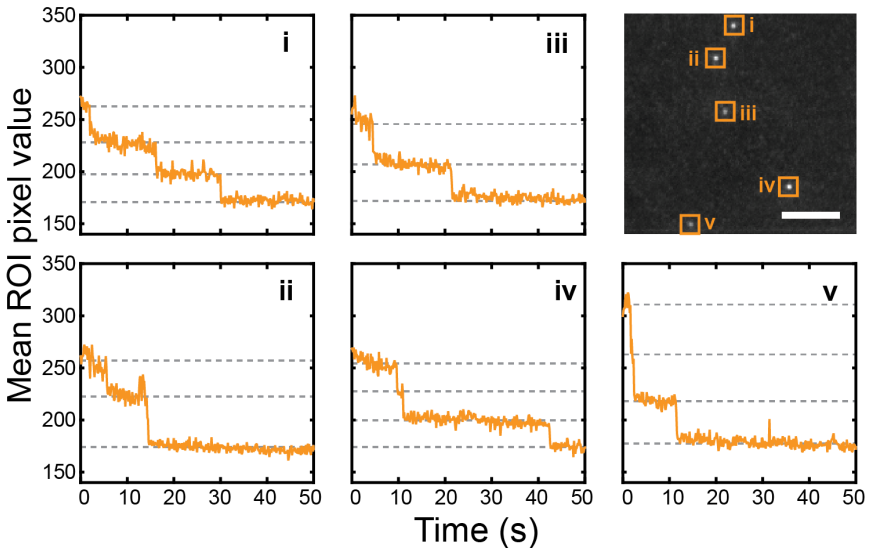

Figure 6.8: Stepwise photobleaching traces of quadruply labelled motors in tubes. Plots show mean pixel values for ROI surrounding 5 representative spots from a single FOV over 50 s. Median projected pixel value for the first 10 seconds is shown in microscope image (top right). Roman numerals match ROIs to photobleaching plots. Fluorophores bleach in discrete uniform steps. A maximum of 3 photobleaching steps are observed per spot, consistent with a motor which can be labelled with up to 4 Cy3 fluorophores. Grey dashed lines show fluorescence levels between which photobleaching events occur. Imaging conditions: 100 ms exposure 200× EM gain. Scale bar 1 μm.

method to increase fluorophore lifetime by reducing laser exposure for the duration of interest for motor observation ($\sim$ hours) would be required to use these fluorescent motors for super-resolution localisation and tracking. Frames could be taken with longer time intervals and shuttering of the laser could be sped up to minimise the amount of time the sample is exposed to laser light. This may be difficult to achieve in practice because of focal drift between sparse time points. Were focussing to be performed in the same channel as observation, any refocussing would involve additional exposure to laser light.

Fluorescent stator tops

An alternative method for motor observation was used to observe a fluorescent track instead of a fluorescent motor. This had the advantage that the extent of the track could be seen in advance of setting the motor off. Tube polymers of 3 or more tubes (360 nm) should no longer appear as diffraction-limited spots, but convolutions of the point spread function with the tube shape, identifiable even in the absence of super-resolution techniques. As designed, we can place fluorophores on stator-tops which should diffuse away once the motor has passed them, leaving a dark wake or asymmetric 'bleaching' of the tube from one

end.

We used this method to attempt to observe similar preparations of AB tubes loaded with motor under three different conditions:

- Positive Control: Buffer added contains 10× excess motor (relative to the quantity of tube stators added to the flow chamber) and enzyme, which should enable the cleavage of all stator tops *without* the requirement for migration of motor along tracks.
- Negative Control: Buffer added contains no enzyme, so stator-tops should not cleave.
- Motor Operation: Buffer added contains enzyme, but no additional motor. The only motors present should be those loaded into the tube polymers via A tubes. Fluroescent stator-top cleavage should be possible only by the movement of motors down the tracks.

A tubes were assembled missing the staple at stator position S_A1, with poly-T left edge and no edge staples along the right edge, with 1:1 stoichiometry of stator to Cy3 stator-top probe. B tubes were assembled with no edge staples along either edge and 1:1 stoichiometry of stator to Cy3 stator-top probe. A tubes were loaded with motor by incubating for 1 hr at 37°C with M:S_A1, at a ratio of tube to motor-stator complex of 0.9:1. M:S_A1 was pre-annealed at a motor:stator ratio of 0.9:1, therefore the maximum achievable loading would be 81%. The stator set used contained the misplaced S2 and stator tops with hairpin secondary structure.

Once prepared, A and B tubes were separately purified with S-300 in Origami Buffer (see appendix A). They were then incubated together in an A:B ratio of 1:20 with a total monomer tube concentration of 20 nM, and a 5× excess of two-domain linkers at 37°C for 1 hr. After this incubation the combined sample was purified again with S-300. We anticipate producing long polymers with a motor loaded A tube followed by a chain of B tubes.

40 μl of sample was divided in half and each half was added to a separate handmade tunnel slide (see appendix D). These were used for the negative control and motor operation samples. The positive control sample was prepared separately, using an identical protocol except Cy3 stator-top probes were added to both A and B *after* the tubes were annealed, at

the same time as the motor was being loaded into A tubes (by incubating for 1 hr at 37°C)[5]. Samples were incubated at 37°C whilst not under microscope observation.

Images were taken of each sample over a period of 8 days. Representitive images of a sample for each of the 8 days are shown in appendix D figure D.2. Unfortunately observations were taken over too long a time interval (between 1-8 days). Using the expected average speed of the motor determined in reference 3 of 0.1 nms^{-1}, the time required to traverse a single tube would be ∼20 minutes, and the time required to traverse a 20mer tube (tube length ∼2 μm) would be ∼8 hours. We therefore expect most of the interesting motor dynamics to have completed within the first day.

Additionally photobleaching of samples was too great to be able to observe movement of motors along tubes, cutting away fluorescent tops. Fortunately we did observe cleaved fluorophore-labelled stator-top probes diffusing which is indicative of motor operation.

'Flashing' of cleaved stator-top probes in TIRFM suggests motor movement

We observed that at first the samples all appeared static, however, in the samples containing enzyme, over time flashing spots started to appear (see figure 6.9), attributable to cleaved Cy3-labelled stator-top fragments diffusing around, into and out of the TIRF observable volume. Videos of samples were collected to compare this effect. For each sample 10 frames with 100 ms exposure, 7.4 Hz frame rate, and 100 EM gain were collected. The first 5 frames are shown in figure 6.9. Videos were subtracted by the median value of a pixel at location xy with time. This should remove features which do not vary with time (grey areas), and highlight spots which are present in that frame, but not in the majority of other frames in white, and spots which are not present in that frame, but are in the majority of other frames in black. Frames and median-subtracted frames are shown in figure 6.9. Spot-detection and ellipse-fitting were performed on these data by Alexandra Lucas using Laplacian of Gaussian (LoG) edge detection (see reference 204).

Results from 10 frames for each of the three samples counted 58, 3 and 595 transient

[5]When stator-top probes are added after assembly they must displace the hairpin secondary structure to bind to stators. If this does not complete during the incubation, remaining unbound probes will be purified away and some stators may be left unlabelled.

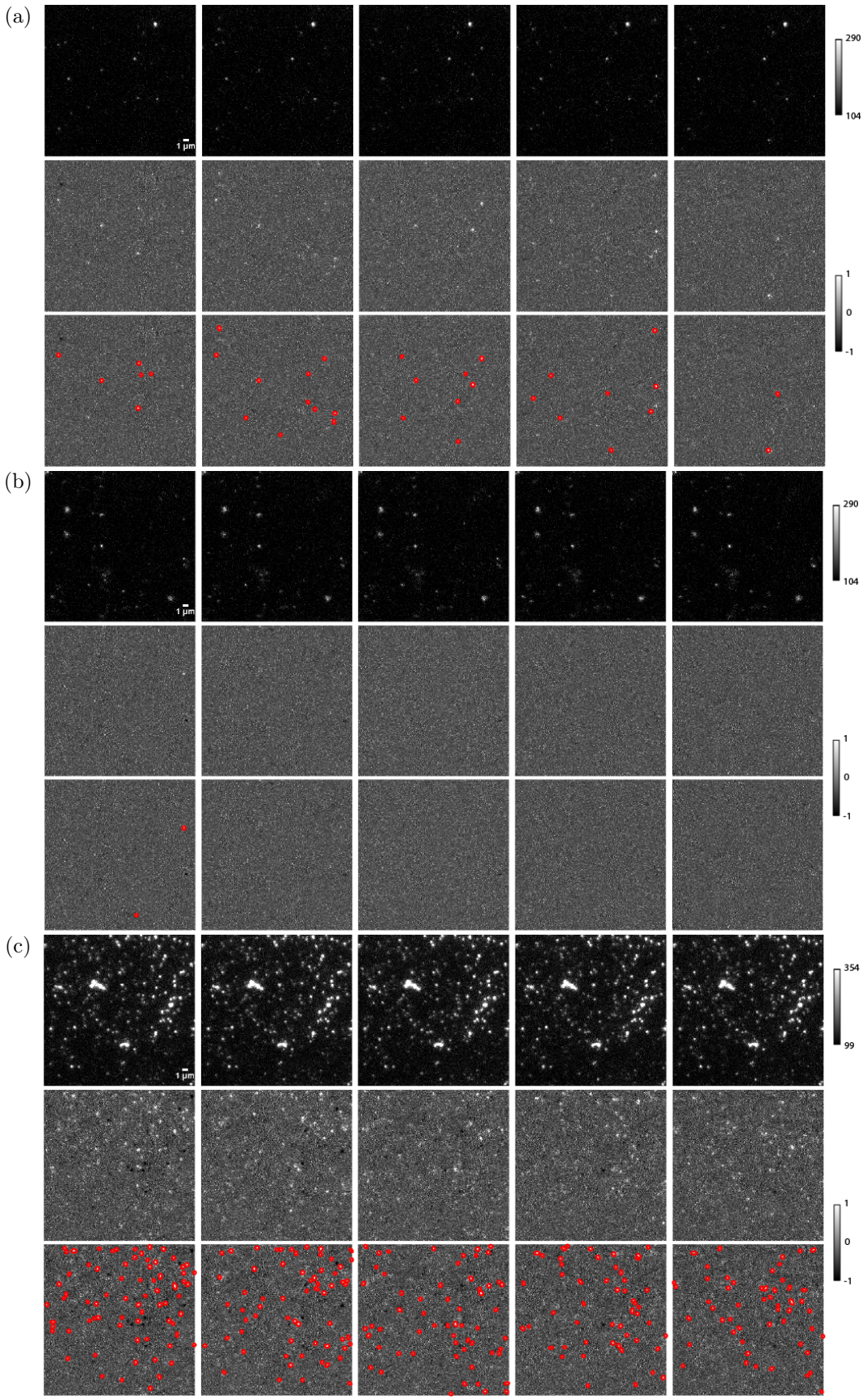

Figure 6.9: Observation of diffusing stator-tops with TIRFM. Top rows of images show consecutive snapshots from **a** motor operation, **b** negative control and **c** positive control samples. Middle row of images show each snapshot subtracted with the median value for each pixel with time and normalised by scaling the lowest 0.1% pixels to -1 and highest 0.1% pixels to 1. Time varying points in the image will appear as black or white regions, while time constant points should appear grey. Bottom row show median subtracted images with spots identified by LoG edge detection and fitted with ellipses (red) by Alexandra Lucas using Wolfram Mathematica. Scale bars 1 μm.

spots in the motor operation, negative control and positive control samples respectively. We see fewer spots in the negative control than in the motor operation sample, despite a higher surface sample density in the negative control sample compared to the motor operation sample (see figure D.2). This would be consistent with the stator-tops being cleaved by the motor as it moves along the tube in the motor operation sample, but no stator-top cleavage and therefore no flashing spots of diffusing stator-tops, when there is no enzyme present.

The positive control sample contained considerably more spots, and these were first observed on day 1 after the addition of excess motor. In the other two samples, the videos analysed in figure 6.9 were collected on day 8, but the positive control sample video was collected on day 1 after excess motor addition. Flashes were clear in the motor operation sample from, at the latest, day 7 onwards, but it is possible that flashes could be observed before this, however no videos were collected for days 1-4 so it is not possible to confirm. We would expect a more rapid generation of freely diffusing cleaved stator-tops in the positive control sample, so the increased number of transient spots and the early detection of such transient spots is in line with our expectations for the positive control. All comparisons of spot numbers should be taken with the caveat that flashing videos were taken on a different day for the positive control sample. Despite this there is a clear increase in diffusing fluorescent species in the motor operation sample compared to the negative control which were collected on the same day. This indicates fluorescent stator-top cleavage is only possible with enzyme, which is consistent with the burnt-bridges mechanism.

6.3 Burnt-Bridges Tracks Across Tile Hetero-Polymers

We have demonstrated that a burnt-bridges motor can pass between DNA origami nanotubes which opens up the possibility of generating more complex tracks for the motor. Complex tracks are required to allow motor-based computation of logical operations [199]. On a standard M13mp18-based origami tile [74] a maximum of 18 stators can be placed along the length of the tile ($\sim$100 nm). This could be extended slightly by changing the aspect ratio of the tile. If multiple distinct tile types could be joined together in a modular fashion, long tracks with multiple junctions could be produced. In this case, unlike the tube case

where we only used two tube types and we produced ABBB polymers with a distribution of lengths, we aim to generate ABCD polymers. Formally polymers are multiple repeats of the same monomeric unit, but in this case each of the subunits differs slightly. We shall refer to them as hetero-polymers, as they are formed form heterogeneous monomers.

One solution to this problem would be to generate a completely unique origami, with completely unique scaffold for every type of tile. The hetero-polymer could then be generated by simple one-pot annealing with suitable staples to join the tiles together. This would be impractical as we commonly use scaffolds of biological origin which therefore must share a large coding proportion of their sequence. Another solution would be to fold tiles with unique staple sets which use permutations of the same scaffold sequence. These origami would need to be folded separately but could be linked together after folding, as the scaffold sequence along each edge to be linked together could be unique.

To also avoid the need for a new staple set for each tile, at the expense of the loss of a unique scaffold sequence at the tile edges, tiles could be folded with different types of edge staples: a pattern of "sticky" and "non-sticky" edge staples which should only interact with a matching partner. It is this scheme which I investigate in the subsequent sections.

6.3.1 Edge adaptations

The base tile used was the minimised-twist tile from reference 87. The simplest edge staples for this tile are called blunt edge staples, which form helices at the edges which have blunt ends. Blunt-ended helices can stack on each other, and in the case of a tile with a row of aligned blunt helices, this can lead to non-specific end-to-end stacking of tiles.

Another flavour of edge staples for this tile include loops of 4T unpaired nucleotides on all of the staples at the edges. This is an attempt to reduce the blunt helical stacking effect.

To join tiles together I formed asymmetric linking staples similar to those used for the nanotubes in section 6.2.2 and figure 6.2. This was done by placing right hand edge of one tile against the left hand edge of the tile and breaking all staple helix crossovers at the edges of the tile and merging the staples across the tiles at the former crossover positions. The

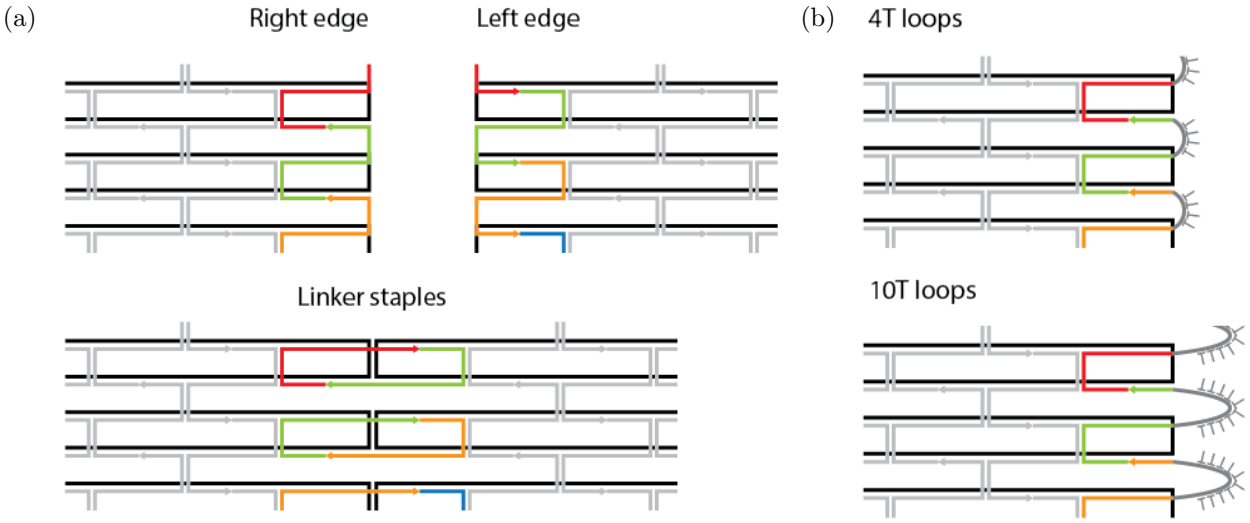

Figure 6.10: Edge staple types for tiles. Scaffold is in black and edge staples in red/green/blue/orange. **a** Linking scheme. Each colour pair of non-polymerising edge staples can be replaced with a like-coloured pair of polymerising staples which form the same base pairs with the scaffold but link pairs of helices across two tiles. **b** Tiles can be prevented from stacking using edge staples with poly-T loops at their terminal crossover. Two lengths of poly-T loop are used, 4T (top) and 10T (bottom).

result was a series of asymmetric staples which are seam-like, (the staple pattern is analogous to the central seam of the origami). As none of the staple end positions had been shifted, all staple ends in the tile still point out of the same tile face. This was an important feature to maintain as it allowed stators to be placed on select edge staples to reduce the distance a motor had to step when crossing tiles.

These sets of edge staples were used to form different sticky edge sequences by using combinations of linking staples and non-linking staples. Pairs of non-linking staples were replaced by matching pairs of linking staples which complete all of the same base-pairs in the origami (see figure 6.10a). There are 24 helices therefore 12 pairs of helices in total. To begin with I divided these into four different edge types (named α-δ) with each edge linking 3 pairs = 6 individual helices. The linked helix pairs were spaced by 6 unlinked helices.

A zig-zag track of burnt-bridge stators was added to the tiles, which would cross between tiles if the tiles were joined with their edges flush. I also added "dead-end" paths where if a motor would keep moving along stators in a straight line rather than zig-zagging, it would reach a dead-end where it could not pass onto the next tile. An individual tile would therefore have track stators arranged in what is almost a Greek letter λ arrangement. When

a stator at the tile edge was required to be part of a linker staple, it was added to the 5' end of that staple, and when an edge stator was part of a non-linker staple, that staple was a blunt type staple, i.e. it did not have any poly-T loop at the edge crossover.

6.3.2 Tile polymerisation attempts

Initially I attempted to fold 5 different tiles which can form a 5-mer as depicted in figure 6.11a. Tiles were separately annealed with edge staples present. The starting tile left edge, and ending tile right edge, had 4T loops to prevent stacking. All linking tile edges contained 3 pairs of linkable helices and helix ends which were non-linking were left blunt. Tile types were 4T:α, α:β, β:γ, γ:δ, δ:4T. The right hand γ-linkage contained a linking stator (stator S1), as did the left hand δ-linkage, (stator S18). Tiles were purified by S-300 columns equilibrated in 1×TAE + 6 mM $MgCl_2$. Purified tiles were mixed in different combinations before incubating for 24 hrs at 20°C and running on agarose gel. Results are shown in figure 6.11b. There is some evidence of a dimer band even in the single tile sample (lane 1), however the dimer band is stronger for the pair of consecutive tiles (lane 2) and there is a hint of a trimer band. Three consecutive tiles produce a stronger trimer band (lane 3), and a small amount of tetramer band. The addition of a 4th and 5th tile (lanes 4 and 5) did not increase the intensity of the trimer or tetramer band beyond that observed for three tiles. Mixing sets of two (lane 6) or three (lane 7) non-consecutive tiles looked similar to mixing sets of two and three consecutive tiles. Increasing the concentration of $MgCl_2$ to 10 mM when mixing the tiles together had no effect on the appearance in a gel (data not shown).

Non-specific interactions dominate with blunt helices between linkable helices

These results indicate that the polymerisation here is non-specific, most likely driven by stacking interactions rather than designed hybridization interactions, because pairs of non-sequential tiles are indistinguishable on a gel from pairs of sequential tiles. The 4T loops on the start tile are clearly having some effect, as the lane containing only this tile shows only monomer and dimer band. I hypothesise that the dimer forms between two such tiles along the linking edge, with one tile flipped relative to the other in order to do so. It may be that the 4T loops are sufficient to prevent stacking but that the non-specific stacking

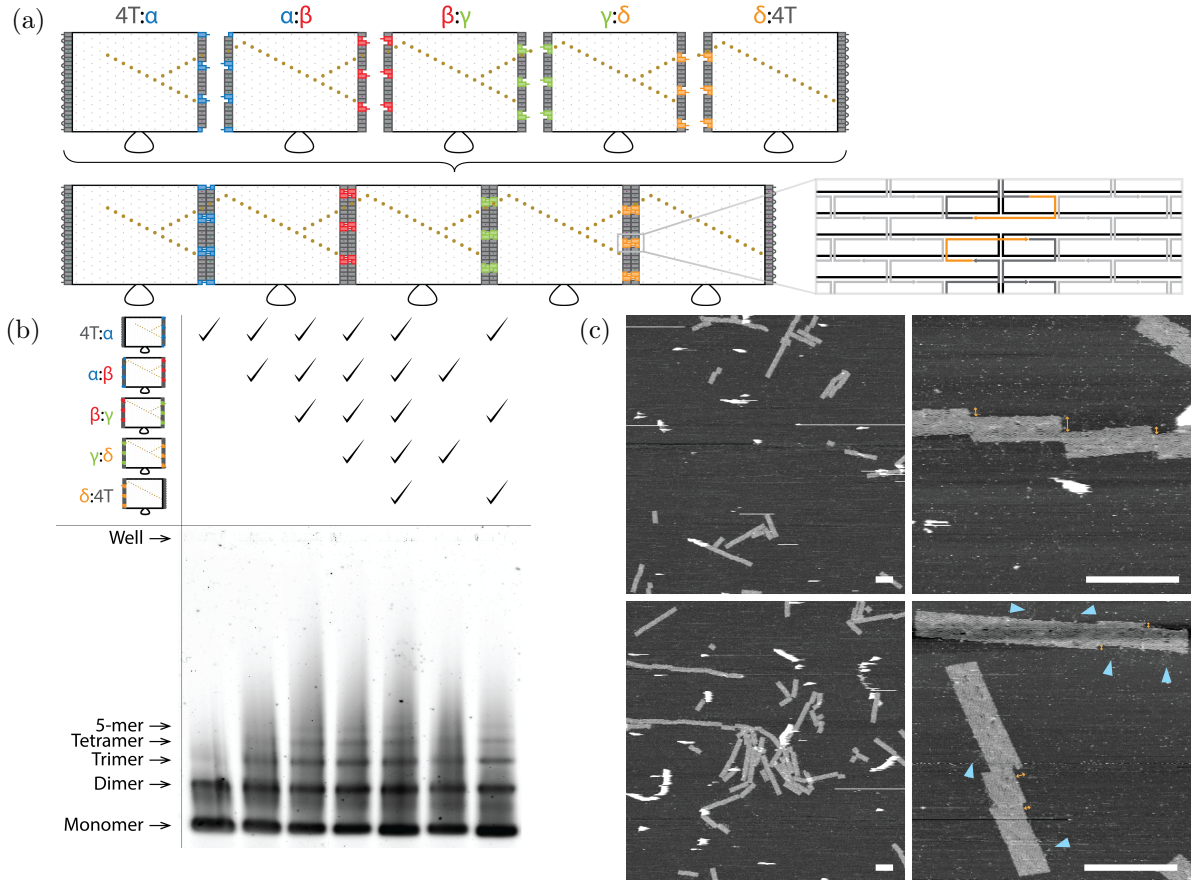

Figure 6.11: 5 tile hetero-polymer attempts with linking edges of 6 linking and 18 blunt helices. **a** Schematic of linking scheme and stators on tile. Stators are indicated by brown circles and the track forms a forked zig-zag arrangement. Linking helices are shown by coloured blocks where the colour corresponds to the link type α-δ, and non-linking helices are shown with grey blocks. Start and end tile edges have 4T loops. A zoomed view of part of an δ-linkage is shown. **b** Agarose gel of polymer attempts after mixing different subsets of tiles with equal stoichiometry. Multimer bands up to 5-mer in length can be seen. Mulitmers are observed in lanes where non-adjacent tiles are mixed indicating non-specific interactions. **c** AFM images of the polymer attempt with all 5 tile types. A distribution of polymer lengths are observed. Some interfaces do not join tiles in a straight polymer, but with offsets along the short tile edge indicated by orange arrows. Blue markers point out the scaffold loops which should be on one side of the polymer if linked correctly, however the polymers show loops on both sides indicating that some tiles are rotated relative to each other.

interactions from 18 blunt helices per tile are too strong for the designed hybridization to dominate polymer formation. Non-specific stacking is likely to be increased by the fact that I have chosen a uniform spacing between linkable helices [94], this means that along any linking edge, there will be the same pattern of 6 adjacent blunt helices, then a gap of 2 helices, then 6 more adjacent blunt helices, then another gap of 2 etc.

To confirm the hypothesis of non-specific stacking I imaged the 5 tile type polymer attempt with AFM. Representative images are shown in figure 6.11c. Here it is clear that many long polymers form, but the polymers are not always connected in straight lines,

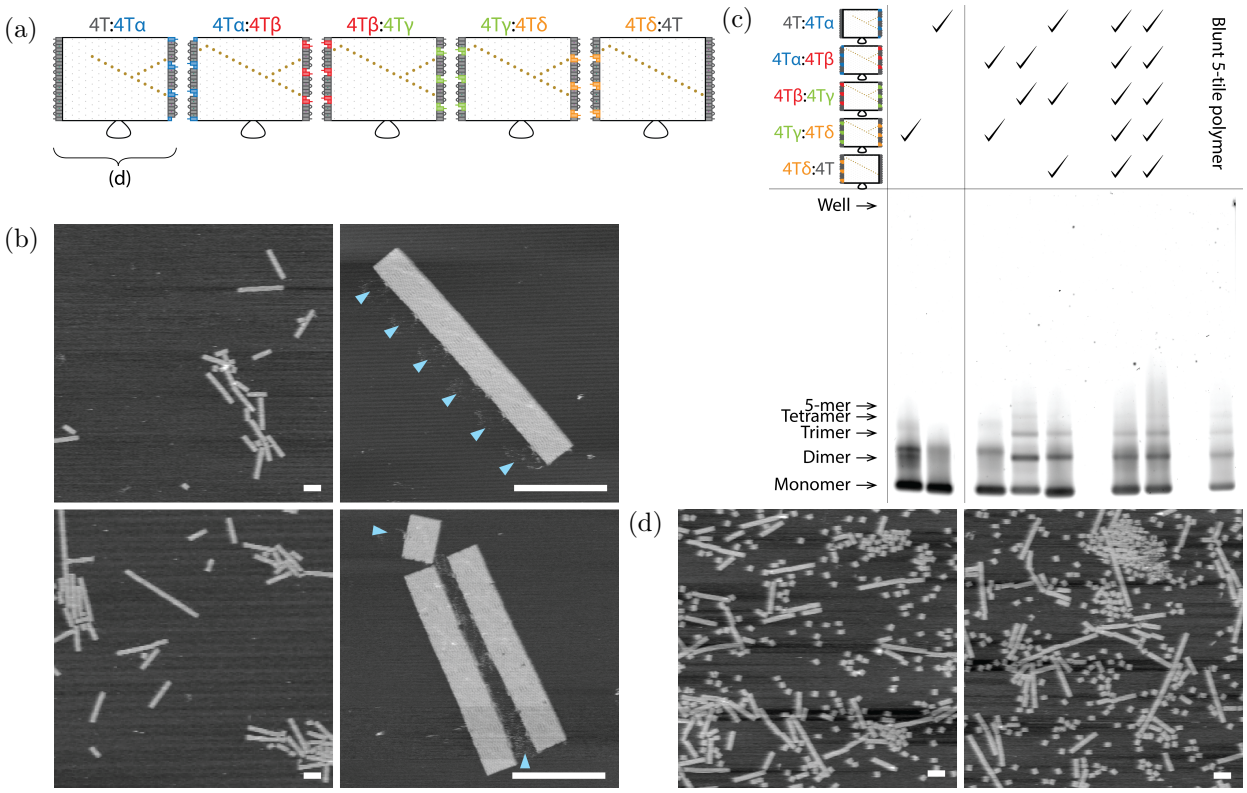

Figure 6.12: 5 tile hetero-polymer attempts with edges of 6 linking helices and 18 helices with 4T loops. **a** Schematic of tiles used. Linking helices are shown by coloured blocks, and non-linking helices by grey blocks. Start and end tile edges have 4T loops. **b** AFM images of the polymer attempt using 5 tiles with edge types 4Tα-4Tδ. A distribution of polymer lengths are observed. All linking interfaces join tiles in a straight polymer. Blue markers indicate scaffold loops (see **a**) which should be on one side of the polymer if linked correctly. This is always the case in observed polymers, indicating that within a polymer all tiles are likely to be the same way round. **c** Agarose gel showing polymer attempts after mixing subsets of tiles with equal stoichiometry. Distinct multimer bands up to 5-mer in length are seen. The final lane shows a 5-tile polymer attempt of the scheme shown in figure 6.11 with blunt helices in-between linking ones for comparison. **d** AFM of 4T:4Tα monomer tile. Sample contains monomers and long polymers. Scale bars 200 nm.

sometimes the short edges of the tile are offset from each other. Tiles would have to be fully aligned for linking staple domains to be correctly hybridized. Additional evidence that non-specific interactions are a problem is that polymer chains of >5 tiles commonly form.

Non-specific interactions persist with 4T loops on non-linking edge staples

In an attempt to reduce stacking along linkable edges I used edge staples with 4T loops on linking edges for any helices not involved in the linking interactions, with the exception of those bearing stators (stator-bearing non-linking edge staples remained blunt). These edge types are named 4Tα, 4Tβ, 4Tγ and 4Tδ.

Results are shown in figure 6.12c. The two single-tile lanes now show only a monomer

band and an upper scaffold band[6], rather than the tile dimer band which migrates faster than the upper scaffold band. This suggests that monomers are not stacking significantly.

When using 5 tiles with linking edge pairs 4Tα-4Tδ, only monomer band and upper scaffold band are observed suggesting that these tiles do not stick together as intended. When two consecutive middle tiles are mixed then a clear dimer band is observed consistent with the desired interaction between the tiles. A trimer band is also seen which indicates an undesired interaction. When three non-adjacent tiles (start middle and end) are mixed the result shows both dimer and trimer bands, though both are lower in intensity than the same bands for two adjacent middle tiles, indicating non-specific interactions. When all 5 tiles are mixed, then the same dimer and trimer bands are observed but nothing higher order.

The conclusion from these results is that 4T loops are insufficient to prevent undesired interactions between tiles. Promisingly monomer quality appears improved when the 4T loops are used, indicating that the non-specific interactions between monomers of the same type have been reduced. There are also signs of specificity where previously there were none, as adjacent middle tiles produce higher order bands than non-adjacent middle tiles.

Similar 4Tα-4Tδ samples were imaged using AFM, however the sample was purified with PEG precipitation in a solution containing 1× TAE + 12.5 mM $MgCl_2$ + 505 mM NaCl and 7.5% PEG 8000, re-suspending in 1× TAE + 12.5 mM $MgCl_2$, rather than by size exclusion S-300 resin equilibrated in 1× TAE + 6 mM $MgCl_2$. Purifying by column transfers your sample into the column buffer so the final magnesium ion concentration in the sample used for AFM is higher than that in the sample used in the gel. This will have the effect of stabilising stacking interactions [93]. AFM results are shown in figure 6.12b. Here, in contrast to figure 6.11c, all polymer edges were aligned and no short edge offset was observed. This is a marked improvement from the previous case as it suggests that one form of off-target interaction, that between offset tiles, presumably due to blunt helical stacking, has been reduced.

One of the monomer tiles used to form the hetero-polymer was also imaged alone to check the monomer quality. AFM images of this sample are shown in figure 6.12d. The majority

[6]Perhaps dimer or double-stranded DNA, it is also present in scaffold only samples (not shown).

of tiles were monomeric consistent with gel results of figure 6.12c, however long tile polymers were also observed. There is still some non-specific polymerisation when 4T loops are used.

Omitting non-linking edge staples on one edge creates mostly homo-dimers

In an attempt to improve monomer quality, the monomer tile from figure 6.12d, was adapted to include no staples on the left hand edge where previously this non-linking edge had been filled with staples with 4T loops. The hypothesis behind this is that undesired polymerisation may be due to stacking, in which case 16 nt scaffold loops, which are longer than 4T staple loops, will cause greater steric hindrance to stacking interactions than relatively short 4T loops. Alternatively the polymerisation might be due to linking staple sticky ends on the right hand side of the tile sticking to their complementary domains on the left of the tile. If this is the mechanism, removal of edge staples may even promote interactions as there is now no edge staple competing for the complementary domain to the sticky end.

AFM of the sample without left edge staples is shown in figure 6.13a. In contrast to figure 6.12d, far fewer long polymers are observed, but proportionally more dimer tiles are observed. Pairs of tiles in a dimer always have scaffold loops on the same side, consistent with the tiles in a dimer having the same orientation. If the stators had been visible in these images I could have confirmed whether tiles had the same orientation or not. I can, however, identify a 4Tα edge on the outer edge of one of the dimers where I can also identify the scaffold loops. Assuming pairs of scaffold loop edges cannot interact, this dimer must be of the form shown in figure 6.13b. The dimer is of two similarly oriented tiles and the link between the pair is between the left hand edge of one and the right hand edge of the other (bottom left of figure 6.13b). Such an interaction may involve hybridization of linker staples between tiles as the domains are correctly aligned.

Undesired polymerisation of monomers is due to linker staples

The previous results led me to hypothesise that monomer polymerisation was possible only in the presence of the linker staples, and therefore due to hybridization of linker staples across edges, even when only one of a pair of linker staples was present, and even when the linker staple sticky end was competing with an edge staple for binding to this domain. To

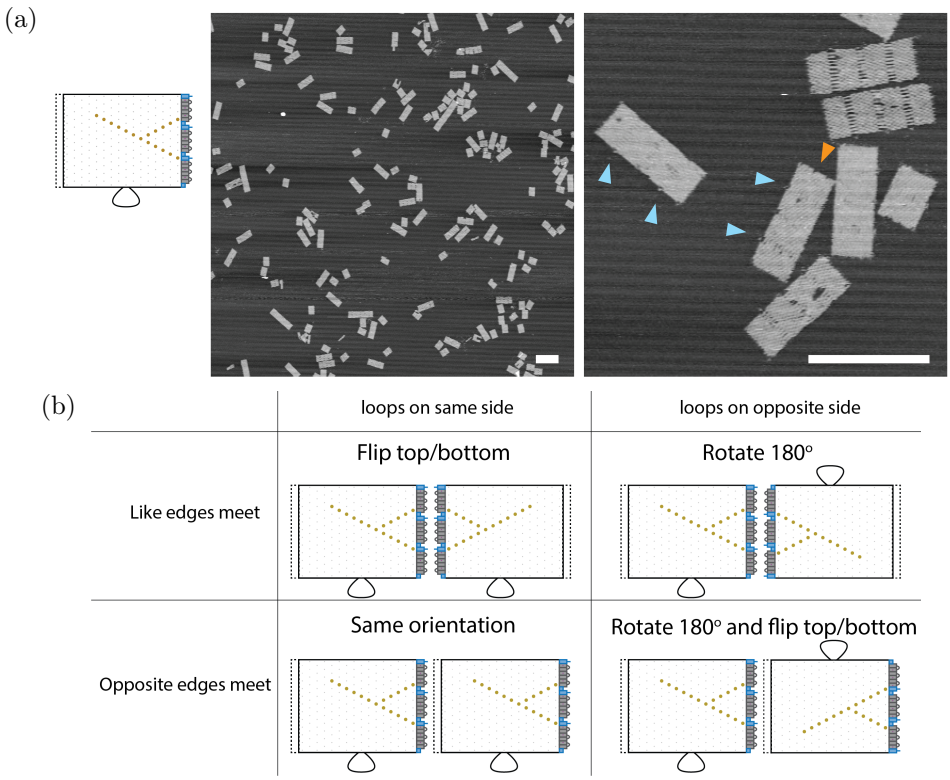

Figure 6.13: **a** AFM image of monomer tile (left) prepared identically to the sample shown in figure 6.12d but with no left hand edge staples used. Instead of forming long polymers, far fewer long polymers are observed, but many dimers. Blue markers indicate scaffold loop positions which are consistently on the same side of the dimer. Orange marker indicates a 4Tα-edge which can be identified. Scale bars 200 nm. **b** Table depicting the possible orientations of tiles in a dimer. Tiles can be flipped upside down, and/or rotated relative to one another to allow different edges to meet. All possible dimer configurations are uniquely identifiable if the stators can be clearly seen. For edge hybridization interactions to occur tiles must have scaffold loops on the same side.

test this I assembled tiles with 10T loop end staples leaving gaps where a linker staple would be required, and purified these by PEG precipitation. Purified tiles were mixed in different combinations with and without the full set of linker staples and run on an agarose gel.

Results are shown in figure 6.14b. There is a general pattern where right hand lane of each pair of lanes, containing linker staples, shows higher order polymers than the left hand lane, without linker staples. Single tiles types, particularly the first three tiles, show a band slightly higher than the position of the dimer band. This was also observed in figure 6.12c. Assuming that this is not true dimer, we see the appearance of true dimer and trimer, and even tetramer, in lanes where linker staples have been added even when these linker staples should not be able to link due to the choice of gap position. This indicates that linker staples are likely to be the cause of monomer type tiles forming polymers.

This method is not a good method for producing hetero-polymers as it seems no better at

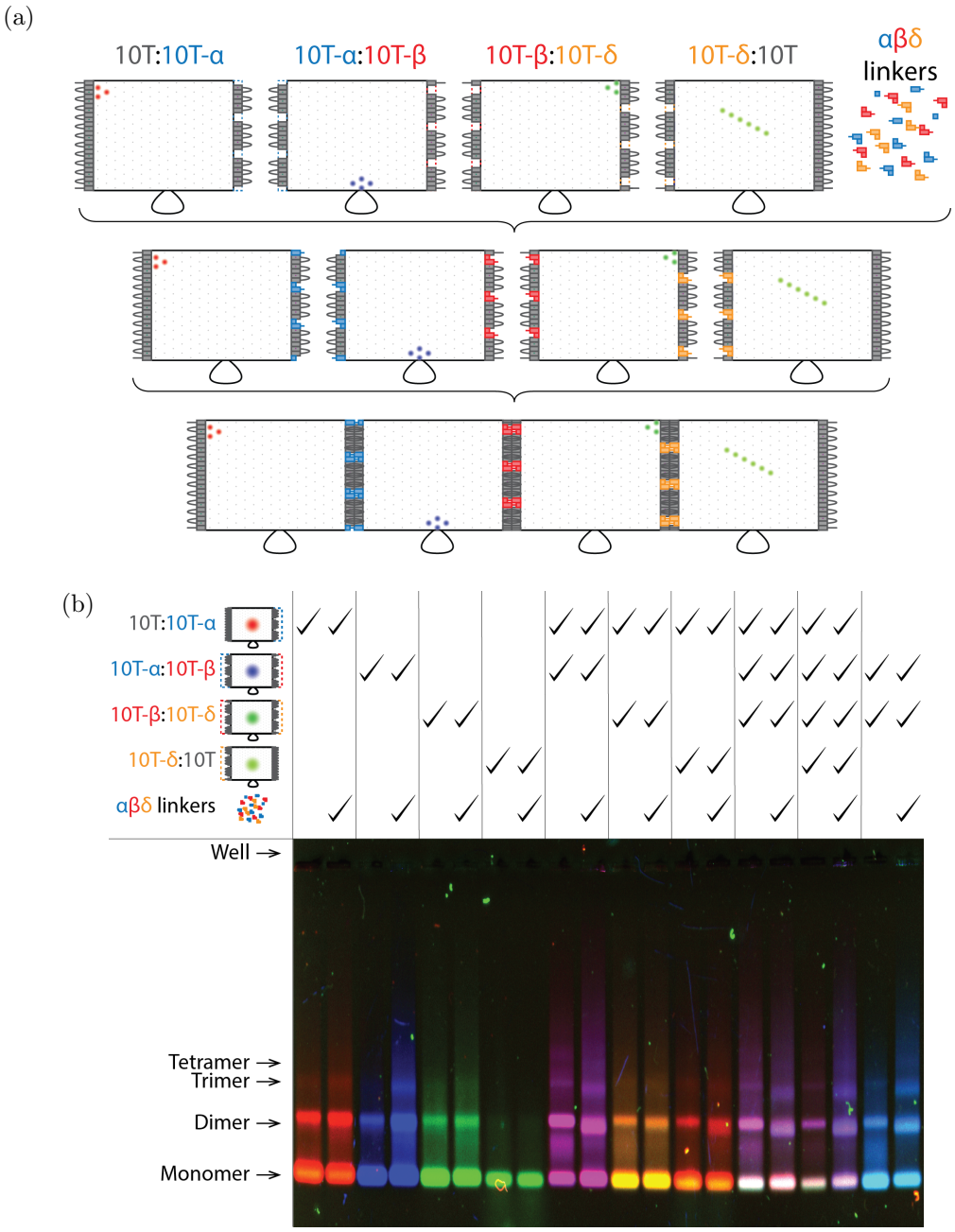

Figure 6.14: AGE of 4 tile hetero-polymer attempts of tiles assembled with linking edges with 10T loops with gaps for linking staples α, β and δ, mixed with 10× excess of linking staples α, β and δ. **a** Schematic of linking scheme and fluorophore markers on tile. Tiles are assembled with 10T loop edge staples leaving gaps corresponding to a certain link type. These tiles, once purified, are mixed together with a 10× excess of linker staples which should fill the gaps and allow polymers to form. 6 × Alexa647 fluorophores are attached to tile 1, 8 × Alexa488 are attached to tile 2, 6 × Alexa594 are attached to tile 3 and 6 × Cy3 are attached to tile 4 via long Cy3 motors on stators (figure 6.5a). **b** RGB fluorescence channel overlay for agarose gel showing polymer attempts after mixing different subsets of fluorophore-labelled component tiles with approximately equal stoichiometry, with and without addition of 10× excess of α, β and δ linker staples. Samples were incubated for 24 hrs at room temperature. Distinct multimer bands up to 4-mer in length can be seen. Multimers can form in the presence of linkers even when the tiles mixed together are not consecutive (e.g. in tiles of a single type).

producing the desired multimers than it is at producing undesired multimers from monomers, and even after 24 hrs very low quantities of multimer can be seen. After a shorter incubation time of 3 hrs, there is no difference between samples with and without linkers.

6.3.3 Assembly with more linked helices on an edge

I attempted to use linkable edges with more links between tiles, each have 8 rather than 6 helices joined, and therefore 16 rather than 18 non-linking helices which should both increase the strength of hybridization between tiles. This produces three edge types named ε, ζ, η.

4 tile types were each assembled with edges 10Tε-10Tη, where non-linking helices have 10T loops at the edge. After assembly and purification the tiles were mixed together and left for 30 hrs at room temperature before running on a gel. Tiles were each labelled with fluorophore markers to help identify the composition of any bands. The first three tiles were labelled with staple extensions which hybridize 18mer oligos labelled with Alexa647, Alexa488 and Alexa 594 respectively. The label used for the 4^{th} tile was a long Cy3 motor (see figure 6.5a) on stators and the Cy3 label, a 11 nt oligo, appears to fall off the origami as noted in section 6.2.7 as a trail of green fluorescence in front of the origami band is observed in lane 4. The gel was scanned in three different channels: Red = Alexa647, Green = Alexa594[7], Blue = Alexa488 (see table B.1 for lasers and filters), and overlaid.

Results from figure 6.15b show that some of the designed edge interactions can be identified, but there are also undesired bands observed. For example lane 5 contains the first two tiles of the polymer which are labelled with red and blue fluorophores respectively. After mixing and incubating the majority band formed is a magenta coloured dimer, suggesting the two tiles have stuck together as intended. The remaining monomer band is mostly red which may indicate the stoichiometry of tiles was not exactly 1:1, but that the red tile was in slight excess. Lanes 6, 7, and 10 are pairs of non-consecutive tiles and all show some undesired bands of dimer and above, however the majority band is monomeric tile.

When attempting to form chains of three consecutive tiles, lanes 11 and 14, monomer band intensity is reduced and there is a clear trimer band. Other lanes also display a trimer

[7]Cy3 fluorescence should also be visible in this channel.

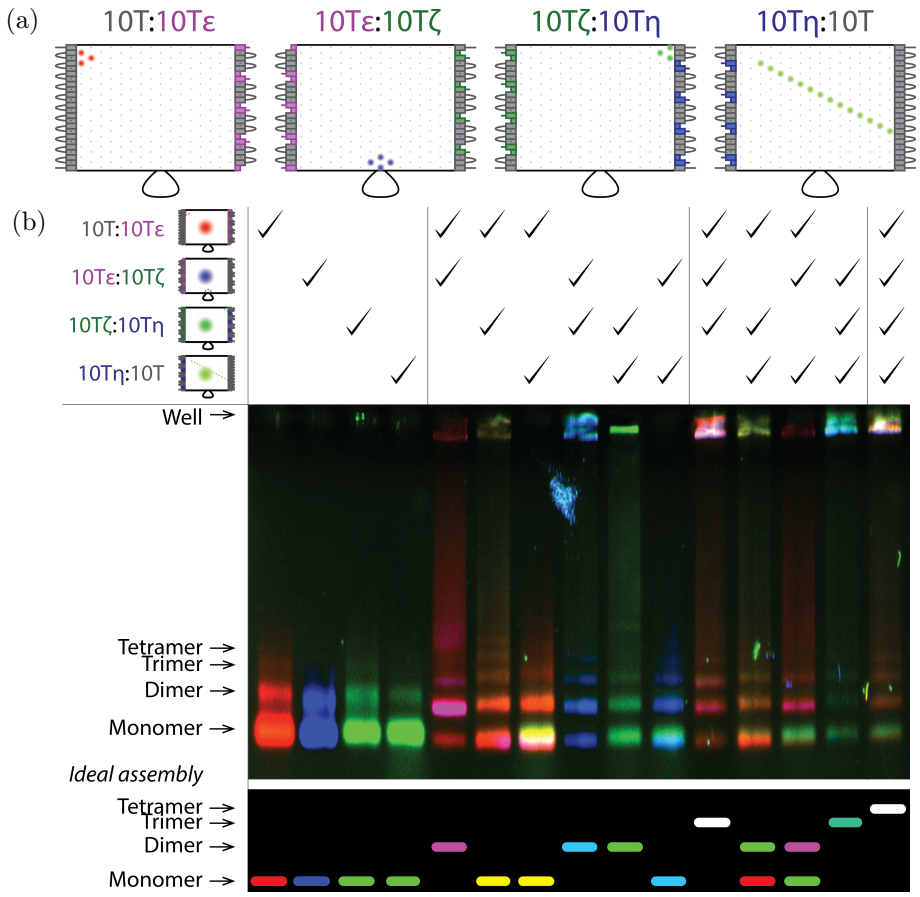

Figure 6.15: AGE of 4 tile hetero-polymer attempts of tiles assembled with linking edges with 10T loops and edge types ε, ζ and η, where 4 pairs of linking helices are spaced by 6 non-linking helices with 10T edge loops. **a** Schematic of tiles used and fluorophore markers on tile. Tiles are assembled with the appropriate edge staples, purified, mixed together, and incubated for 30 hrs. 6 × Alexa647 fluorophores are attached to tile 1, 8 × Alexa488 are attached to tile 2, 6 × Alexa594 are attached to tile 3 and 15 × Cy3 are attached to tile 4 via long Cy3 motors on stators (figure 6.5a). **b** RGB fluorescence channel overlay of agarose gel showing polymer attempts after mixing different subsets of component tiles, with and without addition of 10× excess of α, β, and δ linker staples. Samples were incubated for 30 hrs at room temperature. Distinct multimer bands up to 4-mer in length can be seen. Band positions and colours expected from an ideal assembly are shown below the gel.

band but tend to have relatively more monomer and dimer. Similarly when all 4 tiles are mixed and incubated, a tetramer band can be seen and monomer and dimer bands are feint.

Many of the lanes, particularly those containing consecutive tiles, show fluorescence intensity in the well. Running the same, or similar, samples on different gels with the same recipe and running conditions can lead to more or less of the sample remaining in the well. I expect that some of the fluorescence intensity in the well may be due to short multimers, perhaps even 4-mers, which never enter the gel matrix. It is more difficult for long molecules to enter the gel matrix than more compact ones with the same net charge, and orientation of polymers may also play a part in altering electrophoteric mobility [211]. Additionally

long polymers can become tangled together which may prevent them entering the gel. Even if a band corresponding to a certain size polymer can be observed in the gel, it cannot be assumed that all of the polymers of that length will have entered the gel.

Most tetramers have correct tile composition and ordering, and most undesired links are between monomers of the same type

Assembly with edge types 10Tε-10Tη, though clearly not ideal due to many off target products, does seem to produce some multimers as designed. Tiles polymerised in a similar way to these were also adorned with stators where each tile had a unique pattern of stators to demonstrate a continuous, non repeating track across multiple tiles. The differences between stator patterns on each tile also helps the tile types to be identified in AFM images.

Tiles were assembled with edges, purified by PEG precipitation, then mixed together in pairs, the first two tiles in one pot and the last two in another, and incubating pairs for ∼20 hrs before mixing both pairs together. On a 0.5% agarose gel in 1× TAE + 6 mM $MgCl_2$ (not shown), origami mostly remained in the well. However, the sample when imaged by AFM showed many short polymers. All images collected where stators were clearly visible and stator patterns identifiable are presented in figures 6.16c and 6.16d, with the tile type inferred by stator patterning labelled. All tiles where the stator pattern was visible are deposited face down on the mica, i.e. stators are pointing from the bottom face of the tile.

Polymers where stators were visible which were 4 tiles in length all had stator patterns which could be due to the correct tile ordering (figure 6.16c). Only 4 such 4-mers with clear stators were imaged which is too small a sample size to indicate all 4-mers are ordered as intended, but some clearly are. Other multimers imaged with length less than or greater than 4, showed evidence of two main types of linkage, the first being the desired linkages between consecutive tile types, and the second being links between tiles of the same type (figure **??**). Very few linkages between different types of non-consecutive tiles were observed. The pair of tiles in the top left of figure 6.16d shows a 1-2 dimer, one of the tile pairs mixed in the first step of the pairwise assembly. This might be a precursor for a correct 4-mer. The cluster of tiles in the top left of figure 6.16d also shows tiles from the first pair. All tiles appear to be of type 1 or 2, with all links between different tile types 1-2 and none 2-1.

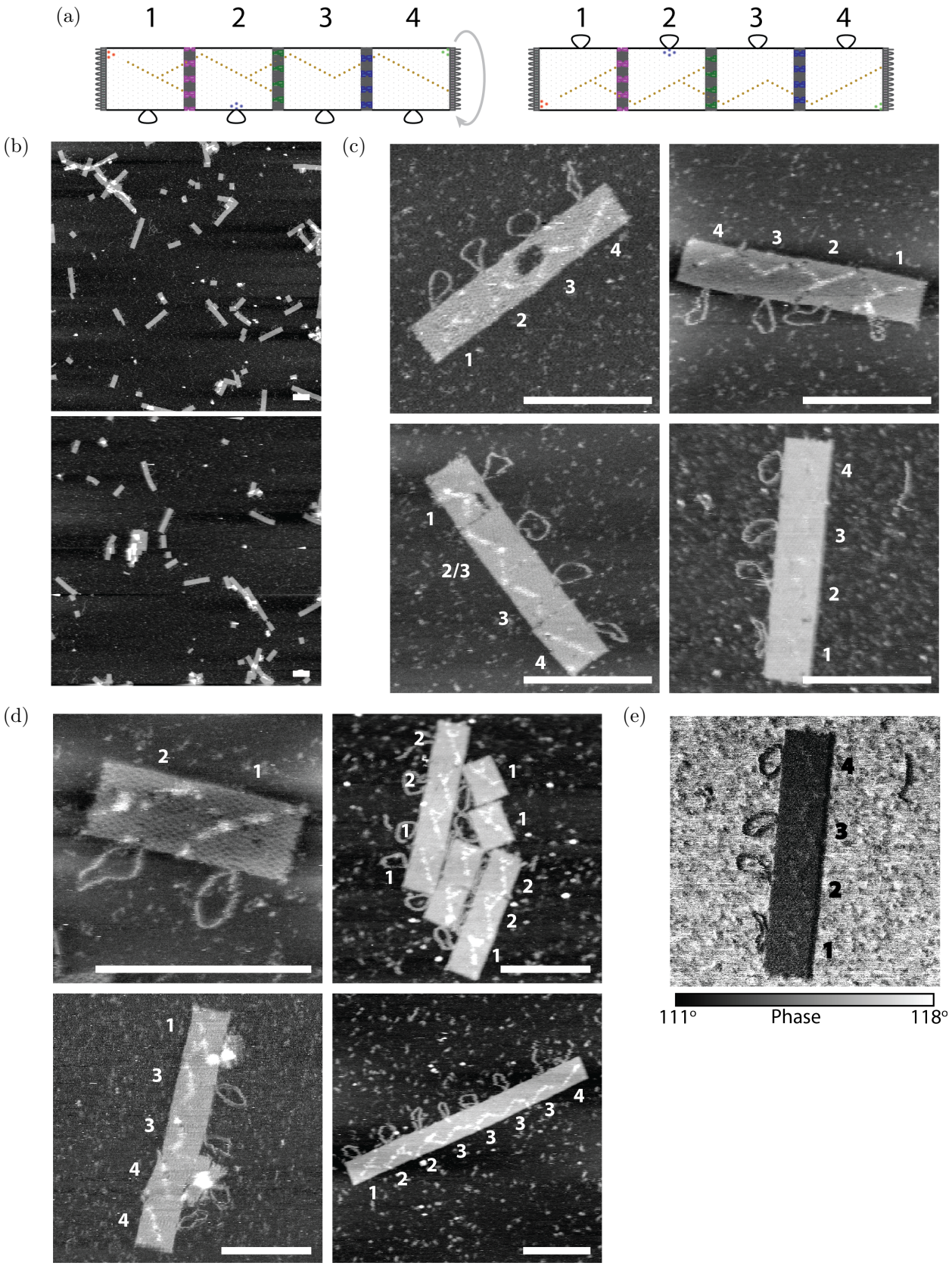

Figure 6.16: AFM images of 4 tile hetero-polymer with edge types 10Tε, 10Tζand 10Tη. When stators are visible, tiles are always stator-side-down, as shown on the right. **a** Schematic of tiles used and stators and fluorophore markers on tile. **b** Large fields of view show many short (∼4-mer) polymers. **c** Polymers with length 4 and visible stators show stator patterns which are consistent with the designed linking of tiles 1-2-3-4. **d** Other polymers with visible stators show links between consecutive tile pairs 1-2, 2-3, 3-4, and links between tiles of the same type 1-1, 2-2, 3-3, 4-4. Few edge links between non-consecutive tiles are seen. **e** The phase image corresponding to bottom right of figure 6.16c where stator pattern is more easily identifiable. Scale bars 200 nm.

Stator patterns are often ambiguous, as stators are small, single-stranded and flexible and therefore sometimes chunks of a pattern may be missing from the image. It is possible that some tile types have been incorrectly identified in these images, for example, a type 2 tile where the straight fork and fluorophore patch close to the loop cannot be seen would look like a type 3 tile.

To better assess tile ordering, each tile should be prepared with a completely different pattern of lumps rather than patterns which are subsets of each other. Given the information above it seems that the correct edges are linking together, however incorrect links between tiles of the same type are still an issue and these may form during the annealing process.

The presence of undesired links between monomers of the same type may be due to the passivating edge staple design, and has probably occured during tile assembly. The poly-T loop in the staple links has a short 8 bp followed by a 16 bp domain on one side of the poly-T loop, and a short 8 bp domain on the other side of the loop. The poly-T loop reduces the effective concentration of this isolated 8 bp domain at its complementary domain on the scaffold, reducing the likelihood of the staple being fully bound, and exposing a scaffold domain which can also bind to the sticky end of the linker staple. If this is the case, using passivating end staples with poly-Ts at the 5' and 3' ends of two-domain staples similar to those used on tubes earlier in the chapter (see figure 6.2 top left) might improve monomer quality and therefore increase subsequent yield of well formed polymers. These staples would also have the advantage of both edge staple domains binding early (at higher temperatures) during annealing as both domains would be the long, 16 bp, type which has a higher melting temperature. This might allow them to seal themselves in place on the origami before the short 8 bp sticky ends of the adajcent tile have a chance to stick.

Length analysis of polymers formed by pairwise assembly of tiles shows most tiles are part of dimers or tetramers

Using large fields of view polymers of different lengths were counted, both by hand, and using MATLAB 2014b. Using the MATLAB image processing toolbox the image is binarised by thresholding filling holes in shapes, any objects smaller than half a tile, or with parts

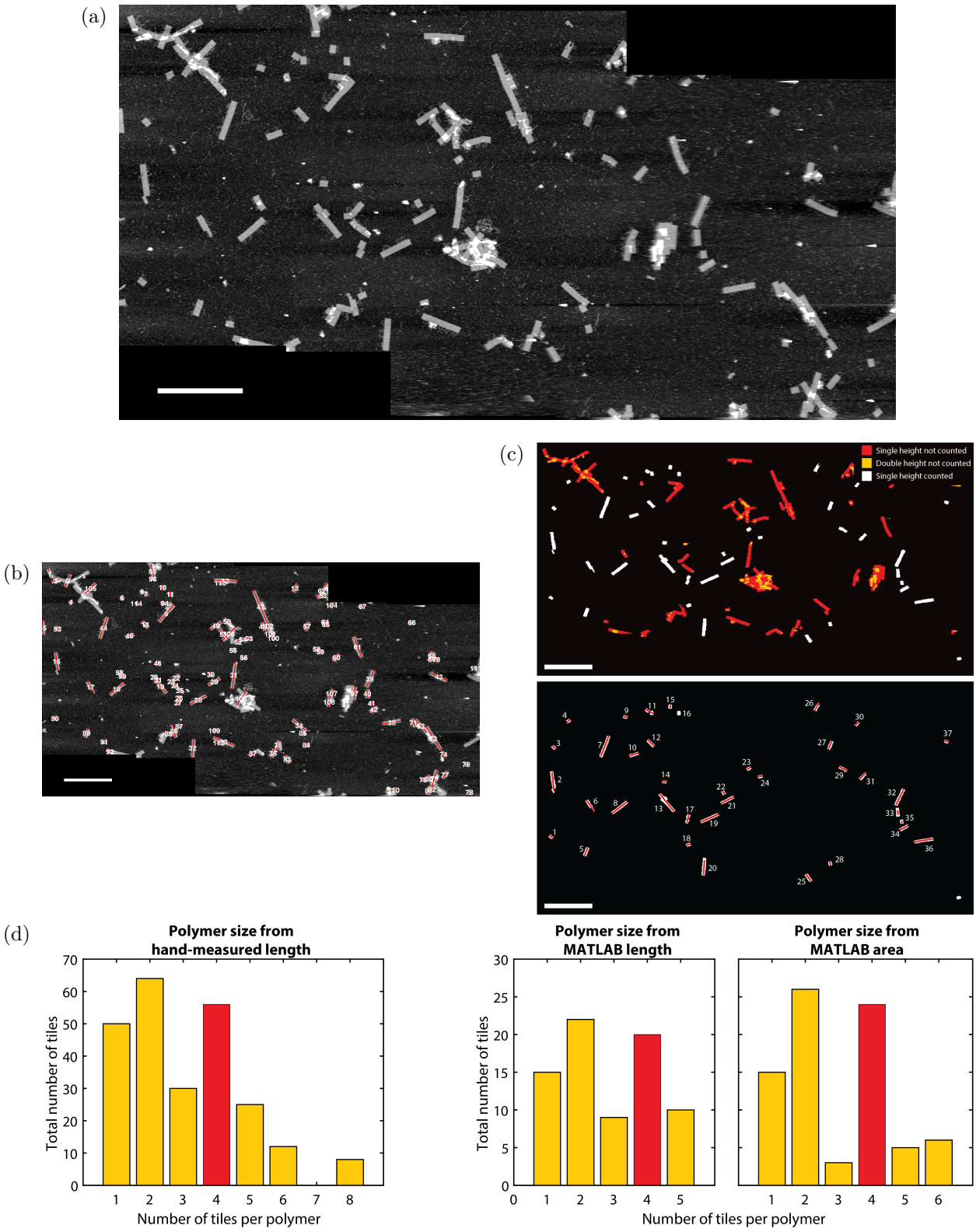

Figure 6.17: Inferring hetero-polymer size from AFM images. **a** Large field of view AFM images stitched together to count tiles in polymers. **b** Same image annotated by hand with red lines along the long axis of any tile polymer. Lengths were hand-measured for 114 tiles or chains of tiles. **c** Same image masked to show polymer tiles which will be analysed in white, and overlapping / aggregated regions which are not analysed in red and yellow (top). Masked image showing only objects analysed and the length markers in red showing length determined by MATLAB script (bottom). **d** Bar plots showing the number of tiles forming part of polymers of each length calculated from hand-measured length (left) or using MATLAB (right) where either polymer length, or masked polymer area were used to determine the number of tiles. Scale bars 1 μm.

higher than 1.5× the height of a flat tile, are not counted. Sometimes polymers can overlap, the hand-counting method includes some of these polymers, where the length is clear and they appear well formed, but the MATLAB counting method excludes them. An ellipse is fitted to the remaining objects in the image, and the major axis is used to define the polymer end-to-end vector. The first and last point within the masked region along the polymer end to end vector are used to calculate the polymer length. Lengths are then binned to give discrete numbers of tiles in a polymer. Alternatively the masked area of the polymer can be used to determine the number of tiles. This is done by calculating the single tile area using the median minor axis length of the polymer-fitted-ellipses as the tile width, and multiplying this by 90 nm, the tile length. By dividing the masked object area by the single tile area and rounding to the nearest integer, an estimate for the number of tiles is found. Out of a total of 37 objects analysed, these two methods produced the same estimated number of tiles in 32 cases. In the cases where they did not agree the reason was usually clear from figure 6.17c. Occasionally area was over-estimated because there was another tile next to it, but not overlapping, which formed a simply-connected mask with the polymer (e.g. 13). Length was also overestimated due to the same effect, especially for pairs of close monomers (e.g. 11). Length was sometimes underestimated (e.g. 20, 16). Despite these differences, the resulting length distributions were similar for all three counting methods. With all methods we see large numbers of monomer and even larger numbers of tiles form dimers. Far fewer tiles form trimers, which is likely due to the pairwise assembly method, precursors to the correctly assembled 4-mer should be dimers assuming polymerisation time scale is shorter than the amount of time the pairs were incubated before mixing together. Large numbers of tiles form 4-mers. The estimated yield of tiles forming 4-mer from the hand-counting method is approximately 18% of all origami deposited on the surface. Fewer tiles form part of 5-mer or longer chains. The 4-mer peak is a obvious deviation from the Flory-Schulz model of polymerisation [212] where each monomer has equal probability of linking to any other monomer. The fraction F of individual tiles which form part of an n-mer using this model is given by the equation $F = n(1-p)^2 p^{n-1}$, where p is the probability of two tiles sticking. Such a function can only have a maximum for $p < 1/4$. The function cannot have

two maxima therefore it cannot explain the measured distribution of polymer lengths in figure 6.17.

Together with the evidence that many 4-mer tracks are composed of correctly ordered monomers, length data demonstrates that I have designed a scheme which will produce a 4-mer burnt-bridges track, however yield is low. In order to take forward such a track for a burnt-bridges motor it would be desirable to either increase the yield, or establish a method for purification of the desired 4-mer product away from other higher and lower order polymers, or both. One method of purification which may allow separation of polymers of different lengths would be size-exclusion chromatography. I have already used size-exclusion chromatography as an origami purification technique, but this used a resin with MWCO far below the molecular weight of an origami structure but above that of an oligo, in order to slow the passage of excess staples through the resin but allow the rapid passage of origami which does not enter the resin. Alternatively size-exclusion chromatography with a much larger MWCO resin could be used such that polymers can enter the resin, but their passage through the resin would be dependent on their size, and given sufficient resolution, different lengths of polymer would be recovered in separate fractions.

6.3.4 DNA PAINT proof-of-concept for track imaging

To answer questions about the behaviour of individual motors on tracks, rather than ensemble averaged results such as those in reference 108, we require single molecule methods of walker observation, with resolution on the length scale of the tracks. Previously this has been achieved with high speed AFM [87], however, high speed AFM is very low throughput and does not allow motors to be observed in parallel as the FOV must be small to achieve high frame rates. As noted previously, tracks where stators were observable tended to be stator-face-own, this may limit access of any stator-bound motor to enzyme[8], if it is trapped on the underside of the tile, and the proximity of mica may interfere with motor migration.

An alternative solution is super-resolution microscopy. I attempted to image track polymers with super-resolution localisation technique DNA PAINT [4, 37]. As in the DNA nan-

[8]The converse may also be true, an enzyme trapped between mica and tile is at high local concentration.

otube scheme, it is the stator-tops that are labelled with fluorophores, though this time transiently with the hybridization of 10 nt of an Alexa532 labelled imager strand. After the motor has passed a stator and cleaved away its PAINT docking top, that stator will no longer be localisable and will not form part of the super-resolution reconstructed image.

Three tile types were used to produce samples for microscopy. The first two were used to form polymer tracks in a similar way to the tube scheme, where a start half-tile was assembled with only staples for the right hand half of the tile with no edge staples along the right hand edge, and only one stator, just before the right hand edge with motor annealed. Separately a polymerisable tile was folded with forked zig-zag track across the full length of the tile with no edge staples. After these two tile types were assembled and purified, they were polymerised by the addition of 5× excess of linker staples to a mixture of start half-tile and polymerisable tile at a ratio of 1:4 and incubating for 1 hr.

The third tile was a marker for fiducial drift. It was assembled with 10T loops along both edges to ensure it remained monomeric and contained extensions to both 5' and 3' ends of 29 staples in a central patch of the tile with a PAINT docking sequence. Due to the high density of docking sites (58 per tile) the patch will be occupied by at least one PAINT imager strand for a significant proportion of the time given an appropriate bulk imager strand concentration. Localised drift-marker-patches can be used for xy alignment of frames. All tiles were modified with biotin on staple extensions pointing from the face opposite the stators (and drift marker docking sites) for adhesion to a streptavidinated surface (see appendix D). This should favour tiles binding to the surface stator-face-up.

Super-resolution localisation was achieved with the ImageJ plug-in ThunderSTORM [213]. Positions used for drift correction were those where a marker was visible 10% of the time. Figure 6.18d shows drift-corrected super-resolution localisation molecular density maps projected onto a false pixel grid with 150 pixels/μm. A single frame (from a total of 24000) used to produce figure 6.18d is shown in figure 6.18c and a diffraction-limited reconstruction is produced by summing all 24000 frames shown in figure 6.18b. A zig-zag pattern can be seen in the super-resolution image but not in the diffraction-limited image.

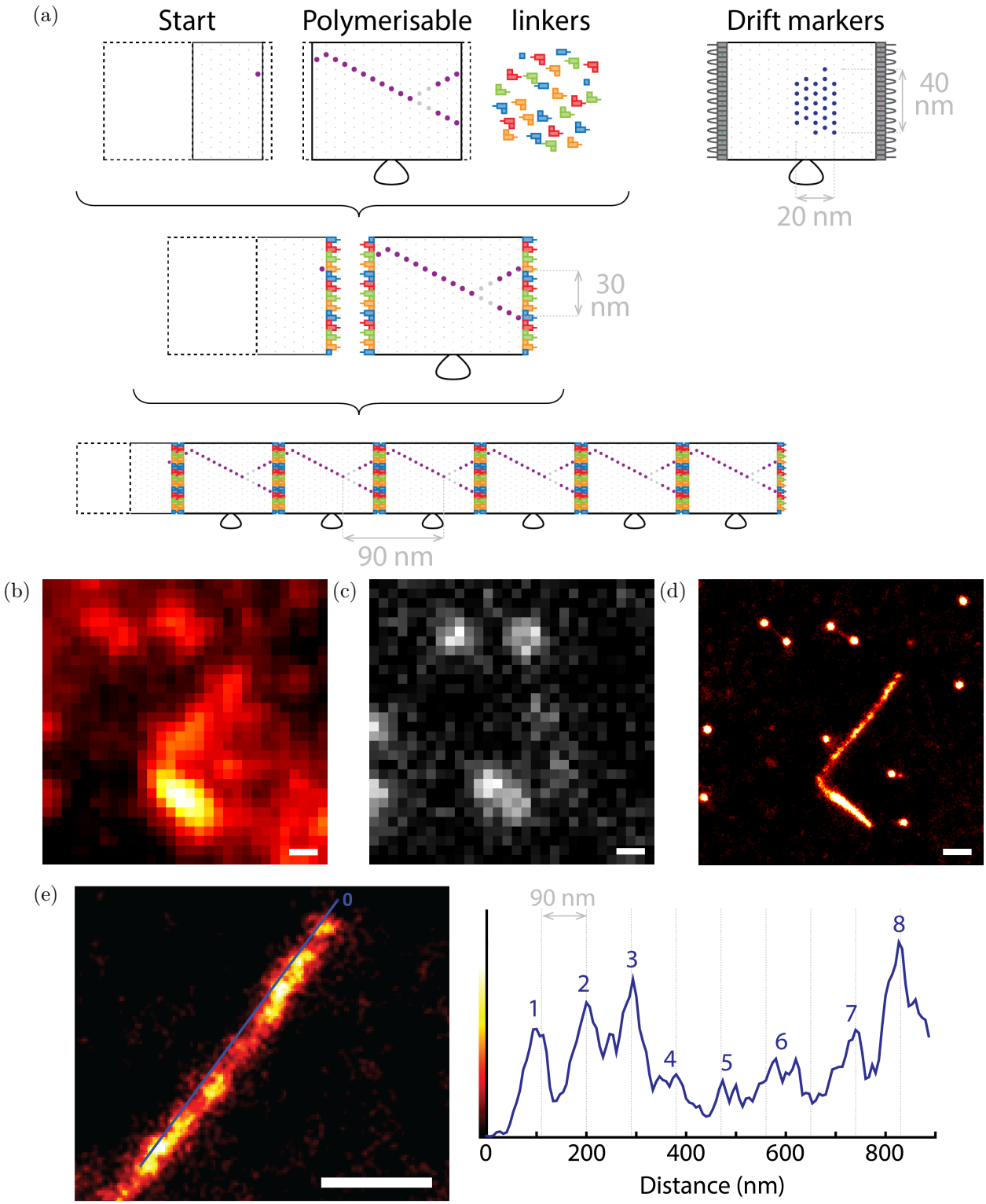

Figure 6.18: PAINT of burnt-bridges polymer track. **a** Schematic of tile types used and polymerisation method. Start half-tiles and polymerisable tiles are assembled at a 1:4 ratio without edge staples, purified then incubated with a 5× excess of linker staples. Stator positions are indicated on the tile diagrams with stators in purple bearing PAINT docking sites. Drift maker tiles have 10T loops on their edges to prevent polymerisation, and PAINT docking sites at positions marked in blue. **b** Diffraction-limited image created by summing all frames collected. **c** Single image frame. Any PSF in each image frame are fitted with integrated Gaussian functions to give localisations. **d** Super-resolution image created using ImageJ plugin ThunderSTORM [213] from 24000 frames each with 100 ms exposure, aligned using drift marker tiles positions. A zig-zag pattern can be seen. **e** Line plot along blue line enables periodicity of zig-zag to be measured. Peaks (numbered) occur with ∼90-100 nm spacing similar to the expected period of ∼90 nm. Scale bars 300 nm.

This pattern is likely to be due to localisation of PAINT docking sites on a forked zig-zag track polymer. Resolution is insufficient to see the track forks which may be because the fork only has three stators with docking sites, whereas the zig-zag section is a run of 14 stators with docking sites. Drift markers appear as spots with FWHM ∼50 nm, which is roughly the size of the patch of PAINT docks on the tile. 11 such markers are visible in figure 6.18d. A line profile obtained along the zig-zag pattern (figure 6.18e) shows a regular pattern of peaks spaced by ∼100 nm, similar to the individual tile length 90 nm.

Track resolution of 50 nm achieved with DNA PAINT

Spatial resolution can be defined by the FWHM of a point image. Two points can be resolved when separated by at least one FWHM. In my case I do not have any isolated PAINT docking positions which could be used to estimate the FWHM. As an alternative method for measuring the FWHM I plot a line profile perpendicular to the approximate position of the track, in this direction the profile should be similar to that of a single point. I found the FWHM for such line profiles to be 50±20 nm, significantly lower than the diffraction limited resolution of ∼250 nm. This would suggest that I can identify features of this size or larger than this in the image. I can see the width of the zig-zag, which should have a scale of ∼30 nm so it is likely that my estimate is on the conservative side.

The best reported resolution for DNA PAINT imaging is 1 nm [214], comparable with other localisation-based super-resolution methods [215]. My estimated resolution is significantly worse. This could be due to properties of the microscope, photo-properties of the fluorophore used or properties of the sample used. One limit on the achievable resolution imposed by the sample is in the fact that the PAINT docking positions are at the top of stators, and able to diffuse around at the end of the ssDNA stator tether. Normally PAINT docks would be placed close to the rigid surface of the origami to reduce the variance of the position of the fluorophore when bound. This would reduce the resolution by ∼2 nm, the RMS end to end length of the stator ssDNA, hence it is not the main cause of the low resolution. In principle, higher resolution PAINT images could be achieved of the same sample type. Additionally resolution could be increased by reducing the size of the fiducial marker.

These results show that it is possible to image polymer tracks spanning multiple tiles. Tracks on a single tile may be more difficult to identify, due to their smaller spatial scales which are of the order of the resolution of the reconstructed image.

I found it difficult to identify ROI to record because regions containing track polymers and regions where no track polymers were visible after super-resolution reconstruction produce very similar looking raw data. To improve this, a rapid form of super-resolution reconstruction should be used, fast enough to enable ROI to be identified during image acquisition.

6.4 Summary

Long tracks for the burnt-bridges motor have been produced either inside DNA origami nanotubes or across multiple DNA origami tiles. Polymerisation can be achieved without increasing the gap between monomers by adapting the edge staples to allow opposite edges to link. Asymmetric staples polymerise nanotubes better than linker staples with symmetric domain lengths sticking to either edge. Burnt-bridges motors are able to cross between DNA nanotubes, demonstrating a potential to walk long distances along multiple origamis. DNA tiles of multiple different types can be connected in a defined order by using a subset of linking edge staples, with remaining helices on edges bearing poly-T loops to prevent stacking. This opens up possibilities for more complex computation using walkers as uniquely addressable track positions are limited to a single tile. Labelling the track is one method for single molecule motor studies. The potential to image polymer origami tracks with super-resolution microscopy has been demonstrated with DNA PAINT, though greater resolution would be required to learn more about motor movement on a single tile.

A major limitation is the low yield of well-formed tracks, both for tubes and for tile polymers, which could be improved by changes to assembly conditions, or by new purification techniques to select for the correctly formed structures. In the case of tile hetero-polymers, reducing interactions between monomers would be key to increasing correct polymer yield as most undesired interactions are between monomers of the same type. This could be attempted by changing the type of non-linking edge staple used to bind more strongly, and early in assembly before undesired interactions have a chance to form. Another limitation

clear from TIRF studies of the motor on nanotubes is photobleaching. The Cy3 fluorophore used shows poor photostability under the laser illumination conditions used, even in anti-bleach buffer, therefore making long duration measurements difficult. This could be improved by methods to reduce laser exposure; such as adding a more rapid shutter to the laser, and focussing in a spectrally separate channel (e.g. IR autofocus) to reduce exposure to the laser soley to the time period over which the sample is being imaged, and not the time over which the focus is being determined. A different, more photo-stable, fluorophore could be tested such as Atto532, or alternatively nanoparticles could be used which are not subject to bleaching in the same way as organic fluorophores [84, 216]. The DNA PAINT technique has innate resistance to photobleaching as the fluorophore labelled strands exchange with a reservoir of identical strands in bulk solution which are not in range of the evanescent TIRF excitation. The limitations encountered with this technique, however, involve the difficulty in identifying ROI from raw data, therefore investing time in imaging a region where there are no clear structures is common. This could be improved by adding an additional label (e.g. a different coloured fluorophore) to motors or tiles, and imaging these in a different fluorescence channel to select a ROI to capture in the channel of the PAINT imager.

Chapter 7

Conclusion

In this thesis I developed structures and mechanisms for synthetic DNA motors. I had three aims. Firstly, I aimed to investigate mismatch repair by strand displacement [38], which forms part of the bipedal motor mechanism in reference 2, and to quantitatively characterise the effect of position of the defect to be repaired on displacement kinetics. My second aim was to create a single-stranded track for the bipedal motor of reference 2 held under tension, demonstrated *in silico* to improve the forward-stepping bias of the motor [123], and to observe motor operation on this taut track. The third aim was to create scaffolds and methods to observe long distance (>100 nm) movement of burnt-bridges motors [3].

In chapter 1 I described key principles and mechanisms used in structural and dynamic DNA nanotechnology. I detailed a selection of DNA motors with different locomotive principles, and explained how the advent of DNA origami enabled better scaffolds for DNA motor tracks to be produced. I also introduced some structures from the literature incorporating tension, which provided inspiration for my own DNA biped track design.

In chapter 2 I addressed the first aim with a systematic study of the effect of position of a mismatch to be repaired by displacement, on toehold-mediated strand displacement kinetics. I tested 14 mismatch positions within a 20 bp duplex which have similar duplex free energies. was fitted assuming second-order irreversible kinetics. The second-order rate constant varied non-monotonically with the distance of the mismatch from the toehold in the substrate duplex. Displacement rate increased as the mismatch was moved from 2 bp to 3 bp from the toehold, but then decreased exponentially beyond 3 bp with increasing distance from the toehold up to 13 bp, where displacement rates reach the equivalent rate without mismatches. Results independently collected using oxDNA coarse-grained simulations showed exactly the same behaviour. These findings were used to inform the design of a DNA-based pulse generator, where two invading strands competed for displacing a common substrate duplex, where one was kinetically favoured by choice of mismatch repair position.

In chapter 3 I addressed the second aim, by designing and constructing a DNA origami tightrope to hold a single-stranded section under a desired tension, adjusted by stretching more, or fewer, single-stranded nucleotides over a gap of fixed size. The design placed the single-stranded sequence to be held under tension into the scaffold strand, so M13 bacteriophage cloning was required, and a method for PCR screening of clones directly from phage-plaque-infected cultures was demonstrated. Tightrope origami assembly was checked by TEM and AFM imaging. The latter was more fruitful, due to the tendency of the tightropes deposit base-down on the imaging surface, allowing the central gap (over which the ssDNA is stretched) to be seen. As verification that the folded tightropes were holding the ssDNA under a designed tension, bulk measurements were taken of FRET probes attached to the ssDNA of tightropes assembled with different tensions. The FRET signal followed the expected trend, of decreasing FRET with increasing designed tightrope tension, due to increased probe separation, confirming the designed order of tensions.

In chapter 4 I looked at the effect of the tightrope origami structure on hybridization kinetics. The bipedal motor studied in chapter 5 undergoes repeated hybridization of feet to track, and therefore the impact of the structure on the rates of hybridization is of great importance. I modelled the hybridization of oligonucleotide probe strands to tightropes using a reversible bimolecular association reaction. Dissociation (measured via a probe dilution method) was fitted well by an exponential decay, as expected from a first-order dissociation process. Association was poorly fit by the bimolecular association process, however, but better fit by a bimolecular association with two populations. The fitted rate constant k_{off} determined from the dissociation fits was found to vary non-monotonically with the number of nucleotides stretched across the tightrope gap, and hence tension, with a minimum at 94-102 nt stretched across the 35 nm gap. The rate constant k_{on}, from single population bimolecular association fits, was maximal for 102-122 nt across the gap, for both 10 nt and 22 nt hybridizing domains. Tightrope-bound duplex stability inferred from the same data was found to be maximum with 122 nt, and significantly reduced ($\Delta\Delta G$ increased by 1.5 kcal mol^{-1}) when 144 nt were used to span the gap. A compromise tension with 102 nt across the gap (estimated $\sim$9 pN) was used for studies of the bipedal motor in chapter 5.

In chapter 5 the bipedal motor of reference 2 was given non-identical foot sequences to reduce motor stalling by overstepping. The motor turned over >16 fuels in 16 hours on a two-site test track, only when the motor orientation matched the fuel used. A scheme for loading the motor specifically onto the start site of the tightrope origami track was tested on olignucleotide tightrope mimics and on tightropes. Gel evidence suggested that all aspects of motor loading worked as intended: reduction of non-site-specific foot-track interactions using blockers; introduction of specific interaction at the start using a loading domain; removal of blockers; and removal of the loading domain. On three-site test tracks, foot-lifting with fuel was observed, but after foot replacement there was no change in the distribution of motors at the start and end of the track. Rearrangement of the motor between track positions was evident from 2D PAGE and bulk fluorescence experiments. On tightropes, though foot-lifting could be observed, there was no accumulation of motors at the track end. I concluded that the motor was unable to make directed progress along the track, which may be due to rearrangement of the motor along the track.

In chapter 6 I tackled the third aim, to observe long distance locomotion of a burnt-bridges motor, using two different track origami designs. The first was a DNA nanotube, with a motor track on the inside. Tube polymerisation with linker staples allowed the track to span multiple tubes. I found linker staples with 6 nt sticky ends to be optimal. Motor crossing between two tubes in a dimer using these linker staples was verified by bulk fluorescence. Motor observation on even longer tubes was attempted with TIRF microscopy, however substantial photobleaching limited measurements. Instead motor movement was inferred from 'flashing' of fluorophores cleaved from the tops of stators as they entered and left the observation volume, which requires the movement of the motor. I also attempted to join multiple, unique, track-bearing origami tiles together in a prescribed order by using subsets of linking staples for each linkable edge, whilst leaving all other edge staples without sticky ends. When 4 monomers were mixed together with 8 linking helices per edge, and 10T loops on non-linking helices, enrichment of the desired 4-mer was observed, and the correct ordering of tiles in 4-mers was confirmed by AFM, though repeating units of a single monomer type were also seen. DNA PAINT imaging of a polymer zig-zag track was performed with 50 nm resolution. This could be used to watch track burning as a motor moves along it.

Appendix A: Origami methods

A.1 Annealing

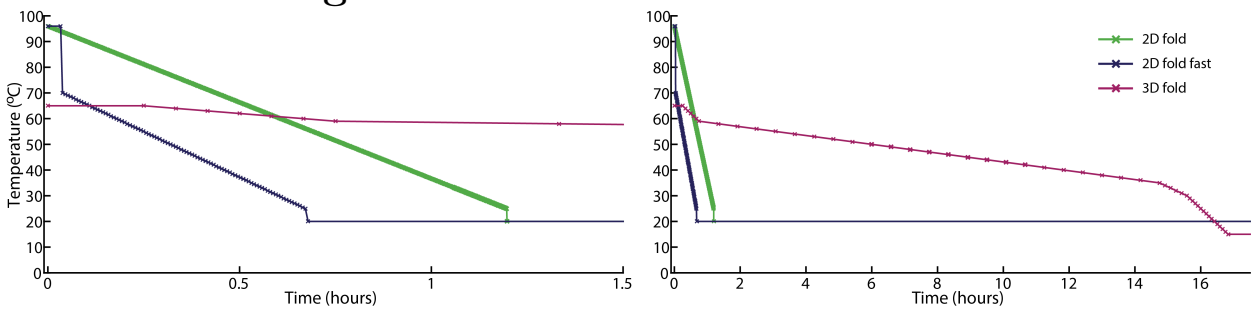

Figure A.1: Thermal ramps for annealing origami. 2D folding (green), rapid 2D folding (blue) and 3D folding (purple) plotted on different time axes.

A.1.1 2D tiles and tubes

Origami tubes were folded with a 5× excess of non-stator staples and 3× excess of stator staples, at a scaffold concentration of between 10 nM and 50 nM in 1× TE + 12.5 mM $MgCl_2$, or 1 × TE + 50 mM NaCl + 10 mM $MgCl_2$. Though DNA nanotubes are technically 3D, they are only one helix thick with similar crossovers to conventional 2D origami. Consequently a 2D origami thermal ramp was used, first heating the sample to 95°C, then cooling linearly to 25°C in steps of -0.1°C every 6 seconds (equivalent cooling rate of 60°C hour^{-1}).

Tiles were folded with a 5× excess of non-stator staples and 3× excess of stator staples, at a scaffold concentration of between 20 nM and 50 nM in Origami buffer. Tiles were annealed with one of two 2D folding ramps: either the ramp used for the nanotubes above, or a rapid version, cooling faster when far from the melting temperature of the tile. This ramp heats the sample to 96°C and hold for 2 minutes, then cools to 70°C and holds for a further minute, before cooling from 70°C to 25°C in steps of -0.5°C every 25 seconds. (equivalent cooling rate of 72°C hour^{-1}) Both folding ramps were observed to produce well-folded tiles.

A.1.2 3D tightropes

Tightropes require longer folding times due to their three-dimensional design [76]. The thermal ramp for the tightrope structures was devised by Dr Robert Schreiber similar to reference 217. The sample was first heated to 65°C held for 15 mins, then cooled in steps of -1°C every 5 mins to 59°C, then cooled in steps of -1°C every 35 mins to 35°C. Between

35-30°C cooling was in steps of -1°C every 10 mins and finally between 30°C and end temperature 15°C in steps of -1°C every 5 mins. Total annealing time is ∼16 hours, and the slowest cooling gradient is equivalent to 1.7°C hour^{-1}, ∼50× slower than 2D ramps.

A.2 Purification

A.2.1 S-300

Equilibrating the resin

25 ml Sephacryl S-300 stock 90% v/v in ethanol (GE Healthcare) was first mixed 1:1 with ddH_2O in a 50 ml centrifuge tube, by slowly inverting, then spun at 1000 rcf for 1 min to sediment the resin. The liquid supernatant was removed, volume restored to 50 ml, and the mixing and sedimentation followed by disposal of supernatant was repeated a further three times with ddH_2O, and four times with Origami buffer (see appendix G) before topping up to 40 ml with Origami buffer, to give 50% v/v S-300 in Origami buffer.

S-300 column origami purification method

Empty Micro Bio Spin columns (Bio Rad) were filled with 500 µl 50% S-300 in Origami buffer. The columns were packed by spinning at 1000 rcf for 4 min in a bench-top microfuge (Eppendorf) to remove excess buffer, adding a second 500 µl 50% S-300 and spinning for a further 4 min. The resulting column has a volume of approximately 500 µl S-300 after the removal of excess buffer. To a S-300 packed column, between 20 and 75 µl unpurified origami was added, and passed through the column by spinning at 1000 rcf for 3 min. The void volume containing origami structures was collected. Staples were retarded by entering the resin. Unless explicitly stated otherwise, origami samples were purified by running through three S-300 columns in series.

A.2.2 PEG precipitation

PEG precipitation was carried out using a similar protocol to reference 86. A 2× stock of precipitation buffer was prepared containing 1 × TAE, 15% w/v PEG 8000, 500 mM NaCl and 12.5 mM $MgCl_2$. This buffer was mixed with unpurified origami in a 1:1 volume ratio to achieve a final PEG concentation of 7.5% w/v. The PEG/origami solution was then spun at 20000 rcf at 20°C for 30 mins in a refrigerated centrifuge (Eppendorf)[1]. Immediately after

[1]Refrigeration ensures origami do not melt due to frictional heating at high rcf.

centrifugation, the supernatant was removed with a pipette, taking care not to dislodge the pellet, by pipetting with the tip pointed on the opposite side of the base of the tube from the pellet. With low quantities of origami (∼0.5 pmoles) the pellet was not clearly visible, but with fluorophore-labelled, or high quantities (∼2.5 pmoles) of origami, the pellet was visible by eye. Pellets were resuspended in buffer (Tightrope buffer for tightropes, Origami buffer for tiles) by pipetting continuously for at least 5 mins. Unless explicitly stated otherwise, precipitation was repeated three times to remove the majority of the staples [86]. In motor preparation protocols, precipitation steps were interspersed with motor loading steps.

A.2.3 Agarose Gel

After running an origami sample on an agarose gel, the band was visualised by UV shadowing or fluorescence scanning for fluorophores co-localised to the origami structure (e.g. loaded motors). The band was cut away from the gel leaving as little excess gel around the band as possible. The cut band was placed on a clean sheet of parafilm and cut into small chunks with a scalpel. The lacerated band was squeezed by slow application of even pressure from above with a parafilm coated microscope slide, until beads of liquid emerged from around the edges of the slide. The beads of liquid contained purified origami in gel running buffer, and were collected from the hydrophobic parafilm surface with a pipette.

Appendix B: Gel Electrophoresis

B.1 PAGE

Native polyacrylamide gel electrophoresis was carried out by diluting a stock solution of 30% w/v 29:1 Acrylamide:bis-acrylamide (Sigma) to the desired percentage in a final concentration of 1× TAE. To polymerise 5 ml of acrylamide solution, 50 µl 10% w/v ammonium persulphate (Sigma) and 5 µl tetramethylethylenediamine (Sigma) were added and the solution was mixed by inverting. The solution was poured between two glass plates with 0.75 mm spacing, a gel comb was added, and the gel was left ∼20 min to polymerise.

1-10 µl of sample containing 2.5% w/v Ficoll (Sigma) or 0.5-1× loading dye NEB purple (New England Biolabs) was added to each well. Gels were run in 1× TAE in a Mini-

PROTEAN Tetra cell vertical gel electrophoresis system (Bio-Rad) in a cold room or refrigerator (∼5°C). After electrophoresis, gels were either scanned in the glass plates, when fluorophore-labeled DNA was present, or removed from the plates and stained for 5 min with 1/10000 Sybr Gold (Life Technologies) and destained for 1 min in ddH_2O before scanning.

B.2 AGE

Agarose gel electrophoresis was performed by preparing a solution containing the appropriate mass of agarose (Fisher Scientific) in gel running buffer (usually 1× TAE with 6-11 mM $MgCl_2$ for origami samples) to achieve the desired % w/v agarose. The agarose was dissolved to form a molten gel in a microwave, cooled to ∼ 55°C to pour ∼0.5 cm deep into a level mould containing a gel comb. Once set, the gel was transferred to a wide Mini-Sub cell GT electrophoresis tank (Bio-Rad) and submerged in running buffer. The comb was removed and 1-20 µl of sample containing 2.5% w/v Ficoll or 0.5-1× loading dye NEB purple was added to each well cast with a standard comb, and up to 100 µl in wells cast with a custom comb. Gels were run either in a cold room (∼5°C), on the bench-top (∼20°C), or in an ice bath (∼0°C). Gels were stained with ∼1/10000 Sybr Gold in used running buffer. Staining time was adjusted for gel percentage. Staining time ranged from 15 min for a 0.5% gel to 60 min for a 2.5% gel with half as long destaining in used running buffer.

B.3 Visualising gel samples

Gels were scanned with a PharosFX molecular imager (Bio-Rad) using filter/laser sets from table B.1, or visualised by UV-shadowing over a TLC sheet.

Fluorophore	Laser λ(nm)	Filter bandpass (nm)
Cy3	532	605
Cy5	635	695
Cy3/Cy5 FRET	532	695
Alexa488	488	530
Alexa647	635	695
Alexa594	532	640
Alexa546	532	605
FAM	488	530
Sybr Gold	488	530
Sybr Green II	488	530

Table B.1: Lasers and filters for fluorescence observation of fluorophores used in PharosFX scanner.

B.4 Summary of gel conditions

Figure	Matrix	%	$MgCl_2$ (mM)	Observation Stain (Fluor)	PD (V)	Time min	Temp (°C)
3.4b	Polyacrylamide	10	0	Sybr Gold	180	60	5
3.7a	Agarose	2	11	(Cy5)	60	150	0
5.5	Polyacrylamide	10	0	(Cy5)	180	60	5
5.6 LHS	Agarose	2	11	(Cy5)	55	180	0
5.6 RHS	Agarose	2	11	(Cy5)	60	150	0
5.8	Polyacrylamide	15	0	Sybr Gold (Cy5, Cy3)	180	120	5
5.9	Polyacrylamide	15	0	Sybr Gold (Cy5, Cy3)	180	30	5
5.10	Agarose	0.7	6	(Cy5)	60	60	20
5.11	Polyacrylamide	10	0	(Cy5)	180	60	5
5.12	Polyacrylamide	10	0	(Cy5, FAM)	180	60	5
5.13	Polyacrylamide	10	0	(Cy5, FAM)	180	60	5
5.14a	Polyacrylamide	15	0	Sybr Gold (Cy5)	180	45	0
5.14b	Polyacrylamide	15	0	Sybr Gold (Cy5)	180	30	0
??	Polyacrylamide	15	0	Sybr Green II (Cy5)	180	30	0
5.15	Polyacrylamide	10	0	(Cy5, FAM)	180	120	0
5.16	Polyacrylamide	10	0	(Cy5, FAM)	180	120	0
5.17a	Polyacrylamide	10	0	(Cy5, FAM)	180	120	0
5.17b	Polyacrylamide	10	0	(Cy5, FAM)	180	120	0
5.18a	Polyacrylamide	10	0	(Cy5, FAM)	180	120	0
5.18b	Polyacrylamide	10	0	(Cy5, FAM)	180	120	0
5.20	Polyacrylamide	10	0	(Cy5, Cy3)	180	120	0
6.3	Agarose	0.5	6	(Cy3)	50	70	20
6.6	Polyacrylamide	15	0	Sybr Gold (Cy3)	200	80	5
6.11b	Agarose	0.7	6	Sybr Gold	55	180	20
6.12c	Agarose	0.7	6	Sybr Gold	55	180	20
6.14b	Agarose	0.7	6	(Alexa647, 488, 594)	55	150	20
6.15b	Agarose	0.5	6	(Alexa647, 488, 594)	55	105	20

Table B.2: Gel conditions used to produce figures in this thesis.

Appendix C: Fluorometer Experiments and Fluorophore labelled DNA

C.1 Fluorometer hardware

Two fluorometers were used to obtain the data presented in this thesis. The fluorometer used to collect data in chapters 2, 4 and chapter 5 was a Jobin Yvon Horiba Spex Fluoromax 3 with water-bath-cooled four-position sample changer. The fluorometer used in

chapters 3 and 6 was a Varian Cary Eclipse with a four-position sample changer and peltier temperature control. The Cary Eclipse has more staple temperature control especially at temperatures further from room temperature. The Fluoromax 3 has a lower fluorescence detection threshold and is more appropriate for looking at small changes in fluorescence or lower concentations of fluorescent sample. It typically exhibits a higher signal to noise ratio than the Cary Eclipse. I will discuss some features of the Fluoromax 3 in the coming sections, as they bear relevance to the collected data.

C.1.1 Fluoromax 3

The Fluoromax 3 sample illumination is via a Xenon lamp (Osram), spectrally filtered using an excitation grating monochromator to pick off a target wavelength. Band-pass filtering of the exciting light is achieved via a physical slit. The grating monochromators of the Fluoromax exhibit a dispersion of 4.25 nm mm^{-1} so to achieve a bandpass of 2 nm, the slit width would be 0.47 mm (bandpass = slitwidth $\times$ dispersion). The emitted light is spectrally filtered via another grating monochromator/slit pair. The excitation and emission slit widths can be independently controlled.

Xenon lamp

The Xenon lamp lifetime is $\sim$ 2200 hours when the lamp is in a vertical orientation. The lamp brightness also varies over its lifetime, and there is variation in intensity between new lamps, mostly due to alignment of the bulb in the mirrored bulb housing. This will have affect photobleaching of samples, though due to the acquisition mode, the magnitude of the recorded signal (which is the signal collected at the emission wavelength divided by the signal from a reference beam diverted from the exciting beam prior to reaching the sample) should be similar. Bulb signal is measured via intensity of a water Raman peak, excited at 350 nm. Peak intensity over the lifetimes of all of the bulbs used varies by $<3\times$.

Slit width software bug

The software used for instrument control and data acquisition (DataMax version 2.2.10A) has a bug where one of the software methods to adjust the widths of the emission slit, would falsely report a change of slit width in the user interface, when no change had been made

within the instrument. As a result, prior to noticing the bug, some of the data in chapter 2 was collected with a variety of emission slit widths (either 1, 2, or 5 nm band-pass), resulting in large variation in the measured signal magnitude. As it is the emission slit, rather than the excitation slit, that is changing, any change in signal occurs after passing through the sample, so photobleaching is unaffected. It is expected that the only effect of this change is to change the magnitude of the recorded signal, though the change in signal may not be linear with slit width, as it will be affected by details of the fluorophore emission spectrum. Consequently, for the analysis and fitting of mismatch repair position data in chapter 2, the data is fitted to an arbitrary peak intensity, determined from each individual dataset. The fitted value of the second-order rate constant in chapter 2 is independent of the fitted maximum intensity which is related to the concentration of displaced V.

C.2 General fluorometer experimental method

C.2.1 Data collection

Fluoromax 3

Data was collected with excitation slit width 2 nm bandpass for chapters 2, 4 and 5 Emission slit width was 2 nm bandpass for data in chapter 2[1], and 3 nm bandpass in chapters 4 and 5. Data was collected with a 2 second integration time, at 60 second intervals.

Fluorophore	Excitation wavelength (nm)	Emission wavelength (nm)
Cy3	550	564
Cy5	648	664
Alexa647	650	668
Cy3/Cy5 FRET	550	664
Cy3/Alexa647 FRET	550	668

Table C.1: Wavelengths used in Fluoromax 3 (chapters 2, 4 and 5).

Cary Eclipse

Data in chapters 3 and 6 was collected with 5 nm bandpass excitation and emission widths. PMT voltage was set to 960 V for chapter 3 and 1000 V for chapter 6. Data was collected with 1 second integration time, at 13.6 second intervals.

[1]Except for experiments affected by software bug (see C.1.1).

Fluorophore	Excitation wavelength (nm)	Emission wavelength (nm)
Cy5	650	670
TET	522	536
Alexa647	650	670
Alexa546	555	571
Alexa546/Alexa647 FRET	555	670

Table C.2: Wavelengths used in Cary Eclipse (chapters 3 and 6).

C.2.2 Cuvettes

Cuvettes used were either 2 ml quartz cuvettes (Hellma 119.004F_QS), or 150 µl quartz cuvettes (Starna 16.100-F Q Z20). Cuvettes were washed before use by 5 washes with distilled water, followed by 2 washes with 100% ethanol. If mineral oil had been used in the sample (usually the case for the smaller volume cuvettes) an additional wash step with 2 washes with 0.5% Hellmanex (Hellma) and 2 washes with 100% ethanol were performed prior to washing with distilled water.

C.2.3 Reduction of evaporation

Evaporation causes artefacts in the fluorescence signal of a sample as the volume of solvent decreases, first increasing observed fluorescence intensity as fluorophore concentration increases, then decreasing signal, as the solution volume decreases below the observed volume. Evaporation artefacts in the two cuvette types were prevented using different methods:

Large cuvettes The 2 ml cuvettes were sealed with bases of 1.5 ml microfuge tubes as lids.

Small cuvettes The 150 µl cuvettes could not be well-sealed with lids, and therefore ∼200 µl Mineral oil (Sigma) was gently added on top of the sample by slowly pipetting the oil along the inner face of the cuvette.

C.2.4 Sample addition and mixing during an fluorometer run

When using the Fluoromax-3, data acquisition was paused and the cuvette removed from the fluorometer prior to an addition. In the Cary Eclipse, the cuvette was removed without pausing data acquisition, causing zero intensity readings whilst the cuvette was not in the light path. Lids (where present) were removed and the sample added (beneath the mineral oil if present) using one pipette, then mixed with a second pipette set to half the total

volume of the solution. Mixing was achieved by pipetting up and down 20×. Care was taken to ensure no bubbles or dispersed mineral oil was introduced into the sample. Lids were replaced, cuvettes returned to the instrument, and data acquisition resumed.

C.3 Handling of fluorescent DNA

Individual strands were stored at between 100 μM and 10 μM in ddH_2O or 1×TE buffer. For long term storage, multi-strand complexes were assembled and stored at 5-10 μM in assembly buffer. PAGE purification was avoided where possible as APS can oxidise fluorophores [193]. Fluorescent complexes were protected from light and stored at 4°C.

C.3.1 Reduction of photobleaching

Photobleaching is a common issue when working with fluorophores. Cy5 bleaches more rapidly than Cy3 [206, 218] under the fluorometer excitation and emission conditions used. Photobleaching was found to be reduced when the exposure of the liquid sample to air was reduced. For experiments in chapter 2, this was achieved by filling the nominal 2 ml volume cuvette with 1960 μl sample solution, then sealing with the base of a 1.5 ml microfuge tube, taking care to leave as little air trapped in the cuvette as possible.

C.3.2 Sticking to tubes

Dilute solutions congaing fluorescent DNA stored in plastic tubes were observed to rapidly decrease in fluorescence when observed in the fluorometer (a 5× decrease in fluorescence intensity was observed after storing for 5 hours). The tubes which previously contained the fluorescent DNA appeared fluorescent when excited in a laser scanner (figure C.1) indicating that the fluorescent DNA had stuck to the tube. After discovering this issue, to reduce the effect of fluorescent DNA sticking to tubes, Complexes or single strands of fluorescent DNA were kept at high storage concentrations of ≥5 μM and diluted immediately prior to use.

C.3.3 Anomalous dilution

When adding large concentrations of DNA to a solution containing a Cy3 labelled DNA strand, the observed fluorescence after addition of the extra DNA was higher than expected. This was observed when adding 30 μl unpurified 50 nM origami with ∼ 120 staples each in

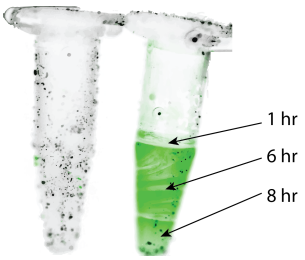

Figure C.1: Cy3-DNA adsorbed onto the surface of an eppendorf microfuge tube. Empty tube, never used (left), and empty tube which formerly contained 700 µl 20 nM Cy3 DNA (right), Cy3-DNA solution was removed from the tube in three steps. 120 µl removed after 1 hr, 480 µl removed after 6 hrs, and final 100 µl removed after 8 hrs. Cy3 excited by laser scanner (green in overlay) can clearly be seen in the tube which previously contained Cy3-DNA. After 6 hours in the tube, the fluorescence intensity observed in a fluorometer dropped by a factor of 5.

10× excess, to a solution containing 120 µl 25 nM 22 nt olgo with 3' Cy3. The presence of extra DNA in solution caused the fluorescence increase beyond the value prior to the addition, despite the addition of 20% non-fluorescent solution.

To test the hypothesis that it was the extra DNA in solution causing this effect, rather than some feature of the origami, 120 µl 25 nM of the same fluorescent oligo was observed, and then diluted with 30 µl of 66 µl 69 nt oligo, which was not expected to interact significantly, although NUPACK predicts the formation of a 6 bp duplex between the two oligos at concentrations of 33 µM. A further dilution in buffer alone was made, adding 37.5 µl buffer to the 150 µl solution, the same volume dilution factor. Results in table C.3 show that the fluorescence was decreased by less than expected when diluted with a solution containing high concentration of DNA. Dilution with buffer produced the expected drop in fluorescence.

	66 µM 69nt oligo	Tightrope buffer
Fluorescence dilution factor	0.95	0.79
Volume dilution factor	0.8	0.8

Table C.3: Anomalous fluorescence after dilution with high concentrations of DNA.

C.4 Two-toehold data processing

Raw two-toehold data shown in figure C.2 was scaled and combined prior to ODE fitting as described in this section.

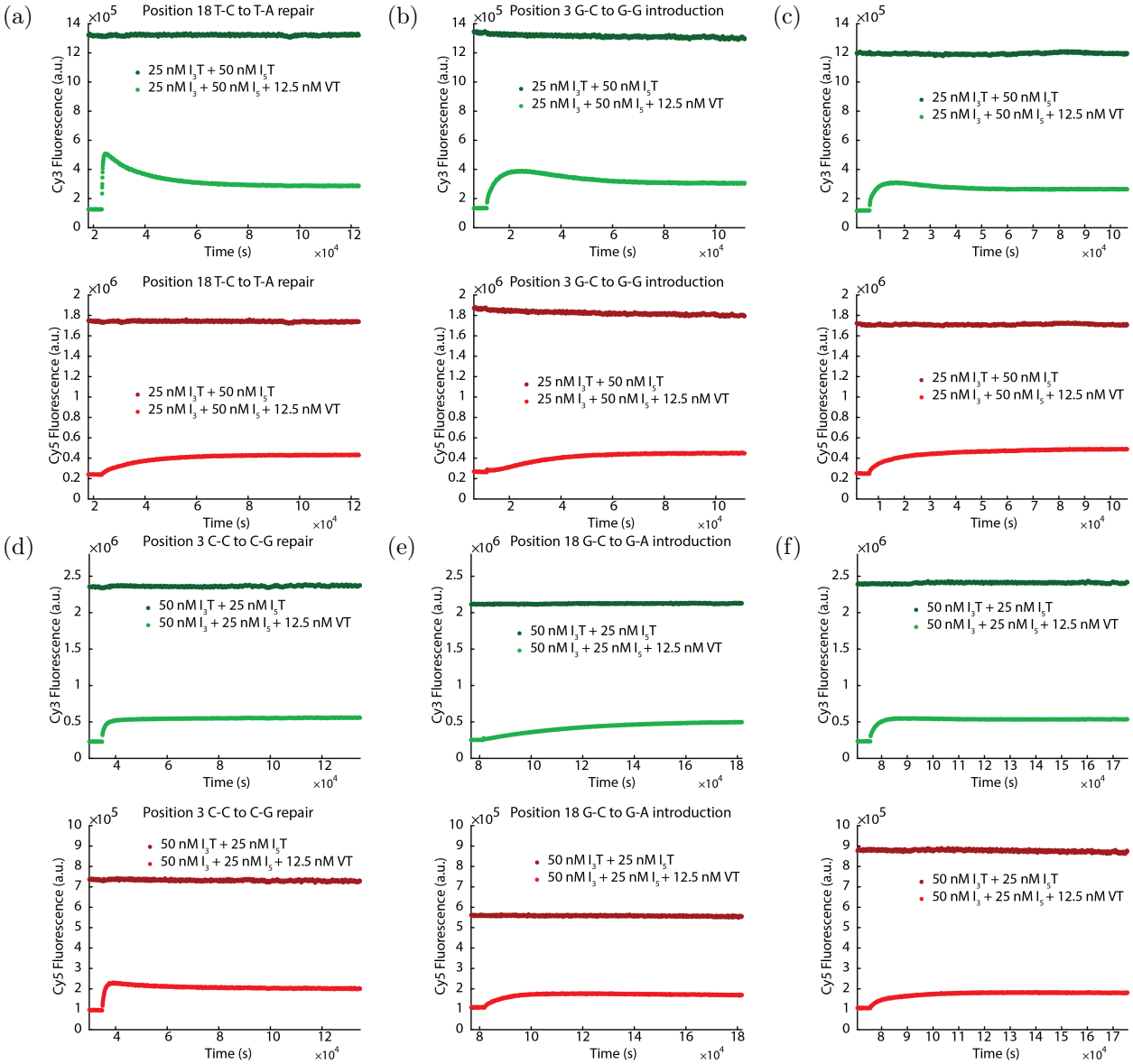

Figure C.2: Raw data used to produce plots in figure 2.9. Cy3 waveset data (top) and Cy5 waveset data (bottom) for experimental traces (light) and target saturated invaders (dark).

Initial scaling

Experimental fluorescence traces $f_X^{\text{raw}}(t')$ were first divided by the fluorescence of the concurrent saturated control sample $s_X^{\text{raw}}(t')$ in each wavelength X to give background divided fluorescence $\hat{f}_X(t') = \frac{f_X^{\text{raw}}(t')}{s_X^{\text{raw}}(t')}$. The data was zeroed by setting $t = 60$ at the first data point collected after mixing (60 s is the measured typical mixing time), $t = t' - t_{\text{start}}$, and subtracting from all fluorescence counts the fluorescence of the last data point collected before addition of VT duplex $\hat{f}_X(t) - \hat{f}_X(t = 0)$. Values of raw invader fluorescence at t=0 $f_X^{\text{raw}}(t = 0)$, and median saturated invader fluorescence $s_X^{\text{sat}}(\text{median})$ were used to scale the background divided traces to give the fraction F of invader strand which is bound to the

target $[I_X T]/[T]_{\text{total}}$

$$F_X = \left(\hat{f}_X(t) - \hat{f}_X(0)\right) \times \frac{s_X^{\text{sat}}(\text{median})}{s_X^{\text{sat}}(\text{median}) - f_X^{\text{sat}}(0)} \times \frac{[I_X]_0}{[VT]_0} \tag{C.1}$$

Where $[VT]_0 = [T]_{\text{total}}$ is the initial concentration of target-output duplex added and $[I_X]_0$ is the initial concentration of free X' invader.

Plateau subtraction

During a fluorometer run, fluorescence of a species can drift upwards or downwards relative to the control sample that it is divided by. The drift is linear on experimental timescales. The value of the plateau reached is important in the fitting process as it determines the ratio of fitted parameters k_{35} and k_{53} therefore a linear fluorescence drift subtraction is performed on the data. This is done by finding the second derivative of fluorescence values smoothed by taking a rolling average of 1000 data points. A median first derivative for all values where the second derivative is smaller than 0.5×10^{-5}n.u.s^{-2} is used as the gradient m of the straight line for linear subtraction. Normalised fluorescence data points are subtracted by this gradient multiplied by the time value $F_X^{\text{flattened}} = F_X - m_X \times t$.

Normalisation

The data is finally normalised such that the sum of the two fluorescence channels is 1 at its maximum by finding the sum of the two fluorescence values with time, averaging over a rolling window of 60 data points (∼1 hr) and taking the maximum. All fluorescence data points are divided by this maximum to give their normalised values.

Appendix D: Microscopy

D.1 AFM

Atomic force microscopy (AFM) was performed with either a Brucker Nanoscope with SNL-10 tips or a Asylum Cypher ES with AC40TS tips. The Nanoscope drives the cantilever with a piezo, and the Cypher drives the cantilever with a blue laser (filtered by a 0.01× ND filter cube). 5 µl sample in AFM buffer (see appendix G) was deposited on freshly cleaved mica at 2.5-10 nM concentration and left for 2 min. After deposition, ∼100 µl AFM buffer was added between the sample and the cantilever for imaging. All images were taken with

tapping mode in fluid. Typical imaging parameter ranges for the both microscopes are given below. Paramaters such as the tip resonant frequency (and hence drive frequency) vary from tip to tip and other parameters can vary from image to image.

D.1.1 Brucker Nanoscope (SNL-10 tips, 0.01× ND filter cube)

Drive frequency: 19 kHz

Free amplitude (in buffer): 200 mV

Amplitude setpoint (engage attempt): 120 mV

Drive amplitude: 14 V

Scan size: 0.5-10 μm

D.1.2 Asylum Cypher ES (AC40TS tips)

Drive frequency: 35-22 kHz

Free amplitude (in buffer): 800-400 mV

Amplitude setpoint (engage attempt): 550-250 mV

Amplitude setpoint (imaging): 480-180 mV

line scan rate: 2-5 Hz

Scan size: 0.35-10 μm

Points per line: 512-1024

D.2 TEM

Transmission electron microscopy (TEM) was performed using a Tecnai 12 microscope with Gatan OneView CMOS camera. 10 μl ∼1 nM samples were deposited on glow-discharged, carbon-coated, Formvar supported, copper grids (Agar scientific) for 2 min, excess sample was removed. Samples were stained by adding 10 μl droplet 2% w/v uranyl acetate (UA) for 20 seconds, removing, and then adding a second 10 μl UA droplet for another 20 seconds and removing. The stained samples were left to dry for 20 min before transfer to the microscope.

D.3 TIRF

Total internal reflection fluorescence (TIRF) microscopy was performed using an IX81 inverted microscope (Olympus), with 100× 1.4 NA oil immersion objective (Olympus) and 1.6× internal magnification, and laser excitation with a 50 mW fiber-coupled 532 nm laser (OZ optics) with 20 mW at the fiber end. Images were captured using an iXon EMCCD camera (Andor). The microscope was fitted with a motorised stage (Prior) with two home-made stage inserts for sample slide dimensions 60 mm ×25 mm or 75 mm × 25 mm. When using the 60 mm × 25 mm stage insert, samples were held with 170 g metal weights. When using the 75 mm × 25 mm stage insert, samples were secured by tightening metal lever arms. This was found to be required to reduce focal drift. Type NVH Immersion oil with refractive index 1.5150 (Cargille) was applied between the sample coverslip and objective.

The microscope and stage were controlled manually or by computer using μManager software [219, 220]. The iXon camera was controlled either using Andor IQ, or μManager.

D.3.1 Filters

Cy3/Cy5 filter cube

The filter cube used had a dual-band emission filter (Chroma) and no emission filter (see figure D.1a). The absorption spectrum for light passing directly through the filter cube dichroic was measured, data is shown in figure D.1b.

Dual-view

The microscope was fitted with a DV2 dual view (Photometrics) beam-splitter for simultaneous multi-channel imaging (see figure D.1c). After passing through the emission filter, but before reaching the camera, light was split by a dichroic filter at 630 nm. Long wavelength light transmitted by the dichroic was passed through a 680±25 nm filter and focussed on one half of the camera chip. Short wavelength light reflected by the dichroic was passed through a 575±20 nm filter and focussed on the other half of the camera chip. All images in this thesis come from this 'green filtered' half of the image.

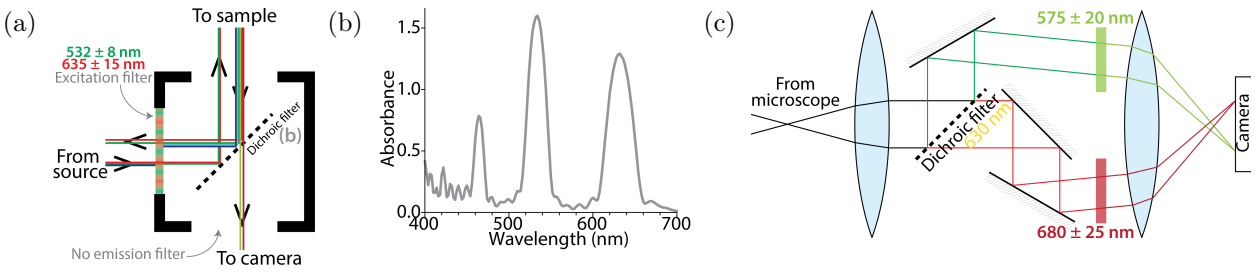

Figure D.1: TIRF microscopy schematics. **a** Cy3/Cy5 filter cube schematic. Light from the 532 nm laser will pass through the excitation filter, reflect at the dichroic and illuminate the sample. Light from the sample will be filtered by the dichroic such that longer wavelength light from emission of Cy3 or Cy5 can pass through. **b** Absorption spectrum showing wavelengths of light able to pass through the dichroic in **a** without reflection (low absorption). **c** Dual View DV2 (Photometrics) schematic. Light is split by a dichroic filter, and further spectrally filtered before focussing the two different colour light paths on the top and bottom of the camera.

D.3.2 Autofocus

After an initial focus had been established, the sample was kept in focus using autofocus software 'Oughtafocus' included with μManager [219, 220].

D.3.3 Sample preparation

Samples were imaged in homemade tunnel slides or μ-slides VI 0.4 uncoated (Ibidi).

Homemade tunnel slides

Slides and coverslips (including the base coverslip for the commercial tunnel slide), were burnt in a furnace at ∼300°C before use. To make tunnel slides, two 1 cm × 2.5 cm strips of parafilm (Bemis) were placed parallel lengthways, separated by ∼ 2 mm, along either a 60 mm × 25 mm No. 1.5 long coverslip (Menzel Gläzer) or 75 mm × 25 mm glass slide (VWR). A 22 mm × 22 mm no 1.5 coverslip (Menzel Gläzer) was placed centrally on top of the parafilm slips to form the tunnel. The slide/coverslip with parafilm was placed on a hot plate ∼ 80°C for 30 sec to melt the parafilm slightly and pressed down gently to seal. The volume of tunnel slides was estimated to be between 10 and 30 μl.

Streptavidin-mediated immobilisation protocol

The preparation steps for samples involve multiple wash and incubation steps with volume relative to the tunnel volume (TV). The sample immobilisation protocol is as follows:

- Wash with 1TV origami buffer
- Incubate 5 mins with 1TV 0.5 mg ml^{-1} biotinylated BSA (Sigma) / 2 mg ml^{-1} BSA (Sigma) solution in origami buffer.

- Wash with 2TV origami buffer
- Incubate 5 mins with 1TV 1 mg ml^{-1} streptavidin (New England Biolabs)
- Wash with 2TV origami buffer
- Incubate 2 mins with 1 nM - 10 nM biotinylated DNA sample
- Wash with 2TV origami buffer
- Wash with 2TV anti-bleach buffer, or DNA PAINT imager strands in origami buffer
- Seal with VALAP [221] for homemade tunnels, or plastic plugs for Ibidi slides.

D.4 Additional TIRFM data for chapter 6

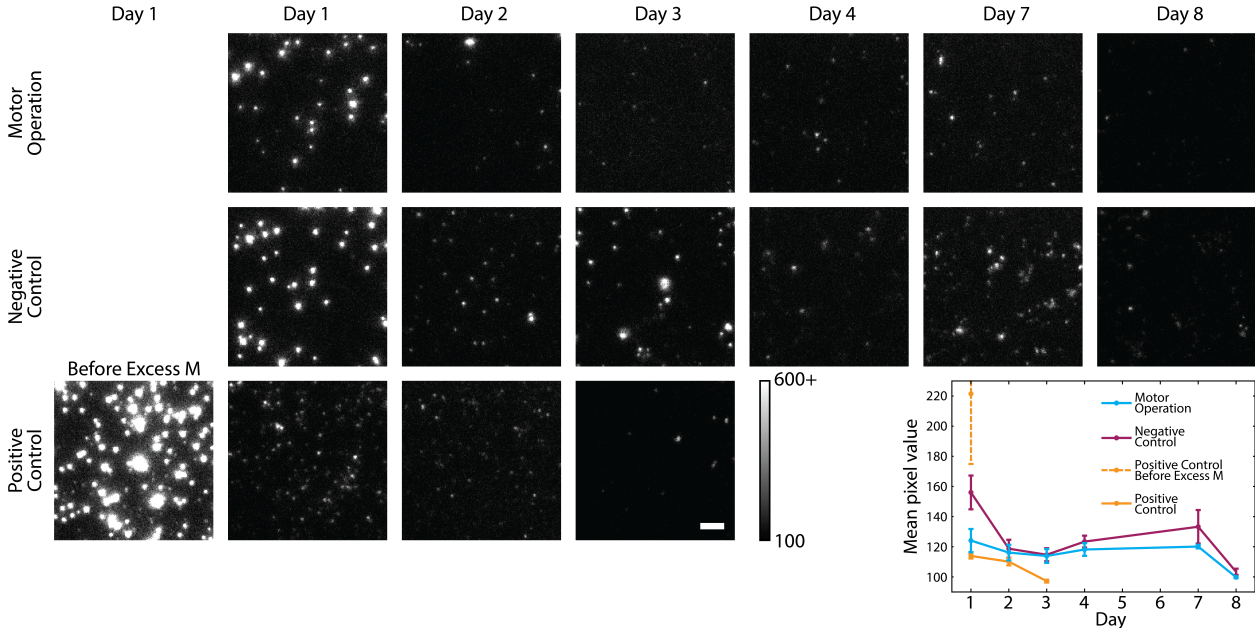

Figure D.2: Representative TIRF images for samples of 1:20 A:B polymerised tubes with motors loaded onto A with Cy3 stator-top probes. Top images show the tubes imaged in anti-bleach buffer containing 0.5 U/ml Nt.BbvCI. This should enable the motor to walk along the track. Middle images show a negative control sample imaged in anti-bleach buffer without enzyme. Stator tops should not be cleaved away. Bottom images show positive control for stator cutting in anti-bleach buffer containing 0.5 U/ml Nt.BbvCI and 42 µM unlabelled motor. Motors can hybridize all stators completing restriction sites for Nt.BbvCI, allowing stator cleavage without motor migration. All images have the same pixel value LUT mapping shown with the colour-bar. Scale bar 3 µm. Lower right polot shows mean pixel value for sets of 5 or 6 images of 1280 $µm^2$ FOV taken each day for each sample. Imaging conditions 100 ms exposure, 100× EM gain. Error bars 1 standard deviation.

Appendix E: mFJC Model of Entropic Spring ssDNA

E.1 Derivation of force-distance relationship for a ssDNA entropic spring using an mFJC model

Consider the partition function z_i for an individual segment of a FJC (NOTE: Here the individual chain segment is NOT a nucleotide - it is a Kuhn segment with length equal to twice the persistance length of the polymer (in this case ssDNA) and is larger than the size of a nucleotide). Each segment i has length b and orientation $\hat{\vec{n}}_i$ where $\left|\hat{\vec{n}}_i\right| = 1$.

$$\text{Single segment partition function} = z_i = \int d^2 b \hat{\vec{n}}_i e^{\frac{-E_i}{k_B T}} \tag{E.1}$$

$$E = -\vec{f} \cdot \vec{R} \tag{E.2}$$

Without loss of generality the force can be considered along the axis $\vec{z}$ such that $\vec{f} = f\hat{\vec{z}}$.

$$\vec{R} = \sum_{j=1}^{X} b\hat{\vec{n}}_j \tag{E.3}$$

$$E = \sum_{i=1}^{X} E_i = -fb \sum_{i=1}^{X} \cos(\theta_i) \tag{E.4}$$

Where θ_i is the angle between unit vectors $\hat{\vec{z}}$ and $\hat{\vec{n}}_i$.

Integrating over all possible orientations of the unit vector we obtain:

$$z_i = \frac{4\pi k_B T}{fb} \sinh\left(\frac{fb}{k_B T}\right) \tag{E.5}$$

As each segment of the FJC contributes indepently to the energy, partition function Z is simply a product of partition functions of individual segments z_i:

$$Z = \prod_{i}^{X} z_i = \left[\frac{4\pi k_B T}{fb} \sinh\left(\frac{fb}{k_B T}\right)\right]^X \tag{E.6}$$

Helmholtz free energy F

$$F = U - TS - fR \tag{E.7}$$

$$dU = -pdV + TdS + fdR \tag{E.8}$$

$$F = -k_B T \ln Z \tag{E.9}$$

$$R = -\left(\frac{\partial F}{\partial f}\right)_{V,T} = k_B T \frac{\partial}{\partial f} lnZ \tag{E.10}$$

$$R = Xb\left[\coth\left(\frac{fb}{k_BT}\right) - \frac{k_BT}{fb}\right] \tag{E.11}$$

Including segments which are themselves elastic rather than rigid:

$$R = Xb\left[\coth\left(\frac{fb}{k_BT}\right) - \frac{k_BT}{fb}\right]\left(1 + \frac{f}{S}\right) \tag{E.12}$$

Where S is the stretch modulus and is related to the spring constant for an individual chain segment k by the relationship $S = bk$.

To find the number of ssDNA nucleotides in the FJC we equate the contour length written in terms of Kuhn segments $L = Xb$ with the contour length written in terms of nucleotides $L = Nl$ where l is the length of one nucleotide and N is the number of nucleotides.

$$R = Nl\left[\coth\left(\frac{fb}{k_BT}\right) - \frac{k_BT}{fb}\right]\left(1 + \frac{f}{S}\right) \tag{E.13}$$

Values for the relevant parameters can be obtained from [16, 113]:

$S = 800 \times 10^{-12}$ N
$b = 1.5 \times 10^{-9}$ m
$T = 310$ K
$k_B = 1.38 \times 10^{-23}$ kg m^{-2} s^{-2} K^{-1}
$l = 0.5 \times 10^{-9}$ m

E.2 Tension of dsDNA and ssDNA length equivalence

Consider the case where the RMS end-to-end distance is equal to the contour length of a DNA duplex $R = 0.34 \times 10^{-9} \times N$ m.

$$0.34 \times 10^{-9} \times N = Nl\left[coth\left(\frac{fb}{k_BT}\right) - \frac{k_BT}{fb}\right]\left(1 + \frac{f}{S}\right) \tag{E.14}$$

N cancels indicating that the result is independent of the number of segments.

$$\frac{0.34}{0.5} = \left[\coth\left(\frac{fb}{k_BT}\right) - \frac{k_BT}{fb}\right]\left(1 + \frac{f}{S}\right) \tag{E.15}$$

The point of equal length occurs when $f = 8.6$ pN as can be seen from a plot of the two halves of equation E.15 in figure E.1.

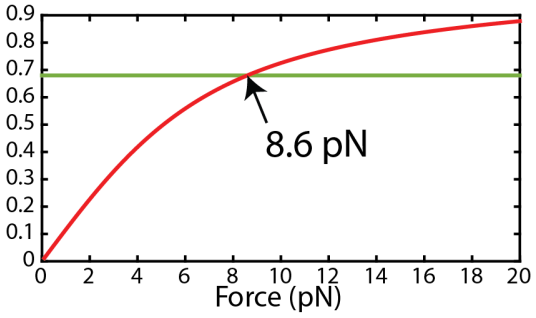

Figure E.1: Green=LHS of equation E.15. Red=RHS of equation E.15. Intersection at $f = 8.6$ pN.

E.3 A ssDNA mFJC with a single dsDNA segment

The result for the end-to-end distance R of a ssDNA freely jointed chain including a single duplex of length D can be obtained following the same methodology assuming the duplex length is much shorter than the persistance length of double stranded DNA and it can therefore be treated as a FJC segment with Kuhn length equal to the duplex length.

$$R = Nl\left[\coth\left(\frac{fb}{k_BT}\right) - \frac{k_BT}{fb}\right]\left(1+\frac{f}{S_{ss}}\right) + D\left[\coth\left(\frac{fD}{k_BT}\right) - \frac{k_BT}{fD}\right]\left(1+\frac{f}{S_{ds}}\right) \quad \text{(E.16)}$$

S_{ss} is the stretch modulus of ssDNA of 800 pN [16]. S_{ds} is the stretch modulus of dsDNA measured as ∼1000 pN in reference 14. D is duplex length in nm. Assuming tension is applied to the same strand of the duplex, for short duplexes the helical repeat (∼16 bp per 1.5 turns) will also be needed to calculate length.

$$D = \sqrt{(md)^2 + \left(2r\sin\left(\frac{m\pi}{h}\right)\right)^2} \quad \text{(E.17)}$$

Where m is the number of base pairs in the duplex; d, the length per base pair, is 0.34 nm bp^{-1}; r, duplex diameter, is 1 nm; and h, the helical repeat of the duplex, is 10.5 bp.

E.4 A particle on a tightrope is subject to a non-symmetric an-harmonic potential

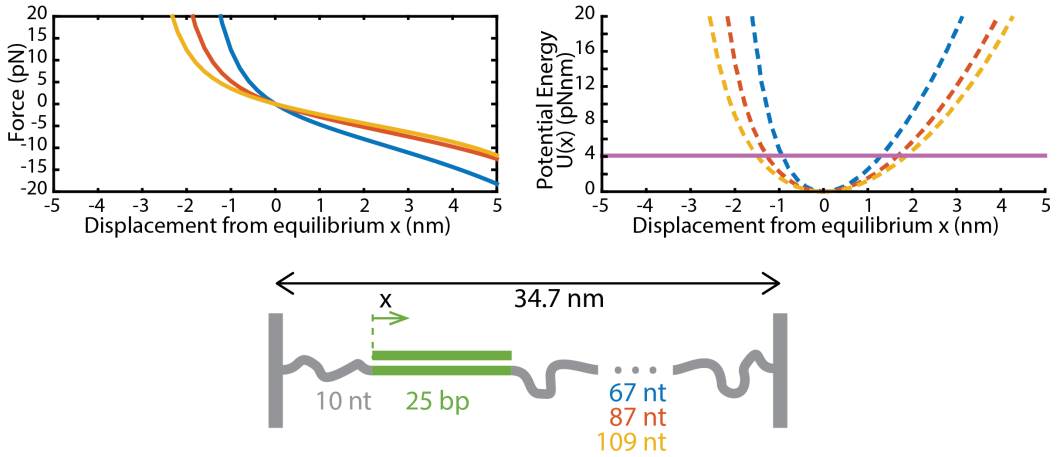

Figure E.2: Top: Restoring force (top left) and potential energy (top right) of a particle displaced from its equilibrium position 10 nt from the left hand side tightrope end held by a fixing duplex of 25 bp with either 67 nt (blue), 87 nt (red) or 109 nt (yellow) between the opposite end of the fixing duplex and the RHS of the tightrope. k_BT at 25°C is shown as a pink line. Corresponding schematic of tightrope system shown (bottom).

For an FJC model, the entropic spring can be considered Hookean in the low force limit where $fb \ll k_BT$. At room temperature $k_BT \approx 4.1$ pNnm however $b = 1.5$ nm therefore the force f must be $\ll 3$ pN to be in the Hookean limit. Outside of this limit the relationship

between force and extension is non-linear and defined by equation E.13. This equation can be numerically solved for f, and used to generate restoring forces for perturbation of a particle attached to the tightrope along the end to end vector of the tightrope. The integral of restoring force can be used to obtain potential energy of a particle displaced from equilibrium. Figure E.2 shows the potential for an asymmetrically placed particle 10 nt from the LHS of a tightrope held by a 25 bp duplex for different lengths of ssDNA across the gap.

E.4.1 Expectation value for FRET efficiency

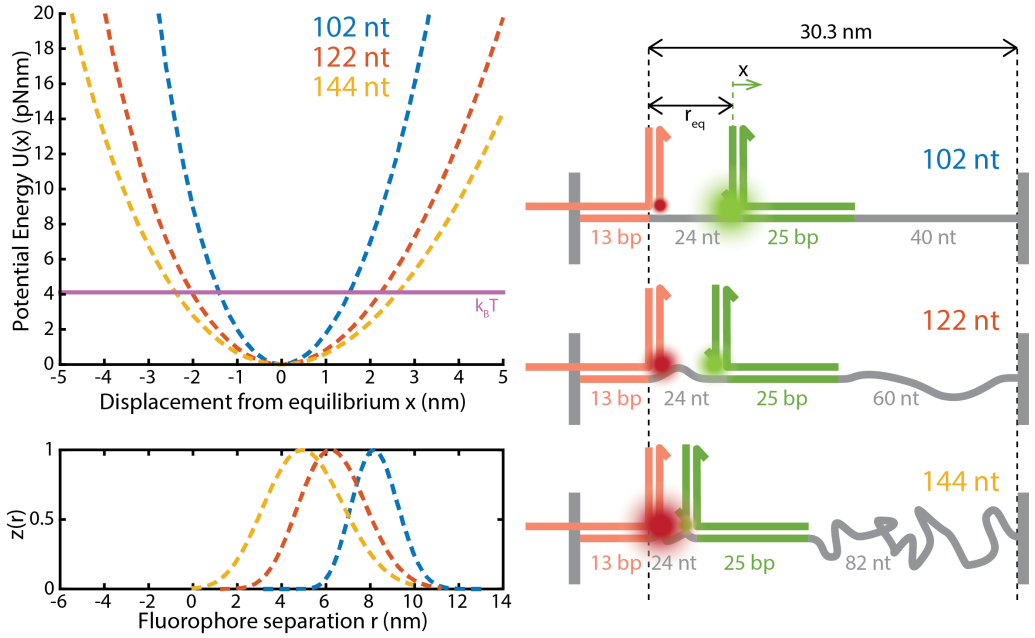

Figure E.3: Potential energy distribution and probability distribution for the bookend FRET experiment in section 3.5.4. Plots of calculated potential energy $U(x)$ (top left) and Boltzmann probability distribution $z(r)$ (bottom left), using distances marked in schematics (right). NB: $r = x + r_{eq}$.

As demonstrated, the potential $U(x)$ for a particle on a tightrope (constrained to 1D motion) can be determined from its displacement from equilibrium position x on the tightrope using numerical methods. This can then be used to calculate the Boltzmann distribution as a function of displacement from equilibrium position $z(x) = e^{-U(x)/k_BT}$.

$U(x)$ and $z(x)$ are calculated for the bookend system in section 3.5.4 using MATLAB 2014b using the ssDNA, dsDNA, and gap lengths shown in figure E.3. The FRET efficiency is a function of the separation between fluorophores r, which is related to x by $r = r_{eq} + x$ where r_{eq} is the expectation value for the equilibrium separation between fluorophores. FRET

efficiency can therefore be written in terms of x according to equation E.18:

$$E(x) = \frac{1}{1 + \left(\frac{r_{eq}+x}{R_0}\right)^6} \tag{E.18}$$

Using equation 3.6, the expectation value of FRET efficiency $\langle E \rangle$ can be found:

$$\langle E \rangle = \frac{\int E(x)z(x)dx}{\int z(x)dx} = \frac{\int \frac{e^{-U(x)/k_BT}}{1+\left(\frac{r_{eq}+x}{R_0}\right)^6}dx}{\int e^{-U(x)/k_BT}dx} \tag{E.19}$$

Appendix F: Bimolecular Association

F.1 Derivation of bimolecular association kinetics

The bimolecular association equilibrium:

$$V + T \underset{k_{\text{off}}}{\overset{k_{\text{on}}}{\rightleftharpoons}} VT \tag{F.1}$$

Can be used to generate an ODE for $[VT]$:

$$\frac{d}{dt}[VT] = k_{\text{on}}[V][T] - k_{\text{off}}[VT] \tag{F.2}$$

Also including conservation equations assuming only V and T initially:

$$[V] + [VT] = [V]_{\text{TOT}} = [V]_0 \tag{F.3}$$

$$[T] + [VT] = [T]_{\text{TOT}} = [T]_0 \tag{F.4}$$

We can rewrite equation F.2 in terms of only variable $[VT]$:

$$\frac{d}{dt}[VT] = -k_{\text{off}}[VT] + k_{\text{on}}\left([V]_0 - [VT]\right)\left([T]_0 - [VT]\right) \tag{F.5}$$

Expanding:

$$\frac{1}{k_{\text{on}}}\frac{d}{dt}[VT] = [VT]^2 - \left(\frac{k_{\text{off}}}{k_{\text{on}}} + \left([V]_0 + [T]_0\right)\right)[VT] + [V]_0[T]_0 \tag{F.6}$$

At the stationary points of $[VT]$, we find roots $r_{1,2}$:

$$r_{1,2} = \frac{\left(k_{\text{off}} + k_{\text{on}}\left([V]_0 + [T]_0\right)\right) \pm \sqrt{\left(k_{\text{off}} + k_{\text{on}}\left([V]_0 + [T]_0\right)\right)^2 - 4k_{\text{on}}^2[V]_0[T]_0}}{2k_{\text{on}}} \tag{F.7}$$

Allowing the factorisation of the expression on the RHS of equation F.6:

$$\frac{d}{dt}[VT] = k_{\text{on}}([VT] - r_1)([VT] - r_2) \tag{F.8}$$

Introduce a substitution $u = [VT] - r_2$:

$$\frac{du}{dt} = k_{\text{on}}\left(u + r_2 - r_1\right)u = k_{\text{on}}u\left(u - \frac{\sqrt{\left(k_{\text{off}} + k_{\text{on}}\left([V]_0 + [T]_0\right)\right)^2 - 4k_{\text{on}}^2[V]_0[T]_0}}{k_{\text{on}}}\right) \tag{F.9}$$

This converts the equation to the form of a Bernoulli equation:

$$\frac{du}{dt} = k_{\text{on}}u^2 - k_{\text{on}}u\beta \tag{F.10}$$

Where $\beta = \sqrt{\left([V]_0 + [T]_0 + \frac{k_{\text{off}}}{k_{\text{on}}}\right)^2 - 4[V]_0[T]_0}$. Perform a further substitution $u = \frac{1}{z}$, $\frac{du}{dt} = -z^{-2}\frac{dz}{dt}$:

$$-z^{-2}\frac{dz}{dt} + k_{\text{on}}\beta z^{-1} = k_{\text{on}}z^{-2} \tag{F.11}$$

$$\frac{dz}{dt} - k_{\text{on}}\beta z = -k_{\text{on}} \tag{F.12}$$

Equation F.12 can be solved using the integrating factor $e^{\int -k_{\text{on}}\beta dt} = e^{-k_{\text{on}}\beta t}$:

$$e^{-k_{\text{on}}\beta t}\frac{dz}{dt} - e^{-k_{\text{on}}\beta t}k_{\text{on}}\beta z = -k_{\text{on}}e^{-k_{\text{on}}\beta t} \tag{F.13}$$

$$e^{-k_{\text{on}}\beta t}z = \frac{e^{-k_{\text{on}}\beta t}}{\beta} + c \tag{F.14}$$

Restoring the origional variables by substituing back $z = \frac{1}{u}$:

$$\frac{e^{-k_{\text{on}}\beta t}}{u} = \frac{e^{-k_{\text{on}}\beta t}}{\beta} + c \tag{F.15}$$

$$e^{-k_{\text{on}}\beta t}\left(\frac{1}{u} - \frac{1}{\beta}\right) = c \tag{F.16}$$

And substituing back $u = [VT] - r_2$:

$$e^{-k_{\text{on}}\beta t}\left(\frac{1}{[VT] - r_2} - \frac{1}{\beta}\right) = c \tag{F.17}$$

At $t = 0$, i.e. $[VT](t = 0) = 0$.

$$-\left(\frac{1}{r_2} + \frac{1}{\beta}\right) = c \tag{F.18}$$

F.1.1 Solution for initial hybridization

$$e^{-k_{\text{on}}\beta t}\left(\frac{1}{[VT] - r_2} - \frac{1}{\beta}\right) = -\left(\frac{1}{r_2} + \frac{1}{\beta}\right) \tag{F.19}$$

$$[VT] = \frac{[V]_0[T]_0(1 - e^{-k_{\text{on}}\beta t})}{\beta + r_2(1 - e^{-k_{\text{on}}\beta t})} \tag{F.20}$$

F.1.2 Calculation of $[VT]_{eq}$

At equilibrium detailed balance gives the relation:

$$k_{\text{on}}[V]_{eq}[T]_{eq} = k_{\text{off}}[VT]_{eq} \tag{F.21}$$

Rearranging and eliminating $[V]_{eq}$ and $[T]_{eq}$ using relation from equations F.3 and F.4:

$$[VT]_{eq} = \frac{k_{\text{on}}}{k_{\text{off}}}\left([V]_0 - [VT]_{eq}\right)\left([T]_0 - [VT]_{eq}\right) \tag{F.22}$$

Solving for $[VT]_{eq}$:

$$[VT]_{eq} = \frac{[V]_0 + [T]_0 + \frac{k_{\text{off}}}{k_{\text{on}}} \pm \sqrt{\left([V]_0 + [T]_0 + \frac{k_{\text{off}}}{k_{\text{on}}}\right)^2 - 4[V]_0[T]_0}}{2} = r_{1,2} \tag{F.23}$$

Expanding and rearranging to find K_D:

$$K_D = \frac{[V]_0[T]_0}{[VT]_{eq}} - [VT]_{eq} - [V]_0 - [T]_0 \tag{F.24}$$

F.1.3 Half completion time for hybridization experiment

$$[VT](t = t_{1/2}) = \frac{[VT]_{eq}}{2} = \frac{r_2}{2} = \frac{[V]_0[T]_0(1 - e^{-k_{\text{on}}\beta t_{1/2}})}{\beta + r_2(1 - e^{-k_{\text{on}}\beta t_{1/2}})} \tag{F.25}$$

Rearranging:

$$\beta r_2 + r_2^2(1 - e^{-k_{\text{on}}\beta t_{1/2}}) = 2[V]_0[T]_0(1 - e^{-k_{\text{on}}\beta t_{1/2}}) \tag{F.26}$$

Collecting terms of $1 - e^{-k_{\text{on}}\beta t_{1/2}}$:

$$\beta r_2 = \left(2[V]_0[T]_0 - r_2^2\right)\left(1 - e^{-k_{\text{on}}\beta t_{1/2}}\right) \tag{F.27}$$

$$e^{-k_{\text{on}}\beta t_{1/2}} = \frac{[V]_0[T]_0}{2[V]_0[T]_0 - r_2^2} \tag{F.28}$$

Taking natural logs and solving for $t_{1/2}$:

$$t_{1/2} = \frac{1}{k_{\text{on}}\beta} \ln \frac{2[V]_0[T]_0 - [VT]_{eq}^2}{[V]_0[T]_0} \tag{F.29}$$

F.2 Dissociation kinetics

Once a steady state $[VT]$ has been achieved, an excess of unlabelled strand V_u can be used to find k_{off} in the regime where $[V] \ll [V_u]$.

$$V + T \underset{k_{\text{off}}}{\overset{k_{\text{on}}}{\rightleftharpoons}} VT \tag{F.30}$$

$$V_u + T \underset{k_{\text{off}}}{\overset{k_{\text{on}}}{\rightleftharpoons}} V_uT \tag{F.31}$$

$$V_u + VT \underset{k_{blunt}}{\overset{k_{\text{blunt}}}{\rightleftharpoons}} V + V_uT \tag{F.32}$$

$$\frac{d}{dt}[VT] = k_{\text{on}}[V][T] - k_{\text{off}}[VT] - k_{blunt}[VT][V_u] + k_{blunt}[V_uT][V] \tag{F.33}$$

The first and last terms will be negligible because of the excess of V_u.

$$\frac{d}{dt}[VT] = -k_{\text{off}}[VT] - k_{blunt}[VT][V_u] \tag{F.34}$$

If the dissociation dominates the process will be well captured by an exponential decay.

$$[VT] = [VT]_{eq}e^{-k_{\text{off}}t} \tag{F.35}$$

Here $[VT]_{eq}$ is the amount of VT after the first equilibrium before any V_u was added.

Appendix G: Buffer Conditions

G.1 Assembly buffer

- 10 mM Tris·HCl pH 8.0
- 10 mM $MgCl_2$
- 50 mM NaCl

G.2 RNAse H reaction buffer + Mg^{2+}

- 50 mM Tris·HCl pH 8.3
- 77 mM KCl
- 15 mM $MgCl_2$
- 10 mM DTT

G.3 Origami buffer

- 1× TAE pH 8.3
- 12.5 mM $MgCl_2$

G.4 Tightrope buffer

- 1× TAE pH 8.3
- 15 mM $MgCl_2$

G.5 NEB4

- 20 mM Tris-Ac pH 7.9
- 50 mM KAc
- 10 mM $MgAc_2$
- 1 mM DTT

G.6 AFM buffer

- 1× TAE pH 8.3
- 12.5 mM $MgAc_2$
- 4 mM $NiCl_2$

G.7 Anti-bleach buffer

- 1× TAE pH 8.3
- 12.5 mM $MgCl_2$
- 2% w/v Glucose
- 0.88 g/l Glucose oxidase
- 3.5×10^{-4} g/l Catalase
- 0.44 g/l Trolox
- 0.28 g/l NBA
- 0.016% COT
- 1.5% DMSO

Appendix H: Sequences

H.1 Mismatch repair sequences

Table H.1: Mismatch repair position sequences.

Name	Sequence	Name	Sequence
Va	ACCACAAACACCCACACCACTCAAACTCCAACTCAC		
Ia_0	CCACTCAAACTCCAACTCACATCA	Ta_0	TGATGTGAGTTGGAGTTTGAGTGG
Ia_5G	CCACTCAAACTCCAAGTCACATCA	Ta_5C	TGATGTGACTTGGAGTTTGAGTGG
Ia_15G	CCACTGAAACTCCAACTCACATCA	Ta_15C	TGATGTGAGTTGGAGTTTCAGTGG
Ia_17G	CCAGTCAAACTCCAACTCACATCA	Ta_17C	TGATGTGAGTTGGAGTTTGACTGG
Vb	ACCACAAACACCCACACCACTCAAACTCCACACACT		
Ib_2G	CCACTCAAACTCCACACAGTATCA	Tb_2C	TGATACTGTGTGGAGTTTGAGTGG
Vc	ACCACAAACACCCACACCACTCAAACTCCACAACTC		
Ic_3G	CCACTCAAACTCCACAAGTCATCA	Tc_3C	TGATGACTTGTGGAGTTTGAGTGG
Vd	ACCACAAACACCCACACCACTCAAACTCCACACTAC		
Id_4G	CCACTCAAACTCCACAGTACATCA	Td_4C	TGATGTACTGTGGAGTTTGAGTGG
Ve	ACCACAAACACCCACACCACTCAAACTCCACTACAC		
Ie_6G	CCACTCAAACTCCAGTACACATCA	Te_6C	TGATGTGTACTGGAGTTTGAGTGG
Vf	ACCACAAACACCCACACCACTCAAACTCACTCACAC		
If_7G	CCACTCAAACTCAGTCACACATCA	Tf_7C	TGATGTGTGACTGAGTTTGAGTGG
Vg	ACCACAAACACCCACACCACTCAAACTACTCCACAC		
Ig_8G	CCACTCAAACTAGTCCACACATCA	Tg_8C	TGATGTGTGGACTAGTTTGAGTGG
Vh	ACCACAAACACCCACACCACTCAAACACTTCCACAC		
Ih_9G	CCACTCAAACAGTTCCACACATCA	Th_9C	TGATGTGTGGAACTGTTTGAGTGG
Vi	ACCACAAACACCCACACCACTCAAAACTCTCCACAC		
Ii_10G	CCACTCAAAAGTCTCCACACATCA	Ti_10C	TGATGTGTGGAGACTTTTGAGTGG
Vj	ACCACAAACACCCACACCACTCAAACTACTCCACAC		
Ij_11G	CCACTCAAAGTACTCCACACATCA	Tj_11C	TGATGTGTGGAGTACTTTGAGTGG
Vk	ACCACAAACACCCACACCACTCAACTAACTCCACAC		
Ik_12G	CCACTCAAGTAACTCCACACATCA	Tk_12C	TGATGTGTGGAGTTACTTGAGTGG
Vl	ACCACAAACACCCACACCACTCACTAAACTCCACAC		
Il_13G	CCACTCAGTAAACTCCACACATCA	Tl_13C	TGATGTGTGGAGTTTACTGAGTGG

Table H.2: Reporter complex sequences.

Name	Sequence	Name	Sequence
rvA	ACACCACAAACACCCACA	rtA	TGAGTGGTGTGGGTGTTTGTGGTGT
rvB	ACCACAAACACCCACACC	rtB	TTTGAGTGGTGTGGGTGTTTGTGGT

Table H.3: Two-toehold sequences.

Name	Sequence	Name	Sequence
Itt_5_3G18A	TCCCACTCAAACTCCACAAGTCGAAGCT	Itt_3_3G18A	TCGAAGCCACTCAAACTCCACAAGTCCT
Ttt_3C18T	CTTCGACTTGTGGAGTTTGAGTGGCTTC	Vtt_3G18A	CCACTCAAACTCCACAAGTC
Ttt_3G18T	CTTCGAGTTGTGGAGTTTGAGTGGCTTC	Vtt_3G18C	CCCCTCAAACTCCACAAGTC
Ttt_3C18G	CTTCGACTTGTGGAGTTTGAGGGGCTTC	Vtt_3C18A	CCACTCAAACTCCACAACTC

H.2 Tightrope sequences

M13NH1 ATAACGCATATGATACTAAACAGGCTTTTTCTAGTAATTATGATTCCGGTGTTTATTCTTATTTAACGCCTT
ATTTATCACACGGTCGGTATTTCAAACCATTAAATTTAGGTCAGAAGATGAAATTAACTAAAATATATTTGAAAAAGTTTTCTCGCG
TTCTTTGTCTTGCGATTGGATTTGCATCAGCATTTACATATAGTTATATAACCCAACCTAAGCCGGAGGTTAAAAAGGTAGTCTCTC
AGACCTATGATTTTGATAAATTCACTATTGACTCTTCTCAGCGTCTTAATCTAAGCTATCGCTATGTTTTCAAGGATTCTAAGGGAA
AATTAATTAATAGCGACGATTTACAGAAGCAAGGTTATTCACTCACATATATTGATTTATGTACTGTTTCCATTAAAAAAGGTAATT
CAAATGAAATTGTTAAATGTAATTAATTTTGTTTTCTTGATGTTTGTTTCATCATCTTCTTTTGCTCAGGTAATTGAAATGAATAATT
CGCCTCTGCGCGATTTTGTAACTTGGTATTCAAAGCAATCAGGCGAATCCGTTATTGTTTCTCCCGATGTAAAAGGTACTGTTACTG
TATATTCATCTGACGTTAAACCTGAAAATCTACGCAATTTCTTTATTTCTGTTTTACGTGCAAATAATTTTGATATGGTAGGTTCTAA
CCCTTCCATTATTCAGAAGTATAATCCAAACAATCAGGATTATATTGATGAATTGCCATCATCTGATAATCAGGAATATGATGATAA
TTCCGCTCCTTCTGGTGGTTTCTTTGTTCCGCAAAATGATAATGTTACTCAAACTTTTAAAATTAATAACGTTCGGGCAAAGGATTT
AATACGAGTTGTCGAATTGTTTGTAAAGTCTAATACTTCTAAATCCTCAAATGTATTATCTATTGACGGCTCTAATCTATTAGTTGTT
AGTGCTCCTAAAGATATTTTAGATAACCTTCCTCAATTCCTTTCAACTGTTGATTTGCCAACTGACCAGATATTGATTGAGGGTTTG
ATATTTGAGGTTCAGCAAGGTGATGCTTTAGATTTTTCATTTGCTGCTGGCTCTCAGCGTGGCACTGTTGCAGGCGGTGTTAATACT
GACCGCCTCACCTCTGTTTTATCTTCTGCTGGTGGTTCGTTCGGTATTTTTAATGGCGATGTTTTAGGGCTATCAGTTCGCGCATTA
AAGACTAATAGCCATTCAAAAATATTGTCTGTGCCACGTATTCTTACGCTTTCAGGTCAGAAGGGTTCTATCTCTGTTGGCCAGAAT
GTCCCTTTTATTACTGGTCGTGTGACTGGTGAATCTGCCAATGTAAATAATCCATTTCAGACGATTGAGCGTCAAAATGTAGGTATT
TCCATGAGCGTTTTTCCTGTTGCAATGGCTGGCGGTAATATTGTTCTGGATATTACCAGCAAGGCCGATAGTTTGAGTTCTTCTACT
CAGGCAAGTGATGTTATTACTAATCAAAGAAGTATTGCTACAACGGTTAATTTGCGTGATGGACAGACTCTTTTACTCGGTGGCCTC
ACTGATTATAAAAACACTTCTCAGGATTCTGGCGTACCGTTCCTGTCTAAAATCCCTTTAATCGGCCTCCTGTTTAGCTCCCGCTCTG
ATTCTAACGAGGAAAGCACGTTATACGTGCTCGTCAAAGCAACCATAGTACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTG
GTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCG
CCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTG
ATTTGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTG
GACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGGCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGGAACCACC
ATCAAACAGGATTTTCGCCTGCTGGGGCAAACCAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGAAGGGCAATCAG
CTGTTGCCCGTCTCACTGGTGAAAAGAAAAACCACCCTGGCGCCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTA
ATGCAGCTGGCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAGGCACC
CCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATTGTGAGCGGATAACAATTTCACACAGGAAACAGCTATGACCA
TGATTACGAATTCGAGCTCGGTACCCGGGGATCCTCTAGAGTCGACCTGCAGGCATG GCTATGGTCTTACTAGGCTATGCTGGTGA
TATGCTATGGTCTTACTAGGCTATGCTGGTGATATGCTATGGTCTTACTAGGCTATGCTGGTGATATGCTATG AGCTTGGCACTGG
CCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGC
GTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGCGCTTTGCCTGGTTTCCGGCAC
CAGAAGCGGTGCCGGAAAGCTGGCTGGAGTGCGATCTTCCTGAGGCCGATACTGTCGTCGTCCCCTCAAACTGGCAGATGCACGGTT
ACGATGCGCCCATCTACACCAACGTGACCTATCCCATTACGGTCAATCCGCCGTTTGTTCCCACGGAGAATCCGACGGGTTGTTACT
CGCTCACATTTAATGTTGATGAAAGCTGGCTACAGGAAGGCCAGACGCGAATTATTTTTGATGGCGTTCCTATTGGTTAAAAAATGA
GCTGATTTAACAAAAATTTAATGCGAATTTTAACAAAATATTAACGTTTACAATTTAAATATTTGCTTATACAATCTTCCTGTTTTTG
GGGCTTTTCTGATTATCAACCGGGGTACATATGATTGACATGCTAGTTTTACGATTACCGTTCATCGATTCTCTTGTTTGCTCCAGA
CTCTCAGGCAATGACCTGATAGCCTTTGTAGATCTCTCAAAAATAGCTACCCTCTCCGGCATTAATTTATCAGCTAGAACGGTTGAA
TATCATATTGATGGTGATTTGACTGTCTCCGGCCTTTCTCACCCTTTTGAATCTTTACCTACACATTACTCAGGCATTGCATTTAAAA
TATATGAGGGTTCTAAAAATTTTTATCCTTGCGTTGAAATAAAGGCTTCTCCCGCAAAAGTATTACAGGGTCATAATGTTTTTGGTA
CAACCGATTTAGCTTTATGCTCTGAGGCTTTATTGCTTAATTTTGCTAATTCTTTGCCTTGCCTGTATGATTTATTGGATGTTAATGC
TACTACTATTAGTAGAATTGATGCCACCTTTTCAGCTCGCGCCCCAAATGAAAATATAGCTAAACAGGTTATTGACCATTTGCGAAA
TGTATCTAATGGTCAAACTAAATCTACTCGTTCGCAGAATTGGGAATCAACTGTTATATGGAATGAAACTTCCAGACACCGTACTTT
AGTTGCATATTTAAAACATGTTGAGCTACAGCATTATATTCAGCAATTAAGCTCTAAGCCATCCGCAAAAATGACCTCTTATCAAAA
GGAGCAATTAAAGGTACTCTCTAATCCTGACCTGTTGGAGTTTGCTTCCGGTCTGGTTCGCTTTGAAGCTCGAATTAAAACGCGATA
TTTGAAGTCTTTCGGGCTTCCTCTTAATCTTTTTGATGCAATCCGCTTTGCTTCTGACTATAATAGTCAGGGTAAAGACCTGATTTTT
GATTTATGGTCATTCTCGTTTTCTGAACTGTTTAAAGCATTTGAGGGGGATTCAATGAATATTTATGACGATTCCGCAGTATTGGAC
GCTATCCAGTCTAAACATTTTACTATTACCCCCTCTGGCAAAACTTCTTTTGCAAAAGCCTCTCGCTATTTTGGTTTTTATCGTCGTC
TGGTAAACGAGGGTTATGATAGTGTTGCTCTTACTATGCCTCGTAATTCCTTTTGGCGTTATGTATCTGCATTAGTTGAATGTGGTA
TTCCTAAATCTCAACTGATGAATCTTTCTACCTGTAATAATGTTGTTCCGTTAGTTCGTTTTATTAACGTAGATTTTTCTTCCCAACG
TCCTGACTGGTATAATGAGCCAGTTCTTAAAATCGCATAAGGTAATTCACAATGATTAAAGTTGAAATTAAACCATCTCAAGCCCAA
TTTACTACTCGTTCTGGTGTTTCTCGTCAGGGCAAGCCTTATTCACTGAATGAGCAGCTTTGTTACGTTGATTTGGGTAATGAATAT
CCGGTTCTTGTCAAGATTACTCTTGATGAAGGTCAGCCAGCCTATGCGCCTGGTCTGTACACCGTTCATCTGTCCTCTTTCAAAGTT
GGTCAGTTCGGTTCCCTTATGATTGACCGTCTGCGCCTCGTTCCGGCTAAGTAACATGGAGCAGGTCGCGGATTTCGACACAATTTA
TCAGGCGATGATACAAATCTCCGTTGTACTTTGTTTCGCGCTTGGTATAATCGCTGGGGGTCAAAGATGAGTGTTTTAGTGTATTCT
TTTGCCTCTTTCGTTTTAGGTTGGTGCCTTCGTAGTGGCATTACGTATTTTACCCGTTTAATGGAAACTTCCTCATGAAAAAGTCTTT
AGTCCTCAAAGCCTCTGTAGCCGTTGCTACCCTCGTTCCGATGCTGTCTTTCGCTGCTGAGGGTGACGATCCCGCAAAAGCGGCCTT
TAACTCCCTGCAAGCCTCAGCGACCGAATATATCGGTTATGCGTGGGCGATGGTTGTTGTCATTGTCGGCGCAACTATCGGTATCAA
GCTGTTTAAGAAATTCACCTCGAAAGCAAGCTGATAAACCGATACAATTAAAGGCTCCTTTTGGAGCCTTTTTTTTGGAGATTTTCA
ACGTGAAAAAATTATTATTCGCAATTCCTTTAGTTGTTCCTTTCTATTCTCACTCCGCTGAAACTGTTGAAAGTTGTTTAGCAAAATC
CCATACAGAAAATTCATTTACTAACGTCTGGAAAGACGACAAAACTTTAGATCGTTACGCTAACTATGAGGGCTGTCTGTGGAATGC
TACAGGCGTTGTAGTTTGTACTGGTGACGAAACTCAGTGTTACGGTACATGGGTTCCTATTGGGCTTGCTATCCCTGAAAATGAGGG
TGGTGGCTCTGAGGGTGGCGGTTCTGAGGGTGGCGGTTCTGAGGGTGGCGGTACTAAACCTCCTGAGTACGGTGATACACCTATTC
CGGGCTATACTTATATCAACCCTCTCGACGGCACTTATCCGCCTGGTACTGAGCAAAACCCCGCTAATCCTAATCCTTCTCTTGAGGA
GTCTCAGCCTCTTAATACTTTCATGTTTCAGAATAATAGGTTCCGAAATAGGCAGGGGGCATTAACTGTTTATACGGGCACTGTTAC
TCAAGGCACTGACCCCGTTAAAACTTATTACCAGTACACTCCTGTATCATCAAAAGCCATGTATGACGCTTACTGGAACGGTAAATT
CAGAGACTGCGCTTTCCATTCTGGCTTTAATGAGGATTTATTTGTTTGTGAATATCAAGGCCAATCGTCTGACCTGCCTCAACCTCC
TGTCAATGCTGGCGGCGGCTCTGGTGGTGGTTCTGGTGGCGGCTCTGAGGGTGGTGGCTCTGAGGGTGGCGGTTCTGAGGGTGGCG
GCTCTGAGGGAGGCGGTTCCGGTGGTGGCTCTGGTTCCGGTGATTTTGATTATGAAAAGATGGCAAACGCTAATAAGGGGGCTATG
ACCGAAAATGCCGATGAAAACGCGCTACAGTCTGACGCTAAAGGCAAACTTGATTCTGTCGCTACTGATTACGGTGCTGCTATCGAT
GGTTTCATTGGTGACGTTTCCGGCCTTGCTAATGGTAATGGTGCTACTGGTGATTTTGCTGGCTCTAATTCCCAAATGGCTCAAGTC
GGTGACGGTGATAATTCACCTTTAATGAATAATTTCCGTCAATATTTACCTTCCCTCCCTCAATCGGTTGAATGTCGCCCTTTTGTCT
TTGGCGCTGGTAAACCATATGAATTTTCTATTGATTGTGACAAAATAAACTTATTCCGTGGTGTCTTTGCGTTTCTTTTATATGTTG
CCACCTTTATGTATGTATTTTCTACGTTTGCTAACATACTGCGTAATAAGGAGTCTTAATCATGCCAGTTCTTTTGGGTATTCCGTTA
TTATTGCGTTTCCTCGGTTTCCTTCTGGTAACTTTGTTCGGCTATCTGCTTACTTTTCTTAAAAAGGGCTTCGGTAAGATAGCTATT
GCTATTTCATTGTTTCTTGCTCTTATTATTGGGCTTAACTCAATTCTTGTGGGTTATCTCTCTGATATTAGCGCTCAATTACCCTCTG
ACTTTGTTCAGGGTGTTCAGTTAATTCTCCCGTCTAATGCGCTTCCCTGTTTTTATGTTATTCTCTCTGTAAAGGCTGCTATTTTCAT
TTTTGACGTTAAACAAAAAATCGTTTCTTATTTGGATTGGGATAAATAATATGGCTGTTTATTTTGTAACTGGCAAATTAGGCTCTG
GAAAGACGCTCGTTAGCGTTGGTAAGATTCAGGATAAAATTGTAGCTGGGTGCAAAATAGCAACTAATCTTGATTTAAGGCTTCAAA
ACCTCCCGCAAGTCGGGAGGTTCGCTAAAACGCCTCGCGTTCTTAGAATACCGGATAAGCCTTCTATATCTGATTTGCTTGCTATTG
GGCGCGGTAATGATTCCTACGATGAAAATAAAAACGGCTTGCTTGTTCTCGATGAGTGCGGTACTTGGTTTAATACCCGTTCTTGGA
ATGATAAGGAAAGACAGCCGATTATTGATTGGTTTCTACATGCTCGTAAATTAGGATGGGATATTATTTTTCTTGTTCAGGACTTAT
CTATTGTTGATAAACAGGCGCGTTCTGCATTAGCTGAACATGTTGTTTATTGTCGTCGTCTGGACAGAATTACTTTACCTTTTGTCG
GTACTTTATATTCTCTTATTACTGGCTCGAAAATGCCTCTGCCTAAATTACATGTTGGCGTTGTTAAATATGGCGATTCTCAATTAAG
CCCTACTGTTGAGCGTTGGCTTTATACTGGTAAGAATTTGT

M13NH2 Same as M13NH1 but with sub-sequence marked in red replaced with CATACCCTCATCCAA
CTCAACATCACGCAGAAAACCCCACTAGACAGCACATTCCTACAATCCCTCTCGGACCAACTCAAATACCATCAACTCACTCACTCA.

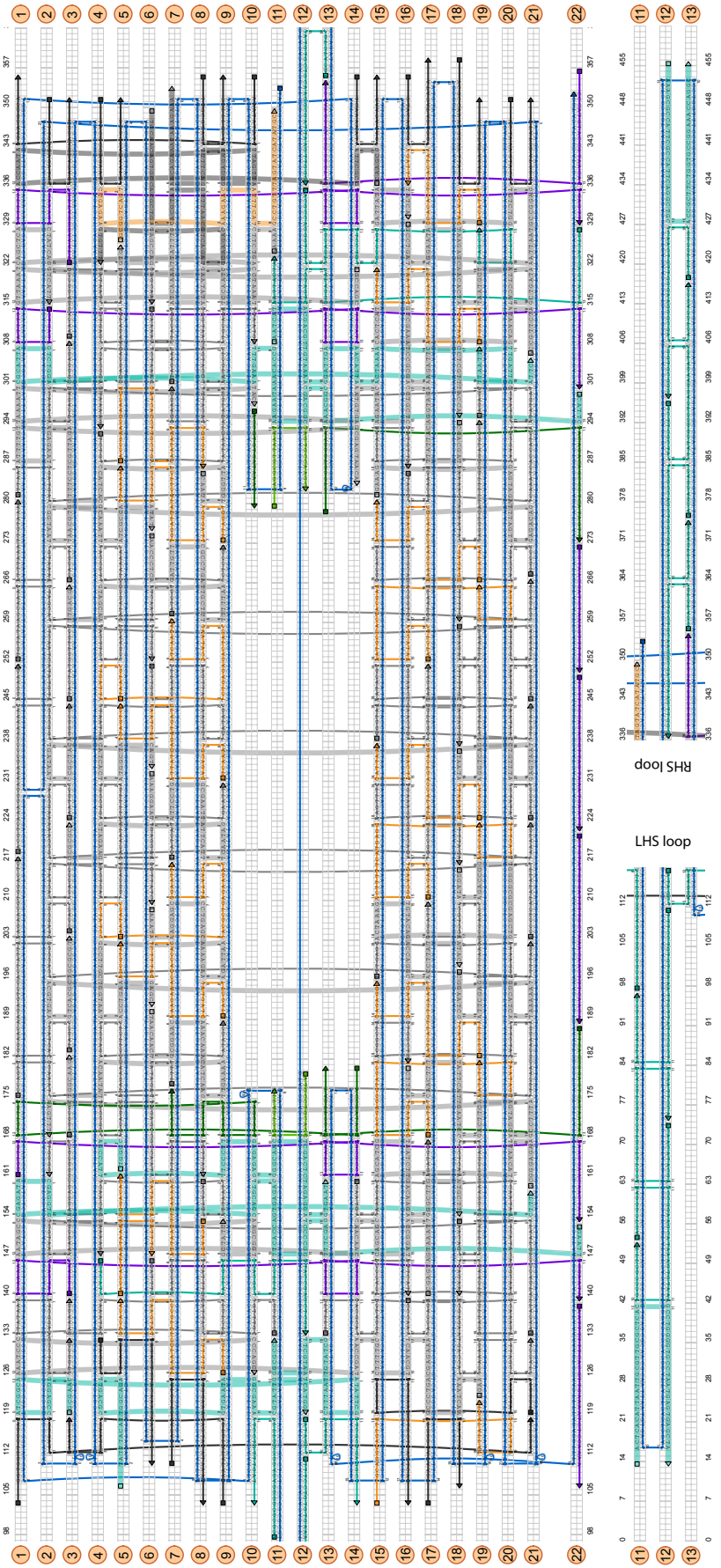

Figure H.1: Tightrope origami layout for 102 nt across the gap with M13NH1 or M13NH2 scaffold.

Table H.4: Tightrope staples for universal tension.

Name	Sequence
1[175]-16[181]	TCCAGTTCCAAATCTTTCAACGCAAGGAATTCTACACGGCTACTGGCTGA
1[218]-21[244]	AACGTCAAAGGGAAAAGCCCCAAAAACCCCCGGTGTGTAGG
1[252]-21[266]	TTGTATAAGCAAATCTAGCATGGTGAGAA
1[281]-16[286]	GTTAATCGATGAATCAAATCACCATCATGCGAACAGGCAAACATGTTAC
11[133]-3[132]	GGGTGGTGGTTTGCTTATAAAATCGGAACCCTAAAGTGGCGA
11[324]-16[330]	AAAAACTTCTGACAGCTCACTATCAGAATTAATGCCGGAGTGTCTGGGATTATATATCAT
14[160]-17[167]	ACAATGAGAAAGACCCAATAAGAGGAAGCCCGAAACCCTCAAAGATGGT
14[321]-5[325]	AGCCAGCGTTAAATCCTAAATTCCAATCGAATATAATCAAAATTATTTGGAGCCAGGTGA
14[133]-18[140]	GTTTATCCACGCATGGGATCGGCAAAGAGGATTGCATCAAAATCAGAAA
16[139]-21[132]	ACCCAAACACCAGAACGAGAATGACCATGCAAAGCATTAGCAGTACCAA
16[180]-2[168]	CCTTCACAACTTTTATTCATTGAATCCGACTTCAAGTAGCAGCCTTTAAAGTTTT
16[285]-19[294]	TTAGCCAACAACAAAAAACCAAAATAGGAGGTCATTTTTGCGGATGGCT
16[329]-2[315]	CGCCTGATAGGAATATAAGAGCAACACTATAATGCTATTCCATGCTGATAGTCATTG
15[196]-17[209]	TTTGAGGAAAAGGTCTTCAAAGCGAACCGCGTCCAGCGATTT
15[238]-17[251]	CGGGTAACAATAACGATTAGAGAGTACCGTTTTGCGAAGAAA
15[281]-6[275]	CGAAAGGAGTAGATTTAGTTGAGACAGCGGTAATTAACGTGGCAAATTATCGCCAATATC
18[153]-16[140]	AAACAGTAGATTAAATCATACAGGCAAGTCACCCTATTCATT
18[197]-3[203]	GCGGAATCGATTCGAGGGCATCATAAAAATACGTGAAGTCACGCTGCGCGTA
18[215]-18[198]	TGGATAAGACCGGTGGGGCGCGAGCTGACTAAAGTGTACAGACCAGGCACCTTATATACT
18[236]-6[233]	GTAATAAGGTCAGCTGTTTAGAGTAATTGATAATCCGCTACCATCACTATGGCTACAGTT
18[258]-1[251]	AAAAGAATTTAATTATTTCGCTCAAAAGGTCAATCATATGTAAGGAAGA
18[294]-10[297]	ACGACGATTTATTACAGGTAGACCTGCTCAGAATACAATGTGATAAAATTTACC
17[140]-9[153]	ACGAGTAGTAAATTACCGGATCAGCAGCCAACAACAATAGCCAAGCGGTGCTATTA
19[122]-10[126]	AGAAAAATCAAGCCCTGACGAGAAATCAACGTCTTTTGCAACCGATAAATCCCCCCAGCA
19[295]-4[293]	TAGAGCCCAATTCATATGATGCAAACAATCAGAGGTAAAAGA
4[292]-7[300]	GTCTGTTATAACTATATGTATTTACATGGAAGGG
21[133]-4[147]	AAACATTATGACCCTGTAATACCGTAAAAAAGCGAGAAAAAC
21[159]-18[154]	GCGGGAGAATTAACATAGCATCGTAATCTTGACAAGAGGGCTTGATGCTTT
21[203]-6[209]	ACCCTCATATATTTCTATCAGACACCCGAACTCAATGGCACAGGTTATCT
21[224]-18[216]	AATGCCTGCTATATAAGTTTCTGAAAGAGGACAGATGGCTCAGTTTAGAC
21[245]-18[237]	TAAAGATAAATGGTAATACGTGGGAACCGAACTGAACGTTGGCAGAGGGG
21[267]-18[259]	AGGCCGTGACCATGGCACCAGGCGCAGACGGTCAGTTAATACTTTTGC
21[306]-18[295]	ACCGTTCTAATAACAGCACTCATGTCGAAATCCGCGAAAGATTCCCTCGTTTACCAG
3[168]-8[161]	AGGGCGCTACCGCCAACAGAGTAGAGCCGTCAATATTTGAGTAATTAAT
3[183]-21[202]	GTAGCGCCATCACTGGAACAAGAGTCCACTATTAGCCCACTTTTTAGA
3[204]-1[217]	ACCACCGGCGATGAAGAACGTGGACTCC
3[224]-21[223]	TAATGCGCAGCGAAAAACCGTTAAATGC
3[245]-6[252]	GTACTATGGTTGCTTACTTCTGCGAACTTCTGGTC
3[266]-1[280]	GCACGTACGTAAAAATTTAAATTGTAAAC
3[309]-6[315]	AAACAGGAGGCCGTATAATCAACAGAGCAGCAAA
6[146]-3[139]	GGATTTAGAACGTTTCATTTGAATTACCTTCTGTAGAAATACGAAAGGA
6[190]-3[182]	CTAACAACTGCGGAACAAACAAACATCAAGTTGAAAAACAATATTGGCAAGT
6[208]-6[191]	AAAATAAAGGAGCCCTGAGCAAAAGAAAGATTAAGCCTTGCTGGTAATACCTGAAGAGCA
6[232]-3[223]	GAAAGGAATATTCCTGGCGAATTATTCATTTCAATAGGTAGAAGCCGCGCT
6[251]-3[244]	AGTTGGCGCAATTCAAGTTACAAAATCGTCATAGGGTAATAAAGGGCGC
6[274]-3[265]	AAACCCTCAGATTGTTCGCCTGATTGCTTTTTTTTAAGTAGCAATTGACGA
6[314]-3[308]	TGAAAAACTACCATCAGTAACAGTACCTAATGCTGCCACCGACGGGAGCT
8[160]-2[161]	TACATTTATTAATTGCAACAGAAGGAGCGGGCGCTTTGGGGT
8[285]-14[283]	ACAATATGGGTTACCATCACCTTTCCTCGTTAGAAGAGAATATTTTGTGCGAGTAACAA
7[177]-9[188]	ATTTTAATAGATATAGAACCCTTCTGATCCAGACATAGCG
7[217]-9[230]	TCATCATTGAGGAAGACAATATTTTTGATGCCTGATGAATTT
7[260]-9[272]	AATCCTATCAATAGATAGCCCTAAAACAACCGTTCCTCCGG
7[301]-10[308]	TTAGAACTCTAAAGCCAGCAGAAGATAAAGTGAGGATGCAAATTAATGG
2[351]-3[351]	TTTTTTTGAGAGATCTAGACAGGAACGGTTTT
1[103]-21[118]	TTTTGGTGGTTCCGAACCCGATTAAGCTAA
14[355]-15[355]	TTTTTCAAAAATACGAAACAAATTTT
16[355]-17[358]	TTTTGTACAACGGTGCAGATACATATTTT
15[336]-7[353]	CCAAGCGATTCGCGACCAATATTAGTTAGCGAGAATTTTCAGAACAGAAATAAAGAAATT
18[358]-19[351]	TTTTACGCCAAAAGGAATTACGATTTAAATATGCATTTT
17[103]-18[106]	TTTTTGAATAAGGCTTAAATCAGGTTTTT
18[351]-21[351]	TTTTACTAAAGTACGGAGGGTAGCTATTTTTT
4[132]-6[110]	CTACATTTTCACCATAGACTTTACAAACAATTCTTTT
4[351]-5[351]	TTTTTACGCCAGAATCATTAACACCGCCTTTT
21[119]-16[103]	ATCGGTTAAATTAAAAGGCCGAACAAAGCTGCTCATTCAGTTTT
6[349]-4[322]	TTTGCAACAGTGCCACGCTGACACGTAAGTTTAACGTCAGATGCAAGACGTGTTTT
8[355]-9[355]	TTTTTTGCGTAGAAACTTTTTCTTTT
7[110]-8[103]	TTTTGACAACTCGTATATGGAAACAGTACATAAATTTTT
10[355]-1[355]	TTTTAAATATATTGGAACGCCATTTT

9[103]-3[118]	TTTTCAATATATGTGATCAATCGGAAAGCC
2[314]-22[300]	CCTGAGAAATTTTTTTTCATCCCGTAATACTACAACGCCTGTA
1[161]-22[153]	GGGTTGATGCGCCGTCTTAAAAAAGGAACAACTAAA
22[138]-22[106]	TAATTTTTTCACGTTGAAAATCTCCAAAATTTT
22[221]-22[188]	TCTGTATGGGATTTTGCTAAACAACTTTCAACAG
22[249]-22[222]	GTCTTTCCAGACGTTAGTAAATGAATTT
22[272]-22[250]	GCGTAACGATCTAAAGTTTTGTC
22[356]-13[354]	TTTTCCCATGTACCGTAACACGATGGGCGCATCGTAACCG
3[140]-22[139]	AGGGAAGGCACTAATCAAAAGCATCGCCAGCTTGCGCGAATAA
3[322]-22[329]	ATTAAAGGGATTTTACAAAGGTTTTTTATCTGGCCTGGTGTATGAGTTT
15[103]-19[121]	TTTTTGCAGGGAGTTAGCAATAATATTATAGTC
17[168]-19[181]	TTAATTTTCAAGAGGAACGAGGGTAGCATAATAGTAATATCG
17[210]-19[223]	TAAGAACTGAACGGACTTTTTCATGAGGTTTCATTAAGCAAA
17[252]-19[265]	AATCTACATCATAAAATGCCACTACGAATAGATACGCTCCTT
19[182]-15[195]	CGTTTTATCATAAAAATCATTGTGAATTGCATAGGCAGAGGC
19[224]-15[237]	CTCCAACGTAAAATTTATACCAGTCAGGCCAACTTCATTAAA
19[266]-15[280]	TTGATAACGAGAGGAAACGAACTAACGGGGAACGAACCTAAAA
19[308]-15[321]	CTGAATATCATAACATCAGTTGAGATTTAAATTGTCTTTGAC
19[329]-14[336]	AACATGTGGCATAGCCACATTCAACTAAAGATTTG
5[140]-6[147]	GACCAGTAATAAAACATTTGA
5[203]-7[216]	GAATACGACTATCGGACGCTGAGAAGAGTCAATTAGGAATTA
5[245]-7[259]	TTAATGCTTGATTATCTGAGAGACTACCGAATACCATCAATAT
5[287]-8[286]	TACCGAACGAACCACATCACCTTGCTGAGAATAATCGGGAGAA
5[326]-11[349]	GGCGGTCAGTCTGAGAAAAAGAACATTTCATGCCTGTTTAGTATCATATGCG
8[153]-5[161]	AACAATTATTAATTTTAAAAGGATAATAGGGACATT
9[126]-5[139]	ACCTTGCTTTTTTTATAAATCCTTTGCCCGAAGTATGTCACAC
9[189]-5[202]	ATAGCTTGATGATGAAAGAAACCACCAGTCTTTAGAGCGTAA
9[231]-5[244]	ATCAAAACGCAGAGATTATCAGATGATGAAATCAATTAGTCT
9[273]-5[286]	CTTAGGTACGGATTTGGATTATACTTCTACCTCAATTAAAAA
14[180]-7[176]	TTTTACCGATAGTGTGTTGTTGGCCCTAGAATCCAAAACAAAACATTATC
13[278]-22[273]	TTTTCTCCGTGGGAACCACAGACAGCCCTCATAGTTA
22[187]-13[180]	TTTCAGCGGAGTGAGAATAGCAGCTTGATTTTT
10[296]-10[279]	GACCGTGTGATAAATTTT
l_12[118]-14[103]	CATAGCTAAGGAGCGTCGCTGAGGCT
l_12[179]-11[176]	CATGCCTGCAGGAGCTGATTG
l_4[146]-12[133]	GCTCATGAATCGTCCCACGCTTTTTCTTGTACCGAGCTCGAA
l_22[152]-21[158]	GGAATTTTTCGAGTCCCCGGTTCACCAGTGAGACTTGCAGCCGAGATACGAGGTGCTTTT
l_3[119]-11[132]	GGCGAACGGGAGCCATCGGCAATATTCGCTTTAATTGTATCGTTCGTAAGGCGCCA
l_5[162]-13[160]	CTGGCCAGCCATTTTCCCTTGAGAGAGGGGCAACTCGACTCTAGAGGAGTGAATT
l_ds_12[111]-11[96]	GTTTCCTGTGTGAAATTGTTATCCGCTCAATGAATCGGCCA
l_ds_12[73]-12[74]	ACAACATACGAACCTGTCGTGCCAGCTGCATTACAATTCCAC
l_ds_11[13]-12[13]	ACTCACATTAATTGCGTTGCGCTCACGCCTGGGGTGCCTAATGAGTGAGCTA
l_ds_11[53]-11[52]	CAGTCGGGAAGCCGGAAGCATAAAGTGTAAATGCCCGCTTTC
l_ds_11[97]-10[103]	ACGCGCGGGGAGAGGCGGTTTGATCCTGTTTGAT
l_ds_5[106]-12[119]	TATTTACATTGGCAGATTGACGCGTGAATAGGCGAAACGTATTGTCATGGT

Table H.5: Tightrope staples for immobilisation.

Name	Sequence
5'_Biotin	AGGTGGTGGAGAGGAG
3'_Biotin	TGCGATATTGGATTAATT
15[103]-19[121]_Immob	TTTTTGCAGGGAGTTAGCAATAATATTATAGTCTTCTCCTCTCCACCACCT
17[168]-19[181]_Immob	TTAATTTTCAAGAGGAACGAGGGTAGCATAATAGTAATATCGTTCTCCTCTCCACCACCT
17[210]-19[223]_Immob	TAAGAACTGAACGGACTTTTTCATGAGGTTTCATTAAGCAAATTCTCCTCTCCACCACCT
17[252]-19[265]_Immob	AATCTACATCATAAAATGCCACTACGAATAGATACGCTCCTTTTCTCCTCTCCACCACCT
Immob_19[182]-15[195]	TTAATCCAATATCGCATTCGTTTTATCATAAAAATCATTGTGAATTGCATAGGCAGAGGC
Immob_19[224]-15[237]	TTAATCCAATATCGCATTCTCCAACGTAAAATTTATACCAGTCAGGCCAACTTCATTAAA
Immob_19[266]-15[280]	TTAATCCAATATCGCATTTGATAACGAGAGGAAACGAACTAACGGGGAACGAACCTAAAA
Immob_19[308]-15[321]	TTAATCCAATATCGCATTCTGAATATCATAACATCAGTTGAGATTTAAATTGTCTTTGAC
Immob_19[329]-16[336]	TTAATCCAATATCGCATTAACATGTGGCATAGCCACATTCAACTAAAGATTTG
Immob_5[140]-6[147]	TTAATCCAATATCGCATTGACCAGTAATAAAACATTTGA
Immob_5[203]-7[216]	TTAATCCAATATCGCATTGAATACGACTATCGGACGCTGAGAAGAGTCAATTAGGAATTA
Immob_5[245]-7[259]	TTAATCCAATATCGCATTTAATGCTTGATTATCTGAGAGACTACCGAATACCATCAATAT
Immob_5[287]-8[286]	TTAATCCAATATCGCATTACCGAACGAACCACATCACCTTGCTGAGAATAATCGGGAGAA
Immob_5[326]-11[349]	TTAATCCAATATCGCATTGGCGGTCAGTCTGAGAAAAAGAACATTTCATGCCTGTTTAGTATCATATGCG
8[153]-5[161]_Immob	AACAATTATTAATTTTAAAAGGATAATAGGGACATTTTCTCCTCTCCACCACCT
9[126]-5[139]_Immob	ACCTTGCTTTTTTTATAAATCCTTTGCCCGAAGTATGTCACACTTCTCCTCTCCACCACCT
9[189]-5[202]_Immob	ATAGCTTGATGATGAAAGAAACCACCAGTCTTTAGAGCGTAATTCTCCTCTCCACCACCT
9[231]-5[244]_Immob	ATCAAAACGCAGAGATTATCAGATGATGAAATCAATTAGTCTTTCTCCTCTCCACCACCT
9[273]-5[286]_Immob	CTTAGGTACGGATTTGGATTATACTTCTACCTCAATTAAAAATTCTCCTCTCCACCACCT

Table H.6: Tension setting staples for M13mp18.

Name	Sequence
NOINS_12[335]19[328]	TCAGGCTGCGCAACGTCACGTTTCCTGTCCCCAGCAAGTTTCGTAGCTC
NOINS_12[403]13[403]	TTTGACTCCAGCCAGCTTGTATCGGCCTCAGGAATTT
NOINS_11[279]12[282]	TTTAAGGCGTTAAATCTTCGCTATTAC
NOINS_11[308]19[307]	GAATCATAAGGGCGATCGGTGGGATTGAAACATTAACTAAAATTGATTCTTAATTG
NOINS_13[355]12[336]	TGCATCTGCGGAAACCAGGCAAAGCGCCATTCGCCAT
NOINS_13[375]13[374]	GACGACGACATCCGGCACCGCTTCTGGTGCCCAGTTTGAGGG
NOINS_22[299]21[305]	GCATTCAAACGGCCGGGCCTAAGAATAAACACCGTTTGAAACGCATTAGTCTGGAATTCA
NOINS_22[328]11[323]	CGTCACCAGTACAAGGGATAGTGTTGGGAATTACTAG
20r_NOINS_12[335]19[328]	GGCAAAGCGCCATTGTCACGTTTCCTGTCCCCAGCAAGTTTCGTAGCTC
20r_NOINS_11[279]12[282]	TTTAAGGCGTTAAATAGGGCGATCGGT
20r_NOINS_11[308]19[307]	GAATCATCAGGCTGCGCAACTGGATTGAAACATTAACTAAAATTGATTCTTAATTG
20r_NOINS_13[355]12[336]	TGCATCTGCCCGGCACCGCTTCTGGTGCCGGAAACCA
20r_NOINS_13[375]13[374]	GACGACGACAGATCGCACTCCAGCCAGCTTTCAGTTTGAGGG
20r_NOINS_22[299]21[305]	GCATTCAAACGGCGTTGGGAAAGAATAAACACCGTTTGAAACGCATTAGTCTGGAATTCA
20r_NOINS_22[328]11[323]	CGTCACCAGTACAAGGGATAGCGCCATTAATTACTAG
42r_NOINS_12[335]19[328]	CCGCTTCTGGTGCCGTCACGTTTCCTGTCCCCAGCAAGTTTCGTAGCTC
42r_NOINS_12[377]13[384]	TATCGGCCTCAGGACAGTTTGAGGGGACGATTTTT
42r_NOINS_11[279]12[282]	TTTAAGGCGTTAAATTCAGGCTGCGCA
42r_NOINS_11[308]19[307]	GAATCATAGGCAAAGCGCCATGGATTGAAACATTAACTAAAATTGATTCTTAATTG
42r_NOINS_13[355]12[336]	TGCATCTGCAGATCGCACTCCAGCCAGCTTTCCGGCA
42r_NOINS_22[299]21[305]	GCATTCAAACGGCTCGCCATAAGAATAAACACCGTTTGAAACGCATTAGTCTGGAATTCA
42r_NOINS_22[328]11[323]	CGTCACCAGTACAAGGGATAGGGAAACCAATTACTAG

Table H.7: Tension setting staples for M13NH1 and M13NH2.

Name	Sequence
r_12[335]-19[328]	AACGCCAGGGTTTTGTCACGTTTCCTGTCCCCAGCAAGTTTCGTAGCTC
r_12[395]-12[396]	CTCTTCGCTATGTATCGGCCTCAGGAAGATCGCGGTGCGGGC
r_12[455]-13[455]	AAAGCGCCATTCGCCATTCAGGCTGCACCGCTTCTGGTGCCGGAAACCAGGC
r_11[279]-12[282]	TTTAAGGCGTTAAATCAGTGCCAAGCT
r_11[308]-19[307]	GAATCATACGACGTTGTAAAAGGATTGAAACATTAACTAAAATTGATTCTTAATTG
r_13[355]-12[336]	TGCATCTGCGATGTGCTGCAAGGCGATTAAGTTGGGT
r_13[375]-13[374]	GACGACGACATACGCCAGCTGGCGAAAGGGGCAGTTTGAGGG
r_13[417]-13[416]	GCTTTCCGGCGCAACTGTTGGGAAGGGCGATCACTCCAGCCA
r_28[299]-21[305]	GCATTCAAACGGCCGACGGCAAGAATAAACACCGTTTGAAACGCATTAGTCTGGAATTCA
r_28[328]-11[323]	CGTCACCAGTACAAGGGATAGCCCAGTCAATTACTAG
20r_12[335]-19[328]	GCAAGGCGATTAAGGTCACGTTTCCTGTCCCCAGCAAGTTTCGTAGCTC
20r_12[395]-12[396]	GAAGGGCGATCGTATCGGCCTCAGGAAGATCGCAACTGTTGG
20r_11[279]-12[282]	TTTAAGGCGTTAAATCGACGTTGTAAA
20r_11[308]-19[307]	GAATCATACGCCAGGGTTTTCGGATTGAAACATTAACTAAAATTGATTCTTAATTG
20r_13[355]-12[336]	TGCATCTGCACGCCAGCTGGCGAAAGGGGGATGTGCT
20r_13[375]-13[374]	GACGACGACAGGTGCGGGCCTCTTCGCTATTCAGTTTGAGGG
20r_13[417]-13[416]	GCTTTCCGGCCCATTCGCCATTCAGGCTGCGCACTCCAGCCA
20r_22[299]-21[305]	GCATTCAAACGGCCCAGTCAAAGAATAAACACCGTTTGAAACGCATTAGTCTGGAATTCA
20r_22[328]-11[323]	CGTCACCAGTACAAGGGATAGTTGGGTAAATTACTAG
42r_12[335]-19[328]	GCTGGCGAAAGGGGGTCACGTTTCCTGTCCCCAGCAAGTTTCGTAGCTC
42r_12[395]-12[396]	ATTCAGGCTGCGTATCGGCCTCAGGAAGATCGGCCATTCGCC
42r_11[279]-12[282]	TTTAAGGCGTTAAATAACGCCAGGGTT
42r_11[308]-19[307]	GAATCATTGCAAGGCGATTAAGGATTGAAACATTAACTAAAATTGATTCTTAATTG
42r_13[355]-12[336]	TGCATCTGCCGGTGCGGGCCTCTTCGCTATTACGCCA
42r_13[375]-13[374]	GACGACGACAGCAACTGTTGGGAAGGGCGATCAGTTTGAGGG
42r_13[417]-13[416]	GCTTTCCGGCGTGCCGGAAACCAGGCAAAGCCACTCCAGCCA
42r_22[299]-21[305]	GCATTCAAACGGCGTTGGGTAAGAATAAACACCGTTTGAAACGCATTAGTCTGGAATTCA
42r_22[328]-11[323]	CGTCACCAGTACAAGGGATAGGATGTGCAATTACTAG
-8r_12[335]19[328]	GGTTTTCCCAGTCAGTCACGTTTCCTGTCCCCAGCAAGTTTCGTAGCTC
-8r_12[395]12[396]	TATTACGCCAGGTATCGGCCTCAGGAAGATCGGCCTCTTCGC
-8r_12[437]12[438]	TGCGCAACTGTACCGCTTCTGGTGCCGGAAACCCATTCAGGC
-8r_12[459]13[459]	TTTCGCCATTCGCAGGCAAAGTTT
-8r_11[308]19[307]	GAATCATGTAAAACGACGGCCGGATTGAAACATTAACTAAAATTGATTCTTAATTG
-8r_13[355]12[336]	TGCATCTGCGCAAGGCGATTAAGTTGGGTAACGCCAG
-8r_13[375]13[374]	GACGACGACACTGGCGAAAGGGGGATGTGCTCAGTTTGAGGG
-8r_13[417]13[416]	GCTTTCCGGCTGGGAAGGGCGATCGGTGCGGCACTCCAGCCA
-8r_22[299]21[305]	GCATTCAAACGGCAGTGCCAAAGAATAAACACCGTTTGAAACGCATTAGTCTGGAATTCA
-8r_22[328]11[323]	CGTCACCAGTACAAGGGATAGCGACGTTAATTACTAG
-8r_11[279]12[282]_M13NH1	TTTAAGGCGTTAAATAGCTCATAGCAT
-8r_11[279]12[282]_M13NH2	TTTAAGGCGTTAAATAGCTTGAGTGAG

-14r_12[335]19[328]	CCCAGTCACGACGTGTCACGTTTCCTGTCCCCAGCAAGTTTCGTAGCTC
-14r_12[395]12[396]	GCCAGCTGGCGGTATCGGCCTCAGGAAGATCGTCGCTATTAC
-14r_12[437]12[438]	ACTGTTGGGAAACCGCTTCTGGTGCCGGAAACAGGCTGCGCA
-14r_12[462]13[462]	TTTCATTCGCCATTCCAGGCAAAGCGCTTT
-14r_11[308]19[307]	GAATCATCGACGGCCAGTGCCGGATTGAAACATTAACTAAAATTGATTCTTAATTG
-14r_13[355]12[336]	TGCATCTGCCGATTAAGTTGGGTAACGCCAGGGTTTT
-14r_13[375]13[374]	GACGACGACAAAAGGGGGATGTGCTGCAAGGCAGTTTGAGGG
-14r_13[417]13[416]	GCTTTCCGGCGGGCGATCGGTGCGGGCCTCTCACTCCAGCCA
-14r_22[328]11[323]	CGTCACCAGTACAAGGGATAGTGTAAAAAATTACTAG
-14r_22[299]21[305]_M13NH1	GCATTCAAACGGCAAGCTCAAAGAATAAACACCGTTTGAAACGCATTAGTCTGGAATTCA
-14r_22[299]21[305]_M13NH2	GCATTCAAACGGCAAGCTTGAAGAATAAACACCGTTTGAAACGCATTAGTCTGGAATTCA
-14r_11[279]12[282]_M13NH1	TTTAAGGCGTTAAATTAGCATATCACC
-14r_11[279]12[282]_M13NH2	TTTAAGGCGTTAAATAGTGAGTGAGTT
-18r_12[335]19[328]	GTCACGACGTTGTAGTCACGTTTCCTGTCCCCAGCAAGTTTCGTAGCTC
-18r_12[395]12[396]	GCTGGCGAAAGGTATCGGCCTCAGGAAGATCGTATTACGCCA
-18r_12[437]12[438]	TTGGGAAGGGCACCGCTTCTGGTGCCGGAAACTGCGCAACTG
-18r_12[464]13[464]	TTTTTCGCCATTCAGGCCAGGCAAAGCGCCATTT
-18r_11[308]19[307]	GAATCATGGCCAGTGCCAAGCGGATTGAAACATTAACTAAAATTGATTCTTAATTG
-18r_13[355]12[336]	TGCATCTGCTAAGTTGGGTAACGCCAGGGTTTTCCCA
-18r_13[375]13[374]	GACGACGACAGGGGATGTGCTGCAAGGCGATCAGTTTGAGGG
-18r_13[417]13[416]	GCTTTCCGGCGATCGGTGCGGGCCTCTTCGCCACTCCAGCCA
-18r_22[328]11[323]	CGTCACCAGTACAAGGGATAGAAACGACAATTACTAG
-18r_22[299]21[305]_M13NH1	GCATTCAAACGGCTCATAGCAAGAATAAACACCGTTTGAAACGCATTAGTCTGGAATTCA
-18r_22[299]21[305]_M13NH2	GCATTCAAACGGCTTGAGTGAAGAATAAACACCGTTTGAAACGCATTAGTCTGGAATTCA
-18r_11[279]12[282]_M13NH1	TTTAAGGCGTTAAATATATCACCAGCA
-18r_11[279]12[282]_M13NH2	TTTAAGGCGTTAAATAGTGAGTTGATG

Table H.8: Cloning sequences.

Name	Sequence
Trk6_α_β	GCTATGGTCTTACTAGGCTATGCTGGTGATATGCTATGGTCTTACTAGGCTATGCTGGTGATAT GCTATGGTCTTACTAGGCTATGCTGGTGATATGCTATG
Trk6_α_βcomplement	AGCTCATAGCATATCACCAGCATAGCCTAGTAAGACCATAGCATATCACCAGCATAGCCTAGTAA GACCATAGCATATCACCAGCATAGCCTAGTAAGACCATAGCCATG
M13 Forward (-41)	CGCCAGGGTTTTCCCAGTCACGAC
M13 Reverse (-48)	AGCGGATAACAATTTCACACAGG

Table H.9: FRET tension verification sequences.

Name	Sequence
25nt_Bookend_LHS	CCCTCTCTTCATTATCTTAGTGCCAAGCTTGCATGCCTGCAGG
25nt_Bookend_RHS	TTGGGTAACGCCAGGGTTTTCCCAGTATCTCACCTCCTACCCT
Spacer_24nt	TCACGACGTTGTAAAACGACGGCC

Table H.10: Hybridization sequences.

Name	Sequence
Low secondary structure insert	CATACCCTCATCCAACTCAACATCACGCAGAAAACCCCACTAGACAGCACATTCCTACAATCCCT CTCGGACCAACTCAAATACCATCAACTCACTCACTCA
Low secondary structure insert complement	AGCTTGAGTGAGTGAGTTGATGGTATTTGAGTTGGTCCGAGAGGGATTGTAGGAATGTGCTGTC TAGTGGGGTTTTCTGCGTGATGTTGAGTTGGATGAGGGTATGCATG
R	GGGTTTTCTGCGTGATGTTGAG
V_{22}	TGTAGGAATGTGCTGTCTAGTG
V_{10}	TCCTCCTCCACCCTGTCTAGTG
T^*_{22}	TTTTTTCTCAACATCACGCAGAAAACCCCACTAGACAGCACATTCCTACAATCCCTTTTT
T^*_{10}	TTTTTTCTCAACATCACGCAGAAAACCCCACTAGACAGGGTGGAGGAGGATTTTTT

H.3 Bipedal motor sequences

Table H.11: Foot sequences. X indicates an abasic site.

Name	Sequence
D16_α_Foot	GTTAAGTGTGTGTGTATTTTCATAGCCGAGTAAGACCATAGC
d16_α_Foot	TACACACACACTTAACTTTTCATAGCCGAGTAAGACCATAGC
d16_β_Foot	TACACACACACTTAACTTTTCATAGCAGATCACCAGCATAGC
D16_γ_Foot	GTTAAGTGTGTGTGTATTTTCATAGCACAGACACAACATAGC
d16_γ_foot	TACACACACACTTAACTTTTCATAGCACAGACACAACATAGC
D16_α_Heel_mm_Foot	GTTAAGTGTGTGTGTATTTTTTTTTTTCGAGTAAGACCATAGC
D16_α_Toe_mm_Foot	GTTAAGTGTGTGTGTATTTTCATAGCCGAGTAAGACTTTTTT
d16_β_Heel_mm_Foot	TACACACACACTTAACTTTTTTTTTTAGATCACCAGCATAGC
d16_β_Toe_mm_Foot	TACACACACACTTAACTTTTCATAGCAGATCACCAGTTTTTT
D16_α_Foot_12[179]-12[167]	GTTAAGTGTGTGTGTATTTTCATAGCCGAGTAAGACCATAGCXATGCCTGCAGG
D16_α_Foot_12[179]-11[176]	GTTAAGTGTGTGTGTATTTTCATAGCCGAGTAAGACCATAGCCXATGCCTGCAGGAGCTGATTG

Table H.12: Test tracks and associated strands.

Name	Sequence
Trk2_α_β	GCTATGGTCTTACTAGGCTATGCTGGTGATATGCTATG
Trk2_β_α	GCTATGCTGGTGATATGCTATGGTCTTACTAGGCTATG
Trk2_α_γ	GCTATGGTCTTACTAGGCTATGTTGTGTCTTTGCTATG
Trk2_γ_α	GCTATGTTGTGTCTTTGCTATGGTCTTACTAGGCTATG
12[179]-11[176]'_Trk2_α_β_11[279]-12[289]'	CAATCAGCTCCTGCAGGCATGGCTATGGTCTTACTAGGCTATGCTGGTGATATGCTATGAGCTTGGCACTGATTTAACG
Trk3_α_β_α	GCTATGGTCTTACTAGGCTATGCTGGTGATATGCTATGGTCTTACTAGGCTATG
Trk3_β_α_β	GCTATGCTGGTGATATGCTATGGTCTTACTAGGCTATGCTGGTGATATGCTATG
Trk3_β_α_γ	GCTATGCTGGTGATATGCTATGGTCTTACTAGGCTATGTTGTGTCTTTGCTATG
Trk3_γ_α_β	GCTATGTTGTGTCTTTGCTATGGTCTTACTAGGCTATGCTGGTGATATGCTATG
Trk3_β_γ_β	GCTATGCTGGTGATATGCTATGTTGTGTCTTTGCTATGCTGGTGATATGCTATG
Trk3_α_β_α	GCTATGGTCTTACTAGGCTATGCTGGTGATATGCTATGGTCTTACTAGGCTATG
Trk3_α_β_γ	GCTATGGTCTTACTAGGCTATGCTGGTGATATGCTATGTTGTGTCTTTGCTATG
Trk3_γ_β_α	GCTATGTTGTGTCTTTGCTATGCTGGTGATATGCTATGGTCTTACTAGGCTATG
Trk3_α_γ_α	GCTATGGTCTTACTAGGCTATGCTGGTGATATGCTATGGTCTTACTAGGCTATG
Trk3_10T_α_β_α_10T	TTTTTTTTTTGCTATGGTCTTACTAGGCTATGCTGGTGATATGCTATGGTCTTACTAGGCTATGTTTTTTTTTT
Trk3_10T_γ_β_α_10T	TTTTTTTTTTGCTATGTTGTGTCTTTGCTATGCTGGTGATATGCTATGGTCTTACTAGGCTATGTTTTTTTTTT
Trk3_10T_α_β_γ_10T	TTTTTTTTTTGCTATGGTCTTACTAGGCTATGCTGGTGATATGCTATGTTGTGTCTTTGCTATGTTTTTTTTTT
Trk3_10T_β_α_β_10T	TTTTTTTTTTGCTATGCTGGTGATATGCTATGGTCTTACTAGGCTATGCTGGTGATATGCTATGTTTTTTTTTT
Trk3_10T_γ_α_β_10T	TTTTTTTTTTGCTATGTTGTGTCTTTGCTATGGTCTTACTAGGCTATGCTGGTGATATGCTATGTTTTTTTTTT
Trk3_10T_β_α_γ_10T	TTTTTTTTTTGCTATGCTGGTGATATGCTATGGTCTTACTAGGCTATGTTGTGTCTTTGCTATGTTTTTTTTTT
12[179]-11[176]'_Trk3_α_β_α	CAATCAGCTCCTGCAGGCATGGCTATGGTCTTACTAGGCTATGCTGGTGATATGCTATGGTCTTACTAGGCTATG
12[179]-11[176]'	CAATCAGCTCCTGCAGGCATG
11[279]-12[282]_END'_Q'	TTTAAGGCGTTAAATCAGTGCCAAGCTTTACCACAAACACCCACA
Q	TGAGTGGTGTGGGTGTTTGTGGT
Displacer	ACCACAAACACCCACACCACTCAAACTCCAACTCAC

Table H.13: Fuel sequences. X indicates an abasic site.

Name	Sequence	Name	Sequence
Fuel_α_Fwd_[8,8]	GTCTTACTXGGCTATG	Fuel_α_Rev_[8,8]	GCTATGGTXTTACTCG
Fuel_β_Fwd_[8,8]	CTGGTGATXTGCTATG	Fuel_β_Rev_[8,8]	GCTATGCTXGTGATCT
Fuel_α_Fwd_[16]	GTCTTACTCGGCTATG	Fuel_α_Rev_[16]	GCTATGGTCTTACTCG
Fuel_β_Fwd_[16]	CTGGTGATCTGCTATG	Fuel_β_Rev_[16]	GCTATGCTGGTGATCT
Fuel_α_Fwd_[RNA]	GUCUUACUCGGCUAUG	Fuel_α_Rev_[RNA]	GCUAUGGUCUUACUCG
Fuel_β_Fwd_[RNA]	CUGGUGAUCUGCUAUG	Fuel_β_Rev_[RNA]	GCUAUGCUGGUGAUCU
Fuel_γ_Fwd_[16]	TTGTGTCTGTGCTATG	Fuel_γ_Rev_[16]	GCTATGTTGTGTCTGT

Table H.14: Loading and release sequences.

Name	Sequence
LoadBlock_α	GTCTTACTCGGCTATGCTCTAAACAAACAAACTACTCACC
Release_α	GGTGAGTAGTTTGTTTGTTTAGAGCATAGCCGAGTAAGACCTCTAAACAAACAAAC
LoadBlock_β	CTGGTGATCTGCTATGTGTTTGTTGTTTATTGTGTTTGGG
Release_β	CCCAAACACAATAAACAACAAACACATAGCAGATCACCAGTGTTTGTTGTTTATTG

H.4 Burnt-bridges sequences

Origami staple sequences for the DNA origami nanotube can be found in reference 198, and staple sequences for the DNA origami tile can be found in reference 156.

H.4.1 Motors

Table H.15: Burnt-bridges motor sequences for both tube and tile experiments.

Name	Sequence
Motor	CGATGTTAGTTGGGCTGAGGTTCC
M_Q	CGATGTTAGTTGGGCTGAGGTTCCTTT
Long_M_{Cy3}	GCATGTTAGTTGGGCTGAGGTTCCTTGAGTGTGATGATTTGAGTGTGATGA
Double_Cy3_Motor_top	TTTCATCACACTCTT
M_polyhairpin	CGATGTTAGTTGGGCTGAGGTTCCTTTTCCACGAAGACGTCTTCGTGG

H.4.2 Tubes

Table H.16: Tube neighbour stalk sequences.

Name	Sequence
A1	CTCATTATATTTCAACTTTAATCTTATCAGCTAAGTGGAGTG
A2	TGACAGGAAAGCCAGAATGGAATCACGTTGAAGTGGAGTG
A3	ATACCGATCGAATAATAATTTTTAGCGCAGTCAAGTGGAGTG
A4	AATTGAGCAATAATAACGGAATGAGCCGCCAAGTGGAGTG
A5	ATGGCTTTGAGCCACCACCCTCAACCCAAAAAAGTGGAGTG
A6	TAATTGAAAAATAATATCCCATAGGGTAATAAGTGGAGTG
A7	AGTATGTTCCTGAACAAAGTCAGCCTAATTTAAAGTGGAGTG
A8	ATGGAAGTACATTTAACAATTTATTCTTACAAGTGGAGTG
A9	GTCTTTCCATATGCGTTATACAACATTTGAATAAGTGGAGTG
A10	ATCCAGAATTAACACCGCCTGCATCAATATAAGTGGAGTG
A11	AATATATGAGATGATGGCAATTCAACAGTGCCAAGTGGAGTG
A12	AGCATAAAAAGCGAAAGGAGCGGCCTGAGTAAAGTGGAGTG
A13	AAAGCATCTAATAACATCACTTGGCGCTAGGGAAGTGGAGTG
A14	CAAGAGAAAAAGCGCCATTCGCTGTTATCCAAGTGGAGTG
A15	AACCACCAGTTTCCTGTGTGAAATCATTCAGGCAAGTGGAGTG
A16	ATTAAGAGAAGCCTTTATTTCACAAAGGCTAAGTGGAGTG
A17	GGCCTCTTATTTTTGAGAGATCTAACGCAAGGAAAGTGGAGTG
A18	ACCTGCTCAGTAAAATGTTTAGACTATTATAAGTGGAGTG
A19	ATGCAATGTCAGGTCTTTACCCTGACTGGATAGAAGTGGAGTG
A20	CACCCTCAAAGAATACACTAAAATTTGTATAAGTGGAGTG
A21	CATTGAATACAAAGTACAACGGAGACACTCATCAAGTGGAGTG

Table H.17: Tube biotin tethering sequences.

Name	Sequence
Tether1	ATGGTTTAACCAGTCAGGACGTGTCTGGAAGTTCTCCTCTCCACCACCT
Tether2	ACACCAGCGTTAATAAAACGAAAACATGTTTTAACTCCTCTCCACCACCT
Tether3	CATTCAGTTTACAGGTAGAAAGAGCTTAATTGCTCTCCTCTCCACCACCT
Tether4	TCATTACCTAGGAATACCACATTTGATAAGAGGTCTCCTCTCCACCACCT
Tether5	ATCAAGAATAACGCCAAAAGGACAGGATTAGAGACTCCTCTCCACCACCT
Tether6	AACCAGGCTCGATGAACGGTAATCAATCGGTTGTACCTCCTCTCCACCACCT
Tether10	AACGCCAGAATCACCATCAATATGGTAAAGATTCAACTCCTCTCCACCACCT
5'_Biotin	AGGTGGTGGAGAGGAG
5'_Biotin+30nt	GGGTAGGGAAGATAAGGTGGTGGAGAGGAG

Table H.18: Tube stator sequences.

Name	Sequence
Tube_S1	TTTACACGATCTACATTGGAACCTCAGCCCAACTAACATTCACTCCACAATGCTTTCGCAAAAAA AAGGCTCCACAAATAAA
Tube_S1i	TTTACACGATCTACATTGGAACCTCAGCCCAACTAACATCGTCACTCCACAATGCTTTCGCAAAA AAAAGGCTCCACAAATAAA
Tube_S1i_Start_TET	ATCCCACTCCCTACACTTCGTGGAACCTCAGCCCAACTAACATCGTCACTCCACAATGCTTTCGC AAAAAAAAGGCTCCACAAATAAA
Tube_S2	TTTACACGATCTACATTGGAACCTCAGCCCAACTAACATTCACTCCACAAAAAATCTAAACGAGT AGTAAATTGAACAGCTTG
Tube_S3	TTTACACGATCTACATTGGAACCTCAGCCCAACTAACATTCACTCCACAATCTGAATTCACCACC AGAGCCGCCGGAAACCG
Tube_S4	TTTACACGATCTACATTGGAACCTCAGCCCAACTAACATTCACTCCACAAACCAGAACTACCGTT CCAGTAACAACTAAA
Tube_S5	TTTACACGATCTACATTGGAACCTCAGCCCAACTAACATTCACTCCACAAGAACTGGCAATATCA GAGAGATAATAAGTCCT
Tube_S5*_Stop_Cy5	CCTTTTACTTCCACTACCCATGGAACTTCAGCCCAACTAACATTCACTCCACAAGAACTGGCAAT ATCAGAGAGATAATAAGTCCT
Tube_S6	TTTACACGATCTACATTGGAACCTCAGCCCAACTAACATTCACTCCACAATGAGCGCTATGATTA AGACTCCAGAACCGC
Tube_S7	TTTACACGATCTACATTGGAACCTCAGCCCAACTAACATTCACTCCACAACGAGCATGAGCCAAC GCTCAACAGAAGAAAACA
Tube_S8	TTTACACGATCTACATTGGAACCTCAGCCCAACTAACATTCACTCCACAACAGTATAATAGAAAC CAATCAAAGAATTAA
Tube_S9	TTTACACGATCTACATTGGAACCTCAGCCCAACTAACATTCACTCCACAATACCTTTTTTGTTTG GATTATACTAGGTGAGG
Tube_S10	TTTACACGATCTACATTGGAACCTCAGCCCAACTAACATTCACTCCACAAAATCCTGATTAATGG AAACAGTAAGCCTGT
Tube_S11	TTTACACGATCTACATTGGAACCTCAGCCCAACTAACATTCACTCCACAAACGCTGAGCAAACTA TCGGCCTTGAGAAAGGA
Tube_S12	TTTACACGATCTACATTGGAACCTCAGCCCAACTAACATTCACTCCACAAGAAGAACTAGCCAGC AGCAAATCATATTCC
Tube_S13	TTTACACGATCTACATTGGAACCTCAGCCCAACTAACATTCACTCCACAACGCTGGCATTCCACA CAACATACGGTGCCGGA
Tube_S14	TTTACACGATCTACATTGGAACCTCAGCCCAACTAACATTCACTCCACAAGCTCACAAAGTGTAG CGGTCACATACTTCT
Tube_S15	TTTACACGATCTACATTGGAACCTCAGCCCAACTAACATTCACTCCACAATGCGCAACATTGCCT GAGAGTCTTAATACTTT
Tube_S16	TTTACACGATCTACATTGGAACCTCAGCCCAACTAACATTCACTCCACAAATCAGGTCTGTTGGG AAGGGCGAATCATGGT
Tube_S17	TTTACACGATCTACATTGGAACCTCAGCCCAACTAACATTCACTCCACAATAAAAATTGCAAAGC GGATTGCATTGCCAGAG
Tube_S18	TTTACACGATCTACATTGGAACCTCAGCCCAACTAACATTCACTCCACAAAGTCAGAATTTAGAA CCCTCATACCGGAGAG
Tube_S19	TTTACACGATCTACATTGGAACCTCAGCCCAACTAACATTCACTCCACAACGTCCAATGATAAAT TGTGTCGATAAAACGAA
Tube_S20	TTTACACGATCTACATTGGAACCTCAGCCCAACTAACATTCACTCCACAACATCGCCTACTGCGG AATCGTCAACCATAAA
Tube_S21	TTTACACGATCTACATTGGAACCTCAGCCCAACTAACATTCACTCCACAATTTGACCCAGAACCG CCACCCTCTGTATCACC
Cy3_stator_top	TATGTAGATCGTGTAAA
TET_stator_top	CGAAGTGTAGGGAGTGGGAT
Cy5_stator_top	TGGGTAGTGGAAGTAAAAGG
Incorrect Tube_S2	TTTACACGATCTACATTGGAACCTCAGCCCAACTAACATTCACTCCACAAAAAATCTAAACGAGT AGTAAATTGAACAGCTTG
Faulty stator	AGGTTCCGATCTACATTGGAACCTCAGCCCAACTAACATTCACTCCACAA

Table H.19: Tube linker and asymmetric linker sequences.

Name	Sequence	Name	Sequence
Linker 1	TGAGACGGGCAACTCACGACGTTG	2SE_2[523]-21[532]	ACGACGTTGAGCTAAACAGGAG
Linker 2	GGTTTTCCCAGAGCTTTCATCAA	2SE_4[524]-3[523]	GAGACAGTCAGGTTTTCCCAGAG
Linker 3	CCTTCCTGTAGCCCGGAGACAGTCA	2SE_6[525]-5[524]	TTTAAACAGTGTGAGAAAGGCCG
Linker 4	GTGAGAAAGGCCGAACGAGTAGA	2SE_8[526]-7[525]	CAAGCGCGAACCCCCTCAAATTT
Linker 5	TTCCCAATTCTGGCTTTAAACAGT	2SE_10[527]-9[526]	CGCCACCCTCCCAGCGATTATAA
Linker 6	CCCCCTCAAATTTATGCGATTTT	2SE_12[528]-11[527]	CCAGCAAAATGAGGTTTAGTGA
Linker 7	ATTGTGAATTACCACCAAGCGCGAA	2SE_14[529]-13[528]	TTAGTTGCTATTGGGAATTAGCT
Linker 8	CCAGCGATTATAATTGTATCGGT	2SE_16[530]-15[529]	TATATAACTACCTTAAATCAAAG
Linker 9	AAAGGAGCCTTTACCGCCACCCTC	2SE_18[531]-17[530]	AATACATTTGGCTTAGGTTGGAT
Linker 10	GAGGTTTAGGATATTCACAAA	2SE_20[532]-19[531]	GCCGATTAAGCCGTCAATAGCT
Linker 11	GACGATTGGCCTTAGCCAGCAAAAT	2SE_3[148]-2[147]	CTTTCATCAATGAGACGGGCAACTC
Linker 12	TTGGGAATTAGCTTTTTAAGAAA	2SE_5[149]-4[148]	AACGAGTAGACCTTCCTGTAGCCCG
Linker 13	TCTTACCGAAGCCGATTAGTTGCTA	2SE_7[150]-6[149]	ATGCGATTTTTTCCCAATTCTGGC
Linker 14	CCTTAAATCAAAGTAATTCTGTC	2SE_9[151]-8[150]	TTGTATCGGTATTGTGAATTACCAC
Linker 15	CGACAAAAGGTAAGTTATATAACTA	2SE_11[152]-10[151]	TATTCACAAAAAAGGAGCCTTTAC
Linker 16	GCTTAGGTTGGATAACGGATTC	2SE_13[153]-12[152]	TTTTAAGAAAGACGATTGGCCTTAG
Linker 17	ATCGGGAGAAACAATAATACATTTG	2SE_15[154]-14[153]	TAATTCTGTCTCTTACCGAAGCCGA
Linker 18	GCCGTCAATAGCTTCTGACCTGA	2SE_17[155]-16[154]	AACGGATTCCGACAAAAGGTAAGT
Linker 19	CAGAGATAGAACCAGGCCGATTAA	2SE_19[156]-18[155]	TCTGACCTGAATCGGGAGAAACAAT
Linker 20	AGCTAAACAGGAGCTGATTGCCC	2SE_21[157]-20[156]	CTGATTGCCCCAGAGATAGAACCAG

4SE_2[523]-21[532]	GACGTTGAGCTAAACAGGAGCT	6SE_2[523]-21[532]	CGTTGAGCTAAACAGGAGCTGA
4SE_4[524]-3[523]	GACAGTCAGGTTTTCCCAGAGCT	6SE_4[524]-3[523]	CAGTCAGGTTTTCCCAGAGCTTT
4SE_6[525]-5[524]	TAAACAGTGTGAGAAAGGCCGAA	6SE_6[525]-5[524]	AACAGTGTGAGAAAGGCCGAACG
4SE_8[526]-7[525]	AGCGCGAACCCCCTCAAATTTAT	6SE_8[526]-7[525]	CGCGAACCCCCTCAAATTTATGC
4SE_10[527]-9[526]	CCACCCTCCCAGCGATTATAATT	6SE_10[527]-9[526]	ACCCTCCCAGCGATTATAATTGT
4SE_12[528]-11[527]	AGCAAAATGAGGTTTAGTGATA	6SE_12[528]-11[527]	CAAAATGAGGTTTAGTGATATT
4SE_14[529]-13[528]	AGTTGCTATTGGGAATTAGCTTT	6SE_14[529]-13[528]	TTGCTATTGGGAATTAGCTTTTT
4SE_16[530]-15[529]	TATAACTACCTTAAATCAAAGTA	6SE_16[530]-15[529]	TAACTACCTTAAATCAAAGTAAT
4SE_18[531]-17[530]	TACATTTGGCTTAGGTTGGATAA	6SE_18[531]-17[530]	CATTTGGCTTAGGTTGGATAACG
4SE_20[532]-19[531]	CGATTAAGCCGTCAATAGCTTC	6SE_20[532]-19[531]	ATTAAGCCGTCAATAGCTTCTG
4SE_3[148]-2[147]	TTCATCAATGAGACGGGCAACTCAC	6SE_3[148]-2[147]	CATCAATGAGACGGGCAACTCACGA
4SE_5[149]-4[148]	CGAGTAGACCTTCCTGTAGCCCGGA	6SE_5[149]-4[148]	AGTAGACCTTCCTGTAGCCCGGAGA
4SE_7[150]-6[149]	GCGATTTTTTCCCAATTCTGGCTT	6SE_7[150]-6[149]	GATTTTTTCCCAATTCTGGCTTTA
4SE_9[151]-8[150]	GTATCGGTATTGTGAATTACCACCA	6SE_9[151]-8[150]	ATCGGTATTGTGAATTACCACCAAG
4SE_11[152]-10[151]	TTCACAAAAAAGGAGCCTTTACCG	6SE_11[152]-10[151]	CACAAAAAAGGAGCCTTTACCGCC
4SE_13[153]-12[152]	TTAAGAAAGACGATTGGCCTTAGCC	6SE_13[153]-12[152]	AAGAAAGACGATTGGCCTTAGCCAG
4SE_15[154]-14[153]	ATTCTGTCTCTTACCGAAGCCGATT	6SE_15[154]-14[153]	TCTGTCTCTTACCGAAGCCGATTAG
4SE_17[155]-16[154]	CGGATTCCGACAAAAGGTAAGTTA	6SE_17[155]-16[154]	GATTCCGACAAAAGGTAAGTTATA
4SE_19[156]-18[155]	TGACCTGAATCGGGAGAAACAATAA	6SE_19[156]-18[155]	ACCTGAATCGGGAGAAACAATAATA
4SE_21[157]-20[156]	GATTGCCCCAGAGATAGAACCAGGC	6SE_21[157]-20[156]	TTGCCCCAGAGATAGAACCAGGCCG
8SE_2[523]-21[532]	TTGAGCTAAACAGGAGCTGATT		
8SE_4[524]-3[523]	GTCAGGTTTTCCCAGAGCTTTCA		
8SE_6[525]-5[524]	CAGTGTGAGAAAGGCCGAACGAG		
8SE_8[526]-7[525]	CGAACCCCCTCAAATTTATGCGA		
8SE_10[527]-9[526]	CCTCCCAGCGATTATAATTGTAT		
8SE_12[528]-11[527]	AAATGAGGTTTAGTGATATTCA		
8SE_14[529]-13[528]	GCTATTGGGAATTAGCTTTTTAA		
8SE_16[530]-15[529]	ACTACCTTAAATCAAAGTAATTC		
8SE_18[531]-17[530]	TTTGGCTTAGGTTGGATAACGGA		
8SE_20[532]-19[531]	TAAGCCGTCAATAGCTTCTGAC		
8SE_3[148]-2[147]	TCAATGAGACGGGCAACTCACGACG		
8SE_5[149]-4[148]	TAGACCTTCCTGTAGCCCGGAGACA		
8SE_7[150]-6[149]	TTTTTTCCCAATTCTGGCTTTAAA		
8SE_9[151]-8[150]	CGGTATTGTGAATTACCACCAAGCG		
8SE_11[152]-10[151]	CAAAAAAGGAGCCTTTACCGCCAC		
8SE_13[153]-12[152]	GAAAGACGATTGGCCTTAGCCAGCA		
8SE_15[154]-14[153]	TGTCTCTTACCGAAGCCGATTAGTT		
8SE_17[155]-16[154]	TTCCGACAAAAGGTAAGTTATATA		
8SE_19[156]-18[155]	CTGAATCGGGAGAAACAATAATACA		
8SE_21[157]-20[156]	GCCCCAGAGATAGAACCAGGCCGAT		

H.4.3 Tiles

Table H.20: Tile edge staple sequences.

Name	Sequence
blunt_r-9t2i	ACTTTTTCTTTTCACGTTGAAAATATTGCGAATAATAATT
blunt_r-9t4e	GGTGTACAATGAGGAAGTTTCCATGACTAAAG
blunt_r-9t6e	ACATTATTGACCAGGCGCATAGGCAGATGAAC
blunt_r-9t8e	GAAAACGAACAGGTAGAAAGATTCACGGAACA
blunt_r-9t10e	ACTAAAGTGAATGACCATAAATCAACAGTTCA
blunt_r-9t12e	CCCTGTAAACGGTGTCTGGAAGTTAATATGCA
blunt_r-9t14e	AAAACTAGTACTTTTGCGGGAGAACATTATGA
blunt_r-9t16e	AGGTCACGCATGTCAATCATATGTGTAATCGT
blunt_r-9t18e	AAACGACGTTGGTGTAGATGGGCGAATGGGAT
blunt_r-9t20e	GGCGGTTTGCCAGTGCCAAGCTTGACGTTGTA
blunt_r-9t22e	TATCACGGGCGTATTGGGCGCCAGCGGGGAGA
blunt_r-9t24j	AACCCTAAAGGGAGCCAAACCGTC
10T_r-9t2i	ACTTTTTCTTTTTTTTTTTTTTTCACGTTGAAAATATTGCGAATAATAATTTTTTTTTTTT
10T_r-9t4e	GGTGTACATTTTTTTTTTATGAGGAAGTTTCCATGACTAAAG
10T_r-9t6e	ACATTATTTTTTTTTTTTGACCAGGCGCATAGGCAGATGAAC
10T_r-9t8e	GAAAACGATTTTTTTTTTACAGGTAGAAAGATTCACGGAACA
10T_r-9t10e	ACTAAAGTTTTTTTTTTTGAATGACCATAAATCAACAGTTCA
10T_r-9t12e	CCCTGTAATTTTTTTTTTACGGTGTCTGGAAGTTAATATGCA
10T_r-9t14e	AAAACTAGTTTTTTTTTTTACTTTTGCGGGAGAACATTATGA
10T_r-9t16e	AGGTCACGTTTTTTTTTTCATGTCAATCATATGTGTAATCGT
10T_r-9t18e	AAACGACGTTTTTTTTTTTTGGTGTAGATGGGCGAATGGGAT
10T_r-9t20e	GGCGGTTTTTTTTTTTTTGCCAGTGCCAAGCTTGACGTTGTA

10T_r-9t22e	TATCACGGTTTTTTTTTTGCGTATTGGGCGCCAGCGGGGAGA
10T_r-9t24j	TTTTTTTTTTAACCCTAAAGGGAGCCAAACCGTC
link_r-9t24j	ACATCGCCAACCCTAAAGGGAGCCAAACCGTC
link_r-9t2i	ACCGCCACTTTTCACGTTGAAAATATTGCGAATAATAATT
link_r-9t4e	GGCTTTTGATGAGGAAGTTTCCATGACTAAAG
link_r-9t6e	CCCCTTATGACCAGGCGCATAGGCAGATGAAC
link_r-9t8e	TTTACCAGACAGGTAGAAAGATTCACGGAACA
link_r-9t10e	GCCCAATAGAATGACCATAAATCAACAGTTCA
link_r-9t12e	GCACCCAGACGGTGTCTGGAAGTTAATATGCA
link_r-9t14e	AAGAAAAATACTTTTGCGGGAGAACATTATGA
link_r-9t16e	AACACCGGCATGTCAATCATATGTGTAATCGT
link_r-9t18e	TTTTCCCTTTGGTGTAGATGGGCGAATGGGAT
link_r-9t20e	GGTTTAACGCCAGTGCCAAGCTTGACGTTGTA
link_r-9t22e	TTAGAAGTGCGTATTGGGCGCCAGCGGGGAGA
link_r-9t2i	ACTTTTTCTTTTCACGTTGAAAATATTGCGAA
blunt_r9t0i	CCTCAGAACCGCCACCCCCTCAGA
blunt_r9t2f	ACCGCCACATGATACAGGAGTGTATCATACAT
blunt_r9t4f	GGCTTTTGTAGCGTTTGCCATCTTGTCATAGC
blunt_r9t6f	CCCCTTATCGCCAAAGACAAAAGGTCATATGG
blunt_r9t8f	TTTACCAGATAAGAGCAAGAAACATGAGTTAA
blunt_r9t10f	GCCCAATACTACAATTTTATCCTGGCTATTTT
blunt_r9t12f	GCACCCAGTAATATCCCATCCTAATCCTGAAC
blunt_r9t14f	AAGAAAAAAATCATAATTACTAGATAAGAATA
blunt_r9t16f	AACACCGGTAGAATCCTTGAAAACTTAATTAA
blunt_r9t18f	TTTTCCCTGTCAGATGAATATACAGATTTTCA
blunt_r9t20f	GGTTTAACATTAGACTTTACAAACTTGAGGAT
blunt_r9t22j	TTAGAAGTATTAAAAATACCGAACGCCCTAAAACATCGCC
10T_r9t0i	TTTTTTTTTTCCTCAGAACCGCCACCCCCTCAGA
10T_r9t2f	ACCGCCACTTTTTTTTTTTATGATACAGGAGTGTATCATACAT
10T_r9t4f	GGCTTTTGTTTTTTTTTTTTAGCGTTTGCCATCTTGTCATAGC
10T_r9t6f	CCCCTTATTTTTTTTTTTTCGCCAAAGACAAAAGGTCATATGG
10T_r9t8f	TTTACCAGTTTTTTTTTTTATAAGAGCAAGAAACATGAGTTAA
10T_r9t10f	GCCCAATATTTTTTTTTTTCTACAATTTTATCCTGGCTATTTT
10T_r9t12f	GCACCCAGTTTTTTTTTTTTAATATCCCATCCTAATCCTGAAC
10T_r9t14f	AAGAAAAATTTTTTTTTTTAATCATAATTACTAGATAAGAATA
10T_r9t16f	AACACCGGTTTTTTTTTTTTAGAATCCTTGAAAACTTAATTAA
10T_r9t18f	TTTTCCCTTTTTTTTTTTTGTCAGATGAATATACAGATTTTCA
10T_r9t20f	GGTTTAACTTTTTTTTTTTATTAGACTTTACAAACTTGAGGAT
10T_r9t22j	TTAGAAGTTTTTTTTTTTTATTAAAAATACCGAACGCCCTAAAACATCGCCTTTTTTTTTT
link_r9t0i	TAATAATTCCTCAGAACCGCCACCCCCTCAGA
link_r9t22j	TTAGAAGTATTAAAAATACCGAACGCCCTAAA
link_r9t2f	ACTTTTTCATGATACAGGAGTGTATCATACAT
link_r9t4f	GGTGTACATAGCGTTTGCCATCTTGTCATAGC
link_r9t6f	ACATTATTCGCCAAAGACAAAAGGTCATATGG
link_r9t8f	GAAAACGAATAAGAGCAAGAAACATGAGTTAA
link_r9t10f	ACTAAAGTCTACAATTTTATCCTGGCTATTTT
link_r9t12f	CCCTGTAATAATATCCCATCCTAATCCTGAAC
link_r9t14f	AAAACTAGAATCATAATTACTAGATAAGAATA
link_r9t16f	AGGTCACGTAGAATCCTTGAAAACTTAATTAA
link_r9t18f	AAACGACGGTCAGATGAATATACAGATTTTCA
link_r9t20f	GGCGGTTTATTAGACTTTACAAACTTGAGGAT
link_r9t22j	TATCACGGATTAAAAATACCGAACGCCCTAAAACATCGCC

Table H.21: Tile biotin anchor staple sequences and adjacent staple replacements.

Name	Sequence
Bio_repl_r-7t2f	AAAAGGCTTACAGAGGCTTTGAGTAAACGGGTAAAATACAACTTTGATT
Bio_repl_r7t16f	CGTGTGATTTAAGACGCTGAGAAGTGAATAACCTTGCTTCCATCGGGTT
Bio_repl_r7t2f	ACTCAGGAACGGGGTCAGTGCCTGCAGTCTCTGAATTTACCGGAACTT
Bio_repl_r-7t16f	TGATAATCCGGCGGATTGACCGTCATCGTAACCGTGCATGGTTTTCCTT
repl_r-7t20f	CAGTCACGCATGCCTGCAGGTCGAGAATCGGCCAACGCGGGTGGTTT
repl_r7t6f	CAGAGCCACAGACTGTAGCGCGTTTGAGGGAGGGAAGGTTAAGTTTA
repl_r7t20f	AGAAACAATGCACGTAAAACAGAATTAAATCCTTTGCCCAATAGATT
repl_r-7t6f	AAGAGGACTGGCTGACCTTCATCATAATAAAACGAACTAATCAGTTG
Bio_repl_r-7t4f	TAAAATACAACTTTGAAAGAGGACTGGCTGACCTTCATCATAATAAATT
repl_r-7t8f	ACGAACTAATCAGTTGAGATTTAGCCTCAAATGCTTTAAAAAATCAG
Bio_repl_r7t18f	CTTGCTTCCATCGGGAGAAACAATGCACGTAAAACAGAATTAAATCTT
repl_r7t22f	CTTTGCCCAATAGATTAGAGCCGTATAAAACAGAGGTGAATGGCTAT

Table H.22: Tile stator sequences, drift marker staple sequences and PAINT imager sequence.

Name	Sequence
PAINT_imager	TATGTAGATC
P_Drift_r-1t10e_P	GATCTACATTTACTCCAACAAAGCGAACCAGACCGCAGCCTTTTGATCTACAT
P_Drift_r1t16f_P	GATCTACATTTCCAATCGCTATATGTAAATGCTGCCAATAGGTTGATCTACAT
P_Drift_r3t18e_P	GATCTACATTTATCGCGCAAATTACCTTTTTTAACTGAGAGATTGATCTACAT
P_Drift_r3t16f_P	GATCTACATTTTTTTCAAATTAACCTCCGGCTTATTAACAATTTGATCTACAT
P_Drift_r1t10e_P	GATCTACATTTTGTTTAACGAATAACATAAAAACTAATAACGTTGATCTACAT
P_Drift_r3t14f_P	GATCTACATTTAGGTAAAGGAATCGCCATATTTAAGAAAACTTTGATCTACAT
P_Drift_r3t10e_P	GATCTACATTTTTTATCCCCGGGAGAATTAACTGAAGTTACCTTGATCTACAT
P_Drift_r-1t14e_P	GATCTACATTTATTCAACCAGTCAAATCACCATCTTTTCGAGTTGATCTACAT
P_Drift_r3t12e_P	GATCTACATTTCAAGTACCCGGTATTCTAAGAACGCCATATTATTGATCTACAT
P_Drift_r1t14f_P	GATCTACATTTCCAGTAATATTTAGGCAGAGGCAAATATGATTTGATCTACAT
P_Drift_r1t12e_P	GATCTACATTTCGTTTTTATAGCAAGCAAATCAGCGATTTTTTTGATCTACAT
P_Drift_r-1t16e_P	GATCTACATTTAACGCCATTCAGCTCATTTTTTAAATGCAAATTTGATCTACAT
P_Drift_r1t8f_P	GATCTACATTTTCCTTATTAAAAGAACTGGCATGATAGCGAGTTGATCTACAT
P_Drift_r3t12f_P	GATCTACATTTGGCTTATCGCACTCATCGAGAACCCGACAAATTGATCTACAT
P_Drift_r3t6f_P	GATCTACATTTCATCGATAGACTTGAGCCATTTGGAAACGTAGTTGATCTACAT
P_Drift_r1t10f_P	GATCTACATTTTACAGAGAGTCAAAAATGAAAATAGGAAGCAATTGATCTACAT
P_Drift_r-1t12e_P	GATCTACATTTTCAATTCTGGGCGCGAGCTGAAATCATTACCTTGATCTACAT
P_Drift_r1t12f_P	GATCTACATTTGCGCCCAATTTTCATCGTAGGAAAGGTGGCATTGATCTACAT
P_Drift_r3t10f_P	GATCTACATTTGCATTAGAAATCCAAATAAGAAAATATAGAATTGATCTACAT
P_Drift_r3t14e_P	GATCTACATTTTTAATTGATAATTCTGTCCAGACTATTAAACTTGATCTACAT
P_Drift_r1t16e_P	GATCTACATTTTATATAACAAGACAAAGAACGCGACAACGCCTTGATCTACAT
P_Drift_r3t8f_P	GATCTACATTTAAAATACAACCGAGGAAACGCAAAGGGAAGCTTGATCTACAT
P_Drift_r-1t8e_P	GATCTACATTTAGGCTTTTACGATAAAAACCAAAATTAAGACTTGATCTACAT
P_Drift_r3t16e_P	GATCTACATTTCTACCTTTTATATTTTAGTTAATTAGTAGGGCTTGATCTACAT
P_Drift_r1t18e_P	GATCTACATTTACCTGAGCACAAAATTAATTACATGGTTGGGTTTGATCTACAT
P_Drift_r1t14e_P	GATCTACATTTAACATGTAAAGAGAATATAAAGTAAAGCAAGCTTGATCTACAT
P_S1/_blunt_r-9t4e	GATCTACATGGAACCTCAGCCCAACTAACATTTTGGTGTACAATGAGGAAGTTTCCATGACTAAAG
P_S1/_link_r9t4f	GATCTACATGGAACCTCAGCCCAACTAACATTTTACTTTTTCATGATACAGGAGTGTATCATACAT
P_S2/_Bio_repl_r-7t4f	GATCTACATGGAACCTCAGCCCAACTAACATTTTTAAAATACAACTTTGAAAGAGGACTGGCTGAC CTTCATCATAATAAATT
P_S3_r-7t2e	GATCTACATGGAACCTCAGCCCAACTAACATTTTGCAACGGCCCAAAAGGAGCCTTTGGAGTGAG
P_S4_r-5t4f	GATCTACATGGAACCTCAGCCCAACTAACATTTTACCTAAAAGACGGTCAATCATAAAGAACCGG
P_S5_r-5t4e	GATCTACATGGAACCTCAGCCCAACTAACATTTTGAGGCGCACGAAAGAGGCAAAAGGTCACCCT
S6i_r-3t6e	GGAACCTCAGCCCAACTAACATCGTTTGAATTACCAGTGAATAAGGCTTGTCGAAATC
P_S6_r-3t6e	GATCTACATGGAACCTCAGCCCAACTAACATTTTGAATTACCAGTGAATAAGGCTTGTCGAAATC
P_S7_r-3t6f	GATCTACATGGAACCTCAGCCCAACTAACATTTTGCTCATTCTTATGCGATTTTAAGCGAGGCAT
P_S8_r-1t8f	GATCTACATGGAACCTCAGCCCAACTAACATTTTACCAGACGGCAAAAGAAGTTTTGTTTTAATT
P_S9_r-1t8e	GATCTACATGGAACCTCAGCCCAACTAACATTTTAGGCTTTTACGATAAAAACCAAAATTAAGAC
P_S10_r1t10f	GATCTACATGGAACCTCAGCCCAACTAACATTTTTACAGAGAGTCAAAAATGAAAATAGGAAGCAA
P_S11*_r7t16e	GATCTACATGGAACTTCAGCCCAACTAACATTTTAGCTTAGAAAATAAGGCGTTAAAAAAAGCCT
P_S11_r1t10e	GATCTACATGGAACCTCAGCCCAACTAACATTTTTGTTTAACGAATAACATAAAAACTAATAACG
P_S12_r3t12f	GATCTACATGGAACCTCAGCCCAACTAACATTTTGGCTTATCGCACTCATCGAGAACCCGACAAA
P_S13/_r3t10e	GATCTACATGGAACCTCAGCCCAACTAACATTTTTTTATCCCCGGGAGAATTAACTGAAGTTACC
P_S13_r3t12e	GATCTACATGGAACCTCAGCCCAACTAACATTTTCAAGTACCCGGTATTCTAAGAACGCCATATTA
P_S14/_r5t10f	GATCTACATGGAACCTCAGCCCAACTAACATTTTGAACAAAGTTACAAAATAAACAGCGAGGCGT
P_S14_r5t14f	GATCTACATGGAACCTCAGCCCAACTAACATTTTTAAACAACAAAGCCAACGCTCAACTCATCTTC
P_S15/_r5t8e	GATCTACATGGAACCTCAGCCCAACTAACATTTTAAAAGTAACAAAGACACCACGGAAAAATATTG
P_S15_r5t14e	GATCTACATGGAACCTCAGCCCAACTAACATTTTACCAGTATATGTTCAGCTAATGCGGCTGTCT
P_S16/_r7t8f	GATCTACATGGAACCTCAGCCCAACTAACATTTTTTTTGTCATATCTTACCGAAGCCCCGCTAATA
P_S16_r7t16f	GATCTACATGGAACCTCAGCCCAACTAACATTTTCGTGTGATTTAAGACGCTGAGAAGTGAATAAC
P_S17/_r7t8f	GATCTACATGGAACCTCAGCCCAACTAACATTTTCAACCGATTTCATCGGCATTTTCGTTCATAAT
P_S17_r7t16e	GATCTACATGGAACCTCAGCCCAACTAACATTTTAGCTTAGAAAATAAGGCGTTAAAAAAAGCCT
P_S18/_blunt_r9t6f	GATCTACATGGAACCTCAGCCCAACTAACATTTTCCCCTTATCGCCAAAGACAAAAGGTCATATGG
P_S18/_link_r-9t6e	GATCTACATGGAACCTCAGCCCAACTAACATTTTCCCCTTATGACCAGGCGCATAGGCAGATGAAC
B1_S13_r3t12e	AACAACACAAAAACAAAACCGGAACCTCAGCCCAACTAACATTTTCAAGTACCCGGTATTCTAAGA ACGCCATATTA
B1_S14_r5t14f	AACAACACAAAAACAAAACCGGAACCTCAGCCCAACTAACATTTTTAAACAACAAAGCCAACGCTC AACTCATCTTC
B2_S13/_r3t10e	CTCTCTCTCTTTTTTACCGTGGAACCTCAGCCCAACTAACATTTTTTTATCCCCGGGAGAATTAACT GAAGTTACC
B2_S14/_r5t10f	CTCTCTCTCTTTTTTACCGTGGAACCTCAGCCCAACTAACATTTTGAACAAAGTTACAAAATAAACA GCGAGGCGT

Bibliography

[1] Green, S. J., Bath, J., & Turberfield, A. J. "Coordinated chemomechanical cycles: A mechanism for autonomous molecular motion." *Phys. Rev. Lett.*, **101**:238101 (2008). doi:10.1103/PhysRevLett.101.238101.

[2] Bath, J., Green, S. J., Allen, K. E., & Turberfield, A. J. "Mechanism for a directional, processive, and reversible DNA motor." *Small*, **5**:1513 (2009). doi:10.1002/smll.200900078.

[3] Bath, J., Green, S. J., & Turberfield, A. J. "A free-running DNA motor powered by a nicking enzyme." *Angew. Chemie*, **44**:4358 (2005). doi:10.1002/anie.200501262.

[4] Jungmann, R., Avendaño, M. S., Woehrstein, J. B., Dai, M., Shih, W. M., & Yin, P. "Multiplexed 3D cellular super-resolution imaging with DNA-PAINT and Exchange-PAINT." *Nat. Methods*, **11**:313 (2014). doi:10.1038/nmeth.2835.

[5] Franklin, R. E. & Gosling, R. G. "Molecular configuration in sodium thymonucleate." *Nature*, **171**:740 (1953). doi:10.1038/171740a0.

[6] Wilkins, M. H. F., Stokes, A. R., & Wilson, H. R. "Molecular structure of nucleic acids: molecular structure of deoxypentose nucleic acids." *Nature*, **171**:738 (1953). doi:10.1038/171738a0.

[7] Watson, J. D. & Crick, F. H. C. "Molecular structure of nucleic acids: A structure for deoxyribose nucleic acid." *Nature*, **171**:737 (1953). doi:10.1038/171737a0.

[8] Watson, J. D. & Crick, F. H. "Genetical implications of the structure of deoxyribonucleic acid." *Nature*, **171**:964 (1953). doi:10.1038/171964b0.

[9] Hunter, C. A. & Sanders, J. K. M. "The nature of π-π interactions." *J. Am. Chem. Soc.*, **112**:5525 (1990). doi:10.1021/ja00170a016.

[10] Saenger, W. *Principles of nucleic acid structure*. Springer Advanced Texts in Chemistry. Springer New York, New York, NY (1984). ISBN 978-0-387-90761-1. doi:10.1007/978-1-4612-5190-3.

[11] Wang, J. C. "Helical repeat of DNA in solution." *Proc. Natl. Acad. Sci.*, **76**:200 (1979). doi:10.1073/pnas.76.1.200.

[12] Rhodes, D. & Klug, A. "Helical periodicity of DNA determined by enzyme digestion." *Nature*, **286**:573 (1980). doi:10.1038/292378a0.

[13] Hagerman, P. J. "Flexibility of DNA." *Annu. Rev. Biophys. Biophys. Chem.*, **17**:265 (1988). doi:10.1146/annurev.bb.17.060188.001405.

[14] Wang, M. D., Yin, H., Landick, R., Gelles, J., & Block, S. M. "Stretching DNA with optical tweezers." *Biophys. J.*, **72**:1335 (1997). doi:10.1016/S0006-3495(97)78780-0.

[15] Brinkers, S., Dietrich, H. R. C., de Groote, F. H., Young, I. T., & Rieger, B. "The persistence length of double stranded DNA determined using dark field tethered particle motion." *J. Chem. Phys.*, **130**:215105 (2009). doi:10.1063/1.3142699.

[16] Smith, S. B., Cui, Y., & Bustamante, C. "Overstretching B-DNA: The elastic response of individual double-stranded and single-stranded DNA molecules." *Science*, **271**:795 (1996). doi:10.1126/science.271.5250.795.

[17] Tinland, B., Pluen, A., Sturm, J., & Weill, G. "Persistence length of single-stranded DNA." *Society*, **30**:5763 (1997). doi:10.1021/ma970381.

[18] Holliday, R. "A mechanism for gene conversion in fungi." *Genet. Res.*, **5**:282 (1964). doi:10.1017/S0016672308009476.

[19] Liu, Y. & West, S. C. "Timeline: Happy Hollidays: 40th anniversary of the Holliday junction." *Nat. Rev. Mol. Cell Biol.*, **5**:937 (2004). doi:10.1038/nrm1502.

[20] Clegg, R. M., Murchie, a. I., & Lilley, D. M. "The solution structure of the four-way DNA junction at low-salt conditions: a fluorescence resonance energy transfer analysis." *Biophys. J.*, **66**:99 (1994). doi:10.1016/S0006-3495(94)80765-9.

[21] Kallenbach, N. R., Ma, R.-I., & Seeman, N. C. "An immobile nucleic acid junction constructed from oligonucleotides." *Nature*, **305**:829 (1983). doi:10.1038/305829a0.

[22] Tsu-Ju, F. & Seeman, N. C. "DNA double-crossover molecules." *Biochemistry*, **32**:3211 (1993). doi:10.1021/bi00064a003.

[23] Winfree, E., Liu, F., Wenzler, L. A., & Seeman, N. C. "Design and self-assembly of two-dimensional DNA crystals." *Nature*, **394**:539 (1998). doi:10.1038/28998.

[24] Chen, J. H. & Seeman, N. C. "Synthesis from DNA of a molecule with the connectivity of a cube." *Nature*, **350**:631 (1991). doi:10.1038/350631a0.

[25] Goodman, R. P., Schaap, I. A. T., Tardin, C. F., Erben, C. M., Berry, R. M., Schmidt, C. F., & Turberfield, A. J. "Rapid chiral assembly of rigid DNA building blocks for molecular nanofabrication." *Science*, **310**:1661 (2005). doi:10.1126/science.1120367.

[26] Kato, T., Goodman, R. P., Erben, C. M., Turberfield, A. J., & Namba, K. "High-resolution structural analysis DNA nanostructure by cryoEM." *Nano Lett.*, **9**:2747 (2009). doi:10.1021/nl901265n.

[27] Kilchherr, F., Wachauf, C., Pelz, B., Rief, M., Zacharias, M., & Dietz, H. "Single-molecule dissection of stacking forces in DNA." *Science*, **353**:aaf5508 (2016). doi:10.1126/science.aaf5508.

[28] SantaLucia, J. "A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics." *Proc. Natl. Acad. Sci.*, **95**:1460 (1998). doi:10.1073/pnas.95.4.1460.

[29] Gotoh, O. & Tagashira, Y. "Stabilities of nearest-neighbor doublets in double-helical DNA determined by fitting calculated melting profiles to observed profiles." *Biopolymers*, **20**:1033 (1981). doi:10.1002/bip.1981.360200513.

[30] Vologodskii, A. V., Amirikyan, B. R., Lyubchenko, Y. L., & Frank-Kamenetskii, M. D. "Allowance for heterogeneous stacking in the DNA helix-coil transition theory." *J. Bomolecular Struct. Dyn.*, **2**:131 (1984). doi:10.1080/07391102.1984.10507552.

[31] Breslauer, K. J., Frank, R., Blöcker, H., & Marky, L. A. "Predicting DNA duplex stability from the base sequence." *Proc. Natl. Acad. Sci.*, **83**:3746 (1986). doi:10.1073/pnas.83.11.3746.

[32] Delcourt, S. G. & Blake, R. D. "Stacking energies in DNA." *J. Biol. Chem.*, **266**:15160 (1991).

[33] Doktycz, M. J., Goldstein, R. F., Paner, T. M., Gallo, F. J., & Benight, A. S. "Studies of DNA dumbbells. I. Melting curves of 17 DNA dumbbells with different duplex stem sequences linked by T4 endloops: Evaluation of the nearest-neighbor stacking interactions in DNA." *Biopolymers*, **32**:849 (1992). doi:10.1002/bip.360320712.

[34] Sugimoto, N., Nakano, S. I., Yoneyama, M., & Honda, K. I. "Improved thermodynamic parameters and helix initiation factor to predict stability of DNA duplexes." *Nucleic Acids Res.*, **24**:4501 (1996). doi:10.1093/nar/24.22.4501.

[35] SantaLucia, J., Allawi, H. T., & Seneviratne, P. A. "Improved nearest-neighbor parameters for predicting DNA duplex stability." *Biochemistry*, **35**:3555 (1996). doi:10.1021/bi951907q.

[36] SantaLucia, J. & Hicks, D. "The thermodynamics of DNA structural motifs." *Annu. Rev. Biophys. Biomol. Struct.*, **33**:415 (2004). doi:10.1146/annurev.biophys.32.110601.141800.

[37] Jungmann, R., Steinhauer, C., Scheible, M., Kuzyk, A., Tinnefeld, P., & Simmel, F. C. "Single-molecule kinetics and super-resolution microscopy by fluorescence imaging of transient binding on DNA origami." *Nano Lett.*, **10**:4756 (2010). doi:10.1021/nl103427w.

[38] Yurke, B., Mills, A. P. J., & Jr., A. P. M. "Using DNA to power nanostructures." *Genet. Program. Evolvable Mach.*, **4**:111 (2003). doi:10.1023/A:1023928811651.

[39] Zhang, D. Y. & Winfree, E. "Control of DNA strand displacement kinetics using toehold exchange." *J. Am. Chem. Soc.*, **131**:17303 (2009). doi:10.1021/ja906987s.

[40] Srinivas, N., Ouldridge, T. E., Sulc, P., Schaeffer, J. M., Yurke, B., Louis, A. A., Doye, J. P. K., & Winfree, E. "On the biophysics and kinetics of toehold-mediated DNA strand displacement." *Nucleic Acids Res.*, **41**:10641 (2013). doi:10.1093/nar/gkt801.

[41] Sha, R., Liu, F., & Seeman, N. C. "Direct evidence for spontaneous branch migration in antiparallel DNA Holliday junctions." *Biochemistry*, **39**:11514 (2000). doi:10.1021/bi0010468.

[42] Reynaldo, L. P., Vologodskii, A. V., Neri, B. P., & Lyamichev, V. I. "The kinetics of oligonucleotide replacements." *J. Mol. Biol.*, **297**:511 (2000). doi:10.1006/jmbi.2000.3573.

[43] Qian, L., Winfree, E., & Bruck, J. "Neural network computation with DNA strand displacement cascades." *Nature*, **475**:368 (2011). doi:10.1038/nature10262.

[44] Qian, L. & Winfree, E. "Scaling up digital circuit computation with DNA strand displacement cascades." *Science*, **332**:1196 (2011). doi:10.1126/science.1200520.

[45] Green, S. J., Lubrich, D., & Turberfield, A. J. "DNA hairpins: fuel for autonomous DNA devices." *Biophys. J.*, **91**:2966 (2006). doi:10.1529/biophysj.106.084681.

[46] Li, B., Ellington, A. D., & Chen, X. "Rational, modular adaptation of enzyme-free DNA circuits to multiple detection methods." *Nucleic Acids Res.*, **39**:e110 (2011). doi:10.1093/nar/gkr504.

[47] Chen, S. X., Zhang, D. Y., & Seelig, G. "Conditionally fluorescent molecular probes for detecting single base changes in double-stranded DNA." *Nat. Chem.*, **5**:782 (2013). doi:10.1038/nchem.1713.

[48] Allen, P. B., Arshad, S. a., Li, B., Chen, X., & Ellington, A. D. "DNA circuits as amplifiers for the detection of nucleic acids on a paperfluidic platform." *Lab on a Chip*, **12**:2951 (2012). doi:10.1039/c2lc40373k.

[49] Dirks, R. M. & Pierce, N. A. "Triggered amplification by hybridization chain reaction." *Proc. Natl. Acad. Sci.*, **101**:15275 (2004). doi:10.1073/pnas.0407024101.

[50] Chandran, H., Rangnekar, A., Shetty, G., Schultes, E. a., Reif, J. H., & LaBean, T. H. "An autonomously self-assembling dendritic DNA nanostructure for target DNA detection." *Biotechnol. J.*, **8**:221 (2013). doi:10.1002/biot.201100499.

[51] Meng, W., Muscat, R. A., McKee, M. L., Milnes, P. J., El-Sagheer, A. H., Bath, J., Davis, B. G., Brown, T., O'Reilly, R. K., & Turberfield, A. J. "An autonomous molecular assembler for programmable chemical synthesis." *Nat. Chem.*, **8**:542 (2016). doi:10.1038/nchem.2495.

[52] Choi, H. M. T., Beck, V. A., & Pierce, N. A. "Next-generation in situ hybridization chain reaction: Higher gain, lower cost, greater durability." *ACS Nano*, **8**:4284 (2014). doi:10.1021/nn405717p.

[53] Yurke, B., Turberfield, A. J., Mills, A. P. J., Simmel, F. C., & Neumann, J. L. "A DNA-fuelled molecular machine made of DNA." *Nature*, **406**:605 (2000). doi:10.1038/35020524.

[54] Goodman, R. P., Heilemann, M., Doose, S., Erben, C. M., Kapanidis, A. N., & Turberfield, A. J. "Reconfigurable, braced, three-dimensional DNA nanostructures." *Nat. Nanotechnol.*, **3**:93 (2008). doi:10.1038/nnano.2008.3.

[55] Shin, J.-S. & Pierce, N. A. "A synthetic DNA walker for molecular transport." *J. Am. Chem. Soc.*, **126**:10834 (2004). doi:10.1021/ja047543j.

[56] Sherman, W. B. & Seeman, N. C. "A precisely controlled DNA biped walking device." *Nano Lett.*, **4**:1203 (2004). doi:10.1021/nl049527q.

[57] Tomov, T. E., Tsukanov, R., Glick, Y., Berger, Y., Liber, M., Avrahami, D., Gerber, D., & Nir, E. "DNA bipedal botor achieves a large number of steps due to operation using microfluidics-based interface." *ACS Nano*, **11**:4002 (2017). doi:10.1021/acsnano.7b00547.

[58] Yin, P., Yan, H., Daniell, X. G., Turberfield, A. J., & Reif, J. H. "A unidirectional DNA walker that moves autonomously along a track." *Angew. Chemie*, **43**:4906 (2004). doi:10.1002/anie.200460522.

[59] Cherepanov, A. V. & De Vries, S. "Kinetics and thermodynamics of nick sealing by T4 DNA ligase." *Eur. J. Biochem.*, **270**:4315 (2003). doi:10.1046/j.1432-1033.2003.03824.x.

[60] Astakhova, I. K., Pasternak, K., Campbell, M. A., Gupta, P., & Wengel, J. "A locked nucleic acid-based nanocrawler: Designed and reversible movement detected by multicolor fluorescence." *J. Am. Chem. Soc.*, **135**:2423 (2013). doi:10.1021/ja311250w.

[61] Owczarzy, R., You, Y., Groth, C. L., & Tataurov, A. V. "Stability and mismatch discrimination of locked nucleic acid-DNA duplexes." *Biochemistry*, **50**:9352 (2011). doi:10.1021/bi200904e.

[62] Tian, Y., He, Y., Chen, Y., Yin, P., & Mao, C. "A DNAzyme that walks processively and autonomously along a one-dimensional track." *Angew. Chemie*, **44**:4355 (2005). doi:10.1002/anie.200500703.

[63] You, M., Huang, F., Chen, Z., Wang, R.-W., & Tan, W. "Building a nanostructure with reversible motions using photonic energy." *ACS Nano*, **6**:7935 (2012). doi:10.1021/nn302388e.

[64] Yeo, Q. Y., Loh, I. Y., Tee, S. R., Chiang, Y. H., Cheng, J., Liu, M., & Wang, Z. "A DNA bipedal nanowalker with a piston-like expulsion stroke." *Nanoscale*, **23**:187 (2017). doi:10.1039/C7NR03809G.

[65] Bath, J. & Turberfield, A. J. "DNA nanomachines." *Nat. Nanotechnol.*, **2**:275 (2007). doi:10.1038/nnano.2007.104.

[66] Cheng, J., Sreelatha, S., Hou, R., Efremov, A., Liu, R., van der Maarel, J. R. C., & Wang, Z. "Bipedal nanowalker by pure physical mechanisms." *Phys. Rev. Lett.*, **109**:238104 (2012). doi:10.1103/PhysRevLett.109.238104.

[67] You, M., Chen, Y., Zhang, X., Liu, H., Wang, R., Wang, K., Williams, K. R., & Tan, W. "An autonomous and controllable light-driven DNA walking device." *Angew. Chemie*, **51**:2457 (2012). doi:10.1002/anie.201107733.

[68] Hatch, K., Danilowicz, C., Coljee, V., & Prentiss, M. "Demonstration that the shear force required to separate short double-stranded DNA does not increase significantly with sequence length for sequences longer than 25 base pairs." *Phys. Rev. E*, **78**:011920 (2008). doi:10.1103/PhysRevE.78.011920.

[69] Kufer, S. K., Puchner, E. M., Gumpp, H., Liedl, T., & Gaub, H. E. "Single-molecule cut-and-paste surface assembly." *Science*, **319**:594 (2008). doi:10.1126/science.1151424.

[70] Strunz, T., Oroszlan, K., Schäfer, R., & Güntherodt, H. J. "Dynamic force spectroscopy of single DNA molecules." *Proc. Natl. Acad. Sci.*, **96**:11277 (1999). doi:10.1073/pnas.96.20.11277.

[71] Cocco, S., Monasson, R., & Marko, J. F. "Force and kinetic barriers to unzipping of the DNA double helix." *Proc. Natl. Acad. Sci.*, **98**:8608 (2001). doi:10.1073/pnas.151257598.

[72] Ho, D., Zimmermann, J. L., Dehmelt, F. A., Steinbach, U., Erdmann, M., Severin, P., Falter, K., & Gaub, H. E. "Force-driven separation of short double-stranded DNA." *Biophys. J.*, **97**:3158 (2009). doi:10.1016/j.bpj.2009.09.040.

[73] Huguet, J. M., Bizarro, C. V., Forns, N., Smith, S. B., Bustamante, C., & Ritort, F. "Single-molecule derivation of salt dependent base-pair free energies in DNA." *Proc. Natl. Acad. Sci.*, **107**:15431 (2010). doi:10.1073/pnas.1001454107.

[74] Rothemund, P. W. K. "Folding DNA to create nanoscale shapes and patterns." *Nature*, **440**:297 (2006). doi:10.1038/nature04586.

[75] Shih, W. M., Quispe, J. D., & Joyce, G. F. "A 1.7-kilobase single-stranded DNA that folds into a nanoscale octahedron." *Nature*, **427**:618 (2004). doi:10.1038/nature02307.

[76] Douglas, S. M., Dietz, H., Liedl, T., Högberg, B., Graf, F., & Shih, W. M. "Self-assembly of DNA into nanoscale three-dimensional shapes." *Nature*, **459**:414 (2009). doi:10.1038/nature08016.

[77] Dietz, H., Douglas, S. M., & Shih, W. M. "Folding DNA into twisted and curved nanoscale shapes." *Science*, **325**:725 (2009). doi:10.1126/science.1174251.

[78] Douglas, S. M., Marblestone, A. H., Teerapittayanon, S., Vazquez, A., Church, G. M., & Shih, W. M. "Rapid prototyping of 3D DNA-origami shapes with caDNAno." *Nucleic Acids Res.*, **37**:5001 (2009). doi:10.1093/nar/gkp436.

[79] Castro, C. E., Kilchherr, F., Kim, D.-N., Shiao, E. L., Wauer, T., Wortmann, P., Bathe, M., & Dietz, H. "A primer to scaffolded DNA origami." *Nat. Methods*, **8**:221 (2011). doi:10.1038/nmeth.1570.

[80] Kim, D. N., Kilchherr, F., Dietz, H., & Bathe, M. "Quantitative prediction of 3D solution shape and flexibility of nucleic acid nanostructures." *Nucleic Acids Res.*, **40**:2862 (2012). doi:10.1093/nar/gkr1173.

[81] Geary, C., Rothemund, P. W. K., & Andersen, E. S. "A single-stranded architecture for cotranscriptional folding of RNA nanostructures." *Science*, **345**:799 (2014). doi:10.1126/science.1253920.

[82] Benson, E., Mohammed, A., Gardell, J., Masich, S., Czeizler, E., Orponen, P., & Högberg, B. "DNA rendering of polyhedral meshes at the nanoscale." *Nature*, **523**:441 (2015). doi:10.1038/nature14586.

[83] Matthies, M., Agarwal, N. P., & Schmidt, T. L. "Design and synthesis of triangulated DNA origami trusses." *Nano Lett.*, **16**:2108 (2016). doi:10.1021/acs.nanolett.6b00381.

[84] Zhang, P., Lee, S., Yu, H., Fang, N., & Ho Kang, S. "Super-resolution of fluorescence-free plasmonic nanoparticles using enhanced dark-field illumination based on wavelength-modulation." *Sci. Rep.*, **5**:11447 (2015). doi:10.1038/srep11447.

[85] Sobczak, J.-P. J., Martin, T. G., Gerling, T., & Dietz, H. "Rapid folding of DNA into nanoscale shapes at constant temperature." *Science*, **338**:1458 (2012). doi:10.1126/science.1229919.

[86] Stahl, E., Martin, T. G., Praetorius, F., & Dietz, H. "Facile and scalable preparation of pure and dense DNA origami solutions." *Angew. Chemie*, **126**:12949 (2014). doi:10.1002/ange.201405991.

[87] Wickham, S. F., Endo, M., Katsuda, Y., Hidaka, K., Bath, J., Sugiyama, H., & Turberfield, A. J. "Direct observation of stepwise movement of a synthetic molecular transporter." *Nat. Nanotechnol.*, **6**:166 (2011). doi:10.1038/nnano.2010.284.

[88] Zhang, H., Chao, J., Pan, D., Liu, H., Huang, Q., & Fan, C. "Folding super-sized DNA origami with scaffold strands from long-range PCR." *Chem. Commun.*, **48**:6405 (2012). doi:10.1039/c2cc32204h.

[89] Marchi, A. N., Saaem, I., Vogen, B. N., Brown, S., & Labean, T. H. "Toward larger DNA origami." *Nano Lett.*, **14**:5740 (2014). doi:10.1021/nl502626s.

[90] Zhao, Z., Yan, H., & Liu, Y. "A route to scale Up DNA origami using DNA tiles as folding staples." *Angew. Chemie*, **49**:1414 (2010). doi:10.1002/anie.200906225.

[91] Suzuki, Y., Endo, M., & Sugiyama, H. "Lipid-bilayer-assisted two-dimensional self-assembly of DNA origami nanostructures." *Nat. Commun.*, **6**:8052 (2015). doi:10.1038/ncomms9052.

[92] Liber, M., Tomov, T. E., Tsukanov, R., Berger, Y., & Nir, E. "A bipedal DNA motor that travels back and forth between two DNA origami tiles." *Small*, **11**:568 (2014). doi:10.1002/smll.201402028.

[93] Gerling, T., Wagenbauer, K. F., Neuner, A. M., & Dietz, H. "Dynamic DNA devices and assemblies formed by shape-complementary, non-base pairing 3D components." *Science*, **347**:1446 (2015). doi:10.1126/science.aaa5372.

[94] Woo, S. & Rothemund, P. W. K. "Programmable molecular recognition based on the geometry of DNA nanostructures." *Nat. Chem.*, **3**:620 (2011). doi:10.1038/nchem.1070.

[95] Liu, W., Zhong, H., Wang, R., & Seeman, N. C. "Crystalline two-dimensional DNA-origami arrays." *Angew. Chemie*, **123**:278 (2011). doi:10.1002/ange.201005911.

[96] Aghebat Rafat, A., Pirzer, T., Scheible, M. B., Kostina, A., & Simmel, F. C. "Surface-assisted large-scale ordering of DNA origami tiles." *Angew. Chemie*, **53**:7665 (2014). doi:10.1002/anie.201403965.

[97] Endo, M., Sugita, T., Katsuda, Y., Hidaka, K., & Sugiyama, H. "Programmed-assembly system using DNA jigsaw pieces." *Chemistry - A Eur. J.*, **16**:5362 (2010). doi:10.1002/chem.200903057.

[98] Endo, M., Sugita, T., Rajendran, A., Katsuda, Y., Emura, T., Hidaka, K., & Sugiyama, H. "Two-dimensional DNA origami assemblies using a four-way connector." *Chem. Commun.*, **47**:3213 (2011). doi:10.1039/c0cc05306f.

[99] Rajendran, A., Endo, M., Katsuda, Y., Hidaka, K., & Sugiyama, H. "Programmed two-dimensional self-assembly of multiple DNA origami jigsaw pieces." *ACS Nano*, **5**:665 (2011). doi:10.1021/nn1031627.

[100] Yin, P., Choi, H. M. T., Calvert, C. R., & Pierce, N. A. "Programming biomolecular self-assembly pathways." *Nature*, **451**:318 (2008). doi:10.1038/nature06451.

[101] He, Y. & Liu, D. R. "Autonomous multistep organic synthesis in a single isothermal solution mediated by a DNA walker." *Nat. Nanotechnol.*, **5**:778 (2010). doi:10.1038/nnano.2010.190.

[102] Muscat, R. A., Bath, J., & Turberfield, A. J. "A programmable molecular robot." *Nano Lett.*, **11**:982 (2011). doi:10.1021/nl1037165.

[103] Liu, M., Cheng, J., Tee, S. R., Sreelatha, S., Loh, I. Y., & Wang, Z. "Biomimetic autonomous enzymatic nanowalker of high fuel efficiency." *ACS Nano*, **10**:5882 (2016). doi:10.1021/acsnano.6b01035.

[104] Šulc, P., Ouldridge, T. E., Romano, F., Doye, J. P. K., & Louis, A. A. "Simulating a burnt-bridges DNA motor with a coarse-grained DNA model." *Nat. Comput.*, **13**:535 (2014). doi:10.1007/s11047-013-9391-8.

[105] Omabegho, T., Sha, R., & Seeman, N. C. "A bipedal DNA Brownian motor with coordinated legs." *Science*, **324**:67 (2009). doi:10.1126/science.1170336.

[106] Gu, H., Chao, J., Xiao, S.-J., & Seeman, N. C. "A proximity-based programmable DNA nanoscale assembly line." *Nature*, **465**:202 (2010). doi:10.1038/nature09026.

[107] Zhou, C., Duan, X., & Liu, N. "A plasmonic nanorod that walks on DNA origami." *Nat. Commun.*, **6**:8102 (2015). doi:10.1038/ncomms9102.

[108] Wickham, S. F., Bath, J., Katsuda, Y., Endo, M., Hidaka, K., Sugiyama, H., & Turberfield, A. J. "A DNA-based molecular motor that can navigate a network of tracks." *Nat. Nanotechnol.*, **7**:169 (2012). doi:10.1038/nnano.2011.253.

[109] Lund, K., Manzo, A. J., Dabby, N., Michelotti, N., Johnson-Buck, A., Nangreave, J., Taylor, S., Pei, R., Stojanovic, M. N., Walter, N. G., Winfree, E., & Yan, H. "Molecular robots guided by prescriptive landscapes." *Nature*, **465**:206 (2010). doi:10.1038/nature09012.

[110] Kopperger, E., Pirzer, T., & Simmel, F. C. "Diffusive transport of molecular cargo tethered to a DNA origami platform." *Nano Lett.*, **15**:2693 (2015). doi:10.1021/acs.nanolett.5b00351.

[111] Cha, T.-G., Pan, J., Chen, H., Salgado, J., Li, X., Mao, C., & Choi, J. H. "A synthetic DNA motor that transports nanoparticles along carbon nanotubes." *Nat. Nanotechnol.*, **9**:39 (2014). doi:10.1038/nnano.2013.257.

[112] Yehl, K., Mugler, A., Vivek, S., Liu, Y., Zhang, Y., Fan, M., Weeks, E. R., & Salaita, K. "High-speed DNA-based rolling motors powered by RNase H." *Nat. Nanotechnol.*, **11**:184 (2015). doi:10.1038/nnano.2015.259.

[113] Liedl, T., Högberg, B., Tytell, J., Ingber, D. E., & Shih, W. M. "Self-assembly of three-dimensional prestressed tensegrity structures from DNA." *Nat. Nanotechnol.*, **5**:520 (2010). doi:10.1038/nnano.2010.107.

[114] Zhou, L., Marras, A. E., Su, H. J., & Castro, C. E. "DNA origami compliant nanostructures with tunable mechanical properties." *ACS Nano*, **8**:27 (2014). doi:10.1021/nn405408g.

[115] Shi, Z., Castro, C. E., & Arya, G. "Conformational dynamics of mechanically compliant DNA nanostructures from coarse-grained molecular dynamics simulations." *ACS Nano*, **11**:4617 (2017). doi:10.1021/acsnano.7b00242.

[116] Iwaki, M., Wickham, S. F., Ikezaki, K., Yanagida, T., & Shih, W. M. "A programmable DNA origami nanospring that reveals force-induced adjacent binding of myosin VI heads." *Nat. Commun.*, **7**:13715 (2016). doi:10.1038/ncomms13715.

[117] Funke, J. J. & Dietz, H. "Placing molecules with Bohr radius resolution using DNA origami." *Nat. Nanotechnol.*, **11**:47 (2015). doi:10.1038/nnano.2015.240.

[118] List, J., Falgenhauer, E., Kopperger, E., Pardatscher, G., & Simmel, F. C. "Long-range movement of large mechanically interlocked DNA nanostructures." *Nat. Commun.*, **7**:12414 (2016). doi:10.1038/ncomms12414.

[119] Powell, J. T., Akhuetie-Oni, B. O., Zhang, Z., & Lin, C. "DNA origami rotaxanes: Tailored synthesis and controlled structure switching." *Angew. Chemie*, **55**:11412 (2016). doi:10.1002/anie.201604621.

[120] Yang, Y., Tashiro, R., Suzuki, Y., Emura, T., Hidaka, K., Sugiyama, H., & Endo, M. "A photoregulated DNA-based rotary system and direct observation of its rotational movement." *Chemistry - A Eur. J.*, **23**:3979 (2017). doi:10.1002/chem.201605616.

[121] Marras, A. E., Zhou, L., Su, H.-J., & Castro, C. E. "Programmable motion of DNA origami mechanisms." *Proc. Natl. Acad. Sci.*, **112**:713 (2015). doi:10.1073/pnas.1408869112.

[122] Maier, B., Bensimon, D., & Croquette, V. "Replication by a single DNA polymerase of a stretched single-stranded DNA." *Proc. Natl. Acad. Sci.*, **97**:12002 (2000). doi:10.1073/pnas.97.22.12002.

[123] Ouldridge, T. E., Hoare, R. L., Louis, A. A., Doye, J. P. K., Bath, J., & Turberfield, A. J. "Optimizing DNA nanotechnology through coarse-grained modeling: a two-footed DNA walker." *ACS Nano*, **7**:2479 (2013). doi:10.1021/nn3058483.

[124] Machinek, R. R. F., Ouldridge, T. E., Haley, N. E. C., Bath, J., & Turberfield, A. J. "Programmable energy landscapes for kinetic control of DNA strand displacement." *Nat. Commun.*, **5**:5324 (2014). doi:10.1038/ncomms6324.

[125] Green, A. A., Kim, J., Ma, D., Silver, P. A., Collins, J. J., & Yin, P. "Complex cellular logic computation using ribocomputing devices." *Nature*, **548**:117 (2017). doi:10.1038/nature23271.

[126] Nguyen, H. K. & Southern, E. M. "Minimising the secondary structure of DNA targets by incorporation of a modified deoxynucleoside: Implications for nucleic acid analysis by hybridisation." *Nucleic Acids Res.*, **28**:3904 (2000). doi:10.1093/nar/28.20.3904.

[127] Zhang, D. Y., Chen, S. X., & Yin, P. "Optimizing the specificity of nucleic acid hybridization." *Nat. Chem.*, **4**:208 (2012). doi:10.1038/nchem.1246.

[128] Song, T., Xiao, S., Yao, D., Huang, F., Hu, M., & Liang, H. "An efficient DNA-fueled molecular machine for the discrimination of single-base changes." *Adv. Mater.*, **26**:6181 (2014). doi:10.1002/adma.201402314.

[129] Gao, Z. F., Ling, Y., Lu, L., Chen, N. Y., Luo, H. Q., & Li, N. B. "Detection of single-nucleotide polymorphisms using an ON-OFF switching of regenerated biosensor based on a locked nucleic acid-integrated and toehold-mediated strand displacement reaction." *Anal. Chem.*, **86**:2543 (2014). doi:10.1021/ac500362z.

[130] Byrom, M., Bhadra, S., Jiang, Y. S., & Ellington, A. D. "Exquisite allele discrimination by toehold hairpin primers." *Nucleic Acids Res.*, **42**:e120 (2014). doi:10.1093/nar/gku558.

[131] Khodakov, D. A., Khodakova, A. S., Huang, D. M., Linacre, A., & Ellis, A. V. "Protected DNA strand displacement for enhanced single nucleotide discrimination in double-stranded DNA." *Sci. Rep.*, **5**:8721 (2015). doi:10.1038/srep08721.

[132] Hwang, M. T., Landon, P. B., Lee, J., Choi, D., Mo, A. H., Glinsky, G., & Lal, R. "Highly specific SNP detection using 2D graphene electronics and DNA strand displacement." *Proc. Natl. Acad. Sci.*, **113**:7088 (2016). doi:10.1073/pnas.1603753113.

[133] Ponder, J. W. & Case, D. A. "Force fields for protein simulations." *Adv. Protein Chem.*, **66**:27 (2003). doi:10.1016/S0065-3233(03)66002-X.

[134] Oostenbrink, C., Villa, A., Mark, A. E., & Van Gunsteren, W. F. "A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6." *J. Comput. Chem.*, **25**:1656 (2004). doi:10.1002/jcc.20090.

[135] MacKerell Jr, A. "Atomistic models and force fields." In *Comput. Biochem. Biophys.* CRC Press (2001). ISBN 978-0-8247-0455-1. doi:10.1201/9780203903827.ch2.

[136] Ouldridge, T. E., Louis, A. A., & Doye, J. P. K. "Structural, mechanical, and thermodynamic properties of a coarse-grained DNA model." *J. Chem. Phys.*, **134**:085101 (2011). doi:10.1063/1.3552946.

[137] Ouldridge, T. E., Louis, A. A., & Doye, J. P. K. "DNA nanotweezers studied with a coarse-grained model of DNA." *Phys. Rev. Lett.*, **104**:178101 (2010). doi:10.1103/PhysRevLett.104.178101.

[138] Šulc, P., Romano, F., Ouldridge, T. E., Rovigatti, L., Doye, J. P. K., & Louis, A. A. "Sequence-dependent thermodynamics of a coarse-grained DNA model." *J. Chem. Phys.*, **137**:135101 (2012). doi:10.1063/1.4754132.

[139] Snodin, B. E., Randisi, F., Mosayebi, M., Šulc, P., Schreck, J. S., Romano, F., Ouldridge, T. E., Tsukanov, R., Nir, E., Louis, A. A., & Doye, J. P. "Introducing improved structural properties and salt dependence into a coarse-grained model of DNA." *J. Chem. Phys.*, **142** (2015). doi:10.1063/1.4921957.

[140] Romano, F., Chakraborty, D., Doye, J. P. K., Ouldridge, T. E., & Louis, A. A. "Coarse-grained simulations of DNA overstretching." *J. Chem. Phys.*, **138**:085101 (2013). doi:10.1063/1.4792252.

[141] Schreck, J. S., Ouldridge, T. E., Romano, F., Šulc, P., Shaw, L. P., Louis, A. A., & Doye, J. P. K. "DNA hairpins destabilize duplexes primarily by promoting melting rather than by inhibiting hybridization." *Nucleic Acids Res.*, **43**:6181 (2015). doi:10.1093/nar/gkv582.

[142] Mosayebi, M., Louis, A. A., Doye, J. P. K., & Ouldridge, T. E. "Force-induced rupture of a DNA duplex: From fundamentals to force sensors." *ACS Nano*, **9**:11993 (2015). doi:10.1021/acsnano.5b04726.

[143] Ding, W., Deng, W., Zhu, H., & Liang, H. "Metallo-toeholds: Controlling DNA strand displacement driven by Hg(II) ions." *Chem. Commun.*, **49**:9953 (2013). doi:10.1039/c3cc45373a.

[144] Genot, A. J., Zhang, D. Y., Bath, J., & Turberfield, A. J. "Remote toehold: a mechanism for flexible control of DNA hybridization kinetics." *J. Am. Chem. Soc.*, **133**:2177 (2011). doi:10.1021/ja1073239.

[145] Li, C., Li, Y., Chen, Y., Lin, R., Li, T., Liu, F., & Li, N. "Modulating the DNA strand-displacement kinetics with the one-sided remote toehold design for differentiation of single-base mismatched DNA." *RSC Adv.*, **6**:74913 (2016). doi:10.1039/C6RA17006D.

[146] Li, F., Tang, Y., Traynor, S. M., Li, X. F., & Le, X. C. "Kinetics of proximity-induced intramolecular DNA strand displacement." *Anal. Chem.*, **88**:8152 (2016). doi:10.1021/acs.analchem.6b01900.

[147] Panyutin, I. G. & Hsieh, P. "Formation of a single base mismatch impedes spontaneous DNA branch migration." *J. Mol. Biol.*, **230**:413 (1993). doi:10.1006/jmbi.1993.1159.

[148] Broadwater, D. W. & Kim, H. D. "The effect of basepair mismatch on DNA strand displacement." *Biophys. J.*, **110**:1476 (2016). doi:10.1016/j.bpj.2016.02.027.

[149] Dunn, K. E., Trefzer, M. A., Johnson, S., & Tyrrell, A. M. "Investigating the dynamics of surface-immobilized DNA nanomachines." *Sci. Rep.*, **6**:29581 (2016). doi:10.1038/srep29581.

[150] Zadeh, J. N., Steenberg, C. D., Bois, J. S., Wolfe, B. R., Pierce, M. B., Khan, A. R., Dirks, R. M., & Pierce, N. A. "NUPACK: Analysis and design of nucleic acid systems." *J. Comput. Chem.*, **32**:170 (2011). doi:10.1002/jcc.21596.

[151] Johansson, M. K., Fidder, H., Dick, D., & Cook, R. M. "Intramolecular dimers: A new strategy to flourescence quenching in dual-labeled oligonucleotide probes." *J. Am. Chem. Soc.*, **124**:6950 (2002). doi:10.1021/ja025678o.

[152] Green, S. J. "DNA nanomachines." Ph.D. thesis, University of Oxford (2006).

[153] Landau, L. D. & Lifschitz, E. M. *Theory of elasticity*. Pergamon press, New York (1986).

[154] Broedersz, C. P. & Mackintosh, F. C. "Modeling semiflexible polymer networks." *Rev. Mod. Phys.*, **86**:995 (2014). doi:10.1103/RevModPhys.86.995.

[155] Revyakin, A., Allemand, J. F., Croquette, V., Ebright, R. H., & Strick, T. R. "Single-molecule DNA nanomanipulation: detection of promoter-unwinding events by RNA polymerase." *Methods Enzymol.*, **370**:577 (2003). doi:10.1016/S0076-6879(03)70049-4.

[156] Wickham, S. F. "DNA origami: A substrate for the study of molecular motors." Ph.D. thesis, University of Oxford (2011).

[157] Kauert, D. J., Kurth, T., Liedl, T., & Seidel, R. "Direct mechanical measurements reveal the material properties of three-dimensional DNA origami." *Nano Lett.*, **11**:5558 (2011). doi:10.1021/nl203503s.

[158] Messing, J. "Cloning in M13 phage or how to use biology at its best." *Gene*, **100**:3 (1991). doi:10.1016/0378-1119(91)90344-B.

[159] Hofschneider, P. H. & Preuss, A. "M13 bacteriophage liberation from intact bacteria as revealed by the electron microscope." *J. Mol. Biol.*, **7**:450 (1963).

[160] Gronenborn, B. & Messing, J. "Methylation of single-stranded DNA in vitro introduces new restriction endonuclease cleavage sites." *Nature*, **272**:375 (1978). doi:10.1038/272375a0.

[161] Messing, J., Crea, R., & Seeburg, P. H. "A system for shotgun DNA sequencing." *Nucleic Acids Res.*, **9**:309 (1981). doi:10.1093/nar/9.2.309.

[162] Norrander, J., Kempe, T., & Messing, J. "Construction of improved M13 vectors using oligodeoxynucleotide-directed mutagenesis." *Gene*, **26**:101 (1983). doi:10.1016/0378-1119(83)90040-9.

[163] Langley, K. E., Villarejo, M. R., Fowler, A. V., Zamenhof, P. J., & Zabin, I. "Molecular basis of beta-galactosidase alpha-complementation." *Proc. Natl. Acad. Sci.*, **72**:1254 (1975). doi:10.1073/pnas.72.4.1254.

[164] Yanisch-Perron, C., Vieira, J., & Messing, J. "Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mpl8 and pUC19 vectors." *Gene*, **33**:103 (1985). doi:10.1016/0378-1119(85)90120-9.

[165] `https://www.neb.com/tools-and-resources/usage-guidelines/cleavage-close-to-the-end-of-dna-fragments` "Cleavage close to the end of DNA Fragments | NEB." *New England Biolabs*, [Accessed 09/03/2017].

[166] Sambrook, J. & Russell, D. W. *Molecular cloning : A laboratory manual*. Cold Spring Harbor Laboratory Press (2001). ISBN 0879695773.

[167] Rapley, R. *The nucleic acid protocols handbook*. Part I. Humana Press (2000). ISBN 0896034593.

[168] Johnson, I. D. & Spence, M. T. Z. *The molecular probes handbook: A guide to fluorescent probes and labeling technologies*. Life Technologies Corporation, 11 edition (2010). ISBN 978-0-9829279.

[169] Nickels, P. C., Wünsch, B., Holzmeister, P., Bae, W., Kneer, L. M., Grohmann, D., Tinnefeld, P., & Liedl, T. "Molecular force spectroscopy with a DNA origami-based nanoscopic force clamp." *Science*, **354**:305 (2016). doi:10.1126/science.aah5974.

[170] Morrison, L. E. & Stols, L. M. "Sensitive fluorescence-based thermodynamic and kinetic measurements of DNA hybridization in solution." *Biochemistry*, **32**:3095 (1993). doi:10.1021/bi00063a022.

[171] Howorka, S., Movileanu, L., Braha, O., & Bayley, H. "Kinetics of duplex formation for individual DNA strands within a single protein nanopore." *Proc. Natl. Acad. Sci.*, **98**:12996 (2001). doi:10.1073/pnas.231434698.

[172] Gao, Y., Wolf, L. K., & Georgiadis, R. M. "Secondary structure effects on DNA hybridization kinetics: A solution versus surface comparison." *Nucleic Acids Res.*, **34**:3370 (2006). doi:10.1093/nar/gkl422.

[173] Ouldridge, T. E., Šulc, P., Romano, F., Doye, J. P. K., & Louis, A. A. "DNA hybridization kinetics: Zippering, internal displacement and sequence dependence." *Nucleic Acids Res.*, **41**:8886 (2013). doi:10.1093/nar/gkt687.

[174] Zhang, J. X., Fang, J. Z., Duan, W., Wu, L., Zhang, A., Dalchau, N., Yordanov, B., Petersen, R., Phillips, A., & Zhang, D. "Predicting DNA hybridization kinetics from sequence." *bioRxiv* (2017). doi:10.1101/149427.

[175] Schoen, I., Krammer, H., & Braun, D. "Hybridization kinetics is different inside cells." *Proc. Natl. Acad. Sci.*, **106**:21649 (2009). doi:10.1073/pnas.0901313106.

[176] Johnson-Buck, A. & Walter, N. G. "Discovering anomalous hybridization kinetics on DNA nanostructures using single-molecule fluorescence microscopy." *Methods*, **67**:177 (2014). doi:10.1016/j.ymeth.2014.02.032.

[177] Pinheiro, A. V., Nangreave, J., Jiang, S., Yan, H., & Liu, Y. "Steric crowding and the kinetics of DNA hybridization within a DNA nanostructure system." *ACS Nano*, **6**:5521 (2012). doi:10.1021/nn301448y.

[178] Whitley, K. D., Comstock, M. J., & Chemla, Y. R. "Elasticity of the transition state for oligonucleotide hybridization." *Nucleic Acids Res.*, **45**:547 (2017). doi:10.1093/nar/gkw1173.

[179] Zhang, Y., Zhou, H., & Ou-Yang, Z.-C. C. "Stretching single-stranded DNA: interplay of electrostatic, base-pairing, and base-pair stacking interactions." *Biophys. J.*, **81**:1133 (2001). doi:10.1016/S0006-3495(01)75770-0.

[180] Montanari, A. & Mézard, M. "Hairpin formation and elongation of biomolecules." *Phys. Rev. Lett.*, **86**:2178 (2001). doi:10.1103/PhysRevLett.86.2178.

[181] Dessinges, M.-N., Maier, B., Zhang, Y., Peliti, M., Bensimon, D., & Croquette, V. "Stretching single stranded DNA, a model polyelectrolyte." *Phys. Rev. Lett.*, **89**:248102 (2002). doi:10.1103/PhysRevLett.89.248102.

[182] Arrhenius, S. "Über die Dissociationswärme und den Einfluss der Temperatur auf den Dissociationsgrad der Elektrolyte." *Zeitschrift für Phys. Chemie*, **4**:96 (1889). doi:10.1515/zpch-1889-0108.

[183] Baumann, C. G., Smith, S. B., Bloomfield, V. A., & Bustamante, C. "Ionic effects on the elasticity of single DNA molecules." *Proc. Natl. Acad. Sci.*, **94**:6185 (1997). doi:10.1073/pnas.94.12.6185.

[184] Strick, T., Allemand, J.-F., Bensimon, D., & Croquette, V. "Behavior of supercoiled DNA." *Biophys. J.*, **74**:2016 (1998). doi:10.1016/S0006-3495(98)77908-1.

[185] Rouzina, I. & Bloomfield, V. A. "Force-induced melting of the DNA double helix 1. Thermodynamic analysis." *Biophys. J.*, **80**:882 (2001). doi:10.1016/S0006-3495(01)76067-5.

[186] Rouzina, I. & Bloomfield, V. A. "Force-induced melting of the DNA double helix. 2. Effect of Solution conditions." *Biophys. J.*, **80**:894 (2001). doi:10.1016/S0006-3495(01)76068-7.

[187] Moreira, B. G., You, Y., Behlke, M. A., & Owczarzy, R. "Effects of fluorescent dyes, quenchers, and dangling ends on DNA duplex stability." *Biochem. Biophys. Res. Commun.*, **327**:473 (2005). doi:10.1016/j.bbrc.2004.12.035.

[188] Friel, C. T. & Howard, J. "Coupling of kinesin ATP turnover to translocation and microtubule regulation: one engine, many machines." *J. Muscle Res. Cell Motil.*, **33**:377 (2012). doi:10.1007/s10974-012-9289-6.

[189] Ouldridge, T. E. "Coarse-grained modelling of DNA and DNA self-assembly." Ph.D. thesis, University of Oxford (2011).

[190] McWilliam, H., Li, W., Uludag, M., Squizzato, S., Park, Y. M., Buso, N., Cowley, A. P., & Lopez, R. "Analysis tool web services from the EMBL-EBI." *Nucleic Acids Res.*, **41**:W597 (2013). doi:10.1093/nar/gkt376.

[191] Sugimoto, N., Nakano, S.-i., Katoh, M., & Matsumura, A. "Thermodynamic parameters to predict stability of RNA/DNA hybrid duplexes." *Biochemistry*, **34**:11211 (1995). doi:10.1021/bi00035a029.

[192] Owczarzy, R., Moreira, B. G., You, Y., Behlke, M. A., & Walder, J. A. "Predicting stability of DNA duplexes in solutions containing magnesium and monovalent cations." *Biochemistry*, **47**:5336 (2008). doi:10.1021/bi702363u.

[193] `https://eu.idtdna.com/pages/support/technical-vault/faq/product/why-is-page-purification-not-available-with-fluorescent-dyes-` "Integrated DNA Technologies FAQ." *Integrated DNA Technologies*, [Accessed 21/07/2017].

[194] Hogrefe, H. H., Hogrefe, R. I., Walder, R. Y., & Walder, J. A. "Kinetic analysis of Escherichia coli RNase H using DNA-RNA-DNA/DNA substrates." *J. Biol. Chem.*, **265**:5561 (1990).

[195] Keck, J. L., Goedken, E. R., & Marqusee, S. "Activation/attenuation model for RNase H: A one-metal mechanism with second-metal inhibition." *J. Biol. Chem.*, **273**:34128 (1998). doi:10.1074/jbc.273.51.34128.

[196] Helenius, J., Brouhard, G., Kalaidzidis, Y., Diez, S., & Howard, J. "The depolymerizing kinesin MCAK uses lattice diffusion to rapidly target microtubule ends." *Nature*, **441**:115 (2006). doi:10.1038/nature04736.

[197] Lu, H., Ali, M. Y., Bookwalter, C. S., Warshaw, D. M., & Trybus, K. M. "Diffusive movement of processive kinesin-1 on microtubules." *Traffic*, **10**:1429 (2009). doi:10.1111/j.1600-0854.2009.00964.x.

[198] Benn, F. "Functional DNA Nanotechnology." Ph.D. thesis, University of Oxford (2016).

[199] Boemo, M. A., Lucas, A. E., Turberfield, A. J., & Cardelli, L. "The formal language and design principles of autonomous DNA walker circuits." *ACS Synth. Biol.*, **5**:878 (2016). doi:10.1021/acssynbio.5b00275.

[200] Maier, A. M., Bae, W., Schiffels, D., Emmerig, J. F., Schiff, M., & Liedl, T. "Self-assembled DNA tubes forming helices of controlled diameter and chirality." *ACS Nano*, **11**:1301 (2017). doi:10.1021/acsnano.6b05602.

[201] Yin, P., Hariadi, R. F., Sahu, S., Choi, H. M. T., Park, S. H., LaBean, T. H., & Reif, J. H. "Programming DNA tube circumferences." *Science*, **321**:824 (2008). doi:10.1126/science.1157312.

[202] Kuzyk, A., Schreiber, R., Fan, Z., Pardatscher, G., Roller, E.-M., Högele, A., Simmel, F. C., Govorov, A. O., & Liedl, T. "DNA-based self-assembly of chiral plasmonic nanostructures with tailored optical response." *Nature*, **483**:311 (2012). doi:10.1038/nature10889.

[203] Thompson, R. E., Larson, D. R., & Webb, W. W. "Precise nanometer localization analysis for individual fluorescent probes." *Biophys. J.*, **82**:2775 (2002). doi:10.1016/S0006-3495(02)75618-X.

[204] Lucas, A. E. "Dynamic DNA motors and structures." Ph.D. thesis, University of Oxford (2016).

[205] Wu, N., Czajkowsky, D. M., Zhang, J., Qu, J., Ye, M., Zeng, D., Zhou, X., Hu, J., Shao, Z., Li, B., & Fan, C. "Molecular threading and tunable molecular recognition on DNA origami nanostructures." *J. Am. Chem. Soc.*, **135**:12172 (2013). doi:10.1021/ja403863a.

[206] Selvin, P. R. & Ha, T. *Single-molecule techniques : A laboratory manual*. Cold Spring Harbor Laboratory Press (2008). ISBN 9780879697761.

[207] Cooper, D., Uhm, H., Tauzin, L. J., Poddar, N., & Landes, C. F. "Photobleaching lifetimes of cyanine fluorophores used for single-molecule Förster resonance energy transfer in the presence of various photoprotection systems." *ChemBioChem*, **14**:1075 (2013). doi:10.1002/cbic.201300030.

[208] Benesch, R. E. & Benesch, R. "Enzymatic Removal of Oxygen for Polarography and Related Methods." *Science*, **118** (1953).

[209] Rasnik, I., McKinney, S. A., & Ha, T. "Nonblinking and long-lasting single-molecule fluorescence imaging." *Nat. Methods*, **3**:891 (2006). doi:10.1038/nmeth934.

[210] Dave, R., Terry, D. S., Munro, J. B., & Blanchard, S. C. "Mitigating unwanted photophysical processes for improved single-molecule fluorescence imaging." *Biophys. J.*, **96**:2371 (2009). doi:10.1016/j.bpj.2008.11.061.

[211] Grossman, P. D. & Soane, S. S. "Orientation effects on the electrophoretic mobility of rod-shaped molecules in free solution." *Anal. Chem.*, **62**:1592 (1990). doi:10.1021/ac00214a011.

[212] Flory, P. J. "Molecular size distribution in linear condensation polymers." *J. Am. Chem. Soc.*, **58**:1877 (1936). doi:10.1021/ja01301a016.

[213] Ovesný, M., Křížek, P., Borkovec, J., Švindrych, Z., & Hagen, G. M. "ThunderSTORM: A comprehensive ImageJ plug-in for PALM and STORM data analysis and super-resolution imaging." *Bioinformatics*, **30**:2389 (2014). doi:10.1093/bioinformatics/btu202.

[214] Schnitzbauer, J., Strauss, M. T., Schlichthaerle, T., Schueder, F., & Jungmann, R. "Super-resolution microscopy with DNA-PAINT." *Nat. Protoc.*, **12**:1198 (2017). doi:10.1038/nprot.2017.024.

[215] Selvin, P. R., Lougheed, T., Tonks Hoffman, M., Park, H., Balci, H., Blehm, B. H., & Toprak, E. "Fluorescence imaging with one-nanometer accuracy (FIONA)." *CSH. Protoc.*, **2007**:pdb.top27 (2007). doi:10.1101/PDB.TOP27.

[216] Verma, N. C., Khan, S., & Nandi, C. K. "Single-molecule analysis of fluorescent carbon dots towards localization-based super-resolution microscopy." *Methods Appl. Fluoresc.*, **4**:044006 (2016). doi:10.1088/2050-6120/4/4/044006.

[217] Schreiber, R., Santiago, I., Ardavan, A., & Turberfield, A. J. "Ordering gold nanoparticles with DNA origami nanoflowers." *ACS Nano*, **10**:7303 (2016). doi:10.1021/acsnano.6b03076.

[218] Gust, A., Zander, A., Gietl, A., Holzmeister, P., Schulz, S., Lalkens, B., Tinnefeld, P., & Grohmann, D. "A starting point for fluorescence-based single-molecule measurements in biomolecular research." *Molecules*, **19**:15824 (2014). doi:10.3390/molecules191015824.

[219] Edelstein, A., Tsuchida, M. A., Amodaj, N., Pinkard, H., Vale, R. D., & Stuurman, N. "Advanced methods of microscope control using μManager software." *J. Biol. Methods*, **1**:10 (2014). doi:10.14440/jbm.2014.36.

[220] Edelstein, A., Amodaj, N., Hoover, K., Vale, R., & Stuurman, N. "Computer control of microscopes using μManager." *Curr. Protoc. Mol. Biol.*, **92**:14.20.1 (2010). doi:10.1002/0471142727.mb1420s92.

[221] "Valap for agar mounts." *CSH. Protoc.*, **2010**:pdb.rec12380 (2010). doi:10.1101/pdb.rec12380.