Control and Observation of DNA Nanodevices

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

Robert R. F. Machinek

Merton College

University of Oxford

Trinity Term 2014

Robert R. F. Machinek: *Control and Observation of DNA Nanodevices*

Trinity Term 2014

Control and Observation of DNA Nanodevices

A thesis submitted for the degree of Doctor of Philosophy by Robert R. F. Machinek, Merton College and Department of Physics, University of Oxford, Trinity Term 2014.

Abstract

The uniquely predictable and controllable binding mechanism of DNA strands has been exploited to construct a vast range of synthetic nanodevices, capable of autonomous motion and computation. This thesis proposes novel ideas for the control and observation of such devices. The first of these proposals hinges on introducing mismatched base pairs into toehold-mediated strand displacement – a fundamental primitive in most dynamic DNA devices and reaction networks. Previous findings that such mismatches can impede strand displacement are extended insofar as it is shown that this impediment is highly dependent on mismatch position. This discovery is examined in detail, both experimentally and through simulations created with a coarse-grained model of DNA. It is shown that this effect allows for kinetic control of strand displacement decoupled from reaction thermodynamics.

The second proposal improves upon a previously presented strand displacement scheme, in which two strands perform displacement cooperatively. This scheme is extended to be cascadable, so that the output of one such reaction serves as input to the next. This scheme is implemented in reaction networks capable of performing fundamental calculations on directed graphs.

The third proposal is exclusively concerned with a novel observation methodology. This method is based on single-molecule fluorescence microscopy, and uses quantum dots, a fluorescent type of semiconductor nanocrystal, as a label. These quantum dots display a set of characteristics particularly promising for single-molecule studies on the time- and length scales most commonly encountered in DNA nanotechnology. This method is shown to allow for highly precise measurements on static DNA devices. Preliminary data for the observation of a complex dynamic device is also presented.

ACKNOWLEDGMENTS

Many people around me contributed to this work directly or indirectly, and I would like to take the opportunity to express my gratitude towards them. First and foremost, I would like to thank my supervisor, Andrew Turberfield. His guidance throughout the past years has had a tremendous impact on me, my experiments, this thesis and the publications derived from it. Jon Bath's assistance has been of a more pragmatic, but similarly significant nature, as he taught me several experimental and computational techniques in addition to providing me with creative solutions to seemingly intractable problems I encountered in my research.

I have had the chance to work with three external collaborators to whom I am greatly indebted: Tom Ouldridge provided important simulation results and theoretical aspects to my work, Chris Thachuk provided insights and ideas from a computer scientist's point of view, and Stephan Uphoff shared Matlab code for microscopy data analysis with me. Working with Natalie Haley and Sam Tusk, two extraordinarily gifted former master's (now DPhil) students, taught me at least as much as I hope I was able to teach them in return. I have been very lucky to be part of the Oxford Physics Department's DNA group, whose past and present members I wish to thank for being such engaging, helpful and friendly colleagues.

Whereas all of the aforementioned provided me with a stimulating academic environment, the members and staff of Merton College provided me with a fantastic social one. The many friends I found there will be dearly missed.

My studies and research would have been impossible without the funding of several institutions whom I would like to thank for their generosity: the Engineering and Physical Sciences Research Council, the Foundation of German Business (Stiftung der deutschen Wirtschaft), the German National Merit Foundation (Studienstiftung des deutschen Volkes) as well as the Bayer Science and Education Foundation.

I would further like to thank my parents: my father for his limitless logistic support as well as his advice on navigating the academic world, and my mother for her advice on the world beyond it. Lastly, I am infinitely grateful to my siblings, who throughout my life have guided me with their ample provision of the fiercest criticism whenever I was wrong, and the staunchest support whenever I was right. In the hope that this remains so in the future, this work is dedicated to them.

Meinen Geschwistern

CONTENTS

ACRONYMS

A	adenine
AFM	atomic force microscopy
ATP	adenosine triphosphate
BME	beta-mercaptoethanol
BSA	bovine serum albumin
C	cytosine
CRN	chemical reaction network
Cy3	Cyanine 3 dye
Cy5	Cyanine 5 dye
DNA	deoxyribonucleic acid
DNA-PAINT	DNA points accumulation for imaging in nanoscale topography
EDTA	ethylenediaminetetraacetic acid
EM	electron microscopy
EMCCD	electron-multiplying charge-coupled device
FIONA	fluorescence imaging with one-nanometre accuracy
FRET	fluorescence resonance energy transfer
FWHM	full width at half maximum

G	guanine
HPLC	high-performance liquid chromatography
I/O	input/output
PALM	photo-activated localisation microscopy
PCR	polymerase chain reaction
PEG	polyethylene glycol
PSF	point spread function
RLC	resistor-inductor-capacitor
SHREC	single-molecule high-resolution colocalisation
SNR	signal-to-noise ratio
STORM	stochastic optical reconstruction microscopy
T	thymine
TIRF	total internal reflection fluorescence
TMSD	toehold-mediated strand displacement
Tris	tris(hydroxymethyl)aminomethane

1

INTRODUCTION

"What are the possibilities of small but movable machines? They may or may not be useful, but they surely would be fun to make."

- R. P. Feynman

In 1959, Richard Feynman gave a seminal talk in which he noted an inconsistency in the evolution of science: researchers had gained access to the extremes of certain properties of physical systems, for example being able to cool materials to within one kelvin of absolute zero, or to create pressures exceeding that of the atmosphere by five orders of magnitude. The ability to control systems at small, i.e. molecular or atomic length scales, however, was still severely limited. Feynman recognised the utility in developing such capacities, not only for improved understanding of naturally occurring processes, but also for new technologies, such as more powerful information storage and processing devices [34].

Feynman also offered the vision of microscopic machines. Although he identified some of the challenges of machine design at such length scales, like the decreased importance of inertia and the rapid dissipation of heat, he drew the inspiration for such devices quite clearly from man-made, macroscopic machinery. We know now that biological molecular machines provide a better guide, as they have evolved their functionalities on the nano- and microscale over billions of years. They are fundamental to some of the most important biological processes, such as organism locomotion, nucleic acid transcription, and intracellular organisation. Due to their high specialisation to certain tasks and environ-

ments, they are difficult to integrate into controllable, complex synthetic systems. Despite this, these proteins provide proof that machinery at such scales is viable; that there exists no fundamental, physical restriction which would prohibit similar, man-made devices.

This knowledge alone, however, does not provide a recipe for how such devices can be constructed. The biological synthesis of molecular machines is a highly complex process that remains the subject of considerable research efforts. Though rapidly evolving, our understanding of these processes is still insufficient to regard artificial, rationally designed proteins as a realistic approach to nanoscale engineering. However, as the following introduction attempts to showcase, biology does provide us with an alternative, somewhat unlikely building material, deoxyribonucleic acid (DNA).

1.1 DNA NANOTECHNOLOGY

1.1.1 *DNA*

DNA is a polymer, which in a biological context assumes the central role of serving as the carrier of genetic information. This information is encoded in the sequence of the monomers, or nucleotides, each of which carries one of four distinct bases, adenine (A), cytosine (C), guanine (G) and thymine (T). A directionality in this sequence can be inferred from the asymmetrical structure of the polymer backbone. The bases have remarkable pairwise binding properties, with A forming two hydrogen bonds with T, and C forming three hydrogen bonds with G. This allows two anti-parallel strands of DNA to form a stable duplex, if their sequences are complementary, as illustrated in Figure 1.1 [139]. Indeed, duplex DNA is the most commonly found configuration in nature. This

structure is intimately connected to the replication of DNA, and thus the inheritance of genetic information: separating the two strands and using each as a template to synthesise a complement yields two copies of the original duplex.

In physiological conditions, the DNA duplex is a double-helical structure with a 3.6 nm pitch. The width of such a duplex is 2 nm, with each individual nucleotide contributing 0.34 nm to its length. This helix is remarkably rigid on the nanoscale, with a persistence length of around 25 nm [105]. For single strands, the persistence length is around 1 nm [131]. This is, of course, a very incomplete picture of the known properties of DNA, but it is sufficient for the introduction DNA as a nanoscale building material.

1.1.2 *Synthetic DNA nanostructures*

The original idea to remove DNA from its biological context to create artificial DNA structures was first promoted by Nadrian Seeman in 1982 [113]. He envisioned that by using synthetic oligonucleotides with predefined sequence, it would be possible to create immobile junctions of DNA duplexes. The sequences of the component strands would be synthesised in such a way that the number of base pairs is maximised when the systems assumes a junction conformation, thus minimising system free energy. This fundamental idea is depicted in Figure 1.2. The most striking advantage of this idea over other approaches to manipulation on the molecular scale (such as the use of a scanning tunneling microscope [30]) is that of self-assembly: instead of having to position each component of the desired structure individually, oligonucleotides with an appropriately designed and synthesised sequence would form such junctions with no outside control. Instead of serial production of indi-

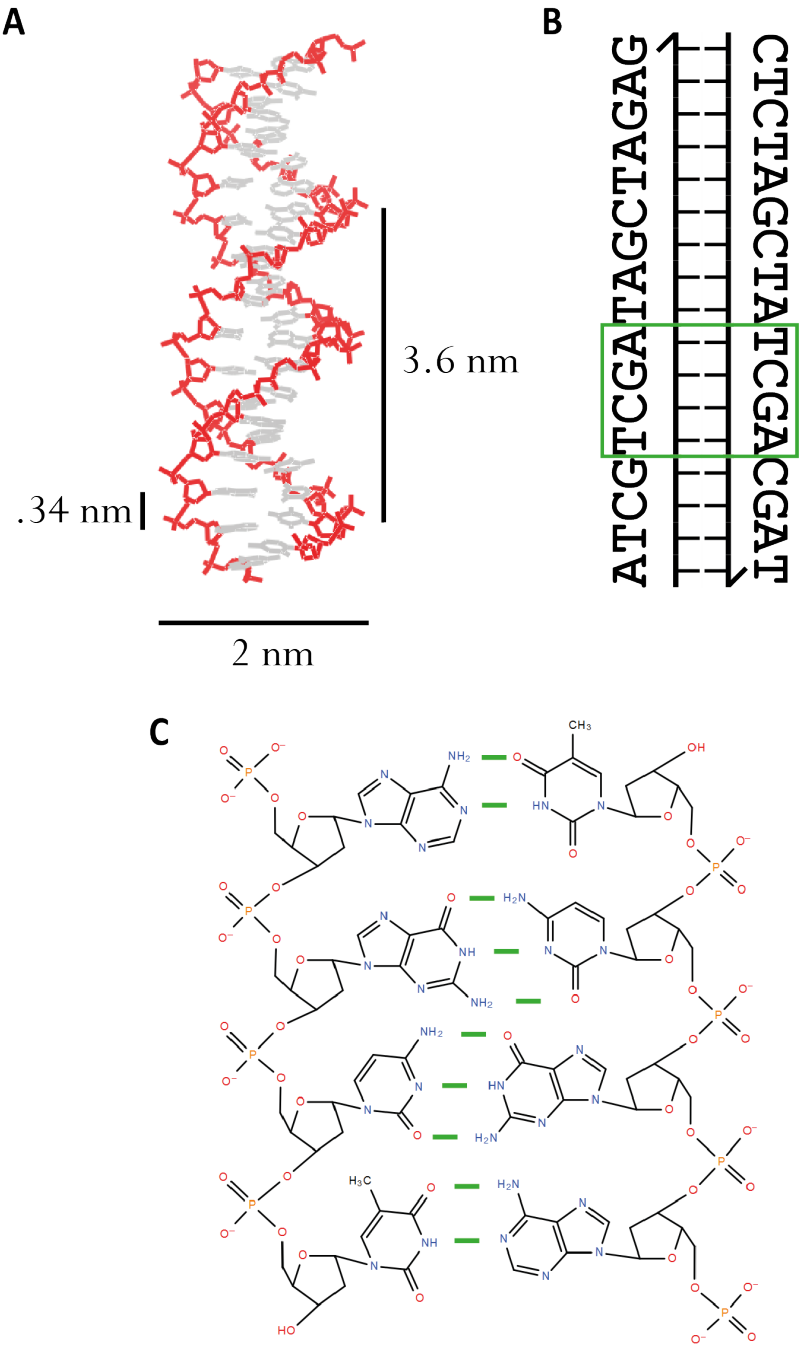

Figure 1.1: **A:** Model of a 17-base pair DNA duplex. Phosphate-deoxyribose backbone is coloured red, the nucleobases grey. Nucleotide-nucleotide distance, duplex width and helical pitch are indicated by scale bars (created with 3D-DART [24] and rendered with Chimera [94]). **B:** Simplified depiction of the same duplex. Structural details are omitted for a clearer depiction of base pairs. Arrowheads indicate the directionality of DNA, pointing out the 3′-end of each oligonucleotide. **C:** Structural formula of the four base pairs enclosed by the green rectangle in the simplified depiction. The base pairs form via hydrogen bonds (green).

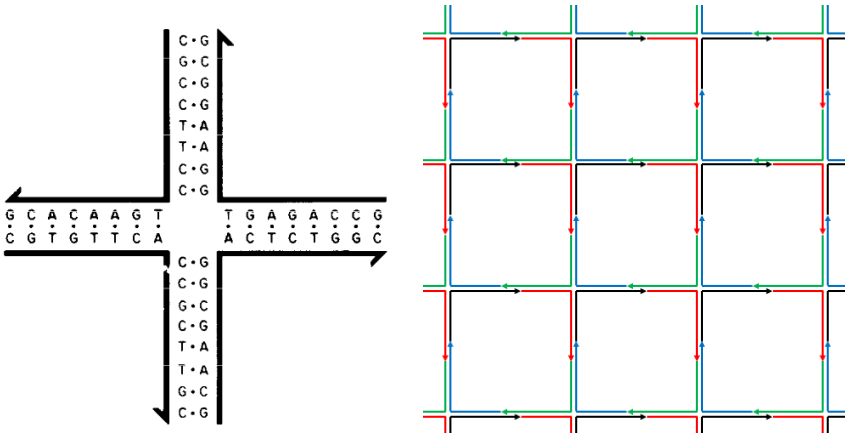

Figure 1.2: **Left:** An immobile junction consisting of four DNA oligonucleotides, as suggested by Seeman (adapted from [113], with permission from Elsevier, ©1982). This is an important primitive for more complex, synthetic DNA nanostructures. **Right:** Modifying the constituent oligonucleotides to form complementary single-stranded overhangs at each branch of the junction allows, at least in theory, for the construction of a 2-dimensional, periodic lattice.

vidual structures, a huge number of structures can assemble in parallel in the same reaction vessel. Furthermore, Seeman noted that these junctions could be connected via complementary sticky ends, single-stranded overhangs from a duplex, forming periodic lattices, which, for example, could be used for the periodic arrangement of proteins and their subsequent crystallographic characterisation.

An experimental demonstration of such junctions soon followed Seeman's original proposal [59]. The creation of lattices with long-range order proved elusive, however, as the originally presented junctions are structurally not very rigid. Robust lattices were achieved in two dimensions in 1998 [144], with three-dimensional DNA crystals following eleven years later [159]. In the meantime, however, DNA nanotechnology had become a well-established field, in which the original objectives set out in Seeman's 1982 paper had been expanded drastically. In 1991, it had already been demonstrated that nanoscale polyhedra can be created out of short strands of DNA alone [16]; several more complex and more robust examples followed (Figure 1.3, [158, 41]).

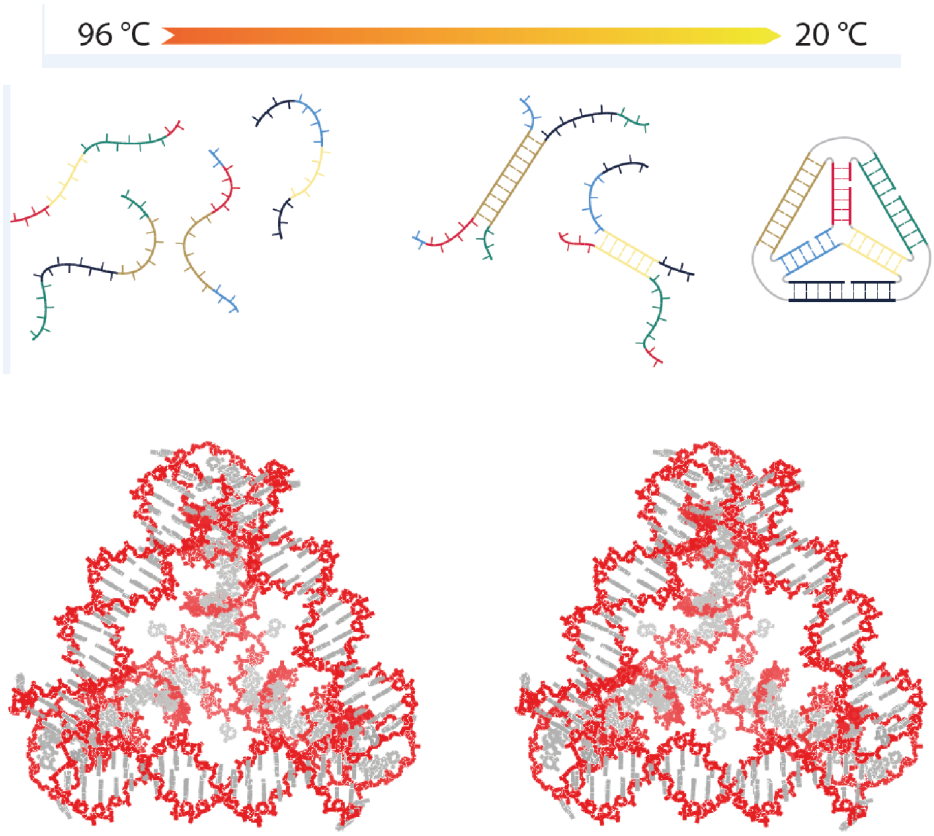

Figure 1.3: **Top:** Self-assembly of a nanoscale DNA tetrahedron from four oligonucleotides, as demonstrated by Goodman and colleagues [41]. The sequences of these oligonucleotides are designed in such a fashion that domains of the same colour on different strands are complementary to each other. The yield is greatly increased by annealing, i.e. heating of the sample to 96 °C, where no bases are paired, and subsequent, slow cooling. This process leads to a comparatively rapid equilibration of the sample, guaranteeing high fidelity in domain hybridisation. Following assembly, the nicks between the strands can be enzymatically ligated. Some graphic elements of this figure were kindly provided by Jonathan Bath (Department of Physics, University of Oxford). **Bottom:** Stereo image of a molecular model of such a tetrahedron (rendered with Chimera [94]). Each edge of the structure measures approximately 8 nm.

In 2006, Paul Rothemund presented with 'DNA origami' a truly revolutionary idea to construct nanostructures [103]. Relying on short oligonucleotides alone makes the production of complex structures at high yield very complicated; creating a long, synthetic sequence which can fold onto itself to create a desired structure is very limited in versatility [118]. Instead, Rothemund used a long, viral 7-kilobase 'scaffold' strand, which was subsequently folded into nanoscale structures with the aid of short 'staple' strands. Changing the sequences of these staples would lead to different sections of the scaffold being connected with each other, yielding a wide variety of structures that could be realised with the same scaffold and the same experimental method. The design of the structure-specific staples can be assisted by specialised software [28].

DNA origami was a particularly impressive achievement not only in the context of the previous studies on DNA nanostructures, but more generally for all of nanotechnology, as this method allows for easy, self-assembly-based realisation of almost arbitrary structures with molecular precision. The many examples, improvements and applications of DNA origami that have been published since Rothemund's study indicate its outstanding impact on the field. Two of these advances are depicted in Figure 1.4, together with examples of the original work.

Although the assembly of such structures is interesting in itself, a greater, and more relevant challenge is to use such structures to organise other functional materials and entities on the nanoscale. DNA origami structures are highly addressable, meaning in this context that spatial relation between different elements of the structure can be controlled with nanometric precision. Through chemical modification of selected staple strands, for example to include a biotin or a thiol, complex architectures of functional particles can be achieved. Particularly impressive studies towards such abilities include the work by Kuzyk and colleagues, in which

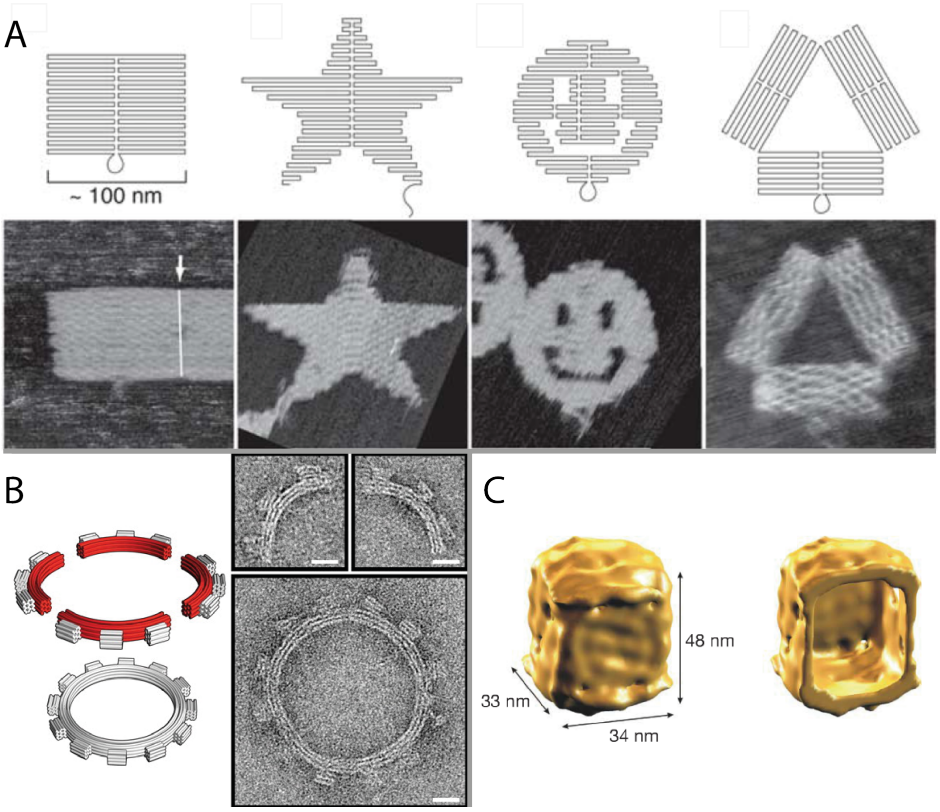

Figure 1.4: Examples of DNA origami. **A:** Two-dimensional structures adapted from Rothemund's original work. In the top row, the path of the folded scaffold strand is depicted schematically . The bottom row depicts AFM recordings (see subsection 1.3.3) of these structures; each micrograph's dimensions are equivalent to 165 nm × 165 nm. **B:** Three-dimensional origami structures mimicking a mechanical cog-wheel, imaged with EM, scale bars 20 nm. **C:** A three-dimensional, hollow nanocontainer imaged with cryo-EM. A adapted from [103], with permission from Macmillan Publishers Ltd, ©2006, B adapted from [23], with permission from AAAS, ©2009, C adapted from [4], with permission from Macmillan Publishers Ltd, ©2009.

DNA origami serves as a template for a plasmonic helical assembly of gold nanoparticles [63]. Nanoassemblies of fluorophores may serve as sophisticated cell-labels in biological studies [69]. And although one of the field's original goals, three-dimensional protein-DNA crystals, have not yet been demonstrated, recent advances in two-dimensional protein-DNA arrays [74, 114] and in the complexity, manufacture and structure of DNA origami [27, 140, 120] suggest that it remains a realistic endeavour. A plethora of further important structures has been created in the last decade, suggesting a broad potential for applications [70]. The focus of this introduction will shift to dynamic DNA systems at this point, however, as they are the central theme of this work.

1.1.3 *Dynamic DNA nanodevices*

In the last two decades, multiple two-state DNA nanodevices switching their structural conformation in response to environmental conditions and external triggers have been presented [61], including switches with sensing capacity for certain proteins [116], light of particular wavelengths [67], salt concentrations [76] or changes in pH [71, 83].

In the context of this work, however, the main interest are devices that rely on dynamic interactions between nucleic acids. Molecular beacons are an early example: these devices consist of a single strand of DNA which is partially complementary to itself, thus forming a double-stranded stem and a single-stranded loop, yielding a hairpin structure. Upon exposure to sequences complementary to the loop region, the hairpin opens. This switching mechanism is highly sensitive to sequence correctness, making these devices reliable nucleic-acid probes [134].

A somewhat more complex approach to conformational switching was described by Yurke and colleagues, who presented nanoscale DNA tweez-

ers which could be opened and closed through the addition of specific instruction strands of DNA [151]. This device, presented in Figure 1.5B, relies on the non-enzymatic displacement of one strand in a duplex by an energetically more favourable one. This toehold-mediated strand displacement (TMSD) is depicted in Figure 1.5A and briefly discussed in the caption, and is revisited in more detail in section 1.2 and throughout this thesis. This device contains a further important design element, utilising the differences in persistence length between single strands and duplexes to connect the relatively rigid arms with a flexible hinge. Due to its structural complexity and its capacity to be operated cyclically using oligonucleotide hybridisation alone, this device is commonly considered to be one of the original DNA nanomachines, which have since received great attention within DNA nanotechnology. Similar hybridisation switches have since been used, for example, to dynamically control the spatial conformation of structures as diverse as a DNA tetrahedron [40], a DNA origami nanocontainer [4] and entire two-dimensional, periodic DNA arrays [33].

DNA strand displacement can also be used to operate nanodevices with a far greater number of states. Shin and Pierce demonstrated this with a two-footed DNA walker, which, upon sequential addition of instruction strands, displayed linear motion along a DNA track (Figure 1.6A) [119]. Sherman and Seeman presented a similar device [117], both being reminiscent of two-footed, biological motor proteins, such as myosin. There is one particularly striking, systematic difference between the natural and synthetic machines, however: each step of these two DNA nanomachines has to be triggered externally, displaying a considerable lack of autonomy. In contrast, myosin, provided with sufficient adenosine triphosphate (ATP) and actin track, is capable of taking on the order of 10^6 directional steps before failing without any further outside instruction [47].

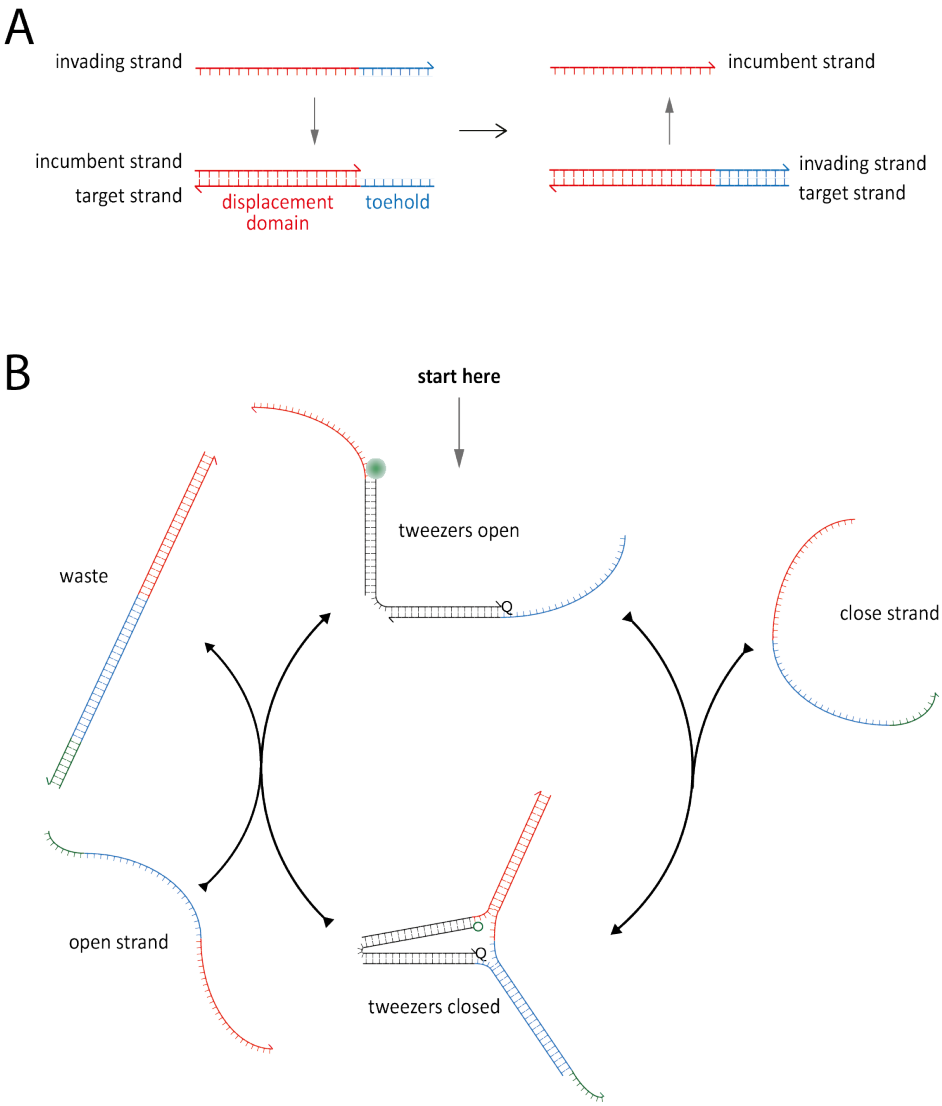

Figure 1.5: **A:** Schematic representation of a TMSD reaction. A target strand's displacement domain is hybridised with a complementary incumbent strand (red, bottom left). the target strand further has a single-stranded overhang, or 'toehold' (blue). An invading strand, which is not only complementary to the displacement domain, but also to the toehold, is then capable of displacing the incumbent from the target. The rate of this reaction is usually strongly dependent on toehold length. This process will be explored in further detail in section 1.2. **B:** The cyclic operation of the nanomechanical tweezers presented by Yurke and co-workers [151] relies on TMSD. Adding a strand which is complementary to single-stranded overhangs of the two arms of the tweezers forces them into a closed conformation. Even when bound to the tweezers, the close strand provides a toehold to a complement, or open strand, which removes the close strand from the tweezers via TMSD. This creates an inert waste duplex and returns the tweezers to their open conformation. This process can be observed using FRET (subsection 1.3.1) between a quencher (Q) and a fluorophore (green circle) attached to either arm.

This challenge was addressed soon after with several DNA machine designs being presented in which sustained, nanoscale linear motion can be achieved without requiring external triggers. In those initial designs for autonomous motors, DNA strand displacement remained of central importance. The energy required, however, was either supplied through enzymatic hydrolysis of nucleic acids [18, 7] (Figure 1.6B) or ligase-catalysed hydrolysis of ATP [149].

Somewhat more elegant motor designs use DNA itself as their energy source. Apart from representing a conceptually interesting feature, such designs do not require environmental conditions to be tailored to the enzymes used. They also provide greater design flexibility, as they do not need to be engineered around a predefined set of enzyme recognition sites. These mechanisms utilise single-stranded loops of DNA, which are formed either as a bulge emerging from a two-strand duplex, or from a domain in which the strand forming the loop binds onto itself, i.e. a hairpin as described above. Assuming sufficient stability of the double-stranded domains, spontaneous, full hybridisation of such a loop to a complementary, second one will be strongly impeded, due to topological constraints [13].

This reaction, which is connected with a considerable thermodynamic gain, can be catalysed with a specialised oligonucleotide which is capable of temporarily resolving one loop structure using TMSD, leading to a reduction of topological constraints, and thus rapid hybridisation of the two complementary loops [133, 44]. A DNA motor capable of catalysing such metastable complexes would thus not be required to rely on enzymes which are arguably much more complex than itself. This conceptual milestone was reached by Green and co-workers [43]. The intriguing design of this autonomous, two-footed walker did not only allow it to rely on DNA hybridisation alone as an energy source, it also left the DNA

track it was walking on intact (Figure 1.6C). This previously unachieved track reusability is also an important feature of biological motor proteins.

With the advent of DNA origami, it also became possible to provide these types of nanomachines with more robust tracks, allowing for more complex tasks and more sophisticated device observation. Gu and colleagues showed that an origami tile can serve as the basis for a molecular assembly line, demonstrating a DNA nanomotor which could sequentially load and transport three gold nanoparticles [46]. Lund and co-workers showed that a similar system does not require linear tracks, but can be controlled to take turns [72]. For the motor system introduced by Green and colleagues [43], Wickham and colleagues showed that DNA origami can also serve as template for a network of tracks, and that the motor can navigate through this network, following external commands [142].

1.1.4 *DNA computation and reaction networks*

DNA sequences can store vast amounts of information, both genetic and man-made [39], using a four-letter alphabet. They can also be made to interact dynamically in a highly controllable fashion, as several of the studies cited above demonstrate. This combination of data storage and controllable interaction capacities suggests that DNA can be used to process and compute information outside of its biological context. Adleman demonstrated one of the first applications of this ability, showing that DNA oligonucleotides, in conjunction with enzymes and sophisticated purification techniques, can be used for a molecular algorithm, capable of solving the Hamiltonian path problem for directed graphs, an important representative of a particular class of problems in computational complexity theory [1]. Further impressive proofs of principle regarding the vi-

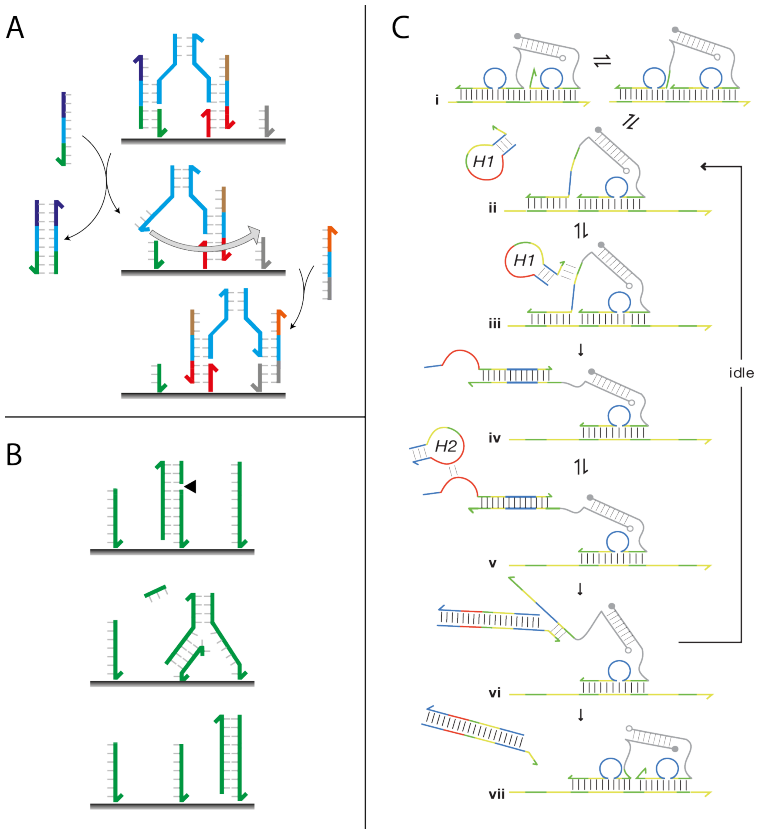

Figure 1.6: Three examples of linear DNA motors, moving from left to right. **A:** A system proposed by Shin and Pierce [119]. A two-footed motor (light blue) is bound to two stators on the surface (red and green) via two connector strands. Adding the complement to the left connector frees the left foot, which can then diffuse forwards and rebind to the surface in the presence of a connector attaching it to the grey stator. The connectors have to be added separately for successful operation. **B:** An autonomous, directional motor by Bath and co-workers [7]. When the single stranded motor (3′ end pointing towards the top of the figure) binds to a stator immobilised on the surface, it creates a recognition site for a nicking enzyme, which subsequently cleaves the stator. The cut stator disintegrates, revealing a toehold for the next, still fully intact downstream stator. **C:** An autonomous, DNA-fuelled motor design by Green and co-workers [43]. Here, the track consists of a single strand of DNA with a periodic sequence (bottom of each panel). The motor (grey) binds to the track with two feet, each of which form a loop when bound to the track. Bound feet compete with each other for one domain on the track, leading to occasional opening of the left foot loop. In its opened state, this loop can hybridise hairpin H1, leading to detachment of the left foot from the track. The now opened hairpin H1 is in turn capable of hybridising hairpin H2, removing H1 from the foot and thus allowing the foot to rebind to the track, either at its original position, or a forward position. A and B were adapted from [8], with permission from Macmillan Publishers Ltd, ©2007, and modified; C was adapted from [43], with permission from the American Physical Society, ©2008.

ability of DNA for computation were presented subsequently, all of which rely on DNA hydrolysis, either through biological enzymes [92, 10, 75] or deoxyribozymes [124]. As with some of the dynamic devices introduced above, the use of enzymes and deoxyribozymes is commonly associated with a number of limitations, as these molecules require very particular conditions to operate as intended, and can lead to design constraints which limit system complexity.

Seelig and co-workers presented an impressive alternative, demonstrating that cascaded, toehold-mediated DNA strand displacement reactions alone can serve as the basis for logic circuits [111]. A few 'DNA-only' reaction systems had been presented before, for example as amplification systems [25, 112]. However, Seelig and co-workers were arguably the first to construct multilayered circuits, in which individual DNA devices, acting as logic gates, controllably interacted with each other to produce output according to a predefined Boolean formula. With a digital, rather than analogue implementation, signal restoration mechanisms could be included in the circuit, thus reducing output ambiguity. The same group later showed that similar DNA devices can be used to implement arbitrary reaction pathways [121]. This theoretical study thus provided a general 'translation' from chemical reaction equations to DNA chemical reaction networks (CRNs). Although experimental support for the more sophisticated proposals of that study has yet to be produced, it still provided an important evaluation of the vast potential of dynamic DNA reactions.

The current pinnacle of experimental implementation of DNA logic circuits is undoubtedly found in the recent work of Qian and colleagues. After devising a versatile DNA strand displacement logic gate motif [98] (Figure 1.7), they demonstrated that complex reaction systems based on this motif are capable of calculating square roots digitally [97] and representing neural networks [99].

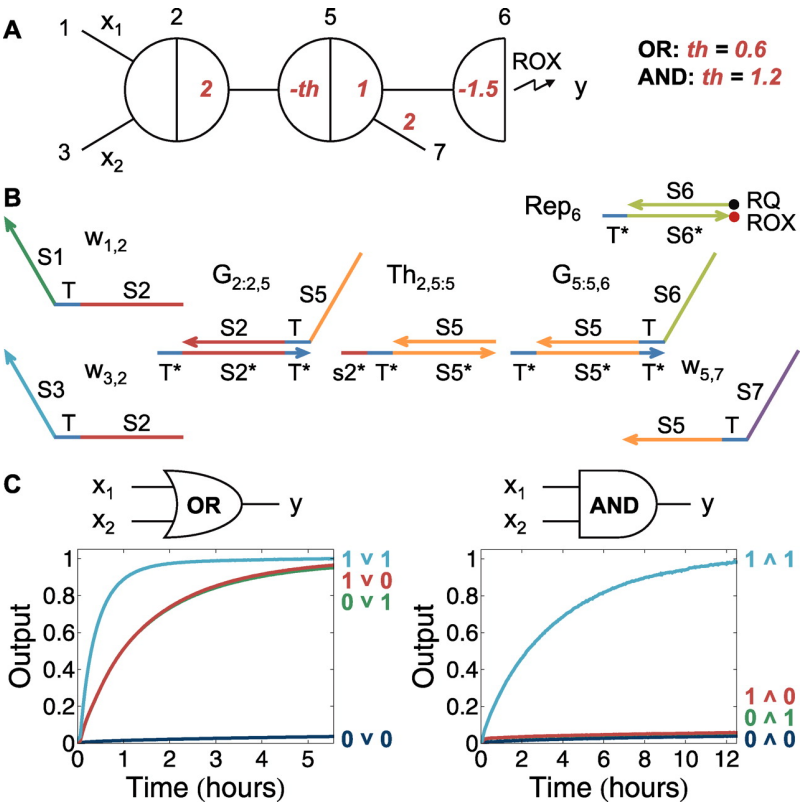

Figure 1.7: Seesaw gate motif by Qian and Winfree. **A:** Schematic depiction of a gate that can act according either to an AND or OR logic to inputs x_1 and x_2, producing output y. **B:** Domain depiction of such a gate. The inputs are realised with strands $w_{1,2}$ and $w_{3,2}$, which can both interact with duplex $G_{2:2,5}$ to displace the incumbent. This incumbent in turn can interact with threshold duplex $Th_{2,5:5}$ and duplex $G_{5:5,6}$. Depending on the initial concentration of threshold duplex relative to initial input concentrations, this system acts either as an AND or an OR gate. **C:** Output is reported using bulk spectrofluorimetry (subsection 1.3.1). Adapted from [98] with permission from AAAS, ©2011.

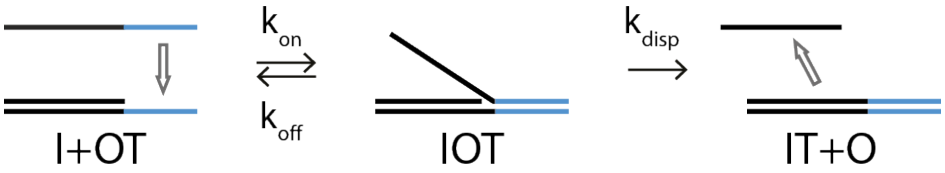

Figure 1.8: A three-state model of TMSD. The invader I can initiate a toehold on the target T, which is in a duplex with incumbent O, forming the intermediate complex IOT. Toehold binding is characterised by a rate constant k_{on}, toehold dissociation by k_{off}. Following toehold initiation, the invading strand can displace the incumbent from the intermediate complex; this process is characterised by k_{disp}.

1.2 CONTROL AND VARIATIONS OF TMSD

1.2.1 *A basic view on TMSD*

The previous section demonstrated the importance of TMSD for a wide variety of synthetic DNA devices and nucleic acid CRNs. Here, a more detailed introduction is given. Yurke and Mills described quantitatively how the rate of DNA strand displacement reactions can be increased significantly if the invading strand is provided with an exclusive toehold. For each nucleotide added to the length of the toehold, they found the reaction rate to increase by approximately one order of magnitude [150]. The role and influence of the toehold was studied in greater detail by Zhang and Winfree, who reported results similar to those of Yurke and Mills and also found that the increase in rate that can be achieved with a longer toehold is limited: lengthening a toehold beyond ~6 nt will generally not increase reaction rate [156].

The main difference between a longer toehold and a shorter one is, of course, that the former provides an energetically more favourable binding site. Accordingly, the reaction rate can not only be controlled via toehold length, but also through toehold composition, i.e. the relative content of G-C and A-T base pair sites [156].

A simple, three-state model of TMSD helps to understand these experimental observations qualitatively (Figure 1.8)[1]. In this picture, the reaction can be described as

$$I + OT \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} IOT \overset{k_{disp}}{\rightarrow} IT + O.$$

Using a quasi-steady-state assumption [156], i.e.

$$\frac{d[IOT]}{dt} = k_{on}[I][OT] - k_{off}[IOT] - k_{disp}[IOT] = 0,$$

the concentration of the intermediate can be described as

$$[IOT] = \frac{k_{on}}{k_{off} + k_{disp}}[I][OT].$$

Thus,

$$\frac{d[IT]}{dt} = \frac{d[O]}{dt} = k_{disp}[IOT] = \frac{k_{disp} \cdot k_{on}}{k_{off} + k_{disp}}[I][OT],$$

which yields an effective rate constant of

$$k = \frac{k_{disp} \cdot k_{on}}{k_{off} + k_{disp}}, \tag{1.1}$$

corresponding to the simple second-order reaction[2]

$$I + OT \overset{k}{\rightarrow} IT + O.$$

If a specific family of displacement systems is described, in which only the length l (in nucleotides), but not the sequence of the toehold is varied, k_{on} can be assumed to be approximately constant [156]. As k_{disp} is not related to the toehold, it can also assumed to be constant for a fixed

1 This model is heavily influenced by the works of Yurke and Mills [150] and Zhang and Winfree [156], who made similar descriptions, as well as discussions with Thomas E. Ouldridge (Department of Physics, University of Oxford).

2 Limitations to these simplifications are encountered and discussed in chapter 2.

displacement domain. Thus, from the equilibrium between the initial and the intermediate state follows

$$k_{off} \propto e^{\Delta G^\circ(l)/RT},$$

with the length-dependent toehold binding energy $\Delta G^\circ(l) < 0$. This binding energy is roughly proportional to $-l$, according to standard nearest-neighbour modelling of DNA [106]. If

$$k_{off} \gg k_{disp},$$

equation 1.1 yields

$$k \propto e^{l},$$

the exponential dependence of the effective TMSD reaction rate on toehold length. Once the toehold is so long that the other extreme

$$k_{off} \ll k_{disp}$$

is reached, however,

$$k \propto k_{on} \approx const.,$$

explaining why for long toeholds, a rate saturation is observed.

1.2.2 *Improving control over TMSD*

Due to the prominent role of this mechanism in the field, several attempts have been made to improve control over its kinetics. As the previous subsection showed, a direct approach is to increase the binding free energy of the toehold. This can easily be achieved by increasing the number of base pairs formed in the toehold domain, or by increasing the

toehold domain's relative G-C content [66, 156, 150]. In the basic model of TMSD presented above, the decrease of k_{off} achieved with a stronger toehold increases the overall rate of displacement, as long as toehold dissociation is the overall rate-limiting step. Strengthening the toehold beyond this threshold does not increase the overall rate of displacement.

Variations of the original TMSD scheme, such as those depicted in Figure 1.9, can also lead to increased control over the displacement process. Zhang and co-workers introduced toehold exchange, in which the incumbent is also provided with an exclusive binding domain on the target, opposite to the invader toehold [155, 156]. This seemingly simple extension to TMSD has had a considerable impact on dynamic DNA nanodevices and nucleic acid CRNs, allowing, for example, for the previously mentioned computation networks by Qian and colleagues [99, 97, 98].

Zhang and co-workers also showed that TMSD is capable of displacing multiple incumbents from the same target simultaneously [155], and that two invading strands can cooperatively displace a long incumbent strand, initiating displacement from two opposite toeholds on the target strand [153]. Chen introduced associative toehold activation, a displacement mechanism in which invading strands do not carry their own toehold-binding domains, but must first hybridise a helper strand containing such a domain to efficiently displace an incumbent [17]. A variation of this mechanism was used by Genot and colleagues to demonstrate a DNA reaction network capable of matrix multiplication [37], which features in chapter 3 of this thesis, where it is discussed in more detail. Genot et al. have also shown that the introduction of an inert spacer between toehold and displacement domain, made of either DNA or polyethylene glycol (PEG), allows increased control of TMSD reactions [38]. The result of increasing the distance between toehold and displacement domain is

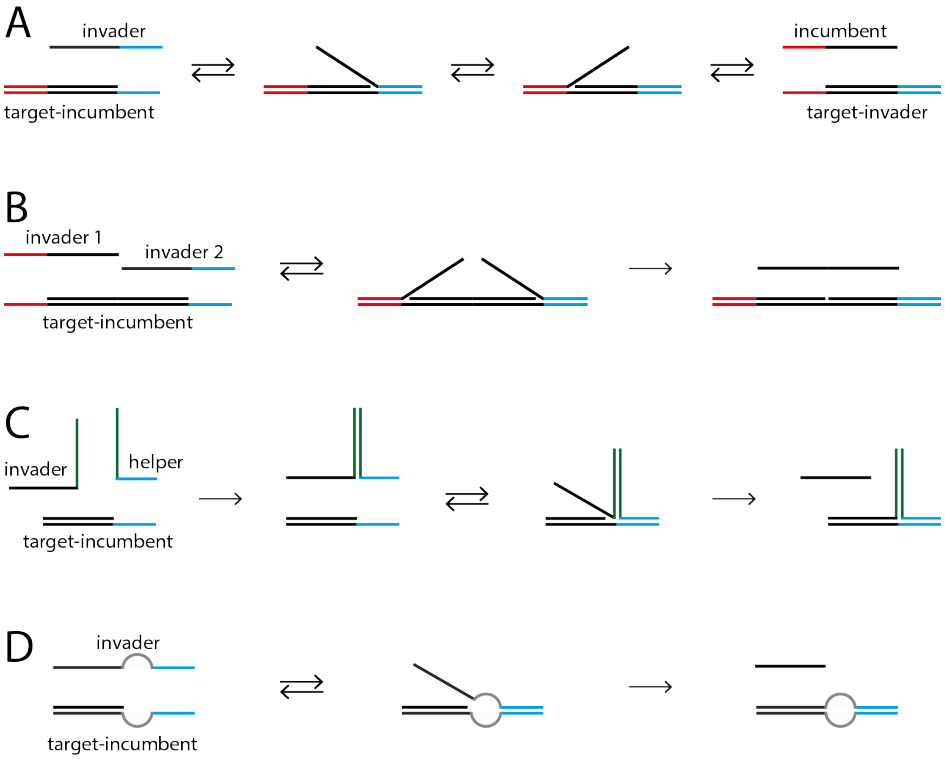

Figure 1.9: Variations and extensions to basic TMSD. **A:** Toehold exchange as introduced by Zhang and co-workers [155, 156]. The invader can bind to a toehold on the target, and initiate a branch migration through the displacement domain. The incumbent, however, has its own exclusive binding domain on the target, making complete displacement of the incumbent less likely, and allowing displaced incumbent to rebind, and occasionally displace the invader. Invader and incumbent toeholds can be chosen to satisfy a wide variety of kinetic and thermodynamic conditions. **B:** Zhang's cooperative displacement [153]. Either invader is too short to displace the incumbent. Only if both are present can displacement occur. **C:** Chen's associative toehold activation [17]. The invader is complementary to the displacement domain, but, lacking a toehold, displaces incumbent only very inefficiently. By introducing a helper strand, which can hybridise with a linking domain on the invader, a two-strand complex is created which can displace the incumbent at high rates. **D:** Remote toehold-mediated strand displacement developed by Genot and colleagues [38]. The toehold and the displacement domain are separated by an inert spacer (grey). The rate of displacement can be controlled through the length of the spacer, with longer spacers providing a greater kinetic impediment than shorter ones.

that the invader requires more time to initiate branch migration, which, in the above model of TMSD, translates into an increase in k_{disp}.

1.3 OBSERVATION OF DNA NANODEVICES

This thesis is not only concerned with the control of DNA nanodevices, but also with the measurement of their operation. For this reason the three most commonly used observation techniques are briefly reviewed here, with a focus on their advantages and disadvantages in the context of DNA nanodevices, instead of describing technical details, for which abundant literature is available elsewhere. This review should elucidate the need for novel observation methodologies to advance the field. An attempt at developing such a technique is described later in this work (chapter 4).

1.3.1 *FRET-based bulk fluorimetry*

One of the most convenient methods to measure time-dependent behaviour of a DNA nanodevice is the use of FRET between fluorophores and quenchers attached to device components. When fluorophore and quencher are in close proximity to each other, the excited fluorophore (donor) transfers energy to the quencher (acceptor) non-radiatively and returns to its ground state without emitting a photon. The quencher can be a second fluorophore, which, as a consequence of the energy transfer, can emit a photon of its own characteristic emission spectrum. Dark quenchers give off energy received from a donor fluorophore as heat. The lack of an emission spectrum makes dark quenchers particularly advantageous in measurements where multiple processes are reported through FRET, and photon emission of an acceptor fluorophore may overlap with

the absorption or emission spectra of the other fluorescent labels in the sample.

The broad applicability of such fluorophore-quencher systems in observing DNA nanodevices is mainly due to the strong distance dependence of FRET on nanometer length scales. FRET requires the donor-acceptor pair to be in close spatial proximity to occur. The fraction of energy transferred from donor to acceptor is $\varepsilon = \frac{1}{1+(r/R_0)^6}$ [125], where r is the donor-acceptor distance and R_0 is a characteristic distance at which $1/2$ of the energy is transferred. Values for R_0 are commonly between 2 nm and 7 nm [146], making FRET an accurate ruler on such length scales. The required labels are readily attached to DNA; this is usually most economically achieved by ordering synthesised oligonucleotides directly with the appropriate modification and subsequent high-performance liquid chromatography (HPLC) purification from an established commercial manufacturer.

Thus, such labels can be integrated into the designs for a DNA nanodevice in a way that different states of the device are associated with different donor-acceptor distances. FRET-induced changes in fluorescence are most conveniently measured with a spectrofluorimeter in bulk, i.e. in $\sim 0.1 - 2$ ml volumes of nano- to micromolar concentrations of sample.

In bulk, the total fluorescence measured represents a sum over a huge number of individual fluorescence signals. From this information, key parameters of device operation can be derived. Apart from the fluorescent labelling, little effort is required to prepare or conduct such experiments. They can be performed in a wide variety of buffers and temperatures with high temporal resolution. Further reactants can be added in controllable quantities during measurement.

With the simplicity of this experimental method also come several restrictions. Due to the instantaneous observation of an entire ensemble,

insight into the operational details and the variety of states of a complex DNA nanodevice is limited. Furthermore, fluorescent labels themselves can be problematic, as they stabilise a DNA duplex thermodynamically similar to an additional base pair, both through interaction with DNA itself [84] and with each other [77]. Both types of interaction may lead to deviations from the intended device behaviour. For nanodevices, this is usually neglected, but for measurements which aim to accurately characterise fundamental DNA reactions, a separate, labelled reporter complex can be introduced, allowing for label-free reactants [156]. Despite these restrictions, a vast majority of previously published DNA nanomachines and reaction networks have been characterised using this technique.

1.3.2 *Single-molecule fluorescence microscopy*

Single-molecule fluorescence microscopy comes in many variations, some of which have yielded tremendous insights into a broad spectrum of previously unexplored biological processes at the nano- and microscale. All these variations share one common goal: By restricting observation to individual molecules, sample heterogeneities and operational details become discernible which would be lost in a bulk measurement. Here, discussion of these different methodologies is limited to those variations previously applied, or directly related to the study of DNA nanodevices.

A study by Steinhauer and colleagues [123] is a sensible starting point for this discussion, as this work demonstrated that complex DNA nanostructures can serve as a calibration standard for super-resolution fluorescence microscopy. As the authors of this study point out, DNA origami structures, in which certain staple strands are fluorescently labelled, allow for easy and cheap mass-production of highly controllable assemblies of fluorophores on a sub-diffraction-limit length scale. Through di-

lution and binding to a surface, these assemblies become individually distinguishable under a microscope. They can then be used to test and demonstrate super-resolution and single-molecule capabilities of experimental set-ups. This study also served as a proof-of-principle that fluorescence microscopy can be applied to analyse DNA nanodevices, though only static origami tiles were observed.

A further study by the same research group soon followed, introducing the DNA points accumulation for imaging in nanoscale topography (DNA-PAINT) method [58], which has since allowed for some of the most remarkable single-molecule fluorescence microscopy measurements of DNA nanostructures to date. In this approach, short fluorescently labelled oligonucleotides bind transiently to complementary docking strands protruding from a surface-immobilised structure at predefined locations. Over time, all docking locations are visited by the label strands in solution, which allows for super-resolution reconstruction of all docking locations in sufficiently long recordings. The basis of DNA-PAINT resembles the fundamental principles of more popular microscopy techniques such as stochastic optical reconstruction microscopy (STORM) [104] and photoactivated localisation microscopy (PALM) [11], in which sub-diffraction-limit resolutions are achieved by isolating individual emitters temporally through controlled activation and deactivation of their fluorescence. Figure 1.10 presents a schematic representation of this principle.

DNA-PAINT reduces the limitations on experimental duration set by photobleaching. By using total internal reflection fluorescence (TIRF) microscopy, an illumination technique in which the excitation laser is aimed at the sample at such an angle that total internal reflection occurs between the sample carrier and the sample itself, only those fluorophores close to the surface are excited [6]. Under these conditions, the vast reservoir of fluorophores in the non-illuminated bulk of the sample guaran-

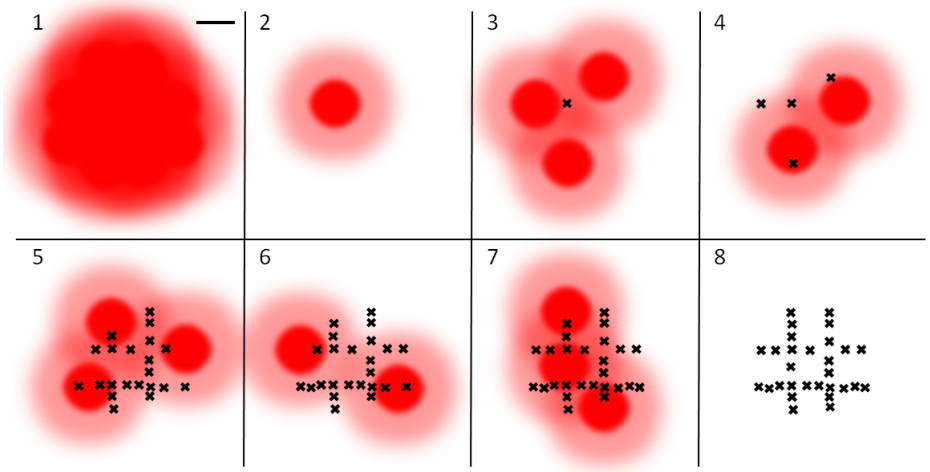

Figure 1.10: Super-resolution microscopy through temporally separated emission. **1:** The spatial relation between multiple, simultaneously emitting fluorophores of the same colour can not be determined if the distances between them are below the diffraction limit of ~250 nm (black scale bar). **2:** If emission can be controlled to be temporary, so that only a single, or a sparse subset of fluorophores (red) emit at certain times, they can be localised individually with nanometric precision. Frames recorded in which fluorophores in sub-diffraction-limit proximity emit simultaneously are usually excluded from data analysis. **3 & 4:** With sufficient time allowed for observation, further fluorophores will transiently emit, allowing for sequential localisation and registration of their respective positions (black crosses). **5 to 8:** This procedure continues, until the sub-diffraction-limit structure of the fluorophore set emerges. The cause of the transient emission behaviour – controlled photoswitching, transient binding to an illuminated surface, inherent fluorescence intermittency etc. – is irrelevant for this general principle.

tees a continuous supply of unbleached label strands. Since its inception, this method has, for example, been shown to be feasible for the imaging of fluorescent DNA origami 'barcodes', which can be used to label cells [69]. Its abilities have recently also been extended to impressively accurate measurements of three-dimensional structures [52].

Though shown to be versatile, so far DNA-PAINT appears to be restricted to static DNA nanotechnology. As it relies on transient, rather than constant fluorescent labelling, it is likely that motion of, say, a molecular motor such as those introduced in subsection 1.1.3 could not be imaged continuously. A recent publication on the state of this technique by Jungmann and co-workers further suggests that very long observation times (~50 min) are required to achieve measurements with nanometric precision on a static structure [57], which is below the temporal resolution required for observing most dynamic DNA nanodevices.

In order to observe with high resolution the operation of a DNA nanomachine, different methods need to be applied. Masoud and colleagues demonstrated microscopy-based observation of a similar motor using single-molecule FRET [78]. Although certain insights into the operational details of the system could be gained, the FRET-based approach prohibited extended tracking of the dynamic device studied, as it is very useful to resolve distances below ~10 nm, but does not allow for the observation of processes on a greater length scale, and is thus restricted to devices operating at a short range.

A different technique, single-molecule high-resolution colocalisation (SHREC), avoids the drawbacks of single-molecule FRET (limited to very small changes in device conformation) and DNA-PAINT (so far only shown to be reliable on static structures). This technique had previously been applied with great success to the study of biological motor proteins [19]. It is based on the fluorescence imaging with one-nanometre accuracy

(FIONA) method pioneered by Paul Selvin and his group [148]. In FIONA, single fluorophore point spread functions (PSFs) are fitted with two-dimensional Gaussians, which allows for localisation of the fluorophore with single-nanometre precision. For this method to work reliably, however, the fluorescent labels need to be sparsely distributed, and sample drift can not directly be compensated for, as no reference label exists. In SHREC, these issues are addressed by labelling the samples with two differently coloured fluorophores, separating their respective emission into two channels using wavelength-specific optical filters, determining their respective positions as in FIONA, and then colocalising the two positions using a previously established mapping function between the two channels. From this information, the spatial relation between the two labels over time can be investigated, making this technique particularly interesting for tracking of dynamic processes, both biological and synthetic, on the sub-micron scale.

Lund and co-workers used this approach to track the movement of a deoxyribozyme-based nanorobot on a ~120 nm DNA origami track [72]. Cyanine 3 dye (Cy3) and Cyanine 5 dye (Cy5) were used as labels of motor and track, respectively. These dyes are convenient labels of oligonucleotides (and were used as such in chapters 2 and 3), and as had been demonstrated in the original SHREC study, they work well in the context of biological motor proteins. The speed of biological molecular motors is several orders of magnitude above that of those synthetic systems based on nucleic acids. As a consequence, only weak illumination of such systems is possible without photobleaching the labels before the device has made a significant advance. The resulting reduced emission is, in turn, detrimental to localisation precision, which decreases with lower signal-to-noise ratio (SNR) [130]. This may have been the reason for the sub-optimal localisation precision in the study by Lund et al., which, as a

consequence, did not allow for resolution of mechanistic details of device operation.

1.3.3 *AFM*

AFM, which allows a surface to be probed with nanometric resolution [12], is an established observation method for nanostructures, and has been applied with great success to the study of static DNA nanotechnology (e.g. DNA origami as in Figure 1.4A). With recent advances in this technique, it has also been shown to be applicable to the study of dynamic devices. The work of the Sugiyama lab and their collaborators is of particular note in this context, made possible through the development of AFM with exceptionally high temporal resolution [5]. Among the devices studied are an opening, three-dimensional DNA cube [31], a rotary device [100] and the motor system illustrated in Figure 1.6B, operating on either a linear [143] or branched [142] DNA origami track. On the linear track, individual, ~6 nm motor steps could be resolved.

Thus, AFM has in principle been shown to be capable of resolving the operational details of dynamic DNA nanodevice operation. In addition to the highly specialised AFM device required for this, however, this method comes with a host of further drawbacks. Among them are the requirements for experimental conditions. The study by Wickham and colleagues [143], for example, contains only little data representing tracking measurements of an individual motor, because the enzyme driving motor operation has to be washed out prior to observation, so that it does not obscure the recorded images. Thus, only those few motors to which an enzyme is attached can be tracked, and only for a few steps before enzyme detachment. The other devices studied successfully are comparatively limited in their abilities, switching conformations once

rather than displaying cyclic operation [31, 100]. Other studies using less sophisticated AFM approaches were either only able to track dynamic devices with very limited temporal resolution [72] or resorted to taking snapshots of different devices of the same type at various stages of device operation [46]. In the former study, it was further assumed that device operation is hindered by the conditions required for AFM.

Even if high temporal resolution is technically feasible, a further disadvantage of this observation method emerges: Due to the interaction of the probing tip and the sample, AFM damages DNA nanostructures [103], an effect that is exacerbated by the repeated scans necessary for the continuous observation of a dynamic nanodevice [141].

1.4 AIMS AND SCOPE OF THIS THESIS

The aim of this work is the development of new paradigms to control and observe DNA nanodevices and reaction networks. My contributions towards this goal are structured as follows: Chapter 2 introduces a novel mechanism for kinetic control of strand displacement reactions. This is achieved through the deliberate integration of mismatched base pairs. This allows for reaction rates to be tuned over more than three orders of magnitude. It is also shown that this approach allows for independent control of strand displacement kinetics and thermodynamics, meaning that reaction rate constants are variable without adjusting the change in free energy driving the reaction. The experimental observations are explored with simulations using a coarse-grained model of DNA. A simple device is presented, in which the competition between two invading strands is influenced by a single mismatch. The findings of this chapter suggest a considerable impact on design capabilities in dynamic DNA nanotechnology and on nucleic acid probes.

Chapter 3 advances a variation of regular toehold-mediated strand displacement. Instead of displacement by a single strand, a previously introduced mechanism in which multiple strands can specifically cooperate to achieve displacement is improved and generalised. This mechanism is shown to allow for DNA reaction network-based calculation of Boolean matrix products. This ability is applied to calculations on directed graphs.

Chapter 4 is different from the previous two chapters in two ways: Firstly, it introduces a technique to observe, rather than control, DNA nanodevices. This technique is based on single-molecule fluorescence microscopy, utilising quantum dots (a type of semiconductor nanocrystal), rather than conventional fluorescent labels. It aims to provide a single-molecule observation methodology tailored to the length- and time scales of DNA nanodevice operation, features which no single of the observation techniques mentioned above displays. Secondly, it presents preliminary results, taken from an ongoing project. Despite the preliminary nature of this effort, some of the proof-of-principle results are very promising. It is shown that due to their remarkable fluorescent properties, quantum dots allow for highly precise optical microscopy on DNA nanostructures. Two subtly different methods for this are presented, one relying on nanometric colocalisation of differently coloured quantum dots, the other utilising the blinking behaviour characteristic for quantum dots to achieve optical resolution below the diffraction limit.

1.5 CONTRIBUTIONS TO THIS THESIS

Much of the work presented in this thesis was part of collaborative efforts. Accordingly, each of the experimental chapters is presented with a paragraph on contributions from colleagues and collaborators. Apart

from the exceptions stated in those paragraphs, or where findings and contributions of others are directly referenced, I, as author of this thesis, am responsible for all content presented. This thesis was written using a LaTeX template based on 'classicthesis', which is kindly provided to the public by Andre Miede (Saarland University of Applied Sciences).

2

CONTROLLING STRAND DISPLACEMENT WITH MISMATCHES

2.1 INTRODUCTION

This chapter describes work on how control over DNA strand displacement processes can be improved by introducing free-energy barriers at different locations within the displacement domain. These barriers are implemented by replacing a Watson-Crick base pair with a base mismatch, which are known to be able to impede strand displacement [93]. It is shown that the kinetics of the displacement process are highly sensitive to changing the location of such a barrier, which does not affect the change in free energy associated with the reaction. This effect is studied experimentally using different displacement systems and reactant concentrations, suggesting a broad potential for application (section 2.2). The process is explored further in section 2.3 by using a state-of-the-art coarse-grained model of DNA. In combination with the experimental findings, simulations based on this model allow for a detailed qualitative discussion of the observed behaviour. Further experimental proof for some of the conclusions derived from this discussion is provided in section 2.4.

In section 2.5, a system is presented that showcases how the impact of mismatch position on reaction rate can be used to influence a competitive strand displacement process in order to display complex kinetic behaviour despite a very small number of individual reactants. The sys-

tem further shows how the introduction of a mismatch allows for control over competing reaction pathways. In section 2.6, it is shown that a single mismatch between an invader and a target strand can improve the addressability and specificity of toeholds in TMSD reactions. This ability could be integrated into CRNs that are highly sensitive to spurious reactions, such as DNA amplification circuits. Lastly, section 2.7 describes the findings of this chapter in the context of displacement-based nucleic acid probes.

Contributions by colleagues and collaborators

Most of the results presented in this chapter have been accepted for publication in *Nature Communications* [73]. I am the lead author of this study. The manuscript was written by myself, Thomas E. Ouldridge and Andrew J. Turberfield, with suggestions from Natalie E. C. Haley and Jonathan Bath (Department of Physics, University of Oxford). In addition to work directly referenced, the following contributions to this chapter have been made by colleagues and collaborators:

- the original idea to study how variations in mismatch positions can influence the kinetics of a TMSD process was mine, but was influenced considerably through discussions with Thomas E. Ouldridge, who developed similar ideas around the same time;
- all authors of [73] were involved in the planning of the experiments;
- oligonucleotide sequence design in subsection 2.2.2 was performed by Natalie E. C. Haley using the design rationale I developed and which is presented there;
- some preliminary experiments were performed by Natalie E. C. Haley, under my direct supervision;
- oxDNA [89, 90] data was provided by Thomas E. Ouldridge;

- the fitting routine applied in subsection 2.2.2 was extended to the routine presented in subsection 2.5.2 with the help of Thomas E. Ouldridge;
- I, Thomas E. Ouldridge and Andrew J. Turberfield interpreted the experimental and simulation results for the aforementioned publication. This discussion, which was also influenced by the suggestions of Natalie E. C. Haley and Jonathan Bath, is included in more detail in this chapter;
- I prepared all figures, some with suggestions from the aforementioned co-authors of [73]. Figure 2.4 contains graphic elements kindly contributed by Jonathan Bath. The simulation snapshots in Figure 2.9A&B were provided by Thomas E. Ouldridge.

2.2 INTRODUCING MISMATCHES TO TMSD

2.2.1 *Direct labelling of reactants*

Experimental implementation

A first series of experiments to test the impact of mismatches in different locations on the kinetics of a TMSD reaction was conducted with a fluorescent Cy5 label on the incumbent strand and a dark quencher on the target strand (depicted in Figure 2.1). In this design, the Cy5 fluorescence is suppressed as long as the incumbent is in a duplex with the target strand. Once it is displaced by an invading strand, the fluorescence increases. This change in fluorescence was measured in bulk, i.e. in volumes of either 0.15 ml or 1.5 ml containing on the order of $1 - 10$ pMole of each reactant. The detailed description of the experimental methods can be found in Appendix A.1.3.

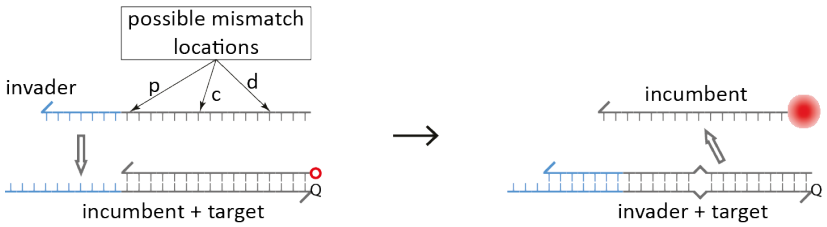

Figure 2.1: A thermodynamically favoured invading strand displaces an incumbent from a target strand via TMSD. Invaders are designed to form one $\overset{\cdots TXA \cdots}{\cdots ACT \cdots}$ mismatch at either the proximal (p), central (c) or distal (d) site, or form no mismatch at all. The reaction can be observed via the interaction of the dark quencher attached to the target (Q) and the Cy5 fluorophore attached to the incumbent (red). Upon displacement, the dark quencher no longer suppresses fluorescence from the Cy5 label, resulting in a measurable increase in emitted photons. In the exemplary reaction depicted here, a 7 nt-toehold invader forms a mismatch at the central site.

With the implementation presented here, three different locations for a $\overset{\cdots TXA \cdots}{\cdots ACT \cdots}$ mismatch, where X represents any of the non-complementary bases C, T or A, and toehold lengths between 0 nt and 10 nt can be tested with the same incumbent-target duplex, whilst only modifying the sequence and length of the invading strand. Within the nearest-neighbour model of DNA thermodynamics, all three possible single-mismatch forming invaders with one specific toehold length form duplexes with target which are thermodynamically equivalent. This was tested with NUPACK [152], which allows for the analysis and sequence design of oligonucleotides based on nearest-neighbour modelling of nucleic acids [107, 95]. The mismatch locations are named according to their relative distance from the toehold. The proximal mismatch is formed directly adjacent to the toehold, the central mismatch is formed 7 nt from the toehold and the distal mismatch is formed 13 nt from the toehold; the displacement domain is 16 nt long.

Most invaders tested here formed a $\overset{\cdots TCA \cdots}{\cdots ACT \cdots}$ mismatch at either of the three possible positions. This type of mismatch was tested because it is one of the energetically most unfavourable ones, destabilising duplexes

by approximately 6 kcal/mol relative to a fully complementary one [95]. It was thus expected to show the strongest effect on the displacement process. Mismatched invaders had toehold lengths ranging from 6 nt to 10 nt. For toeholds shorter than 6 nt, the duplex of target with invading strand including a $\overset{\cdots TCA\cdots}{\cdots ACT\cdots}$ mismatch is not thermodynamically favourable enough for the reverse reaction to be negligible [152].

Invaders with a thermodynamically more favourable $\overset{\cdots TTA\cdots}{\cdots ACT\cdots}$ in the central position were also tested with toehold lengths ranging from 6 nt to 10 nt. For comparison, equivalent measurements were also performed with non-mismatched invaders with toehold lengths between 4 nt and 10 nt.

Rate constant determination

Under the assumption that the reaction $I + TO \xrightarrow{k} IT + O$ (where I: invader, T: target, O: incumbent, XY: duplex consisting of strands X and Y and k: reaction rate constant) is irreversible and can be described with second-order kinetics, the fluorescence signal recorded against time can be fitted using an analytic solution to the differential equation describing the time-dependent increase in concentration of O:

$$\frac{d\left[O\right]}{dt} = k\left[I\right]\left[TO\right].$$

The analytic solution for equal initial concentrations of I and TO (i.e. $[I]_0 = [TO]_0$) is

$$\left[O\right](t) = [I]_0 - \frac{[I]_0}{1 + [I]_0 kt}. \tag{2.1}$$

The second-order reaction rate constant k as well as a scaling factor, which relates fluorescent signal in arbitrary units to $[O](t)$ in Moles per litre, were used as fitting parameters. The results of these experiments and the subsequent fitting are shown in Figure 2.2.

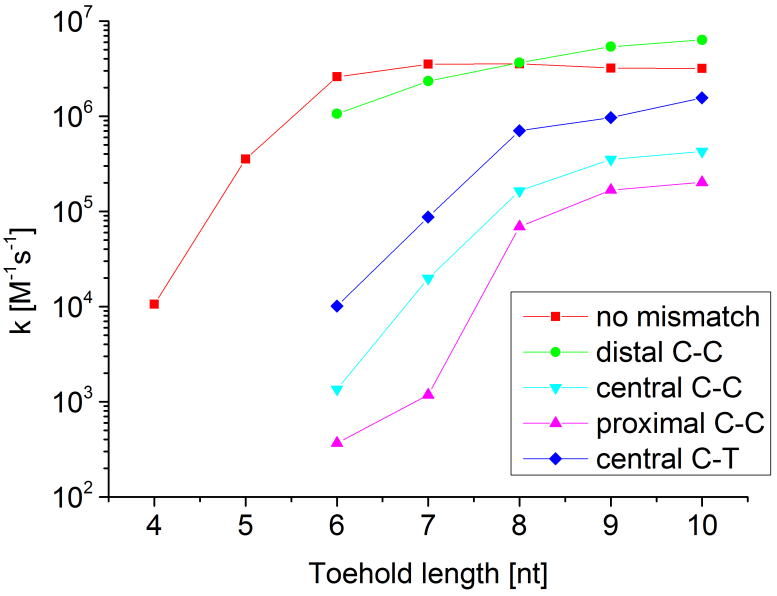

Figure 2.2: TMSD measurements using labelled incumbent and target strands. k values for all 27 invading strands tested with the same incumbent-target duplex. Here, C-C is shorthand for a $\overset{\cdots TCA \cdots}{\cdots ACT \cdots}$ mismatch, C-T represents a $\overset{\cdots TTA \cdots}{\cdots ACT \cdots}$ mismatch.

The rate constants determined for non-mismatched invaders are consistent with previous findings [150, 156]. As expected, such invaders with toehold lengths $\gtrsim 6$ were all measured to have very similar rate constants, as their high toehold binding energy suppresses k_{off}, making the toehold-independent k_{disp} the rate-limiting reaction step. For shorter, non-mismatched invading strands, the effective rate constant is reduced drastically.

For those invading strands forming a $\overset{\cdots TCA \cdots}{\cdots ACT \cdots}$ mismatch, a clear dependence on mismatch location is visible. Distal mismatches far from the toehold have very little impact on the rate.

The rate constants for invaders forming the same type of mismatch at a central or proximal location, however, are orders of magnitude lower than those of the non-mismatched invaders. This dependence on mismatch location is considerably more pronounced for invading strands

with shorter toeholds. It is also visible that a $^{\cdots TTA\cdots}_{\cdots ACT\cdots}$ mismatch in the central location has a lesser effect on reaction rate than a thermodynamically more unfavourable $^{\cdots TCA\cdots}_{\cdots ACT\cdots}$ mismatch in the same position.

Although some coarse patterns emerge from this data, a more detailed study is introduced below. Together with simulation results in section 2.3, fundamental mechanisms underlying the phenomena observed are then explored systematically.

Error estimation

In order to estimate the experimental error attached to the reaction rate constants, the displacement process of the 7 nt toehold invader forming a distal $^{\cdots TCA\cdots}_{\cdots ACT\cdots}$ mismatch was measured ten times. Using the same methods for rate extraction from the experimental data as above, a mean rate constant and standard deviation of $(2.1 \pm 0.2) \times 10^6\,\mathrm{M^{-1}s^{-1}}$ were determined. The standard deviation of $\pm 10\,\%$ gives a measure for the experimental error attached to all data points in Figure 2.2. Variations in reactant quantities from pipetting are assumed to be the main contribution to the overall measurement error. A further source of error, which was presumed to be small and was thus not tested for, is that samples of unpurified oligonucleotides, due to synthesis imperfections, may contain a varying percentage of strands which deviate from the intended sequence [157]. An overall systematic error, which, for example, could arise from the use of miscalibrated pipettors, is assumed to be small, as the rate constants determined for non-mismatched invaders are comparable to those determined in other studies [150, 156]. Errors on the rate constants determined for individual systems may also arise from slightly different stock concentrations of invading strands. This would likely be due to pipetting errors when preparing the stock from lyophilised oligonucleotides, and, although assumed to be small, might explain why invading strands

forming a distal mismatch appear to displace faster than those which do not form a mismatch at all.

Validity of the second-order approximation

Figure 2.3 presents exemplary fits for two different invaders from which the respective values of k in Figure 2.2 were derived. The fit shown at the top of the figure, which is for a 6 nt-toehold invader forming a distal mismatch, describes the data well. Apart from minor deviations that were likely due to thermal fluctuations, this is also the case for most other invaders tested. For 10 nt-toehold invaders with either a central or a proximal mismatch, however, systematic deviations from a second-order analytic fit were observed (bottom of Figure 2.3). This is assumed to be due to a failure of the two-state reaction model forming the foundation of how TMSD processes were reduced to second-order reactions. If the intermediate states in which all three strands are bound in one complex are not resolved quickly, either through incumbent displacement or toehold detachment, they can no longer be assumed to be negligible within the overall reaction scheme. This appears to be the case for the 10 nt-toehold invaders with a central or proximal mismatch, which have very low rates of toehold detachment due to their high toehold binding energy, and also have low rates of displacing incumbent from a three-strand complex due to the presence of a mismatch. As toehold hybridisation is comparatively very fast at the concentrations used in the measurements presented here, resolution of the three-strand complex via incumbent displacement, a first-order reaction, dominates overall kinetics. Thus, a fit according to first order kinetics

$$[O](t) = [I]_0 - [I]_0 e^{-k^* t} \tag{2.2}$$

where k^* represents a first-order reaction rate constant, describes the data better.

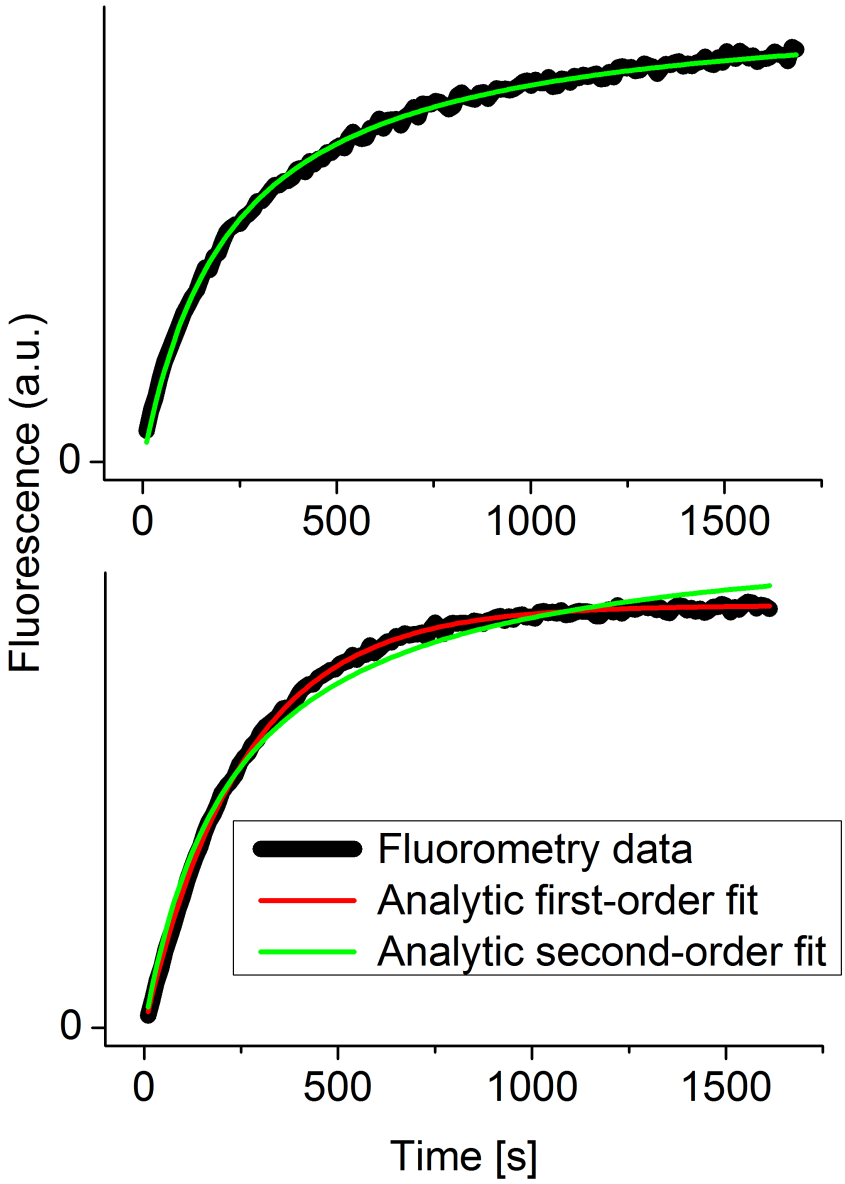

Figure 2.3: Comparison of fit quality for two different invading strands. **Top:** Fluorimetry data from a measurement performed with a 6 nt-toehold invader which forms a $^{\cdots TCA\cdots}_{\cdots ACT\cdots}$ mismatch in the distal position of the displacement domain (black). To determine k, a fit using the analytic equation 2.1 is performed (green). **Bottom:** Fluorimetry data from a measurement performed with a 10 nt-toehold invader which forms a $^{\cdots TCA\cdots}_{\cdots ACT\cdots}$ mismatch in the central position of the displacement domain (black). As above, the green line represents a second-order fit according to equation 2.1, which was used to determine the value of k. The red line represents a fit according to equation 2.2, which is derived from first-order kinetics.

Invader	$k_{5\,nM}/k_{avg}$	$k_{10\,nM}/k_{avg}$	$k_{20\,nM}/k_{avg}$
6 nt toehold, distal C-C	0.88	1.05	1.07
7 nt toehold, central C-T	0.96	1.00	1.04
8 nt toehold, proximal C-C	1.05	0.94	1.02
8 nt toehold, central C-C	1.52	0.84	0.64
9 nt toehold, central C-C	1.22	1.01	0.76
9 nt toehold, distal C-C	1.08	1.06	0.87
10 nt toehold, proximal C-C	1.49	0.93	0.58
10 nt toehold, central C-T	1.25	1.01	0.74
7 nt toehold, no mismatch	1.11	0.86	1.03

Table 2.1: Second order rate constant k from experiments conducted at different reactant concentrations, divided by the average $k_{avg} = {}^1/_3\,(k_{5\,nM} + k_{10\,nM} + k_{20\,nM})$ for different, exemplary invading strands.

The limited adequacy of the second-order approximation for some reactions also becomes evident from the data in Table 2.1. This table describes the concentration robustness of several mismatch displacement processes. They were tested at half (5 nM) and double (20 nM) the concentration at which they were tested to determine the values of k stated in Figure 2.2. For most of the systems, the rate constant does not depend on concentration, as expected from a second-order process. The deviations are commensurate with the measurement errors determined previously. For the 10 nt-toehold invaders and two of the 8 nt- and 9 nt-toehold invaders, however, there seems to be an inverse correlation between rate constant and concentration, hinting at the possibility that a second-order approximation is less appropriate for these systems. Considering, however, that even for these systems, second-order data fits still described the data reasonably well, their use to provide with k a very general measure for how fast reactions progress is justifiable.

2.2.2 *Using a separate reporter complex*

Design rationale and experimental implementation

There are several reasons why it is advantageous to use a system which reports on the progress of a TMSD reaction without labelling the reactants of interest directly. As noted by Zhang and Winfree [156], whose design for a reporter system was closely followed here, fluorescent labels can influence the kinetics and thermodynamics of a strand displacement reaction [84, 77]. Direct labelling of reactants may thus interfere with an accurate quantification of such reactions. Another reason mentioned by Zhang and Winfree is that the presence of such a reporter system suppresses a reverse reaction in which previously displaced incumbent at least temporarily rehybridises a target strand occupied by an invader. A separate reporter system can thus serve as a 'sink' for displaced incumbent. This allows for an isolated measurement of the forward rate of a displacement reaction, even if the invader-target duplex is not thermodynamically stable enough for the reverse reaction to be negligible.

Two further reasons make the use of a reporter complex particularly useful in the context of this work. Firstly, it allows for increased system design flexibility. As can be seen by the design of the system used in the previous subsection, direct labelling of reactants only allowed for a limited number of positions for thermodynamically equivalent mismatches. The introduction of further sites for a $\substack{\cdots TCA \cdots \\ \cdots ACT \cdots}$ mismatch in the above system would have made the sequence of the displacement domain quite repetitive. Oligonucleotides with repetitive sequences should be avoided, however, as they carry the risk of forming a wide variety of metastable, thermodynamically suboptimal duplexes. Use of a reporter system allows free variation of the displacement domain, with the exception of the seven distal bases, which are determined by the reporter toehold.

Such variation in a system, such as the one used in the previous subsection, would require the production of a corresponding variety of fluorescently labelled strands, which would be uneconomical. Secondly, using a separate reporter system allows for the important proof-of-principle that mismatch-controlled TMSD can be integrated seamlessly into nucleic acid CRNs.

In order to test a wide variety of different positions for a $^{\cdots TCA \cdots}_{\cdots ACT \cdots}$ mismatch whilst keeping the different invader/target/incumbent systems as thermodynamically equivalent as possible, a specific sequence design strategy was implemented. In all systems, the target has a 10 nt toehold domain (blue, Table 2.2, Figure 2.4) and a 17 nt displacement domain (black). The incumbent is fully complementary to this displacement domain, and further has a domain which triggers the reporter system if the incumbent is free in solution (italic). With the toehold for this reporter displacement process (black, italic) initially sequestered in the duplex with target, the single-stranded section of incumbent which corresponds to the reporter complex displacement domain (green, italic) is practically inert on the time scales of the experiments presented here. The invading strands have a 7 nt toehold (blue) and a 17 nt section binding to the target's displacement domain (black).

In the first system in Table 2.2, the invading strand does not form a mismatch with the target. In order to introduce a $^{\cdots TCA \cdots}_{\cdots ACT \cdots}$ mismatch, the second base from the invader toehold was changed from a G to a C (mismatch position: 2, black, bold), whilst target and incumbent strand remained unchanged from the non-mismatched system. In order to test the same mismatch triplet in other positions without drastically altering the change in free energy due to displacement ($\Delta\Delta G^{\circ}_{disp}$), the three bases of which it is composed were moved one nucleotide at a time, with the neighbouring nucleotides changing position accordingly. For the mis-

MM pos.	Reactant	Sequence 5′ to 3′	$\Delta\Delta G^{\circ}_{disp}$ [kcal/mol]
none	target	CCC TCC ACA TTC AAC CTC AAA CTC ACC	−10.0
	incumbent	*TGG TGT TTG TGG GTG TGG TGA GTT* TGA GGT TGA	
	invader	GGT GAG TTT GAG GTT GAA TGT GGA	
2	target	CCC TCC ACA T**TC A**AC CTC AAA CTC ACC	−4.2
	incumbent	*TGG TGT TTG TGG GTG TGG TGA GTT* TGA GGT TGA	
	invader	GGT GAG TTT GAG GT**T CA**A TGT GGA	
3	target	CCC TCC ACA TA**T CA**C CTC AAA CTC ACC	−3.1
	incumbent	*TGG TGT TTG TGG GTG TGG TGA GTT* TGA GGT GAT	
	invader	GGT GAG TTT GAG G**TC AT**A TGT GGA	
4	target	CCC TCC ACA TAC **TCA** CTC AAA CTC ACC	−3.1
	incumbent	*TGG TGT TTG TGG GTG TGG TGA GTT* TGA GTG AGT	
	invader	GGT GAG TTT GAG **TCA** GTA TGT GGA	
⋮	⋮	⋮	⋮
10	target	CCC TCC ACA TAC CTC AAA **TCA** CTC ACC	−3.1
	incumbent	*TGG TGT TTG TGG GTG TGG TGA GTG* ATT TGA GGT	
	invader	GGT GAG **TCA** TTT GAG GTA TGT GGA	
12	target	CCC TCC ACA TTC AAC CTC AA**A CT**C ACC	−4.2
	incumbent	*TGG TGT TTG TGG GTG TGG TGA GTT* TGA GGT TGA	
	invader	GGT G**AC T**TT GAG GTT GAA TGT GGA	
14	target	CCC TCC ACA TTC AAC CTC AAA C**TC A**CC	−4.2
	incumbent	*TGG TGT TTG TGG GTG TGG TGA GTT* TGA GGT TGA	
	invader	GG**T CA**G TTT GAG GTT GAA TGT GGA	

Table 2.2: Oligonucleotide sequences depicting design rationale for the systems using a separate reporter complex. Sequence colours were chosen to correspond with the reaction scheme shown in Figure 2.4.

match position 3 target, for example, the mismatch site (bold) moved one nucleotide towards the 3'-end relative to the mismatch position 2 target, with the A that was previously to its right skipping to its left. According sequence changes were made to the incumbent and the invader. The sequence permutations due to shifting of the mismatch triplet were continued one nucleotide at a time in all three strands until mismatch position 10. For higher mismatch positions, target and incumbent sequences could not be changed without interfering with the region of incumbent that interacts with the reporter (italic). Instead, the same incumbent and target as for the system without a mismatch were used, which within that region contained a native site for a $\substack{\cdots TCA \cdots \\ \cdots ACT \cdots}$ mismatch, as well as a site for a thermodynamically equivalent $\substack{\cdots ACT \cdots \\ \cdots TCA \cdots}$ mismatch. The sequence of the invader alone could then be adjusted so that a mismatch is formed at these sites. With this approach, variations in $\Delta\Delta G^{\circ}_{disp}$ between different mismatch positions were minimised. As in the previous subsection, the reactions presented here were observed using bulk fluorimetry.

Rate constant determination

As in subsection 2.2.1, it is assumed here that both the displacement of the incumbent as well as the displacement of reporter incumbent are second-order reactions. It is further assumed that displacement of the reporter incumbent is effectively irreversible and much faster than any reverse reaction of free incumbent with invader-target duplex, making the primary displacement reaction irreversible as well. With these assumptions, the systems can then be described as $I + TO \xrightarrow{k} IT + O$ and $O + RQ \xrightarrow{k_{rep}} R + OQ$ (where I: invader, T: target, O: incumbent, R: reporter incumbent, Q: reporter target, XY: duplex consisting of strands X and Y, k: second-order reaction rate constant for invader displacing incumbent and k_{rep}: second-order reaction rate constant for incumbent trig-

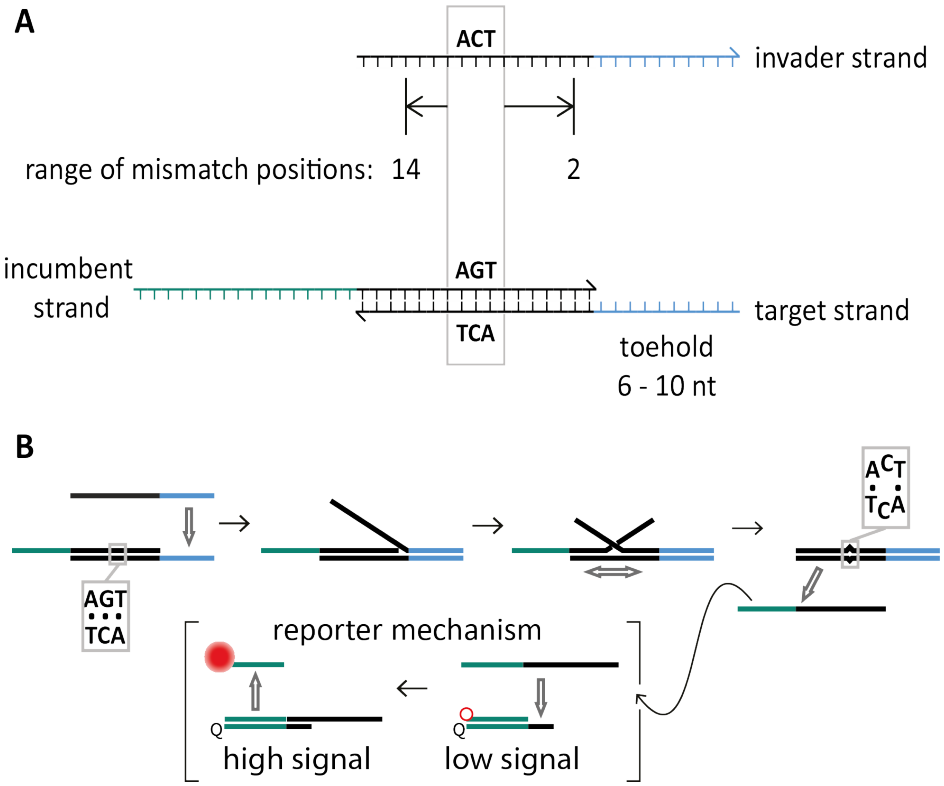

Figure 2.4: Schematic representation of the systems using a separate reporter complex. **A:** An invading strand is capable of displacing an incumbent from a target strand, using a toehold (blue). During displacement, the invader can form a $\overset{\cdots TCA \cdots}{\cdots ACT \cdots}$ mismatch with the target in the displacement domain (black). The mismatch position can be varied between position 2, where the C-C pairing occurs two nucleotides from the toehold, to position 14, which is four nucleotides from the invader's 5' end. Furthermore, the toehold length can be varied. **B:** Incumbent displaced by the invader (top reaction) triggers a reporter complex, in which a strand with an initially quenched fluorophore is released, increasing fluorescence emission. This can be used to track progress of the reaction. Release of reporter fluorescence is negligible when no invader is added, as the toehold the incumbent uses for the reporter displacement reaction is sequestered in the initial incumbent-target duplex.

gering the reporter), yielding the following system of differential equations:

$$\begin{aligned}
\frac{d[I]}{dt} &= \frac{d[TO]}{dt} = -\frac{d[IT]}{dt} = -k[I][TO] \\
\frac{d[O]}{dt} &= k[I][TO] - k_{rep}[O][RQ] \\
\frac{d[R]}{dt} &= \frac{d[OQ]}{dt} = -\frac{d[RQ]}{dt} = k_{rep}[O][RQ]
\end{aligned} \tag{2.3}$$

As above, the parameter of interest is k. The additional parameter k_{rep} was determined through experiments in which the different incumbent strands reacted with the labelled reporter complex in isolation. Fluorescence data was processed as in subsection 2.2.1. As expected, all incumbents behaved very similarly when reacting with reporter complex (Figure 2.5), yielding an average k_{rep} of $(2.1 \pm 0.4) \cdot 10^6\,\mathrm{M}^{-1}\mathrm{s}^{-1}$. The individual k_{rep} of each incumbent species was used for subsequent data analysis.

Fitting was achieved by minimising the squared differences between the recorded data and numerical solutions of equation system 2.3 with variable k and a scaling factor to relate fluorescent signal in arbitrary units to $[R](t)$ in Moles per litre. Like the introduction of a separate reporter complex, this method to determine k closely follows the work of Zhang and Winfree [156]. Figure 2.6 shows the rate constants thus determined.

The reaction rate constants in Figure 2.6 support some of the observations made in subsection 2.2.1, in particular the dependence of k on mismatch position and the different impact a mismatch can have on the displacement process of invaders depending on toehold length. The finer resolution of mismatch positions achieved with the introduction of a separate reporter system also allows for some additional observations. One of the most striking of these is that for the 6 and 7 nt toehold invaders, there is only a very gradual increase in rate constants when moving the

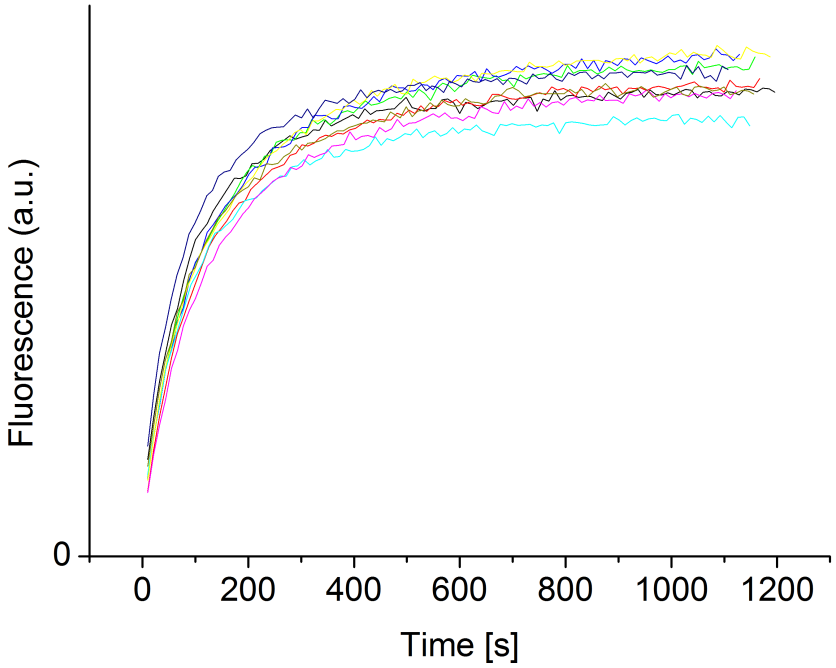

Figure 2.5: Reporter fluorescence increases as the different incumbent species displace a fluorescently labelled strand from the reporter complex. These traces were used to characterise the reporter response, yielding a k_{rep} for each incumbent species.

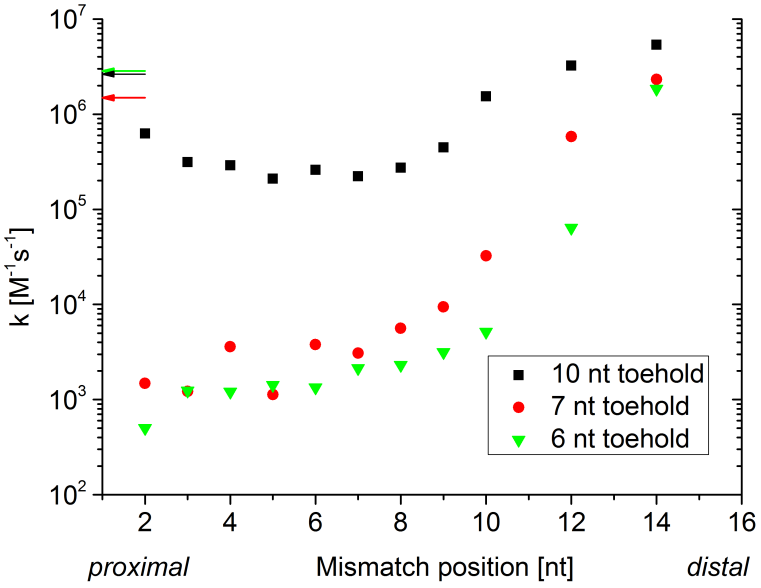

Figure 2.6: Effective rate constants of displacement determined for 33 invaders with different toehold lengths and mismatch positions. Mismatch positions are defined by the distance of the mismatch from the toehold in nucleotides. They are roughly classified as proximal (close to the toehold), central and distal (far from the toehold). Rate constants measured for non-mismatched invaders of all three toehold lengths are indicated with arrows.

mismatch from position 2 to position 8, but from positions 9 to 14, the dependence is exponential. The presence of these two regimes could not be inferred from the data in Figure 2.2 alone, although it was already visible there that the difference in rate constant between the proximal and the central $^{\cdots TCA\cdots}_{\cdots ACT\cdots}$ mismatch systems is much smaller than that between the central and the distal mismatch systems. A further remarkable observation in this context is that for invaders with the 10 nt toehold, rate constants are greater for mismatches very close to the toehold than for those in positions 5 to 8, whereas this is not the case for invaders with shorter toeholds.

Error estimation

As in the previous subsection, one system was measured ten times to get an estimate of the error attached to the rate constants presented here. The system chosen was the one in which a 7 nt toehold invader forms a $^{\cdots TCA\cdots}_{\cdots ACT\cdots}$ mismatch at position 10. The mean rate constant and standard deviation were determined to be $(3.6 \pm 0.8) \cdot 10^4\,\mathrm{M^{-1}s^{-1}}$. The standard deviation of $\pm 22\,\%$ is higher than it was for the directly labelled system. This can be explained by the presence of a greater number of reactant species. As the quantities of each reactant have a certain pipetting error attached to them, a greater number of reactant species in one reaction yields greater uncertainty in the quantities derived from the reaction data. As before, systematic errors are assumed to be small.

Validity of the second-order approximation

As for the reporter-free assay (subsection 2.2.1), the adequacy of assuming displacement reactions to be of second order was investigated. As for most systems that were tested to derive the data in Figure 2.6, the exemplary fits presented in Figure 2.7 describe the fluorimetry traces well.

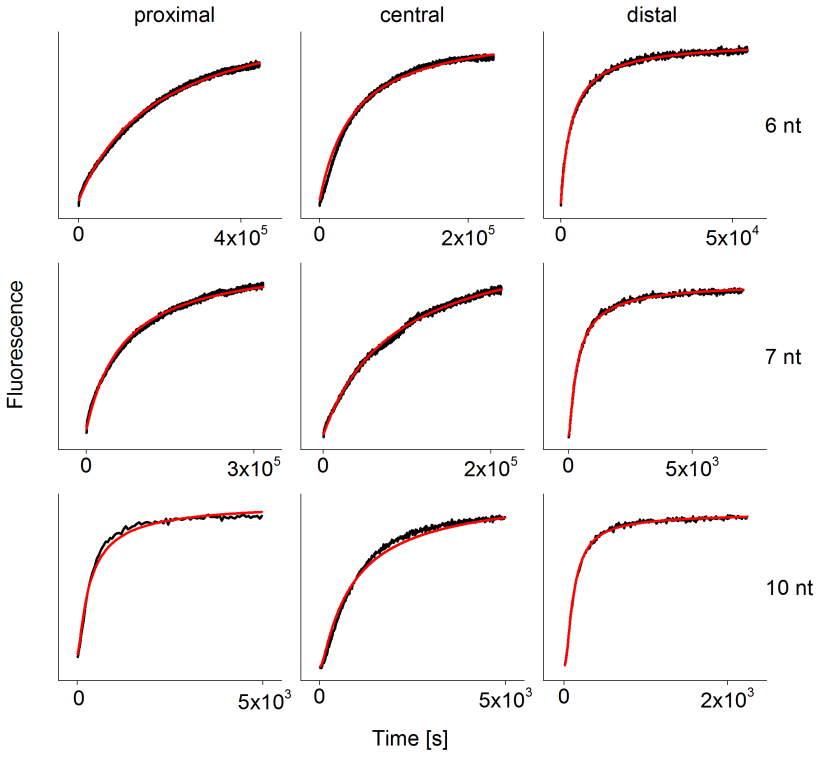

Figure 2.7: Exemplary fits to fluorescence traces recorded for invaders with three distinct mismatch positions and the three toehold lengths tested. The proximal systems formed a $\overset{\cdots TCA \cdots}{\cdots ACT \cdots}$ mismatch in position 2, the central systems in position 7, and the distal systems in position 12.

However, fits deviate systematically from the traces for the proximal- and central-mismatch invaders with 10 nt toeholds. As above, it is assumed that the reason for this is the comparatively slow resolution of the intermediate invader-target-incumbent complex for these systems. Despite these deviations, reaction speed is still characterised well for all systems with the fitting approach used, which allows for comparison of reaction rates across the different mismatch positions and toehold lengths tested.

Further approaches to data analysis

In addition to the numerical, differential-equation approach outlined above (Figure 2.8), two further approaches to derive a rate constant from the raw data were tested. In the first of these, fits were performed using the analytic approximation

$$[R](t) = [I]_0 - \frac{[I]_0}{1 + [I]_0 kt}, \tag{2.4}$$

implicitly assuming the reporter reaction to progress much faster than the initial displacement reaction. Fits using this approximation were implemented as for the directly-labelled systems. The second approach did not require any fitting; instead, the rate constant estimate was based on reaction half-completion times. As the raw fluorescence traces are in arbitrary units, application of this method first required a scaling factor to convert the fluorescence intensity to the concentration of free fluorescent reporter strand in Moles per litre. This factor was estimated from the saturation fluorescence of two displacement reactions performed with the non-mismatched, 10 nt-toehold invader within the complete reaction scheme as depicted in Figure 2.4. This allowed for the identification of the first point in time in each fluorescence trace in which the reaction had just surpassed half-completion. The respective half-completion times $t_{1/2}$ were then estimated through linear interpolation between this data

point and the previous one, in which half-completion had not yet been reached. Again assuming second-order kinetics and fast reporter reaction compared to the initial displacement reaction, the rate constant was then estimated as

$$k = \frac{1}{t_{1/2}\,[I]_0}.$$

The resulting rate constants for all reporter-based systems tested are plotted in Figure 2.8, comparing them to the rate constants derived with the original approach. It is immediately visible that all three methods yield very similar results, both qualitatively and quantitatively. Noteworthy disparities are mostly confined to very fast reactions (i.e. $k \gtrsim 10^6\,\mathrm{M^{-1}s^{-1}}$), which makes sense considering that the assumption of a much faster reporter reaction fails for these systems. Overall, these results confirm that the effects observed are robust against variations in data analysis methodology.

2.3 EXPLORING THE PROCESS WITH INCREASED RESOLUTION

The introduction of a single mismatch destabilises a duplex by $\Delta\Delta G^\circ_{mm} \approx 2\text{-}7\,\mathrm{kcal/mol}$ relative to a fully complementary duplex [95, 152]. In the picture of a three-state displacement process as depicted in Figure 1.8, there is no reason to assume that such a defect would have a different impact on reaction rate depending on its location in the displacement domain. This position dependence, however, was one of the most striking features of the experimental data presented in the previous section. The solution to this paradox lies, of course, in the fact that the step from the three-strand intermediate to incumbent displacement is dominated by a sequential branch migration process starting at the toehold, rather than by pathways in which the invader forms base pairs in the displace-

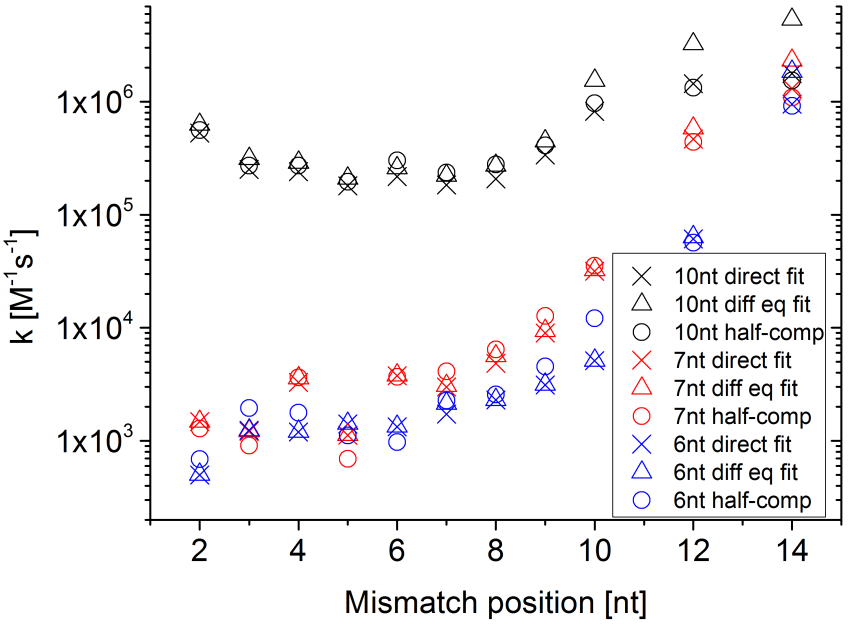

Figure 2.8: Comparison between rate constants determined from the data with a numerical fit based on a system of differential equations ('diff eq fit', equations 2.3), a simpler analytic description which assumes that the entire reaction network can be approximated as a bimolecular, second-order process as in equation 2.4 ('direct fit'), and an approach based on half-completion times ('half-comp'). All mismatched systems demonstrated in Figure 2.6 were tested with the two alternative methods.

ment domain in random order. Although this is well-known (e.g. [122]), it has played virtually no role in the design of previous applications relying on TMSD reactions. In order to explain the dependence of reaction kinetics on mismatch position, it becomes necessary to use a model with higher resolution, through which the ordered manner in which strand displacement occurs is appropriately described.

2.3.1 *Using a coarse-grained model of DNA*

The need for a model capable of describing DNA strand displacement involving mismatches in different locations can be satisfied with the use of oxDNA. OxDNA is a coarse-grained model of DNA, which has previously been applied successfully to the study of fundamental processes such as hybridisation [91] and overstretching [102] of DNA, as well as complex synthetic DNA nanodevices [89, 128, 29]. More importantly for this work, it has also been used to simulate TMSD [122], accurately replicating reaction rate constants determined previously by Yurke and Mills as well as Zhang and Winfree [150, 156], offering novel qualitative insights into this process.

This model has been applied for similar purposes here. Initially, oxDNA was used to simulate a set of four TMSD systems which had also been tested experimentally[1]. These systems all comprise invaders with 5 nt toeholds, which can form a proximal, central or distal $\overset{\cdots TCA \cdots}{\cdots ACT \cdots}$ mismatch, or no mismatch at all ('perfect'). The quantitative results of these simulations, presented in Table 2.3, were found to be in good agreement with the experimentally determined values.

1 As these simulations were conducted by Thomas E. Ouldridge, the exact methods are not presented here, but rather the results in the context of experimental findings. For full details on simulation methodology, the reader is referred to [73].

Invader	$k/k_{perfect} \pm SE$ (oxDNA)	$k/k_{perfect}$ (experiment)
Perfect (no mismatch)	1.00 ± 0.07	1
Distal mismatch	0.57 ± 0.07	0.4
Central mismatch	0.0093 ± 0.0022	0.008
Proximal mismatch	0.0035 ± 0.0015	0.001

Table 2.3: Comparison between rate constants determined through simulation and experiment for the reactions described in the text. Rate constants are expressed relative to the non-mismatched system; mapping simulation to experimental time scales as required for comparison of absolute rate constants is non-trivial [29]. The standard errors of the values determined via oxDNA were determined as described in [73]. Errors on the experimental values can be assumed to be of similar magnitude as those determined in subsection 2.2.2.

The quantitative agreement between experiment and simulation suggests that the oxDNA model can be used to further the qualitative understanding of the measurements presented earlier. For this purpose, oxDNA was configured to produce free energy landscapes with single base pair resolution for those systems introduced in the previous paragraph.

These free-energy landscapes are shown in Figure 2.9. For all landscapes, the state in which the invader has formed five base pairs with the target, whilst the incumbent maintains 16, was found to have the lowest free energy of all states sampled. The free energy of this state was set to zero, so that all other free energy values depicted are relative to this state[2].

In the landscape for the non-mismatched (or 'perfect') system, a free-energy valley is discernible which runs diagonally from state $(5, 16, 0)$

2 Note that only those states describing the three-strand intermediate invader-target-incumbent are included. Further note that in some states, the overall number of base pairs formed by the invader and the incumbent with the target exceeds 21, the total number of bases available for hybridisation. This is due to a relatively low free-energy threshold used in the model to categorise a binding interaction as a full base pair. The choice of the threshold can lead to an overestimation of the total number of base pairs formed.

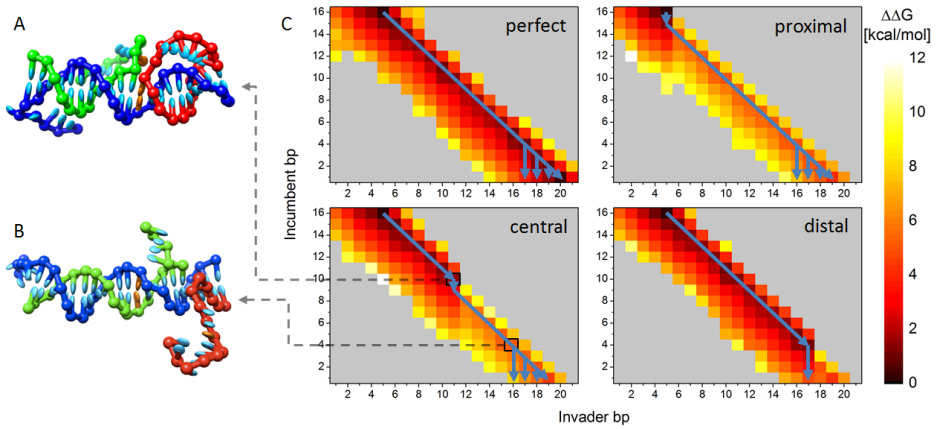

Figure 2.9: Free energy landscapes and representative states for TMSD processes, derived from oxDNA simulations. **A & B:** Representative simulation snapshots of two different possible microstates of a displacement process where a central $\overset{\cdots TCA\cdots}{\cdots ACT\cdots}$ mismatch (gold) is formed between the invading strand (green) and the target (blue), whilst the non-mismatched incumbent (red) is displaced. **C:** free energy landscapes for invading strands with varying mismatch position, where each state is defined by the number of base pairs formed between the incumbent and the target (vertical axes, 'incumbent bp') and the number of base pairs formed between invader and target (horizontal axes, 'invader bp'). Change in free energy $\Delta\Delta G$ of each state relative to the state invader bp $= 5$, incumbent bp $= 16$ is indicated by colour. Grey states are unsampled. Blue arrows indicate representative reaction pathways from toehold initiation to displacement.

to state $(20, 1, 0.3\,\text{kcal/mol})$[3]. These landscapes can be synthesised with what is already known about TMSD, such as the dependence of reaction rates on toehold length and the random walk branch migration process dominating displacement.

State $(5, 16, 0)$ represents the invading strand maintaining a fully hybridised toehold with the target, with strand displacement not yet initiated. This explains why this particular state has the lowest free energy of all states depicted:

1. the maximum number of base pairs are formed, as is the case for all states along the diagonal $(5+n, 16-n, \ldots)$, $n \in [0, 15]$;

2. there is only one single-stranded overhang protruding from the intermediate complex which gives this state a uniquely low free energy even compared to the other states on the diagonal, all of which have two single-stranded overhangs [122].

Starting from state $(5, 16, 0)$, the system can transition along a pathway consisting of energetically similar states in which all bases are paired to ultimately arrive at a state where the incumbent maintains very few base pairs with the target, which can lead to incumbent detaching completely. This is congruent with the established branch migration model of strand displacement, if we assume that these diagonal states $(5+n, 16-n, \ldots)$ predominantly represent full toehold initiation with the invader having formed no base pairs beyond a branch migration point n nucleotides from the toehold[4].

This suggested dominant pathway for strand displacement is indicated in the landscape for the 'perfect' system in Figure 2.9 by the di-

3 Here and in the following, states within these landscapes are abbreviated as (invader bp, incumbent bp, $\Delta\Delta G$ [kcal/mol]).

4 The single base-pair resolution is sufficient to explore the experimental results of this work. The reader is referred to the work of Srinivas et al. [122] for a more detailed insight into the biophysics of branch migration and a discussion on how displacement processes transition from one state to the next via intermediates.

agonal blue arrow. It is, of course, possible and indeed likely for such reactions to deviate from this pathway to nearby states, or to travel back and forth along it. Note that the ability to enter states adjacent to those described by the diagonal arrow also allows for incumbent detachment before the invader has formed all possible base pairs with the target (vertical arrows at the bottom of the landscape). It further explains why shorter toeholds yield decreased overall rates of displacement, as states in which the invader maintains very few base pairs with the target, and is thus prone to abortive detachment, become more readily accessible from the dominant pathway.

Having established the use of this type of free energy landscape to describe the displacement process, the systems in which a mismatch is created between invader and target can now be explored. The free energy landscape for the proximal mismatch system differs quite drastically from the perfect system. Upon toehold initialisation, the system must reduce the total number of base pairs to initiate branch migration (indicated by the short vertical arrow at the top of the 'proximal' landscape). Thus, for the invader to displace incumbent, a free-energy barrier of $\Delta\Delta G > 5\,\text{kcal/mol}$ must be overcome first. Such a hurdle forces the system to remain in an invader bp ≤ 5 state for much longer compared to the non-mismatched system. This in turn increases the likelihood of the invading strand detaching from the target entirely, before any displacement of incumbent has allowed it to form a greater, more stabilising number of base pairs with the target strand.

It can thus be assumed that, on average, the invading strand in a proximal mismatch system has to go through a much greater number of toehold initialisation and detaching cycles before successful displacement compared to a non-mismatched invader, decreasing the overall reaction rate. Considering that detachment rate of the toehold is approximately

exponentially dependent on toehold length, this immediately explains why mismatches as barriers to displacement were generally found to impede invading strands with 6 nt or 7 nt toeholds considerably more than the 10 nt toehold invaders, even though the non-mismatched invaders of all three toehold lengths showed very similar rate constants. It is also worth noting that once the mismatch is enclosed by the invading strand, the probability of the system returning to a pre-mismatch (in this case, toehold-only) state will still be much greater than the probability of proceeding with further displacement.

If the same mismatch is formed in the central instead of the proximal position, the according landscape reveals that the system now can initiate strand displacement from the toehold-only state just as readily as for the non-mismatched system. It can thus enter a wider variety of states before the mismatch is formed. As some of these states allow the incumbent to share up to eleven base pairs with the target, the rate of spontaneous invader detachment is reduced from the proximal mismatch case. Furthermore, once the mismatch is enclosed, the mean number of transitions required to achieve full incumbent displacement is reduced. Both of these effects lead to the slight increase in overall reaction rate observed both in experiment and simulations.

There is one exception to this, which was observed in the experiments with 10 nt toehold invaders. In Figure 2.6, it is quite evident that such a system with a central mismatch, e.g. mismatch position ~8, does not have a significantly higher reaction rate constant as a system with a proximal mismatch. In fact, the proximal mismatch system seems to display a higher k than those with mismatch positions between 5 and 8. This effect, which may at first appear rather counter-intuitive, can be explained with the exceptionally high stability of a 10 nt toehold. If the toehold of the invader is so stable that the rate for toehold detachment can be considered

negligible compared to the other transition rates in the overall process, increasing invader-target stability before mismatch enclosure by shifting the mismatch from a proximal to a central position will have a much lesser benefit on k than for systems with shorter, unstable toeholds.

Instead, offering a greater number of pre-mismatch enclosure states to the system now reduces the likelihood of the system to be in a state from which it can directly overcome the mismatch. For a proximal mismatch, however, a 10 nt toehold invader can almost incessantly attempt to negotiate the barrier posed by the mismatch. It is worth noting that this is a rather subtle observation, which was only made for systems using a separate reporter complex and not the directly labelled ones. Due to the possible influence of the labels on the latter reactions, however, it is reasonable to put more trust into those observations made with unlabelled reactants.

Lastly, the free energy landscape of the system with a distal mismatch (Figure 2.9, bottom right) appears quite similar to the system without a mismatch. In comparison to the proximal- and central-mismatch systems, further pre-mismatch states are readily accessible, making a return to a toehold-only state once branch migration is underway less likely. Once the mismatch has been enclosed, the distance to full displacement is also shortened relative to the other two mismatched systems. Further exploration is required, however, to understand if this alone explains why displacement in such systems is almost as fast as in non-mismatched ones, and why the rate constant dependence on mismatch location becomes so much more sensitive as the mismatch is moved towards the distal end of the displacement domain.

The oxDNA simulations reveal that there exists a displacement 'shortcut'. As branch migration progresses towards the distal end, it is now the incumbent strand that is increasingly likely to detach spontaneously.

Thus, as the mismatch is moved closer to the distal end, once branch migration reaches that barrier, the incumbent may detach entirely whilst or even before the invading strand attempts to enclose the mismatch. Just like the toehold off-rate, the rate for spontaneous detachment of incumbent can be assumed to be exponentially dependent on the number of base pairs the strand maintains with the target, hence the exponential dependence for distal mismatches and the two distinct regimes in Figure 2.6.

For the distal mismatch system presented here, oxDNA calculated that in 80 % of the cases in which branch migration reaches the mismatch site, the incumbent detaches without the mismatch actually being formed. For the central- and proximal-mismatch systems, the incumbent maintains up to 10 or 16 base pairs with the target before mismatch enclosure by the invading strand, respectively. Thus, the probability of spontaneous incumbent detachment without prior enclosure of the mismatch is virtually zero. As indicated by the vertical blue arrows at the bottom of each of the free energy landscapes, it can be assumed that spontaneous incumbent detachment is a possible pathway to reaction completion in all systems presented, but it is only the distal mismatch system for which reaction rate is strongly influenced by this.

A further point should be made regarding these free energy landscapes: the states beyond mismatch enclosure are very similar in terms of free-energy difference to the toehold-only states in all landscapes involving a mismatch. This further supports the earlier claim that the remarkable dependence of reaction rate constants on mismatch position is not due to different product thermodynamics, but is rather a purely kinetic effect due to the sequential nature of strand displacement reactions.

2.4 ELONGATED DISPLACEMENT DOMAIN

To investigate the above conclusion that the relatively weak impact of a distal mismatch on reaction rate is largely due to spontaneous incumbent dissociation before the mismatch is enclosed, a further series of experiments was conducted. For these experiments, the displacement domain was elongated by eight additional nucleotides. Subsequently, the reaction rate constants for appropriate 7 nt-toehold invaders which form a mismatch in different locations along this displacement domain were measured as in subsection 2.2.2. Figure 2.10 shows the results of these measurements, comparing them to the relative rate constants of 7 nt-toehold invader systems with the shorter, 17 nt displacement domain presented above.

As expected, this data proves that the rate constant is mostly dependent on mismatch distance from the distal, not the proximal end of the displacement domain. Otherwise, the rate constants for the elongated displacement domain systems could be expected to be as high for mismatch positions 10 to 15 as they are for the short displacement domain systems. Instead, the strong dependence in mismatch location becomes relevant for both types of systems towards the distal end of their respective displacement domains. In fact, mismatches with the same distance from the distal end lead to virtually equal relative rates for both systems. It has to be noted, however, that the long displacement domain data presented here comes with a caveat: the reporter response to the incumbent species of these systems was not as uniform as it was for the short displacement domain systems. The scaling factor used as fitting parameter together with k_{rep} displayed a standard deviation of $\pm 5\,\%$ for the 17 nt displacement domain incumbents, whereas this was $\pm 27\,\%$ for the 25 nt displacement domain incumbents. There was a further noticeable

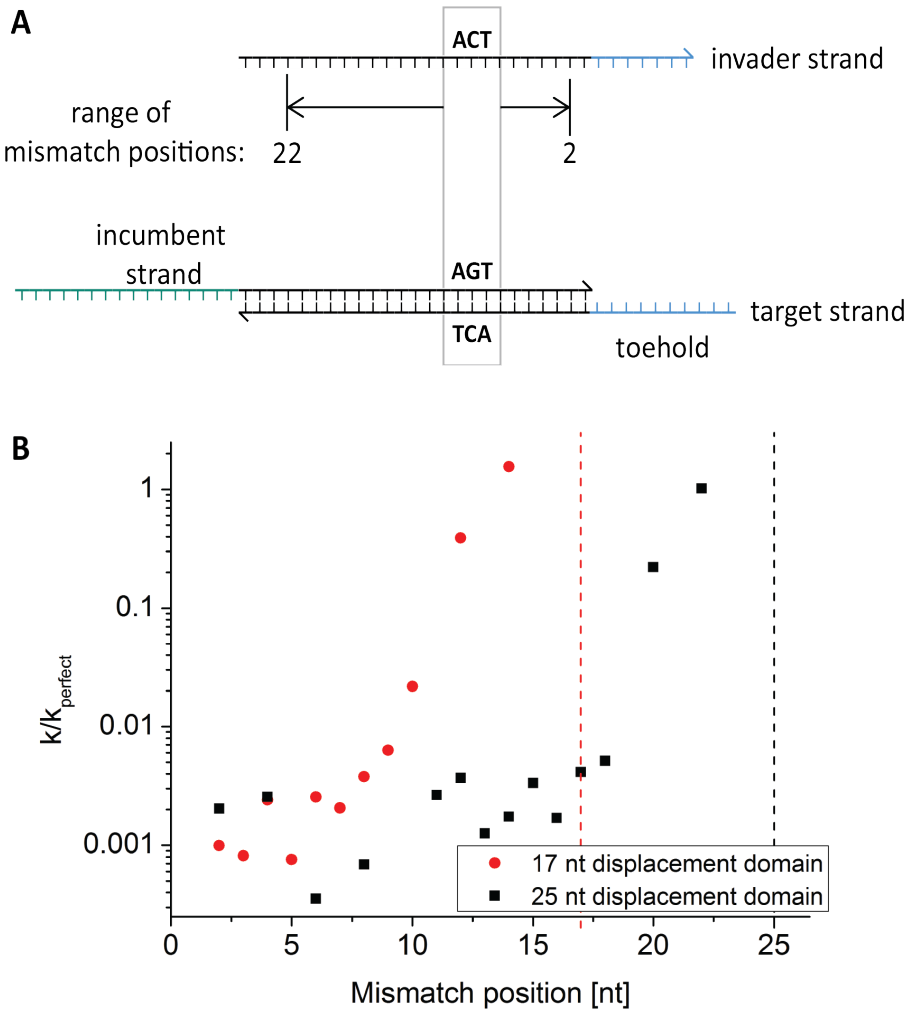

Figure 2.10: Comparison to TMSD involving an elongated displacement domain. **A:** Schematic depiction of the elongated displacement domain system. In this system, the mismatch position can vary between 2 and 22, instead of 2 and 14 as in the original system (Figure 2.4A). **B:** Reaction rate constants for both systems with varying mismatch position (horizontal axis), relative to reaction rate constants of correspondent, non-mismatched systems. The original system has a 17 nt displacement domain (red dots), the other an elongated, 25 nt displacement domain (black squares). For both, the invading strand initiated displacement via a 7 nt toehold. The dashed lines represent the distal ends of either displacement domain.

difference in mean reporter rate constants for both types of systems, as $\overline{k_{rep}^{25\,\mathrm{nt}}} = (1.5 \pm 0.3) \cdot 10^6\,\mathrm{M^{-1}s^{-1}}$ and $\overline{k_{rep}^{17\,\mathrm{nt}}} = (2.1 \pm 0.4) \cdot 10^6\,\mathrm{M^{-1}s^{-1}}$. This can be attributed to the decreased synthesis reliability for longer oligonucleotides, as the long displacement domain incumbents had a length of 41 nt. Purification of these incumbent strand samples would likely lead to reduced variation of the scaling factor, and an increase in mean k_{rep}, if analyses with greater accuracy than those presented here were required.

2.5 COMPLEX KINETIC BEHAVIOUR USING MISMATCHES

2.5.1 *Competitive displacement*

The previous sections explored how the position of a mismatch can influence TMSD kinetics. To provide an experimental proof that this mechanism provides independent control of kinetic and thermodynamic bias, a simple device was designed in which two invaders compete for the same target strand, which is initially occupied by an incumbent. The invading strands initiate their respective displacement processes from toeholds at opposite ends of the displacement domain, similar to a system presented by Zhang (Figure 1.9B, [153]). In contrast to Zhang's system, however, in the device presented here, one invader can hybridise the entire displacement domain, making this a competitive, rather than a cooperative reaction.

The kinetic competition can be biased rather trivially by giving one of the invading strands a longer toehold than the other. In the context of the above findings, a more interesting possibility is the introduction of a mismatch which is formed by both invading strands with the target at the same location within the displacement domain. If this location is closer to one toehold than the other, it could be expected that the reaction

would be kinetically biased towards one invader, without necessarily biasing the equilibrium in the same direction.

Instead, this kinetically favoured invader can be made to be at a thermodynamic disadvantage, as the competing invader is capable of forming more base pairs with the target. The design of this system and the resulting kinetics are shown in Figure 2.11.

In this system, an incumbent-target duplex provides one 6 nt toehold to each invading strand. The 5′ invader faces a C-T mismatch 4 nt from its toehold, whereas the 3′ invader faces the same mismatch 13 nt from its toehold. Thus, the 3′ invader has a kinetic advantage over the 5′ invader when competing for the target. As the 5′ invader is capable of forming three additional base pairs with the target in the 3′ invader's toehold domain, however, it is thermodynamically favoured over the 3′ invader. The invaders are labelled with spectrally separate fluorophores Cy3 and Cy5, which are suppressed by quenchers on the opposite end of the invader as long as the strand is not bound to target. The fluorescence signals depicted in Figure 2.11 are the result of adding the equivalent of 10 nM incumbent-target duplex at $t \approx 0$ s to a mixed sample of 5′ invader and 3′ invader (10 nM each). Normalisation was achieved by adding an excess of free target strand and measuring the resulting signals, representing full hybridisation of each invader. Since the flourimeter used can only measure one pair of excitation/emission wavelengths at a time and requires a few seconds to switch between wavelengths, two identical experiments were performed for the data shown. Only one of the two invader signals was observed in each of these experiments and the resulting traces superimposed into one plot. This yields a higher temporal resolution without influencing the data otherwise.

A control experiment in which both traces were recorded by switching between wavelengths showed the same characteristics with lower tempo-

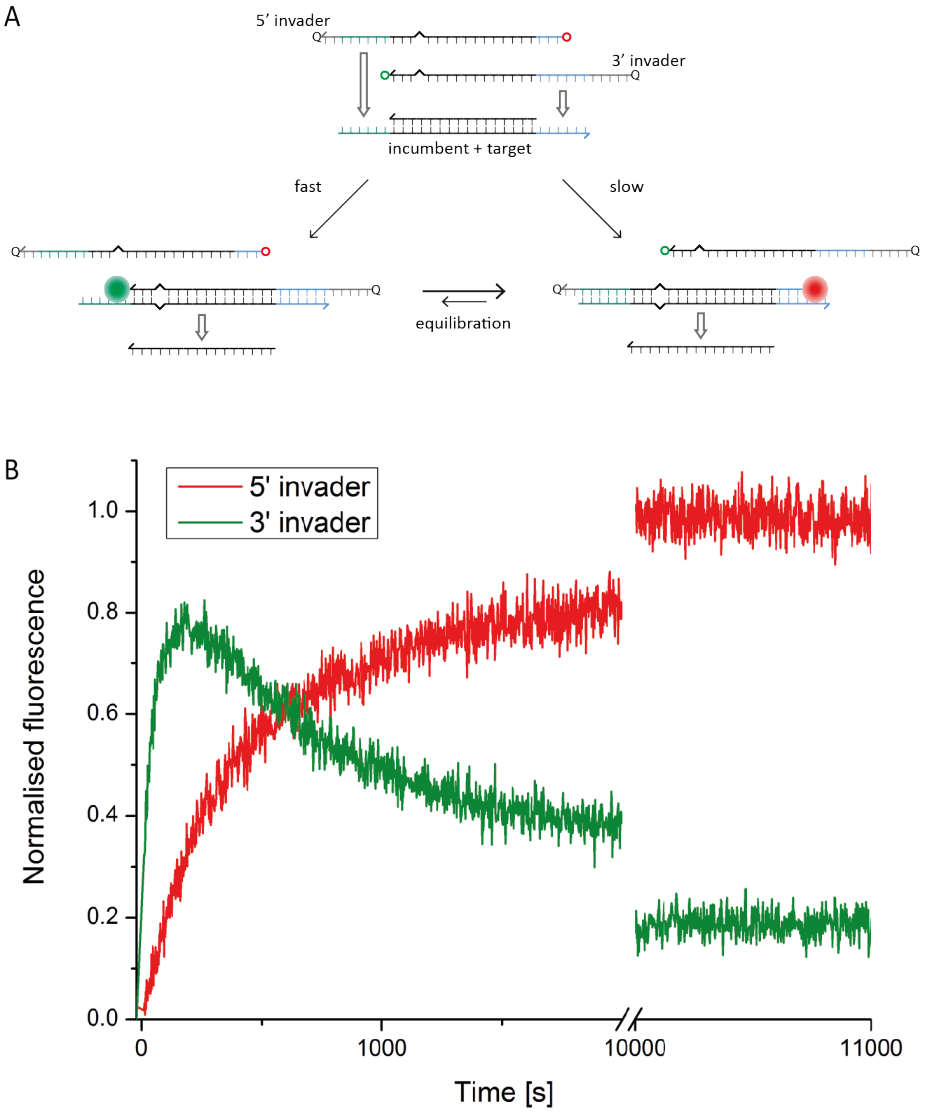

Figure 2.11: A reaction showing a transient response controlled by an asymmetrically positioned mismatch. **A:** Schematic representation of the reaction device. **B:** Fluorescence signals corresponding to the labels on each invading strand, normalised to the ratio of invader species bound to the target.

ral resolution and is depicted in Figure 2.12. Despite the lower temporal resolution, system behaviour is virtually identical. This figure contains a further pair of traces, in which the roles of the invaders have been reversed from the original system. To achieve this, the target-incumbent duplex was designed to create a mismatch closer to the 3′ invader toehold, and to give that invader a thermodynamic advantage. Although the transient response of the alternative system is less pronounced than in the original, it is still clearly discernible. These additional traces support the assumption that the behaviour observed is due to the intended functionality of the system, and not other transient effects, such as improper sample mixing. Note that these traces represent raw data to explain how normalisation of the data in 2.11 had been achieved (Figure 2.12 caption).

2.5.2 *Modelling competitive displacement*

Assuming that the relevant reactions of the competitive displacement process described above are

$$I_5 + OT \xrightarrow{k_5} I_5T + O$$

$$I_3 + OT \xrightarrow{k_3} I_3T + O$$

$$I_5 + I_3T \underset{k_{5\to3}}{\overset{k_{3\to5}}{\rightleftharpoons}} I_5T + I_3$$

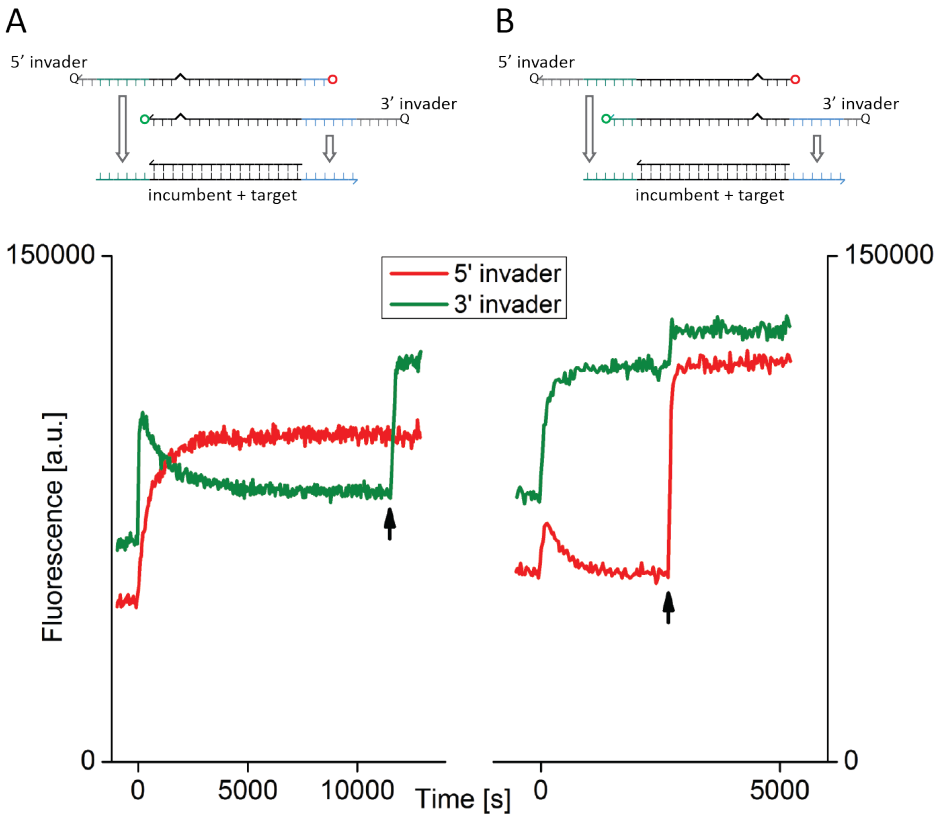

Figure 2.12: Further traces of competitive strand displacement biased by a mismatch. **A:** identical reaction to that depicted in Figure 2.11, with both traces having been recorded simultaneously. **B:** a system in which the roles of the two invaders had been reversed by altering the sequences of target and incumbent. To normalise these traces as in Figure 2.11, the initial signal (i.e. $t < 0\,\mathrm{s}$) of each trace would have been subtracted as a baseline. The baseline-corrected traces would then be divided by the fluorescent signal after addition of an excess of free target strand (indicated by the black arrows), assumed to represent full hybridisation of both invader species.

and further assuming that these reactions are effectively second order, they can be described with the following system of differential equations:

$$\begin{aligned}
\frac{d[OT]}{dt} &= -k_5[OT]([I_5]_0 - [I_5T]) - k_3[OT]([I_3]_0 - [I_3T]) \\
\frac{d[I_5T]}{dt} &= k_5[OT]([I_5]_0 - [I_5T]) + k_{3\to5}[I_3T]([I_5]_0 - [I_5T]) \\
&\quad - k_{5\to3}[I_5T]([I_3]_0 - [I_3T]) \\
\frac{d[I_3T]}{dt} &= k_3[OT]([I_3]_0 - [I_3T]) + k_{5\to3}[I_5T]([I_3]_0 - [I_3T]) \\
&\quad - k_{3\to5}[I_3T]([I_5]_0 - [I_5T])
\end{aligned} \tag{2.5}$$

where I_5 represents the 5′ invader, I_3 the 3′ invader, T the target and O the incumbent strand; k_x is the second-order rate constant for displacement of incumbent by I_x from the target strand, $k_{x\to y}$ is the second-order rate constant for displacement of I_x by I_y from the target strand, and $[I_3]_0$, $[I_5]_0$ and $[OT]_0$ are the initial concentrations.

A best-fit numerical solution for the device shown in Figure 2.11 can then be determined by extending the differential equation fitting method described in subsection 2.2.2. Instead of scaling parameters, the initial concentrations $[I_3]_0$, $[I_5]_0$ and $[OT]_0$ were chosen as fitting parameters in addition to the reaction rates. The initial concentrations were supposed to be equal at 10 nM each, as described above. It was observed experimentally, however, that the 3′-invader signal was equilibrating at a normalised fluorescence of ~0.2, and the 5′-invader signal at ~1, suggesting that there was a slight initial excess of OT in the solution[5]. As a consequence, the fitting routine was implemented to allow for small variations of initial concentrations from the intended value. This yields

5 This is consistent with the traces for the same system presented in Figure 2.12 (left). For both experiments, the incumbent-target duplex came from the same stock, suggesting that this stock solution may have been at a slightly greater concentration than intended. An incumbent-target excess is not observed for the alternative system, depicted on the right side of Figure 2.12.

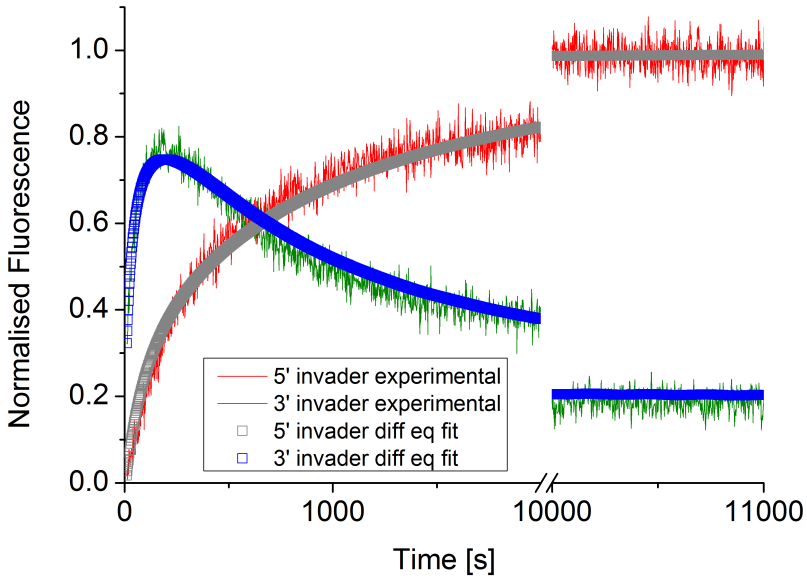

Figure 2.13: Data as in Figure 2.11. The blue and gray points represent solutions to equation 2.5 with the parameters stated in Table 2.4.

Rate constant	k_5	k_3	$k_{5\to3}$	$k_{3\to5}$
$[10^4\,\mathrm{M^{-1}s^{-1}}]$	48	240	0	13
Concentration	$[OT]_0$	$[I_3]_0$	$[I_5]_0$	
[nM]	11.8	9.6	10.1	

Table 2.4: Rate constants and initial concentrations derived from the fit in Figure 2.13.

the fit in Figure 2.13, with the corresponding second-order rate constants and initial concentrations given in Table 2.4.

The rate constants are consistent with the expectation based on the results presented earlier in this chapter. k_3 is comparable to the rate constants determined for a 6 nt toehold invader with a late mismatch. Due to the asymmetric position of the mismatch, k_5 is significantly smaller than k_3.

From the data presented in section 2.2, one might expect k_5 to be smaller. There are, however, two important differences in the free-energy landscape:

- both invaders are labelled with fluorophores at their respective toeholds, thus strengthening the toehold almost similar to an additional base pair [84];
- the mismatch in this system is a C-T mismatch, which provides less of an energetic barrier than a C-C mismatch, and thus has a reduced impact on reaction rate, as determined in section 2.2.

Closer investigation of the fitted 5′-invader trace also reveals that it rises faster than the experimental trace, suggesting that the fitting routine overestimates k_5.

The equilibration rate constants $k_{3\to5}$ and $k_{5\to3}$ reflect the thermodynamic advantage of the 5′ invader over the 3′ invader, which the 5′ invader derives from its three additional base pairs with the target. $k_{5\to3} = 0$, allowing for no free 5′ invader in solution at equilibrium, is consistent with the experimental observation that the fluorescence corresponding to the 5′ invader did not increase once an excess target was added. Note that the fitted initial concentrations confirm the above assumption that OT was in solution at an unintended excess.

2.6 TOEHOLD PROBING

The reaction rate of a TMSD process cannot be increased any further by increasing the toehold binding energy once a certain threshold has been reached. For such toeholds, k_{off} is so low that branch migration is rate-limiting, in contrast to weaker toeholds (usually $\lesssim 6\,\text{bp}$) where $k_{off} > k_{disp}$. When using toeholds beyond that threshold, even invaders with a partially mismatched toehold are capable of displacement at a rate similar to that of the fully complementary invader.

Previous sections explored how a single mismatch between invader and target can force a TMSD process to spend more time in states in which the invader has not formed any base pairs with the target beyond the mismatch. Placing a mismatch proximal to the toehold drastically reduces the overall displacement rate for short toeholds, because it now takes the invader more time to create enough base pairs with the incumbent to be safe from spontaneous detachment, which would reset the entire displacement process.

This kinetic confinement to certain states can be used to increase toehold specificity. By introducing a proximal energy barrier in the shape of a C-C mismatch, a TMSD process can artificially be forced to spend more time in states in which the invader is bound to the target only in the toehold domain, or not at all. As k_{disp} is thus increased, the higher k_{off} of a spurious-toehold invader impairs the overall k considerably more than if the barrier were not present.

This was tested experimentally using two different 8 nt invaders to react with the labelled incumbent-target complex introduced in subsection 2.2.1. The first invader has a displacement domain fully complementary to the target. In the toehold, however, it forms a (relatively stable) T-T mismatch with the incumbent. The second invader is identical, but forms an additional, proximal C-C mismatch in the displacement domain. Figure 2.14 allows for a comparison to invaders whose toeholds are fully complementary to the target.

The displacement reactions of the two invaders with a fully complementary displacement domain proceed at very similar rates (Figure 2.14, top, and Table 2.5). As described above, k_{off} is much smaller than k_{disp}, so the overall reaction is similarly fast despite the fact that one of them has to initiate the reaction using a mismatched toehold. The introduction of the C-C mismatch slows down the displacement reaction, even if the

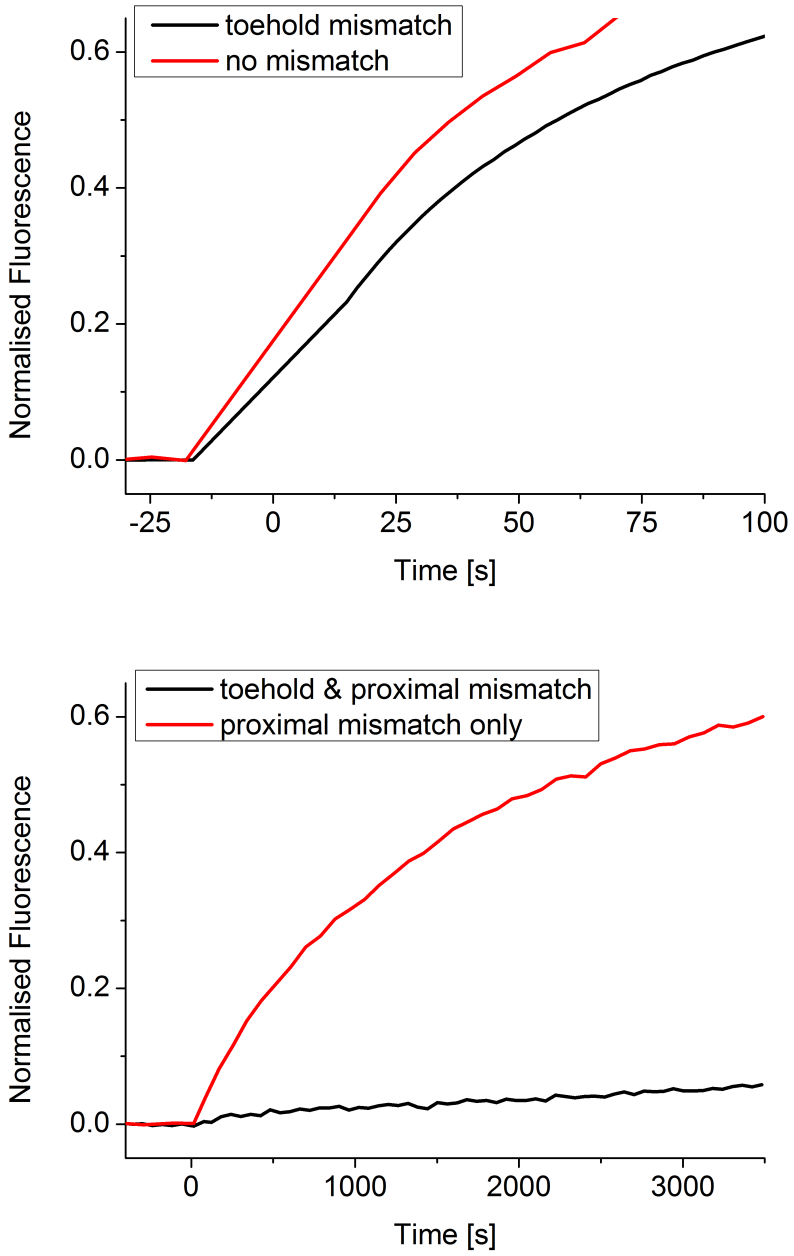

Figure 2.14: Four different invading strands displace fluorescently labelled incumbent from a target strand labelled with a quencher. **Top:** Little difference is visible in the reactions of an invader that is fully complementary and one which forms a mismatch in the toehold region. **Bottom:** After the introduction of a proximal C-C mismatch, the invader forming a mismatch in the toehold is impeded kinetically much stronger than the one with the complementary toehold. Traces were normalised by dividing by the saturation signal after an excess of non-mismatched, 10 nt toehold invader had been added to the sample.

	No MM	TH MM	Prox MM	Prox, TH MM
k [$M^{-1}s^{-1}$]	$3.5 \cdot 10^6$	$2.0 \cdot 10^6$	$6.9 \cdot 10^4$	$1.3 \cdot 10^3$

Table 2.5: Second-order reaction rate constants associated with the displacement of incumbent from target by four different invading strands. The 'No MM' strand is fully complementary to the target. 'TH MM' is the same, except for a T-T mismatch it forms in the toehold region. 'Prox MM' forms a proximal C-C mismatch in the displacement domain, and 'Prox, TH MM' forms both the T-T mismatch in the toehold domain and the proximal C-C mismatch in the displacement domain. Values for k were derived from measured data using a fit based on equation 2.1.

toehold is fully complementary. For the invader with the mismatched toehold, however, the effect is much more drastic. Although the mismatched toehold only has a relatively small defect, which thermodynamically still allows for full displacement of the incumbent [152], a strong kinetic discrimination becomes visible.

The discrimination can be quantified in both cases with a factor $d = {}^{k^c}/_{k^s}$, where k^c is the effective displacement rate of the invader with the correct toehold and k^s that of the invader with the spurious one. With the steady-state expression of $k = \frac{k_{disp}k_{on}}{k_{off}+k_{disp}}$ (section 1.2), and under the assumption that both k_{on} and k_{disp} are equal for both invaders, the discrimination factor becomes $d = \frac{k^s_{off}+k_{disp}}{k^c_{off}+k_{disp}}$. If k_{disp} is high in comparison to the off-rates, $d \to 1$, and thus no discrimination occurs, as expected. If k_{disp} is negligible compared to the off-rates, $d \to {}^{k^s_{off}}/_{k^c_{off}} = e^{\Delta\Delta G^\circ_{s-c}/RT}$, where $\Delta\Delta G^\circ_{s-c}$ is the difference in change in free energy associated with mismatched invader displacing the incumbent and the equivalent for the complementary invader. For the systems presented here, this yields an upper bound of $d_{max} = 60$, using a NUPACK-estimated $\Delta\Delta G^\circ_{s-c} = 2.4\,\text{kcal/mol}$ [152].

Experimentally, the discrimination factors were determined to be $d_{comp} = 1.8$ for the invaders with a fully complementary displacement domain and $d_{C-C} = 53$ for the invaders with the proximal C-C mismatch. This

shows that the introduction of a single proximal mismatch allows for greatly increased kinetic discrimination. Similar to the introduction of an inert spacer between toehold and displacement domain serving as a kinetic barrier [38], this toehold probing may thus serve to expand the sequence space available when designing toeholds in complex, TMSD-based CRNs.

2.7 IMPACT ON EXISTING NUCLEIC ACID PROBES

Several strand displacement-based systems have been designed in the past to screen nucleic acids for a specific sequence. They aim to detect oligonucleotides that are fully complementary to a probe strand, and reject spurious sequences on the basis of the defects these are forced to form when attempting to trigger the probe. If these probes rely exclusively on the kinetic impediment of such a defect in order to 'ignore' spurious inputs, such as the probes presented by Li and colleagues [66] as well as Subramanian and colleagues [126] (Figure 2.15), the findings presented in this chapter suggest that the probes' performance will be poor if such defects occur close to the distal end of the displacement domain. For a wide range of spurious invaders, these probes will thus create false positives[6].

The strand-displacement probe design by Zhang and colleagues [154] offers a viable alternative to these probes, by operating on thermodynamic, rather than kinetic discrimination. This is achieved by providing the incumbent strand with an exclusive binding domain on the target strand, which acts as a thermodynamic counterbalance to the toehold the invader has at its disposal to initiate displacement. Remarkable

6 Note that the terminology used in the studies referenced here usually describe what is called invading strand in this thesis as the 'target' of the probe. In order to prevent confusion, the nomenclature of the previous sections is maintained here.

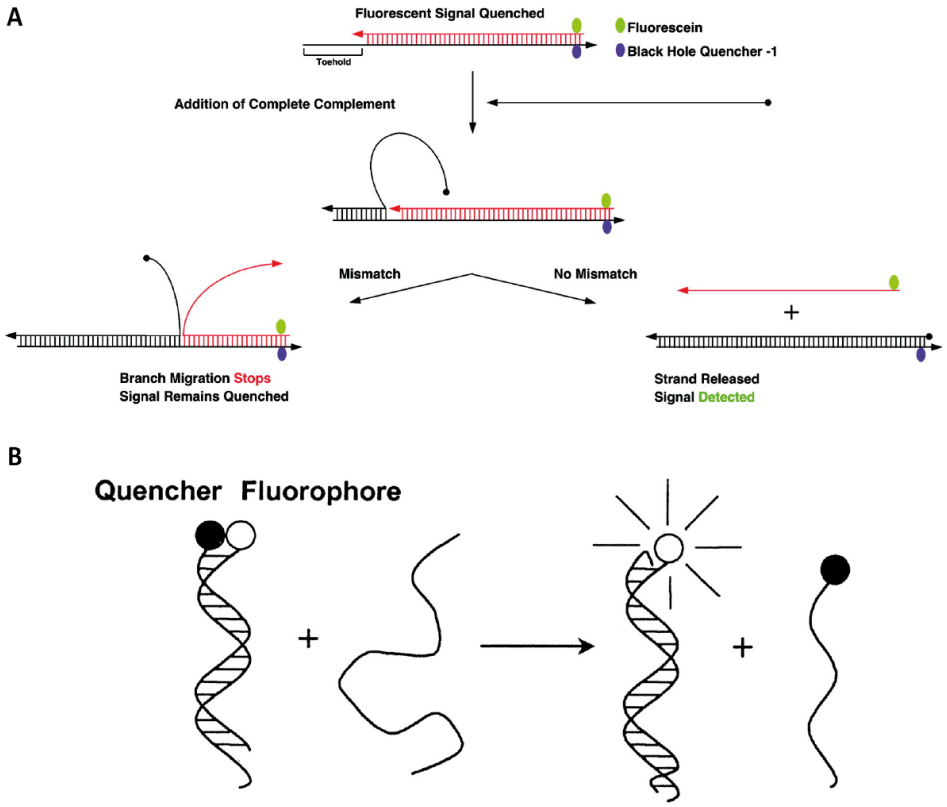

Figure 2.15: Two previously published nucleic acid probe designs based on TMSD. **A:** Probe design by Subramanian and colleagues. The invading strand (black) binds to the toehold and initiates branch migration. If the invader would form a mismatch with the target, branch migration is assumed to stop, leading to no displacement of incumbent strand (red). In the case of full complementarity, the incumbent is displaced and a change in fluorescence observed (adapted from [126], with permission from the American Chemical Society, ©2011). **B:** An earlier, similar concept by Li and co-workers relying on the same basic principles as the Subramanian probe (adapted from [66], with permission from Oxford University Press, ©2002). Neither of these designs were presented with a systematic analysis of probe reliability in the case of distal mismatches.

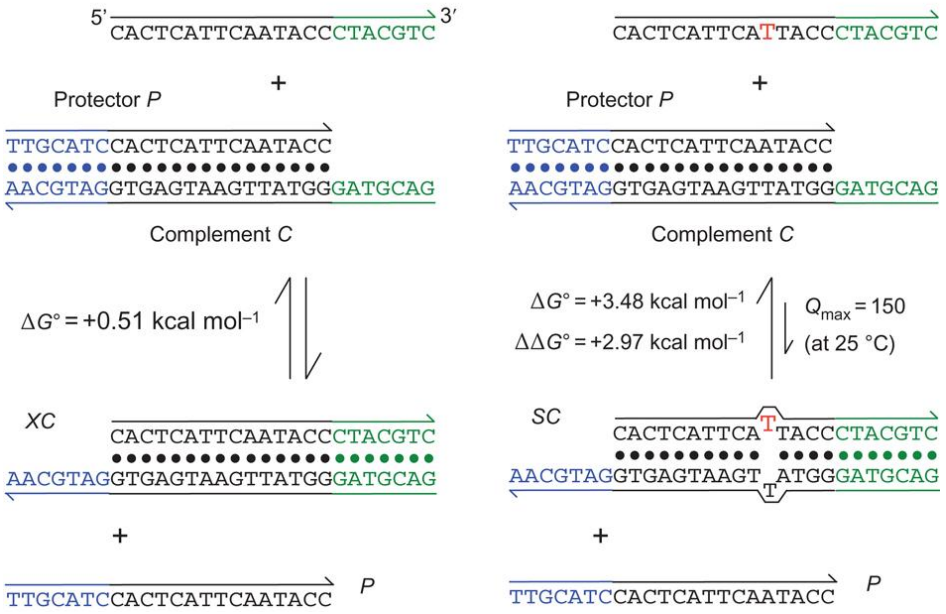

Figure 2.16: A nucleic acid probe design by Zhang and colleagues, which relies on thermodynamic, rather than kinetic discrimination between a fully complementary invader (left) and a spurious one (right). An exclusive binding domain (blue) is available to the incumbent ('Protector') on the target ('Complement'). This makes incumbent displacement by an invading strand thermodynamically less favourable, which in turn allows for greater discrimination against spurious invaders in equilibrium, independent of mismatch location (adapted from [154], with permission from Macmillan Publishers Ltd, ©2012).

discrimination factors are achieved this way, practically independent of defect location. A problem may arise, however, when this equilibrium-discrimination probe is not operated in isolation, but in biological contexts. The probe relies on invaders being incapable of hybridising any part of the incumbent's exclusive domain, which can not be guaranteed in samples with such complex and variable sequence backgrounds.

One way to achieve a truly exclusive incumbent binding domain is the introduction of synthetic bases (such as those presented in reference [62]), which would stop biological sequences from forming base pairs in pre-definable locations. One example is the pairing of isomeric G and C, which do not bind to natural C and G, respectively. These Iso-G/Iso-C bases can be readily integrated in synthetic DNA oligos [129], and would

thus likely be capable of providing universal, invader-independent mismatches. Note that this would also solve the rather obvious problem of the toehold-probing mechanism introduced in the previous section, in which otherwise invaders with a spurious toehold which do not form the proximal mismatch would quickly displace the incumbent. Iso-G/Iso-C bases have, for example, already been shown to improve the sensitivity of branched DNA hybridisation assays, a different type of nucleic acid probe [20], which is restricted to *in vitro* probing.

2.8 CONCLUSIONS & OUTLOOK

This chapter explored how mismatches can serve as controllable free-energy barriers for TMSD reactions. One of the most striking results is the dependence of the reaction rate constant on the position of such a barrier in the displacement domain. Generally, it was observed that a mismatch close the toehold can reduce the rate constant by several orders of magnitude, whereas an identical mismatch close to the distal end of the displacement domain has little or no effect on reaction kinetics. This core observation was made both with a reaction scheme in which reactants were labelled directly with fluorophores and quenchers, as well as a further scheme in which these labels were removed from the main reactants by the use of a distinct reporter complex. The findings were robust across different methods of data analysis.

To explain the observed behaviour, a deeper insight into the mechanistic details of TMSD processes was required. Through the use of oxDNA, a coarse-grained model of DNA, the displacement reactions in which a mismatch is created could be explored with single-base pair resolution free-energy landscapes. These landscapes served as the basis of a qualitative discussion of the observed phenomena. Mismatches in the displacement

domain can lead to a considerable bias against the invading strand in the random walk of branch migration. This bias has a stronger effect when it is implemented closer to the toehold. Moving the barrier further into the displacement domain usually leads to a lesser impediment of the overall reaction, because the probability of the invading strand to be in a state where it is prone to spontaneous, complete detachment from the target is reduced.

If the mismatch is very close to the distal end of the displacement domain, another reaction pathway begins to dominate the overall process. Within this alternative reaction pathway, the mismatch is not enclosed before displacement of incumbent. Depending on the exact mismatch location, this can lead to the mismatch not providing a measurable impediment to displacement relative to a non-mismatched system. The rate dependence on mismatch location was found to be a purely kinetic phenomenon, with systems of equal changes in free energy associated with displacement showing reaction rates several orders of magnitude apart.

The discussion of the results was supported by an additional experimental series using an elongated displacement domain, as well as quantitative simulation results derived using oxDNA. The remarkable accuracy with which oxDNA simulations replicated experimental findings demonstrate that this model is capable of predicting complex nucleic-acid interactions, suggesting its use in rational design and optimisation of DNA nanodevices.

In addition to these core results, it was shown that:

- an energetically less unfavourable type of mismatch has a lesser impact on displacement rate;
- for sufficiently unstable toeholds, i.e. with a length $\lesssim 8\,\mathrm{nt}$, the reaction rates of TMSD processes involving a mismatch scale linearly

with reactant concentrations, as expected for second-order reactions;

- for invaders with toeholds stable over the experimental time scale, a subtle effect emerges which makes the reaction rates of invaders forming a mismatch very close the toehold faster than those forming the mismatch in the centre of the displacement domain;
- TMSD processes involving a mismatch can be integrated seamlessly into existing nucleic-acid reaction network designs.

These results suggest that the effects observed might serve as the basis of a novel mechanism to control the kinetics of TMSD reactions, which serve as a foundation of virtually all dynamic DNA nanodevices and nucleic-acid CRNs. In contrast to most previous mechanisms, changing the position of a mismatch to fix the reaction rate does not require changes in thermodynamic reactant or product stabilities or external reaction conditions. To my knowledge, it would be the only kinetic control mechanism that specifically makes use of the sequential process underlying strand displacement.

In the work presented here, the TMSD free-energy landscapes were only modified by two types of mismatches between the invading strand and the target. Further studies are required to reveal how other types of mismatches, or generally other types of defects, shape these landscapes. Such defects could also be used in the incumbent-target duplex to bias the free-energy landscape in favour of an invading strand.

In addition to the qualitative discussion of the phenomena observed, a solid quantitative description would be desirable. Although oxDNA can accurately predict reaction rate constants, it does not allow for the rapid testing of a wide variety of reactions, as is often required in the design of *de novo* devices and networks. Instead, rate constant guidelines as

those derived by Zhang and Winfree in the context of toehold exchange [156] would be practical. In combination with experimental data on other types of mismatches and defects, this may allow for predictable fine-tuning of rate constants in nucleic acid reaction networks.

This chapter further demonstrated the potential for applications of strand displacement involving mismatches. A device in which the competition of two simultaneous reactions was influenced with an asymmetrically positioned mismatch served as a proof-of-principle for independent control of reaction kinetics and thermodynamics. Through this, it was possible to exert considerable control over the reaction pathways in such a fashion that the overall reaction transiently occupied a state to a much greater extent than would have been possible with, for example, kinetic control via toehold binding energy alone.

The shape of the time dependence of this temporary state is strikingly similar to other systems which provide short pulses of certain quantities, such as certain cellular transcription network motifs [3]. Another, non-biological device showing similar behaviour is the current response of an overdamped linear resistor-inductor-capacitor (RLC) circuit. Such a circuit can provide pulsed discharges, for example in devices as common as spark plugs for combustion engines [53]. A somewhat comparable role is conceivable for the competitive displacement device presented here, as it could serve as a transient provider of a pre-defined DNA duplex in a nucleic acid CRN. Inverting this functionality, one could also consider this device to be capable of transient storage of a specified oligonucleotide reactant.

What makes these abilities particularly noteworthy is that they do not require any external control. The kinetics of uptake and release are controlled through predetermined parameters of the device, such as mismatch location and mismatch type as well as the binding strengths of

the toeholds and the length of the displacement domain. Similar pre-programmable, complex behaviour in a DNA reaction could likely be achieved with a different approach [108, 121], but would likely require the presence of far more individual reactant species, which would increase the risk of unwanted side reactions.

Section 2.2 showed how a mismatch placed in close proximity to the toehold can significantly slow down a TMSD reaction. In the simplistic three-state model introduced above, it does so by lowering k_{disp}, the rate constant describing displacement after the toehold is bound. Section 2.6 showed how this allows for invaders forming long toeholds to be kinetically discriminated by the strength of the toehold. As a proof-of-principle, invading strands were introduced which form a single defect in their 8 nt toehold. The displacement reaction rate of these invaders was significantly impeded compared to an invader forming a fully complementary toehold only when a proximal mismatch was present. In the absence of such a mismatch, the invader forming a spurious toehold displaced the incumbent strand at a similar rate as its fully complementary equivalent. Apart from its use in complex CRNs, the toehold probing mechanism may either serve as the basis of a nucleic acid probe on its own, or might be included into existing designs, improving their ability to detect a desired sequence with high discrimination.

In its mechanism, the mismatch-assisted toehold probing presented here is somewhat reminiscent of kinetic proofreading, a process which allows for high discrimination between thermodynamically similar pathways in natural enzyme systems, most prominently featured in protein synthesis [49, 87]. Such systems, however, are capable of reaching discrimination factors far greater than the thermodynamically fixed $d_{max} = e^{\Delta\Delta G^\circ/RT}$. This is due to the use of an external source of energy in the form of ATP, which allows for repeated checking of a substrate before the re-

action proceeds, and thus exponentially increased discrimination factors. It would be interesting, from a conceptual point of view, to implement a system capable of such specificity with nucleic acids alone. It has already been shown that energy capable of fuelling cyclic processes can be stored in DNA hairpins [43, 88] and similar metastable structures [133]. These might be used for such a purpose, to provide synthetic nucleic acid sensors with capabilities comparable to their natural counterparts.

The last section of this chapter put the findings related to mismatches into a more general context with strand-displacement nucleic-acid probes. Particularly for spurious invaders only forming distal defects, it is questionable whether kinetically discriminating probes are capable of reliable sensing performance. Experiments regarding this potential weakness have not been presented previously. Due to these findings, thermodynamically discriminating probes appear more robust, depending on the application. In the context of such probes, the introduction of non-natural bases may be a sensible extension of design capabilities and robustness. Future studies into how such bases influence kinetic behaviour of displacement reactions otherwise using the native DNA alphabet could be based on the methodology used here.

Further applications of the results described in this chapter are also conceivable, particularly for systems in which mismatches have previously been included for other purposes, such as modifying an enzyme recognition domain in a DNA nanomachine [143] or reducing leak reactions in DNA amplification circuits [55]. Overall, departing from full base complementarity, previously somewhat of a rarely questioned dogma, may offer exciting new design paradigms in DNA nanotechnology and related fields.

3

COMBINATORIAL, CASCADABLE STRAND DISPLACEMENT

3.1 INTRODUCTION

This chapter explores an extension to regular TMSD, in which the incumbent is displaced from a duplex cooperatively by two instead of one invading strand. This general scheme, which has been introduced before (see the work by Chen [17] and Figure 1.9C), requires the two invaders to hybridise to each other before they are able to displace incumbent efficiently. Genot and co-workers described how this can be used to multiply a Boolean 2×2 matrix with a Boolean 2×1 vector [37]. This work is revisited in more detail in this introduction.

Section 3.2 describes how the system design by Genot et al. can be altered to allow for multiplication of binary matrices of higher dimensionality. Section 3.3 shows that this altered reaction scheme is cascadable. Through this cascadability, chain multiplication of binary matrices becomes possible, allowing for DNA-reaction based computations on simple, directed graphs. An extension from two-invader displacement to multi-invader displacement is also explored (section 3.4).

Contributions by colleagues and collaborators

The results of this chapter are currently being prepared for publication by myself as lead author, Chris Thachuk (formerly Department of Computer Science, University of Oxford, now California Institute of Technology),

Jonathan Bath and Andrew J. Turberfield. In addition to work directly referenced, the following contributions to this chapter have been made by colleagues and collaborators:

- Jonathan Bath suggested the use of a quantitative polymerase chain reaction (PCR) machine for fast assaying of the systems presented;
- the idea to use combinatorial strand displacement for calculations on graphs was developed in discussions with Chris Thachuk.

3.1.1 *Associative toeholds and matrix multiplication*

Consider two matrices: M has dimensions $i \times j$, with i denoting the number of rows and j the number of columns. N has dimensions $k \times l$. In order for the dot product $P = M \cdot N$ to exist, j must equal k. P has dimensions $i \times l$. For each element of P,

$$P_{ab} = \sum_{z=1}^{j} M_{az} N_{zb},$$

where X_{ab} represents the element in row a and column b.

For Boolean matrix multiplication, in which all elements of M, N and P are either 0 or 1,

$$P_{ab} = \begin{cases} 1 & \text{if } \sum_{z=1}^{j} M_{az} N_{zb} > 0 \\ 0 & \text{if } \sum_{z=1}^{j} M_{az} N_{zb} = 0 \end{cases},$$

or, in a logic symbol representation

$$P_{ab} = \bigvee_{z=1}^{j} M_{az} \wedge N_{zb}. \tag{3.1}$$

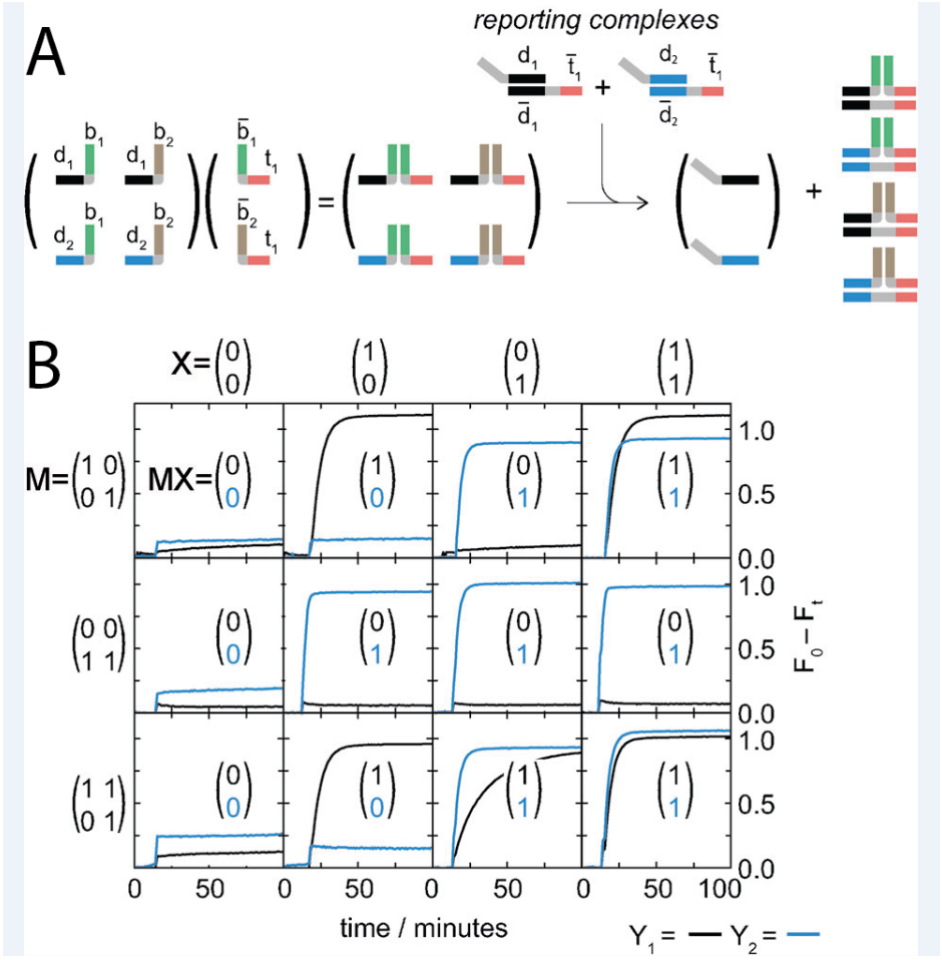

Figure 3.1: Genot and co-workers' scheme for Boolean matrix multiplication using associative toehold displacement. **A:** Reaction scheme of a multiplication of a 2×2 matrix ($M = 1_{2\times 2}$) with a 2×1 matrix ($X = 1_{2\times 1}$). In M, displacement domains d_1 and d_2 correspond to the row index, and linking domains b_1 and b_2 correspond to the column index. For X, the linking domains $\overline{b_1}$ and $\overline{b_2}$ correspond to the row index, with $\overline{b_1}$ denoting complementarity with b_1, and $\overline{b_2}$ denoting complementarity with b_2. Due to this complementarity, the complexes $m_{11}x_{11}$, $m_{12}x_{21}$, $m_{21}x_{11}$ and $m_{22}x_{21}$ form. The former two trigger a reporting complex r_{11} (left), the latter two trigger r_{21} (right), thus yielding the Boolean product $P = 1_{2\times 1}$. **B:** Experimental proof-of-concept. Different combinations of possible M and X matrices were multiplied with each other. Zeros in either matrix were created by omitting the corresponding strand species. Spectrally separated fluorophore-quencher interaction on either reporting complex was recorded to measure displacement of reporter complex incumbent. Adapted from [37], with permission from John Wiley & Sons, ©2013.

Genot and colleagues described how a DNA strand displacement mechanism very similar to Chen's associative toehold activation can be used to perform above calculation [37]. This mechanism is briefly explained here and in Figure 3.1. M and N are Boolean matrices. Each matrix is represented by a mixture of oligonucleotides. One strand species represents exactly one matrix element, e.g. m_{23} represents matrix element M_{23}. If a matrix element is equal to 1, the corresponding strand species is present in the mixture; if the element is equal to 0, it is absent. Each row and each column in M and N is associated with one particular sequence domain. Strand species thus require two domains to identify which matrix element they represent. The sequences of the row domains for N, 1 through $k = j$, are complementary to the column domains of M which have the same index. As a consequence, if the mixture of strand species representing M is added to that of N, complexes of the form $m_{az}n_{zb}$ self-assemble.

Using the associative toehold mechanism, such complexes are capable of triggering a displacement reaction with a read-out duplex r_{ab}. This duplex is designed in such a fashion that the two single-stranded overhangs of $m_{az}n_{zb}$, the domain associated with row a of matrix M, and the domain associated with column b of matrix N, are complementary to displacement domain and toehold of r_{ab}, respectively. Neither the strand species representing M, nor those representing N can perform this displacement reaction efficiently on their own. M-strands lack a toehold, making displacement slow. N-strands are only able to hybridise the toehold domain, but can not displace the incumbent of r_{ab}. This type of displacement reaction can thus be regarded to represend an AND-like logic, as it requires two compatible inputs, m_{az} and n_{zb}, to be present in order to yield an output. It does not, however, distinguish between complexes that differ in linking domain: as long as the linking domain is complementary and forms a sufficiently stable duplex, displacement

will occur on a time scale similar to that of regular TMSD. If we regard the occurrence of displacement on duplex r_{ab} as equivalent to $P_{ab} = 1$, and no (or negligible) displacement as equivalent to $P_{ab} = 0$, this scheme allows for a nucleic-acid representation of equation 3.1.

There are several limitations to this scheme, which make it suboptimal for application. The most obvious is the small dimensions of the matrices multiplied successfully, i.e. 2×2 and 2×1. One reason for this limitation becomes apparent in Figure 3.1B: Despite the small matrix dimensions, leak reactions lead to a signal intensity of up to 25 % of saturation fluorescence, where 0 was expected. This, in turn, is likely due to certain design features of the system. A contributing factor might have been that the displacement domains d_1 and d_2 on the reporter complex incumbents were truncated by one nucleotide in comparison to d_1 and d_2 of the strand species representing M, giving the latter a 1 nt toehold. This allows M-strands to spuriously displace reporter-incumbent ~10× faster compared to blunt-end strand displacement [156], even if no appropriate N-strands are present.

Another reason for this limitation is the way displacement is reported. The two reporter complexes have to rely on spectrally separate fluorophores. For each additional element of the product matrix, a further fluorophore with minimal emission overlap with the other fluorophores would have to be found. Common fluorophores still emit with a quarter of their peak intensity at wavelengths 30 nm to 50 nm from their emission peak, severely limiting the maximum number of reporters with only negligible crosstalk. The optical, rather than chemical reporting also leaves the question unanswered whether such a multiplication can be integrated into a larger system as a component, or whether it is only capable of stand-alone operation.

3.2 DISPLACEMENT MECHANISM AND DESIGN

3.2.1 *Design principles*

In the combinatorial displacement mechanism adopted here, systems were designed in a way so that neither of the invading strands is thermodynamically equivalent or superior to the read-out incumbent in binding to the read-out target. As the schematic in Figure 3.2A shows, the implemented design clearly fulfills this criterion: the read-out incumbent hybridises 23 bases of read-out target: 4 in the proximal invader domain (red), 12 in the distal invader domain (blue) and 7 exclusive bases, which serve as a toehold for a reporter complex upon displacement of incumbent (black). As either invader can only form 12 base pairs each with the target, the incumbent is at a considerable thermodynamic advantage, unless the invaders are able to hybridise via a stable linking domain (12 nt, green), and form an invader complex. The overall reaction yielding in increase in fluorescent signal is then driven by the difference in free energy deriving from a greater number of base pairs being formed: displacement of incumbent by an invader complex corresponds to a net gain of 1 base pair, subsequent triggering of the reporter complex to a net gain of 6 base pairs.

Genot and colleagues observed that introduction of inert TT spacers (grey) around the junction accelerates displacement kinetics [37]. This observation was confirmed in preliminary experiments for the systems used here. The incumbent also includes a TT spacer, and thus forms a double T-T mismatch with the target at the junction site. Preliminary experiments showed that for a fully complementary incumbent, displacement was strongly hindered, as expected considering the findings in chapter 2.

The design implemented here contains some further features: Instead of a quasi-irreversible strand displacement reaction, toehold exchange [156] is employed. As discussed in chapter 2 in the context of nucleic acid probes [154], this allows for stronger discrimination against spurious invader complexes. This toehold exchange further allows a design of displacement reactions that are completely independent of the reporter complex (i.e., toehold- and displacement domains share no sequence overlap with the reporter), allowing the use of the same complex to report the triggering of different read-out duplexes, which will be used in subsection 3.2.2.

As shown in Figure 3.2B, this design, at least in isolation, works well: only the correct combination of invading strands can trigger a significant displacement reaction.

3.2.2 *Multiplication of Boolean* 3×3 *matrices*

From reaction scheme to reaction network

In the introduction to this chapter, two major barriers to scaling up Genot's matrix multiplication scheme were identified. The first were the occasionally significant leak reactions, the second the use of a different fluorophore for each element of the product matrix. A redesign of the displacement mechanism can make these types of reactions virtually leak-free (at least in isolation), as demonstrated in Figure 3.2B.

The second barrier can be resolved by using non-fluorescent read-out duplexes, which can trigger a universal reporter. The lack of spectral separation between the fluorescent signals corresponding with each product matrix element requires a spatial separation of reactions: Instead of mixing the input matrix elements with all read-out duplexes in one reaction vessel, each product matrix element is associated with its own reaction

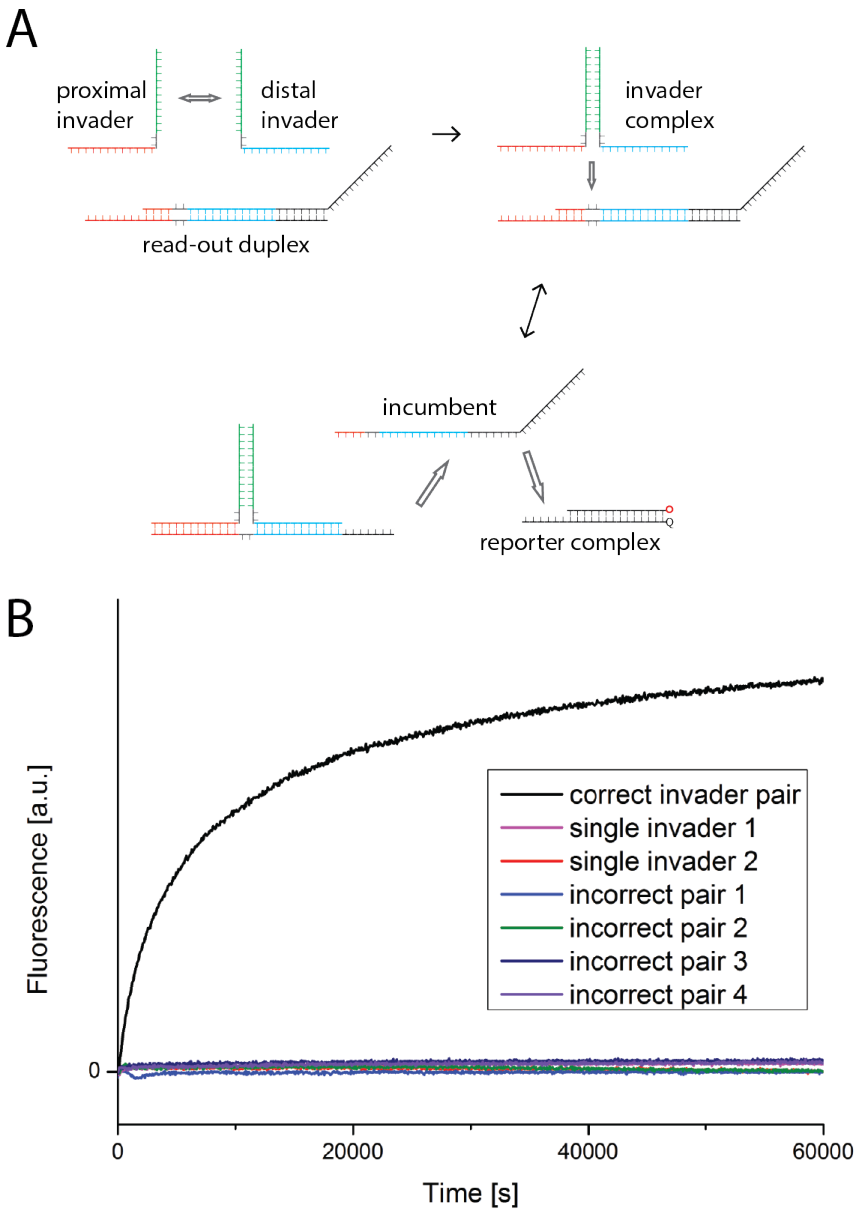

Figure 3.2: Design and proof-of-principle experiments for pairwise, combinatorial strand displacement used in this work. **A:** Reaction scheme. Two invading strands can hybridise via a 12 nt linking domain (green), forming an invader complex. This complex can interact with a read-out duplex, and, using an 8 nt toehold, form a three-arm junction with the read-out target and release an incumbent strand. The displaced incumbent, in turn, is capable of triggering a fluorescently labelled reporter complex via standard TMSD, using a toehold that was previously sequestered. **B:** The response of the reporter complex is measured upon addition of a read-out duplex and different combinations of invading strands at $t \approx 0$. In the correct invader pair, both a proximal invader and a distal invader are present, and they have a complementary linking domain, as depicted in A. On their own, neither invader is capable of triggering a response (single invader 1 and 2). If the linking domain is not complementary (incorrect pair 1), the distal invader carries a spurious displacing domain (incorrect pair 2), the proximal invader carries a spurious toehold domain (incorrect pair 3) or both the toehold domain and the displacing domain are spurious (incorrect pair 4), no displacement occurs.

vessel, in which, in addition to the input matrix elements and the reporter complex, only the read-out duplex associated with that product matrix element is present. Thus, for a product matrix of dimensions $a \times b$, ab individual reactions are measured, each with the same mix of input strands and reporter complex, but a unique index-specific read-out duplex.

In most fluorimeters, this approach would lead to an efficiency problem for bigger matrices, as most of these devices can only measure between 1 and 4 samples at the same time. Quantitative PCR machines, however, usually contain 96-well plate readers, and, with very few programming adjustments, can be converted to 96-sample flourimeters (the details of this method are described in Appendix A.1.3). Such a device was used here to measure the DNA-based calculation of several Boolean dot products of two 3×3 matrices, employing the reaction scheme described above.

To scale this reaction scheme to 3×3 matrices, nine different distal invaders represent the first matrix (with strand presence representing '1', and absence representing '0'), nine proximal invaders the second, and nine read-out duplexes each element in the product matrix. To fulfil the criteria for matrix multiplication outlined in the introduction of this chapter, distal invaders representing elements in the same row have an identical displacing domain, those representing elements in the same column have an identical linking domain. Proximal invaders representing elements of the same row have an identical linking domain, those representing elements in the same column have an identical displacing domain. The displacing domains of the invaders and the sequence of the universal reporter complex fully determine the sequence of the read-out duplexes. Thus, only nine unique, 12 nt sequences are required to create this reaction network, in addition to the reporter sequence.

These sequences were designed with the goal of creating reaction network components with as little spurious behaviour as possible. For this purpose an algorithm implemented in Matlab was used. This algorithm creates random, 12 nt sequences, which only contain bases A, T and C, and accepts or rejects them based on a list of absolute and comparative criteria. The absolute criteria were chosen to be:

1. either A or T may not occur more than four times in a row;
2. C may not occur more than three times in a row;
3. the number of Cs must be between four and eight.

Based on the above criteria, one initial sequence was found. Each further sequence was additionally scrutinised by the following comparative criteria:

1. the sequence may not share an identical subsequence of more than five nucleotides with any of the previously accepted sequences;
2. the sequence may only have an identical base in up to six positions, when compared to any of the previously accepted sequences.

These criteria are similar to those used, for example, by Qian and Winfree in their work on complex circuits of DNA logic gates [98, 97, 99], but, considering the smaller number of domains required here, were chosen to be stricter.

The restriction to a three-letter alphabet aims to limit secondary-structure formation within the individual sequences [82]. Limiting the length of subsequences consisting of only one type of base is assumed to reduce suboptimal hybridisation to a complementary sequence. Control of the C-content ensures that hybridisation free energies are similar for all sequences. The comparative criteria aim to limit crosstalk between sequences

which are not intended to interact. For each 'C-sequence' thus created, a complementary 'G-sequence' was found as well.

Nine of these sequences and their complements were than mapped onto the reaction network component domains. The proximal invaders each consist of two G-sequences, joined by a TT spacer. The distal invaders each consist of one G-sequence joined by a TT spacer with one C-sequence; the latter representing the linking domain. This allows for the read-out duplexes and the reporter to be designed in a fashion that the incumbent contains no Cs. Otherwise, displaced incumbents may form a variety of stable secondary structures, which would likely limit uniformity in reporter response to the different incumbents. The specific sequences assigned to the displacing and linking domains of the invaders were selected to avoid two potential flaws, linking domain promiscuity and displacing domain interaction. The former describes unwanted hybridisation interaction of linking domains, and would allow the formation of invader complexes capable of triggering incorrect read-out duplexes. The latter describes substantial base pairing between the displacement domains of any of the 27 possible invading complexes, which would lead to preferential formation of certain invading complexes over others, as well as impeded displacement of the read-out incumbent. Different sequence combinations were tested using NUPACK [152], until a satisfactory set of invading strands was found (Appendix Table A.10).

Implementation and results

Eight multiplications of 3×3 matrices were tested. The first of these is depicted in Figure 3.3, which reiterates the experimental mechanism and implementation. The first of the two matrices is represented with distal invaders, the second with proximal invaders (Figure 3.3B).

Prior to fluorimetry, all invaders were mixed at an equimolar concentration of 4 μM and annealed, to form invader complexes. 9 wells in a 96-well plate were filled with a universal reporter complex (15 nM) and an index-specific read-out duplex (100 nM). For these read-out duplexes, the domain interacting with distal invader is invariant in each row, and the domain interacting with proximal invader is invariant in each column (Figure 3.3C). The read-out duplexes were prepared with a target-to-incumbent ratio of 1.1 : 1 and annealed, to prevent presence of free incumbent and misformed duplexes.

Upon addition of input mix to all wells (at 100 nM), an invading complex consisting of distal invader d_{ab} and proximal invader p_{bc} ($a, b, c \in [1, 2, 3]$) can displace the incumbent from read-out duplex r_{ac}, and trigger a fluorescent response in the well with index (a, c). As a result, a strongly increasing signal can be expected in all wells that have an index that corresponds to a '1' in the product matrix (Figure 3.3D). In this exemplary measurement this is indeed the case: the one well which corresponds to a '0' in the product matrix shows a distinctly weaker response than all others.

In order to calculate the products of different input matrices, only the composition of the input mix has to be altered; read-out duplexes and reporter complex stay the same. The results of seven further multiplications are depicted in Figures 3.4 and 3.5. In each case, the system works as desired, showing a fast growth in fluorescent signal for those matrix elements expected to yield ones, and a distinctly slower growth for zeros. These examples showcase some important features:

- squares of matrices can be calculated (A).
- multiplication with this system is non-commutative (compare B and C);

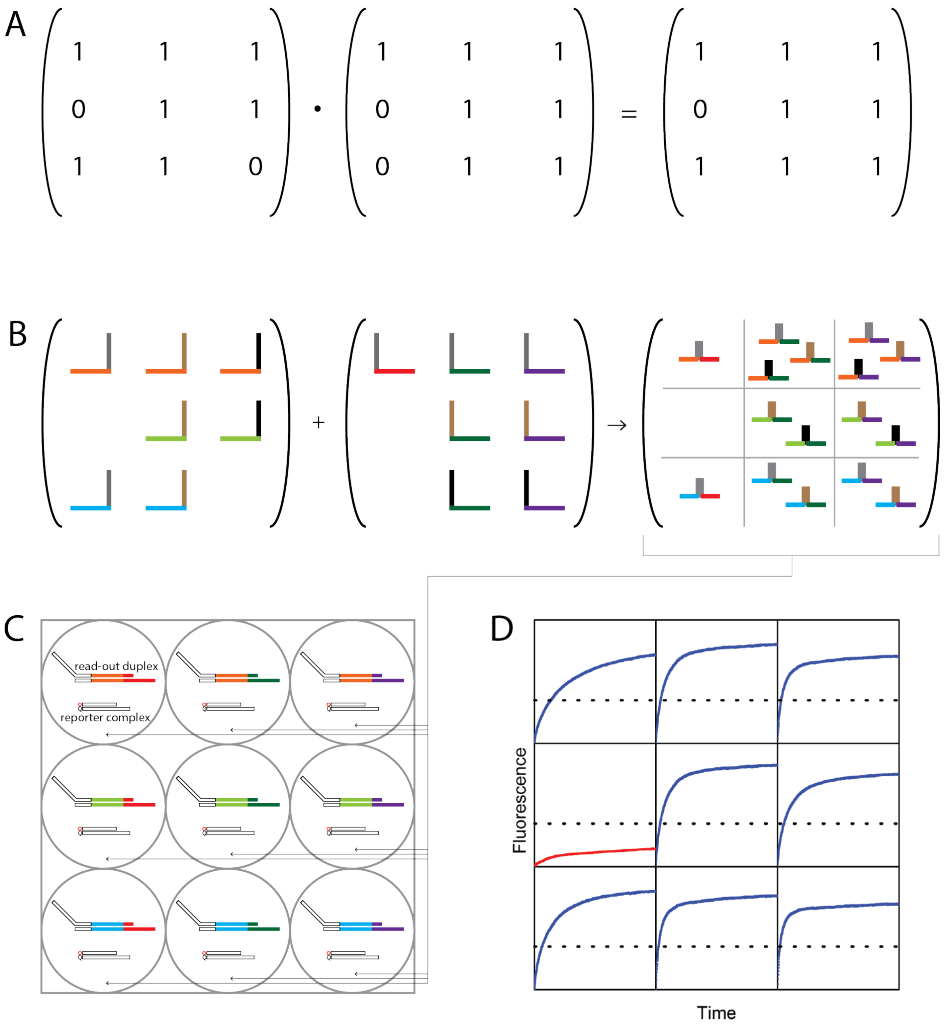

Figure 3.3: Example of a 3×3 Boolean matrix multiplication. **A:** Two matrices are multiplied with each other, according to equation 3.1. **B:** DNA representation of these matrices. Upon mixing of the DNA inputs, 16 stable invading complexes form. These are shown grouped according to the effective displacing domain created. **C:** Nine wells of a plate each contain one index-specific read-out duplex as well as a universal reporter complex. The mix of all invading complexes is added to each of the nine wells, where they can trigger a fluorescent change, depending on whether invading complexes formed that are appropriate for the read-out duplex present. **D:** Fluorescence response traces for the particular matrix multiplication depicted. Fluorescence in each subplot ranges from 0, defined as the fluorescence at the first measured point after sample mixing, to the equivalent of complete consumption of reporter complex, i.e. 15 nM displaced reporter incumbent. Time ranges from the first point measured after sample mixing to 1400 min. Traces expected to correspond to a '1' are blue, the trace expected to correspond to a '0' is red. A cut-off fluorescence can be identified (dotted line) which allows for the unambiguous binary classification of each response (subsection 3.3.3).

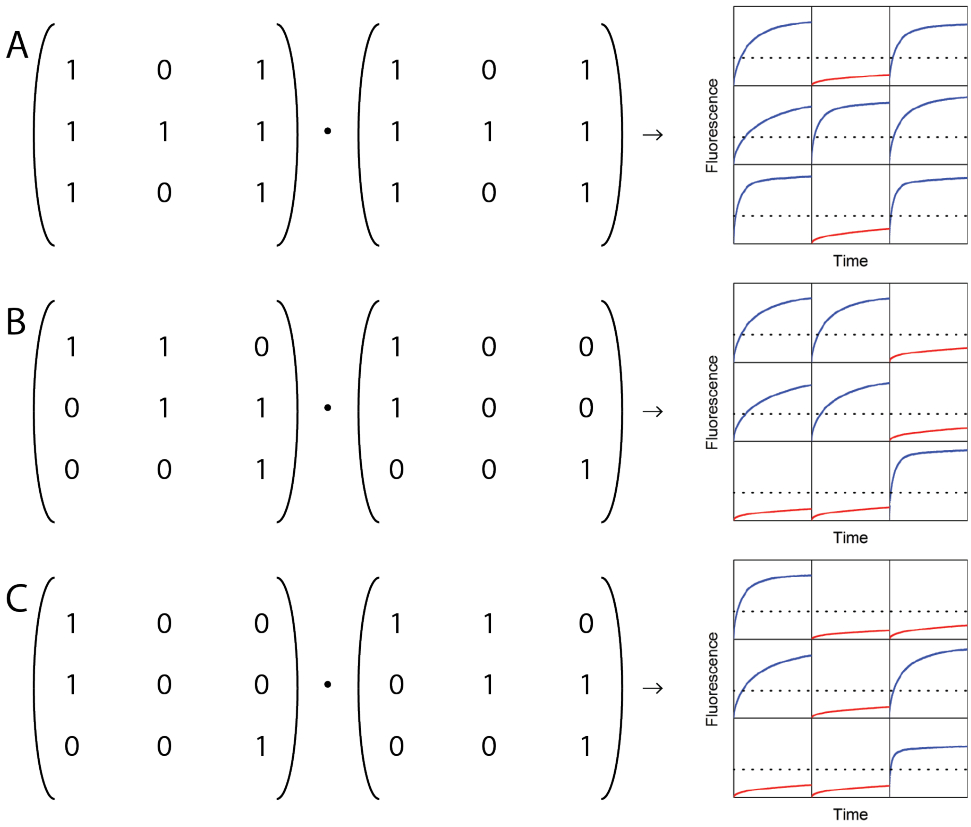

Figure 3.4: Fluorescence response traces according to different 3×3 Boolean matrix multiplications (A, B, C). Traces are coloured and normalised as in Figure 3.3. Further examples are depicted in Figure 3.5.

- some matrices have an inverse (D);
- an identity matrix can be implemented (E);
- permutation matrices can be implemented, which exchange columns when multiplied from the left (F) and rows when multiplied from the right (G).

It is worth noting that not all traces representing a '1' show uniform kinetic behaviour. The main reason for this is assumed to be the variation of invader complex concentration depending on input matrix composition. As an example, if the Boolean multiplication

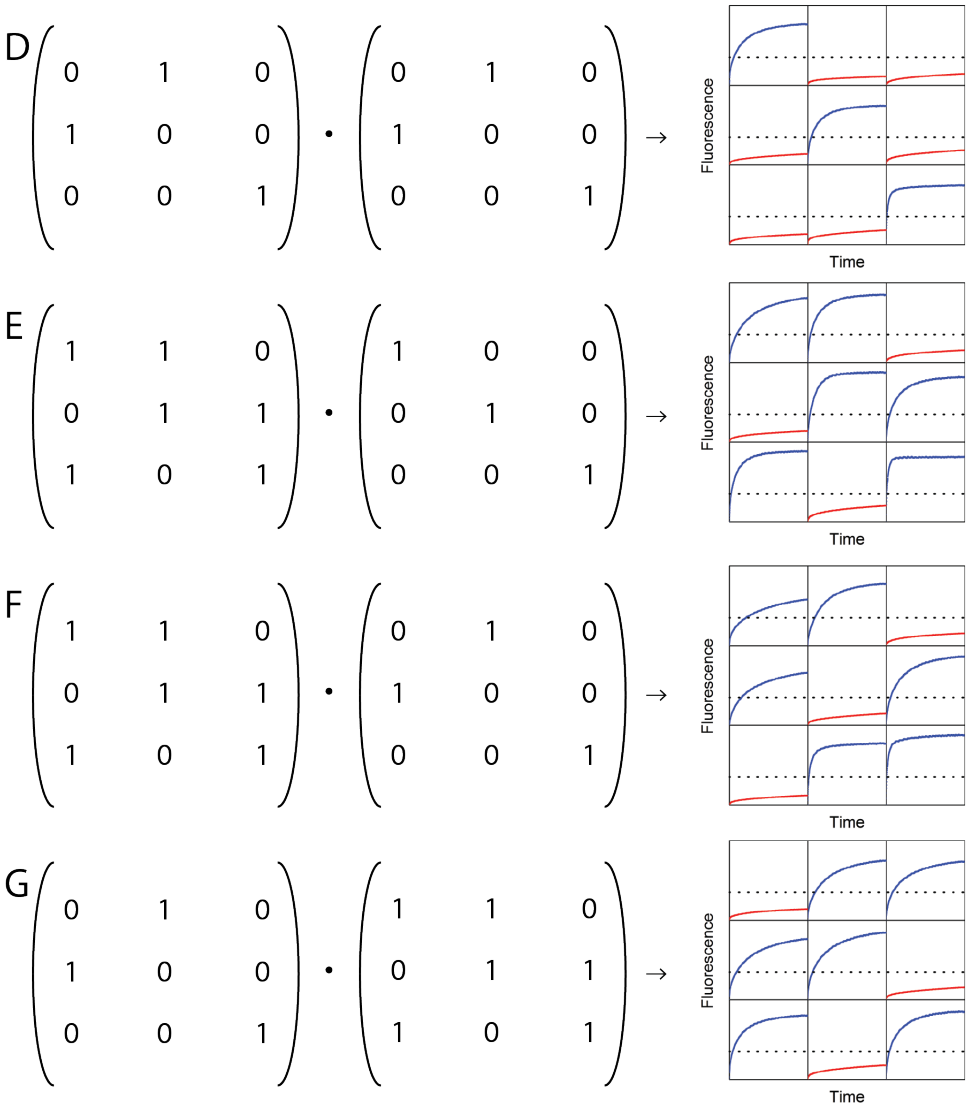

Figure 3.5: Further examples of 3×3 Boolean matrix multiplication (D, E, F, G).

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

were performed, and the appropriate invader strands mixed at concentration c, three invader complexes would form ($d_{11}p_{11}$, $d_{12}p_{21}$ and $d_{13}p_{31}$), all capable of displacing incumbent from read-out duplex r_{11}. The combined concentration of these complexes would be $3c$. Consider as an alternative example

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} .$$

Here, distal invader d_{11} would be divided between three invader complexes, $d_{11}p_{11}$, $d_{11}p_{12}$ and $d_{11}p_{13}$. If all form with equal concentration (a reasonable assumption considering the equality of their respective linking domains), the concentration of the one complex that interacts with r_{11} is now only $c/3$. This difference in invader complex concentration can be expected to manifest itself in the reaction rate. Sequence differences between the read-out duplexes, as well as competition for toeholds between correct invader complexes, unpaired proximal invaders and invader complexes with a spurious distal invader may also contribute to rate variations.

3.3 COMBINATORIAL STRAND DISPLACEMENT CASCADES

3.3.1 *Proof of principle*

The last section showed how combinatorial strand displacement can be scaled up to allow for the multiplication of matrices with greater dimensions. Here, it is shown that combinatorial strand displacement reactions can also be cascaded, meaning that the output of one such reaction triggers the next. For this purpose, the distal invader for a read-out duplex is sequestered as an incumbent in an intermediate duplex, from which it can be released by a combinatorial strand displacement reaction. This process is depicted schematically in Figure 3.6A. The proof-of-principle measurements shown in Figure 3.6B are broadly consistent with the expectation that significant triggering of a fluorescent response is conditional on the presence of all invading strands.

This cascadability increases the complexity of calculations that can be performed by a network of such reactions. The elements of a product matrix resulting from a chain multiplication of three Boolean matrices $M \cdot N \cdot O$ with matrix dimensions $i \times j$, $j \times k$ and $k \times l$, respectively, are

$$P_{ab} = \begin{cases} 1 & \text{if } \sum_{y=1}^{j} \sum_{z=1}^{k} M_{ay} N_{yz} O_{zb} > 0 \\ 0 & \text{if } \sum_{y=1}^{j} \sum_{z=1}^{k} M_{ay} N_{yz} O_{zb} = 0 \end{cases}.$$

As before, this can also be expressed in a logic symbol representation:

$$P_{ab} = \bigvee_{y=1}^{j} \bigvee_{z=1}^{k} M_{ay} \wedge N_{yz} \wedge O_{zb}, \tag{3.2}$$

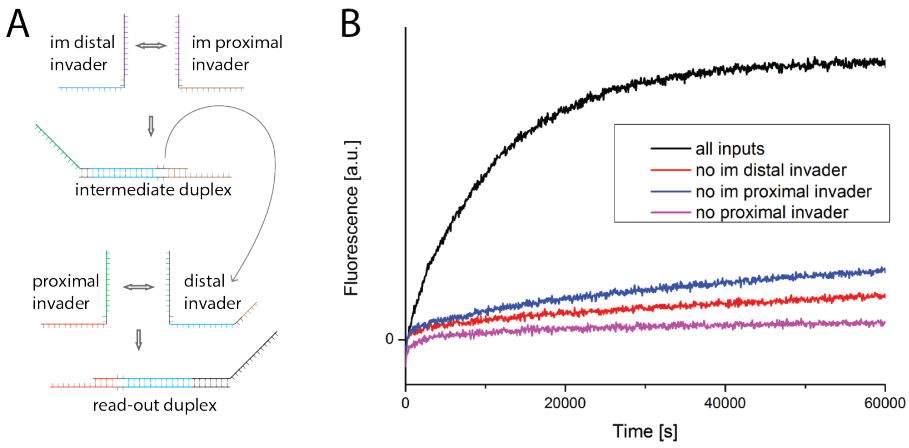

Figure 3.6: Cascaded combinatorial strand displacement. **A:** Reaction scheme. An intermediate (im) distal invader and intermediate proximal invader can hybridise to form an invading complex capable of displacing incumbent from an intermediate duplex. This incumbent, when released, can serve as a distal invader to a further combinatorial strand displacement reaction, as illustrated in Figure 3.2A. **B:** Fluorescent response of reporter complex upon addition of different combinations of input to a mixture of intermediate duplex, read-out duplex and reporter complex (all at 20 nM). If both intermediate invaders and the proximal invader are present, fluorescence increases much faster compared to a reaction in which one of these inputs is missing.

which, if we introduce the intermediate product matrix defined by the elements

$$P^*_{az} = \bigvee_{y=1}^{j} M_{ay} \wedge N_{yz} \tag{3.3}$$

can be written as

$$P_{ab} = \bigvee_{z=1}^{k} P^*_{az} \wedge O_{zb}. \tag{3.4}$$

It was shown above that calculations of the type shown in equations 3.1, 3.3 and 3.4 can be represented using the reaction scheme of Genot and colleagues [37] or that in Figure 3.2A. Cascading this scheme as in Figure 3.6A then allows for an implementation of equation 3.2: The intermediate distal invaders are chosen to represent M and the intermediate proximal invaders N. When mixed, they will interact with each other and the intermediate duplex according to equation 3.3, the output representing the intermediate product matrix P^*. This output can act as distal invaders to the previously introduced read-out duplex, with the proximal invaders chosen to represent O. The interaction between the strands representing P^*, those representing O and the read-out duplexes proceeds in accordance with equation 3.4, thus completing the chain multiplication[1].

3.3.2 *Reachabilty calculations on directed graphs*

A directed, unweighted graph with j nodes can be represented through a characteristic $j \times j$ Boolean matrix, in which the element with index pair (a, b) is 1 if there exists a vertex from graph node a to b, and 0 otherwise.

1 It should be noted that the order in which the individual matrix multiplications are performed, i.e. $(M \cdot N) \cdot O$, as presented here, or $M \cdot (N \cdot O)$, is irrelevant for the result, as matrix multiplication is associative. The order is, however, generally relevant for the efficiency of the calculation (e.g. [50, 51]). As in this work, chain multiplication is only applied to cases in which $M = N = O$ (subsection 3.3.2), such considerations are not necessary here.

Such a matrix is commonly referred to as an adjacency matrix. Trivially, this matrix allows immediate recognition of which node can be reached from which with a single step, i.e. without passing through another node. Multiplied with itself, the resulting product matrix reveals which node can be reached from which in exactly two steps, etc. Below, it is shown how such reachability matrices can be calculated for 3-node graphs using the combinatorial strand displacement mechanisms introduced above.

Implementation

The implementation to calculate a 1-step reachability matrix for a 3-node graph is trivial: the elements of the original adjacency matrix A are represented by invading strands which can directly trigger read-out duplexes of identical index, without having to interact with any other reaction network component. The ability to calculate a 2-step reachability matrix was essentially demonstrated in subsection 3.2.2, as such a calculation merely requires the multiplication of a 3×3 adjacency matrix with itself, i.e. A^2. For both systems, experiments were prepared in such a fashion that the reporter complex was present at a concentration of 15 nM, and all other components at 100 nM.

For the 3-step reachability matrix, the cascaded system introduced above was used to calculate A^3. For this, nine intermediate duplexes were designed, with corresponding sets of nine intermediate distal invaders and nine intermediate proximal invaders. These intermediate duplexes were prepared in the same fashion as the read-out duplexes. Input mixes were selected from the three sets of invaders (intermediate distal, intermediate proximal, proximal) correspondent to the adjacency matrix of the graph for which the reachability calculation was to be performed[2].

2 n-step reachability is defined here as a node-to-node connection in exactly n steps, not up to n steps.

Read-out duplexes and reporter were distributed across nine plate wells, as for the simple, non-cascaded combinatorial displacement systems. In addition, each well contained three of the nine intermediate duplexes, as for a given read-out duplex, there are only three intermediate duplexes containing a distal invader capable of interacting with the read-out duplex. For example, read-out duplex r_{11} can only be triggered in the presence of distal invaders d_{11}, d_{12} or d_{13} (Figure 3.3B,C). As only intermediate duplexes i_{11}, i_{12} and i_{13} contain these particular distal invaders, presence of the other six intermediate duplexes is not required in well $(1,1)$.

As before, the identical, annealed input mix was added to all wells, where reporter complex was present at 15 nM. In preliminary experiments, all other components were chosen to be at concentrations of 100 nM each, as for the other two classes of systems. This was found to yield ambivalent results, however.

The experiments in subsection 3.2.2 showed that although 0-signals and 1-signals are readily distinguishable when multiplying two 3×3 matrices, different indices of the product matrix showed varying reaction rates. This effect was found exacerbated for the cascaded system, to a point where, in preliminary attempts to calculate 3-step reachability matrices for the graphs depicted in Figures 3.9A, 3.10A and 3.11A, the two types of signals became occasionally indistinguishable (Figure 3.7).

To quantify this undesirable inhomogeneity in kinetic behaviour, the responses to input representing

$$1^3_{3\times 3} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}^3 = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

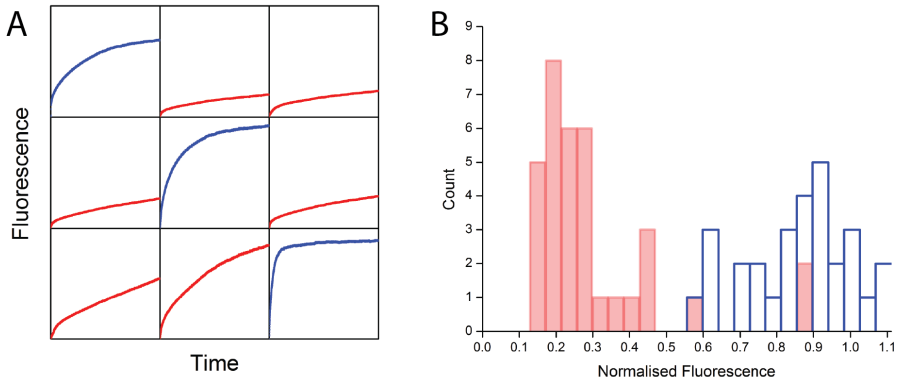

Figure 3.7: Ambivalent results from preliminary experiments. **A:** 3-step result matrix for the graph on the left in Figure 3.10A. Despite the fact that node 2 cannot be reached in three steps from node 3, the corresponding index $(3,2)$ displays a fluorescent response similar to those indices expected to represent a '1'. For these plots, fluorescence was normalised as previously; time ranges from 0 to 2400 min. **B:** Final fluorescence values for the cascaded combinatorial displacement when no concentration adjustment is performed for all seven graphs tested (shown below in Figures 3.9A, 3.10A and 3.11A). As the distributions for ones and zeros overlap, no unambiguous binary classification can be made.

were determined. This input combination was chosen not only because it produces a '1' in every index of the product matrix, but also because the concentration of invader complex interacting with each read-out duplex is expected to be the same.

Figure 3.8A shows that these responses were, as expected, non-uniform. Half-completion time estimates were determined for each trace by identifying the first data point to cross a normalised fluorescence value of 0.5. The average was found to be $\overline{t_{1/2}} = (186.2 \pm 66.9)$ min.

The way the cascaded system is designed, the overall rate of a reaction is limited by the displacement of the read-out incumbent. It is certainly slower than the linking domain hybridisations. Zhang and Winfree's work on toehold exchange [156] also suggests that, due to the 7 nt exclusive binding domain offered to incumbent, it is orders of magnitude slower than displacement of distal invader from the intermediate duplex and displacement of reporter incumbent from reporter complex, where

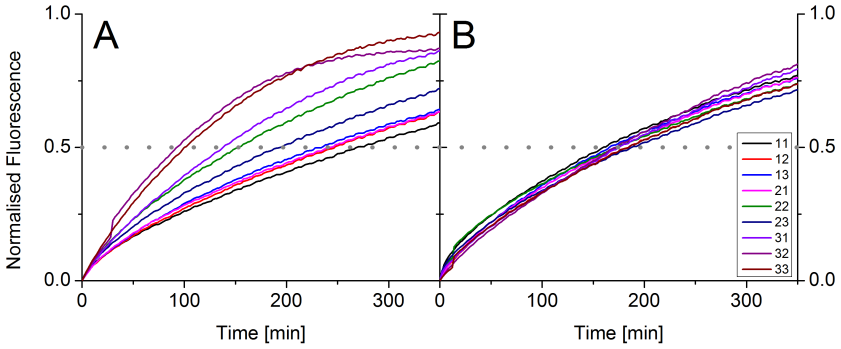

Figure 3.8: Homogenising fluorescent response across product matrix indices by adjusting reaction network component concentrations. **A:** Native response to inputs corresponding to $1^3_{3\times3}$. The times at which the traces reach a normalised fluorescence value of 0.5 (grey dotted line) varies strongly. **B:** Upon concentration adjustment based on half-completion times, the fluorescent response is uniform.

such a reverse toehold is either very short (2 nt, intermediate duplex) or not present at all (reporter complex).

More accurately, it can be assumed that the rate-limiting step is the resolution of a complex consisting of both invaders and the read-out duplex, which produces free read-out incumbent capable of interacting with the reporter complex. This would not be a reasonable assumption if this reaction were a regular displacement reaction with an 8 nt toehold, which can be approximated well as a second-order reaction (chapter 2, [156]). Here, however, NUPACK [152] predicts that due to the double T-T mismatch the incumbent forms with the target close to the toehold, the proximal invader toehold is effectively as stable as a 10 nt toehold. Recalling the measurement with correct inputs present in Figure 3.2, however, the combinatorial displacement reaction is much slower than those direct displacement reactions with a 10 nt toehold measured in chapter 2 and by others [156]. Approximate half-completion times are over two orders of magnitude apart. This only leaves the above conclusion that it is not displacement reaction initiation, but rather displacement completion which limits rate.

This process is a first-order reaction, in which the rate depends linearly on a rate constant determined by system design (reverse toehold strength, junction, displacement domain sequence, experimental conditions), and the concentration of the invader-read-out-duplex compound. This concentration, in turn, depends on multiple factors: depending on input composition, either the input strands, the intermediate duplex or the read-out duplex itself can be concentration-limiting. It can thus be expected that changing *all* reaction components by a certain factor leads to a corresponding change in overall reaction rate by the same factor. This was confirmed experimentally, and used to homogenise the fluorescent response to equivalent input in each well.

To achieve kinetic uniformity for all indices of the product matrix, the component concentrations were individualised for each index (a, b) according to their half-completion times determined for $1^3_{3\times3}$:

$$c_{ab} = \frac{t_{1/2}\,(a,b)}{\overline{t_{1/2}}} \cdot 100\,\text{nM}.$$

These concentrations ranged from 50 nM to 145 nM across the nine wells, but were kept constant across experiments, so that for each graph, the 3-step reachability calculation was conducted under the same conditions, independent of input composition. $1^3_{3\times3}$ was measured again for these individualised concentrations; the resulting kinetic response was found to be very homogenous (Figure 3.8B), with the half-completion times averaging $\overline{t_{1/2}} = (175.8 \pm 9.5)$ min.

It should be reiterated that this procedure can not be expected to make all '1' traces look exactly alike, nor all '0' traces. This can not be achieved with this system design, as concentrations of invader complexes will still vary from well to well, depending on input composition. It only aims to minimise response variations that are due to 'design flaws', i.e. differences in domain binding strengths, propensity for transient spurious

duplex formation etc. It can be best understood as an attempt to approximate the response of an ideal system, in which all the hybridisation and displacement reactions behave exactly alike. In order to absolve the system design of its response variability that is due to input composition, sophisticated signal restoration devices would be required; a possibility that is briefly discussed in the conclusions (section 3.5).

Results

Figure 3.9A shows a directed, 3-node graph with adjacency matrix

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

Using the above implementation, matrices corresponding to 1-step reachability ($A^1 = A$, Figure 3.9B), 2-step reachability (A^2, Figure 3.9C). and 3-step reachability (A^3, Figure 3.9D) were calculated. Six further examples of equivalent calculations performed by these reaction networks are shown in Figures 3.10 and 3.11. All 3-step reachability calculations were performed with component concentrations adjusted as described above.

3.3.3 *System fidelity*

For most graphs, as well as the matrices shown in subsection 3.2.2, fluorescence traces representing ones are readily distinguishable from those representing zeros. In the following, binary classification of the output of the reaction networks presented in this chapter is explored in a more quantitative fashion, and the cut-off values chosen for each class of system are justified.

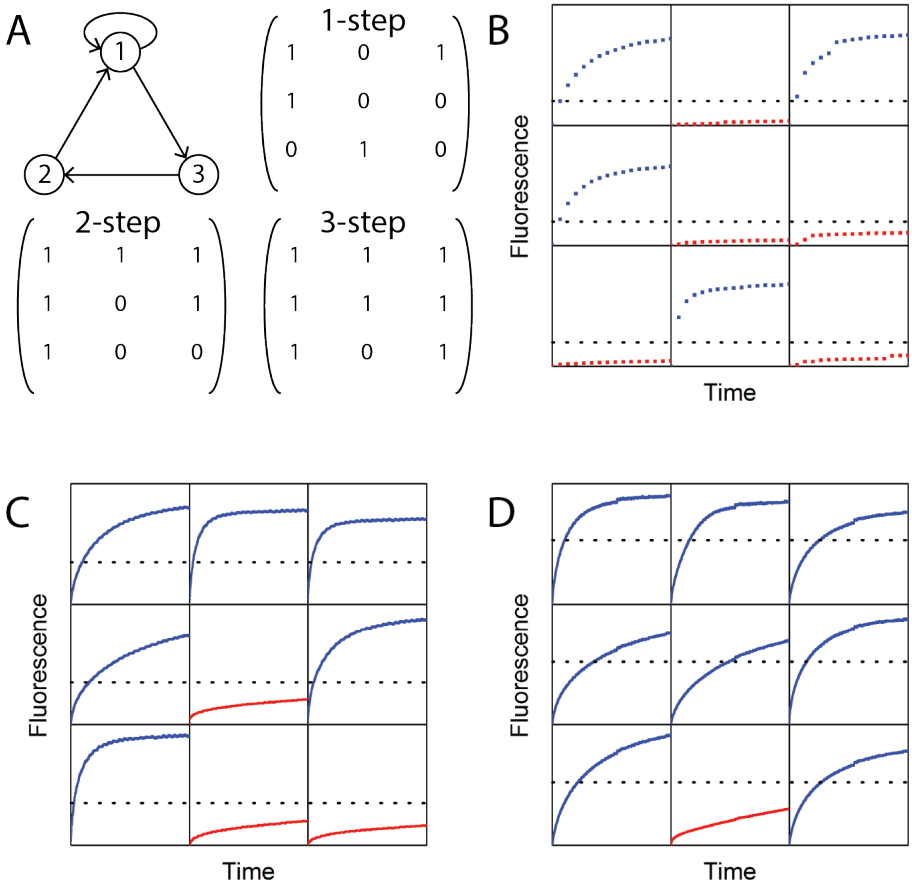

Figure 3.9: Reachability calculations on a directed graph. **A:** An exemplary 3-node graph, with corresponding 1-, 2-, and 3-step reachability matrices. **B:** Reachability for a single step, calculated with a direct displacement system. **C:** Reachability for two steps, calculated with a simple combinatorial displacement system. **D:** Reachability for three steps, calculated with a cascaded combinatorial displacement system. Traces are coloured and normalised as before. Time ranges from $0 - 30$ min for the direct displacement, $0 - 1400$ min for the simple combinatorial displacement, and $0 - 2400$ min for the cascaded combinatorial displacement. As for the simple combinatorial displacement systems, a constant cut-off can also be found for both other classes of systems which unambiguously separates zeros and ones (subsection 3.3.3).

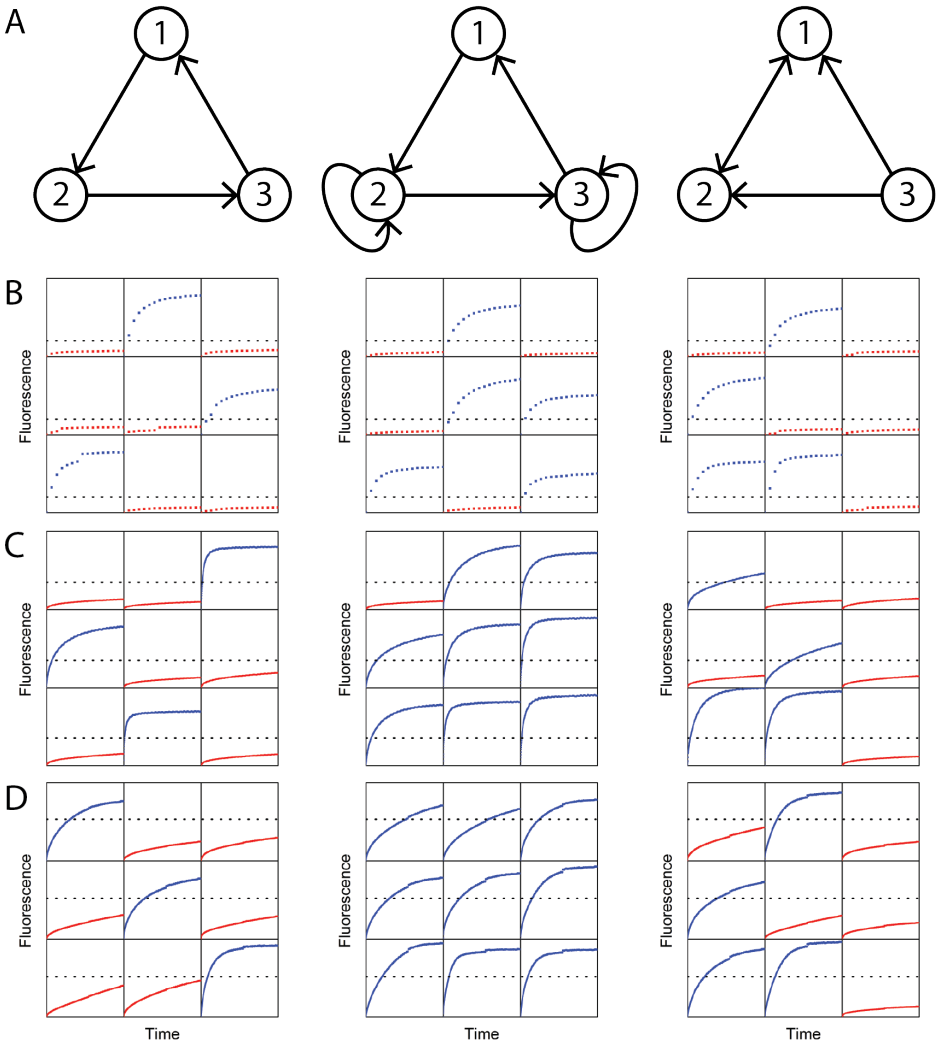

Figure 3.10: Further examples of reachability calculations on graphs (part 1/2). **A:** Graphs translated to system input. **B:** 1-step reachability/adjacency matrix. **C:** 2-step reachability. **D:** 3-step reachability. Definitions, fluorescence and time axis ranges and cut-off values as in Figure 3.9.

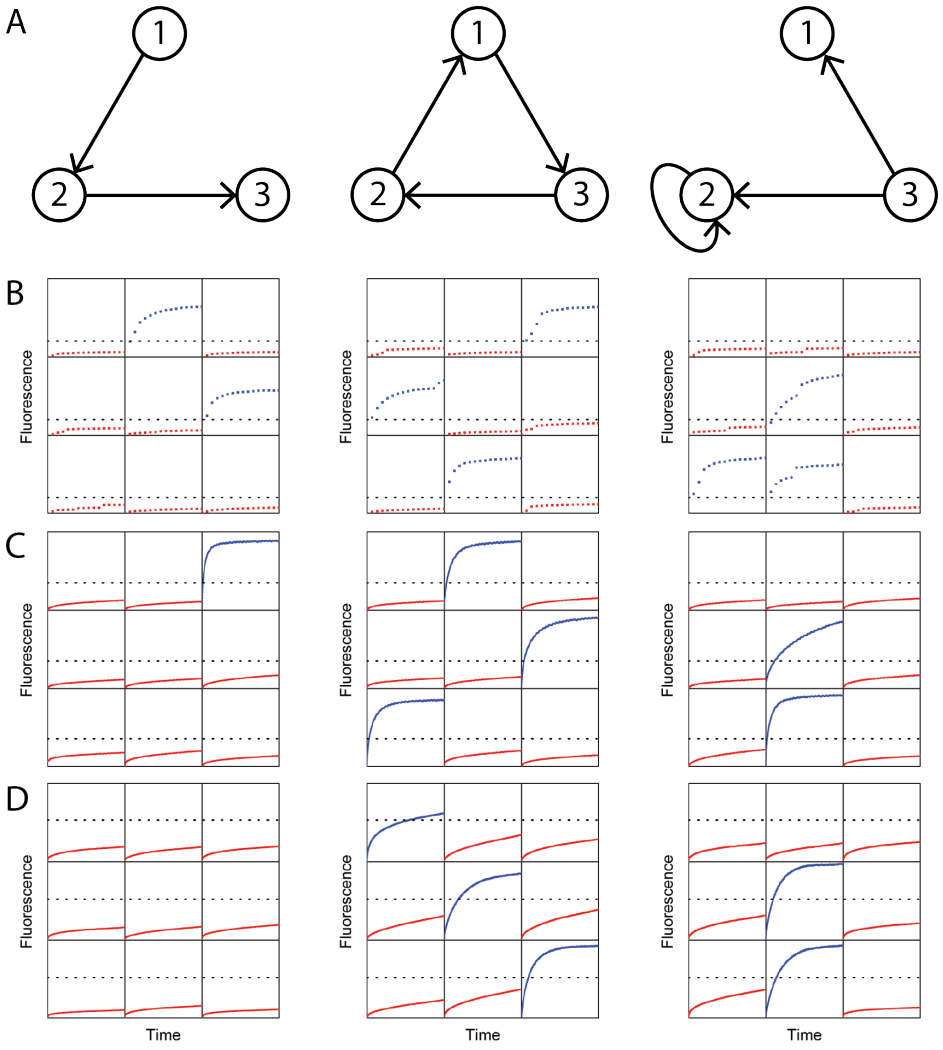

Figure 3.11: Further examples of reachability calculations on graphs (part 2/2).

Figure 3.12 shows three column plots; one for each of the classes of reaction networks presented above. For each class, the final fluorescence values of those traces expected to represent ones and those expected to represent zeros were collated into histograms. In each of the classes, a dichotomy between the two distributions becomes apparent. A cut-off fluorescence in each class can thus be identified, which divides traces unambiguously into ones and zeros, yielding the correct, expected result for every trace measured[3].

In addition, this data can be used to estimate overall system reliability, i.e. how likely each class is to yield false positives (ones that were expected to be zeros) or false negatives (zeros expected to be ones) given inputs different from those tested in above experiments. For this purpose, it was assumed that within each class, the final fluorescence values form two normal distributions. Further assuming that the graphs and matrices measured are generally representative of all other input combinations, Gaussian fits allow for an estimate of an 'optimal' cut-off fluorescence value[4]. If false positives and false negatives are given the same cost, i.e. if the sum of the probability of measuring a false positive and the probability of measuring a false negative is to be minimised, the optimal cut-off lies at the intersection of the two distributions [85]. The probability of measuring a false positive was estimated accordingly as the area to the right of this fluorescence value f_{co} under the Gaussian corresponding to the '0'-distribution:

$$P(fp) = 1/2 \cdot \left(1 - erf\left(\frac{|f_{co} - \mu_0|}{\sqrt{2}\sigma_0}\right)\right),$$

3 It is worth comparing the '1' distributions of Figures 3.7B and 3.12C, which depict final fluorescence values for the cascaded systems for all graphs tested. The latter distribution, using the concentration adjustment, is narrower than the one without.

4 Finding optimal cut-off points for binary classification in distributions has received particular attention within statistical medicine, where it is of great relevance in the context of clinical decision-making. As a result, there exists a wide range of approaches to finding such cut-off values (e.g. [79]). Here, preference is given to a fairly straight-forward and intuitive method, which is entirely sufficient to characterise the systems studied.

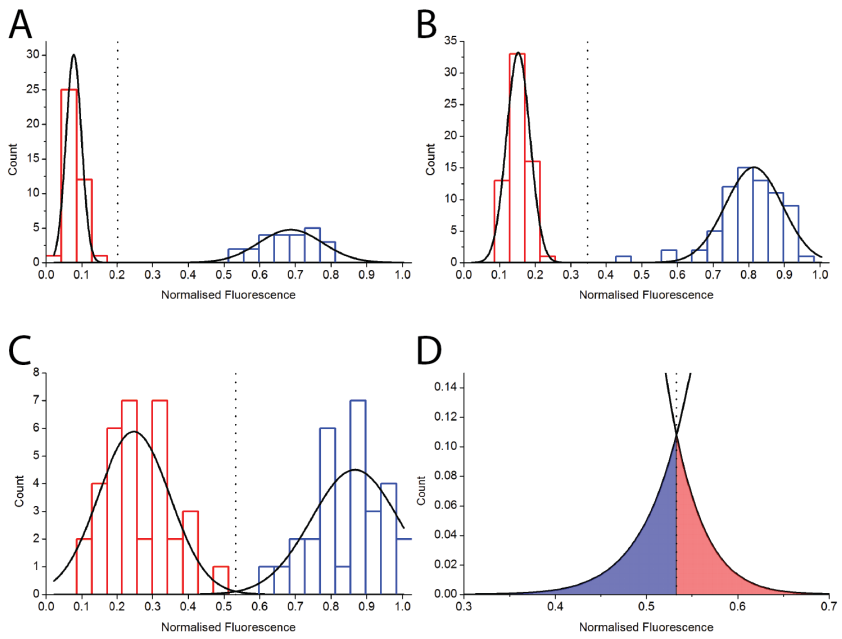

Figure 3.12: Final fluorescence values of each trace measured collated in histograms. **A:** Histograms for traces measured for the direct displacement class of systems, used to determine 1-step reachability for the graphs in subsection 3.3.2. The red histogram contains final fluorescence values of all traces expected to represent a '0'; the blue histogram contains final fluorescence values of all traces expected to represent a '1'. Gaussian fits to each of the histograms are also depicted (black). The dotted line represents the optimal cut-off value for these distributions. **B:** Correspondent plot for the simple combinatorial strand displacement class of systems, used to calculate the matrix products in subsection 3.2.2 and the 2-step graph reachability. **C:** Correspondent plot for the cascaded combinatorial strand displacement class of systems, which were used to calculate 3-step graph reachability. **D:** Section of the Gaussian fits in C, highlighting the area where the two fitted distributions overlap. The cut-off value is at the intersection of the curves; the area to the left under the curve (blue) represents the probability for measuring false negatives, the area to the right (red) the probability for false positives.

Reaction network	f_{co}	$\frac{\lvert f_{co}-\mu_0\rvert}{\sigma_0}$	$\frac{\lvert f_{co}-\mu_1\rvert}{\sigma_1}$	$P(fp)$	$P(fn)$
Direct displacement	0.20	5.7	5.4	4.7×10^{-9}	3.0×10^{-8}
Simple combinatorial	0.35	6.0	5.8	1.2×10^{-9}	2.8×10^{-9}
Cascaded combinatorial	0.53	2.8	2.7	2.3×10^{-3}	3.2×10^{-3}

Table 3.1: Probabilities of measuring false positives or false negatives for the three reaction network classes.

where μ_0 is the fitted mean of the '0' distribution, σ_0 the fitted standard deviation, and $erf(x)$ the Gauss error function. The probability of measuring a false negative was estimated as the area to the left of f_{co} under the Gaussian corresponding to the '1'-distribution:

$$P(fn) = 1/2 \cdot \left(1 - erf\left(\frac{|f_{co} - \mu_1|}{\sqrt{2}\sigma_1}\right)\right).$$

Table 3.1 contains the results of these calculations.

From these values, both the direct displacement and the simple combinatorial displacement systems appear to be extremely reliable. Reliability of the cascaded systems is high, but several orders of magnitude lower than for the other systems. This is assumed to be mainly due to two reasons. Firstly, reactions yielding a '1' in the cascaded systems can be slower than for the other two systems, as the variable invader complex concentration issue described at the end of section 3.2.2 can be exacerbated. Certain types of leak reactions, such as blunt-end strand displacement of reporter incumbent through hybridised read-out incumbent thus become more relevant.

Secondly, in many wells in which a '0' is expected, both the correct proximal and distal invaders for the correspondent read-out duplex are present, with the distal invader hybridised in an intermediate duplex. It is conceivable that in such wells, the proximal invader, by hybridising both the read-out duplex toehold region as well as the distal invader linking domain, is capable of bringing the read-out duplex and the inter-

mediate duplex into close proximity, increasing the rate of an otherwise unlikely strand exchange process.

3.4 MULTI-STRAND COMBINATORIAL DISPLACEMENT

Zhu and co-workers [160] demonstrated that three strands instead of two can form an invader complex as well: in their scheme, two of the invading strands are linked by a third, instead of binding to each other directly. This could conceivably be expanded to invader complexes comprising a greater number of component strands.

Such a mechanism, as depicted schematically in Figure 3.13A, can be designed to follow the exact same set of rules as cascaded two-strand combinatorial displacement. For a chain multiplication of n matrices, an invader complex would consist of n strands, each with two domains representing a fixed matrix index (or vertex, in the context of graphs) which link the previous strand in the complex to the next.

In preliminary experiments, the two-strand combinatorial displacement was extended in a straight-forward fashion. The components of the invader complexes, consisting of two to five individual strand species, were first annealed to ensure optimal complex formation. Subsequently, their respective ability to displace an incumbent from a read-out duplex was tested. For this purpose, the response of reporter complex was observed upon addition of the read-out duplex and invader complex (all components at 20 nM). The change in fluorescence for each type of invader complex is depicted in Figure 3.13B.

There is a clear correlation between the number of strands in the invader complex and the time it takes to displace a given amount of incumbent. The three-stranded complex is much slower than the two-stranded complex, and the four-stranded complex is hardly faster than the leak re-

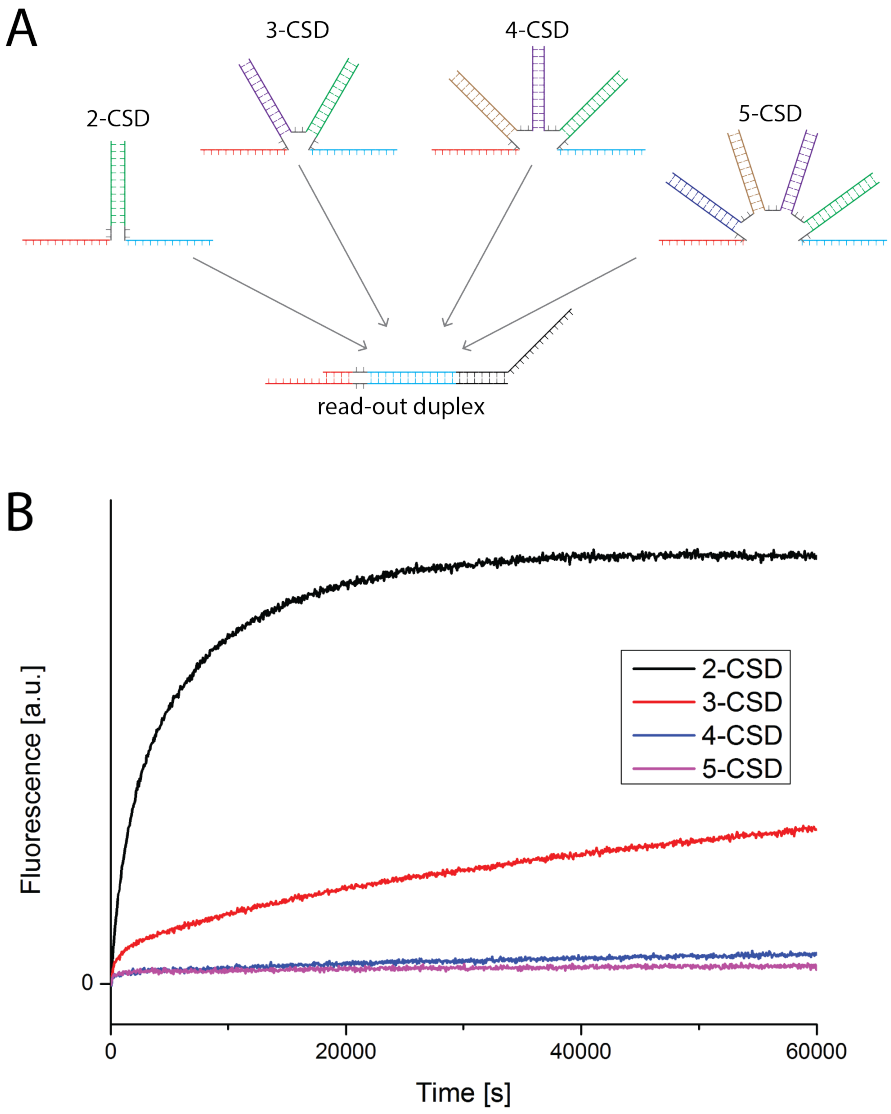

Figure 3.13: Using invader complexes consisting of between two ('2-CSD') and five ('5-CSD') strands. **A:** Schematic depiction of invader complexes, which are all designed to interact with the same read-out duplex. **B:** Fluorescent response to invader complex addition to read-out duplex and reporter complex.

actions observed in Figure 3.2, with the five-stranded complex not showing any measurable displacement activity at all. This strongly restricts this mechanism's immediate applicability in Boolean matrix chain multiplication, as for at least more than three matrices, a cascaded system would be required.

Given the correlation observed, it can be assumed that with a growing number of component strands, the complexes are less capable of bringing the toehold- and the displacing domains into favourable relative positions for successful displacement. Thus, redesigning the invader complexes and/or read-out duplexes may make this approach more feasible for application in DNA logic circuits.

3.5 CONCLUSIONS & OUTLOOK

This chapter investigated DNA strand displacement through complexes of invading strands. It was shown that previous designs of this combinatorial strand displacement can be improved by reducing their propensity for unwanted side reactions. This allowed for implementation of a DNA-based calculator for Boolean matrix products which displayed remarkable reliability. Cascadability of combinatorial strand displacement was also demonstrated. A reaction network based on this idea was shown to be capable of performing matrix chain multiplication. In conjunction with a simple strand displacement design, the two combinatorial displacement networks were applied to calculating reachability properties of small, directed graphs. The last section demonstrated that complexes of more than two invaders may also be feasible for combinatorial displacement reactions, although a clear inverse correlation between displacement efficiency and number of invader complex components was discernible.

Experimentally, this chapter proposed to move away from differentiating reaction network output channels by selecting a corresponding number of spectrally separate fluorophores, as this creates a relatively modest ceiling for output complexity. Instead, a spatial separation was proposed, in which each output channel is represented by a specialised read-out duplex which translates to a universal reporter complex, each in their own, dedicated reaction vessel. This was made experimentally feasible by using a plate reader, which is more commonly used for quantitative PCR.

Overall, combinatorial strand displacement provides a powerful alternative to DNA devices and reaction networks based solely on regular strand displacement. The latter could have been used as well to achieve the applications presented in this chapter, as all logical operations required for Boolean matrix calculations as described in equation 3.1 can be implemented. The strength of the combinatorial approach, however, lies in the fact that it requires far fewer reactive components for this task and can be scaled with greater efficiency.

An interesting next step would be to synthesise combinatorial strand displacement with reaction network components originally designed for regular strand displacement. In the context of the variance in invader complex concentrations, an implementation of signal restoration mechanisms would be particularly desirable. These would optimally guarantee that output of an intentionally triggered combinatorial displacement reaction would be amplified to a fixed concentration, whereas small amounts of erroneously released output would be transformed into non-reactive compounds. Designs for such thresholding intermediates could be inspired by the impressive systems developed by Seelig and co-workers [111] as well as Qian and Winfree [98]. Even without such additions, at least the non-cascaded combinatorial displacement network can be as-

sumed to be scalable to greater matrix dimensions, considering its pronounced separation between ones and zeros.

Despite the low displacement efficiencies of invading complexes consisting of more than two strands, this approach to combinatorial displacement should not be discounted entirely. A more sophisticated design of such complexes, in which the distance between the two displacing domains is decreased, may reduce this drawback (Figure 3.14). Considering the advent of scaffold-free DNA origami [140, 60], it may even be possible to assemble a great number of individual strands into a large invading complex – in contrast to regular origami engineering, however, such structures would be designed to be highly sensitive to the absence of single components. In the context of matrix chain multiplication, each additional component strand raises the chain length by one, requiring only a single read-out duplex, whereas such an increase in chain length for a cascaded system requires a further intermediate complex, which likely increases the impact of leak reactions.

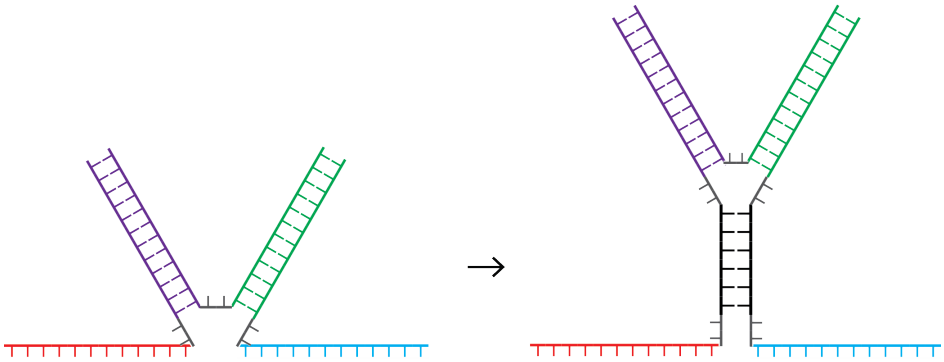

Figure 3.14: A proposal for a multi-strand combinatorial displacement mechanism in which the number of strands in the invading complex is assumed to have a reduced impact on reaction rate. The junction between the invading strands is moved away from the displacing domains. Instead, it is connected to the displacing domains through a linker (black) which is too weak to provide a stable connection between proximal and distal invader on its own, but in the presence of a third, connecting strand (top) brings the two displacing domains into close proximity. This spacer could be identical across all invaders, and would thus not impact the computational rule set of combinatorial strand displacement.

4

QUANTUM DOT MICROSCOPY FOR DNA NANODEVICE OBSERVATION

4.1 INTRODUCTION

This chapter describes work towards the ability to image synthetic DNA nanodevices using quantum dots, a type of semiconductor nanocrystal, as labels for single-molecule fluorescence microscopy. In comparison to classical fluorophores, quantum dots can go through a much greater number of excitation-emission cycles before bleaching, allowing for nanoscale localisation for extended periods of time. This is an important quality particularly in the context of dynamic DNA nanodevices, which commonly operate on long time scales and sub-diffraction-limit length scales.

The first section of this chapter extends the introduction to single-molecule microscopy provided in chapter 1, focusing on studies in which quantum dots were used. The following section describes experimental implementation. The methodological details can be found in Appendix A.1. Section 4.3 describes how DNA origami tiles were labelled with two differently coloured quantum dots and the sub-diffraction-limit distance between the labels determined. Section 4.3.2 demonstrates that such distance measurements are also possible with labels of the same colour, taking advantage of the unique fluorescent behaviour of quantum dots. The findings of section 4.3 are applied in the context of a DNA nanomotor

in section 4.4, in an attempt to elucidate details of the operation of this particular device.

Contributions by colleagues and collaborators

In addition to work directly referenced, the following contributions to this chapter have been made by colleagues and collaborators:

- Andrew J. Turberfield suggested the idea of using quantum dots as fluorescent labels to track the operation of dynamic DNA nanodevices;
- Stephan Uphoff (Department of Biochemistry, University of Oxford) kindly provided me with Matlab code, capable of detection and single-molecule analysis of fluorescent signals in microscopy recordings. I augmented this code to allow for the channel mapping described in subsection 4.2.2, the distance determination method of section 4.3.2 and to optimise data analysis for the entire chapter.

4.1.1 *Quantum dots as fluorescent labels*

Section 1.3 gave an introduction to the most common observation techniques previously applied to the study of DNA nanodevices. It became clear that there exists no accomplished methodology for single-molecule studies that combines the spatial and temporal resolution required for observation of detailed DNA nanodevice operation whilst remaining easily accessible, and applicable to a variety of device designs and operational conditions. Those techniques which fulfil the first requirement (high-speed AFM, single-molecule FRET microscopy) fail the second, and those which fulfil the second (bulk FRET, single-molecule colocalisation microscopy) fail the first. The colocalisation-microscopy study by Lund

and co-workers [72] showed promise in overcoming this divide, particularly because it allowed for fast testing of system response to different conditions and design modifications. Localisation precision, however, was only sufficient to derive results that did not extend far beyond those accessible through simple bulk measurements. Here, an attempt at improving this approach is made by using quantum dots as fluorescent labels, to achieve sustained, high SNRs on the timescale of DNA nanomachine operation.

Quantum dots are semiconductor crystals with diameters commonly between $5 - 20\,\text{nm}$ [2], which display remarkable fluorescent behaviour [101]. Depending on size, composition and ambient conditions, quantum dots have a fluorescent lifetime exceeding that of conventional fluorophores by several orders of magnitude [127] and also display highly tunable emission in conjunction with broad excitation spectra. Due to these qualities, they have been applied broadly in the biological sciences [81].

Two such studies are briefly presented here, as they are particularly relevant to the use of quantum dots in this work. Dahan and co-workers tracked individual, quantum-dot-labelled glycine receptors in living neurons with localisation precisions below 10 nm for up to 20 min, suggesting that single quantum dots can be localised with high precision for extended periods of time [22]. Using two differently coloured species of quantum dots, Warshaw and colleagues investigated the stepping behaviour of myosin, a motor protein, achieving nanometric colocalisation, thus demonstrating that mechanistic details of highly dynamic processes can be analysed reliably using such labels [138].

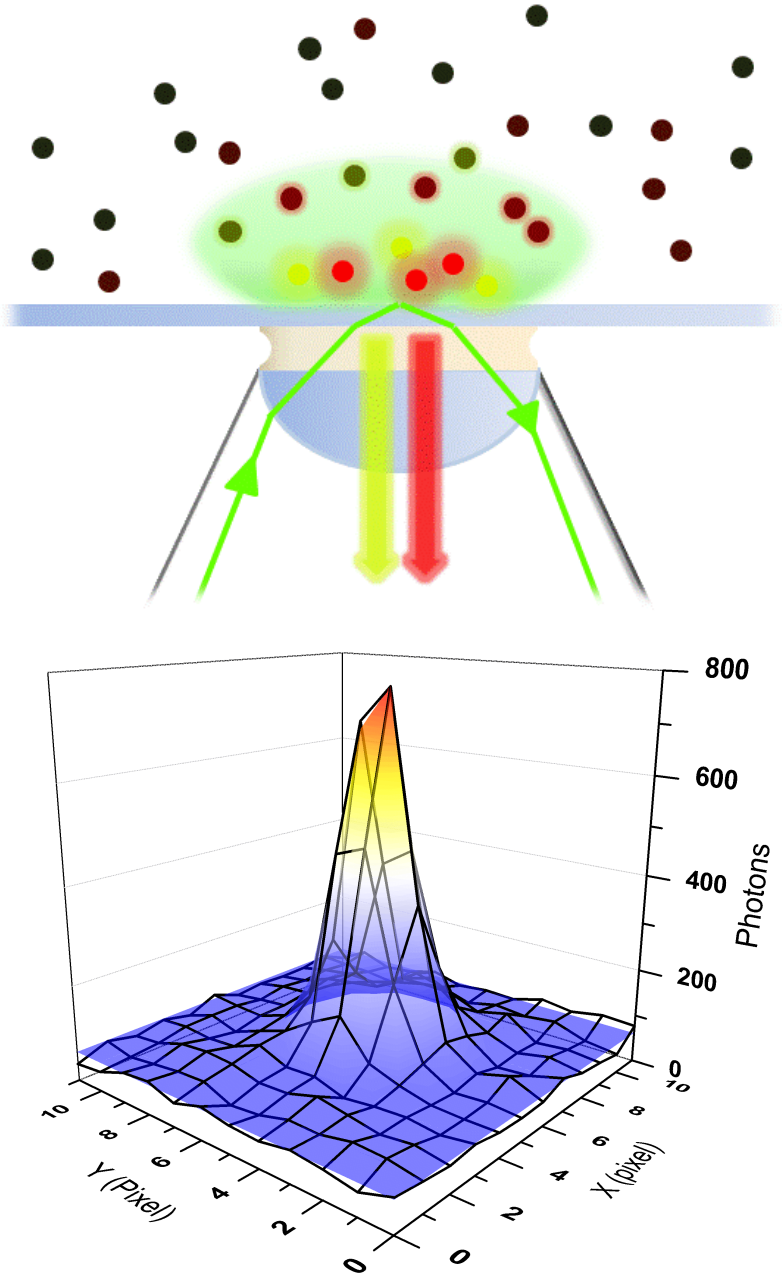

Figure 4.1: **Top:** Schematic representation of TIRF with a sample containing two species of quantum dots (not to scale). The excitation laser (green arrow), which is delivered to the sample through the objective, is reflected at the glass-buffer interface of the cover slip (horizontal) and the sample. A resulting evanescent wave in the sample only excites those quantum dots closest to the cover slip surface. The emitted fluorescent signals can be spectrally separated from each other and from stray excitation light. The evanescent wave does not reach quantum dots far from the cover slip surface (i.e. $\gtrsim 200\,\mathrm{nm}$), reducing background fluorescence. **Bottom:** Exemplary, low-background PSF of a single quantum dot (black grid), superimposed with a fitted two-dimensional Gaussian peak (coloured surface), which allows for nanometric localisation of the quantum dot centre.

4.2 IMPLEMENTATION

4.2.1 *Experimental setup*

An inverted microscope served as the principal experimental platform for the work presented in this chapter. This microscope was operated with TIRF illumination (Figure 4.1), an electron-multiplying charge-coupled device (EMCCD) camera and a two-channel emission beam splitter. The microscope and its extensions are described in Appendix A.1.5.

Quantum dots consisting of a cadmium selenide core, a zinc sulfide shell and a streptavidin coating (approximately 20 nm diameter, 5 to 10 streptavidin per quantum dot) were supplied by Life Technologies in two variations: one with peak emission at 565 nm, the other with peak emission at 655 nm.

4.2.2 *Nanometric colocalisation*

In order to compare individual positions of two spectrally separated channels, there must exist a mapping function, which assigns each point in the first channel a unique point in the other. In order to find such a function, Churchman and colleagues, when introducing SHREC [19], demonstrated the use of multi-fluorescent beads, which are visible in both channels simultaneously. In their work, a single bead is driven in 0.5 μm steps in a grid-like pattern through the field of view, using a piezo microscopy stage. In each position of the bead, its position is determined with nanometric precision, creating 312 fiduciary point pairs. From these pairs, the mapping function is inferred using a local weighted mean approach, which accounts for potential local distortions [42]. This process is depicted on the left in Figure 4.2.

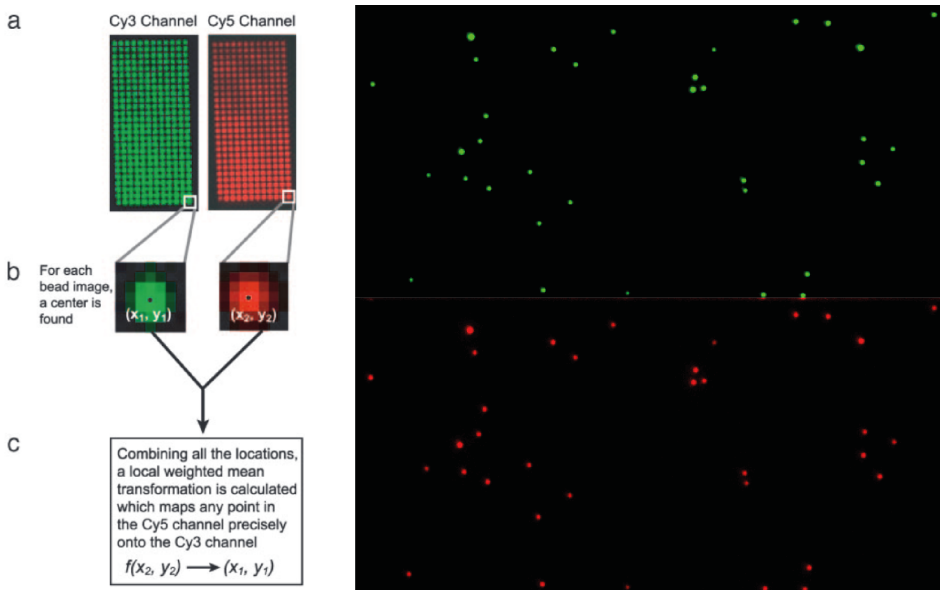

Figure 4.2: **Left:** Mapping two spectrally separated channels onto each other. A fiduciary grid is created by driving a multi-fluorescent bead through the field of view. Each position of the bead is subsequently determined in either channel by fitting the PSF. The resulting point pairs are used to calculate a mapping function. Adapted from [19], with permission from the National Academy of Sciences, USA, ©2005. **Right:** An exemplary two-channel recording of a sample of multi-fluorescent beads randomly distributed on the cover slip surface.

For the purposes of this chapter, the SHREC methodology was slightly altered, in order to be useable with the experimental set-up available. In lieu of a fiduciary grid of bead positions, multiple random bead distributions were used. This approach does not require a high-precision automated microscopy stage, thus making SHREC possible on simpler systems than that used by Churchman and colleagues. A representative two-channel view of such a random bead distribution is depicted on the right in Figure 4.2.

To test this alternative method for acquiring fiduciary point pairs, 18 bead distributions were recorded, each containing on average 27 beads. A varying number of these distributions (between one and 18), and thus a varying number of bead positions, was used to calculate a mapping function. The resulting mapping functions were tested on six in-

dependent recordings of beads. The resulting relative coordinates $x_c = x_{channel\,1} - x_{channel\,2}$ and $y_c = y_{channel\,1} - y_{channel\,2}$ of the approximately 140 beads were recorded in histograms, and Gaussians fitted to these histograms, as shown in Figure 4.3.

These analyses reveal that generally, a higher density of fiducial point pairs reduces both systematic and stochastic deviations from the expectation of $x_c = y_c = 0$. A mapping function created with a low pair density carries the risk of averaging out too many local distortions. If a sample is then in an area far away from a fiducial point pair used to create the map, a relatively high deviation can be expected. Increased pair densities reduce this risk, with the use of 231 and 489 fiducial pairs both showing generally good colocalisation and few outliers. Subsequently, those experiments presented below relying on this method were usually conducted using mapping functions based on approximately 250 fiducial pairs recorded directly prior to the main experiment. Further details on the mapping procedure are described in Appendix A.1.5.

4.2.3 *Suppressing quantum dot blinking*

For reasons still not fully understood, fluorescent quantum dots display spontaneous intervals of emission intermittency, also known as blinking [35]. When using quantum dots as labels to track micro- or nanoscale motion, these unpredictable dark periods can make significant aspects of the process go unrecorded, thus requiring a greater number of repeated experiments to construct a complete picture. In parts of this work, where two differently coloured quantum dots are used in conjunction to measure distances and resolve nanoscale motion, the potentially detrimental effects of this blinking are increased further, as both quantum dots must

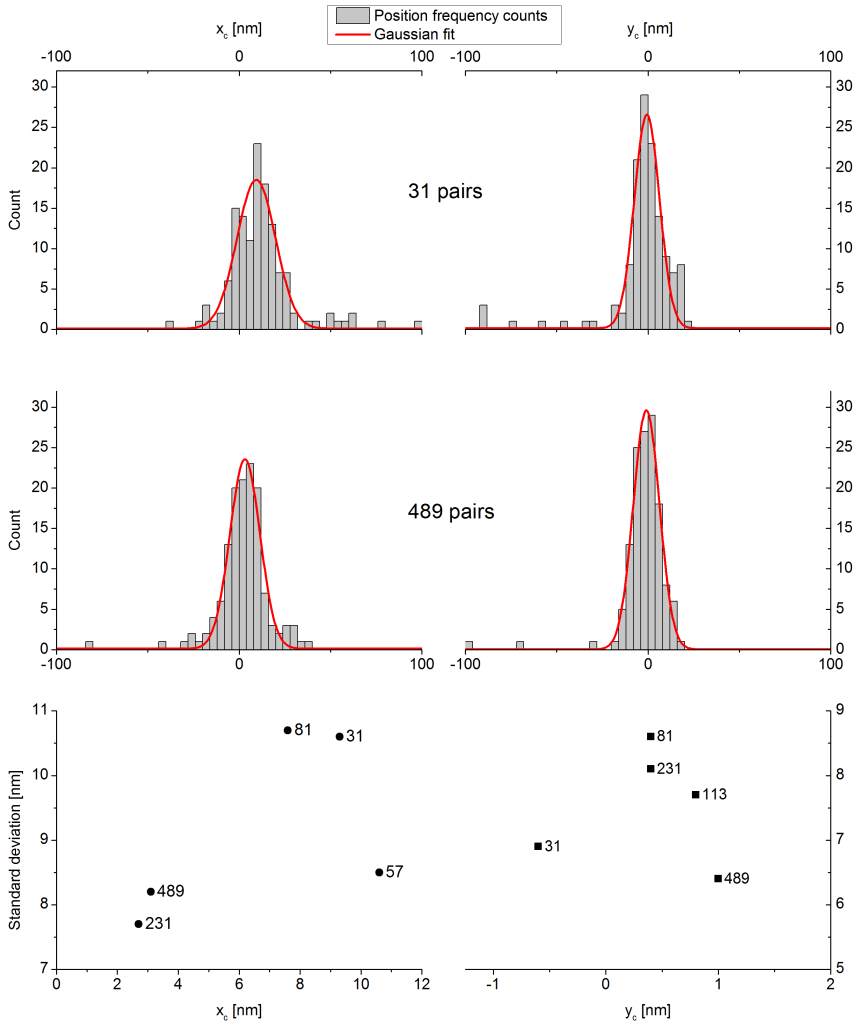

Figure 4.3: Analysis of mapping accuracy and precision depending on number of fiduciary bead pairs. Six mapping functions were created based on accumulated bead distributions containing between 31 and 489 bead pairs. Subsequently, each of these functions is used to measure the distance in independent recordings of 135 beads, expected to be zero. In the top plots, the resulting distributions for the mapping function based on 31 pairs are depicted. Fitting a Gaussian to these distributions yields a mean and a standard deviation. The same procedure is performed for the other maps; the resulting distributions for the map based on 489 pairs is shown in the central plots. Outliers for which either $|x_c| > 100\,\text{nm}$ or $|y_c| > 100\,\text{nm}$ are not shown in the distributions, but were included in the fits. For the x and y distributions based on the 31-bead-pair map, there were 2 and 3 such outliers, respectively, and 1 in both distributions based on the 489-bead-pair map. The bottom plots show the mean for x and y and their respective standard deviations for all maps tested. Point labels are according to the number of bead pairs used to create the respective mapping function.

be emitting simultaneously to allow for a distance measurement between them.

Hohng and Ha showed that this blinking can be suppressed for some quantum dots by adding beta-mercaptoethanol (BME) to the buffer in which the quantum dots are observed [48]. This was tested for both types of quantum dots used in this work and found to have a similar impact.

For one type of quantum dot, this phenomenon was studied in detail. It was determined that in the presence of 1 % v/v BME in the regular microscopy buffer, a quantum dot will be found in an emitting state with a probability of 90 % when measuring at a random time, whereas the probability will merely be 45 % when no BME is used. The probability distributions displayed in Figure 4.4 show that the effect of BME on the length of the emitting periods of quantum dots is substantial, whereas the length of the non-emitting periods is not affected. This increase in signal stability is not always desirable, however: As will be shown in section 4.3.2, quantum dot blinking can also be exploited for a different approach to super-resolution microscopy.

4.2.4 *Surface treatment*

The flow chambers used in microscopy experiments were prepared to allow for controlled binding of DNA tiles. This was achieved by washing the flow chamber surfaces with bovine serum albumin (BSA) and biotinylated BSA, and then successively streptavidin and biotin-functionalised DNA oligonucleotides prior to the addition of DNA tile samples. BSA primarily served as a blocking agent against non-specific binding of both quantum dots and DNA tiles. The biotinylated BSA provided binding sites for streptavidin, which in turn allows for binding of biotinylated DNA

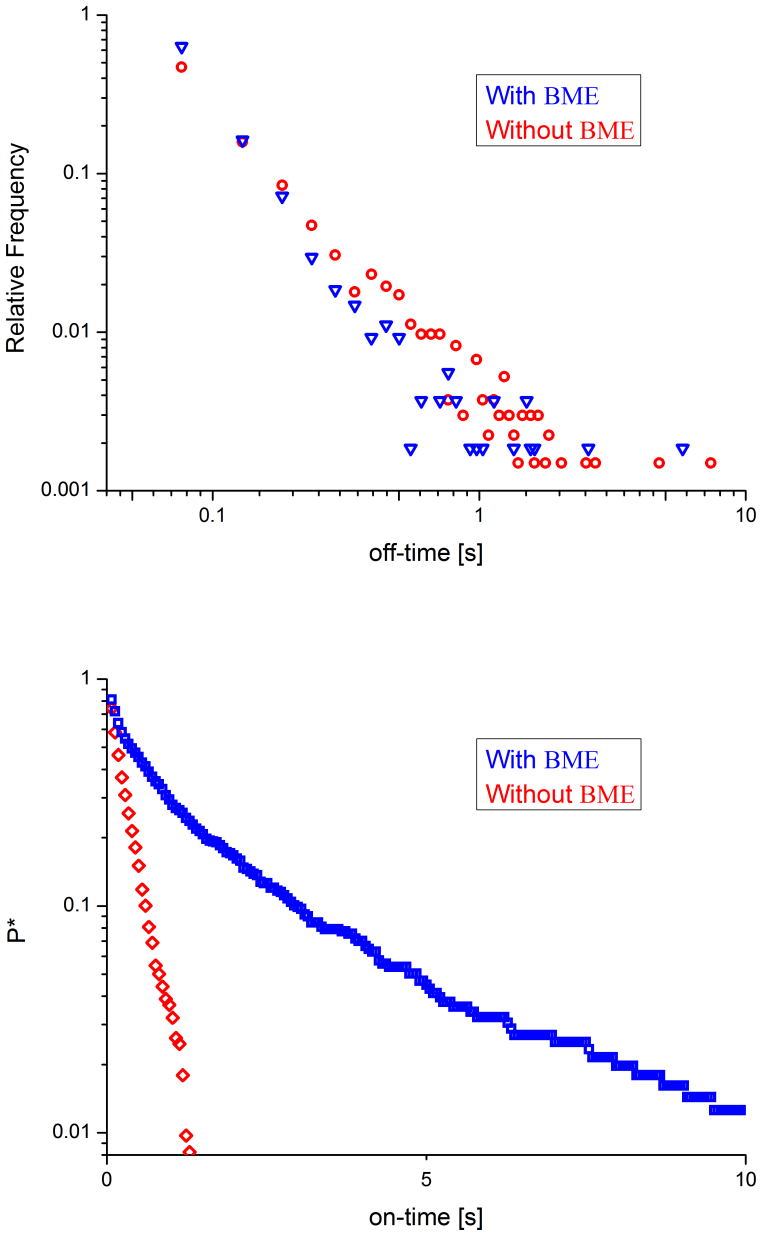

Figure 4.4: Quantification of quantum dot blinking and influence of BME. **Top:** Relative frequency of 'off-times', i.e. lengths of periods in which the quantum dot does not emit. **Bottom:** Probability P^* of a quantum dot remaining in an emitting period after a specified 'on-time'.

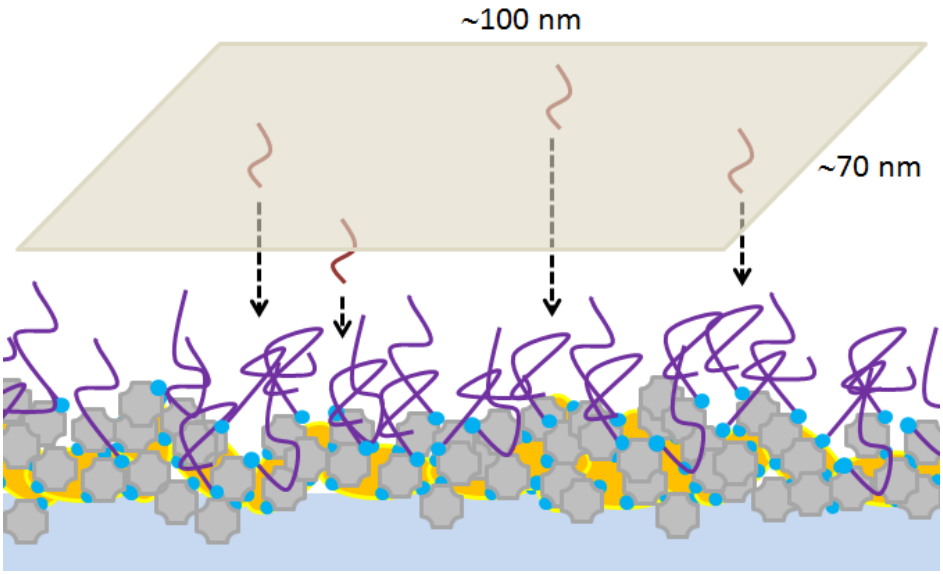

Figure 4.5: Schematic representation of an origami tile binding specifically to the cover slip surface. The surface is coated with BSA (yellow), some of which is functionalised with biotin (blue). Streptavidin (grey) can bind to the surface via the biotinylated BSA, and thus provides surface binding sites to biotinylated DNA anchor strands (purple). These anchor strands are complementary to four strands protruding from the tile surface (brown).

strands, which serve as anchoring for the DNA tiles. This surface treatment protocol is depicted schematically in Figure 4.5.

In contrast to simple adsorption of DNA tiles to an untreated glass surface, the binding to a surface treated with above procedure displayed two important characteristics. Firstly, DNA tile binding was found to be highly specific, i.e. only structures including staple strands complementary to the DNA anchors bound to the surface to an appreciable extent (Figure 4.6). Binding of tiles without the ability to hybridise these anchors was reduced by several orders of magnitude. Including the anchor-binding staples in a way that they are all on the same side of the tile, as in Figure 4.5, aims to reduce interference of the surface with functional tile components on the other side.

Secondly, this procedure was shown to reduce surface binding of quantum dots. Particularly when quantum dots are used to track the dynamic

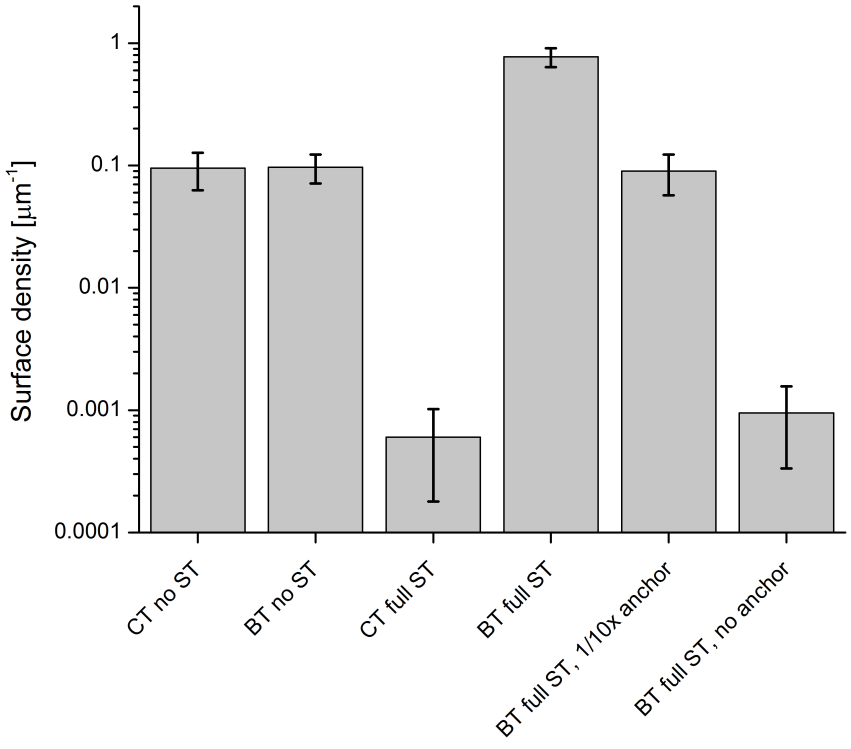

Figure 4.6: Surface density of DNA origami tiles depending on surface treatment (ST). No surface treatment denotes that a sample of tiles was added to an otherwise unprepared flow chamber, which, after 5 minutes of sample incubation, was washed with microscopy buffer containing 1× Sybr Gold to stain any tiles on the surface. This yields similar surface densities for both the control tile, which has no surface binding staples and the binding tile, which incorporates such specialised staples. Subjecting the flow chamber to a full surface treatment including BSA, biotinylated BSA, streptavidin and DNA anchor strand, reduces density of the control tile by two orders of magnitude, whereas the density of the binding tile increases by one order of magnitude. Reducing anchor concentration impairs immobilisation of the binding tile, and omitting the anchor strand entirely brings the surface density down to similarly low levels as the control tile.

behaviour of a DNA nanodevice (below), such non-specific interaction of the quantum dots with the surface is a highly undesirable quality, as strong interaction with nearby surfaces may hinder or prevent entirely the motion of the device. This was also found to be useful for static structures, as the proportion of fully assembled structures relative to isolated quantum dots in each field of view was increased, thus improving experimental efficiency. Many of the results presented below were achieved with an evolving choice of parameters for the surface treatment. Appendix A.1.5 offers a complete compilation of those parameters for each experiment presented here.

4.3 DISTANCE MEASUREMENTS ON A STATIC DNA STRUCTURE

4.3.1 *Using two-channel colocalisation*

The two-channel mapping method was initially used to study a static sample, i.e. a DNA tile in which two staples in the corners of the tile are biotinylated (Figure 4.7). Dual-labelling of this tile was achieved by adding both red and green quantum dots (i.e. with peak emission at 655 nm and 565 nm, respectively) in 2 : 1 excess simultaneously and in the presence of 1 mg/ml BSA to a freshly prepared tile sample, and incubating at room temperature for 30 min. This strategy was intended to lead to ~50 % of tiles carrying two monochrome labels. In experiments, it was found to yield a reasonable ratio of dual-labelled structures and is considerably simpler than post-assembly integration of labelled staples, a strategy described below. An ion exchange protocol based on work by Wickham [141] and Carstairs et al. [14] was used to wash out quantum dots not attached to a tile. The relative amount of dual-labelled structures

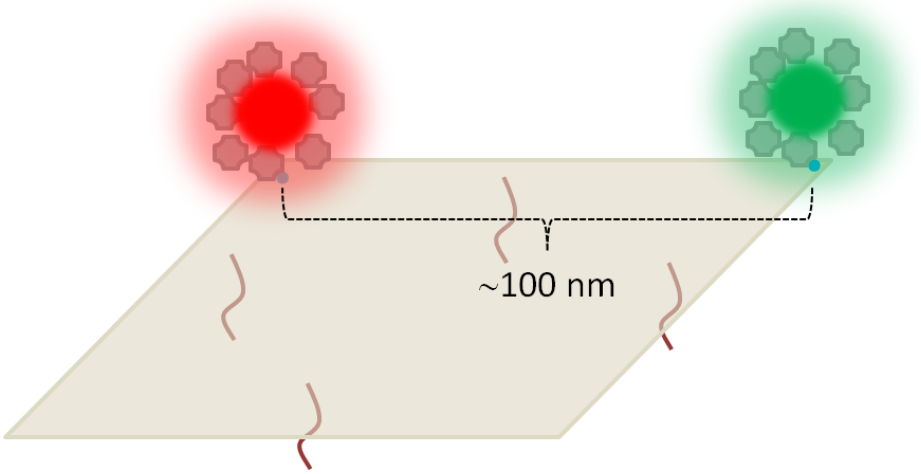

Figure 4.7: Schematic depiction of the tile design for two-channel colocalisation experiments. By using two biotinylated staple strands to prepare the tile, two binding sites for quantum dots are created, approximately 100 nm apart from each other.

visible under the microscope was usually found to be between $30 - 40\,\%$ of all fluorescent signals visible in any field of view.

Several fields of view of the sample were recorded for 50 frames with 1 s integration time each. Upon channel mapping, PSFs of signals showing fluorescence in both channels (representing a structure carrying both types of quantum dots) were fitted with two-dimensional Gaussians. Frames in which either signal was dim or yielded a poor fit were not included in the further analysis. This fitting routine yielded up to 50 relative positions for each structure.

An unexpectedly wide variety of distances between these relative positions was found, examples of which are shown in Figure 4.8. The quality of the individual position distributions was determined by two factors, accuracy and precision. An individual distribution was deemed accurate if the mean relative position had a distance of (100 ± 10) nm from the origin, as the tile was designed to place the two labels at a distance of ~100 nm from each other. The standard deviations from the mean relative

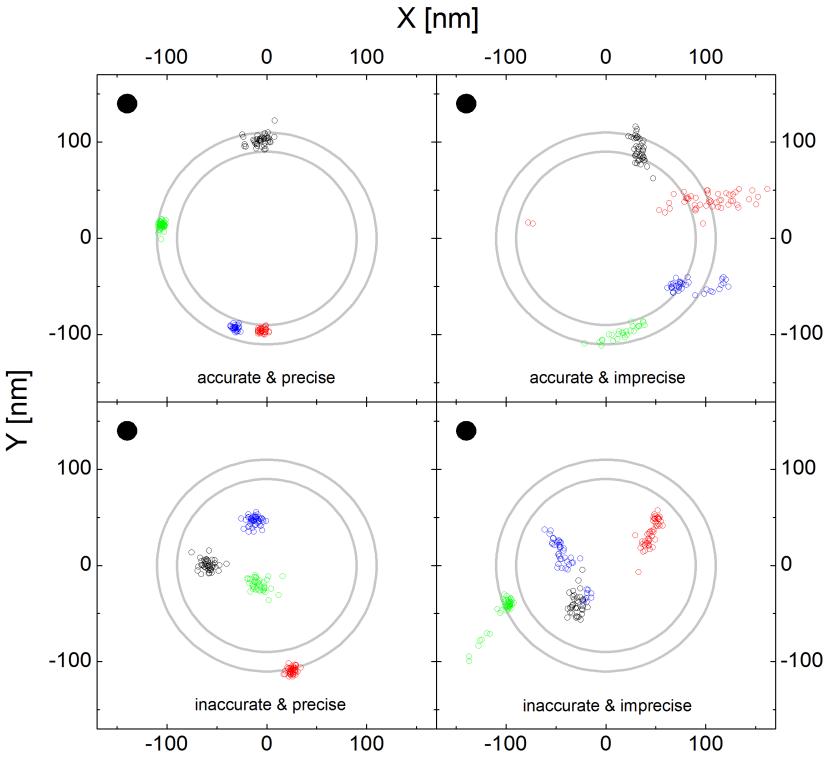

Figure 4.8: Examples of relative position data measured on a dual-labelled tile. Each plot shows four exemplary data sets. The plots are categorised by accuracy and precision, as defined in the text. The grey rings depict the area in which points have a distance between 90 nm and 110 nm from the origin, symbolising the criterion chosen for accuracy. The solid black circles in the top left corner of each plot have a radius equivalent to 10 nm, giving an impression of the criterion chosen for precision.

position in both dimensions x and y served as a basis for the precision criterion

$$\sigma_{xy} = \sqrt{\sigma_x^2 + \sigma_y^2} < 10\,\text{nm}.$$

The distances from the origin of the mean relative position from each individual distribution was integrated into a collective distribution (Figure 4.9, top left). It was determined that the accuracy and precision criteria are slightly correlated. Excluding imprecise individual distributions from the collective distribution of distances yields a refined, narrower distribution in which 38 % of the structures imaged fulfill the criterion for accuracy, compared to 30 % in the raw collective distribution. Overall, the resulting collective distribution is in accordance with the structure design, although its variance could be improved upon.

There are several possible reasons for why the variance in this distribution is greater than anticipated. It is, for example, conceivable that the quantum dots bind in variable conformations to the tile. Due to the size of a dot (~20 nm), the position of its centre is not perfectly determined by the location of the biotinylated staple strand it binds to, as can be assumed for small dye molecules. Variable conformations of the origami tiles themselves on the surface might also play a role. Overall, however, a significant ratio of individual distance measurements suggests that the use of quantum dots indeed allows for high-precision colocalisation on DNA nanostructures.

4.3.2 *Using quantum dot blinking*

Blinking is one of the few undesirable fluorescence properties associated with quantum dots. If constant emission is required for a particular experiment, other fluorophores, or types of quantum dots known to reduce

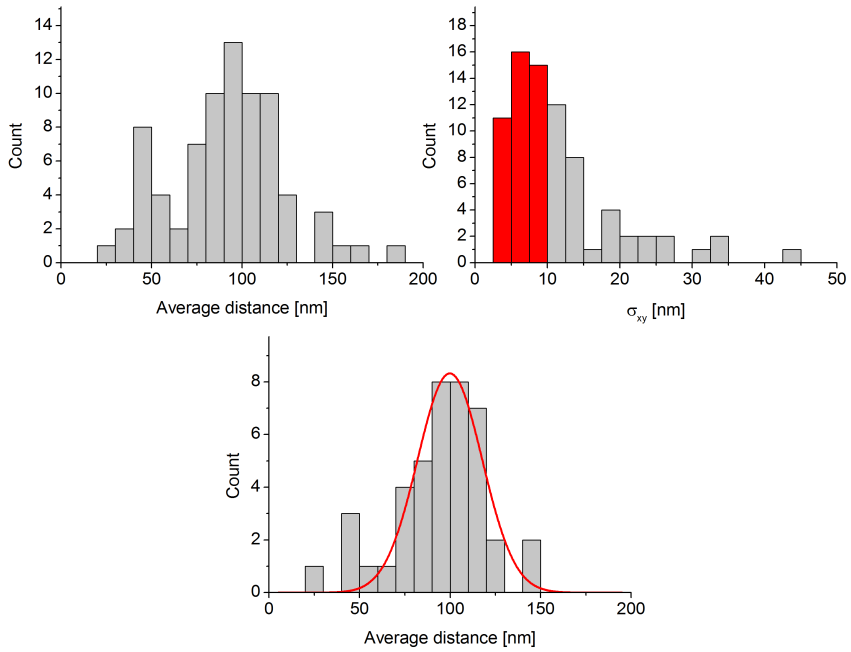

Figure 4.9: **Top left:** raw collective distribution of distances measured between red and green quantum dots attached to two corners of a DNA origami tile $(n = 77)$. **Top right:** collective distribution of the deviation from the mean σ_{xy} in each individual distribution of positions. Considering only those individual distributions for which $\sigma_{xy} < 10\,\text{nm}$ (red columns) reduces the top left distribution to the collective distribution at the **bottom** $(n = 42)$. Fitting this 'refined' collective distribution with a Gaussian yields a mean of 100 nm with a standard deviation of 18 nm.

blinking in the presence of BME, must be used. Here, the application of a method to measure sub-diffraction limit distances is presented which *requires* the fluorescent labels to display fluorescence intermittency.

Lidke et al. first explored the potential of quantum dot blinking statistics for super-resolution microscopy [68]. Wang and co-workers recently provided experimental proof that microscopy based on quantum-dot blinking can be used to study three-dimensional quantum dot arrangements on polystyrene microbeads [137]. Note that the aforementioned DNA-PAINT method relies on the same fundamental principle (depicted in Figure 1.10), except that it is not the fluorescence itself that is being controlled. Instead, fluorescence signals from different emitters are temporally separated through transient binding close to the cover slip surface, the only area of the sample in which excitation of the fluorophores is possible. Using quantum dots, neither active control nor transient binding are required to achieve temporally separate emission, as they are constantly switching between emitting and non-emitting states.

In the work presented here, this was used to resolve sub-diffraction-limit distances on DNA origami tiles between two quantum dots emitting at the same wavelength of 655 nm. As before, the labels' spatial relation was controlled by including two biotinylated tile staples in the origami assembly process. TIRF microscopy recordings were made, in which the PSFs were fitted with two-dimensional Gaussians. For those tiles which carried two functional quantum dots, the resulting distribution of determined locations was found to be stretched in one direction. The top left plot in Figure 4.10 shows such a distribution after PSF fits from frames in which both quantum dots were off were excluded.

Such a distribution gives an impression of the distance between the two labels, but an accurate measurement is not possible from this information alone. If both of the quantum dots are emitting simultaneously,

a PSF fit was assumed to yield a localisation between the actual positions of the labels. A better understanding of this can be derived from the intensity profile of the recording in the centre of Figure 4.10.

This profile can roughly be divided into three regimes. In the first regime, neither of the quantum dots was emitting. This regime is represented by those intensity values close to the horizontal axis of the plot. The fits derived from the frames in the first regime were excluded from the position distribution *a priori.* The intensity values around the dashed line represent a second regime, which is assumed to mostly represent frames of the recording in which only one of the two quantum dots is emitting. Frames in which both quantum dots emit simultaneously contribute high-intensity spikes, the third regime.

Such frames were excluded by introducing a further brightness cut-off filter, only allowing PSFs up to a certain intensity to contribute to the position distribution. In the recording in Figure 4.10, this cut-off intensity is represented by the dashed line in the intensity profile. In the filtered position distribution (top right), two separate clusters become apparent as a consequence. It should be noted that this filtered position distribution served as a guide to choosing the cut-off intensity rather than the intensity profile itself, as the different regimes were not always as clearly discernible as in the example shown in Figure 4.10. This is mainly attributed to quantum dots switching states during, rather than in between, the recording of single frames and to a considerable inconsistency in brightness between different quantum dots of the same type.

There are numerous ways to determine the distance between the labels from the filtered position distribution. Here, the single points in the distribution were projected onto the axis connecting the two estimated quantum dot locations, yielding a one-dimensional histogram as in the bottom right plot of Figure 4.10. As expected, two peaks are dis-

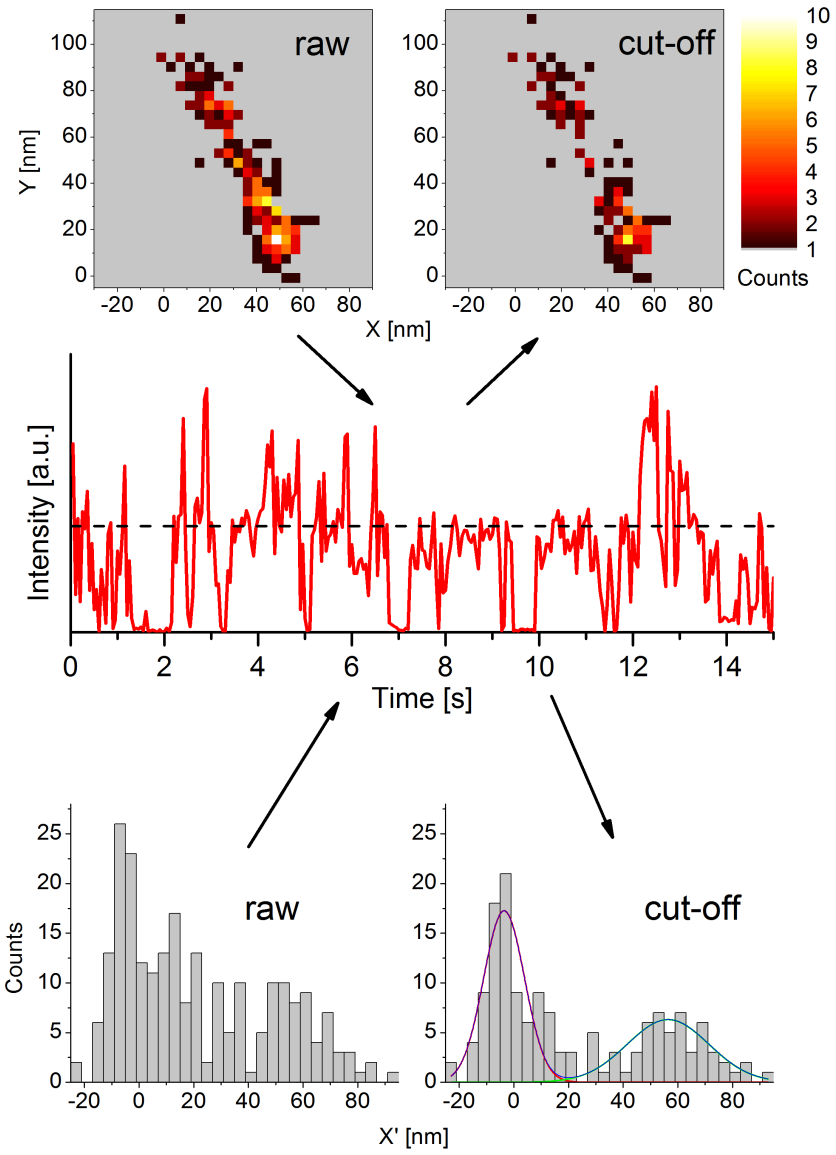

Figure 4.10: An exemplary measurement of a fixed distance on a DNA origami tile using the blinking of two quantum dots of the same colour. **Top left:** The raw distribution of PSF centres is not suitable to determine the distance between the two labels. **Bottom left:** Projecting this distribution onto one dimension shows few discernible features. **Centre:** An intensity profile reveals states in which both quantum dots are emitting simultaneously. **Top right:** By excluding these states, i.e. all PSFs with a brightness above the cut-off (dashed line in the intensity profile), the distinct locations of the quantum dots become apparent. **Bottom right:** Fitting Gaussians to a one-dimensional projection of the resulting distribution reveals the distance between the two labels to be ~60 nm. The distance had been designed to be ~50 nm.

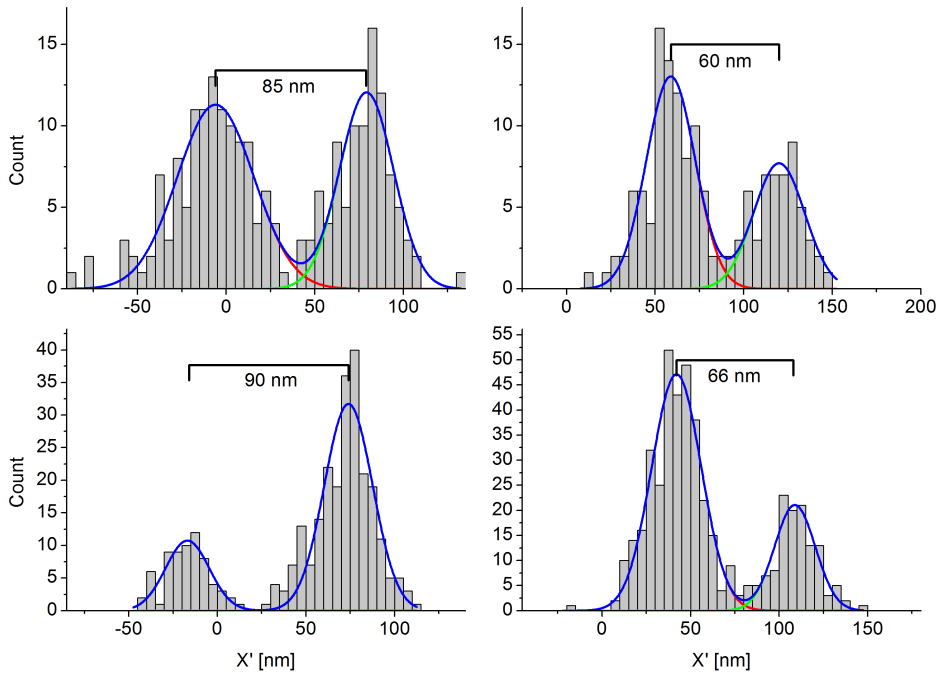

Figure 4.11: Using quantum dot blinking to determine a static distance on the tile design previously introduced in subsection 4.3.1. Due to the increased distance between the quantum dots, the peaks in the distributions of signal positions can be separated more clearly as in the example shown in Figure 4.10.

cernible in this distribution, which was fitted with two Gaussians. Note that performing the same projection for the original distribution yields a histogram in which no clear peaks are discernible (bottom left of Figure 4.10).

As a further proof-of-principle, this method was also successfully tested on the tile design used previously in the two-colour measurements in section 4.3. Examples of this measurement are shown in Figure 4.11. Like the measurements in the previous section, some of the samples deviated noticeably from the designed distance of ~100 nm, an observation which will be discussed in more detail below.

Overall, this method was found to be an interesting alternative to measurements using spectrally separate labels. The Gaussians fitted to the projected distributions of signal positions were commonly found to have full width at half maximums (FWHMs) of around 35 nm, which gives an

estimate for the smallest distance resolvable with this method. However, this method has some noteworthy weaknesses as well, when comparing to the colocalisation technique used in subsection 4.3.1. The most striking of these is that there is no inherent control for drift. Whereas non-rotational drift cancels out when determining the spatial relationship between differently coloured quantum dots over time, this is not the case here. Only experimental setups with a highly sensitive external control for drift would allow for the quantum dot intermittency method to be applicable over extended periods of time.

A further challenge is the exclusion of structures that are labelled incompletely. These structures can be immediately excluded from analysis when using differently coloured labels, simply because such a structure will only be visible in one channel. When using monochrome labels, however, all recorded signals must be evaluated first to determine whether they comprise one or two fluorophores. A potential way to make data analysis more efficient might include a pre-analysis filter, in which signal PSFs are excluded if they do not display a certain degree of eccentricity. Using PSF brightness as an initial filtering parameter represents a further option, but would likely be less reliable as quantum dots were found to have variable emission intensity, and may also not be illuminated homogenously across the entire field of view.

4.4 MEASUREMENTS OF A DYNAMIC DNA DEVICE

4.4.1 *Dynamic measurements*

After establishing that differently coloured quantum dots fixed to static points on a DNA origami tile can be colocalised with high precision, attempts were made to transfer this procedure onto a linear DNA motor

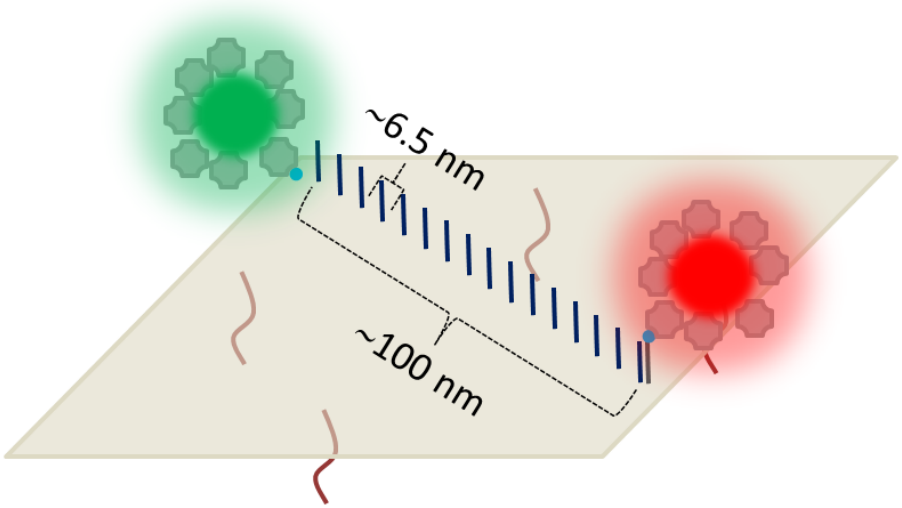

Figure 4.12: Schematic depiction of the tile design for measurements of a DNA motor system. The tile initially contains one biotinylated staple in one of its corners, as well as 15 staples with stator extensions providing a linear track. The biotinylated, quantum-dot-labelled motor is introduced to this tile in a duplex with the starting stator. Motion of the motor can be discerned by a decreasing distance between the two quantum dot labels.

presented by Bath and colleagues ([7], Figure 1.6B). This device had previously been shown to operate reliably on a DNA origami tile track, using both bulk fluorimetry and high-speed AFM by Wickham and colleagues [143]. In their work, this motor moved across a tile on a track with a total length of ~100 nm.

Considering both the longevity and high SNR of quantum dot emission, it could be expected that operation of this device should also be discernible by attaching one label to the motor itself and another to the tile, serving as a point of reference. For this purpose, a tile very similar to that presented by Wickham and colleagues was designed which hosts a 16-stator track (Figure 4.12). The motor, initially bound to the first stator, is labelled with a quantum dot. Adjacent to the last stator is an independent site for a further quantum dot label. Four single strands protruding from the bottom of the tile were included to immobilise the tile on the flow chamber surface, as described in section 4.2.

In contrast to plain tiles, the assembly of a labelled tile carrying a labelled motor is slightly more involved. The tile sample was prepared with a straightforward annealing protocol, containing all components apart from the reference quantum dot, the stator on which the motor starts and the motor itself. If complete tiles were annealed with the motor in solution, control over the motor starting point would be compromised [141]. After annealing and staple removal, a stoichiometric amount of green quantum dots was added. Separately, a previously annealed duplex of biotinylated motor and first stator was mixed with a stoichiometric amount of red quantum dots at 500 nM. Both samples were incubated in the presence of BSA for 1 h at room temperature. In order to include the missing stator carrying the labelled motor in the tile, both samples were combined and incubated for 1 h at 37° C as described by Wickham et al. [143]. Upon application to a surface-treated flow chamber and preliminary microscopy imaging, approximately $10 - 20$ % of signals in one channel were found to have a corresponding signal in the other.

After addition of 1 : 100 v/v dilution of Nt.BbvCI nicking endonuclease, which drives motor operation, and sealing of the flow chamber, initial measurements were performed by recording 20-frame movies of one particular field of view in fixed time intervals. In between recordings, sample excitation was interrupted to reduce the likelihood of photobleaching.

The first experiments of this kind revealed no significant motor motion over the course of 80 minutes. An exemplary recording of this experimental series is depicted in Figure 4.13. For the particular experiment that this recording is derived from, motors were generally found to be in a position consistent with being located at the beginning of the track. For follow-up experiments, however, the distribution of starting posi-

tions often showed much greater variance, although the same protocol was followed to assemble, label and image the track tiles.

An increase in enzyme concentration to 1 : 50 v/v led to a few measurements suggestive of the intended motor function, i.e. recordings in which the position averages of each 20-frame movie were found to move towards the origin, representing the end of the track. Only one such measurement appears to depict motor motion along the complete track, however (Figure 4.14). Although the colocalisation in that particular measurement is less precise as had been shown to be possible previously, there are several factors which suggest that this data indeed represents motor operation. Firstly, the data points recorded are consistent with the expected position of the track. The overall distribution of motor positions is approximately linear. Fitting a straight line to the distances from the origin plotted against time reveals a slope of $-3.2\,\mathrm{nm\ min^{-1}} \mathrel{\widehat{=}} -0.053\,\mathrm{nm\ s^{-1}}$, which is congruent with the speed of ~$0.1\,\mathrm{nm\ s^{-1}}$ determined by Wickham and colleagues [143].

The vast majority of motors were still found to be immobile, however. A further five-fold increase in enzyme concentration aimed to allow for further measurements of motor operation, but instead yielded considerable sample degradation. This degradation is assumed to be due to non-specific cleaving activity of the enzyme, which, according to the manufacturer, can occur at high ($\gtrsim 1 : 10\,\mathrm{v/v}$) enzyme concentration.

4.4.2 *Static measurements*

In previous works by Wickham and colleagues, the device studied here had only been proven to operate reliably without a separate cargo [143, 142]. Considering the lack of successful live measurements of motor mov-

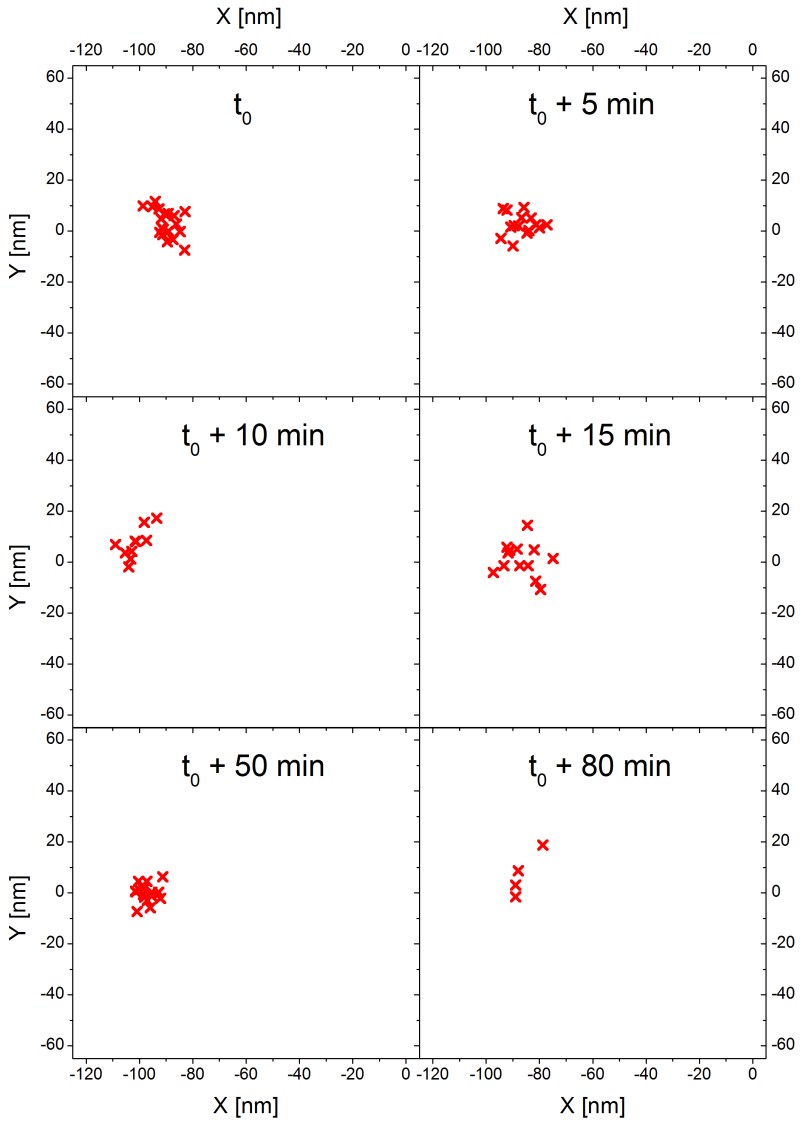

Figure 4.13: Exemplary data for a motor measurement as described in the text. The motor is found in a starting position approximately 90 nm from the origin, which is in good agreement with the design of the system. Over the course of 80 minutes, the motor does not, however, move discernibly towards the origin. High colocalisation precision was preserved.

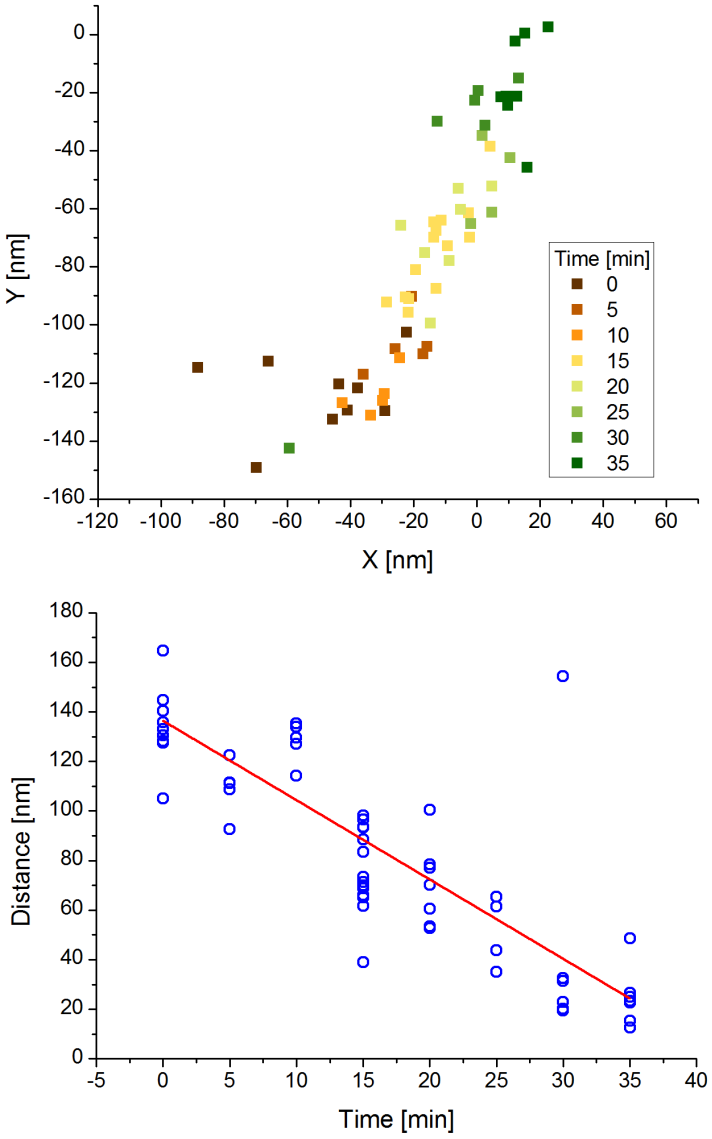

Figure 4.14: A recording of an enzyme-driven motor in motion. **Top:** Motor positions relative to a reference quantum dot. Colour indicates time from the beginning of the observation. As time progresses, the motor appears to move towards the origin. **Bottom:** A plot of the distance $\left(\sqrt{X^2+Y^2}\right)$ between origin and motor for the eight 20-frame recordings against time (blue circles). A linear fit of this data reveals an estimate for motor speed of 0.053 nm s^{-1}, consistent with previously determined behaviour.

ing its quantum dot cargo along the track, the following experiments tested whether the motor is capable of transporting such a cargo at all.

The same tile as for the dynamic measurements was assembled, and part of the sample observed immediately to determine the initial distance between motor and reference label site. The rest of the sample was incubated at room temperature for 24 hours with Nt.BbvCI, and subsequently subjected to the same imaging procedure. The resulting collective distance distributions are depicted at the top of Figure 4.15.

The collective distribution for the freshly assembled sample is broader than anticipated, even though individual structures displaying very widely scattered position distributions $\left(\sigma_{xy} \geq 15\,\mathrm{nm}\right)$ were excluded from further analysis. This may in part be due to unspecific loading of the motor. Although the motor strand is meant to be integrated into the tile as part of a duplex with the starting stator ~100 nm from the reference label, as described above, it is conceivable that there is still unhybridised motor in solution which can bind to an arbitrary stator. A further possibility is that even when the motor is integrated correctly, the elevated temperature required for motor-stator duplex integration may encourage motor diffusion along the track without presence of a nicking enzyme, prior to imaging.

The collective distribution for the sample incubated with nicking enzyme is slightly different. Whereas initially, only 31 % of the tile assemblies were determined to have a motor-reference label distance of $\leq$ 50 nm, this became 44 % after incubation with enzyme. However, a great number of individual position distributions was still determined to be very far from the end of the track. This compares particularly unfavourably to the collective distribution at the bottom of Figure 4.15. This distribution was acquired with a modified tile, in which the motor was integrated into the tile in a duplex with the last stator of the track, and

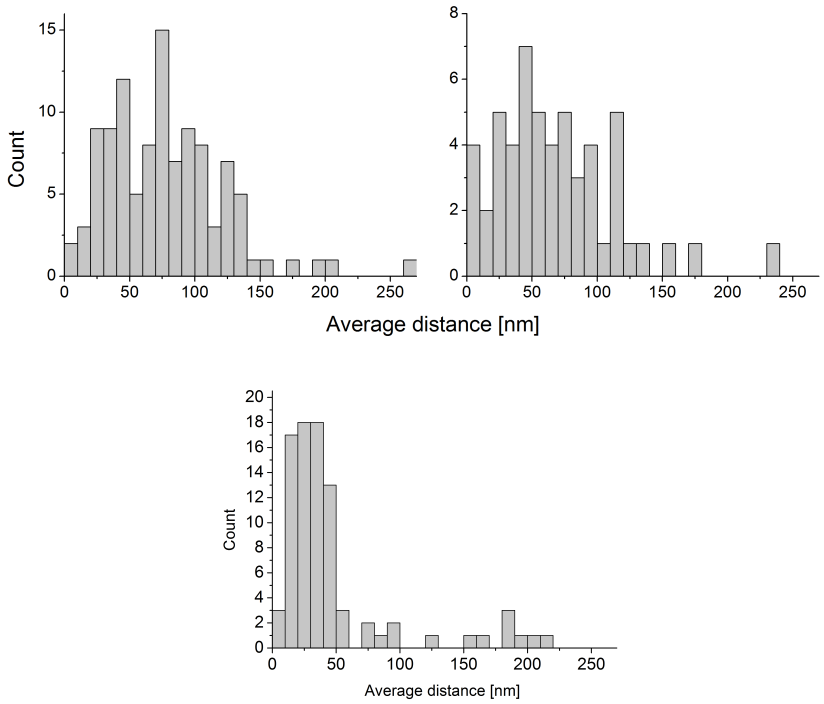

Figure 4.15: Collective distance distributions for tiles carrying a motor. **Top left:** Distance distribution directly after tile assembly $(n = 85)$. **Top right:** The distribution measured after 24 h of incubation with nicking enzyme in a sample tube, $(n = 36)$. **Bottom:** This distribution was derived from a sample in which only the last stator was present on the tile $(n = 77)$. Individual position distributions with very poor precision $(\sigma_{xy} \geq 15\,\text{nm})$ were excluded from all collective distributions.

all other stators having been replaced with unmodified staple strands. This distribution, in which 87 % of the motor-reference label distances are $\leq 50\,\text{nm}$, is a good indicator of what the collective position distribution of the complete system should optimally look like once the motor has reached the end of the track. This data suggests that the quantum dot may offer a greater impediment to motor operation than originally anticipated.

4.5 CONCLUSIONS & OUTLOOK

This chapter presented a preliminary study in the use of quantum dots as fluorescent labels in single-molecule microscopy of synthetic DNA nanodevices. It was shown that even with a relatively simple microscopy setup, single quantum dots can be localised and colocalised with a precision on the nanometre scale for extended periods of time. Previously published findings on the suppression of the characteristic blinking of quantum dots were confirmed with the type of fluorophore used here. However, intermittent emission was also shown to be potentially useful, as it allowed for the sub-diffraction limit resolution of two quantum dots of the same colour on a DNA origami tile.

Further experiments were performed using two types of quantum dots with separable emission spectra. Colocalisation of the two corresponding imaging channels was achieved using a simplified version of an established method based on multi-fluorescent beads. It was demonstrated that despite the experimental simplification of the original method, which makes it applicable on relatively basic microscope setups, mapping functions were of a similar quality. Two-channel recordings of dual-labelled DNA origami tiles were performed using such mapping functions. Despite the high mapping function quality, the distances determined were scattered broadly, i.e. with a standard deviation of 18 nm around the expected value of 100 nm. This is assumed to be due to the sample studied rather than methodology.

A possible source of the variability in measured distances is the potential conformational freedom of the quantum dots on the tiles. It may be possible to restrict this through the use of several, instead of just one, biotinylated staple strand per labelling site. It may also be feasible to use more sophisticated DNA nanostructures as substrates for these

types of experiments. Two-dimensional DNA origami assemblies are assumed to be structurally quite flexible, considering the previously shown possibility to fold [27] or roll [32] two-dimensional origami motifs into three-dimensional structures. Such three-dimensional structures appear to have greater rigidity on the length scales studied here [96], and may thus provide a more homogeneous distribution of label-to-label distances in future experiments. Despite the shortcomings of the collective distance measurements on DNA tiles in section 4.3, however, it is worth remembering that the quality of many *individual* measurements, both in terms of precision and accuracy, was sufficiently high to suggest that generally, this method may allow for detailed observations on the length- and time scales typical for a broad range of DNA nanodevices.

Nanometric colocalisation was also applied to study a dynamic DNA nanomotor. With this system, the precision and accuracy of quantum dot colocalisation was shown to be stable over extended periods of time. However, data suggesting that this system is capable of transporting a quantum dot was severely limited. Of several dozen live recordings of such a motor, only one showed the expected behaviour in all aspects. A separate study, in which the motor system was allowed to operate in solution rather than on a surface suggested that operation of a motor labelled with a quantum dot is severely impaired. This may, for example, be due to quantum dots providing a steric hindrance to enzyme operation or the TMSD process required for directed motion of the system studied.

Both of these issues, if confirmed to be the cause of the negligible motor movement, may be alleviated by changing the design of the system. Whereas in the design used here, the quantum dots were attached directly to the motor sequence, the introduction of a poly-T or PEG linker may remove the quantum dot far enough to reduce interference with

motor operation to negligible levels. In this context it is worth noting that Cha and co-workers have recently used a similar methodology to prove that a different type of DNA nanomotor is capable of transporting a quantum dot along a carbon nanotube; they did not, however, make use of nanometric localisation to investigate detailed motor properties [15].

It is also conceivable to implement a method similar to DNA-PAINT using quantum dots instead of regular fluorophores. Through this, it may be possible to achieve very high localisation precision in short periods of time, whilst binding of the quantum dot would only be transient. Considering the remarkable measurements demonstrated with regular DNA-PAINT, an implementation using quantum dots could provide a significant step forwards in the study and application of dynamic DNA nanodevices in single-molecule fluorescence microscopy. The most obvious advantage compared to traditional DNA-PAINT would be that localisation precision could be improved. A further advantage is that quantum dots of different emission wavelengths can all be excited efficiently with the same, short-wavelength laser; whereas using a variety of regular fluorophores in fluorescence microscopy usually requires very sophisticated, multi-laser excitation set-ups.

The surface treatment used in this work may also require further improvements. Although it was estimated that potential binding sites on the biotinylated BSA were blocked off almost entirely by the excess of streptavidin added in the subsequent step, individual sites may still interfere with the correct surface binding of a quantum dot-labelled tile. For similar experiments in the future, it appears advisable to use a surface protocol based on different binding mechanisms. One possible example has recently been used by Scheible and colleagues [109]. Gold spots on a glass substrate were used to bind individual DNA origami tiles via thi-

olated staple strands. This served as a basis for single-molecule fluorescence microscopy using common fluorophores. It should also be noted that Wu and colleagues recently showed that oligonucleotides protruding from one side of a tile can thread through gaps in the tile to the opposite side, which may have reduced the amount of tiles immobilised in the correct orientation in the experiments presented here. This problem may also be alleviated by using a gold-thiol immobilisation technique, since it allows for very short surface linkers which would be unable to thread through the tile [145].

Further in the future, the exploration of other types of labelling agents and imaging methods may yield deeper insights into the inner workings of synthetic molecular machines. Nanodiamonds, which show even higher photostability than quantum dots [36] could allow for very long, continuous observation with extremely high precision. Considering recent advances in the controlled binding of gold nanoparticles to synthetic DNA structures [110], and the highly successful application of such nanoparticles as light-scattering labels in dark-field microscopy analyses of biological motor proteins [147], further options for sustained nanometric measurements of DNA nanomachines are certainly conceivable. Such advances would facilitate the study of different device designs in high-throughput experiments, and would allow for detailed investigations even if such synthetic systems should reach similar speeds to biological motors.

Further advances in the preparation and functionalisation of DNA nanomachines could allow for the measurement of more subtle parameters of machine operation, such as the stall force. Quantum dots labelling such machines could be manipulated with an optical trap, whilst allowing for simultaneous, high-precision tracking [54]. It would also be conceivable to study grouped machines for emergent cooperative effects as those

described for some biological motor proteins [56]. This would not only further the understanding of DNA nanomachine operation, but may also guide their future design towards improved functionality.

5

SUMMARY & OUTLOOK

5.1 SUMMARY OF WORK PRESENTED IN THIS THESIS

The objective of this thesis was to advance the field of DNA nanotechnology through the design and investigation of novel control and observation concepts for dynamic nanodevices. The work presented here towards this objective comprises three parts:

In chapter 2, the impact of mismatched base pairs on DNA toehold-mediated strand displacement was studied. It was discovered that the rate of such displacement reactions does not necessarily depend on the presence of a mismatch, but much more on its position within the displacement domain. A proximal mismatch close to the toehold was generally found to be a strong kinetic impediment against displacement, whereas a mismatch located close to the distal end of the displacement domain made little or no difference compared to a reaction with an equivalent, but non-mismatched system. This effect, which to my knowledge has not been described before, was explored qualitatively using free-energy landscapes simulated with a coarse-grained model of DNA. These landscapes revealed that the observed dependence of rate on mismatch position is mainly due to the sequential nature of TMSD, which allows for alternative, fast displacement pathways only for distal mismatches. This model was also capable of reproducing experimentally determined relative rates quantitatively, emphasising its potential for application in a wide variety of complex DNA reactions and processes. In

a simple proof-of-principle system, the core findings of chapter 2 were used to demonstrate that the kinetics of a displacement reaction can be biased kinetically with a mismatch, whilst reaction thermodynamics can be controlled independently.

Chapter 3 offered an alternative to straight-forward TMSD by exploring the possibility of using two or more cooperatively interacting invading strands to achieve displacement. A previously published design was improved by giving individual or non-cooperating invading strands a more pronounced thermodynamic disadvantage compared to the incumbent. This improvement allowed for the scheme's implementation into greater reaction networks, capable of calculating the Boolean product of two 3×3 matrices. It was further shown that the improved reaction scheme is cascadable, i.e. that displacement reactions of this type can be linked together to increase reaction network complexity and computational capability. Reaction networks based on such cascades were used to perform fundamental calculations on directed graphs.

In order to improve experimental efficiency, most of the approximately 350 individual time-dependent fluorescence measurements were performed in a quantitative PCR machine, which was found to be readily convertible to a 96-channel fluorimeter. Though a subtle detail of this particular chapter, it is certainly worth mentioning, simply because this approach to fluorimetry allows for the simultaneous assaying of a much greater number of dynamic DNA systems than traditional approaches, thus removing a common bottleneck from fluorimetry studies of DNA nanodevices.

Chapter 4 explored single-molecule fluorescence microscopy as a tool for observing DNA nanodevices. Quantum dots were introduced as fluorescent markers, mainly because they can endure the excitation intensities necessary for nanoscale resolution for the typical time scales of dynamic nanodevice motion. For static nanostructures, they allow for high-

precision measurements in a comparatively short time. The latter was demonstrated using a DNA origami tile decorated with quantum dots in fixed positions. Both two-colour recordings with subsequent nanometric colocalisation, as well as recordings with a single type of quantum dots differentiated by their blinking were shown to return the spatial relation between the labels for distances between approximately $20 - 100\,\mathrm{nm}$. The results for dynamic devices using this methodology were ambivalent: on the one hand, a few measurements displayed behaviour commensurate with the expected device behaviour; on the other, there was considerable evidence that presence of a quantum dot impedes device operation.

5.2 IMPACT AND FUTURE WORK

The findings of chapter 2 will be of particular interest to those researchers interested in the design of dynamic DNA nanodevices and reaction networks. For them, the rational introduction of mismatches into displacement reactions may provide a new paradigm for kinetic control. The oxDNA model was shown to be a powerful tool to predict the rates of defect-modified displacement reactions, and could thus be used to rationally design isolated, and perhaps even networked reactions with certain desired features. Chapter 2 will also be of interest to those who study nucleic acid probes, for example for diagnostic purposes, as the experimental findings presented there raise concerns regarding the reliability of displacement-based probes.

A greater density of experimentally determined data points would also be desirable. The experimental methods used in chapter 3, which allow for the assaying of hundreds of displacement reactions in a matter of weeks, would allow for vast additions of data to that presented

in chapter 2 with relatively little effort[1]. More data could allow for two advances: firstly, a phenomenological rate-predicting model could be derived (comparable to that of Zhang and Winfree for different toehold exchange [156]). Secondly, it could likely be shown that combinations of different defects and toehold lengths allow for extremely high-resolution tunability of displacement rates. Both advances combined may provide an excellent resource for researchers who require precisely tailored rates, for example for DNA reaction networks.

The results presented in chapter 3 may also attract some interest. The ability of a combinatorial strand displacement-based logic to calculate reachability properties of graphs could conceivably be implemented as part of guidance system for DNA robots operating on graph-like tracks [142, 86]. With the implementation of such a system, a motor could make autonomous decisions on how to reach a certain target with the fewest number of steps, or calculate a more complex route under specified constraints. Further improvements in system fidelity would appear sensible before increasing the complexity of such a calculator.

The work in chapter 4 certainly still requires more work to find truly conclusive results. With the data presented, it is unclear whether quantum-dot based single-molecule microscopy has the potential to become a staple method for researchers interested in the mechanistic details of DNA dynamic device operation. Although the preliminary results are promising, there are concerns regarding the potentially disruptive influence quantum dots may have on such delicate nanodevices. Due to the great advantages quantum dots have over common fluorophores, a successful attempt at their integration into, for example, a linear DNA motor system, would likely be rewarded with unparalleled temporal and spatial resolu-

1 The data in chapter 2 had been acquired before fluorimetry with a PCR machine had been developed to work reliably.

tion, in turn allowing for the identification of new rules and guidelines for motor design.

5.3 THE FUTURE OF DNA NANOTECHNOLOGY

Three decades after its humble beginnings, DNA nanotechnology has in recent years begun to reach a certain level of maturity [115]. After an initial boom in nanostructures, dynamic devices and reaction networks, the field now displays a greater orientation towards technological application, as well as interdisciplinary integration. Along the way, some fantastic ideas were replaced by more realistic ones. However, those goals that were abandoned are still greatly outnumbered by regular demonstrations of newly discovered and engineered features.

Despite the inherent unpredictability of scientific research, several trends are emerging. One of them is the application of DNA nanotechnology for biological studies. As described in the introduction to this thesis, this was one of the original goals of the field, but only recently is the toolkit that synthetic DNA structures provide being applied by a broader community of researchers. Examples include the study of proteins using cryo-EM [114], the production of versatile nanopores [64, 9], and sophisticated multifluorescent labels [69].

In chemistry, DNA reaction cascades may provide a template for the one-pot synthesis of complex molecules [80]. Vipergen, one of the first commercial enterprises to rely on DNA nanotechnology, aims to use DNA-guided synthesis for drug discovery. An artificial version of the ribosome, a marvelous factory of biological machinery, provides an ambitious goal for nucleic-acid directed assembly of complex molecules. From a conceptual point of view, nanodevices that autonomously assemble copies of themselves would also be very interesting. Self-replication is a funda-

mental property of life, and has received an according level of attention, both within science and outside. On the macroscale, many important advances towards machines that can replicate themselves have been made (e.g. [45, 161]), a development that will likely be accelerated with the rise of novel manufacturing methods such as 3D printing. On the nanoscale, despite the fact that interaction and behaviour of DNA molecules are highly controllable, advances in this direction are still limited to structures with limited functionality, which require constant human intervention to self-replicate (e.g. [136]). If such devices are implemented successfully, it may even be possible to wilfully include sources for device mutation, thus allowing for nanoscale, evolutionary robotics. A simple device design, for example based on the nanotweezers by Yurke and colleagues ([151], Figure 1.5B) would likely be most accessible for proof-of-principle experiments.

There is also considerable potential for application of DNA nanotechnology in a biomedical context. Although the nanoscale surgeon to swallow that Richard Feynman envisioned [34] may remain confined to science-fiction for some time to come, promising advances are being made towards transferring increasingly complex DNA nanotechnology from the test tube to the syringe.

One of the most promising directions for biomedical application is found in the area of drug delivery [65]. DNA nanodevices can readily be functionalised with cargo molecules, which they can carry into cells (e.g. [135]) or release only in the vicinity of pre-defined target cells (e.g. [26]). Though impressive demonstrations of such devices have been published, realistically many years, if not decades are still required until they could be implemented in marketable therapeutics. The application of synthetic DNA machines capable of reporting environmental conditions from inside cells (e.g. [83]) for diagnostic purposes may arrive much sooner.

Such devices for treatment and diagnosis might achieve particularly impressive levels of ability when combined with DNA computation systems. DNA is too slow and error-prone to compete with electronic computers for the future of computation in general. Applications in biomedical contexts appear sensible, however, as it is there that DNA computation systems can play to their unique strengths, such as functionality in physiological conditions, biocompatibility, and the relative ease with which such systems can be designed to include biologically relevant input/output (I/O) interfaces. So far, however, the conceptually most impressive synthetic DNA computation CRNs have been operated in isolation. Although careful progress has been made towards integrating such systems into biological contexts [10, 111], further efforts are required if such nucleic acid reaction networks are ever to compute diagnostically relevant information to produce a therapeutic response in a complex organism.

A

APPENDIX

A.1 MATERIALS AND METHODS

A.1.1 *Buffers*

The following buffers were used for the experiments presented in this thesis (all prepared with Milli-Q-purified water):

- TE buffer: 10 mM tris(hydroxymethyl)aminomethane (Tris), 1 mM ethylenediaminetetraacetic acid (EDTA), pH 7.4;
- fluorimetry buffer: 1× Tris, 50 mM NaCl, 10 mM $MgCl_2$, pH 8.0;
- PCR machine buffer: fluorimetry buffer with 0.1 mg/ml BSA;
- microscopy buffer: TE buffer with 50 mM NaCl, 10 mM $MgCl_2$, pH 7.4;
- origami folding buffer: 40 mM Tris, 20 mM acetic acid, 12.5 mM magnesium acetate, pH 8.3.

A.1.2 *Oligonucleotide storage and duplex preparation*

All oligonucleotides were ordered from Integrated DNA Technologies. Synthesis scales were usually between 25 nmole and 100 nmole. Unlabelled strands were delivered desalted and lyophilised, except for the basic staples used in chapter 4, which were delivered desalted and in

a Tris-EDTA buffer. Fluorescently labelled strands were delivered HPLC-purified and lyophilised. Lyophilised strands were resuspended either in Milli-Q-purified water or TE buffer to a concentration of 100 μM, following manufacturer information.

In order to prepare duplexes, the two constituent strands were mixed at concentrations between 1 μM and 49 μM in buffer containing at least 10 mM $MgCl_2$. This mixture was annealed from 95 °C to 20 °C at $-3\,°C \cdot min^{-1}$.

A.1.3 *Spectrofluorimetry*

Two fluorimeters were used: For chapter 2, all experiments were conducted in a Fluoromax-3 JY Horiba fluorimeter. This fluorimeter was also used for the characterisation measurements in chapter 3, Figures 3.2B, 3.6B and 3.13B. However, for all other measurements in chapter 3, an Agilent Mx3005P quantitative PCR machine was used. Members of the Cyanine family of fluorophores (Cy3 and Cy5) were chosen as they are not quenched by nucleobases [21, 132].

Fluoromax

In the Fluoromax, all experiments were conducted in Hellma Analytics or Starna quartz cuvettes. All samples were measured in fluorimetry buffer. Most experiments were conducted in a volume of 1.5 ml, with the exception of five experiments in subsection 2.2.1 (those in sequence table A.5 with concentrations $\geq$ 50 nM), which were conducted in a volume of 0.15 ml. Samples were kept at a constant temperature throughout experiments using a Thermo NESLAB water bath (23 °C for experiments in chapter 2, 25 °C for the experiments in chapter 3, since this is the minimum temperature for the PCR machine, and consistency in experimental conditions was desirable).

For both fluorophores, optimal excitation and emission wavelenghts were determined. For Cy3, an excitation wavelength of 548 nm and an emission wavelength of 563 nm were used. For Cy5, the respective wavelengths were 649 nm and 664 nm. Data points were recorded every 2.1 s for fast reactions and up to every 60 s for slow reactions. In all experiments, the integration time was chosen to be 2 s. Excitation and emission bandpass were set to 2 nm, except for the experiments in section 2.5, where they were set to 1 nm to reduce crosstalk between the two fluorescent signals. With 2 nm bandpass, crosstalk had been determined to be present, but negligible; with 1 nm bandpass no crosstalk was measurable at the concentrations and integration times used.

Reactants were added to the cuvettes in the order described in the respective sections. Reactant addition was always performed in the same fashion, leading to the first data point to be acquired approximately 10 s after the addition. Rapid pipetting was used to mix reactants. To prevent sample evaporation, the cuvettes were covered with airtight lids during experiments. I found that molecular biology grade mineral oil, which is commonly used for this purpose, occasionally led to visible distortions in the time series recorded.

Cuvettes were cleaned before each experiment by washing the inside surfaces twice with absolute ethanol, five times with Milli-Q-purified water and twice again with absolute ethanol. The outside surfaces were wiped with a lens tissue soaked in absolute ethanol.

Spectrofluorimetry data was processed using Excel, OriginPro and Matlab. Fitting routines were implemented with custom Matlab code, some of which was based on the ideas by Zhang and Winfree [156], as described in chapter 2.

Mx3005P quantitative PCR machine

In the PCR machine, all experiments were conducted in single-use, 96-well plates (Starlab). All experiments were conducted at a total volume of 100 μl. Samples were covered with 8-cap strips (Life Technologies).

Pre-programmed PCR protocols were changed to perform cycles at a constant temperature of 25 °C. Cycle time was programmed to 2 min, with Cy5 fluorescence measured three times and then averaged at the end of each cycle using the correspondent internal filter set, without additional gain. Plates were prepared with the desired concentration of reporter complex in PCR machine buffer. Once the fluorescent signal was stable, the plates were removed for reactant addition. Mixing was found to be most conveniently performed with an 8-channel, 100 μl pipettor (Eppendorf). Mixing was consistently timed in a way that the first data acquisition would take place directly after.

Note that the PCR machine buffer in which all experiments were conducted deviates from the fluorimetry buffer used in the Fluoromax as it contains 0.1 mg/ml BSA. This was found to be necessary as with the regular fluorimetry buffer, reporter fluorescence would show strongly time-dependent behaviour even before reactant addition, an observation not made for the experiments in the Fluoromax. It is assumed that this is due to reporter complex adsorbing on the surface of the plastic well plate. Although the signal does stabilise eventually, it is assumed that a considerable proportion of reporter complex (and probably other reaction network components) would not participate in the designed reaction pathways, which may lead to undesired system behaviour. As a consequence, BSA was tested as a surface blocking agent (similar to its use in chapter 4) and found to yield satisfactory results (Figure A.1). As it is not known to interfere with DNA reactions, the systems studied in the PCR machine are unlikely to have been influenced in a negative fashion.

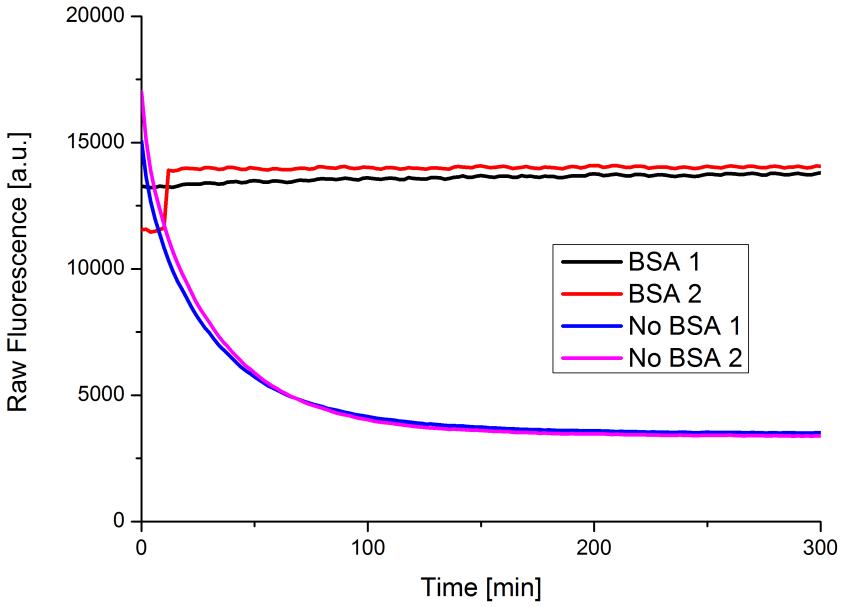

Figure A.1: Raw fluorescence traces of 100 nM reporter complex added to a total of four wells at $t \approx 0$ in PCR machine buffer (fluorimetry buffer containing 0.1 mg/ml BSA, traces BSA 1 and BSA 2) and in regular fluorimetry buffer (traces No BSA 1 and No BSA 2). No baseline has been subtracted from these traces, i.e. 0 fluorescence denotes complete absence of fluorescent signal. The initial jump in the BSA 2 trace is characteristic of a surface bubble bursting.

The recorded cycle-against-fluorescence output was converted to the time-against-fluorescence traces shown in chapter 3 by determining the first cycle after which all reaction components had been added, subtracting its cycle number and fluorescence measured from all later data points corresponding to that well, effectively setting both its cycle number and fluorescence to 0. Subsequently, cycle number was converted to time in minutes for each data point. This is a subtly different procedure than that used in the Fluoromax, in which the baseline fluorescence was determined through an average of data points before the addition of non-fluorescent reaction components. This approach was found to be unfeasible for the PCR machine, as traces would often begin at significantly lower or higher fluorescence values as a previously determined baseline after reactant addition and mixing. This difference to the Fluoromax is likely due to the way fluorescence is measured in the two devices: whereas the Fluoromax excites samples with a beam aimed at the bottom of the cuvette, the beam in the PCR machine passes through the sample surface, which is susceptible to surface effects such as bubbles. The baseline fluorescence approach was only used in the PCR machine for the fluorescence normalisation process, in which an excess of reporter target complement was added to 15 nM reporter complex; a very fast reaction in which the first data point would have deviated quite strongly from the actual starting point due to the unavoidable time delay between sample mixing and first data point acquisition. In order to minimise the impact of surface effects on the normalisation, this experiment was conducted three times, and the average difference between baseline and saturation fluorescence taken to be a good representation of fluorescent change corresponding to displacement of 15 nM reporter complex incumbent.

A.1.4 *Origami tile design and preparation*

Both design and preparation of DNA origami tiles for chapter 4 were based on previous work in the research group notably references [141] and [143]. 216 staple strands were used to fold a single-stranded, M13mp18 scaffold into a ~100 nm × 70 nm rectangle, which is schematically depicted in Figure A.2. Different tile designs were implemented through the preparation of a basic staple mix, containing 193 staples that would be present in all tiles. For each individual tile design, a further specialised staple mix was prepared that would be added to the basic staple mix during the preparation of the DNA origami. Through this approach, different tile designs could be prepared relatively quickly, without the need for the preparation of a full set of staples from scratch. The specialised staple mixes are listed in Table A.15. The full list of basic staples can be found in reference [141].

To prepare a tile sample, 6.7 μl 1 $\mu g \cdot \mu l^{-1}$ M13mp18 was mixed with 20.8 μl 0.72 μM basic staple mix, 7.6 μl 2 μM specialised staple mix, 6 μl 10× origami buffer concentrate and 6.7 μl Milli-Q-purified water. This corresponds to final concentrations of 50 nM scaffold strand, 250 nM of each staple strand and 12.5 mM Mg^{2+} ions in a total volume of 60 μl. This mixture was rapidly heated to 95 °C and subsequently annealed with a thermal gradient of $-1\,°C \cdot min^{-1}$ to room temperature. Occasionally, the mixture was scaled to total volumes between 30 μl and 120 μl, if different quantities of sample were required. Prior to further use, excess staples were removed from annealed samples using hand-packed size-exclusion columns, as described previously [141].

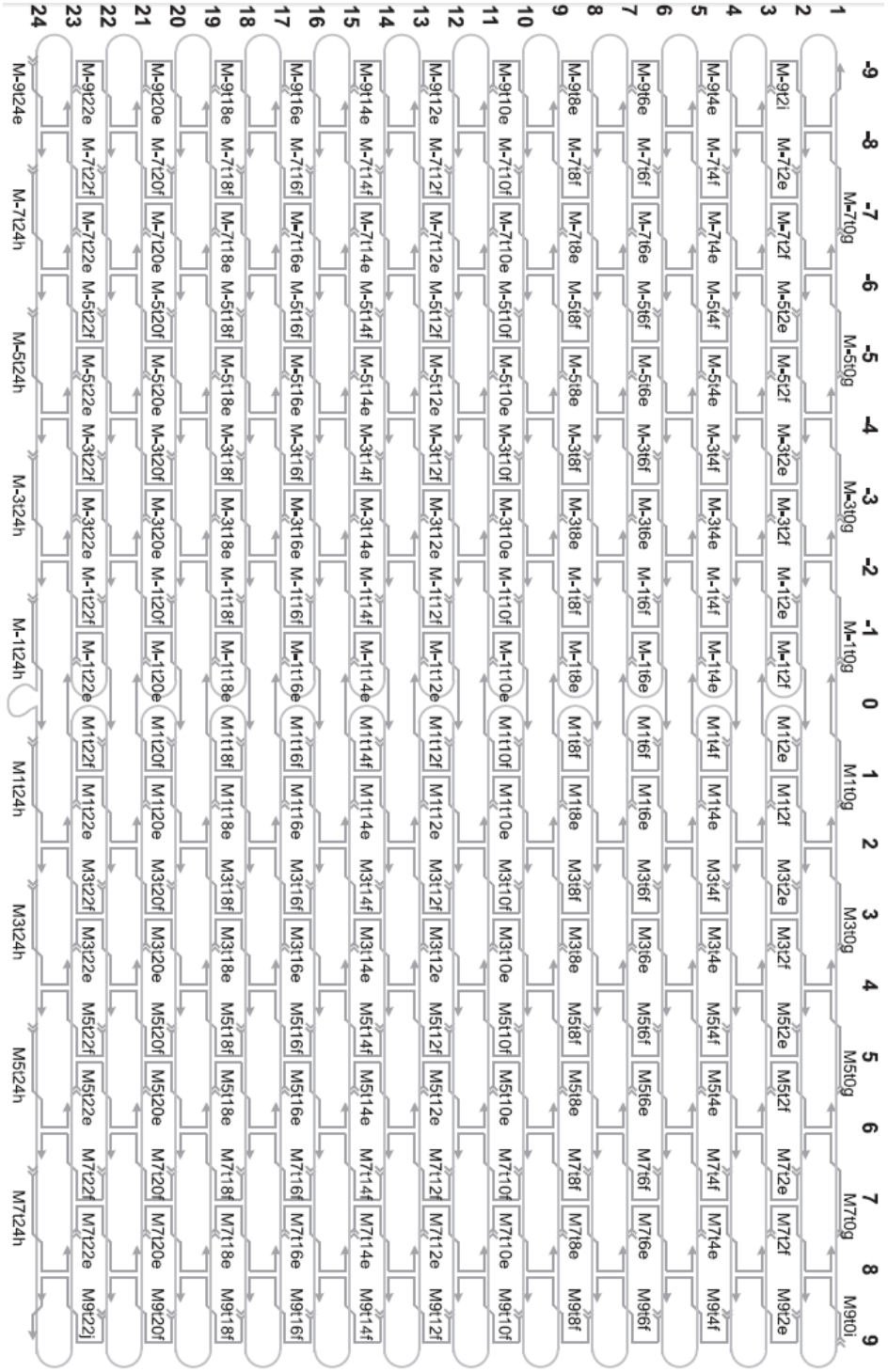

Figure A.2: The origami tile design. The staple strands are shown as arrows, with arrowheads depicting 3′ ends. The naming convention for the staples follows Rothemund [103]. Adapted from the doctoral thesis of Shelley Wickham [141].

Tile designs and specialised staple mixes

Specialised staple	Location	Tile design 1	2	3	4	5	6
label top right	M9t0i	x					
label top left	M-9t2i	x		x	x		
stator 1	M-7t2f		o	x	o	o	o
stator 2	M-7t2e		x	x		x	x
stator 3	M-5t4f		x	x		x	x
stator 4	M-5t4e		x	x		x	x
stator 5	M-3t6f			x		x	x
biotin stator 5	M-3t6f		x				
stator 6	M-3t6e		x	x		x	x
stator 7	M-1t8f		x	x		x	x
stator 8	M-1t8e		x	x		x	x
stator 9	M1t10f		x	x	o	x	x
stator 10	M1t10e		x	x		x	x
stator 11	M3t12f		x	x		x	x
stator 12	M3t12e		x	x		x	x
stator 13	M5t14f			x		x	x
biotin stator 13	M5t14f		x				
stator 14	M5t14e		x	x		x	x
stator 15	M7t16f		x	x		x	x
stator 16	M7t16e		x	o		x	x
sl top 1	M-1t2e	x		x	x	x	
sl top 2	M-1t2e	x		x	x	x	
sl left 1	M-7t12f	x		x	x	x	
sl left 2	M-7t12f	x		x	x	x	
sl right 1	M7t12f	x		x	x	x	
sl right 2	M7t12f	x		x	x	x	
sl bottom 1	M1t22e	x		x	x	x	
sl bottom 2	M1t22e	x		x	x	x	

sl: surface linker

Table A.1: Specialised staple mixes for the tile designs described in the text. The staple location can be inferred from Figure A.2. Sequences of the specialised staples can be found in Tables A.15 and A.16. An 'x' denotes that the tile design included the specialised staple instead of the regular one. Blank denotes that for this location, the regular staple was used, which has no functionality besides scaffold hybridisation. An 'o' denotes that both were omitted before annealing the sample, to allow for later integration of a stator-motor duplex.

As mentioned above, slightly different tile designs were implemented by adding specialised staple mixes to the basic staple mix. These specialised staple mixes are listed in Table A.1.

Tile design 1 includes two biotinylated staples on the corners of a 100 nm edge of the assembled tile, label top left and label top right. It further includes strands to enable surface modification via hybridisation of an anchor strand. Note that for this purpose, four regular staples were replaced with two specialised staples each. This is necessary to ensure correct orientation of the tile on the surface, i.e. so that all functional domains and labels are on the side of the tile opposite the surface-binding domains [141]. This tile design was used for the static distance measurements using two-channel colocalisation as well as the 100 nm-distance measurements using quantum dot blinking (section 4.3).

Tile design 2 contains two biotinylated stator staples which are approximately 50 nm apart. Although originally intended for a different purpose, this tile design served as the basis of the quantum dot blinking experiment presented in Figure 4.10.

Tile design 3 contains a full track, with the exception of stator 16, and a biotinylated staple in the vicinity of stator 1. This design was used to perform the live motor measurements in section 4.4, as well as the static measurements for the two plots at the top of Figure 4.15. Motors were included in duplex with stator 16 and were expected to move towards stator 1.

Tile design 4 contains post-assembly integration sites for stators 1 and 9, and the same labelling site as tile design 3 and was used for the measurement presented in the bottom plot of Figure 4.15.

Tile designs 5 and 6 were used to characterise the surface treatment as presented in Figure 4.6, with 5 expected to bind to the treated surface, and 6 not, due to its lack of specialised surface binding staples.

A.1.5 *Fluorescence microscopy*

Microscopy setup

All microscopy experiments were performed using an Olympus IX-81 inverted microscope. Images were recorded using an Andor iXon EMCCD camera. Samples in TIRF experiments were illuminated using an OZ Optics 532 nm laser attached to an Olympus module for objective-based TIRF. Two Olympus oil-immersion objectives suitable for TIRF in aqueous samples were used, one with $60\times$ and another with $100\times$ magnification, with the latter showing slightly better localisation capabilities. Further internal $1.6\times$ magnification was used for all experiments.

Emission was led through an MAG Biosystems DV-2 device, which splits the signal into two channels which are optically filtered separately with a (565 ± 20) nm and a (665 ± 20) nm filter, respectively. The spectrally separated signals were projected onto 512 px $\times$ 256 px halves of the 512 px $\times$ 512 px camera chip surface. Andor iQ was used as imaging and microscope control software.

Flow chambers

Flow chambers were constructed from two cover slips (25 mm $\times$ 60 mm and 22 mm $\times$ 22 mm, respectively) and Parafilm. Two strips of Parafilm were placed on the big cover slip to form a 2 mm $\times$ 25 mm channel, upon which the small cover slip was placed. After 1 min on a hot plate at 80 °C, the cover slips were pressed together to yield a flow chamber of approximately 5 μl volume.

As mentioned in subsection 4.2.4, surfaces of the flow chamber were prepared with an evolving protocol for the binding of DNA origami tiles. The different parameters are listed in Table A.2. Surface protocol 1 was used for the quantum dot blinking experiment presented in Figure 4.10.

The origami tile sample was allowed to incubate for 2 min in the flow chamber, which was subsequently washed with an excess of buffer. Protocol 2 was used for the quantum dot blinking characterisation in subsection 4.2.3. After washing with 10 μl of 2 mg/ml BSA and 0.5 mg/ml biotinylated BSA mixture in microscopy buffer, 5 min incubation and a subsequent wash with microscopy buffer, 10 μl of 100 pM quantum dots were added to the flow chamber and allowed to incubate for 2 min before the flow chamber was washed with microscopy buffer. To test the effect of BME on the blinking behaviour of the quantum dots, a final wash was performed with 10 μl 1 % BME in microscopy buffer. Surface protocol 3 was used for the two-colour distance measurement between a labelled motor on the last stator and a reference label on the tile, with the resulting collective distance distribution depicted in the bottom plot of Figure 4.15. Protocol 4 was used in all measurements in which a labelled motor was observed, with the aforementioned exception. Protocol 5 was used for all remaining two-colour distance measurements, and the 100 nm-distance measurements using quantum dot blinking.

For protocols 3 to 5, the procedure to prepare the flow chamber surfaces was as follows:

1. wash with 10 μl of 2 mg/ml BSA and 0.5 mg/ml biotinylated BSA mixture in microscopy buffer, incubate for 5 min, wash with microscopy buffer;

2. wash with 10 μl of 1 mg/ml streptavidin in microscopy buffer, incubate for 5 min, wash with microscopy buffer;

3. wash with 10 μl of anchor strand in microscopy buffer (concentration as stated in Table A.2), incubate for 5 min, wash with microscopy buffer;

4. wash with 10 μl of sample, incubate for 5 min, wash with microscopy buffer.

Substance	Surface protocol				
	1	2	3	4	5
BSA	-	2 mg/ml	2 mg/ml	2 mg/ml	2 mg/ml
Biotinylated BSA	-	0.5 mg/ml	0.5 mg/ml	0.5 mg/ml	0.5 mg/ml
Streptavidin	-	-	1 mg/ml	1 mg/ml	1 mg/ml
Anchor strand	-	-	10 nM	10 nM	100 nM
Origami tile	p, 1% v/v	-	up, 0.05% v/v	p, 10 – 50% v/v	p, 10 – 50% v/v

p: ion-exchange purified, up: unpurified

Table A.2: Surface protocol parameters.

Channel mapping

Fiducial bead distributions (as described in subsection 4.2.2) were acquired by first preparing a 3 % v/v dilution of 100 nm diameter Life Technologies TetraSpeck Microspheres in microscopy buffer. 10 μl of this dilution were washed into a fresh, untreated flow chamber, leading to non-specific adsorption of the beads on the chamber surfaces. When necessary, adjustments were made to the dual view module so that

$$(x_{channel\,1},\, y_{channel\,1}) = (x_{channel\,2},\, y_{channel\,2} - 256\,\mathrm{px})\,,$$

a very coarse approximation to the actual mapping function, was approximately satisfied. Different sections of the flow chamber were imaged to yield approximately 250 positions of fiduciary signals.

In the resulting images, bead PSFs were identified and fitted using Matlab code provided by Stephan Uphoff. Signals were identified by finding pixels above a certain brightness threshold. The area around such a local maximum (15 px $\times$ 15 px) was then cut out of the image and fitted with a Gaussian, unless a second local maximum was present in the cutout as well. In such a case, neither signal was evaluated.

This yielded an unpaired list of fiduciary bead positions for both channels. Using my own Matlab code, this array was subsequently ordered by performing a nearest-neighbour search between bead positions of the two channels. The mapping function was derived from this list using the Matlab 'local weighted mean' transformation. This channel mapping routine was performed before each series of two-colour microscopy experiments presented in this thesis, although it was found to be stable for several weeks if the microscope setup remained unused in the meantime.

Image analysis for duochrome samples

The recordings made with quantum dots were treated in a similar fashion to the recordings made with beads for the channel mapping procedure. Local maxima were identified with the same code in an average over all single frames of a stack, again with the restriction that no other signal was found in the vicinity. As with the channel mapping procedure, these peaks were fitted with Gaussians in both channels to retrieve preliminary localisation data. The peak positions registered in the red channel were subsequently transformed with the previously determined mapping function. If a corresponding signal was found in the green channel within a 3 px radius of a transformed red signal position, this pair was considered to represent a dual-labelled structure.

In both channels, the areas around such signals were cut out of the stack, and the PSFs fitted in each frame of these substacks. Positions derived from frames in which either the brightness was comparatively low ($< 10\,\%$ of the maximum brightness of that particular signal) or the fit was poor (normalised $\chi^2 > 10$) were filtered out. Position lists from both channels were manipulated to retrieve positions within the original stack, rather than positions within the substack. This was necessary for the proper application of the mapping function, which was subsequently

used to transform the red channel position lists. The now colocalised position lists were further transformed from being in units of pixels to nanometres, using conversion factors determined with a microscope calibration grid. This concluded colocalisation, from which further analysis followed as described in chapter 4.

Image analysis for monochrome samples

Analysis of image stacks for subsection 4.3.2 was very similar to that of duochrome samples, with the major difference being the lack of a colocalisation process. Upon signal identification and fitting, positions derived from low-brightness frames were discarded. The data was then treated as described in subsection 4.3.2.

A.2 DNA STRAND SEQUENCES AND CONCENTRATIONS

A.2.1 *Chapter 2*

Systems using reporter complex, part 1/2

Strand		Sequence ($5' \rightarrow 3'$)	Conc. [nM]
rep incumbent		Cyanine 5 - TGG TGT TTG TGG GTG T	duplex, 2×
rep target		ACT CAC CAC ACC CAC AAA CAC CA-IowaBlack RQ	
MM pos	Species		
n, 2,	incumbent	TGG TGT TTG TGG GTG TGG TGA GTT TGA GGT TGA	duplex, 1×
12, 14	target	CCC TCC ACA TTC AAC CTC AAA CTC ACC	
	6 nt inv	GGT GAG TTT GAG GTT GAA TGT GG	5
n	7 nt inv	GGT GAG TTT GAG GTT GAA TGT GGA	5
	10 nt inv	GGT GAG TTT GAG GTT GAA TGT GGA GGG	5
	6 nt inv	GGT GAG TTT GAG GTT CAA TGT GG	10
2	7 nt inv	GGT GAG TTT GAG GTT CAA TGT GGA	10
	10 nt inv	GGT GAG TTT GAG GTT CAA TGT GGA GGG	5
	6 nt inv	GGT GAC TTT GAG GTT GAA TGT GG	5
12	7 nt inv	GGT GAC TTT GAG GTT GAA TGT GGA	5
	10 nt inv	GGT GAC TTT GAG GTT GAA TGT GGA GGG	5
	6 nt inv	GGT CAG TTT GAG GTT GAA TGT GG	5
14	7 nt inv	GGT CAG TTT GAG GTT GAA TGT GGA	5
	10 nt inv	GGT CAG TTT GAG GTT GAA TGT GGA GGG	5
	incumbent	TGG TGT TTG TGG GTG TGG TGA GTT TGA GGT GAT	duplex, 1×
	target	CCC TCC ACA TAT CAC CTC AAA CTC ACC	
3	6 nt inv	GGT GAG TTT GAG GTC ATA TGT GG	10
	7 nt inv	GGT GAG TTT GAG GTC ATA TGT GGA	10
	10 nt inv	GGT GAG TTT GAG GTC ATA TGT GGA GGG	5
	incumbent	TGG TGT TTG TGG GTG TGG TGA GTT TGA GTG AGT	duplex, 1×
	target	CCC TCC ACA TAC TCA CTC AAA CTC ACC	
4	6 nt inv	GGT GAG TTT GAG TCA GTA TGT GG	10
	7 nt inv	GGT GAG TTT GAG TCA GTA TGT GGA	10
	10 nt inv	GGT GAG TTT GAG TCA GTA TGT GGA GGG	5
⋮	⋮	⋮	⋮

rep: reporter, MM pos: mismatch position, inv: invader, n: no mismatch

Table A.3: Sequences and experimental concentrations for the different systems tested in subsection 2.2.2. Strand species are classified according to mismatch position. Incumbent-target duplex concentration was chosen to match the concentration of the invading strand in each experiment, reporter complex was always at twice that concentration. Continued in Table A.4.

Systems using reporter complex, part 2/2

Strand MM pos	Species	Sequence (5′ → 3′)	Conc. [nM]
	incumbent	TGG TGT TTG TGG GTG TGG TGA GTT TGA TGA GGT	duplex, 1×
	target	CCC TCC ACA TAC CTC ATC AAA CTC ACC	
5	6 nt inv	GGT GAG TTT GAT CAG GTA TGT GG	10
	7 nt inv	GGT GAG TTT GAT CAG GTA TGT GGA	10
	10 nt inv	GGT GAG TTT GAT CAG GTA TGT GGA GGG	5
	incumbent	TGG TGT TTG TGG GTG TGG TGA GTT TGT GAA GGT	duplex, 1×
	target	CCC TCC ACA TAC CTT CAC AAA CTC ACC	
6	6 nt inv	GGT GAG TTT GTC AAG GTA TGT GG	10
	7 nt inv	GGT GAG TTT GTC AAG GTA TGT GGA	10
	10 nt inv	GGT GAG TTT GTC AAG GTA TGT GGA GGG	5
	incumbent	TGG TGT TTG TGG GTG TGG TGA GTT TTG AGA GGT	duplex, 1×
	target	CCC TCC ACA TAC CTC TCA AAA CTC ACC	
7	6 nt inv	GGT GAG TTT TCA GAG GTA TGT GG	10
	7 nt inv	GGT GAG TTT TCA GAG GTA TGT GGA	5
	10 nt inv	GGT GAG TTT TCA GAG GTA TGT GGA GGG	5
	incumbent	TGG TGT TTG TGG GTG TGG TGA GTT TGA TGA GGT	duplex, 1×
	target	CCC TCC ACA TAC CTC ATC AAA CTC ACC	
8	6 nt inv	GGT GAG TTT CAT GAG GTA TGT GG	10
	7 nt inv	GGT GAG TTT CAT GAG GTA TGT GGA	5
	10 nt inv	GGT GAG TTT CAT GAG GTA TGT GGA GGG	5
	incumbent	TGG TGT TTG TGG GTG TGG TGA GTT GAT TGA GGT	duplex, 1×
	target	CCC TCC ACA TAC CTC AAT CAA CTC ACC	
9	6 nt inv	GGT GAG TTC ATT GAG GTA TGT GG	10
	7 nt inv	GGT GAG TTC ATT GAG GTA TGT GGA	5
	10 nt inv	GGT GAG TTC ATT GAG GTA TGT GGA GGG	5
	incumbent	TGG TGT TTG TGG GTG TGG TGA GTG ATT TGA GGT	duplex, 1×
	target	CCC TCC ACA TAC CTC AAA TCA CTC ACC	
10	6 nt inv	GGT GAG TCA TTT GAG GTA TGT GG	10
	7 nt inv	GGT GAG TCA TTT GAG GTA TGT GGA	5
	10 nt inv	GGT GAG TCA TTT GAG GTA TGT GGA GGG	5

rep: reporter, MM pos: mismatch position, inv: invader, n: no mismatch

Table A.4: Continuation of Table A.3.

Systems with labelled incumbent and target

Strand	Sequence (5′ → 3′)	Conc. [nM]
incumbent	Cyanine 5 - GGT GAG TTT GAG GTT G	duplex, 1×
target	CCC TCC ACA TCA ACC TCA AAC TCA CC-IowaBlack RQ	
10 nt inv n	GGT GAG TTT GAG GTT GAT GTG GAG GG	10
10 nt inv d	GGT CAG TTT GAG GTT GAT GTG GAG GG	10
10 nt inv c	GGT GAG TTT CAG GTT GAT GTG GAG GG	10
10 nt inv c*	GGT GAG TTT TAG GTT GAT GTG GAG GG	5, 10, 20
10 nt inv p	GGT GAG TTT GAG GTT CAT GTG GAG GG	5, 10, 20
9 nt inv n	GGT GAG TTT GAG GTT GAT GTG GAG G	10
9 nt inv d	GGT CAG TTT GAG GTT GAT GTG GAG G	5, 10, 20
9 nt inv c	GGT GAG TTT CAG GTT GAT GTG GAG G	5, 10, 20
9 nt inv c*	GGT GAG TTT TAG GTT GAT GTG GAG G	10
9 nt inv p	GGT GAG TTT GAG GTT CAT GTG GAG G	10
8 nt inv n	GGT GAG TTT GAG GTT GAT GTG GAG	10
8 nt inv d	GGT CAG TTT GAG GTT GAT GTG GAG	10
8 nt inv c	GGT GAG TTT CAG GTT GAT GTG GAG	5, 10, 20
8 nt inv c*	GGT GAG TTT TAG GTT GAT GTG GAG	10
8 nt inv p	GGT GAG TTT GAG GTT CAT GTG GAG	5, 10, 20
8 nt inv n th	GG TGA GTT TGA GGT TGA TGT GGT G	10
8 nt inv p th	GG TGA GTT TGA GGT TCA TGT GGT G	10
7 nt inv n	GGT GAG TTT GAG GTT GAT GTG GA	5, 10, 20
7 nt inv d	GGT CAG TTT GAG GTT GAT GTG GA	10
7 nt inv c	GGT GAG TTT CAG GTT GAT GTG GA	10
7 nt inv c*	GGT GAG TTT TAG GTT GAT GTG GA	5, 10, 20
7 nt inv p	GGT GAG TTT GAG GTT CAT GTG GA	100
6 nt inv n	GGT GAG TTT GAG GTT GAT GTG G	10
6 nt inv d	GGT CAG TTT GAG GTT GAT GTG G	5, 10, 20
6 nt inv c	GGT GAG TTT CAG GTT GAT GTG G	100
6 nt inv c*	GGT GAG TTT TAG GTT GAT GTG G	50
6 nt inv p	GGT GAG TTT GAG GTT CAT GTG G	100
5 nt inv n	GGT GAG TTT GAG GTT GAT GTG	10
4 nt inv n	GGT GAG TTT GAG GTT GAT GT	100

inv: invader, n: no mismatch, d: distal C-C, c: central C-C, c*: central C-T, p: proximal C-C

Table A.5: Sequences and experimental concentrations for the different systems tested in subsection 2.2.1. The invading strands are classified according to their toehold length (between 6 nt and 10 nt), and mismatch type and location. Incumbent-target duplex concentration was chosen to match the concentration of the invading strand in each experiment.

Systems w/ elongated displacement domain part 1/2

Strand		Sequence ($5' \rightarrow 3'$)	Conc. [nM]
n, 2,	incumbent	TGGTGTTTGTGGGTGTGGTGAGTTTGAGGTTGAATGGATGA	duplex, 1×
20, 22	target	CCCTCCACATTCATCCATTCAACCTCAAACTCACC	
n	inv	GGTGAGTTTGAGGTTGAATGGATGAATGTGGA	5
2	inv	GGTGAGTTTGAGGTTGAATGGATCAATGTGGA	10
20	inv	GGTGACTTTGAGGTTGAATGGATGAATGTGGA	5
22	inv	GGTCAGTTTGAGGTTGAATGGATGAATGTGGA	5
	incumbent	TGGTGTTTGTGGGTGTGGTGAGTTTGAGGTTGAATGTGAGA	duplex, 1×
4	target	CCCTCCACATTCTCACATTCAACCTCAAACTCACC	
	inv	GGTGAGTTTGAGGTTGAATGTCAGAATGTGGA	10
	incumbent	TGGTGTTTGTGGGTGTGGTGAGTTTGAGGTTGAATGATGGA	duplex, 1×
6	target	CCCTCCACATTCCATCATTCAACCTCAAACTCACC	
	inv	GGTGAGTTTGAGGTTGAATCATGGAATGTGGA	10
	incumbent	TGGTGTTTGTGGGTGTGGTGAGTTTGAGGTTGTGAAATGGA	duplex, 1×
8	target	CCCTCCACATTCCATTTCACAACCTCAAACTCACC	
	inv	GGTGAGTTTGAGGTTGTCAAATGGAATGTGGA	10
	incumbent	TGGTGTTTGTGGGTGTGGTGAGTTTGAGGTGATTGAATGGA	duplex, 1×
11	target	CCCTCCACATTCCATTCAATCACCTCAAACTCACC	
	inv	GGTGAGTTTGAGGTCATTGAATGGAATGTGGA	10
	incumbent	TGGTGTTTGTGGGTGTGGTGAGTTTGAGTGAGTTGAATGGA	duplex, 1×
12	target	CCCTCCACATTCCATTCAACTCACTCAAACTCACC	
	inv	GGTGAGTTTGAGTCAGTTGAATGGAATGTGGA	10
	incumbent	TGGTGTTTGTGGGTGTGGTGAGTTTGATGAGGTTGAATGGA	duplex, 1×
13	target	CCCTCCACATTCCATTCAACCTCATCAAACTCACC	
	inv	GGTGAGTTTGATCAGGTTGAATGGAATGTGGA	10
⋮	⋮	⋮	⋮

Table A.6: Sequences and experimental concentrations for the elongated displacement domain. Continued in Table A.7.

Systems w/ elongated displacement domain part 2/2

Strand		Sequence (5′ → 3′)	Conc. [nM]
14	incumbent	TGGTGTTTGTGGGTGTGGTGAGTTTGTGAAGGTTGAATGGA	duplex, 1×
	target	CCCTCCACATTCCATTCAACCTTCACAAACTCACC	
	inv	GGTGAGTTTGTCAAGGTTGAATGGAATGTGGA	10
15	incumbent	TGGTGTTTGTGGGTGTGGTGAGTTTTGAGAGGTTGAATGGA	duplex, 1×
	target	CCCTCCACATTCCATTCAACCTCTCAAAACTCACC	
	inv	GGTGAGTTTTCAGAGGTTGAATGGAATGTGGA	10
16	incumbent	TGGTGTTTGTGGGTGTGGTGAGTTTGATGAGGTTGAATGGA	duplex, 1×
	target	CCCTCCACATTCCATTCAACCTCATCAAACTCACC	
	inv	GGTGAGTTTCATGAGGTTGAATGGAATGTGGA	10
17	incumbent	TGGTGTTTGTGGGTGTGGTGAGTTGATTGAGGTTGAATGGA	duplex, 1×
	target	CCCTCCACATTCCATTCAACCTCAATCAACTCACC	
	inv	GGTGAGTTCATTGAGGTTGAATGGAATGTGGA	5
18	incumbent	TGGTGTTTGTGGGTGTGGTGAGTGATTTGAGGTTGAATGGA	duplex, 1×
	target	CCCTCCACATTCCATTCAACCTCAAATCACTCACC	
	inv	GGTGAGTCATTTGAGGTTGAATGGAATGTGGA	5

Table A.7: Sequences and experimental concentrations for the elongated displacement domain.

Competitive displacement

Strand	Sequence ($5' \rightarrow 3'$)
$5'$ invader	Cyanine 5-CCT TAC CTT CTT CCA TTC CTC ATC TAT-IowaBlack RQ
$3'$ invader	IowaBlack FQ-TAT CTA CTC CTT ACC TTC TTC CAT TCC-Cyanine 3
incumbent t: $5'$, k: $3'$	TA CCT TCT TCC AGT CC
target t: $5'$, k: $3'$	A GAT GAG GAC TGG AAG AAG GTA AGG AGT
incumbent t: $3'$, k: $5'$	CC TGA CCT TCT TCC AT
target t: $3'$, k: $5'$	T GAG GAA TGG AAC AAG GTC AGG AGT AGA

Table A.8: Sequences for the two competitive displacement systems tested in section 2.5. For the first incumbent-target pair (t: $5'$, k: $3'$), the $5'$ invader is thermodynamically and the $3'$ invader kinetically favored; for the second pair (t: $3'$, k: $5'$) this is reversed.

A.2.2 *Chapter 3*

Read-out duplexes and reporter complex

Strand	Sequence ($5' \rightarrow 3'$)
reporter incumbent	Cyanine 5 - TGG TTG TGG GTG TG
reporter target	CAC CAT CA CAC CCA CAA CCA - Iowa Black RQ
rd 11 target	CTA ATC TTC CAC TT CTC CCA TTC ATT CAC CAT C
rd 11 incumbent	TGG TTG TGG GTG TGA TGG TG AAT GAA TGG GAG TT GTG G
rd 12 target	CTC ACT ATT TCC TT CTC CCA TTC ATT CAC CAT C
rd 12 incumbent	TGG TTG TGG GTG TGA TGG TG AAT GAA TGG GAG TT GGA A
rd 13 target	CTT CTA CCC TAC TT CTC CCA TTC ATT CAC CAT C
rd 13 incumbent	TGG TTG TGG GTG TGA TGG TG AAT GAA TGG GAG TT GTA G
rd 21 target	CTA ATC TTC CAC TT TAT TCC TCA TCT CAC CAT C
rd 21 incumbent	TGG TTG TGG GTG TGA TGG TG AGA TGA GGA ATA TT GTG G
rd 22 target	CTC ACT ATT TCC TT TAT TCC TCA TCT CAC CAT C
rd 22 incumbent	TGG TTG TGG GTG TGA TGG TG AGA TGA GGA ATA TT GGA A
rd 23 target	CTT CTA CCC TAC TT TAT TCC TCA TCT CAC CAT C
rd 23 incumbent	TGG TTG TGG GTG TGA TGG TG AGA TGA GGA ATA TT GTA G
rd 31 target	CTA ATC TTC CAC TT TTA CTT ACT CTC CAC CAT C
rd 31 incumbent	TGG TTG TGG GTG TGA TGG TG GAG AGT AAG TAA TT GTG G
rd 32 target	CTC ACT ATT TCC TT TTA CTT ACT CTC CAC CAT C
rd 32 incumbent	TGG TTG TGG GTG TGA TGG TG GAG AGT AAG TAA TT GGA A
rd 33 target	CTT CTA CCC TAC TT TTA CTT ACT CTC CAC CAT C
rd 33 incumbent	TGG TTG TGG GTG TGA TGG TG GAG AGT AAG TAA TT GTA G

rd: read-out duplex

Table A.9: Sequences for reporter complex and read-out duplexes used in chapter 3. The read-out duplex for index $(3, 3)$, comprised of rd 33 target and rd 33 incumbent, was used in the proof-of-principle experiments depicted in Figures 3.2, 3.6 and 3.13.

Invader strands for simple Boolean matrix multiplication

Strand	Sequence ($5' \rightarrow 3'$)
distal invader 11	AAT GAA TGG GAG TT AAC CAC CTT ATC
distal invader 12	AAT GAA TGG GAG TT CCC AAA CCT AAC
distal invader 13	AAT GAA TGG GAG TT CCA CAA CAA CTT
distal invader 21	AGA TGA GGA ATA TT AAC CAC CTT ATC
distal invader 22	AGA TGA GGA ATA TT CCC AAA CCT AAC
distal invader 23	AGA TGA GGA ATA TT CCA CAA CAA CTT
distal invader 31	GAG AGT AAG TAA TT AAC CAC CTT ATC
distal invader 32	GAG AGT AAG TAA TT CCC AAA CCT AAC
distal invader 33	GAG AGT AAG TAA TT CCA CAA CAA CTT
proximal invader 11	GAT AAG GTG GTT TT GTG GAA GAT TAG
proximal invader 12	GAT AAG GTG GTT TT GGA AAT AGT GAG
proximal invader 13	GAT AAG GTG GTT TT GTA GGG TAG AAG
proximal invader 21	GTT AGG TTT GGG TT GTG GAA GAT TAG
proximal invader 22	GTT AGG TTT GGG TT GGA AAT AGT GAG
proximal invader 23	GTT AGG TTT GGG TT GTA GGG TAG AAG
proximal invader 31	AAG TTG TTG TGG TT GTG GAA GAT TAG
proximal invader 32	AAG TTG TTG TGG TT GGA AAT AGT GAG
proximal invader 33	AAG TTG TTG TGG TT GTA GGG TAG AAG

Table A.10: Sequences for invader strands used for simple Boolean matrix multiplication. Distal invader 33 and proximal invader 33 were used in the proof-of-principle experiments depicted in Figures 3.2 and 3.13. Proximal invader 33 was additionally used in the proof-of-principle experiment depicted in Figure 3.6.

Intermediate read-out duplexes

Strand	Sequence ($5' \rightarrow 3'$)
imrd 11 target	AA CTC CCA TTC ATT TT TCC TAT CTC TTT
imrd 11 incumbent	A GGA TT AAT GAA TGG GAG TT GTG GAA GAT TAG
imrd 12 target	AA CTC CCA TTC ATT TT ACA ATA TCC TTC
imrd 12 incumbent	T TGT TT AAT GAA TGG GAG TT GGA AAT AGT GAG
imrd 13 target	AA CTC CCA TTC ATT TT CAT TAT TAA CCC
imrd 13 incumbent	A ATG TT AAT GAA TGG GAG TT GTA GGG TAG AAG
imrd 21 target	AA TAT TCC TCA TCT TT TCC TAT CTC TTT
imrd 21 incumbent	A GGA TT AGA TGA GGA ATA TT GTG GAA GAT TAG
imrd 22 target	AA TAT TCC TCA TCT TT ACA ATA TCC TTC
imrd 22 incumbent	T TGT TT AGA TGA GGA ATA TT GGA AAT AGT GAG
imrd 23 target	AA TAT TCC TCA TCT TT CAT TAT TAA CCC
imrd 23 incumbent	A ATG TT AGA TGA GGA ATA TT GTA GGG TAG AAG
imrd 31 target	AA TTA CTT ACT CTC TT TCC TAT CTC TTT
imrd 31 incumbent	A GGA TT GAG AGT AAG TAA TT GTG GAA GAT TAG
imrd 32 target	AA TTA CTT ACT CTC TT ACA ATA TCC TTC
imrd 32 incumbent	T TGT TT GAG AGT AAG TAA TT GGA AAT AGT GAG
imrd 33 target	AA TTA CTT ACT CTC TT CAT TAT TAA CCC
imrd 33 incumbent	A ATG TT GAG AGT AAG TAA TT GTA GGG TAG AAG

imrd: intermediate read-out duplex

Table A.11: Sequences for intermediate read-out duplexes used in chapter 3. The intermediate read-out duplex for index 33 was used in the proof-of-principle experiments depicted in Figure 3.6.

Intermediate invader strands for cascaded matrix multiplication

Strand	Sequence ($5' \rightarrow 3'$)
im distal invader 11	AGT GTG TGT ATT TT AAT GAA TGG GAG
im distal invader 12	TGG AGT TAT TTG TT AAT GAA TGG GAG
im distal invader 13	TAG TTT AAA GGG TT AAT GAA TGG GAG
im distal invader 21	AGT GTG TGT ATT TT AGA TGA GGA ATA
im distal invader 22	TGG AGT TAT TTG TT AGA TGA GGA ATA
im distal invader 23	TAG TTT AAA GGG TT AGA TGA GGA ATA
im distal invader 31	AGT GTG TGT ATT TT GAG AGT AAG TAA
im distal invader 32	TGG AGT TAT TTG TT GAG AGT AAG TAA
im distal invader 33	TAG TTT AAA GGG TT GAG AGT AAG TAA
im proximal invader 11	AAA GAG ATA GGA TT AAT ACA CAC ACT
im proximal invader 12	GAA GGA TAT TGT TT AAT ACA CAC ACT
im proximal invader 13	GGG TTA ATA ATG TT AAT ACA CAC ACT
im proximal invader 21	AAA GAG ATA GGA TT CAA ATA ACT CCA
im proximal invader 22	GAA GGA TAT TGT TT CAA ATA ACT CCA
im proximal invader 23	GGG TTA ATA ATG TT CAA ATA ACT CCA
im proximal invader 31	AAA GAG ATA GGA TT CCC TTT AAA CTA
im proximal invader 32	GAA GGA TAT TGT TT CCC TTT AAA CTA
im proximal invader 33	GGG TTA ATA ATG TT CCC TTT AAA CTA

Table A.12: Sequences for intermediate (im) invader strands used for cascaded Boolean matrix multiplication. Intermediate distal invader 33 and intermediate proximal invader 33 were used in the proof-of-principle experiment depicted in Figure 3.6.

Direct displacement invader strands

Strand	Sequence ($5' \rightarrow 3'$)
dd invader 11	AAT GAA TGG GAG TT GTG GAA GAT TAG
dd invader 12	AAT GAA TGG GAG TT GGA AAT AGT GAG
dd invader 13	AAT GAA TGG GAG TT GTA GGG TAG AAG
dd invader 21	AGA TGA GGA ATA TT GTG GAA GAT TAG
dd invader 22	AGA TGA GGA ATA TT GGA AAT AGT GAG
dd invader 23	AGA TGA GGA ATA TT GTA GGG TAG AAG
dd invader 31	GAG AGT AAG TAA TT GTG GAA GAT TAG
dd invader 32	GAG AGT AAG TAA TT GGA AAT AGT GAG
dd invader 33	GAG AGT AAG TAA TT GTA GGG TAG AAG

Table A.13: Sequences for direct displacement (dd) invader strands used for 1-step reachability calculations.

Multi-strand combinatorial displacement invader strands

Strand	Sequence ($5' \rightarrow 3'$)
3/4/5-CSD 1	GAG AGT AAG TAA TT GGT GAA AAT GTT
3/4/5-CSD 2	AAC ATT TTC ACC TT CCA ATT ACC AAT
3-CSD 3	ATT GGT AAT TGG TT GTA GGG TAG AAG
4/5-CSD 3	ATT GGT AAT TGG TT GAA GGA TAT TGT
4-CSD 4	ACA ATA TCC TTC TT GTA GGG TAG AAG
5-CSD 4	ACA ATA TCC TTC TT CAT TAT TAA CCC
5-CSD 5	GGG TTA ATA ATG TT GTA GGG TAG AAG

Table A.14: Sequences for the multi-strand combinatorial displacement invader strands. The 2-CSD complex consists of distal invader 33 and proximal invader 33 in Table A.10. The 3-CSD complex consists of 3/4/5-CSD 1, 3/4/5-CSD 2 and 3-CSD 3. The 4-CSD complex consists of 3/4/5-CSD 1, 3/4/5-CSD 2, 4/5-CSD 3 and 4-CSD 4. The 5-CSD complex consists of 3/4/5-CSD 1, 3/4/5-CSD 2, 4/5-CSD 3, 5-CSD 4 and 5-CSD 5.

A.2.3 *Chapter 4*

Specialised staple replacements part 1/2

Strand	Staple	Sequence (5′ → 3′)
label top right	M9t0i	Biotin - TTT TCC TCA GAA CCG CCA CCC CCT CAG A
label top left	M-9t2i	Biotin - ACT TTT TCT TTT TTT TCA CGT TGA AAA...
		...TAT TGC GAA TAA TAA TTT TTT
stator 1	M-7t2f	GGA ACC TCA GCC CAA CTA ACA TCG TTT AAA AGG...
		...CTT ACA GAG GCT TTG AGT AAA CGG G
stator 2	M-7t2e	GGA ACC TCA GCC CAA CTA ACA TTT TGC AAC GGC...
		...CCA AAA GGA GCC TTT GGA GTG AG
stator 3	M-5t4f	GGA ACC TCA GCC CAA CTA ACA TTT TAC CTA AAA...
		...GAC GGT CAA TCA TAA AGA ACC GG
stator 4	M-5t4e	GGA ACC TCA GCC CAA CTA ACA TTT TGA GGC GCA...
		...CGA AAG AGG CAA AAG GTC ACC CT
stator 5	M-3t6f	GGA ACC TCA GCC CAA CTA ACA TTT TGC TCA TTC...
		...TTA TGC GAT TTT AAG CGA GGC AT
biotin stator 5	M-3t6f	Biotin - TTGGA ACC TCA GCC CAA CTA ACA TTT TGC...
		...TCA TTC TTA TGC GAT TTT AAG CGA GGC AT
stator 6	M-3t6e	GGA ACC TCA GCC CAA CTA ACA TTT TGA ATT ACC...
		...AGT GAA TAA GGC TTG TCG AAA TC
stator 7	M-1t8f	GGA ACC TCA GCC CAA CTA ACA TTT TAC CAG ACG...
		...GCA AAA GAA GTT TTG TTT TAA TT
stator 8	M-1t8e	GGA ACC TCA GCC CAA CTA ACA TTT TAG GCT TTT...
		...ACG ATA AAA ACC AAA ATT AAG AC
stator 9	M1t10f	GGA ACC TCA GCC CAA CTA ACA TTT TTA CAG AGA...
		...GTC AAA AAT GAA AAT AGG AAG CAA
stator 10	M1t10e	GGA ACC TCA GCC CAA CTA ACA TTT TTG TTT AAC...
		...GAA TAA CAT AAA AAC TAA TAA CG
⋮	⋮	⋮

Table A.15: Staple-replacing strands for functional tile designs. Scaffold-hybridising domains are black, track domains (i.e. those that hybridise the motor strand) are blue, domains binding anchor strand for surface immobilisation are red, and non-hybridising poly-T linkers are green. The second part of this table can be found in Table A.16.

Specialised staple replacements part 2/2

Strand	Staple	Sequence (5′ → 3′)
stator 11	M3t12f	GGA ACC TCA GCC CAA CTA ACA TTT TGG CTT ATC...
		...GCA CTC ATC GAG AAC CCG ACA AA
stator 12	M3t12e	GGA ACC TCA GCC CAA CTA ACA TTT TCA AGT ACC...
		...CGG TAT TCT AAG AAC GCC ATA TTA
stator 13	M5t14f	GGA ACC TCA GCC CAA CTA ACA TTT TTA AAC AAC...
		...AAA GCC AAC GCT CAA CTC ATC TTC
biotin stator 13	M5t14f	Biotin - TTGGA ACC TCA GCC CAA CTA ACA TTT TTA...
		...AAC AAC AAA GCC AAC GCT CAA CTC ATC TTC
stator 14	M5t14e	GGA ACC TCA GCC CAA CTA ACA TTT TAC CAG TAT...
		...ATG TTC AGC TAA TGC GGC TGT CT
stator 15	M7t16f	GGA ACC TCA GCC CAA CTA ACA TTT TCG TGT GAT...
		...TTA AGA CGC TGA GAA GTG AAT AAC
stator 16	M7t16e	GGA ACC TCA GCC CAA CTA ACA TTT TAG CTT AGA...
		...AAA TAA GGC GTT AAA AAA AGC
sl top 1	M-1t2e	ACA ATG ACC GGT CGC
sl top 2	M-1t2e	GAC CAT GAT TAC GCC AAG CTT AAA AGT ATT ATT...
		...TTT GAG GCT TCG ATT ATA
sl left 1	M-7t12f	TAT AAC AGC GGC TGT
sl left 2	M-7t12f	GAC CAT GAT TAC GCC AAG CTT AAA AGT ATT ATT...
		...TTA CCA AAA AGC CTT TAT
sl right 1	M7t12f	TGA AGC CTA ACC AAT
sl right 2	M7t12f	GAC CAT GAT TAC GCC AAG CTT AAA AGT ATT ATT...
		...TTC AAT AAT CAG AAC GCG
sl bottom 1	M1t22e	TCA CCT TGA TCT GGT
sl bottom 2	M1t22e	...GAC CAT GAT TAC GCC AAG CTT AAA AGT ATT ATT...
		...TTC AGT TGG CAA AGG AGC G

sl: surface linker

Table A.16: Continuation of Table A.15.

BIBLIOGRAPHY

[1] L. M. Adleman. Molecular computation of solutions to combinatorial problems. *Science*, 266(5187):1021–1024, Nov. 1994.

[2] A. P. Alivisatos. Semiconductor clusters, nanocrystals, and quantum dots. *Science*, 271(5251):933 –937, Feb. 1996.

[3] U. Alon. Network motifs: theory and experimental approaches. *Nature Reviews Genetics*, 8(6):450–461, June 2007.

[4] E. S. Andersen, M. Dong, M. M. Nielsen, K. Jahn, R. Subramani, W. Mamdouh, M. M. Golas, B. Sander, H. Stark, C. L. P. Oliveira, J. S. Pedersen, V. Birkedal, F. Besenbacher, K. V. Gothelf, and J. Kjems. Self-assembly of a nanoscale DNA box with a controllable lid. *Nature*, 459(7243):73–76, May 2009.

[5] T. Ando, N. Kodera, E. Takai, D. Maruyama, K. Saito, and A. Toda. A high-speed atomic force microscope for studying biological macromolecules. *Proceedings of the National Academy of Sciences of the United States of America*, 98(22):12468–12472, Oct. 2001.

[6] D. Axelrod. Cell-substrate contacts illuminated by total internal reflection fluorescence. *The Journal of Cell Biology*, 89(1):141, Apr. 1981.

[7] J. Bath, S. J. Green, and A. J. Turberfield. A free-running DNA motor powered by a nicking enzyme. *Angewandte Chemie International Edition*, 117(28):4432–4435, July 2005.

[8] J. Bath and A. J. Turberfield. DNA nanomachines. *Nature Nanotechnology*, 2(5):275–284, May 2007.

[9] N. A. W. Bell, C. R. Engst, M. Ablay, G. Divitini, C. Ducati, T. Liedl, and U. F. Keyser. DNA origami nanopores. *Nano Letters*, 12(1):512–517, Jan. 2012.

[10] Y. Benenson, B. Gil, U. Ben-Dor, R. Adar, and E. Shapiro. An autonomous molecular computer for logical control of gene expression. *Nature*, 429(6990):423–429, May 2004.

[11] E. Betzig, G. H. Patterson, R. Sougrat, O. W. Lindwasser, S. Olenych, J. S. Bonifacino, M. W. Davidson, J. Lippincott-Schwartz, and H. F. Hess. Imaging intracellular fluorescent proteins at nanometer resolution. *Science*, 313(5793):1642–1645, Sept. 2006.

[12] G. Binnig, C. F. Quate, and C. Gerber. Atomic force microscope. *Physical Review Letters*, 56(9):930–933, Mar. 1986.

[13] J. S. Bois, S. Venkataraman, H. M. T. Choi, A. J. Spakowitz, Z.-G. Wang, and N. A. Pierce. Topological constraints in nucleic acid hybridization kinetics. *Nucleic Acids Research*, 33(13):4090–4095, Jan. 2005.

[14] H. M. J. Carstairs, K. Lymperopoulos, A. N. Kapanidis, J. Bath, and A. J. Turberfield. DNA monofunctionalization of quantum dots. *ChemBioChem*, 10(11):1781–1783, June 2009.

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

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

[17] X. Chen. Expanding the rule set of DNA circuitry with associative toehold activation. *Journal of the American Chemical Society*, 134(1):263–271, Jan. 2012.

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

[19] L. S. Churchman, Z. Okten, R. S. Rock, J. F. Dawson, and J. A. Spudich. Single molecule high-resolution colocalization of cy3 and cy5 attached to macromolecules measures intramolecular distances through time. *Proceedings of the National Academy of Sciences of the United States of America*, 102(5):1419, Feb. 2005.

[20] M. L. Collins, B. Irvine, D. Tyner, E. Fine, C. Zayati, C.-a. Chang, T. Horn, D. Ahle, J. Detmer, L.-P. Shen, J. Kolberg, S. Bushnell, M. S. Urdea, and D. D. Ho. A branched DNA signal amplification assay for quantification of nucleic acid targets below 100 molecules/ml. *Nucleic Acids Research*, 25(15):2979–2984, Jan. 1997.

[21] J. P. Cooper and P. J. Hagerman. Analysis of fluorescence energy transfer in duplex and branched DNA molecules. *Biochemistry*, 29(39):9261–9268, Oct. 1990.

[22] M. Dahan, S. Levi, C. Luccardini, P. Rostaing, B. Riveau, and A. Triller. Diffusion dynamics of glycine receptors revealed by single-quantum dot tracking. *Science*, 302(5644):442–445, Oct. 2003.

[23] H. Dietz, S. M. Douglas, and W. M. Shih. Folding DNA into twisted and curved nanoscale shapes. *Science*, 325(5941):725–730, July 2009.

[24] M. v. Dijk and A. M. J. J. Bonvin. 3d-DART: a DNA structure modelling server. *Nucleic Acids Research*, 37(suppl 2):W235–W239, Jan. 2009.

[25] R. M. Dirks and N. A. Pierce. Triggered amplification by hybridization chain reaction. *Proceedings of the National Academy of Sciences of the United States of America*, 101(43):15275–15278, Oct. 2004.

[26] S. M. Douglas, I. Bachelet, and G. M. Church. A logic-gated nanorobot for targeted transport of molecular payloads. *Science*, 335(6070):831–834, Feb. 2012.

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

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

[29] J. P. K. Doye, T. E. Ouldridge, A. A. Louis, F. Romano, P. Sulc, C. Matek, B. E. K. Snodin, L. Rovigatti, J. S. Schreck, R. M. Harrison, and W. P. J. Smith. Coarse-graining DNA for simulations of DNA nanotechnology. *Physical Chemistry Chemical Physics*, 15(47):20395–20414, Nov. 2013.

[30] D. M. Eigler and E. K. Schweizer. Positioning single atoms with a scanning tunnelling microscope. *Nature*, 344(6266):524–526, Apr. 1990.

[31] M. Endo, K. Hidaka, and H. Sugiyama. Direct AFM observation of an opening event of a DNA cuboid constructed via a prism structure. *Organic & Biomolecular Chemistry*, 9(7):2075–2077, Mar. 2011.

[32] M. Endo, S. Yamamoto, T. Emura, K. Hidaka, N. Morone, J. E. Heuser, and H. Sugiyama. Helical DNA origami tubular structures

with various sizes and arrangements. *Angewandte Chemie International Edition*, 53(29):7484–7490, May 2014.

[33] L. Feng, S. H. Park, J. H. Reif, and H. Yan. A two-state DNA lattice switched by DNA nanoactuator. *Angewandte Chemie*, 115(36):4478–4482, Sept. 2003.

[34] R. Feynman. There's plenty of room at the bottom, Dec. 1959.

[35] P. Frantsuzov, M. Kuno, B. Janko, and R. A. Marcus. Universal emission intermittency in quantum dots, nanorods and nanowires. *Nature Physics*, 4(7):519–522, July 2008.

[36] C.-C. Fu, H.-Y. Lee, K. Chen, T.-S. Lim, H.-Y. Wu, P.-K. Lin, P.-K. Wei, P.-H. Tsao, H.-C. Chang, and W. Fann. Characterization and application of single fluorescent nanodiamonds as cellular biomarkers. *Proceedings of the National Academy of Sciences of the United States of America*, 104(3):727–732, Jan. 2007.

[37] A. J. Genot, J. Bath, and A. J. Turberfield. Combinatorial displacement of DNA strands: Application to matrix multiplication and weighted sums. *Angewandte Chemie International Edition*, 52(4):1189–1192, 2013.

[38] A. J. Genot, D. Y. Zhang, J. Bath, and A. J. Turberfield. Remote toehold: A mechanism for flexible control of DNA hybridization kinetics. *Journal of the American Chemical Society*, 133(7):2177–2182, Feb. 2011.

[39] N. Goldman, P. Bertone, S. Chen, C. Dessimoz, E. M. LeProust, B. Sipos, and E. Birney. Towards practical, high-capacity, low-maintenance information storage in synthesized DNA. *Nature*, 494(7435):77–80, Feb. 2013.

[40] R. P. Goodman, M. Heilemann, S. Doose, C. M. Erben, A. N. Kapanidis, and A. J. Turberfield. Reconfigurable, braced, three-dimensional DNA nanostructures. *Nature Nanotechnology*, 3(2):93–96, Feb. 2008.

[41] R. P. Goodman, I. A. T. Schaap, C. F. Tardin, C. M. Erben, R. M. Berry, C. F. Schmidt, and A. J. Turberfield. Rapid chiral assembly of rigid DNA building blocks for molecular nanofabrication. *Science*, 310(5754):1661–1665, 2005.

[42] A. Goshtasby. Image registration by local approximation methods. *Image and Vision Computing*, 6(4):255–261, Nov. 1988.

[43] S. J. Green, J. Bath, and A. J. Turberfield. Coordinated chemomechanical cycles: A mechanism for autonomous molecular motion. *Physical Review Letters*, 101(23):238101, Dec. 2008.

[44] S. J. Green, D. Lubrich, and A. J. Turberfield. DNA hairpins: Fuel for autonomous DNA devices. *Biophysical Journal*, 91(8):2966–2975, Oct. 2006.

[45] S. Griffith, D. Goldwater, and J. M. Jacobson. Robotics: Self-replication from random parts. *Nature*, 437(7059):636–636, Sept. 2005.

[46] H. Gu, J. Chao, S.-J. Xiao, and N. C. Seeman. A proximity-based programmable DNA nanoscale assembly line. *Nature*, 465(7295):202–205, May 2010.

[47] H. Hess and E. L. P. Dumont. Fatigue failure and molecular machine design. *Small*, 7(12):1619–1623, June 2011.

[48] S. Hohng and T. Ha. Near-complete suppression of quantum dot blinking in ambient conditions. *Journal of the American Chemical Society*, 126(5):1324–1325, 2004.

[49] J. J. Hopfield. Kinetic proofreading: a new mechanism for reducing errors in biosynthetic processes requiring high specificity. *Proceedings of the National Academy of Sciences of the United States of America*, 71(10):4135–4139, Oct. 1974.

[50] T. Hu and M. Shing. Computation of matrix chain products. part i. *SIAM Journal on Computing*, 11(2):362–373, May 1982.

[51] T. Hu and M. Shing. Computation of matrix chain products. part II. *SIAM Journal on Computing*, 13(2):228–251, May 1984.

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

[53] J. D. Irwin and R. M. Nelms. *Engineering Circuit Analysis*. John Wiley and Sons, New York, international student version, 10th edition, 2011.

[54] L. Jauffred, A. C. Richardson, and L. B. Oddershede. Three-dimensional optical control of individual quantum dots. *Nano Letters*, 8(10):3376–3380, Oct. 2008.

[55] Y. S. Jiang, S. Bhadra, B. Li, and A. D. Ellington. Mismatches improve the performance of strand-displacement nucleic acid circuits. *Angewandte Chemie International Edition*, 53(7):1845–1848, Feb. 2014.

[56] F. Juelicher and J. Prost. Cooperative molecular motors. *Physical Review Letters*, 75(13):2618–2621, Sept. 1995.

[57] R. Jungmann, M. S. Avendano, J. B. Woehrstein, M. Dai, W. M. Shih, and P. Yin. Multiplexed 3d cellular super-resolution imaging with DNA-PAINT and exchange-PAINT. *Nature Methods*, 11(3):313–318, Mar. 2014.

[58] R. Jungmann, C. Steinhauer, M. Scheible, A. Kuzyk, P. Tinnefeld, and F. C. Simmel. Single-molecule kinetics and super-resolution microscopy by fluorescence imaging of transient binding on DNA origami. *Nano Letters*, 10(11):4756–4761, Oct. 2010.

[59] N. R. Kallenbach, R.-I. Ma, and N. C. Seeman. An immobile nucleic acid junction constructed from oligonucleotides. *Nature*, 305(5937):829–831, Oct. 1983.

[60] Y. Ke, L. L. Ong, W. M. Shih, and P. Yin. Three-dimensional structures self-assembled from DNA bricks. *Science*, 338(6111):1177–1183, Nov. 2012.

[61] Y. Krishnan and F. C. Simmel. Nucleic acid based molecular devices. *Angewandte Chemie International Edition*, 50(14):3124–3156, Mar. 2011.

[62] A. T. Krueger and E. T. Kool. Redesigning the architecture of the base pair: Toward biochemical and biological function of new genetic sets. *Chemistry & Biology*, 16(3):242–248, Mar. 2009.

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

[64] M. Langecker, V. Arnaut, T. G. Martin, J. List, S. Renner, M. Mayer, H. Dietz, and F. C. Simmel. Synthetic lipid membrane channels formed by designed DNA nanostructures. *Science*, 338(6109):932–936, Nov. 2012.

[65] J. Li, C. Fan, H. Pei, J. Shi, and Q. Huang. Smart drug delivery nanocarriers with self-assembled DNA nanostructures. *Advanced Materials*, 25(32):4386–4396, 2013.

[66] Q. Li, G. Luan, Q. Guo, and J. Liang. A new class of homogeneous nucleic acid probes based on specific displacement hybridization. *Nucleic Acids Research*, 30(2):e5, Jan. 2002.

[67] X. Liang, H. Nishioka, N. Takenaka, and H. Asanuma. A DNA nanomachine powered by light irradiation. *ChemBioChem*, 9(5):702–705, Mar. 2008.

[68] K. Lidke, B. Rieger, T. Jovin, and R. Heintzmann. Superresolution by localization of quantum dots using blinking statistics. *Optics Express*, 13(18):7052–7062, Sept. 2005.

[69] C. Lin, R. Jungmann, A. M. Leifer, C. Li, D. Levner, G. M. Church, W. M. Shih, and P. Yin. Submicrometre geometrically encoded fluorescent barcodes self-assembled from DNA. *Nature Chemistry*, 4(10):832–839, Aug. 2012.

[70] V. Linko and H. Dietz. The enabled state of DNA nanotechnology. *Current Opinion in Biotechnology*, 24(4):555–561, Aug. 2013.

[71] D. Liu and S. Balasubramanian. A proton-fuelled DNA nanomachine. *Angewandte Chemie International Edition*, 42(46):5734–5736, Dec. 2003.

[72] K. Lund, A. J. Manzo, N. Dabby, N. Michelotti, A. Johnson-Buck, J. Nangreave, S. Taylor, R. Pei, M. N. Stojanovic, N. G. Walter, E. Winfree, and H. Yan. Molecular robots guided by prescriptive landscapes. *Nature*, 465(7295):206–210, May 2010.

[73] R. R. F. Machinek, T. E. Ouldridge, N. E. C. Haley, J. Bath, and A. J. Turberfield. Programmable energy landscapes for kinetic control of DNA strand displacement. *Nature Communications*, 5:5324, Nov. 2014.

[74] J. Malo, J. C. Mitchell, C. Venien-Bryan, J. R. Harris, H. Wille, D. J. Sherratt, and A. J. Turberfield. Engineering a 2d protein-DNA crystal. *Angewandte Chemie International Edition*, 44(20):3057–3061, May 2005.

[75] C. Mao, T. H. LaBean, J. H. Reif, and N. C. Seeman. Logical computation using algorithmic self-assembly of DNA triple-crossover molecules. *Nature*, 407(6803):493–496, Sept. 2000.

[76] C. Mao, W. Sun, Z. Shen, and N. C. Seeman. A nanomechanical device based on the b-z transition of DNA. *Nature*, 397(6715):144–146, Jan. 1999.

[77] S. A. E. Marras, F. R. Kramer, and S. Tyagi. Efficiencies of fluorescence resonance energy transfer and contact-mediated quenching in oligonucleotide probes. *Nucleic Acids Research*, 30(21):e122, Jan. 2002.

[78] R. Masoud, R. Tsukanov, T. E. Tomov, N. Plavner, M. Liber, and E. Nir. Studying the structural dynamics of bipedal DNA motors with single-molecule fluorescence spectroscopy. *ACS Nano*, 6(7):6272–6283, July 2012.

[79] M. Mazumdar and J. R. Glassman. Categorizing a prognostic variable: review of methods, code for easy implementation and applications to decision-making about cancer treatments. *Statistics in Medicine*, 19(1):113–132, Jan. 2000.

[80] M. L. McKee, P. J. Milnes, J. Bath, E. Stulz, R. K. OReilly, and A. J. Turberfield. Programmable one-pot multistep organic synthesis using DNA junctions. *Journal of the American Chemical Society*, 134(3):1446–1449, Jan. 2012.

[81] X. Michalet, F. F. Pinaud, L. A. Bentolila, J. M. Tsay, S. Doose, J. J. Li, G. Sundaresan, A. M. Wu, S. S. Gambhir, and S. Weiss. Quantum dots for live cells, in vivo imaging, and diagnostics. *Science*, 307(5709):538–544, Jan. 2005.

[82] K. U. Mir. A restricted genetic alphabet for DNA computing. *DNA Based Computers II*, pages 243–246, 1999.

[83] S. Modi, C. Nizak, S. Surana, S. Halder, and Y. Krishnan. Two DNA nanomachines map pH changes along intersecting endocytic pathways inside the same cell. *Nature Nanotechnology*, 8(6):459–467, June 2013.

[84] B. G. Moreira, Y. You, M. A. Behlke, and R. Owczarzy. Effects of fluorescent dyes, quenchers, and dangling ends on DNA duplex stability. *Biochemical and biophysical research communications*, 327(2):473–484, Feb. 2005.

[85] E. A. Murphy and H. Abbey. The normal range - a common misuse. *Journal of Chronic Diseases*, 20(2):79–88, Feb. 1967.

[86] R. A. Muscat, J. Bath, and A. J. Turberfield. A programmable molecular robot. *Nano Letters*, 11(3):982–987, Jan. 2011.

[87] J. Ninio. Kinetic amplification of enzyme discrimination. *Biochimie*, 57(5):587–595, July 1975.

[88] T. Omabegho, R. Sha, and N. C. Seeman. A bipedal DNA brownian motor with coordinated legs. *Science*, 324(5923):67–71, Mar. 2009.

[89] T. E. Ouldridge, A. A. Louis, and J. P. K. Doye. DNA nanotweezers studied with a coarse-grained model of DNA. *Physical Review Letters*, 104(17):178101, Apr. 2010.

[90] T. E. Ouldridge, A. A. Louis, and J. P. K. Doye. Structural, mechanical, and thermodynamic properties of a coarse-grained DNA model. *The Journal of Chemical Physics*, 134(8):085101, Feb. 2011.

[91] T. E. Ouldridge, P. Sulc, F. Romano, J. P. K. Doye, and A. A. Louis. DNA hybridization kinetics: zippering, internal displacement and sequence dependence. *Nucleic Acids Research*, 41(19):8886–8895, Jan. 2013.

[92] Q. Ouyang, P. D. Kaplan, S. Liu, and A. Libchaber. DNA solution of the maximal clique problem. *Science*, 278(5337):446–449, Oct. 1997.

[93] I. G. Panyutin and P. Hsieh. Formation of a single base mismatch impedes spontaneous DNA branch migration. *Journal of Molecular Biology*, 230(2):413–424, Mar. 1993.

[94] E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, and T. E. Ferrin. UCSF chimera–a visualization system for exploratory research and analysis. *Journal of Computational Chemistry*, 25(13):1605–1612, Oct. 2004.

[95] N. Peyret, P. A. Seneviratne, H. T. Allawi, and J. SantaLucia. Nearest-neighbor thermodynamics and NMR of DNA sequences with internal a-a, c-c, g-g, and t-t mismatches. *Biochemistry*, 38(12):3468–3477, Mar. 1999.

[96] E. Pfitzner, C. Wachauf, F. Kilchherr, B. Pelz, W. M. Shih, M. Rief, and H. Dietz. Rigid DNA beams for high-resolution single-molecule mechanics. *Angewandte Chemie International Edition*, 52(30):7766–7771, 2013.

[97] L. Qian and E. Winfree. Scaling up digital circuit computation with DNA strand displacement cascades. *Science*, 332(6034):1196–1201, Mar. 2011.

[98] L. Qian and E. Winfree. A simple DNA gate motif for synthesizing large-scale circuits. *Journal of The Royal Society Interface*, 8(62):1281–1297, July 2011.

[99] L. Qian, E. Winfree, and J. Bruck. Neural network computation with DNA strand displacement cascades. *Nature*, 475(7356):368–372, July 2011.

[100] A. Rajendran, M. Endo, K. Hidaka, and H. Sugiyama. Direct and real-time observation of rotary movement of a DNA nanomechanical device. *Journal of the American Chemical Society*, 135(3):1117–1123, Jan. 2013.

[101] U. Resch-Genger, M. Grabolle, S. Cavaliere-Jaricot, R. Nitschke, and T. Nann. Quantum dots versus organic dyes as fluorescent labels. *Nature Methods*, 5(9):763–775, Aug. 2008.

[102] F. Romano, D. Chakraborty, J. P. K. Doye, T. E. Ouldridge, and A. A. Louis. Coarse-grained simulations of DNA overstretching. *The Journal of Chemical Physics*, 138(8):085101, Feb. 2013.

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

[104] M. J. Rust, M. Bates, and X. Zhuang. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nature Methods*, 3(10):793–796, Oct. 2006.

[105] M. Salomo, K. Kegler, C. Gutsche, M. Struhalla, J. Reinmuth, W. Skokow, U. Hahn, and F. Kremer. The elastic properties of single double-stranded DNA chains of different lengths as measured with optical tweezers. *Colloid and Polymer Science*, 284(11):1325–1331, May 2006.

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

[107] J. SantaLucia and D. Hicks. The thermodynamics of DNA structural motifs. *Annual Review of Biophysics and Biomolecular Structure*, 33(1):415–440, June 2004.

[108] C. C. Santini, J. Bath, A. J. Turberfield, and A. M. Tyrrell. A DNA network as an information processing system. *International Journal of Molecular Sciences*, 13(4):5125–5137, Apr. 2012.

[109] M. B. Scheible, G. Pardatscher, A. Kuzyk, and F. C. Simmel. Single molecule characterization of DNA binding and strand displacement reactions on lithographic DNA origami microarrays. *Nano Letters*, 14(3):1627–1633, Feb. 2014.

[110] R. Schreiber, J. Do, E.-M. Roller, T. Zhang, V. J. Schueller, P. C. Nickels, J. Feldmann, and T. Liedl. Hierarchical assembly of metal nanoparticles, quantum dots and organic dyes using DNA origami scaffolds. *Nature Nanotechnology*, 9(1):74–78, Jan. 2014.

[111] G. Seelig, D. Soloveichik, D. Y. Zhang, and E. Winfree. Enzyme-free nucleic acid logic circuits. *Science*, 314(5805):1585–1588, Dec. 2006.

[112] G. Seelig, B. Yurke, and E. Winfree. Catalyzed relaxation of a metastable DNA fuel. *Journal of the American Chemical Society*, 128(37):12211–12220, Sept. 2006.

[113] N. C. Seeman. Nucleic acid junctions and lattices. *Journal of Theoretical Biology*, 99(2):237–247, Nov. 1982.

[114] D. N. Selmi, R. J. Adamson, H. Attrill, A. D. Goddard, R. J. C. Gilbert, A. Watts, and A. J. Turberfield. DNA-templated protein arrays for single-molecule imaging. *Nano Letters*, 11(2):657–660, Feb. 2011.

[115] R. F. Service. DNA nanotechnology grows up. *Science*, 332(6034):1140–1143, Mar. 2011.

[116] W. Shen, M. F. Bruist, S. D. Goodman, and N. C. Seeman. A protein-driven DNA device that measures the excess binding energy of proteins that distort DNA. *Angewandte Chemie International Edition*, 43(36):4750–4752, Sept. 2004.

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

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

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

[120] J.-P. J. Sobczak, T. G. Martin, T. Gerling, and H. Dietz. Rapid folding of DNA into nanoscale shapes at constant temperature. *Science*, 338(6113):1458–1461, Dec. 2012.

[121] D. Soloveichik, G. Seelig, and E. Winfree. DNA as a universal substrate for chemical kinetics. *Proceedings of the National Academy of Sciences of the United States of America*, 107(12):5393–5398, Mar. 2010.

[122] N. Srinivas, T. E. Ouldridge, P. Sulc, J. M. Schaeffer, B. Yurke, A. A. Louis, J. P. K. Doye, and E. Winfree. On the biophysics and kinet-

ics of toehold-mediated DNA strand displacement. *Nucleic Acids Research*, 41(22):10641–10658, Sept. 2013.

[123] C. Steinhauer, R. Jungmann, T. L. Sobey, F. C. Simmel, and P. Tinnefeld. DNA origami as a nanoscopic ruler for super-resolution microscopy. *Angewandte Chemie International Edition*, 48(47):8870–8873, Nov. 2009.

[124] M. N. Stojanovic and D. Stefanovic. A deoxyribozyme-based molecular automaton. *Nature Biotechnology*, 21(9):1069–1074, Sept. 2003.

[125] L. Stryer. Fluorescence energy transfer as a spectroscopic ruler. *Annual Review of Biochemistry*, 47(1):819–846, 1978.

[126] H. K. K. Subramanian, B. Chakraborty, R. Sha, and N. C. Seeman. The label-free unambiguous detection and symbolic display of single nucleotide polymorphisms on DNA origami. *Nano Letters*, 11(2):910–913, Feb. 2011.

[127] A. Sukhanova, J. Devy, L. Venteo, H. Kaplan, M. Artemyev, V. Oleinikov, D. Klinov, M. Pluot, J. H. M. Cohen, and I. Nabiev. Biocompatible fluorescent nanocrystals for immunolabeling of membrane proteins and cells. *Analytical Biochemistry*, 324(1):60–67, Jan. 2004.

[128] P. Sulc, T. E. Ouldridge, F. Romano, J. P. Doye, and A. A. Louis. Simulating a burnt-bridges DNA motor with a coarse-grained DNA model. *Natural Computing*, pages 1–13, Aug. 2012.

[129] C. Switzer, S. E. Moroney, and S. A. Benner. Enzymatic incorporation of a new base pair into DNA and RNA. *Journal of the American Chemical Society*, 111(21):8322–8323, Oct. 1989.

[130] R. E. Thompson, D. R. Larson, and W. W. Webb. Precise nanometer localization analysis for individual fluorescent probes. *Biophysical Journal*, 82(5):2775–2783, May 2002.

[131] B. Tinland, A. Pluen, J. Sturm, and G. Weill. Persistence length of single-stranded DNA. *Macromolecules*, 30(19):5763–5765, Sept. 1997.

[132] M. Torimura, S. Kurata, K. Yamada, T. Yokomaku, Y. Kamagata, T. Kanagawa, and R. Kurane. Fluorescence-quenching phenomenon by photoinduced electron transfer between a fluorescent dye and a nucleotide base. *Analytical Sciences*, 17(1):155–160, Jan. 2001.

[133] A. J. Turberfield, J. C. Mitchell, B. Yurke, A. P. Mills Jr, M. I. Blakey, and F. C. Simmel. DNA fuel for free-running nanomachines. *Physical Review Letters*, 90(11):118102, Mar. 2003.

[134] S. Tyagi and F. R. Kramer. Molecular beacons: probes that fluoresce upon hybridization. *Nature Biotechnology*, 14(3):303–308, Mar. 1996.

[135] A. S. Walsh, H. Yin, C. M. Erben, M. J. A. Wood, and A. J. Turberfield. DNA cage delivery to mammalian cells. *ACS Nano*, 5(7):5427–5432, July 2011.

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

[137] Y. Wang, G. Fruhwirth, E. Cai, T. Ng, and P. R. Selvin. 3d super-resolution imaging with blinking quantum dots. *Nano Letters*, 13(11):5233–5241, Nov. 2013.

[138] D. M. Warshaw, G. G. Kennedy, S. S. Work, E. B. Krementsova, S. Beck, and K. M. Trybus. Differential labeling of myosin v heads

with quantum dots allows direct visualization of hand-over-hand processivity. *Biophysical Journal*, 88(5):L30–L32, May 2005.

[139] J. D. Watson and F. H. Crick. Molecular structure of nucleic acids. *Nature*, 171(4356):737–738, Apr. 1953.

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

[141] S. F. J. Wickham. *DNA Origami: A Substrate for the Study of Molecular Motors*. DPhil thesis, University of Oxford, Oxford, United Kingdom, 2011.

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

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

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

[145] N. Wu, D. M. Czajkowsky, J. Zhang, J. Qu, M. Ye, D. Zeng, X. Zhou, J. Hu, Z. Shao, B. Li, and C. Fan. Molecular threading and tunable molecular recognition on DNA origami nanostructures. *Journal of the American Chemical Society*, 135(33):12172–12175, Aug. 2013.

[146] P. G. Wu and L. Brand. Resonance energy transfer: Methods and applications. *Analytical Biochemistry*, 218(1):1–13, Apr. 1994.

[147] R. Yasuda, H. Noji, M. Yoshida, K. Kinosita, and H. Itoh. Resolution of distinct rotational substeps by submillisecond kinetic analysis of f1-ATPase. *Nature*, 410(6831):898–904, Apr. 2001.

[148] A. Yildiz, J. N. Forkey, S. A. McKinney, T. Ha, Y. E. Goldman, and P. R. Selvin. Myosin v walks hand-over-hand: Single fluorophore imaging with 1.5-nm localization. *Science*, 300(5628):2061 –2065, June 2003.

[149] P. Yin, H. Yan, X. G. Daniell, A. J. Turberfield, and J. H. Reif. A unidirectional DNA walker that moves autonomously along a track. *Angewandte Chemie International Edition*, 43(37):4906–4911, Sept. 2004.

[150] B. Yurke and A. P. Mills Jr. Using DNA to power nanostructures. *Genetic Programming and Evolvable Machines*, 4(2):111–122, June 2003.

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

[152] J. N. Zadeh, C. D. Steenberg, J. S. Bois, B. R. Wolfe, M. B. Pierce, A. R. Khan, R. M. Dirks, and N. A. Pierce. NUPACK: Analysis and design of nucleic acid systems. *Journal of Computational Chemistry*, 32(1):170–173, Jan. 2011.

[153] D. Y. Zhang. Cooperative hybridization of oligonucleotides. *Journal of the American Chemical Society*, 133(4):1077–1086, Feb. 2011.

[154] D. Y. Zhang, S. X. Chen, and P. Yin. Optimizing the specificity of nucleic acid hybridization. *Nature Chemistry*, 4(3):208–214, Mar. 2012.

[155] D. Y. Zhang, A. J. Turberfield, B. Yurke, and E. Winfree. Engineering entropy-driven reactions and networks catalyzed by DNA. *Science*, 318(5853):1121–1125, Nov. 2007.

[156] D. Y. Zhang and E. Winfree. Control of DNA strand displacement kinetics using toehold exchange. *Journal of the American Chemical Society*, 131(47):17303–17314, Dec. 2009.

[157] D. Y. Zhang and E. Winfree. Robustness and modularity properties of a non-covalent DNA catalytic reaction. *Nucleic Acids Research*, 38(12):4182–4197, Jan. 2010.

[158] Y. Zhang and N. C. Seeman. Construction of a DNA-truncated octahedron. *Journal of the American Chemical Society*, 116(5):1661–1669, Mar. 1994.

[159] J. Zheng, J. J. Birktoft, Y. Chen, T. Wang, R. Sha, P. E. Constantinou, S. L. Ginell, C. Mao, and N. C. Seeman. From molecular to macroscopic via the rational design of a self-assembled 3d DNA crystal. *Nature*, 461(7260):74–77, Sept. 2009.

[160] J. Zhu, L. Zhang, S. Dong, and E. Wang. Four-way junction-driven DNA strand displacement and its application in building majority logic circuit. *ACS Nano*, 7(11):10211–10217, Nov. 2013.

[161] V. Zykov, E. Mytilinaios, B. Adams, and H. Lipson. Robotics: Self-reproducing machines. *Nature*, 435(7039):163–164, May 2005.