Tethered Fluorophore Motion Studies of DNA Segregation Machinery

Peter F J May

Balliol College
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy

Trinity 2015

Tethered fluorophore motion (TFM) is a single-molecule fluorescence technique, used to monitor an effective length along a biopolymer, such as DNA. In this work, I start by characterising TFM, focussing on the theoretical basis and potential applications, and then use TFM to study the bacterial site specific recombination system, XerCD-*dif*. I present a description of the sources of noise in TFM and derive an analytical expression for the resolution, which is compared to simulations and experiments. The work demonstrates that length changes as low as 100 bp of double-stranded DNA are distinguishable using TFM. The optimum pixel size for TFM is derived and validated experimentally.

XerCD-*dif* recombination is responsible for chromosome decatenation in most bacteria with circular chromosomes. It is activated by the DNA translocase FtsK. I describe the application of TFM, in combination with Förster resonance energy transfer (FRET) and protein induced fluorescence enhancement, to observe the formation and activation of XerCD-*dif* synaptic complexes. The work followed the reaction, as Holliday junctions formed by XerD-mediated strand exchange isomerised and were resolved to product by XerC, and measured the lifetimes of all the intermediates observed.

To analyse the role of FtsK in the activation of recombination, and to monitor its assembly, translocation and behaviour when encountering XerCD-*dif*, TFM was extended with the use of two spectrally distinct fluorophores, allowing simultaneous monitoring of two effective lengths along the same DNA. FtsK was observed to assemble on DNA, stepwise, into a single hexamer, and begin translocation rapidly ($\sim$0.25 s). Single hexamers approached XerCD-*dif*, and resided for 0.5 s to 1 s before dissociating, irrespective of whether XerCD-*dif* was synapsed or recombination was activated.

The final section of this thesis details the design, construction and evaluation of a three-colour microscope. Alternating laser excitation was implemented in the green, red and near-infrared spectral channels with three laser sources, and an inexpensive automatic focus system was implemented using the back reflection of the excitation lasers and a complementary metal oxide semiconductor camera.

Tethered Fluorophore Motion Studies of DNA Segregation Machinery

Peter F J May

Balliol College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity 2015

Acknowledgements

I would like to thank: Achillefs Kapanidis for his kind supervision, and his willingness to indulge my desire to gain experience in hardware construction; Dave Sherratt and Lidia Arciszewska for their immense biological knowledge and their eagerness to share it with me; Pawel Zawadzki and Justin Pinkney for developing tethered fluorophore motion, and for teaching me how to perform single-molecule experiments; the MathWorks for their financial support and Jos Martin for his mentoring during my internships.

I would like to also thank all members, past and present, of the Kapanidis and Sherratt groups who have made my time so enjoyable over the past three years, and especially the members of the downstairs office who endured, and even grown to accept, my company. Also deserving of a mention are the people who have facilitated my education so far: Mr King, Keith Grainge, Laura Kormos and Ulrich Keyser to name but a few.

I'd like to thank my parents, who have invested (in many ways) so much in my education, and my friends, including those whom I met during my undergraduate studies, and those whom I met during my time in the Balliol MCR. I am grateful to Ioanna, for her support and patience.

Tethered Fluorophore Motion Studies of DNA Segregation Machinery

Peter F J May

Balliol College
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy

Trinity 2015

Tethered fluorophore motion (TFM) is a single-molecule fluorescence technique, used to monitor an effective length along a biopolymer, such as DNA. In this work, I start by characterising TFM, focussing on the theoretical basis and potential applications, and then use TFM to study the bacterial site specific recombination system, XerCD-*dif*. I present a description of the sources of noise in TFM and derive an analytical expression for the resolution, which is compared to simulations and experiments. The work demonstrates that length changes as low as 100 bp of double-stranded DNA are distinguishable using TFM. The optimum pixel size for TFM is derived and validated experimentally.

XerCD-*dif* recombination is responsible for chromosome decatenation in most bacteria with circular chromosomes. It is activated by the DNA translocase FtsK. I describe the application of TFM, in combination with Förster resonance energy transfer (FRET) and protein induced fluorescence enhancement, to observe the formation and activation of XerCD-*dif* synaptic complexes. The work followed the reaction, as Holliday junctions formed by XerD-mediated strand exchange isomerised and were resolved to product by XerC, and measured the lifetimes of all the intermediates observed.

To analyse the role of FtsK in the activation of recombination, and to monitor its assembly, translocation and behaviour when encountering XerCD-*dif*, TFM was extended with the use of two spectrally distinct fluorophores, allowing simultaneous monitoring of two effective lengths along the same DNA. FtsK was observed to assemble on DNA, stepwise, into a single hexamer, and begin translocation rapidly ($\sim$0.25 s). Single hexamers approached XerCD-*dif*, and resided for 0.5 s to 1 s before dissociating, irrespective of whether XerCD-*dif* was synapsed or recombination was activated.

The final section of this thesis details the design, construction and evaluation of a three-colour microscope. Alternating laser excitation was implemented in the green, red and near-infrared spectral channels with three laser sources, and an inexpensive automatic focus system was implemented using the back reflection of the excitation lasers and a complementary metal oxide semiconductor camera.

Disclaimer

I hereby declare that the work in this thesis is my own, except where indicated in the text and as described below:

- The two-colour total internal reflection fluorescence microscope, Half-dome, used in this thesis was designed and built by Ludovic Le Reste and Johannes Hohlbein.
- The analysis software, used for two-colour image processing, was written by Seamus Holden, and the software used for hidden Markov modelling was written by Kristofer Gryte.
- The work in Chapter 2 forms the basis of previously published work [1], and the contributions to this work are described in Section 2.5.
- The work in Chapter 3 forms the basis of previously published work [2], and the contributions to this work are described in Section 3.6.
- The work in Chapter 4 forms the basis of previously published work [3], and the contributions to this work are described in Section 4.6.

Contents

List of Figures

Chapter 1

Introduction

1.1 Introduction

The current understanding of how molecular machines function has required a combination of biochemical, structural and biophysical approaches. This thesis focusses on one recently developed biophysical technique, tethered fluorophore motion (TFM). Chapter 2 presents a characterisation of the theoretical basis, the resolution and the potential applications of this technique. Chapters 3 and 4 describe the expansion and application of the technique to further the understanding of the XerCD-*dif*-FtsK chromosome maintenance and segregation machinery, found in bacteria. Finally, Chapter 5 describes the design, construction and evaluation of an instrument intended to permit the continued expansion of the technique, and the study of XerCD-*dif*-FtsK and other multi-protein molecular machines in increasing detail.

1.2 Single-molecule biophysics

In recent years, single-molecule techniques have allowed the observation of biological reactions in real-time, free from the constraints of ensemble averaging [4, 5]. They include: optical trapping [6], magnetic tweezers [7], tethered particle motion (TPM) [8], atomic

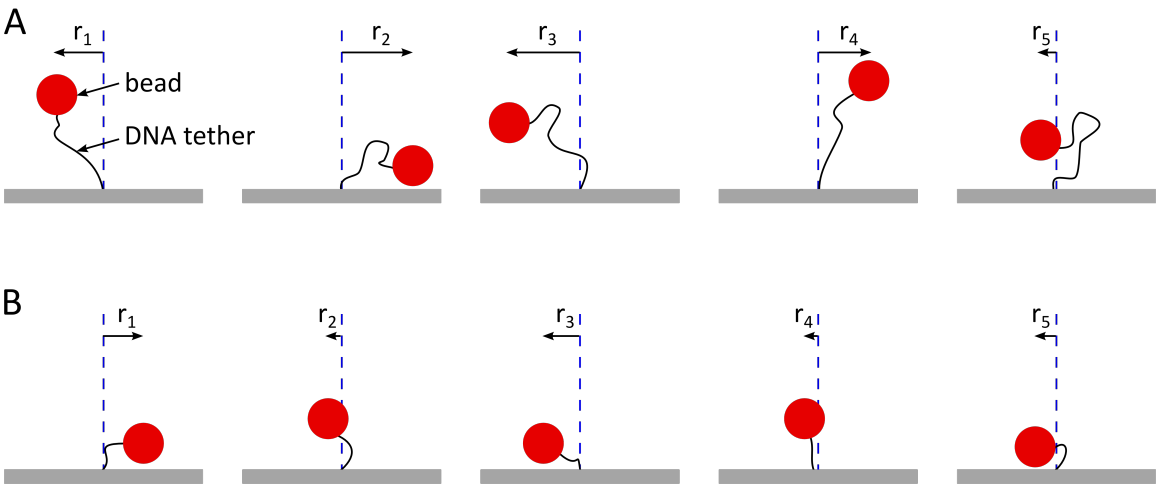

Figure 1.1: Tethered particle motion uses the diffusional motion of a bead to report on the effective length of a DNA tether. **(A)** A longer DNA gives the bead relative freedom to move around in the vicinity of the tether point. The mean excursion can be computed from a series of images of the bead taken using an optical microscope. **(B)** A shorter DNA constrains the bead movement, and reduces the mean excursion.

force microscopy [9], and fluorescence techniques such as those utilising Förster resonance energy transfer (FRET) [10] and the in vivo tracking of labelled biomolecules [11, 12]. The techniques can be broadly distinguished into force-based and fluorescence-based [13]. The force-based category includes magnetic tweezers, optical trapping and atomic force microscopy.

1.2.1 Tethered particle motion

One technique that neither actively applies a force, nor necessarily uses fluorescence, is TPM. It is used to observe conformational changes in biopolymers, such as DNA. TPM experiments typically observe a bead (20 nm to 50 nm in diameter) attached to the end of a long ($>$ 400 bp) DNA molecule, which is tethered at its other end to a surface [8]. The apparent length of the DNA molecule can be inferred from the Brownian motion of the bead (Figure 1.1) [14]. TPM has been previously used to study protein-DNA interactions including: *lac* repressor induced DNA looping [15], transcription [8], restriction enzymes [16], and the polymer dynamics of DNA [17, 18].

Decreasing the size of the bead in TPM has allowed measurements with precision as low as 40 bp [19]. However, TPM is still limited by the entropic stretching force due to the bead, which alters the distribution of DNA conformations [20], and the hydrodynamic drag acting on the bead, which reduces enzymatic rates [21]. In a practical sense, the bead represents a large surface that requires passivation to prevent biomolcules sticking to it, and it can attach multiple DNA tethers simultaneously.

1.2.2 Increasing the bandwidth

Since biological reactions often involve multiple proteins and nucleic acids interacting with each other, methods capable of tracking several observables simultaneously have been established. Previous work has combined optical trapping and FRET to simultaneously monitor distances within a DNA construct under an applied force [22], and has combined FRET with patch-clamping to study ion channel conformational changes [23]. Often combination measurements are made difficult by the need for apparatus for every technique to interact with the sample. However, experiments using only FRET have been able to simultaneously monitor several distances within a biological complex by utilising three or four spectrally distinct labels [24, 25, 26].

1.3 Fluorescence

Fluorescent molecules can absorb light over a specific range of wavelengths and enter an excited state. Once excited, they decay back to their ground state and can emit a photon, whose wavelength is usually longer due to vibrational relaxation. Since there is this difference in excitation and emission wavelengths, termed the Stokes shift, the intense illumination light can be filtered using dichroic mirrors and the fluorescence emission collected. Unfortunately, biological molecules do not usually exhibit significant fluorescence, so, to be able to visualise them, they need to be labelled with a fluorophore. Specific

biomolecules within cells and tissues can be fluorescently labelled using many different approaches, including with DNA intercalators, fluorescein conjugated antibodies, or through conjugation with fluorescent proteins when expressed within the cell. In vitro, proteins and DNA can be labelled at specific reactive sites, introduced during their synthesis, with the use of reactive derivatives (e.g. NHS-esters or maleimides) of organic fluorophores, such as Cyanine 3B (Cy3B), chosen for their photostability and brightness [27, 28].

1.3.1 Photophysics

Organic fluorophores commonly used for single molecule experiments in vitro present practical difficulties. They exhibit an irreversible cessation of fluorescence, termed photobleaching, due to photo-chemical disruption of the dye. Photobleaching is thought to commonly occur due to the reaction of a dye with oxygen radicals in solution [29]. However, these oxygen radicals are thought to quench triplet states of the dye, which are responsible for long-lived dark states (photoblinking) [29]. Hence, to optimise dye performance, buffers containing a system that removes oxygen from solution, and an alternative triplet state quencher, are used. One such system for oxygen-scavenging consists of two enzymes, glucose oxidase and catalase [30]. Buffers containing Trolox (a water-soluble analog of vitamin E) and Trolox-quinone can be used as a triplet state quencher [31].

1.3.2 Tethered fluorophore motion

To combine the ability of TPM to monitor an effective length along DNA with single-molecule fluorescence, TFM has been introduced, and implemented to study site-specific recombination by Cre-*loxP* [32]. TFM uses an organic fluorophore in the place of the bead in TPM. These fluorophores are small (e.g. Cy5, a commonly used fluorophore, has a 1 nm backbone length [33]), and hence their use minimises hydrodynamic drag and entropic stretching. Fluorophores can be attached using a range of covalent chemistries, which reliably attach one DNA per fluorophore. Importantly, the fluorophore used in TFM can

also be used for other fluorescence techniques, ranging from simple colocalisation to FRET.

1.3.3 FRET

Förster resonance energy transfer (FRET) is the non-radiative transfer of energy from an excited, donor, fluorophore to an acceptor fluorophore. Since it is dipole-dipole coupling and, using Fermi's golden rule (for final state ϕ_{DA^*} and initial state ϕ_{D^*A}), the transition rate is proportional to $|\langle\phi_{DA^*}|\hat{H}'|\phi_{D^*A}\rangle|^2$, where $\hat{H}'$, the perturbing Hamiltonian for fluorophores separated by a distance r, scales as $\frac{1}{r^3}$ for a dipole perturbation, the transition rate is proportional to $\frac{1}{r^6}$. Hence, the efficiency of energy transfer, E, can be defined as [34]:

$$E = \frac{1}{1 + \left(\frac{r}{R_0}\right)^6} \tag{1.1}$$

with the Förster radius, R_0 characterises the length scale over which FRET occurs significantly, and usually is in the range 5 nm to 7 nm for organic fluorophores [35]. The Förster radius depends on the overlap between the emission spectrum of the donor and the absorbtion spectum of the acceptor, the relative orientations of the two fluorophores and the quantum yield of the donor.

To use FRET as a single-molecule technique, biomolecules can be labelled with a donor and an acceptor fluorophore. The donor is typically chosen to absorb and emit at shorter wavelengths than the acceptor, and so a light source (e.g. a laser) can be chosen that preferentially excites the donor. The degree of energy transfer can be inferred from the relative brightness of the donor and acceptor fluorophores. The apparent FRET efficiency, E^* is defined:

$$E^* = \frac{DA}{DD + DA} \tag{1.2}$$

where DA is the acceptor intensity under excitation of the donor, and DD is the

donor intensity. Unfortunately, due to crosstalks in the imaging system, and differing probabilities of observing an excitation of the donor and the acceptor, it is non-trivial to relate Equation (1.1) and Equation (1.2) [36]. This conversion of an apparent FRET efficiency to a distance between fluorophores is described in Section 3.2.8 of Chapter 3. However, even without knowing the precise distances involved, FRET is still useful for labelling structural states in single-molecules and observing transitions and subpopulations.

With only a single excitation source, capable of exciting the donor, it is difficult to distinguish a molecule in a low FRET state from one where the acceptor is not photo-active (due to photobleaching or blinking). Alternating laser excitation (ALEX) [37] uses a second excitation source, capable of preferentially exciting the acceptor, that is turned on and off in antiphase with the other source. This allows the stoichiometry (i.e. how many of each fluorescent label it carries) of an imaged biomolecule to be determined, hence allowing low FRET states to be distinguished from non-photoactive acceptors. Chapter 5 describes the implementation of ALEX with three separate excitation lasers, chosen to preferentially excite three spectrally distinct channels.

1.4 Total internal reflection fluorescence microscopy

A total internal reflection fluorescence (TIRF) microscope can simultaneously image many ($\sim$100) biomolecules that are tethered to the surface of a coverslip [38]. To attach to the surface, biotin-neutravidin interactions and a solution concentration of labelled biomolecules of 50 pM to 1 nM can be used. In objective-type TIRF (Figure 1.2), the illumination beam emerges collimated from the objective and is incident on the coverslip-medium interface at greater than the critical angle, θ_c:

$$\sin\theta_c = \frac{n_{medium}}{n_{coverslip}} \tag{1.3}$$

where n_{medium} is the refractive index of the medium, typically $\sim$1.33 for aqueous

buffers, and $n_{coverslip}$ is the refractive index of the glass coverslip, typically ∼1.52. The numerical aperture (NA) of the objective must be sufficient to allow this criterion to be met (NA $> n_{medium}$). For these objectives, oil, with a refractive index approximately matching the coverslip, is placed between the objective and the coverslip to reduce reflections from those interfaces. When the conditions of TIRF are met, there is an evanescent excitation field inside the sample with the profile:

$$I \propto \exp\left(-z/d\right) \tag{1.4}$$

where z is the depth into the sample (away from the surface). The TIRF decay constant, d, is given by:

$$d = \frac{\lambda}{4\pi}\left(n_{coverslip}^2 \sin\theta_i - n_{medium}^2\right)^{-0.5} \tag{1.5}$$

where λ is the illumination wavelength, and θ_i is the angle of incidence. Since the biomolecules of interest are attached to the surface of the coverslip, they are in a region of high excitation and they fluoresce brightly. Out-of-focus light, from deep in the sample, caused by other labelled biomolecules in solution, or by scattering from water molecules, is minimised in TIRF, improving the signal to noise ratio of images obtained.

1.4.1 Resolution

When a source is imaged by an optical system, its image is broadened due to the diffraction of light. The image of a point source can be approximated as a two-dimensional Gaussian of standard deviation $0.42\frac{\lambda}{\mathrm{NA}}$. To recover the number of photons emitted by an emitter, such a Gaussian function can be fit to the emission pattern. Since, in superresolution experiments, it is important to recover the underlying position of the emitter, an analytical expression approximating the uncertainty inherent in this fitting method has been derived [39], and shows that resolutions down to ∼10 nm are routinely possible. However, the

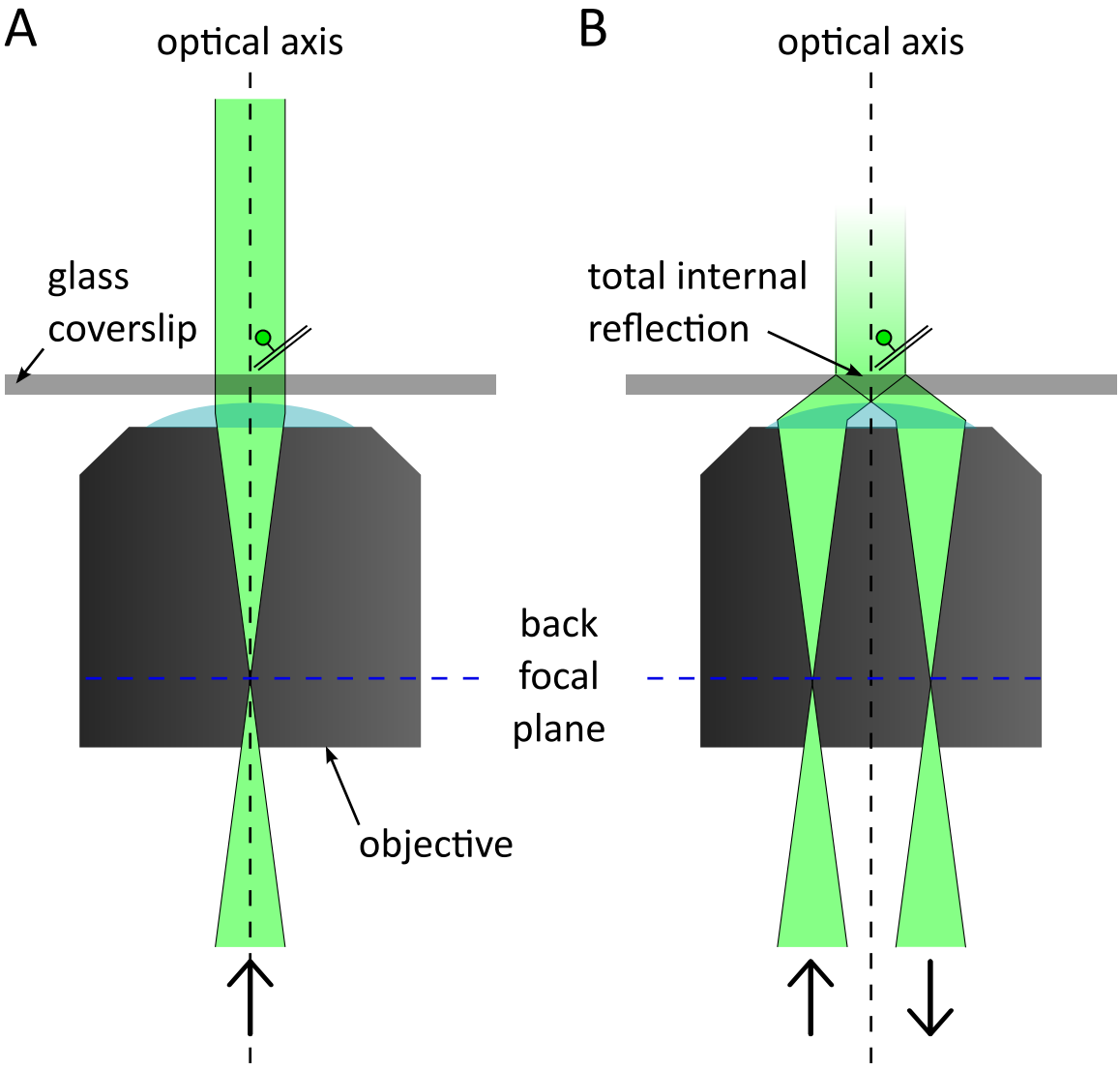

Figure 1.2: Implementing total internal reflection fluorescence using illumination through the objective. **(A)** Epifluorescence. Excitation light, from a laser source, is focussed into the back focal plane of the objective, and emerges collimated. The emitted fluorescence can be collected using the same objective, however, in epifluoresence, the illumination volume is much larger than that necessary to excite the labelled biomolecules tethered to the surface, hence the signal can be hard to distinguish due to background emission (e.g. due to Raman scattering by the buffer). **(B)** By displacing the illumination beam from the optical axis, the excitation beam emerges collimated, but is incident at an angle to the coverslip-medium (i.e. the buffer used for imaging) interface. If the objective NA is sufficiently great to allow the beam to meet the interface at greater than the critical angle, then total internal reflection occurs, and an evanescent field is formed in the sample, decaying away from the interface. Using a typical NA of $\sim$1.4, a decay length as short as $\sim$100 nm can be obtained, reducing the magnitude of background emission, allowing single fluorescent molecules to be easily imaged.

analytical expression relies on weighted least-squares fitting, which (unless subject to ad hoc correction) tends to give arbitrarily high weight to pixels with no observed photons, throwing the estimate off-target [40]. Improving on this, fitting using maximum likelihood estimation approaches the theoretical (Cramér-Rao) lower bound when it utilises an improved point-spread function for the estimation [40]. The improved point-spread (microscope response) function takes into account that the fluorophore imaged is not a point source; it is a dipole emitter near a dielectric interface (the coverslip-medium interface) [41].

A similar analysis has been performed for the uncertainty in the apparent FRET value for a dye pair imaged using TIRF [42], and for the width of the image of a single fluorophore [43]. However, in both cases, weighted least-squares fitting has been assumed. In Chapter 2, I derive an expression for the uncertainty in the image width for ordinary least-squares fitting in the context of TFM, and compare it to experimental data, and simulations based on realistic assumptions about the shape of the microscope response function and the distribution of fluorophore positions coming from a worm-like chain model of DNA.

1.5 Biological motivation

So far, I have discussed the methods of single-molecule biophysics, with a particular emphasis on fluorescence techniques. However, the importance of single-molecule techniques lies in their application to the study of biological systems. In this thesis, I describe the application and development of single-molecule fluorescence techniques to study the XerCD-*dif*-FtsK chromosome segregation machinery.

1.5.1 Site-specific recombination

Site-specific, as oppose to homologous, recombination refers to a process where DNA strand exchange occurs between DNA segments that possess a particular short sequence, or site. Depending on the residue involved in catalysis, most site-specific recombinases are

categorised into two families: tyrosine-family or serine-family [44]. Tyrosine recombinases play key roles in the propagation of viruses and plasmids, and in the horizontal transfer of genetic material [44, 45, 46].

1.5.2 Decatenation and unlinking

Two such tyrosine recombinases are the closely related XerC and XerD, which act in chromosome and plasmid segregation and stability, by resolving chromosome dimers, and by performing the decatenation of linked chromosomes [44, 45, 47, 48]. Linked and dimerised chromosomes are produced during replication or via homologous recombination (Figure 1.3). XerC and XerD bind the *dif*-site located in the terminus region (*ter*) of the *Escherichia coli* chromosome. Recombination by XerCD-*dif* is activated by a DNA translocase, FtsK [49, 50]. The work in Chapter 3 elucidates the mechanism of site-specific recombination by XerCD-*dif*, and compares the structure of the recombining complex to another tyrosine recombinase system Cre-*loxP* [51].

1.5.3 FtsK

FtsK is a DNA translocase with an N-terminal domain that anchors it to the division septum [52, 53, 54], and a C-terminal motor domain [55]. In crystal structures, it forms hexameric rings as part of a double ring head-to-head conformation (Figure 1.4) [56]. In addition to its role activating XerCD-*dif*, it is thought to play a direct role in the segregation of *ter* [57], potentially by removing MatP from *matS*-sites as it translocates, allowing sister *ter* regions to separate. FtsK recognises an 8 bp sequence, FtsK orientating polar sequence (KOPS), which guides its translocation towards the *dif*-site [58, 59, 60]. The activation of XerCD-*dif* recombination requires direct contact between the γ subdomain of the FtsK motor domain and XerD [61, 62]. Chapter 4 discusses the behaviour of FtsK in detail, summarises previous work, and presents experiments focussed on understanding its mechanism of assembly and translocation, and its recognition of XerCD-*dif*.

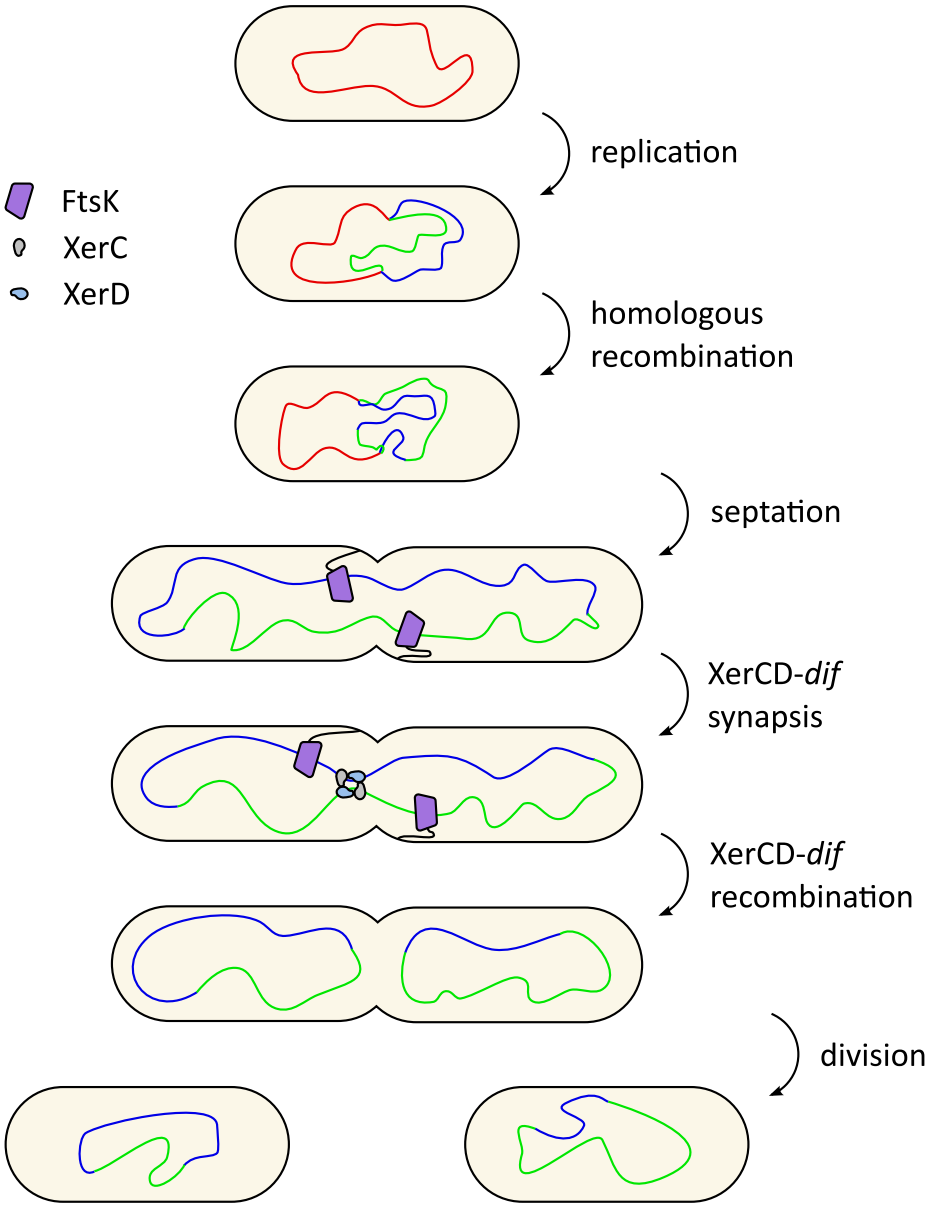

Figure 1.3: XerCD-*dif* recombination. If a chromosome dimer forms, XerCD-*dif* is responsible for its resolution. XerC and XerD bind the *dif* site in the *ter* region. Recombination by XerCD-*dif* is activated by FtsK, a DNA translocase that associates to the division septum.

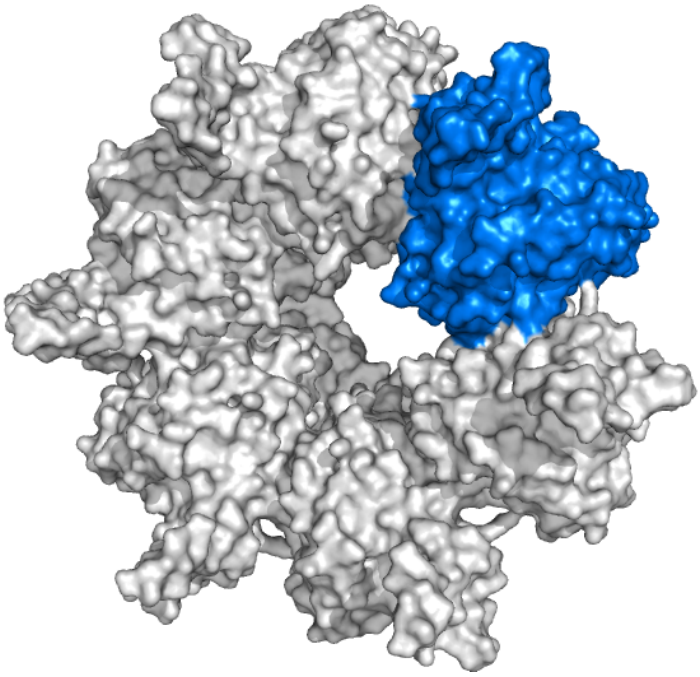

Figure 1.4: Crystal structure of a single hexameric ring of FtsK (2IUU). One monomer is highlighted (blue). The central pore is large enough to accommodate double-stranded DNA.

1.5.4 Assaying recombination

The action of XerCD at *dif* has been shown on circular plasmids containing two *dif*-sites [63]. When the *dif*-sites were placed in direct repeat, productive recombination was apparent as the production of two smaller plasmids that could be distinguished by gel electrophoresis. Knotted plasmids of the same length run at different apparent lengths on agarose gels. This has been used to demonstrate the ability of XerCD-*dif*-FtsK to remove catenation, knotting and dimerisation by recombination [64]. 'Suicide substrates' with nicks on either the top or bottom strand in the central region of the *dif*-site have been used to demonstrate that XerC is responsible for cleavage of the top strand and XerD responsible for cleavage of the bottom strand [62]. These substrates have a double strand break introduced if the non-nicked strand is cut, and hence depending on which strand exchange occurs, the product differs in size and can be distinguished on a gel.

Chapter 2

Defining the limits of tethered fluorophore motion

2.1 Introduction

Tethered fluorophore motion (TFM) has been mentioned briefly in the previous chapter. It uses the same principles as tethered particle motion (TPM) to monitor an effective length along a biopolymer, but employs a single fluorophore in place of the bead, allowing TFM to be combined with existing fluorescence techniques. In this chapter, I describe the characterisation of TFM, focussing on the theoretical basis and applications of the technique. Since TFM is limited in observation time and photon count by photobleaching, I build on previous work [39, 40, 43] and derive an expression describing the resolution of TFM, and, comparing this with Monte Carlo simulations and experimental data, I show that length changes of 100 bp of double-stranded DNA (dsDNA) are distinguishable using TFM, making it comparable to TPM. I also show that the commonly recommended pixel size for single-molecule fluorescence approximately optimises signal to noise in TFM, enabling facile combination of TFM and other fluorescence techniques. Finally, I describe the application of TFM to two biological systems: polymerisation by the Klenow fragment

(KF) of DNA polymerase I [65], and synapsis formation in Cre-*loxP* [66]. In the latter case, TFM was combined with FRET using only a single excitation laser, demonstrating the ease-of-use of this multi-observable combination fluorescence technique.

2.1.1 Principle of TFM

When a single immobile emitter is imaged onto a camera, the image produced is much larger than the emitting molecule because of the diffraction of light through the optical system. This image is dependent on the emission spectrum of the fluorophore, and, because fluorophores are dipole emitters, its orientation. Any aberrations present in the imaging system will also distort the image. In a typical single-molecule total internal reflection fluorescence experiment, fluorophores are attached to biomolecules, themselves attached to a slide surface using biotin-neutravidin (or similar) interactions. These fluorophores, despite their ability to rotate locally around their attachment point, are well approximated by immobile emitters, and hence the width of their image corresponds to the diffraction limit. However, if the fluorophore is not immobilised, and is allowed to move (on scales $> 10\,\text{nm}$) within a single camera frame, the width of this fluorophore image, which we term the FIW, is broader.

For DNA strands with end-to-end distances that are short compared to their diffraction limited image widths (e.g. dsDNA $< 100\,\text{bp}$ has a contour length $< 35\,\text{nm}$, compared to a diffraction limited image width of $\sim$150 nm on a fluorescence microscope), we observe FIWs close to the diffraction limit (Figure 2.1A, left). For longer DNAs, with end-to-end distances longer than the FIW of a single immobile emitter (Figure 2.1A, right), diffusion of the free end and attached fluorophore, about the surface tether point, significantly increases the FIW (Figure 2.1B and C).

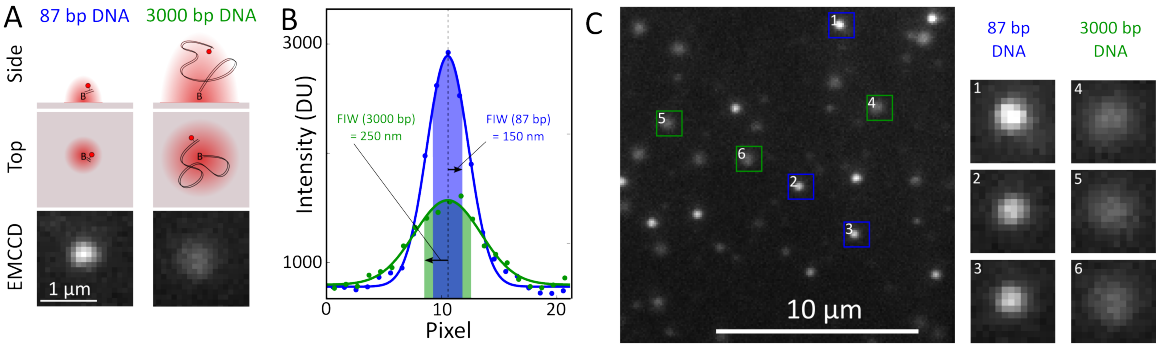

Figure 2.1: Principle of TFM. **(A)** Cartoon and images of single emitters. A Cy5 fluorophore, shown as a red circle was attached to DNA, and tethered to a coverslip using a biotin-neutravidin interaction. An 87 bp DNA results in a diffraction limited image (red in the cartoons); but a 3000 bp DNA allows the fluorophore to diffuse about the tether point during a frame, causing the image to appear broader. The bottom panels show images from an electron-multiplying charge-coupled device (EMCCD) camera. **(B)** Gaussian fits to the pixel intensity values, in digital camera units (DU), from cross sections through the EMCCD data demonstrate the difference in FIW (highlighted region and arrow). **(C)** A section of our field of view: 87 bp DNAs appear as bright, narrow spots, and 3000 bp DNA appear as dimmer, broader spots. Example regions of interest corresponding to 87 bp and 3000 bp DNA are shown in the right panels.

2.2 Methods

In this section, I describe the techniques used for DNA and protein preparation, single-molecule experiments, data analysis, Monte Carlo simulations. I also present the derivation of an expression for the noise due to the stochastic nature of photon detection and the fitting procedure used to extract the FIW from images of fluorophores.

2.2.1 DNA and Protein Preparation

DNA substrates were produced by polymerase chain reaction (PCR), using Phusion High-Fidelity polymerase (New England Biolabs, Ipswich, USA) and a plasmid template. One oligonucleotide primer was fluorophore labelled, and the other was biotin (and, for the Cre substrate, fluorophore) labelled (Figure 2.2). Oligonucleotides were synthesised, labelled and high performance liquid chromatography purified by ATDBio (Southampton, UK). Gel-purification was used to extract the product after PCR.

We chose to use Cyanine dyes because of their brightness and photostability [67].

Cy5 was used for singly-labelled constructs because the red spectral channel is relatively background-free (from auto-fluorescence from optics or dirt).

The KF DNA substrate (Figure 2.2B) was produced using a PCR with a 20-fold excess of labelled strand, which produced an excess of single-stranded DNA (ssDNA). Gel purification was performed after the binding of streptavidin (1 mg/ml) to retard dsDNA, and allow the collection of ssDNA, which was unaffected. To produce the final substrate, biotinylated oligonucleotides were annealed (Figure 2.3). KF was expressed and purified using standard methods [68].

The Cre DNA substrate was produced using PCR, and was 1000 bp long with two *loxP*-sites in an antiparallel orientation. The surface distal *loxP*-site was labelled with Cy3B, and the surface proximal site was labelled with Cy5 (Figure 2.2C). These fluorophores were chosen because of their large spectral separation, which gave a low cross-talk between emission channels, and because they have been previously shown to give FRET at these labelling positions during synapsis formation [32]. The mutant of Cre, A36V, previously shown to give stable synapsis without complete recombination, was expressed and purified using standard techniques [32].

2.2.2 Instrumentation

Single-molecule total internal reflection fluorescence (TIRF) experiments were performed on a custom-built objective-type TIRF microscope. A green (532 nm Cobolt Samba, Cobalt, Solna, Sweden) and red (635 nm Cube, Coherent, Santa Clara, USA) laser were combined using a dichroic mirror and coupled into a fibre optic cable. The output of the fibre was focussed into the back focal plane of the objective (100× oil-immersion, NA 1.4, $f/26.5$, UPlanSApo, Olympus, Tokyo, Japan) and displaced perpendicular to the optical axis such that laser light was incident at the slide-solution interface at greater than the critical angle, creating an evanescent excitation field. Excitation powers were set to 1 mW for both lasers, which corresponds to an approximate power density of 0.4 $\mu W/\mu m^2$ incident

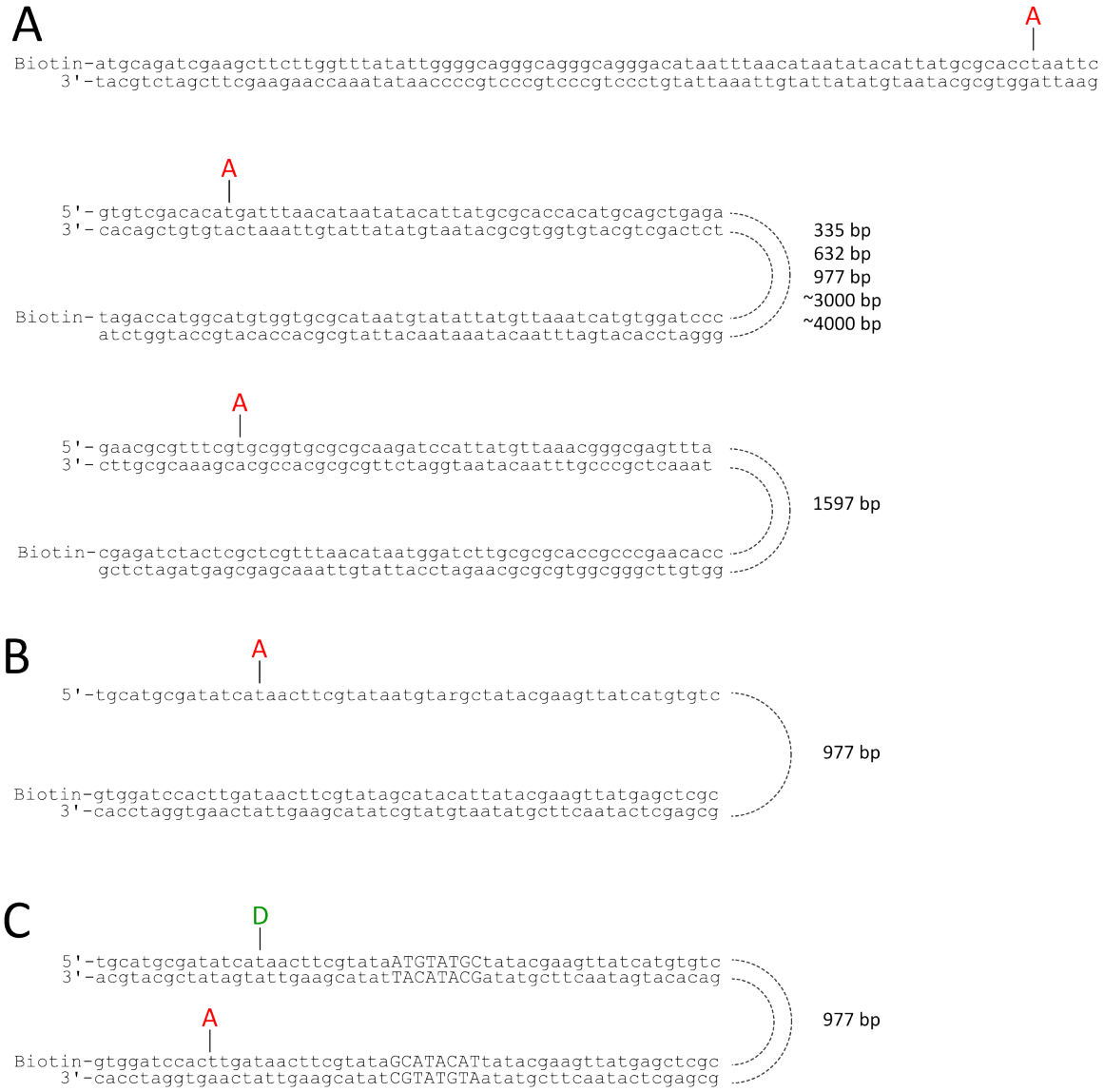

Figure 2.2: DNA primer sequences used. Cy3B and Cy5 are indicated as D and A respectively. Intervening DNA indicated with dashed lines. **(A)** Singly labelled DNA with lengths from 87 bp to ∼4 kb. **(B)** DNA used as a template for KF polymerisation. **(C)** Doubly labelled DNA used in Cre experiments. Captial letters indicate central region of *loxP*-site.

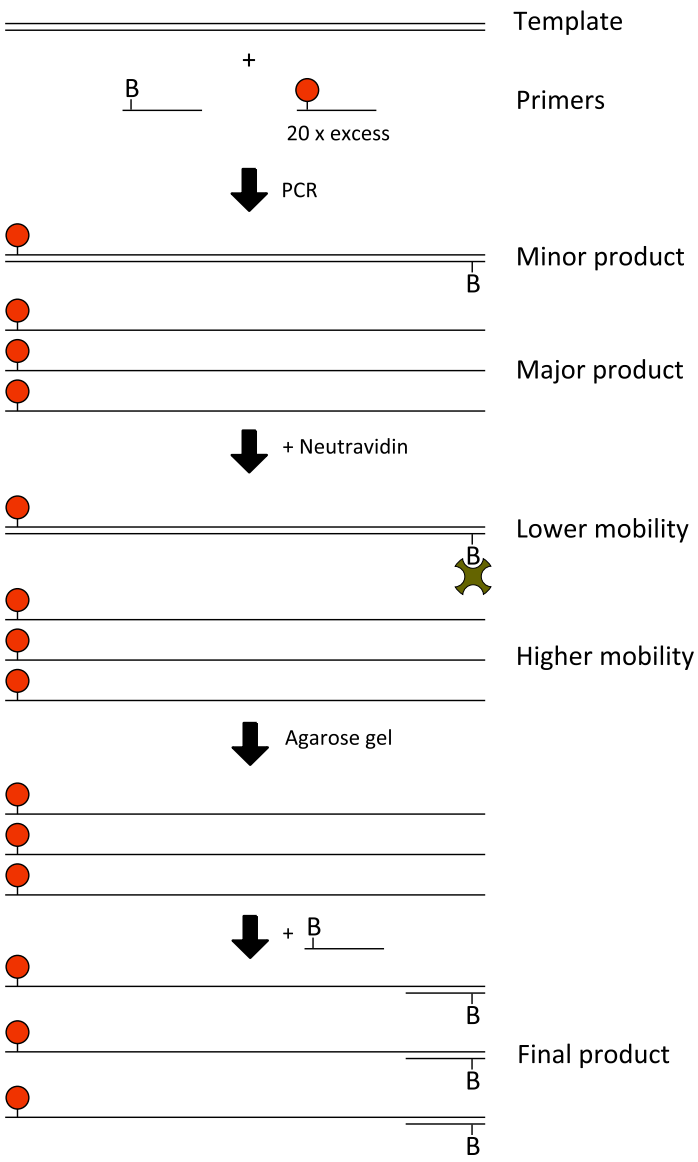

Figure 2.3: Production of KF DNA substrate. A PCR with a 20-fold excess of labelled primer produced mostly labelled single stranded DNA (ssDNA), along with biotinylated double stranded DNA (dsDNA). Addition of neutravidin before gel purification ensured good separation between ssDNA and dsDNA of the same backbone length. The single stranded, labelled DNA was purified from the gel fragment and annealed to an excess of the biotinylated primer. The partially double stranded product was purified on an affinity-binding column.

on the 50 µm × 50 µm field of view. Fluorescence emission was collected by the objective and separated from the excitation light by a dichroic (545 nm/650 nm, Semrock Rochester, USA) and cleanup filters (545 nm long-pass, Chroma, Bellows Falls, USA, and 633/25 nm notch filter, Semrock). The emission signal was focussed onto a rectangular slit to crop the image, and spectrally separated using a dichroic (630 nm long-pass, Omega, Brattleboro, USA) into two emission channels, which were focussed side-by-side onto an electron multiplying charge-coupled device camera (EMCCD, iXon 897, Andor, Belfast, UK). The EMCCD camera was set to an EM gain of 300, corresponding to an approximate real gain of 3.75 counts per photon. Each pixel on the camera corresponded to a 96 nm × 96 nm in the sample plane. A CRIFF (ASI, Eugene, USA) autofocus system was used throughout this work to ensure focus stability over the course of data acquisition.

2.2.3 Data analysis

TwoTone software was used to fit elliptical Gaussian models to fluorescence images [42]. The fluorophore image width (FIW) was obtained from the mean width of the fitted Gaussian, and the FRET efficiency, E^*, was calculated from the fluorescence emission from the acceptor during donor excitation, DA, and the fluorescence emission from the donor, DD, as follows:

$$E^* = \frac{DA}{DA + DD} \tag{2.1}$$

2.2.4 DNA simulations

We performed worm-like chain (WLC) simulations according to a previously implemented scheme [69], which has been applied to TPM by others [14, 20]. The DNA chain was discretised into rigid segments of length, $l = 3.4$ nm. We assumed that twisting energy and the long-range interactions of a chain with itself could be ignored. Hence, the energy cost of a particular orientation of a segment relative to its neighbour can be described by the

bending energy, parametrised by the persistence length, P. We simulated conformations of the DNA chain by choosing the angle between sequential segments, θ from the distribution:

$$p(\theta) = C \exp\left(\frac{-P\theta^2}{2l}\right) \sin(\theta) \tag{2.2}$$

and choosing ϕ, the azimuthal angle, from the uniform distribution (0 to 2π). C is a normalisation constant. To validate our implementation of this model, we generated many unconstrained chain conformations and observed good agreement between the mean end-to-end distance for different chain lengths and the known analytical expression for the WLC.

To simulate tethered DNA, we consider the first link to be attached to the surface at $x = y = z = 0$ and free to point in any direction $z > 0$, and discard any chain that crosses the plane $z = 0$ (Figure 2.4A). This scheme is identical to one producing many free chains and discarding any that cross the surface at any point. We computed many ($> 10^6$) fluorophore positions from independent WLC DNA simulations to adequately sample the accessible volume (Figure 2.4B), and used this distribution of fluorophore positions to calculate the expected FIW, as described in the next section. To find the appropriate persistence length, P, to use in Equation (2.2), we simulated FIWs for a range of persistence lengths, and computed the mean square deviation from experimentally determined FIWs for chain lengths from 87 bp to 4000 bp, and chose the best fit, $P = 48$ nm (Figure 2.4C). For simulations of DNA as it was being polymerised, we used the fitted persistence length for dsDNA and a literature value of 2 nm for ssDNA [70] (Figure 2.5). For the single-stranded segment, because of the reduced persistence length, we reduced the segment length for the discretisation to 0.34 nm.

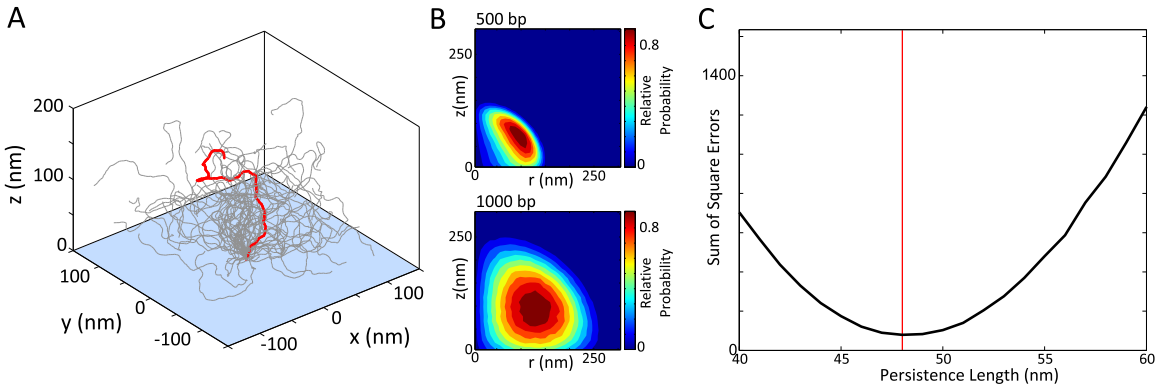

Figure 2.4: Worm-like chain simulations. **(A)** DNA conformations generated using WLC simulations, with a backbone length of 1000 bp. Red: a single conformation of chain. **(B)** Probability of finding a fluorophore at a position relative to the tether point ($z = r = 0$). Cylindrical symmetry has been used to show the distributions in two dimensions. **(C)** Matching the persistence length in the WLC simulations to experimental data. An optimal fit occurs at a persistence length 48 nm. All other simulations in this work use this value for dsDNA.

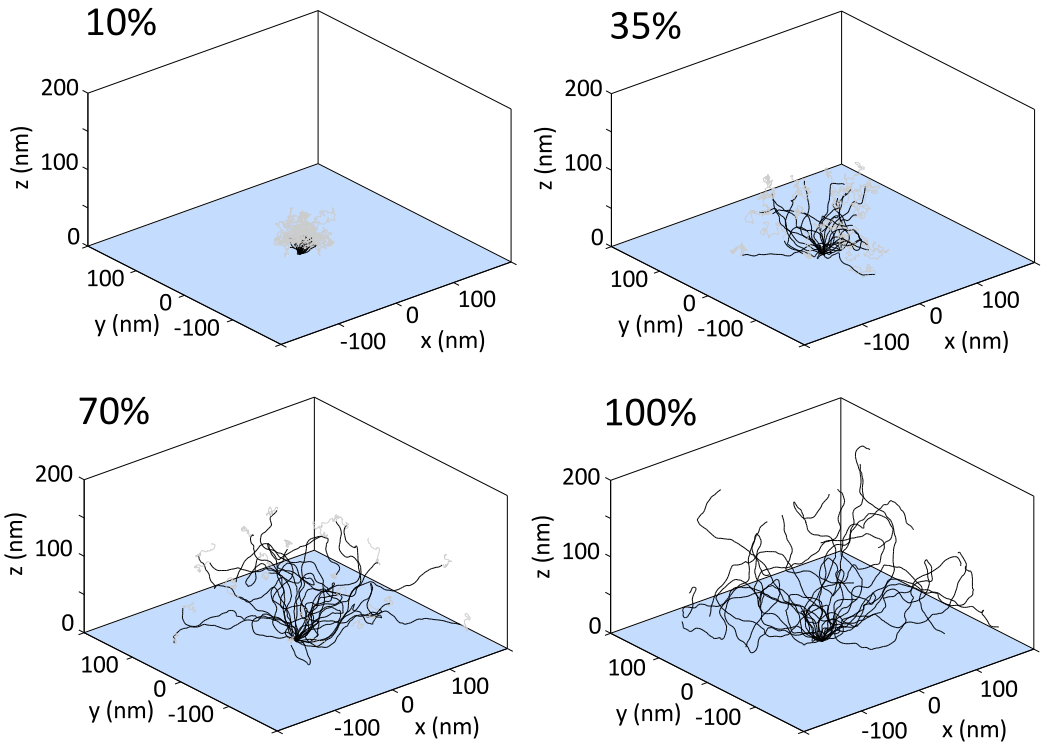

Figure 2.5: Simulations of 1087 bp DNA with increasing double stranded fractions. Single stranded DNA is shown in grey, double stranded in black. As the DNA is polymerised the FIW increases.

2.2.5 Camera simulations

The probability distributions described in the previous section were used in Monte Carlo simulations of EMCCD images. These simulations included the effects of finite pixel size, shot noise, electron multiplying gain and background noise. The response function of the microscope was generated using a Python[1] method, written by Kim Mortensen [40]. We assumed left-hand circularly polarised light incident at the maximum angle of incidence allowed by our objective NA (1.4) This response function was numerically convolved with the fluorophore position probability distributions, weighted by the evanescent TIRF illumination profile, producing a theoretical image. This scheme assumes that the DNA can explore all available conformation during the camera integration time. The image of a subdiffraction fluorescent bead was chosen as our theoretical response function, because the fluorescence anisotropy of our substrates, measured using a fluorimeter (Photon Technologies International, Edison, USA), was near the limiting value of 0.4 (e.g. the 1000 bp substrate had an anisotropy of 0.35). This implies that the fluorophore is roughly stationary during its emission lifetime, but free to be in any orientation over the course of a 100 ms to 500 ms camera frame. In reality the emission pattern from the dipole changes with depth into the medium [71], and the image will be defocussed depending on the fluorophore location.

We found that the width of the image of a single fluorophore attached to a short DNA fragment (found using the same scheme as TwoTone) was 150 nm, rather than the calculated 126 nm. We attribute this broadening to the effect of the biotin-neutravidin and poly-ethylene glycol linker, and to aberrations in the optical system. It was accounted for by convolving the theoretical image with a Gaussian kernel, chosen to match the experimental image width. The poly-ethylene glycol used for surface passivation and for biotin-neutravidin attachment had a molecular weight of around 5 kDa. Given a persistence length of 0.37 nm [72] and an approximate 0.3 nm of backbone length per repeating unit

[1] python.org

in the poly-ethylene glycol structure, the root mean square end-to-end distance for the poly-ethylene glycol used was 5 nm (from the WLC model). Depending on the density of surface attachment the poly-ethylene glycol on the surface may form a polymer brush. The presence of the poly-ethylene glycol will prevent the DNA from exploring conformations close to the surface, reducing the excitation intensity experienced. However, past the poly-ethylene glycol layer, the intensity distribution will remain exponential in form, and the effect of the reduced intensity will be apparent as a reduced photon count, but an unchanged emission distribution shape. Given an approximate depth of 5 nm for the poly-ethylene glycol layer, and a most likely position into the medium of around 100 nm for a 1000 bp DNA (Figure 2.4B) we note that the effect of the poly-ethylene glycol layer on the image shape should be minimal.

The camera properties (gain and read-noise) were determined experimentally using a previously described scheme [73]. Briefly, we took short videos of fields of view under several different uniform low illumination powers, and the digital camera count distributions were fit simultaneously with an expression parametrised by the mean photon count per pixel in each video, the camera gain, and the read-noise. We performed the same procedure, including an empty region in a field of view at a given illumination intensity to determine the background photon count. Simulated images ($> 10^4$ per plotted point, using > 100 different tether locations within a pixel), produced in MATLAB (The MathWorks, Natick, USA), were fit with two-dimensional Gaussian distributions using the same ordinary least-squares scheme as TwoTone [42].

2.2.6 Noise theory

Since TFM is limited by the photon budget of the fluorophore(s) used, we derived an expression for the uncertainty in the width of a fitted Gaussian, as a function of the number of photons captured by the camera. Previous work focussed on weighted least squares fitting, and underestimated the noise by 57 % [43]. In this work, we choose to

focus on ordinary least-squares (OLS) fitting, since it is robust [40, 42], fast to implement, computationally efficient and commonly utilised (e.g. in TwoTone, the software used for image analysis in this work [42]). We assume the image of the fluorophore tethered to DNA is well approximated by a two-dimensional Gaussian, with width characterised by the standard deviations, s_x and s_y. We perform fitting by minimising:

$$\chi^2 = \sum_{i}^{\text{all pixels}} (n_i - N_i)^2 \tag{2.3}$$

with respect to the parameters being fit, where n_i is the observed photon count, and N_i is the expected photon count at the same pixel for the fit. Performing this minimisation on Equation (2.3) with respect to s_x, we obtain:

$$2 \sum_{i}^{\text{all pixels}} (n_i - N_i)\, N_i' = 0 \tag{2.4}$$

where $N_i' = \frac{dN_i}{ds_x}$. Expanding N_i to first order in $\Delta s_x = s_x - s_{x,0}$, the deviation from the 'true' standard deviation ($s_{x,0}$) of the image (i.e. the uncertainty in the fitted width), and defining $\Delta n_i = n_i - N_i\,(s_{x,0})$, we get:

$$\Delta s_x = \frac{\sum_{i}^{\text{all pixels}} \Delta n_i N_i'}{\sum_{i}^{\text{all pixels}} (N_i')^2} \tag{2.5}$$

Squaring Equation (2.5), and averaging over many camera frames:

$$\left\langle \Delta s_x^2 \right\rangle = \frac{\sum_{i,j}^{\text{all pixels}} N_i' N_j' \left\langle \Delta n_i \Delta n_j \right\rangle}{\left(\sum_{i}^{\text{all pixels}} (N_i')^2 \right)^2} \tag{2.6}$$

where $\langle \ldots \rangle$ denotes the expectation value over many camera frames. The noise is uncorrelated between pixels; therefore, we can write $\langle \Delta n_i \Delta n_j \rangle = \delta_{i,j} \sigma_i^2$, where σ_i^2 is the photon count variance due to all sources of noise (shot-noise, background, read-noise and

EMCCD gain), giving:

$$\left\langle \Delta s_x^2 \right\rangle = \frac{\sum_i^{\text{all pixels}} \left(N_i'\right)^2 \sigma_i^2}{\left(\sum_i^{\text{all pixels}} \left(N_i'\right)^2\right)^2} \tag{2.7}$$

with

$$\sigma_i^2 = 2\left(N_i + \langle b \rangle\right) + \frac{\sigma_r^{*2}}{M^2} \tag{2.8}$$

where $\langle b \rangle$ is the expected background photon count, σ_r^* is the standard deviation of the read-noise (in digital units) and M is the electron multiplying gain factor (i.e. how many digital units are recorded on average per photon). The factor of 2 comes from the electron multiplication process, and is approximately the square of the excess noise factor [42]. We can define $\sigma_b^2 = 2\langle b \rangle + \frac{\sigma_r^{*2}}{M^2}$, which groups terms by their dependence on N_i. Equation (2.7) can be solved numerically, but to proceed analytically, we approximate the sums as integrals. However, unlike with weighted least-squares fitting [43], we do not need to interpolate between the high photon count and high background limits. Since we are assuming the image is Gaussian, we use the photon distribution:

$$p\left(x, y\right) = \frac{N}{2\pi s_x s_y} \exp\left(-\frac{x^2}{2s_x^2} - \frac{y^2}{2s_y^2}\right) \tag{2.9}$$

N is the total number of photons detected on average (for this molecule) per frame. The effect of evanescent illumination is included, since this is modelled in the photon count: a longer DNA, or one imaged with a different TIRF angle will have a different photoemission rate. We approximate σ_i^2 as:

$$\sigma_i^2 = 2p\left(x, y\right) + \frac{\sigma_b^2}{a^2} \tag{2.10}$$

where the pixel size, a, has been included to match the contribution to the integral,

when integrating over a single pixel, to the appropriate term in the sum in Equation (2.7). Hence, we get:

$$\left\langle \Delta s_x^2 \right\rangle = \frac{256 s_{x,0}^2}{81N} + \frac{16\pi s_{x,0}^3 s_{y,0} \sigma_b^2}{3a^2 N^2} \tag{2.11}$$

where we have integrated over all space in x, y and evaluated the expression at $s_x = s_{x,0}$ and $s_y = s_{y,0}$ (i.e. we have assumed the fit is unbiased). Equation (2.11) can be modified to include the effect of finite pixel size, by replacing $s^2 \to s^2 + \frac{a^2}{12}$. This accounts for the increased uncertainty in the detection position of a photon due to pixelation [39, 43]. Substituting this in, we get the analytical expression for noise in TFM:

$$\left\langle \Delta s_x^2 \right\rangle = \frac{256 \left(s_{x,0}^2 + \frac{a^2}{12} \right)}{81N} + \frac{16\pi \left(s_{x,0}^2 + \frac{a^2}{12} \right)^{\frac{3}{2}} \left(s_{y,0}^2 + \frac{a^2}{12} \right)^{\frac{1}{2}} \sigma_b^2}{3a^2 N^2} \tag{2.12}$$

2.2.7 Single-molecule experiments

DNA was immobilised to the surface of a coverslip using biotin-neutravidin interactions. The coverslip was passivated using poly-ethylene glycol, and sealed using a silicone gasket (Grace Bio-labs, Bend, USA) and a second coverslip as a lid. A buffer consisting of 50 mM Tris-HCl (pH 7.5), 50 mM NaCl, 5 mM $MgCl_2$, 100 µg/ml BSA, and 1 mM UV-treated Trolox, was used for imaging. An enzymatic oxygen scavenging system consisting of 1 mg/ml glucose oxidase, 40 µg/ml catalase and 1.4 % (w/v) glucose was added prior to sealing the sample before image acquisition. Experiments were performed at a temperature of 21 °C.

2.2.8 Klenow fragment experiments

Single-molecule experiments were performed as described in the previous section, in a different buffer (50 mM Tris pH 7.5 , 10 mM $MgCl_2$, 1 mM DTT, 10 % glycerol, 100 µg/ml BSA, and 1 mM UV-treated Trolox). KF (24 nM) and nucleotides (10 µM each) were added

during data acquisition and the wells were not sealed. Data was collected at 2 Hz, until the majority of fluorophores had photobleached or no further polymerisation could be observed. To convert from FIW to number of bases polymerised, we performed WLC simulations of 1087 bp DNA with varying fractions of dsDNA (Section 2.2.4).

2.2.9 Cre experiments

The 532 nm laser was used to excite the FRET donor fluorophore, Cy3B. The FIW from the same fluorophore also reported on the effective length of the DNA. Cre was added to immobilised DNA at a final concentration of 20 nM, and movies were captured at 10 Hz with continuous wave excitation. Hidden Markov modelling (HMM) was used to automate the extraction of dwell times as described previously [74]. Dwells corresponding to synaptic complexes were used to extract the E^* histogram. Data immediately before or after a transition were discarded to minimize camera integration time effects [75].

2.3 Results and discussion

2.3.1 FIW and mean excursion

Depending on the relative timescales of DNA motion and camera integration, two regimes are apparent in TFM experiments. If the DNA does not move appreciably during a camera frame, a mean excursion can be computed from the individual snapshots of the fluorophore positions (Figure 2.6). On the other hand, if the integration time is long, compared to the DNA relaxation time, the image will be blurred by the motion of the DNA during a single camera frame and the FIW will be increased (Figure 2.6). Since the FIW does not increase significantly with camera frame times above 100 ms (Figure 2.6C), we note that our approximation that the DNA is free to sample the whole set of allowed conformational states during each camera integration is reasonable for DNAs at the length used (3000 bp) at frame times longer than 100 ms. EMCCD cameras offer limited fields of view at high

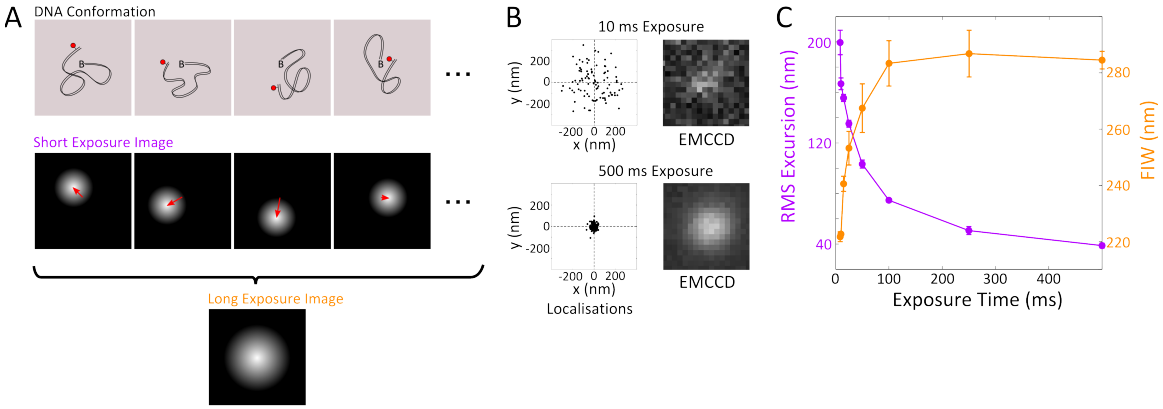

Figure 2.6: Observables in TFM **(A)** Cartoon of two different imaging schemes. Using short exposures, we see a diffraction limited FIW and the DNA length is apparent as a mean excursion (red arrows). Using long exposures, the averaging is done during the camera acquisition and we see an increased FIW, fixed in the image plane. **(B)** Single-molecule localisations from sequential frames of a 100 frame movie of 3000 bp DNA, taken with 10 ms or 500 ms exposure times. Localisations (i.e. the position of the peak of the fitted Gaussian) are shown in black, and the mean of all localisations is indicated by the dashed lines. Representative images at the two exposure times are shown. The brightness of the images has been adjusted since the 500 ms exposure gives a much brighter image. **(C)** FIW and root mean square (RMS) excursion for 3000 bp DNAs at frame times from 8.5 ms to 500 ms. There is a decrease in FIW and increase in mean excursion as the frame time becomes comparable to the DNA relaxation time. Error bars are the standard error in mean from 5 molecules at each exposure time. Localisation error is less than 10 nm in all cases.

($>$ 50 Hz) acquisition rates, in this work, we focus on the use of FIW, rather than mean excursion, to report on effective length.

2.3.2 Differentiating DNA lengths

Due to the evanescent excitation profile of TIRF illumination, fluorophores attached to long DNAs experience a reduced average excitation intensity, and hence show reduced intensities, compared to ones tethered tightly to the surface. We imaged dsDNA with lengths of 87 bp and 1000 bp (Figure 2.7), observing many molecules in each field of view (Figure 2.7A and B, left). The molecules can be distinguished by both their FIW (Figure 2.7C) and their intensity (Figure 2.7D).

Although we observe differences in intensity between DNAs of different lengths, this

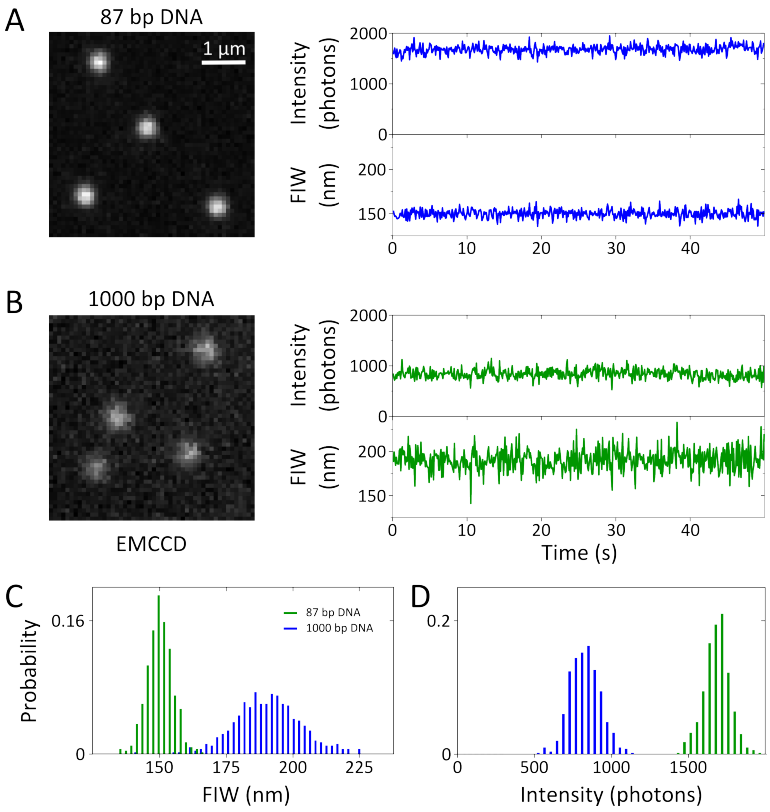

Figure 2.7: TFM timetraces **(A)** Example camera image of 87 bp DNA and timetrace showing relatively high intensity, and narrow FIW. Data taken with 1 mW illumination intensity and 10 Hz frame rate. **(B)** Example camera image of 1000 bp DNA and timetrace showing lower intensity, and increased FIW. **(C)** Histogram of FIW showing the two populations are resolvable using FIW **(D)** Histogram of intensities. The 87 bp and 1000 bp DNA are also distinguishable by their intensities (under these conditions).

intensity difference is sensitive to many factors, including the illumination intensity pattern, the underlying intensity of each fluorophore, and the TIRF angle. Hence, in this work, we focus on the FIW, which is not as sensitive to these confounding factors. Changes in intensity within a single molecule will not be subject to all of these effects, but to determine a change in effective length from a change in intensity would require a TIRF angle dependent calibration, which would potentially need to be generated for every experimental session.

We generated a calibration between FIW and DNA tether length by imaging DNAs with lengths 87 bp to 4000 bp and measuring the mean FIW at each length. The standard deviation from three independent measurements (of ∼10 molecules in separate fields of view) at each DNA length was less than 10 nm (Figure 2.8A). To compare the observed FIWs to a physical model of DNA, we performed discrete WLC simulations of tethered DNA [14, 20]. As outlined in Sections 2.2.4 and 2.2.5, we generated static distributions of fluorophore position and related these to FIWs. Fitting the persistence length used in

the simulations to match the FIW for the DNAs imaged, we obtain a best fit $P = 48\,\text{nm}$ (Figure 2.8A). This is in agreement with experimentally determined values measured under similar conditions (Table VIII, $P = 35\,\text{nm}$ to $55\,\text{nm}$ in [76]), which gives us confidence in our ability to predict TFM observables. This agreement supports the use of a simple microscope response function that does not vary with fluorophore position in the camera simulations (Section 2.2.5).

2.3.3 Noise in TFM

To quantify the precision of TFM measurements and validate our understanding of the sources of noise in TFM, we measured the standard deviation of the FIW from many molecules over a range of excitation intensities for DNAs with lengths of 87 bp and 3000 bp (Figure 2.8). These experimentally determined standard deviations were compared with predictions from analytical theory (Equation (2.12)) and from camera simulations (Section 2.2.5). The analytical expression derived here differs from the weighted least-squares expression [43] in the first term, and in that it is not an interpolation between the high-background and photon-shot-noise dominated regimes. The experimental data was spread along the vertical (standard deviation) axis because of variations in FIW, due to defocussing across the FOV, and variations in the background photon count for each molecule.

For 87 bp DNA, our results show agreement between the experimental data, the predictions made using the analytical expression, and the simulations. Over the range of photon counts (illumination intensities 1 mW to 50 mW), the uncertainty was roughly a power law of the intensity, indicated by straight lines on a logarithmic plot (Figure 2.8B, dashed lines). The power law fit has an exponent of -0.47, consistent with a shot-noise dominated regime. Standard deviations as low as 5 nm, corresponding to the ability to resolve a change from 87 bp to 237 bp in a single frame (Figure 2.8A), can be obtained with photon counts of several thousand per frame, which are readily achievable experimentally

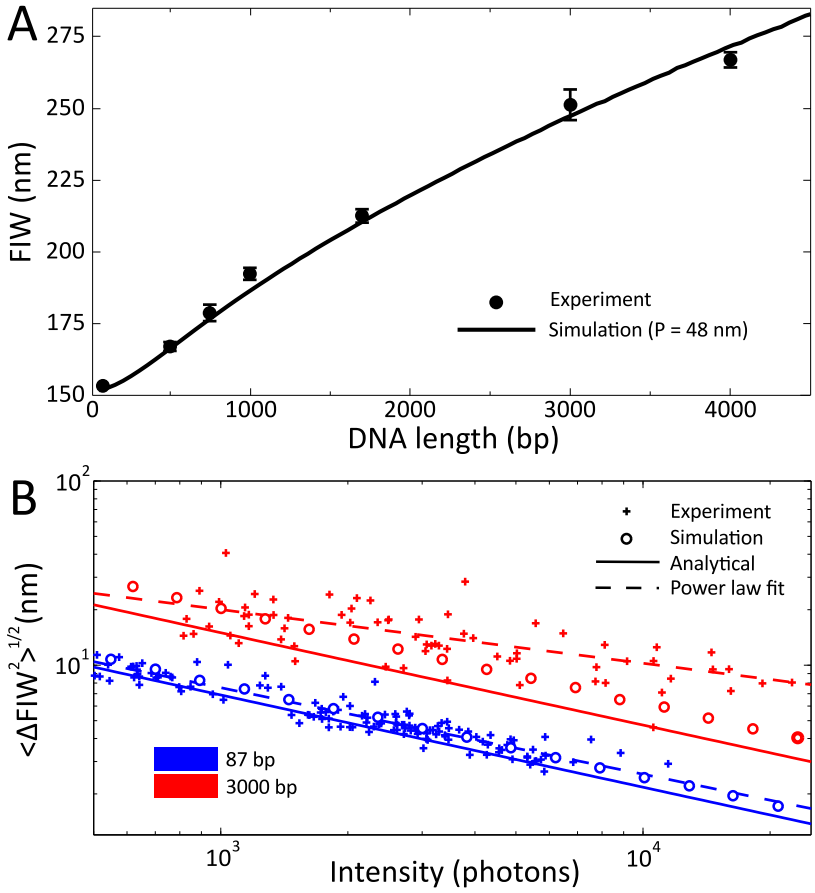

Figure 2.8: Physical models of TFM **(A)** Experimental measurements of the FIW for double stranded DNAs 87 bp to 4000 bp in backbone length. Data were acquired at 1 Hz, and error bars indicate the standard deviation of three independent measurements for each sample. The solid line is the best fit of persistence length to experimental data ($P = 48$ nm). We observe transient sticking of the DNA end to the surface in less than 1 % of traces. **(B)** The resolution of TFM. Three methods were used to quantify the uncertainty in FIW for DNAs of 87 bp and 3000 bp. The experimentally determined uncertainties for individual molecules, from 100 frame movies recorded at 2 Hz with illumination powers in the range 1 to 50 mW, are shown as crosses. The intensity plotted is the photon count for a single molecule. Dashed lines are power-law fits as guides for the eye (exponents: -0.47 for the 87 bp DNA and -0.30 for the 3000 bp DNA). Equation (2.12) was used to plot the analytical approximation (solid lines). A simulation, using the parameters of our microscope and the theoretical images of the DNA, generated as described in Materials and Methods, was used to produce the data shown as circles. We observe a good agreement between experiment, simulations, and our analytical expression, with the approximations made in deriving the analytical expression tending to underestimate the noise. For the 3000 bp DNA, deviations can be observed at intensities $>$ 3000 photons, which we attribute to incomplete sampling of the conformations available to the DNA during a single camera frame. The background count in experiments was found to vary linearly with the observed photon count, with 5.3×10^{-3} background photons per recovered photon for the 87 bp DNA and 9.4×10^{-3} for the 3000 bp DNA; this was used in our simulations and in the plotting of the analytical approximation.

[67].

(For 87 bp DNA) closer inspection of the data shows that the analytical expression underestimates the noise by $\sim$10 %, in part due to the approximations made in its derivation (Section 2.2.6). The simulation is in excellent agreement with the data, showing that the better approximation to the microscope response function, and a better model of noise per pixel, can account for the deviation between theory and data.

For the 3000 bp DNA, the FIW was broader, and hence at a given photon count, the uncertainty in FIW is larger. Here, the analytical expression underestimates the noise by $\sim$40 %, since most of the approximations made in its derivation become less valid. However, below a photon count of 4000 photons/frame, the simulation agrees well with the data, suggesting that the simulation describes the behaviour of the system well. Above 4000 photons/frame, the simulation underestimates the uncertainty, which we speculatively attribute to DNA dynamics on timescales similar to the exposure time, leading to incomplete sampling of DNA conformational space within the 500 ms frames. The underestimation is only apparent at higher photon counts, because it is only when the fitting noise (well-modelled by the simulation) diminishes, that the noise due to incomplete sampling dominates. Our simulations and analytical model take the decay of the TIRF field into account because the models are parametrised by the observed photon count, and, to generate each point for the simulation or from the analytical expression, we input the observed photon count and the experimentally determined background count.

Having discussed the effect of changing illumination intensity on the resolution of TFM, we now move to discuss more practical matters: the dependence of the uncertainty in the FIW on the microscope design. For single-molecule tracking, there is a well-known optimum pixel size close to the FIW [39]. Using binning on our EMCCD, we experimentally varied our pixel size in the range 96 nm to 384 nm. The camera parameters, σ_b, M, and $\langle b \rangle$, were determined separately for each pixel size. The background photon count, $\langle b \rangle$, was found to vary linearly with the pixel area (Figure 2.9A), and σ_b and M remained

approximately constant. This variation in the background count with pixel size means we cannot simply minimise Equation (2.12) with respect to the pixel size, but we can evaluate it numerically. The same is likely true for previous work [39], and has the effect of slightly reducing the optimum pixel size.

To match the observed photon counts in simulations, we had to input an excess of photons in our simulations. This underestimation was likely due to the OLS Gaussian fit, and we saw that at shorter DNA chain lengths this underestimation was more significant (Figure 2.9B) [42]. Between this and the effect of the evanescent illumination field, we note that intensity and FIW are correlated in a non-trivial manner.

Numerically, the theoretical expression predicts an optimum pixel size of 45 nm, smaller than the simulated minimum at 100 nm (Figure 2.9C), which we attribute to the finite fitting window, not modelled in the theory. The experimental data matches the simulations well, with a small underestimation, which can be attributed to the incomplete sampling of of DNA conformations already discussed. For DNA of length 1000 bp, the uncertainty in the FIW is relatively insensitive to pixel size from 100 nm to 300 nm (Figure 2.9C). We suggest that a pixel size approximately equal to the FIW is optimum for TFM, despite different DNA lengths having different optima, because the shallow gradient of the uncertainty in this region allows an experimentally useful range of FIWs to be studied without changing the microscope construction.

2.3.4 Resolution

TFM and TPM rely on similar physical principles to measure an effective length along DNA. Hence, to give some context to the resolution and practical considerations of TFM, we compared it to TPM. TFM can be implemented on a fluorescence microscope utilising an EMCCD camera, whereas TPM is typically performed on a dark-field microscope utilising a camera capable of high frame rates [17]. We anticipate that when designing an experiment, the ease of combining TFM with existing fluorescence methods, the equipment available,

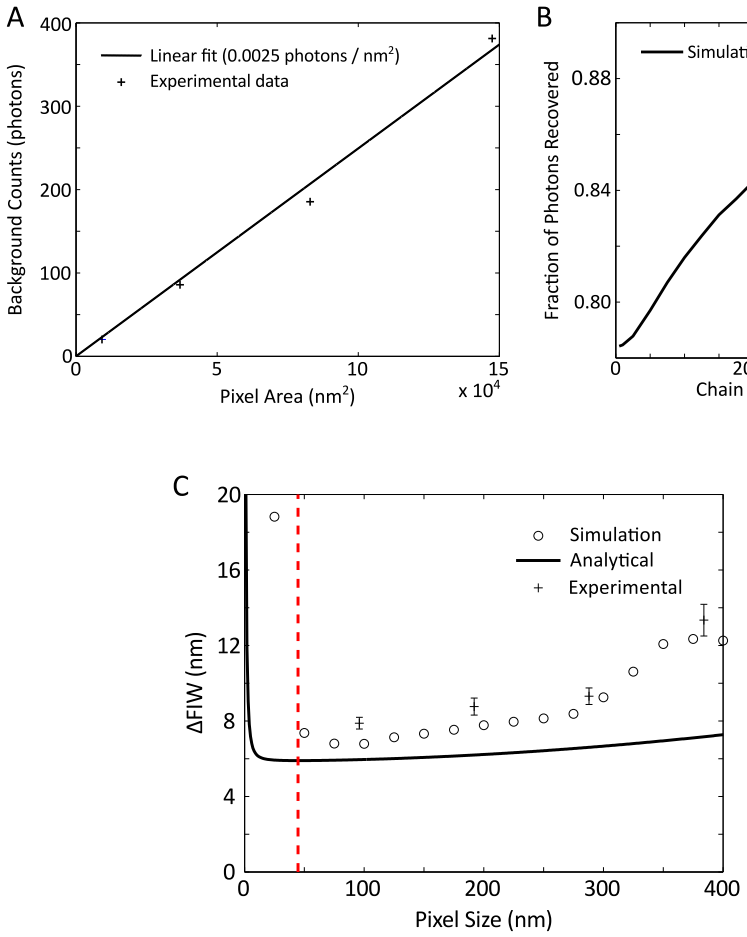

Figure 2.9: Noise in TFM **(A)** Dependence of background photon count on the area covered by each pixel in the image plane. Four different pixel areas were achieved using binning on our EMCCD camera. A linear fit was constrained to go through the origin. This dependence was taken into account in our simulation and evaluation of our analytical expression when deducing an optimum pixel size. **(B)** Underestimation of photon count by ordinary least-squares Gaussian fitting to simulated images. We simulated 80 000 images per chain length using a persistence length of 48 nm, pixel size of 96 nm, $N = 1000$, and a gain, read noise and background count chosen to give $\sigma_b = 5$. As the chain length was increased, the underestimation became less significant, which coupled with intensity changes due to the evanescent excitation field in TIRF makes recovered photon count an unreliable reporter of DNA length. **(C)** Optimum pixel size, using $N = 2250$, and 1000 bp DNA with a FIW of 188 nm. The analytical expression has a minimum at a pixel size of 45 nm (red dashed line), which is smaller than the simulation minimum at around 100 nm. As with the dependence on intensity, the analytical expression appears to underestimate the amount of noise present; the underestimation at larger pixel sizes can be explained by our approximation of sums as integrals in the derivation. We have taken the underestimation of photon count, by our image analysis algorithm, into account by matching recovered counts with our simulation. Error bars are the standard error in the mean from $>$ 15 molecules. Data were taken at 2 Hz.

and the desired observation time will be deciding factors.

The temporal resolution of both techniques is fundamentally limited by the relaxation time of the DNA-probe system (where the probe is either a bead in TPM or an organic fluorophore in TFM). In TPM, even at a small bead size (20 nm), the relaxation time of a 400 bp DNA is almost doubled, giving a temporal resolution of 20 ms for a 400 bp looping event, in a 800 bp molecule [18]. Since, in TFM, a small fluorophore is used, the drag exerted on the DNA will be smaller and temporal resolution will be improved. However, photobleaching of the fluorophore(s) used will limit the observation time. If a mean excursion were calculated using short ($\sim$1 ms) frames in a 20 ms window, with a sufficient photon count of $\sim$1000 photons/frame, corresponding to a localisation precision of $<$ 10 nm [39], we would expect a Cy3B fluorophore to last $\sim$30 s before photobleaching, while at least matching the temporal resolution of TPM.

TPM can typically detect 200 bp changes in an 800 bp DNA using an averaging window of 4 s [77]. Using TFM, $\sim$2500 photons would be needed in 4 s to match this spatial resolution (Equation (2.12) and Figure 2.8B). This photon count is readily achievable (Figure 2.7A has 1000 photons per 100 ms frame), hence, such requirements allow an observation time of several tens of minutes before photobleaching.

2.3.5 Polymerisation by Klenow fragment

Having looked at the theoretical basis of TFM, and compared its resolution to TPM, we now move to discussing practical applications of TFM. We used TFM to monitor the progress of polymerisation by the Klenow fragment (KF) of *Escherichia coli* DNA polymerase I (lacking the 5'-3' nuclease domain) [68]. The substrate consisted of a 55 bp double-stranded segment, tethered to the surface, followed by a 1032 base single-stranded segment, which acted as the template strand, with an internal Cy5 label at the surface distal end (Figure 2.2B). As polymerisation proceeds, the double-stranded segment is extended and the FIW increases (Figure 2.10A). To establish a calibration, WLC simulations of DNA

at various stages of polymerisation were performed, and an approximately linear relation between the polymerisation rate and the rate of change of the FIW (Section 2.2.4 and Figure 2.10B) was obtained, with a gradient of 25 bases/nm.

To analyse the data, we manually selected molecules that did not bleach during data acquisition and showed increasing FIW consistent with polymerisation initiating, progressing and terminating (Figure 2.10C). The gradient of the FIW with respect to time was estimated using linear least squares fitting in MATLAB, and this was converted into a polymerisation rate using the calibration. The end of polymerisation was evident as a drop in fluorescence intensity (Figure 2.10C, dashed blue lines). This can be attributed to the change in the local environment of the fluorophore as the DNA in its vicinity is converted from single-stranded to double-stranded. The linear fits were performed from 50 s, which allows enough time after the addition of KF and nucleotides for polymerisation to have initiated, to $\sim$20 s before the drop in fluorescence intensity. In 50 % of the timetraces, a small fluorescence enhancement is evident before the large drop in fluorescence, which we attribute to protein induced fluorescence enhancement (PIFE) due to the proximity of the KF [78]. PIFE is an established fluorescence technique, which occurs when a protein reaches the proximity (1 nm to 2 nm) of a fluorophore, such as Cy5, that undergoes a cis-trans isomerisation. The presence of the protein reduces the rate of isomerisation into the photo-inactive cis state. In our experiments, PIFE may only have been visible in some timetraces because of the low frame rate chosen, or because some events were obscured by noise.

We measure a mean polymerisation rate of 4.5 ± 0.7 bases/s (mean $\pm$ standard error in mean, from the six molecules in Figure 2.10C), at a nucleotide concentration of 10 µM. This is slower than the 6 bases/s previously measured using a 1 pN stretching force and a 100 µM nucleotide concentration [79]. We note, however, that our experiments have been performed at a lower nucleotide concentration and under no applied force. Without this stretching force, it is likely that the formation of transient secondary structure slows

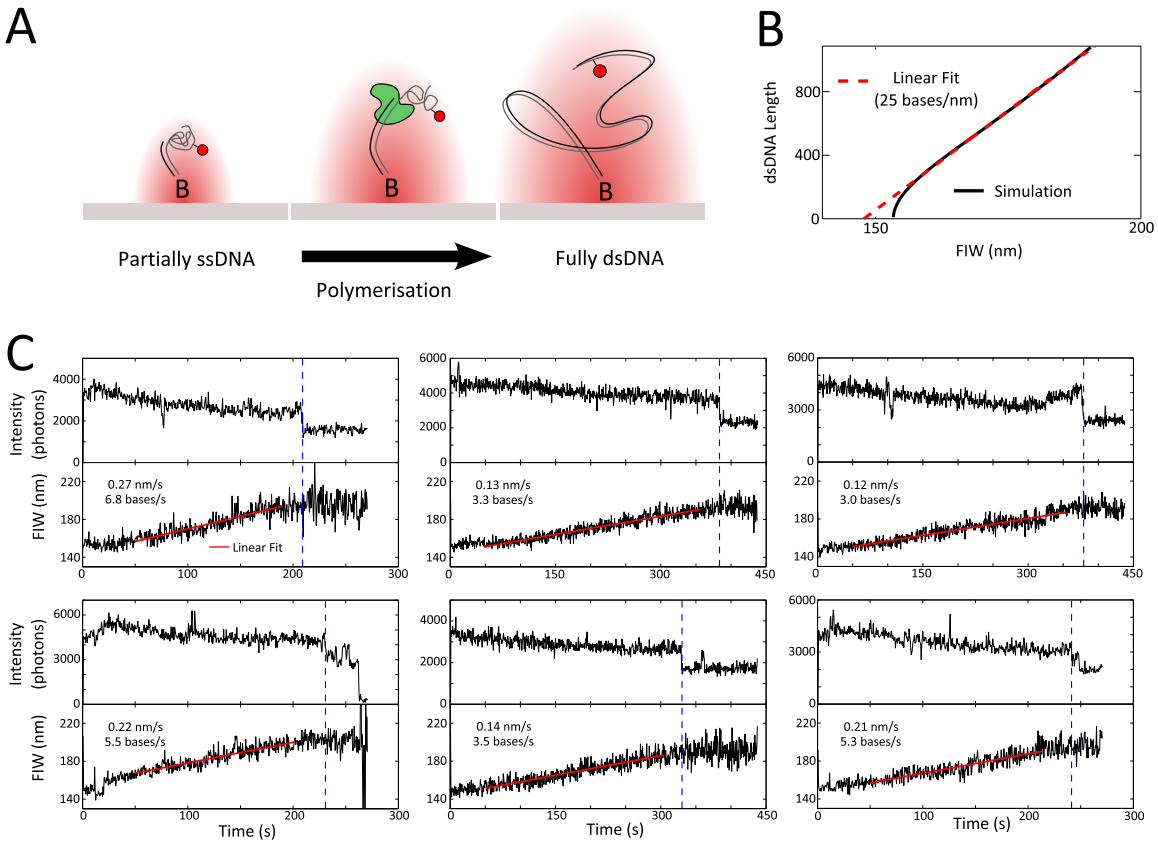

Figure 2.10: Polymerisation by KF **(A)** Schematic. Initially the predominantly ssDNA substrate has a diffraction limited FIW. As the DNA is converted to double stranded during polymerisation the mean persistence length increases, broadening the FIW until polymerisation is complete. **(B)** Predicted FIW during polymerisation. A linear fit to the region from 200 bp onwards, extracts an approximate conversion of 25 bases/nm. **(C)** Time traces from six polymerising molecules. Blue dashed lines indicate the drop in fluorescence associated with the completion of polymerisation, due to the conversion from single to double stranded of the DNA in the vicinity of the fluorophore. The red lines are linear least squares fits to the actively polymerising segments of the time traces, defined as from 50 s after the start of data acquisition until 20 s before the end of polymerisation. The gradients of these fits can be used to extract a mean polymerisation rate of 4.5 ± 0.7 bases/s. Molecules were selected to have not bleached before the end of the movie. In some (50 %) of the time traces a small PIFE is evident before the completion of polymerisation, which can be attributed to the proximity of KF to the fluorophore.

down the progress of the polymerase. The rate obtained using TFM is also slower than the previously determined 14 bases/s at a 100 µM nucleotide concentration [80], since the method used here does not seek to resolve short (∼10 s) pauses, and hence these get averaged into the rate. Longer pauses (∼1000 s) [79] were not apparent in our data because the observation time was limited to ∼400 s before photobleaching.

2.3.6 DNA looping by Cre

Section 2.3.5 used a gradual change in FIW to report on the progress of a biological process. Here, we use TFM to observe a step change in effective DNA length, while simultaneously monitoring the conformation of a nucleoprotein complex using FRET.

Cre is a well-studied tyrosine recombinase of the bacteriophage P1 of *Escherichia coli* that forms synapses consisting of four Cre monomers and two *loxP*-sites [66]. Previous work [32], using alternating laser excitation (ALEX) [37], has looked at synapsis formation by Cre A36V. This work uses a single excitation laser to demonstrate the simplicity of TFM-FRET. We used a 1000 bp doubly labelled DNA substrate (Figure 2.2C and Figure 2.11A), where the Cy3B fluorophore acted as a FRET donor and its FIW reported on the diffusional freedom of the DNA and any synapse formation. The fluorophore positions were chosen so that, when a synapse formed, they would be close together enough for FRET.

Synapsis was apparent as a decrease in the FIW of the donor, coincident with an increase in the apparent FRET efficiency, E^* (Figure 2.11B, 110 s to 135 s). We segmented the timetraces using HMM based on the E^*, since this showed superior signal to noise compared to the FIW. The E^* from all frames during synapsis were binned into a histogram (Figure 2.11C), recovering $E^* = 0.30$ for the complexes formed. Dwell times were fit with a single exponential model using maximum likelihood estimation in MATLAB, extracting a dwell time of 32 ± 4 s (Figure 2.11D), broadly in agreement with the previously measured time of 52.1 s [32].

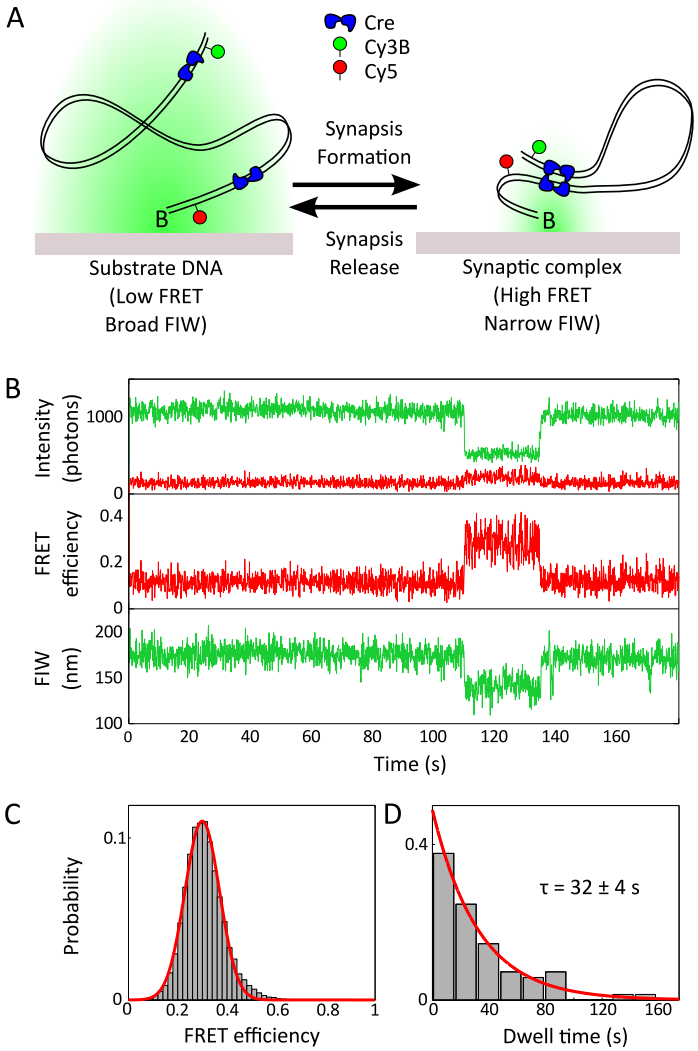

Figure 2.11: Cre-*loxP* synapsis formation **(A)** Cartoon. A 1 kb DNA is labelled with Cy5 and Cy3B, 10 bp and 990 bp from the biotin end respectively. The substrate DNA displays a broad Cy3B FIW. Upon addition of Cre, synaptic complex formation between *loxP*-sites located close to the labelling positions leads to a looping of the DNA and a decrease in FIW, as well as bringing the Cy3B and Cy5 fluorophores into close proximity, resulting in FRET. **(B)** Representative time trace of reversible synaptic complex formation in the presence of Cre, with the intensities corresponding to the emission of the donor, in green, and acceptor, in red, under 532 nm continuous excitation. Synapsis is evident between 110 s and 135 s. **(C)** Histogram of FRET efficiency, measured within the looped complex formed, with mean $E^* = 0.299 \pm 0.004$, determined using an unweighted least-squares fit to the data. **(D)** Histogram of dwell times of individual looping events, fit using maximum likelihood to a one parameter single exponential with a dwell time of 32 ± 4 s. The uncertainty in fit parameters is defined as the 1 σ confidence interval of the fit ($n = 52$). Movies were taken with a laser power of 1 mW and a frame rate of 10 Hz.

2.4 Conclusions

TFM uses the diffusional characteristics of a fluorophore tethered to a slide to report on the effective length of the polymer used as the tether. The use of a single fluorophore (rather than a bead) enables TFM to be combined with a variety of single-molecule fluorescence methods and observables, such as FRET and PIFE. The absence of a large colloid ensures that the motion of the DNA, rather than the bead, determines the dynamics of the system, and eliminates the entropic stretching force exerted on the DNA. Sample preparation is simplified, since there is only one non-covalent attachment chemistry and hence no risk of multiply tethered particles. TFM can be readily implemented on any existing single-molecule wide-field microscope; conveniently, the optimum signal-to-noise for TFM requires a pixel size very similar to that for single-molecule localisation microscopy. Despite these advantages, TFM is limited in observation time to the lifetimes of the fluorophores used, and signal-to-noise has to be traded against this lifetime. However, we have shown that in conditions typical for single-molecule experiments, length resolutions down to 100 bp are achievable, and we note that in principle the time resolution of TFM can exceed that of TPM. More photostable fluorophores, such as Atto 647N (ATTO-TEC GmbH, Seigen, Germany) would allow longer observations.

We have also demonstrated the use of TFM to follow processes that change the conformation of DNA either gradually or in a single step. We have monitored the motion of KF as it polymerises DNA along a template strand, and we have watched reversible synapsis formation by Cre. Although our KF experiments used a fluorophore attached to the DNA, a labelled protein can be used as the TFM reporter, enabling much longer observation times, as the turnover of protein binding to the DNA would continually replace bleached fluorophores. In this case, the second fluorophore on KF would allow direct measurement of KF processivity, and a FRET scheme [68], along with alternating laser excitation, could be used to observe protein conformational changes and correlate them to the progress of polymerisation. TFM can also be combined with another single-molecule fluorescence

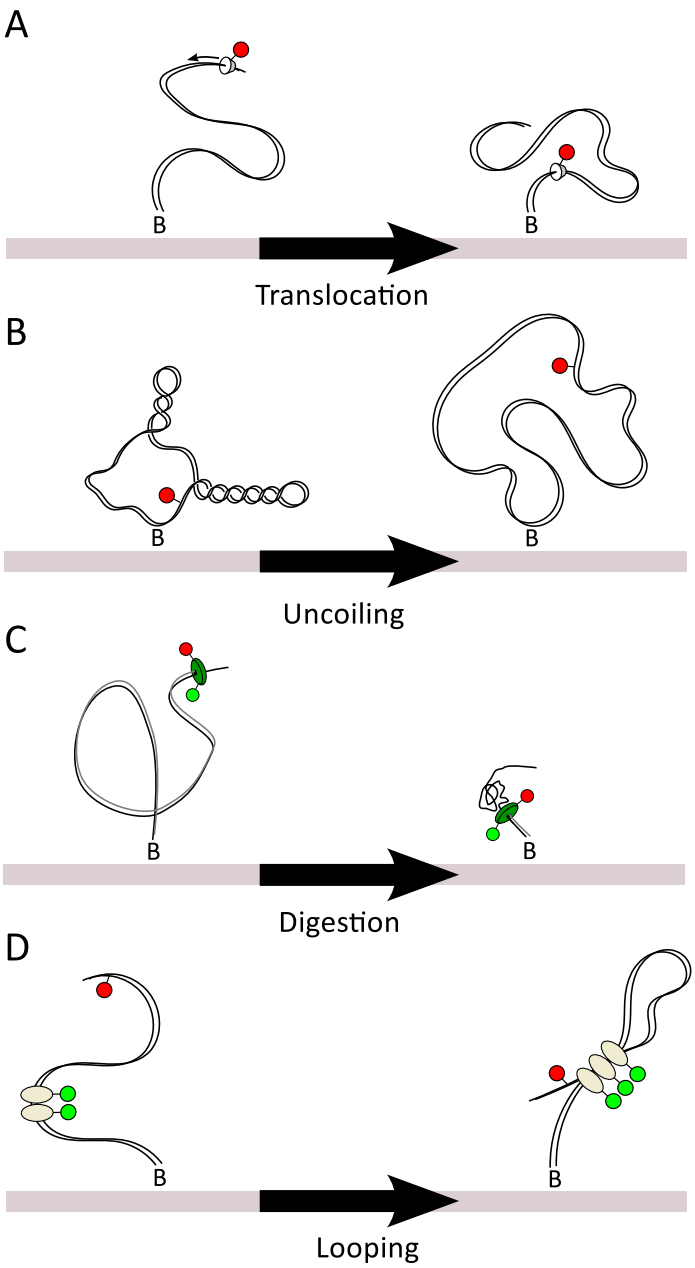

Figure 2.12: Ideas for TFM use **(A)** Labelled translocase on a DNA track. As the translocase moves towards the surface the FIW decreases. **(B)** Topological simplification of DNA. Supercoiled, or knotted DNA, would show an increase in FIW as an enzyme simplifies their topology. **(C)** Digestion of DNA. An enzyme, such as lambda exonuclease, that digests DNA would reduce the effective persistence length, reducing the FIW. A FRET pair can be used as the TFM reporter whilst simultaneously monitoring protein conformation. **(D)** Looping by a protein of unknown stoichiometry. The FIW of the green fluorophore can be used to determine where along the DNA the proteins have bound, and their stoichiometry can be deduced from the green intensity. Looping can be inferred by a decrease in red FIW.

method, PIFE, to follow interactions between multi-protein molecular machinery.

Recombination by XerCD is activated by the DNA translocase, FtsK, and the labelling of such translocases and accessory proteins offers another set of TFM experiments (Figure 2.12A). The FIW on translocase binding can be used to determine where along the DNA it binds; any subsequent motion along a linear substrate can be followed as changes in the FIW. Circular substrates also offer the possibility of watching topological simplification of DNA by topoisomerases (Figure 2.12B) [81]. Plectonemes have lengths of several kb [82] and so their presence would be apparent as a reduced FIW; unknotting or uncoiling would increase the FIW. Analogously to polymerisation, digestion by lambda exonuclease [83] could be monitored using TFM-FRET (Figure 2.12C). Finally, in contrast with TPM, multiple spectrally separate TFM probes can be attached to the same DNA, to simultaneously measure the effective length at multiple sites along DNA (Figure 2.12D).

2.5 Contributions

The work in this chapter has formed the basis of a publication [1]. Figures and text have been adapted from this publication, and were produced (for the publication) in collaboration with Justin Pinkney. The text in the publication was edited by Achillefs Kapanidis. Pawel Zawadzki purified proteins and produced the DNA substrates. With the exception of the KF experiments and the data in Figure 2.7, which were generated by Justin Pinkney and Geraint Evans, I performed all experiments in this chapter. I performed all the analysis, simulations and the derivations.

Chapter 3

Activation of XerCD-*dif* recombination

3.1 Introduction

In this chapter, I describe the application of tethered fluorophore motion (TFM), in combination with two other single-molecule fluorescence techniques, Förster resonance energy transfer (FRET) and protein induced fluorescence enhancement (PIFE), to track the conformational transitions in individual protein-nucleic acid complexes. The system studied, XerCD-*dif*-FtsK, plays a key role in chromosome unlinking in most bacteria with circular chromosomes. We observed the formation of XerCD-*dif* synapses and deduced a five-step pathway for the recombination reaction. We compared the structures of the states in the pathway to those of a well-studied recombinase, Cre, and conclude that the substrate for FtsK activation does not resemble the Cre-*loxP* structure. Using PIFE, we observed the proximity of FtsK to the recombining complex.

3.1.1 XerCD recombination

Tyrosine recombinases, such as XerC and XerD, share a common reaction mechansim whereby a nucleophillic tyrosine residue catalyses recombination via a Holliday junction (HJ) intermediate [45, 51, 84]. Four recombinase monomers bind two DNA sites together, and then reciprocal activation-inactivation of pairs of recombinases occurs after a first round of strand exchanges, during isomerisation of the HJ intermediate. This model for the action of tyrosine recombinases comes from biochemical and structural studies of the Cre-*loxP* recombination system [45, 51, 85, 86, 87] (Figure 3.1A).

Two monomers of XerC and two monomers of XerD bind their respective half-sites. Each pair (of XerC or XerD) mediates the exchange of a particular strand: 'top' strands by XerC and 'bottom' strands by XerD (Figure 3.1B). XerCD act at a variety of sites, found naturally on plasmids and chromosomes. One particular site, *dif*, is located in the terminus region of the chromosome. Recombination at *dif* requires activation by FtsK, a DNA translocase. In the absence of FtsK, XerC can mediate strand exchanges to form HJs, but no product (which would be formed by a subsequent XerD-mediated strand exchange) is produced [88, 89]. However, in the presence of FtsK, productive recombination proceeds with an initial XerD-mediated strand exchange to form a HJ, which is then resolved to product by XerC [88, 90]. It has remained unclear how FtsK gives rise to synaptic structures poised for XerD-mediated strand exchange, although it has been proposed that FtsK actively remodels a synaptic complex that was poised for XerC-mediated strand exchange [90]. If the Cre-*loxP* model holds for XerCD-*dif*, then such remodelling would require the breaking of protein-protein or protein-DNA contacts, or the complete reassembly of the synaptic complex.

3.1.2 Combination of fluorescence methods

To monitor the progression of the recombination reaction in vitro, we used three independent observables from three fluorescence methods: FRET, PIFE and TFM. FRET allowed us

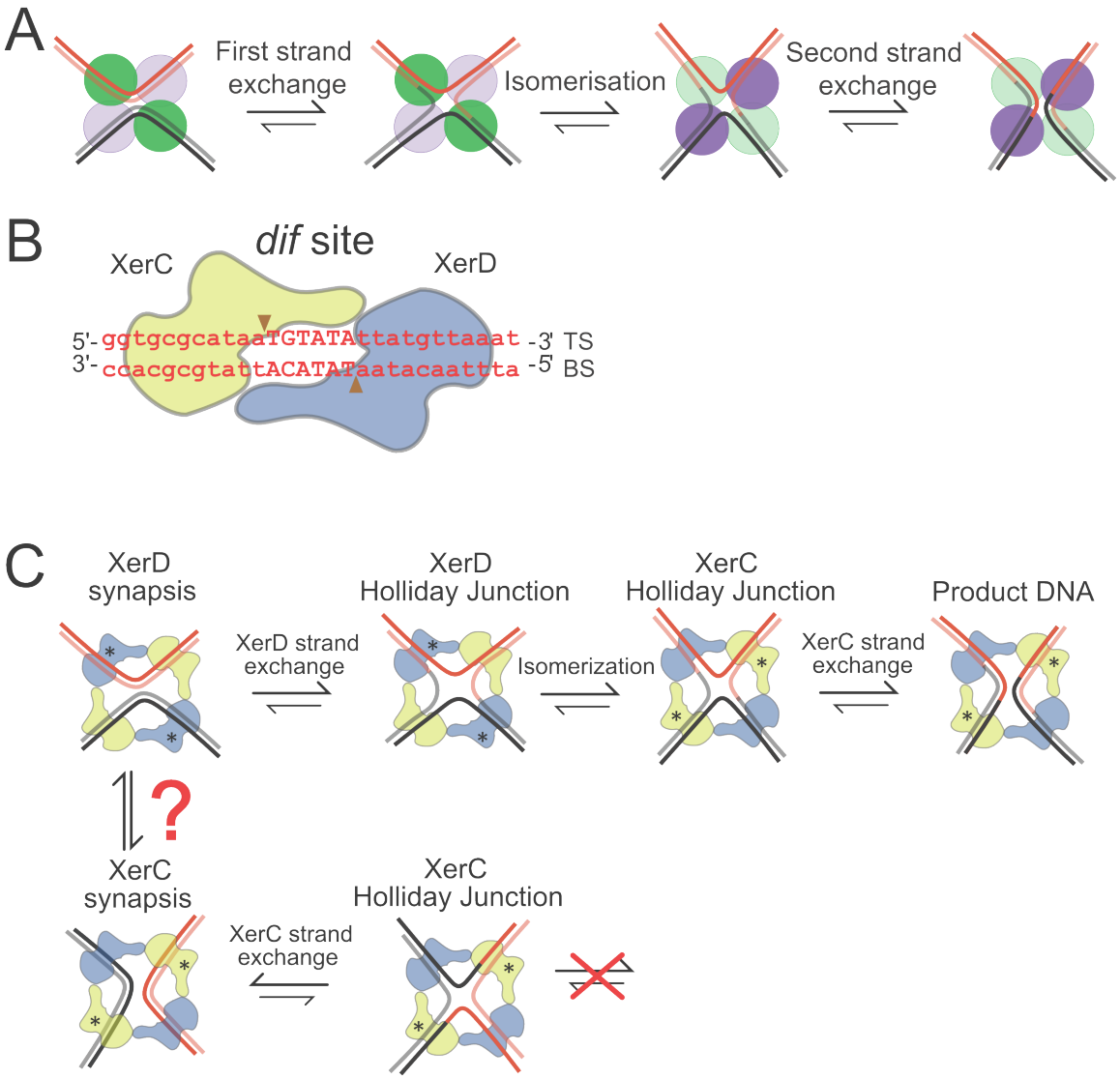

Figure 3.1: Site-specific recombination. **(A)** Recombination pathway for tyrosine recombinases. Green and violet circles are recombinase proteins. Darker colour indicates activity. Upon HJ isomerisation the catalytic activity switches from one pair of recombinases to the other. **(B)** XerC and XerD binding sites. Arrows indicate the phosphate cut during catalysis. Top strand (TS) and bottom strand (BS) are marked. **(C)** Predicted XerCD-*dif* recombination pathway, based on the Cre pathway. XerC or XerD synapsis forms then undergoes strand exchange. Conversion may be possible between the two isoforms, but would require the breaking of nucleoprotein interactions. HJs produced by an initial round of XerC mediated strand exchange cannot isomerise and proceed to product (bottom), but those produced by XerD can (top).

to monitor nanoscale changes in conformation within the XerCD-*dif* complexes, TFM allowed us to monitor the formation of synapses and the completion of recombination, and PIFE allowed us to monitor the proximity of FtsK to the recombining complex. A detailed description of the methods used for single-molecule experiments, data analysis and the conversion of apparent FRET efficiencies into distances are given in the next section (Section 3.2), accompanied by Figures 3.2 to 3.4 and Tables 3.1 to 3.3. The results of single-molecule experiments are presented in Section 3.3.

3.2 Methods

3.2.1 Distance prediction

To compare the structure of the XerCD-*dif* complex to that of a well-studied recombinase system, Cre-*loxP*, we predicted interfluorophore distances using the Cre-*loxP* crystal structure, 2HOI [66]. Other distances derived using this structure have been compared to those derived using accurate FRET in Cre-*loxP* complexes, and have been found to be in close agreement [32]. We performed three-radii accessible volume calculations using FRETnps software [33], which computes the physically accessible space that the fluorophore can occupy (constrained by the six-carbon linker). The modelling is performed three times, with the fluorophore modelled as a sphere with radius equal to one of the three principal radii of the fluorophore. The position of the dye is given as the mean position of the three modelling runs. DNA arms in the crystal structure were extrapolated using linear B-DNA, and the labelling position was found by counting from the scissile phosphate in the crystal structure the correct number of bases in the correct direction (Table 3.1). The predicted dye distances for the substrate DNA conformation (i.e. no strand exchanges have occurred), with the labelling scheme from Substrate I, were 10.4 nm if XerD-mediated strand exchange was poised to occur, and 6.5 nm if XerC-mediated strand exchange was poised to occur (Figure 3.2).

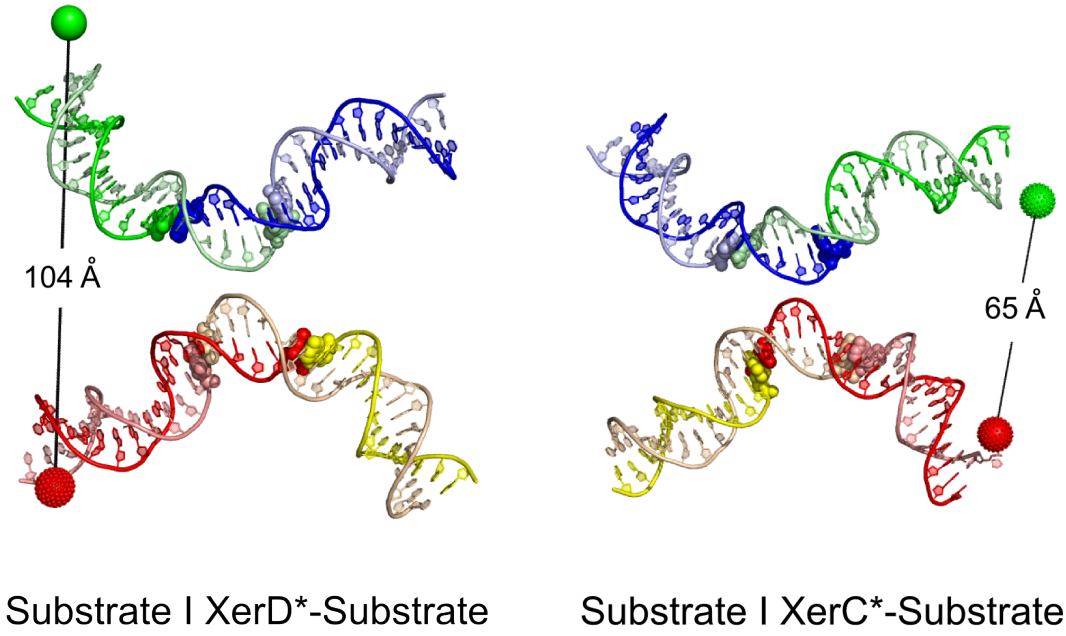

Figure 3.2: Predicted interfluorophore distances in Substrate I, using the crystal structure, 2HOI, for Cre-*loxP* [66]. The DNA has not undergone any rounds of strand exchange. Right: structure poised for cleavage by XerD. Left: structure poised for cleavage by XerC. XerC binding sites are shown in green and yellow, and XerD binding sites are shown in blue and red. Residues either side of a scissile phosphate shown as spheres. Given an R_0 of around 7 nm for the dyes used, we predict a FRET efficiency of 0.6 for an XerC* synapsis and 0.1 for an XerD* synapsis.

Table 3.1: Distance predictions using Cre-*loxP* structure.

	Distance (nm)			
Substrate	XerC*-Substrate	XerD*-Substrate	XerC*-Product	XerD*-Product
Substrate I	6.5	10.4	10.7	8.2
Substrate II	9.5	8.5	9.3	10.6
Substrate III	10.7	8.2	6.5	10.4

Distances between fluorophores were predicted using three-radii accessible volume calculations and the crystal structure of Cre-*loxP*.

3.2.2 Protein induced fluorescence enhancement

The proximity of a protein to an organic fluorophore can enhance or quench its fluorescence. Tryptophan residues can participate in a photo-induced electron transfer process, quenching fluorescence [91]. On the other hand, for dyes such as Cy5, which undergo a cis-trans isomerisation to a photo-inactive cis state, the proximity of a protein can reduce the rate of isomerisation to the photo-inactive cis state, increasing the quantum-yield of the fluorophore [78].

3.2.3 DNA and protein preparation

The 4 kb DNA substrates were produced a polymerase chain-reaction using Phusion High-fidelity DNA polymerase (New England Biolabs, Ipswich, USA). The oligonucleotides used as primers were fluorescently labelled, and the plasmid template contained two directly repeated *dif*-sites, separated by a 1 kb KmR gene cassette (pRB10). The oligonucleotides used were as follows: Substrate I forward 5'-GTGTCGACACAXGATTTAACATAAT-3' and reverse 5'-CTCTAGACCATGGAXCATGTGGTGCGCATA-3', Substrate II forward 5'-CATGTGTCGACACATGATTTAACATAATATACATTATGCGCACCACATGXAGCT GAGATCTG-3' and reverse 5'-AGACCATGGCATGTGGTGCGCATAATGTATATTATGT TAAATCXTGTGGATCCAC-3', Substrate III forward 5'-TAGCGTCGACCTACAXGATT TAACATAATATACATTATGCGCACCAAATGATTCGCAGCTGAGATC-3' and reverse 5'-GATCGATCTCAGCTGCGAAXCATTTGGTGCGCATAATGTATATTATGTTAAAT CATGTAGGTCGACGCTA-3'. X indicates the position of the fluorophore attachment, via six-carbon linker attached to a thymine base. Forward oligonucleotides were labelled with Cy5 and reverse oligonucleotides were labelled with Cy3B. Oligonucleotides were synthesised, high-pressure liquid-chromatography purified, and Cy5 labelled (ATDBio, Southampton, UK). Cy3B labelling was performed as previously described [32]. The PCR products were NcoI digested and ligated to a biotinylated 200 bp extension. After digestion with SalI, a 2.8 kb tail was ligated to the opposite end. The finished substrate was

gel-purified. Substrate III was prepared with unlabelled oligonucleotides, otherwise identical to those used for Substrate I. The distal *dif*-site was removed using a BglII digestion, and a fluorescently labelled site ligated back. Proteins were purified using established techniques [56, 92].

3.2.4 Instrumentation

Single-molecule total internal reflection fluorescence (TIRF) experiments were performed on a custom-built objective-type TIRF microscope. A green (532 nm Cobolt Samba, Cobalt, Solna, Sweden) and red (635 nm Cube, Coherent, Santa Clara, USA) laser were combined using a dichroic mirror and coupled into a fibre optic cable. The output of the fibre was focussed into the back focal plane of the objective (100× oil-immersion, NA 1.4, $f/26.5$, UPlanSApo, Olympus, Tokyo, Japan) and displaced perpendicular to the optical axis such that laser light was incident at the slide-solution interface at greater than the critical angle, creating an evanescent excitation field. Alternating laser excitation was implemented by directly modulating the lasers and all data were taken using a 10 Hz alternation rate (20 Hz for the Substrate II XerC^{KQ} experiments). Excitation powers were set to 1 mW for both lasers, which corresponds to an approximate power density of 0.4 µW/µm^2 incident on the 50 µm × 50 µm field of view. Fluorescence emission was collected by the objective and separated from the excitation light by a dichroic (545 nm/650 nm, Semrock Rochester, USA) and cleanup filters (545 nm long-pass, Chroma, Bellows Falls, USA, and 633/25 nm notch filter, Semrock). The emission signal was focussed onto a rectangular slit to crop the image, and spectrally separated using a dichroic (630 nm long-pass, Omega, Brattleboro, USA) into two emission channels, which were focussed side-by-side onto an electron multiplying charge-coupled device camera (EMCCD, iXon 897, Andor, Belfast, UK). The EMCCD camera was set to an EM gain of 300, corresponding to an approximate real gain of 4.55 counts per photon. Each pixel on the camera corresponded to a 96 nm × 96 nm in the sample plane.

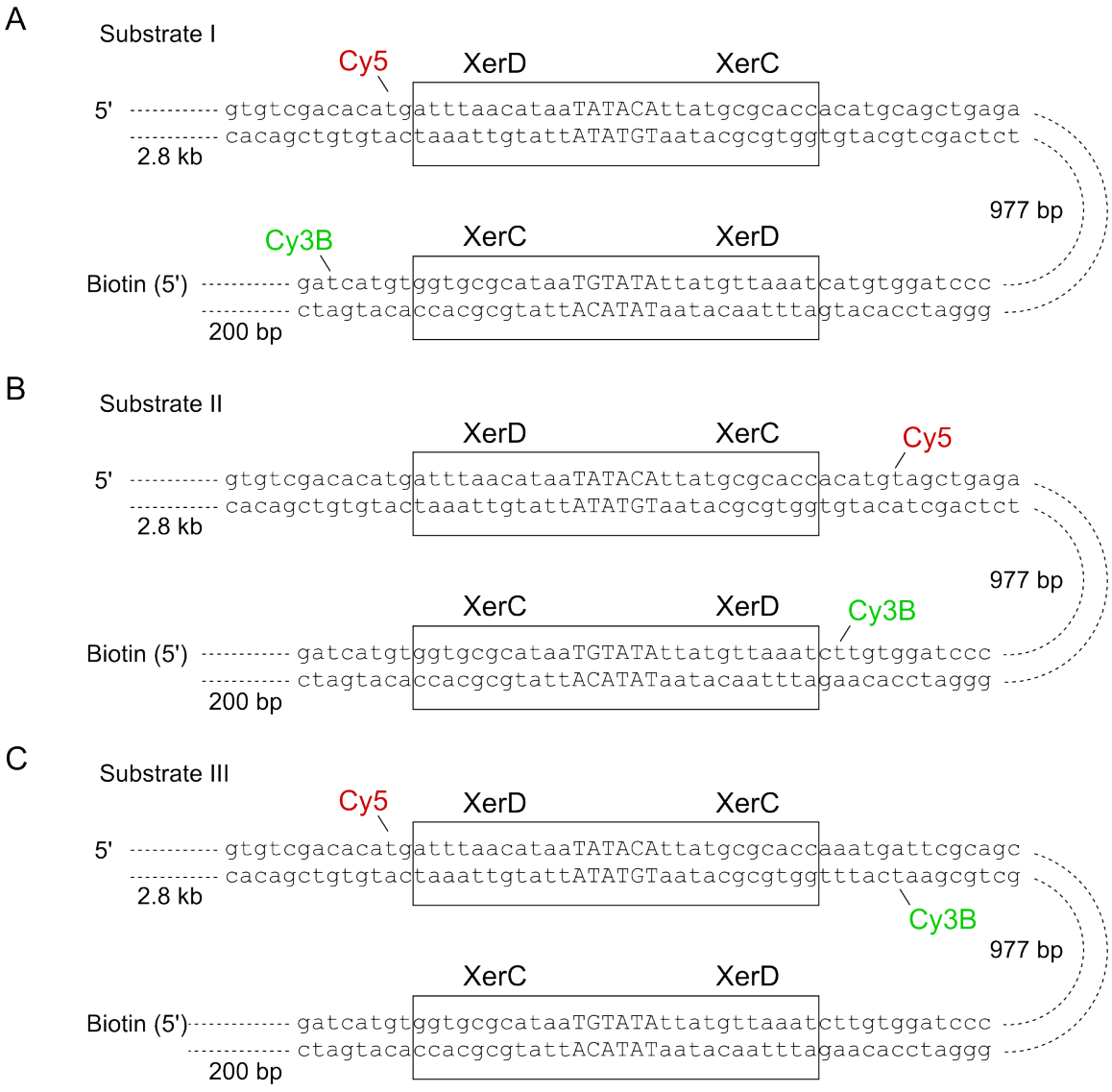

Figure 3.3: DNA used in this chapter. The *dif* sites are indicated by boxes, and XerC and XerD binding sites are labelled. Fluorophores were attached using a six-carbon linker. **(A)** Substrate I **(B)** Substrate II **(C)** Substrate III

3.2.5 Sample preparation

DNA was immobilised to the surface of a coverslip using biotin-neutravidin interactions. The coverslip was passivated using poly-ethylene glycol, and sealed using a silicone gasket (Grace Bio-labs, Bend, USA) and a second coverslip as a lid. A buffer consisting of 50 mM Tris-HCl (pH 7.5), 50 mM NaCl, 5 mM $MgCl_2$, 100 µg/ml BSA, and 1 mM UV-treated Trolox, was used for imaging. An enzymatic oxygen scavenging system consisting of 1 mg/ml glucose oxidase, 40 µg/ml catalase and 1.4 % (w/v) glucose was added prior to sealing the sample before image acquisition. Experiments were performed at a temperature of 21 °C.

3.2.6 Data analysis

Previously described TwoTone software was used to extract fluorescence intensities from images [42]. We calculated the apparent FRET using:

$$E^* = \frac{DA}{DD + DA} \tag{3.1}$$

where DD is the area under the Gaussian fit to the donor emission under donor excitation, and DA is the area under the Gaussian fit to the acceptor emission under acceptor excitation. The FIW in all channels was the mean width of the fitted elliptical Gaussian.

3.2.7 HMM for data extraction

A combination of manual and hidden Markov modelling (HMM) based timetrace segmentation was used to generate E^* histograms, which were fit with Gaussian distributions using non-linear least-squares fitting in MATLAB (Figure 3.4). HMM can be used as a means of resolving FRET states present in noisy data [93]. In a Markov model, the probability of a transition only depends on the present state of the system. HMM refers to

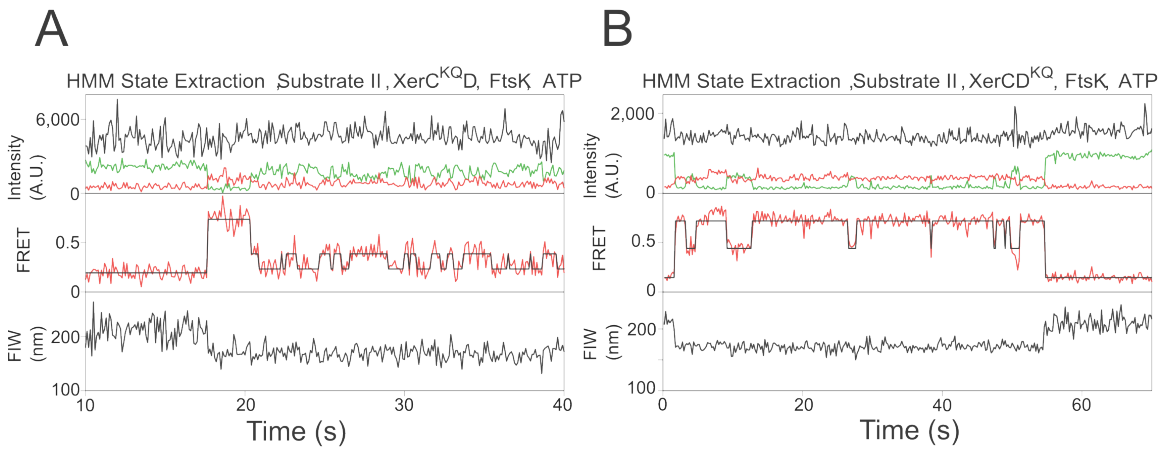

Figure 3.4: HMM used for state extraction. **(A)** Timetrace, in the presence of XerC^{KQ}, XerD, FtsK$_C$ and ATP, shows oscillations between XerD-HJ ($E^* = 0.37$) and XerC-HJ ($E^* = 0.18$). Black line (middle panel): HMM-idealised timetrace, used to identify transitions between states. **(B)** Repeated activation of synaptic complex is apparent as transitions to $E^* = 0.44$. Black line (middle panel): HMM-idealised timetrace, used to identify and segment these transitions.

the hidden nature of the underlying state which can produce a range of outputs (apparent FRET efficiencies) at each time point sampled. To perform the HMM on each timetrace, we used Seneca, a package written in MATLAB [94, 95]. The extracted states were categorised according to their FRET efficiency and the state that preceded them. Using this categorisation, data belonging to particular structural states were plotted together in histograms and fit (Table 3.2). Data immediately before or after a transition were discarded to minimize camera integration time effects [75].

3.2.8 Accurate FRET

So that we could extract a distance that corresponded to each FRET state, we followed the previously established technique for correcting FRET efficiencies [36]. Defining the donor leakage into the acceptor emission channel as l, the direct excitation of the acceptor as a, and the ratio of detection efficiencies and quantum yields between the fluorophores as γ, the accurate FRET efficiency is:

Table 3.2: Parameters of Gaussian fits to FRET efficiency histograms

Population	Mean	SEM	SD	SEM of SD
Substrate I				
Synapsis	0.549	0.001	0.082	0.001
Recombination	0.571	0.002	0.111	0.002
Substrate II (XerCD)				
Synapsis	0.720	0.002	0.086	0.002
Initial complex	0.717	0.002	0.101	0.002
XerD*	0.372	0.003	0.106	0.003
XerC-HJ/XerC-P	0.174	0.001	0.050	0.001
Substrate II ($XerC^{KQ}D$)				
Initial start	0.727	0.002	0.095	0.002
Initial end	0.718	0.002	0.097	0.002
XerD-HJ	0.366	0.002	0.079	0.002
XerC-HJ	0.217	0.001	0.061	0.001
Substrate II ($XerCD^{KQ}$)				
Initial complex	0.718	0.001	0.081	0.001
XerD*	0.438	0.004	0.094	0.004
Substrate III				
XerC-HJ	0.559	0.003	0.096	0.003
XerC-P	0.400	0.002	0.078	0.002

Data extracted using a combination of manual and HMM timetrace segmentation, then fit with a single Gaussian distribution for each population/state. The mean and standard deviation (SD) are from the fit, and the standard error in the mean (SEM) is defined as the 1σ confidence interval on the parameter fit.

$$E = \frac{DA_{correct}}{\gamma DD + DA_{correct}} \tag{3.2}$$

where,

$$DA_{correct} = DA - lDD - dAA \tag{3.3}$$

and

$$\gamma = \frac{\Delta DA_{correct}}{\Delta DD} \tag{3.4}$$

where $\Delta DA_{correct}$ is the change in $DA_{correct}$ either side of a change in FRET state, AA is the acceptor emission under acceptor excitation, and ΔDD is the change in DD either side of the same transition. Transitions, rather than photobleaching events [96], were chosen since they are more frequent in our data, and they allow us to confirm that the variation in γ between states is small, supporting our use of a mean γ for all populations in each substrate. The values for l, d, and γ were 16 %, 8 %, and 0.38 respectively, averaging across all substrates.

The Förster radius, R_0 was calculated from:

$$R_0 = 0.211 \left(\kappa^2 n^{-4} Q_D J\right)^{\frac{1}{6}} \tag{3.5}$$

using the overlap integral, J, which is given by:

$$J = \int f_D(\lambda)\epsilon_A(\lambda)\lambda^{-4}\,\mathrm{d}\lambda \tag{3.6}$$

where: n is the refractive index of the medium, $n = 1.33$; $f_D(\lambda)$ is the normalised fluorescence emission spectrum of the donor; $\epsilon_A(\lambda)$ is the acceptor molar extinction coefficient; Q_D is the quantum yield of the donor; and κ^2 is the orientation factor. The emission and absorption spectra were measured on a fluorometer (Photon Technology International, Edison, USA) for singly labelled central regions of Substrate I and Substrate II, in the presence of wild-type XerCD and in the same buffer as the recombination experiments,

but without Trolox or the oxygen scavenging system. A literature value of 0.67 was assumed for Q_D [28]. A Perrin plot of the anisotropy was used to estimate the freedom of each fluorophore relative to its attachment point, and, using FRETnps software, we confirmed that $\kappa^2 = 2/3$ was a reasonable approximation, and we estimated the 67 % confidence interval [97]. The errors were propagated [98], and distances calculated (Table 3.3). The average R_0 for the three substrates was 6.5 nm, consistent with the expected value from similar organic fluorophores [35].

Table 3.3: Accurate FRET efficiencies and derived distances

Population	Accurate FRET	SEM	Distance (nm)	Uncertainty in distance (nm)
Substrate I				
Synapsis	0.68	0.02	5.86	0.88
Recombination	0.70	0.02	5.76	0.87
Substrate II (XerCD)				
Synapsis	0.84	0.01	5.16	0.77
Initial complex	0.82	0.02	5.32	0.80
XerD*	0.39	0.06	7.35	1.14
XerC-HJ/XerC-P	< 0.2		> 9.0	
Substrate II ($XerC^{KQ}D$)				
Initial start	0.83	0.01	5.24	0.78
Initial end	0.83	0.01	5.26	0.79
XerD-HJ	0.42	0.05	7.20	1.09
XerC-HJ	< 0.2		> 9.0	
Substrate II ($XerCD^{KQ}$)				
Initial complex	0.83	0.01	5.24	0.78
XerD*	0.53	0.03	6.68	1.01
Substrate III				
XerC-HJ	0.72	0.01	5.67	0.86
XerC-P	0.54	0.02	6.47	0.98

The uncertainty in the each of the accurate FRET distances is ∼1 nm, the majority of which can be attributed to the uncertainty in the orientation factor, κ^2. The use of multiple independent labelling schemes has allowed us to mitigate this, and to investigate structure despite the relatively large orientational uncertainty [99].

The smallest accurate FRET that could be reliably distinguished from the background due to leakage and direct excitation of the acceptor was established using a Raleigh criterion and a kinetic Monte Carlo Markov chain simulation. The simulation explicitly simulated photon emission times in each channel with the appropriate (experimentally determined) cross-talks, electron multiplying gain by the EMCCD, read noise and the least-squares fitting to the image. The photon count was chosen to match the width of the $E^* = 0.72$ population in Substrate II, which accounts for the excess heterogeneity found in previous studies [42]. It was found that $E = 0.20$ was just resolvable against the background, corresponding to a maximum measureable distance of around 9 nm (Figure 3.5). Hence, any states that showed no apparent FRET correspond to an interfluorophore distance > 9 nm.

3.2.9 Rate determination

The transition rates, $k = 1/\tau$, quoted in this work (Figure 3.18) have been extracted using maximum likelihood fitting in MATLAB. The probability density distribution for a dwell, t, was:

$$p(t) = k \exp(-kt) \tag{3.7}$$

In Substrate III, reverse transitions were observed (Figure 3.14):

$$\mathrm{XerC{-}HJ} \underset{}{\overset{k_+, k_-}{\rightleftharpoons}} \mathrm{XerC{-}P} \xrightarrow{k_2} \mathrm{Dissociation}$$

where k_+ and k_- refer to the forward and backwards rates for the conversion of

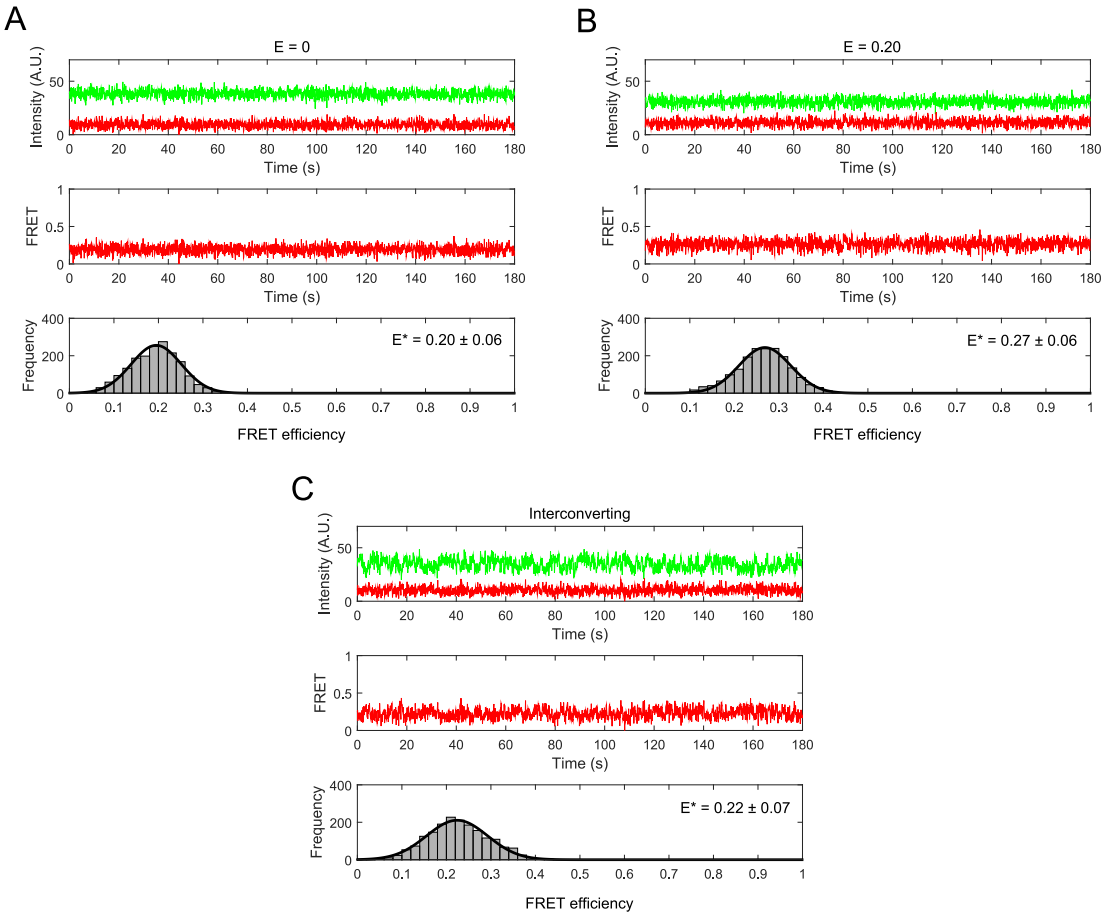

Figure 3.5: Simulation of minimum resolvable FRET. Using the measured cross-talks and relative detection efficiencies, we simulated images of fluorophores and generated time traces. **(A)** The top panel shows the simulated intensity in the DD (green) and DA (red) channel. The middle panel shows the apparent FRET efficiency of each frame. The bottom panel shows the distribution of apparent FRET efficiencies. With no simulated FRET, the apparent FRET due to leakage has a mean of 0.20 and a standard deviation (SD) of 0.06. **(B)** With a simulated 'real' FRET efficiency of 0.20, the apparent FRET distribution has a mean of 0.27 and an SD of 0.06. According to a Rayleigh criterion, these two populations are just distinguishable, and so $E = 0.20$ represents the smallest FRET efficiency that can be resolved against the background due to cross-talks. **(C)** As a sanity check, a system with two states with one at $E = 0$ and the other at $E = 0.20$ was simulated. The interconversion rate was $1\,\mathrm{s}^{-1}$ in both directions, which is close to the rates found in XerCD-*dif* recombination. Interconversions between states are just resolvable by eye in the simulated FRET timetrace.

XerC-HJ to XerC-Product (XerC-P), and k_2 is the rate of dissociation of XerC-P. The summed dwells in XerC-P, $t_{\Sigma\mathrm{XerC-P}}$, which are the total time each molecule spends in XerC-P before dissociating, should follow the distribution:

$$p\left(t_{\Sigma\mathrm{XerC-P}}\right) = k_2 \exp\left(-k_2 t_{\Sigma\mathrm{XerC-P}}\right) \tag{3.8}$$

The dwell time in XerC-HJ before each transition to XerC-P, $t_{\mathrm{XerC-HJ}}$, should follow:

$$p\left(t_{\mathrm{XerC-HJ}}\right) = k_+ \exp\left(-k_+ t_{\mathrm{XerC-HJ}}\right) \tag{3.9}$$

and the dwell time spent in XerC-P before either a transition to XerC-HJ or a dissociation, $t_{\mathrm{XerC-P}}$, should follow:

$$p\left(t_{\mathrm{XerC-P}}\right) = \left(k_2 + k_-\right) \exp\left(-\left(k_2 + k_-\right) t_{\mathrm{XerC-P}}\right) \tag{3.10}$$

These distributions were fit using maximum likelihood estimation in MATLAB, and the individual rates were deduced.

To place an upper bound on the reverse transition rates (Figure 3.18), we noted that the probability of observing fewer than m reverse transitions is:

$$p(<m) = \sum_{0<i<m} \left(\frac{k_f}{k_b + k_f}\right)^{N-i} \left(\frac{k_b}{k_b + k_f}\right)^{i} \frac{N!}{(N-i)!\,i!} \tag{3.11}$$

where k_f is the forward transition rate, k_b is the reverse rate, and N transitions have been observed in total. The limiting reverse rate was chosen such that if the 'true' reverse rate were faster than the chosen limit, our data (fewer than m reverse transitions) would have been observed in less than 5 % of datasets of our size (N total transitions). Setting $p = 5\,\%$, $m = 4$, $N = 249$, 134, or 111 for Substrate I, Substrate II and Substrate III respectively, and k_f to the appropriate forward rate, we can solve Equation (3.11) for k_b.

3.3 Results

3.3.1 XerCD-*dif* synaptic complexes

Three DNA substrates were produced, as described (Section 3.2.3 and Figure 3.3). They consisted of 4 kb DNA molecules, containing two *dif*-sites separated by 1 kb. They were tethered to poly-ethylene glycol passivated coverslips using biotin-neutravidin interactions. The 2.8 kb DNA tail contained an FtsK loading site at its distal end. Substrates were labelled, near to the *dif*-sites, with a donor (Cy3B, green) and an acceptor (Cy5, red) fluorophore (Figures 3.3, 3.6, 3.8 and 3.13). In addition to its role as a FRET acceptor, the red fluorophore was also used as the TFM reporter. It was separated from the surface by 1.2 kb of linear DNA; looping caused by synapsis formation would reduce this effective tether length, and so was apparent as a decreased fluorescence image width (FIW). We used a covalent trimer of the motor domain of FtsK, $FtsK_C$ [100], since monomers have been observed to aggregate [101].

Substrate design was informed by the crystal structure, 2HOI, of Cre-*loxP* [66] (Figure 3.2 and Table 3.1). If a synapsis formed that resembled the Cre-*loxP* structure with the XerC protomers poised to perform strand exchange (XerC*), then the interfluorophore distance would be 6.5 nm, giving a FRET efficiency of around 0.6 (given a Förster radius, R_0, for these dyes of around 7 nm [32]). It must be noted, however, that one distance is not sufficient to confirm a particular structure for any complex formed.

When Substrate I was imaged in the presence of XerCD, but in the absence of $FtsK_C$, we observed transient narrowing of FIW, accompanied by FRET ($E^* = 0.55$, Figure 3.6B). The narrowing of the FIW was consistent with formation and disassembly of XerCD-*dif* synaptic complexes that did not complete recombination, since FtsK is required for the activation of recombination. The lifetime of these synaptic complexes, τ, was 8.6 s (Figure 3.6B). Using hidden Markov modelling (HMM, Section 3.2.7) to look for transitions in FIW, it was determined that a small minority (< 9 %) of narrowing events were not

accompanied by FRET. We attribute these to non-specific sticking of DNA or proteins to the slide surface.

3.3.2 Activation of recombination

Substrate I was imaged in the presence of XerCD, FtsK$_C$, and ATP, and we observed events where the FIW remained narrow after FRET disappeared (Figure 3.6C). This is consistent with a productive recombination, and the lifetime of the FRET state in these complexes was shorter ($\tau = 4.0$ s) than that of non-productive synapses ($\tau = 8.6$ s). Not all synaptic complexes productively recombined in the presence of FtsK$_C$ (Figure 3.7).

Substrate II was designed to give FRET during the middle stages of the recombination reaction (Table 3.1 and, before HJ isomerisation Figure 3.3B). In addition, when the product of recombination dissociated, both fluorophores would be carried away from the surface, and disappear from view simultaneously (Figure 3.8A). Since the labelling position was different, the FRET efficiency was higher than in Substrate I ($E^* = 0.72$, Figure 3.8B), but the lifetime remained relatively unchanged ($\tau = 8.1$ s, Figure 3.9). This FRET efficiency corresponds to a distance of 5.2 ± 0.8 nm (Table 3.3), inconsistent with the 9.5 nm predicted for a synapsis formed with XerC poised to catalyse strand exchange (XerC*-Substrate, Table 3.1) in this substrate, suggesting that the intial state does not follow the Cre-*loxP* model.

In the presence of FtsK$_C$ and ATP, transitions from the initial synaptic complex ($E^* = 0.72$) to an intermediate that retained a narrow FIW, but had a reduced FRET efficiency ($E^* = 0.37$, Figure 3.8C) were apparent. This intermediate was short lived ($\tau = 1.3$ s, Figure 3.9), and was followed by a low FRET state ($E^* = 0.18$), which in turn was followed by the simultaneous loss of both fluorophores, indicating the dissociation of product after completed recombination. Using the Cre-*loxP* structure, we predict that a synapsis where XerD is active and poised to catalyse strand exchange (XerD*), and the HJ intermediate that follows it (XerD-HJ), would have an interfluorophore distance of

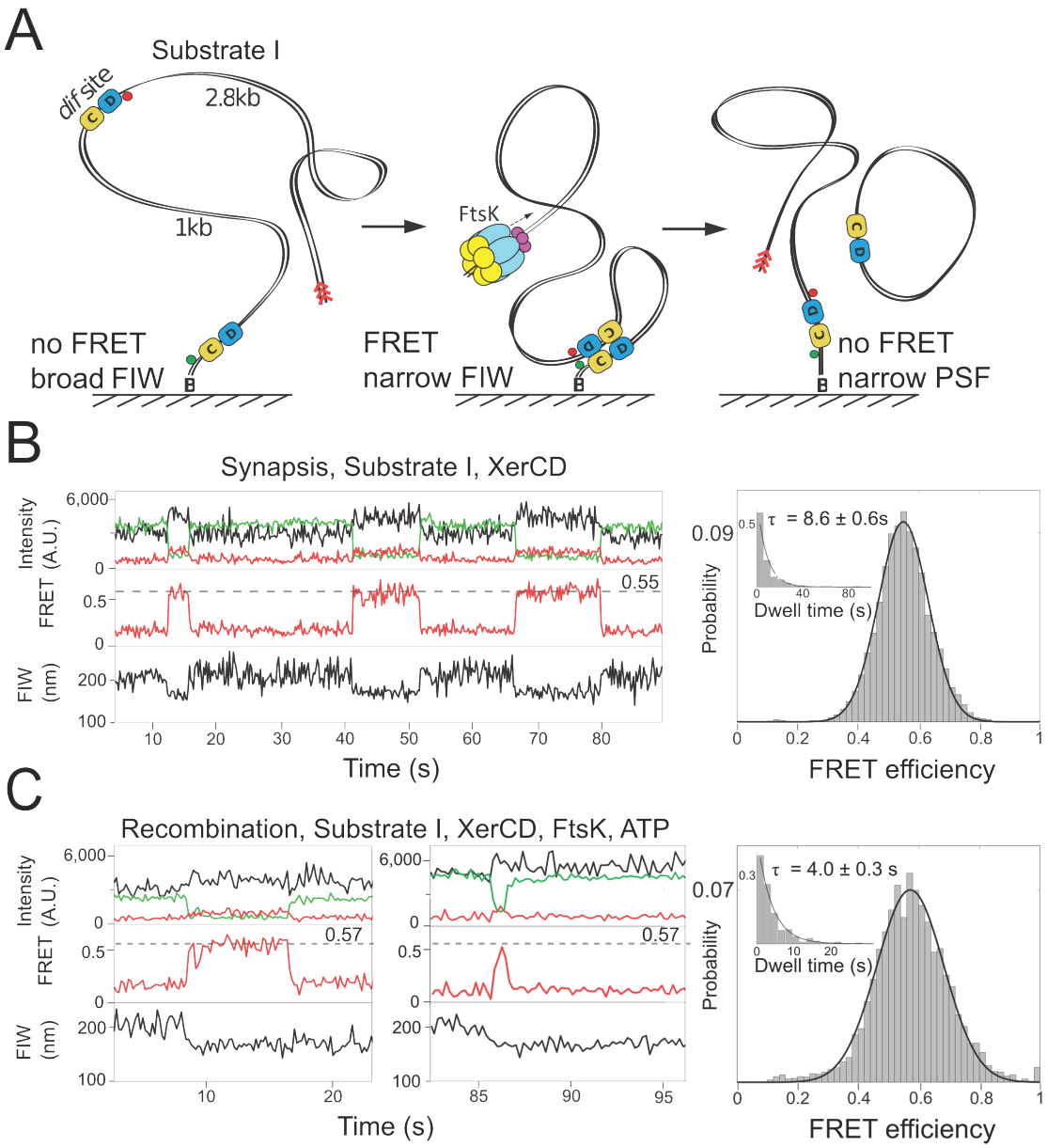

Figure 3.6: XerCD-*dif* synapsis and recombination. **(A)** Cartoon of the recombination reaction in Substrate I. Fluorophores (donor: green circle, and acceptor: red circle) were attached near both *dif*-sites. Synapsis and recombination can be monitored using the FIW and the FRET efficiency. FtsK orientating polar sequence (KOPS) indicated with red arrows. **(B)** Formation and disassembly of synaptic complexes. Top: donor intensity under donor excitation (DD, green), acceptor intensity under acceptor excitation (AA, black), and acceptor intensity under donor excitation (DA, red). Middle: FRET efficiency. Bottom: acceptor FIW. Non-productive events show a broadening of the FIW coincident with the end of FRET. Data acquired using 100 ms ALEX frames (unless otherwise stated). Right: FRET efficiency and synapse lifetime ($n = 380$). All dwell times fit with a single exponential distribution using maximum likelihood fitting in MATLAB. **(C)** Recombination. Recombination was apparent as a permanent reduction in FIW. Right: FRET efficiency and dwell times ($n = 249$).

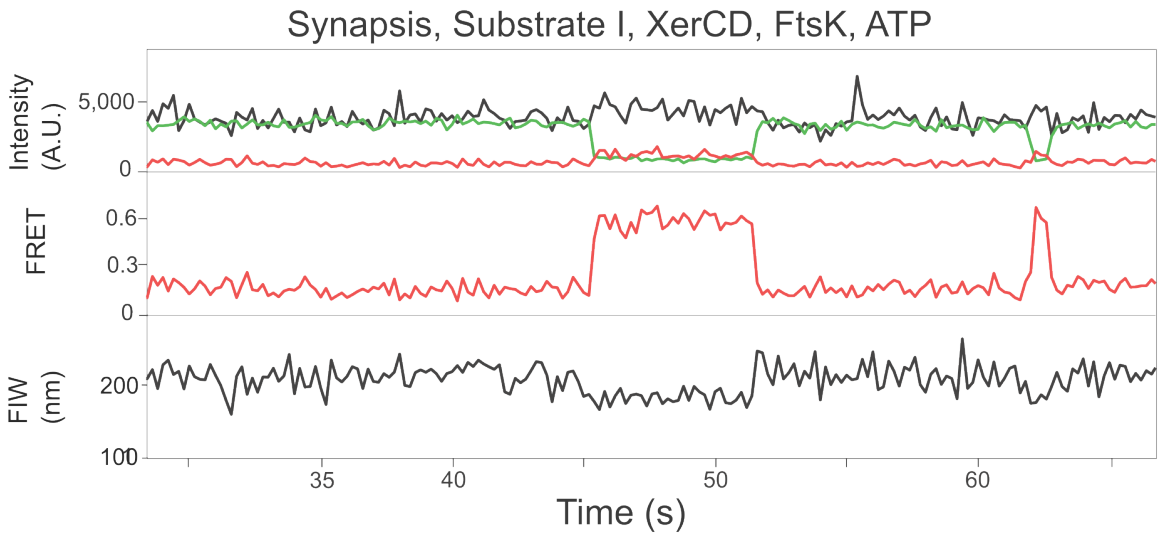

Figure 3.7: Reversible synapsis formation in the presence of $FtsK_C$ and ATP.

8.5 nm in this substrate (Table 3.1). Converting the FRET efficiency, $E^* = 0.37$, into a distance (Section 3.2.8), we calculated an interfluorophore distance of 7.4 ± 1.1 nm, consistent with the XerD*-Substrate Cre-*loxP* structure. Hence, we suggest that the $E^* = 0.37$ state is XerD*/XerD-HJ, and we label the state which preceeds it as the 'initial synapsis'. Since the predicted distance for the HJ following the isomerisation, and the product following a second round of strand exchanges by XerC (XerC-HJ and XerC-P, XerC*-Product in Table 3.1) is > 9 nm, we expect no FRET to be resolvable against the background ($E^* = 0.17$). Hence, the low FRET state in this substrate is consistent with the predicted XerC-HJ/XerC-P distance. Again, in the presence of $FtsK_C$, not all synaptic complexes were activated (Figure 3.10).

To confirm this assignment of states, we imaged Substrate II in the presence of mutants of XerC and XerD. Both mutants were incapable of catalysing their respective strand exchanges. Synaptic complexes in the presence of $XerC^{KQ}$ (K172Q) and XerD, and in the presence of XerC and $XerD^{KQ}$, formed with an initial apparent FRET indistinguishable from wild-type XerCD complexes ($E^* = 0.72$, Figure 3.11A and B). In the presence of $XerC^{KQ}$, XerD, $FtsK_C$ and ATP, we observed two rapidly interconverting FRET states (Figure 3.12A). Segmenting the timetraces using HMM, these states were deduced to

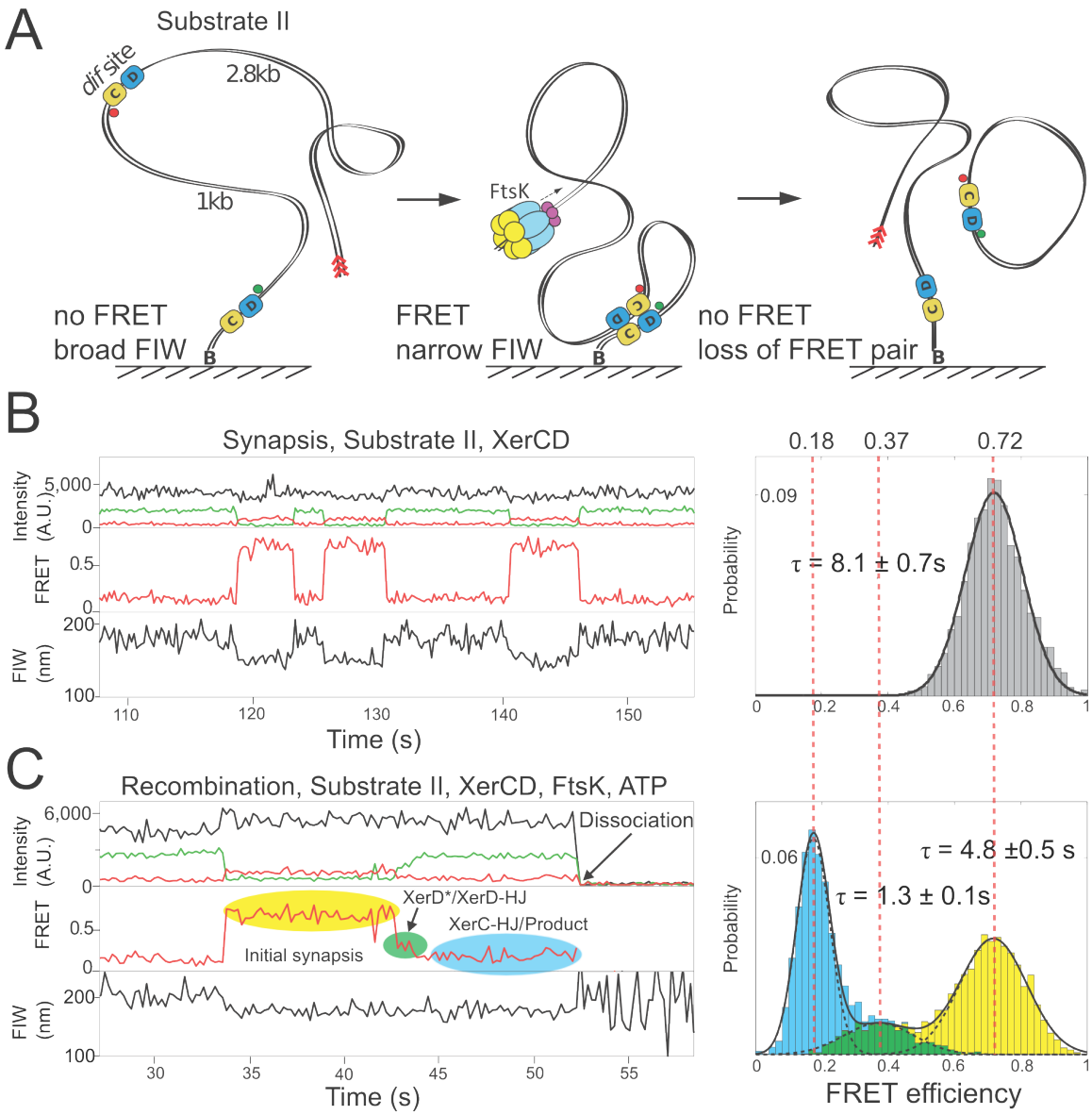

Figure 3.8: XerCD-*dif* recombination intermediates **(A)** Cartoon of the recombination reaction in Substrate II. This substrate monitors dissociation of the product of recombination: the circular DNA product produced carries both fluorophores away when it dissociates. **(B)** Formation and disassembly of synaptic complexes in Substrate II. Non-productive events show a broadening of the FIW coincident with the end of FRET. Right: FRET efficiency and synapse lifetime ($n = 382$). Dwell time, τ, recovered using maximum likelihood fitting of a single exponential distribution in MATLAB. **(C)** Recombination. Intermediates in the recombination pathway can be identified by their FRET efficiencies. Yellow: initial synaptic complex. Green: XerD*, and/or XerD-HJ. Cyan: XerC-HJ and XerC-P. Right panel: colour-coded FRET efficiency histograms and the corresponding dwell time ($n = 134$). Dwell times are for individual steps, not the entire reaction.

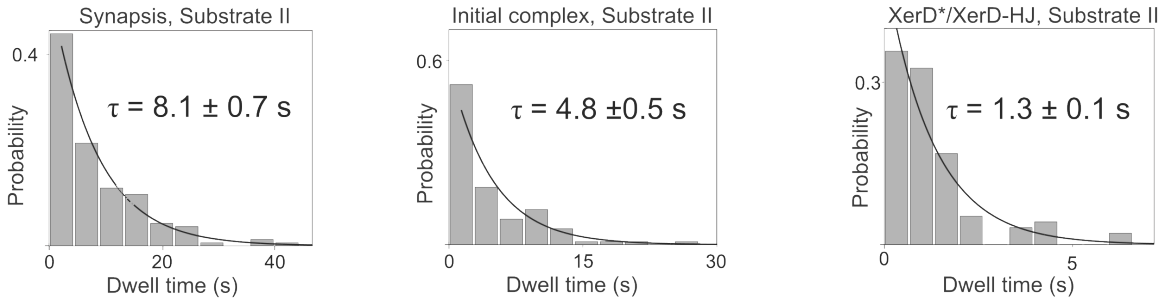

Figure 3.9: Substrate II transition rates. Left: dwell time for synapsis in Substrate II (Figure 3.8B). Middle: dwell time for initial synapsis when activated by $FtsK_C$ (Figure 3.8C). Right: dwell time for XerD*/XerD-HJ (Figure 3.8C). The dwell time, τ, was obtained using a one-parameter exponential fit using maximum likelihood estimation in MATLAB.

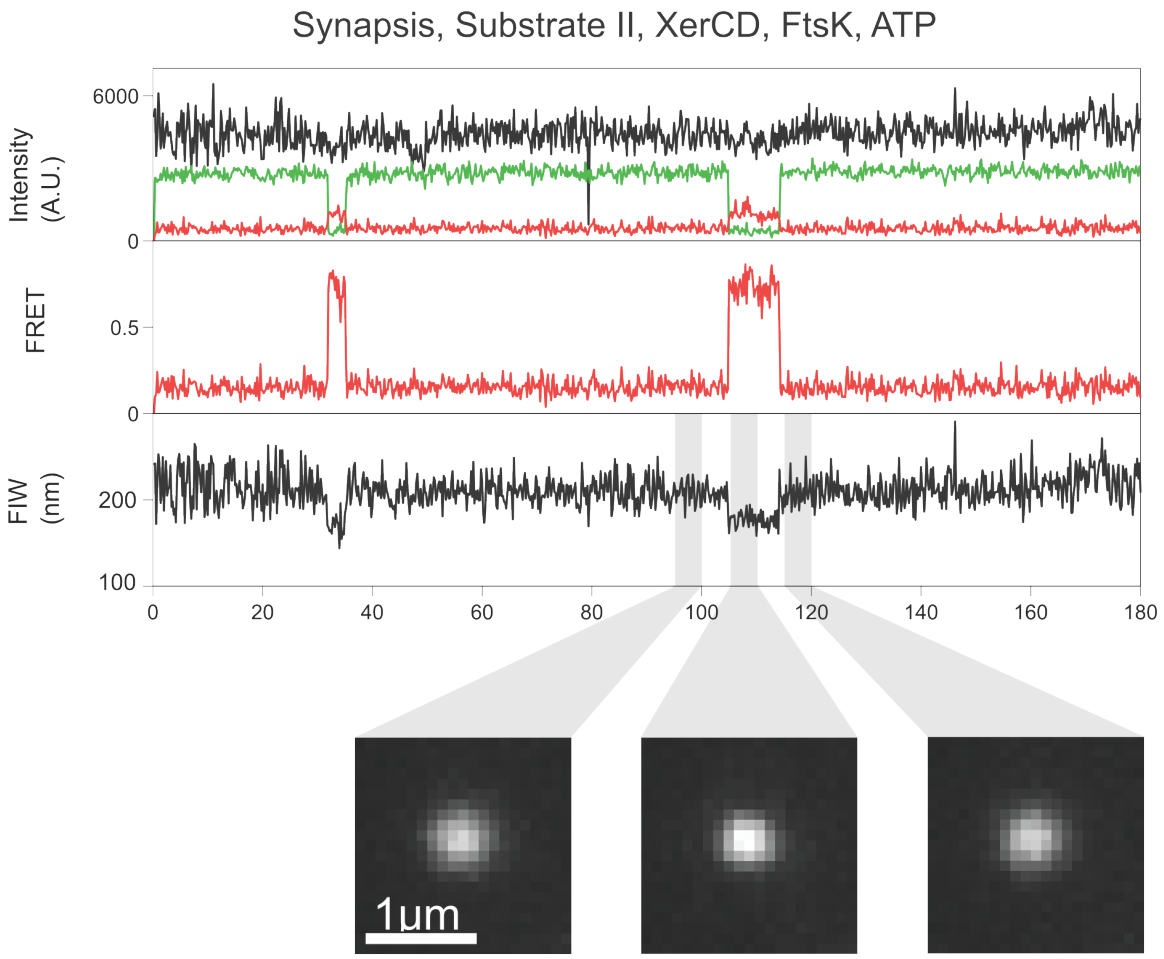

Figure 3.10: XerCD-*dif* synapsis in the presence of $FtsK_C$. Images of fluorophores made by averaging 25 consecutive frames (grey). During synapsis, the FIW is reduced, which is our TFM observable ($n > 100$).

be at $E^* = 0.37$ and $E^* = 0.22$ (Figure 3.4A). We interpret these interconversions as isomerisation between XerD-HJ and XerC-HJ, since XerC^{KQ} cannot resolve HJs to product in bulk assays [89]. Disassembly of the synapsis (judged by an increase in FIW) was always via the high FRET state ($E^* = 0.72$, Figure 3.11A, purple), suggesting that the initial synapsis is an obligatory intermediate in the recombination pathway, and that XerD is capable of HJ resolution (when accompanied by XerC^{KQ}).

When Substrate II was imaged in the presence of XerC, XerD^{KQ} (K172Q), FtsK$_C$ and ATP, we observed transitions from the intial synapsis ($E^* = 0.72$) to a lower FRET state ($E^* = 0.44$, Figure 3.12B). Since XerD^{KQ} cannot catalyse strand exchange, this state must correspond to XerD*. FtsK$_C$ activates the intial synapsis, but the reaction can only proceed as far as XerD*. The $E^* = 0.44$ assigned here to XerD* is significantly different to the $E^* = 0.37$ observed using wild-type XerCD in this substrate. We propose that $E^* = 0.37$ mostly represents XerD-HJ, with XerD* being transient in wild-type complexes. Converting these FRET efficiencies into distances (Section 3.2.8 and Table 3.3), we see that the interfluorophore distance differs by 0.5 ± 1.5 nm, with the uncertainty dominated by the uncertainty in κ^2, the orientation factor.

3.3.3 Proximity of FtsK$_C$ to the recombining complexes

Substrate III was designed (again, with reference to the Cre-*loxP* crystal structure) to give FRET in the final stages of the recombination pathway (Table 3.3). Dissociation of product would be apparent as the disappearance of the donor fluorophore (Figure 3.13A and Figure 3.3C). In the absence of FtsK$_C$, we observed transient narrowing of the FIW, with no associated FRET, consistent with assembly of the initial synapsis (Figure 3.13B).

When FtsK$_C$ and ATP were also present, in addition to non-productive synapsis, we observed complexes consistent with productive recombination (transition to a permanent narrow red FIW). These productive complexes showed transitions to $E^* = 0.56$, then transitions to $E^* = 0.40$, followed by product dissociation (Figure 3.13C). Occasionally,

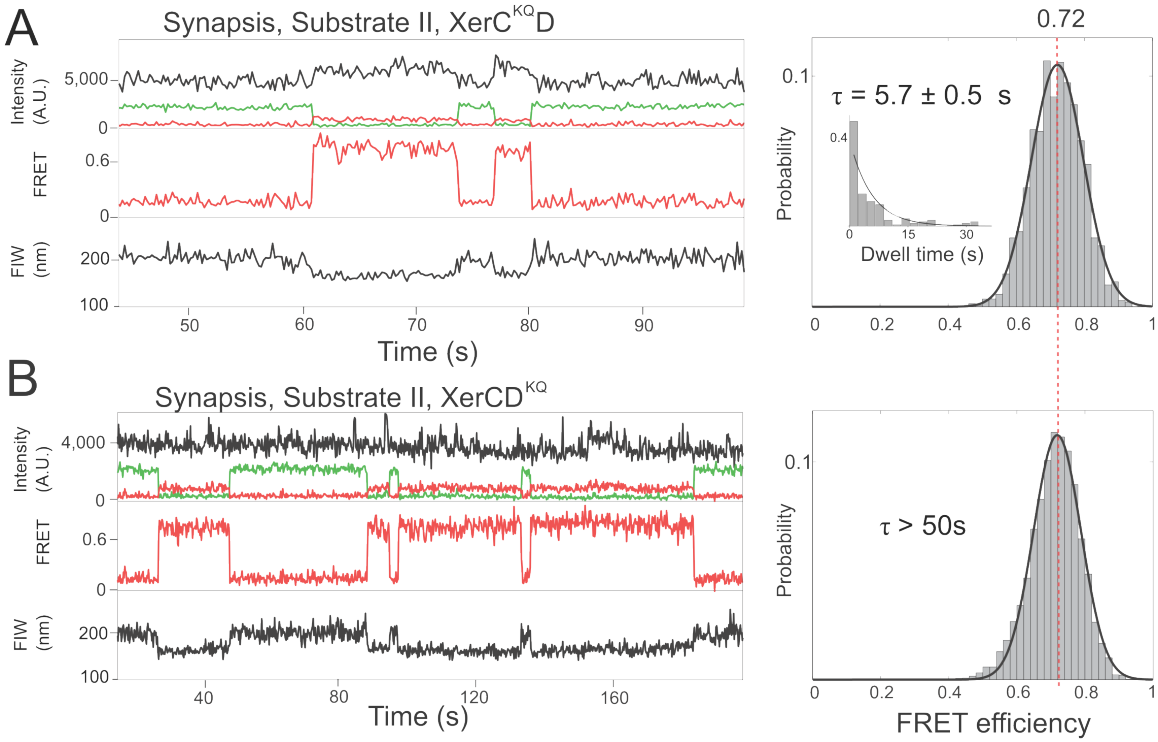

Figure 3.11: Synapses in the presence of catalytic mutants. Complexes formed are similar to those formed by wild-type XerCD. **(A)** XerC^{KQ} and XerD ($n > 100$). **(B)** XerC and XerD^{KQ} ($n > 100$).

transitions back to the $E^* = 0.56$ state from the $E^* = 0.40$ state were apparent (Figure 3.14A). Since XerC-HJ preceeds XerC-P in the recombination pathway, we assign $E^* = 0.56$ to XerC-HJ and $E^* = 0.40$ to XerC-P. This is confirmed using XerC^{KQ} in this substrate, where transitions to $E^* = 0.50$ are observed (Figure 3.15). Since XerC^{KQ} cannot catalyse 'top' strand exchange, this $E^* = 0.50$ state must correspond to XerC-HJ. The apparent FRET efficiency of this state is underestimated, since rapid transitions (within a frame) to XerD-HJ will be missed by the HMM and will bias the apparent FRET towards the FRET efficiency for XerD-HJ ($E^* = 0.20$), hence this XerC-HJ state in the mutant experiments corresponds the the $E^* = 0.56$ state in the wild-type experiments.

From this analysis, we conclude that the reverse transitions in the wild-type Substrate III experiments (Figure 3.14A) correspond to XerC-mediated catalysis from XerC-P to XerC-HJ. Using HMM, we can extract the dwell times (Figure 3.14B), and the FRET efficiencies for these (Figure 3.13C). Converting the FRET efficiencies for these states ($E^* = 0.56$

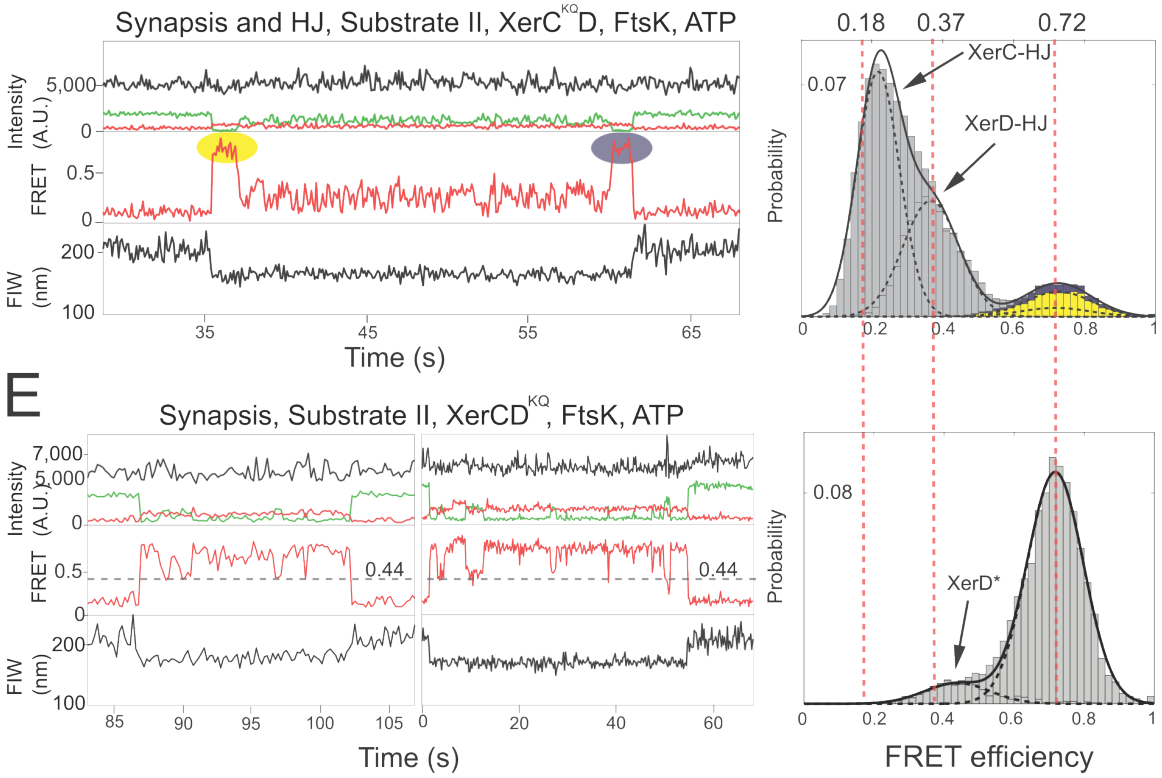

Figure 3.12: Transitions in the presence of catalytic mutants. **(A)** XerC^{KQ}. Data taken using 50 ms ALEX frame time. Since XerC^{KQ} is unable to perform catalysis, complexes are trapped at the HJ intermediate stage, and rapidly interconvert between XerC-HJ and XerD-HJ ($E^* = 0.22$ and $E^* = 0.37$). Even in the presence of XerC^{KQ}, the complexes assembled into the initial conformation (yellow highlight), and adopted the same conformation before dissociation (purple). Right panel: FRET efficiency and dwell time ($n = 84$). **(B)** XerD^{KQ}. Multiple reductions in FRET efficiency are apparent ($E^* = 0.44$), which suggests the transient formation of XerD*. Right panel: FRET efficiencies ($n = 66$).

and $E^* = 0.40$) into distances, we found that the interfluorophore distance changed by 0.8 ± 1.3 nm (Table 3.3), again, with uncertainty dominated by the uncertainty in κ^2. Together with the change in distance between XerD* and XerD-HJ, this demonstrates that structural rearrangement occurs after strand exchange by either XerC or XerD.

During recombinations in Substrate III, in 77 % of events, the transition from initial synaptic complex to XerC-HJ was accompanied by a PIFE of the Cy5 fluorophore (Figure 3.13C). PIFE indicates the close proximity ($<$2 nm) of a protein to a fluorophore [78], and we interpret this as evidence of the proximity of $FtsK_C$ to the recombining complex. PIFE began after the formation of the intial synapsis (Figure 3.16A), suggesting that FtsK acts on a pre-formed intial synapsis, rather than promoting de novo formation of an XerD* synapsis. The disappearance of PIFE was roughly coincident with the appearance of FRET marking the transition to XerC-HJ (Figure 3.16B), suggesting that FtsK may dissociate or otherwise disengage from the recombining complex. Observations of PIFE in Substrate I followed a similar sequence of events (Figure 3.17).

3.4 Discussion

In this chapter I have presented the combination of TFM with two established single-molecule fluorescence methods, FRET and PIFE, and described their use to follow the progress of XerCD-*dif* recombination. TFM has enabled us to watch the formation of synapses, and to distinguish their fate upon activation by FtsK. FRET has enabled us to measure distances within each step in the recombination pathway and compare these to predictions made using a model based on a related tyrosine recombination system, Cre-*loxP*. Finally, PIFE has enabled us to monitor the proximity of $FtsK_C$ to the recombining complex.

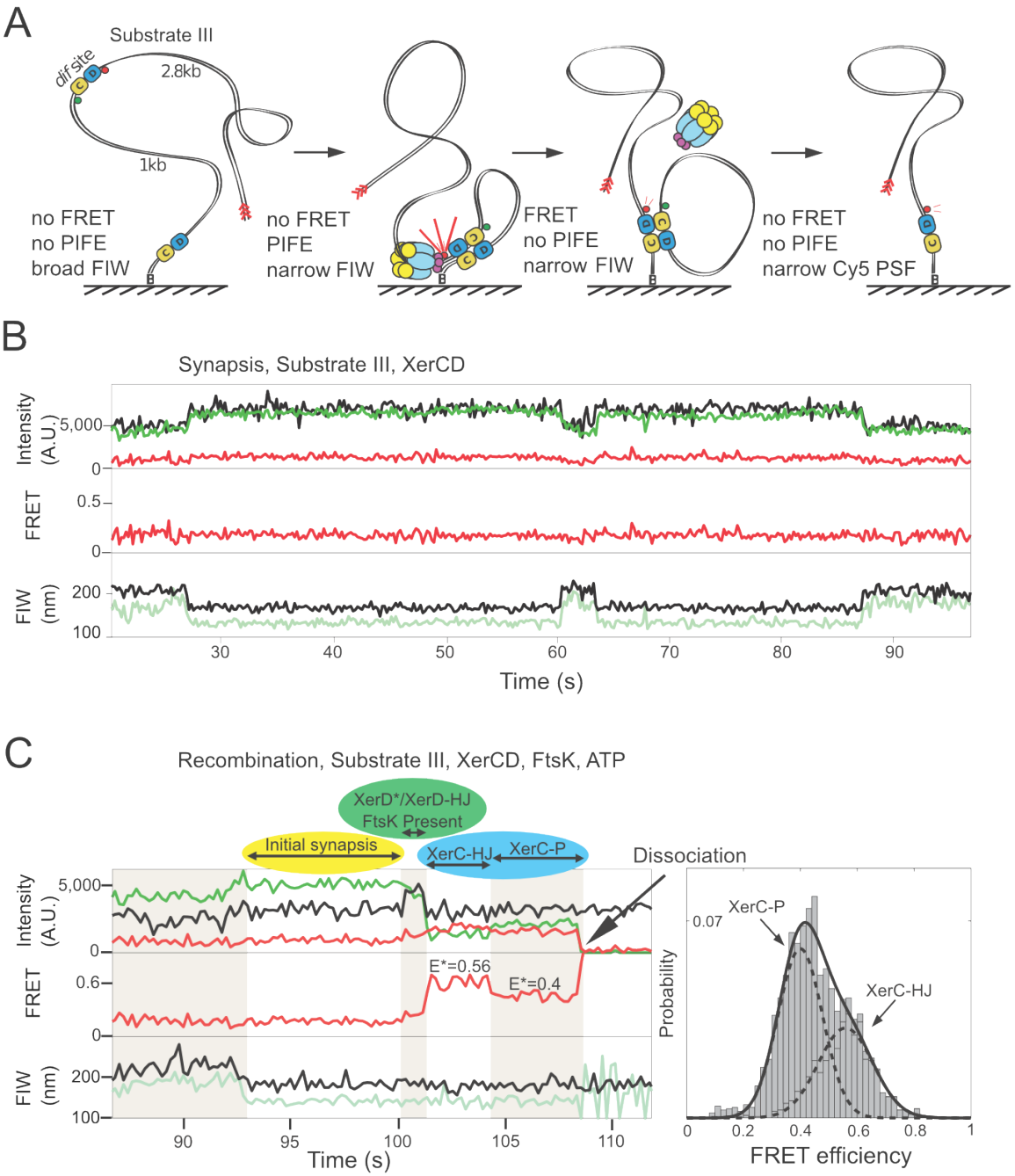

Figure 3.13: Late stages in recombination reaction monitored using Substrate III. **(A)** Cartoon of the recombination reaction in Substrate III. Both fluorophores flank the distal *dif*-site. Dissociation of product is apparent as a loss of green fluorescence. **(B)** Synapsis in Substrate III. Both fluorophores report on the same effective length along DNA. Simultaneous narrowing indicates synapsis formation. The fluorophores in Substrate III are positioned to give no FRET in the initial synaptic conformation ($n > 100$). **(C)** Recombination in Substrate III. Colour coding matches Figure 3.8. An increase in acceptor intensity at ∼100 s, due to PIFE, indicates the presence of $FtsK_C$ at XerCD-*dif*. Right panel: FRET efficiencies for XerC-HJ and XerC-P were extracted using HMM (Section 3.2.7) and fit with two separate Gaussians ($n = 111$ including timetraces with reverse transitions).

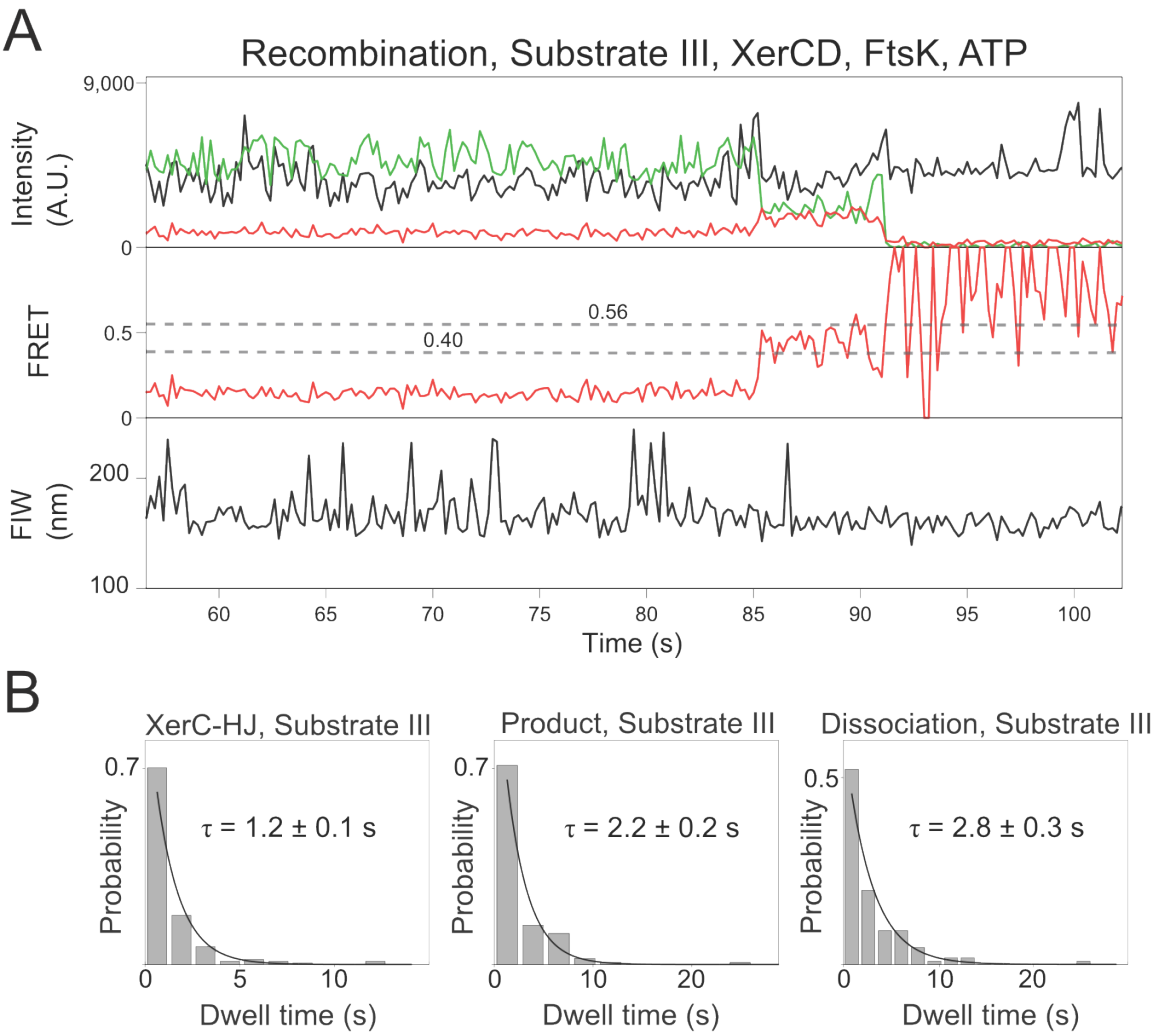

Figure 3.14: Transitions between XerC-HJ and XerC-P in Substrate III **(A)** Example time trace. The time trace shows transitions between XerC-HJ ($E^* = 0.56$) and XerC-P ($E^* = 0.40$) before dissociation (monitored by the loss of green fluorescence). To quantify the rate of this reverse transition (XerC-P to XerC-HJ), timetraces were segmented using HMM. **(B)** Dwell times in Substrate III. Left: dwell times for XerC-HJ converting to XerC-P. Middle: dwell times for XerC-P transitioning to either product or back to XerC-HJ. Right: dwell times for XerC-P dissociating. The rate of conversion of XerC-P to XerC-HJ was estimated as the difference between the rate of transition out of XerC-P (to either XerC-HJ or dissociation) and the rate of dissociation of XerC-P (Section 3.2.9).

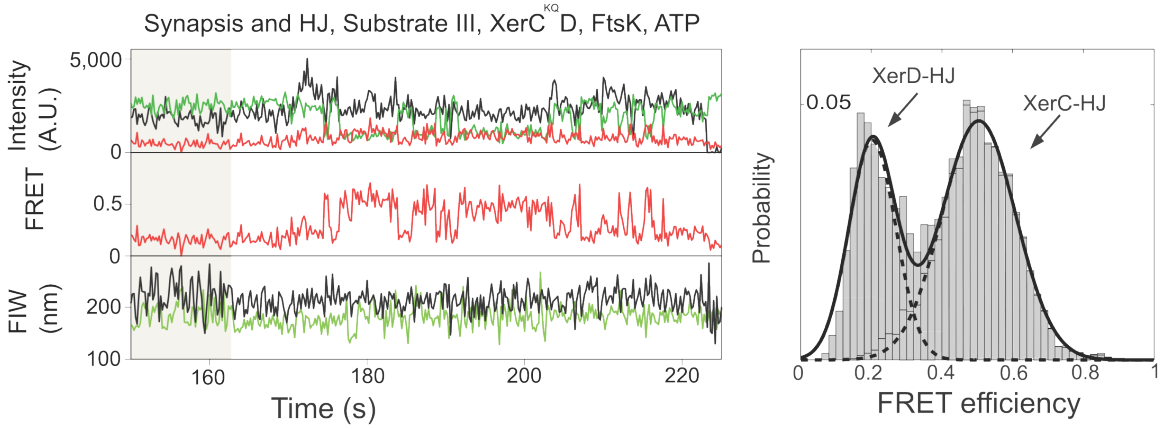

Figure 3.15: HJ isomerisation in Substrate III in the presence of $XerC^{KQ}$, XerD, $FtsK_C$ and ATP. Synapsis formation is apparent, as a reduction in the donor and acceptor FIW, at ~162 s (end of grey shading). Transitions between XerD-HJ ($E^* = 0.20$), and XerC-HJ ($E^* = 0.50$) are apparent. The apparent FRET efficiency for XerC-HJ here ($E^* = 0.50$) is lower than in the wild-type complexes ($E^* = 0.56$), since rapid transitions to XerD-HJ and back in a single frame reduce the apparent FRET efficiency ($n = 26$).

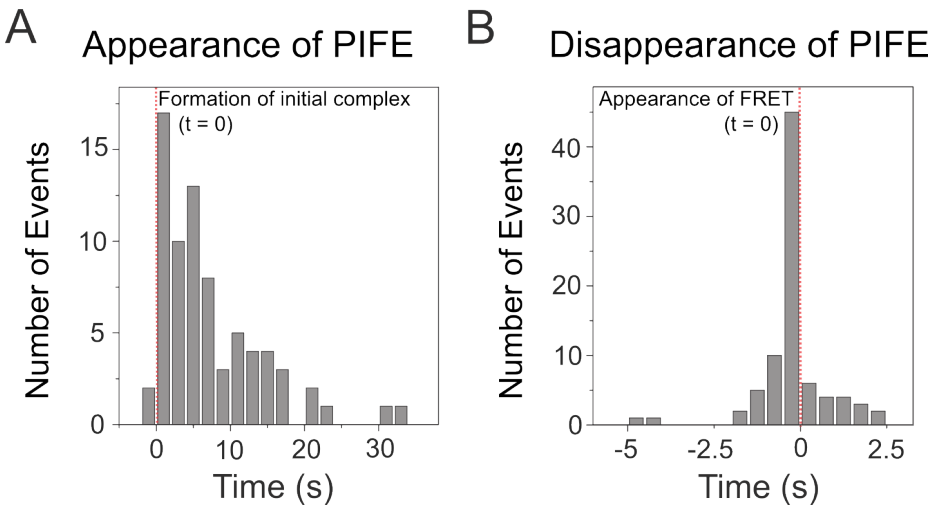

Figure 3.16: Timings of PIFE in Substrate III. **(A)** Histogram of the timing of the appearance of PIFE, relative to the formation of the initial synapsis. PIFE can be used to monitor the proximity of $FtsK_C$ to the XerCD-*dif* complex. The distribution implies that $FtsK_C$ arrives after the formation of the initial complex, indicating that the initial synapsis is the substrate for FtsK activation. **(B)** Histogram of the timing of the disappearance of PIFE, relative to the isomerisation of XerD-HJ to XerC-HJ that appears as a transition from background to $E^* = 0.56$.

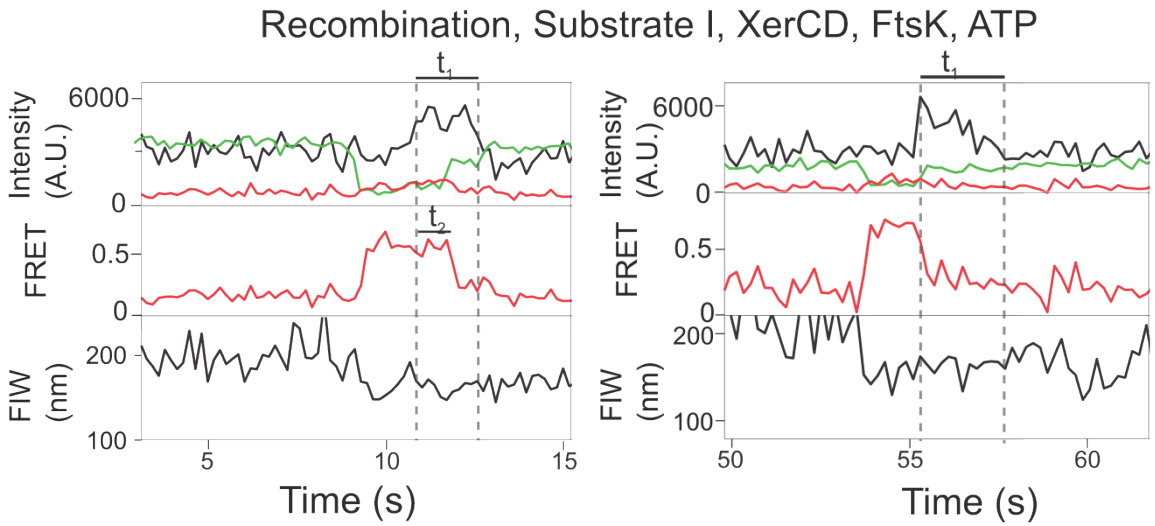

Figure 3.17: PIFE in Substrate I. Time t_1, the duration of PIFE, corresponds to $FtsK_C$'s proximity to the synaptic complex. Time, t_2, is the time between the start of $FtsK_C$ proximity and the conformational switch from initial synapsis to XerD*. Around 50 % of recombination events in Substrate I were accompanied by PIFE.

3.4.1 Recombination pathway

The results presented in the previous section lead us to propose the following model of XerCD-*dif* recombination: spontaneously formed initial synapses are acted upon directly by FtsK, prompting a switch to a conformation where XerD is poised to perform strand exchange (XerD*); strand exchange by XerD forms the first HJ intermediate (XerD-HJ), which isomerises to a conformation in which XerC is poised to resolve this HJ (XerC-HJ); after this second strand exchange, the recombinant product (XerC-P) is formed, and the synaptic complex dissociates (Figure 3.18).

3.4.2 Structural comparison to Cre-*loxP*

With the exception of the intial synapsis, the architecture of the recombination pathway is broadly similar to the Cre-*loxP* model (Tables 3.1 and 3.3). The initial synapse matches the XerC*-Substrate distances predicted in Substrate I (predicted: 6.5 nm; measured: 5.9 ± 0.9 nm) and Substrate III (predicted: 10.7 nm; measured: >9 nm), but does not match in Substrate II (predicted: 9.5 nm; measured: 5.2 ± 0.8 nm). This disagreement

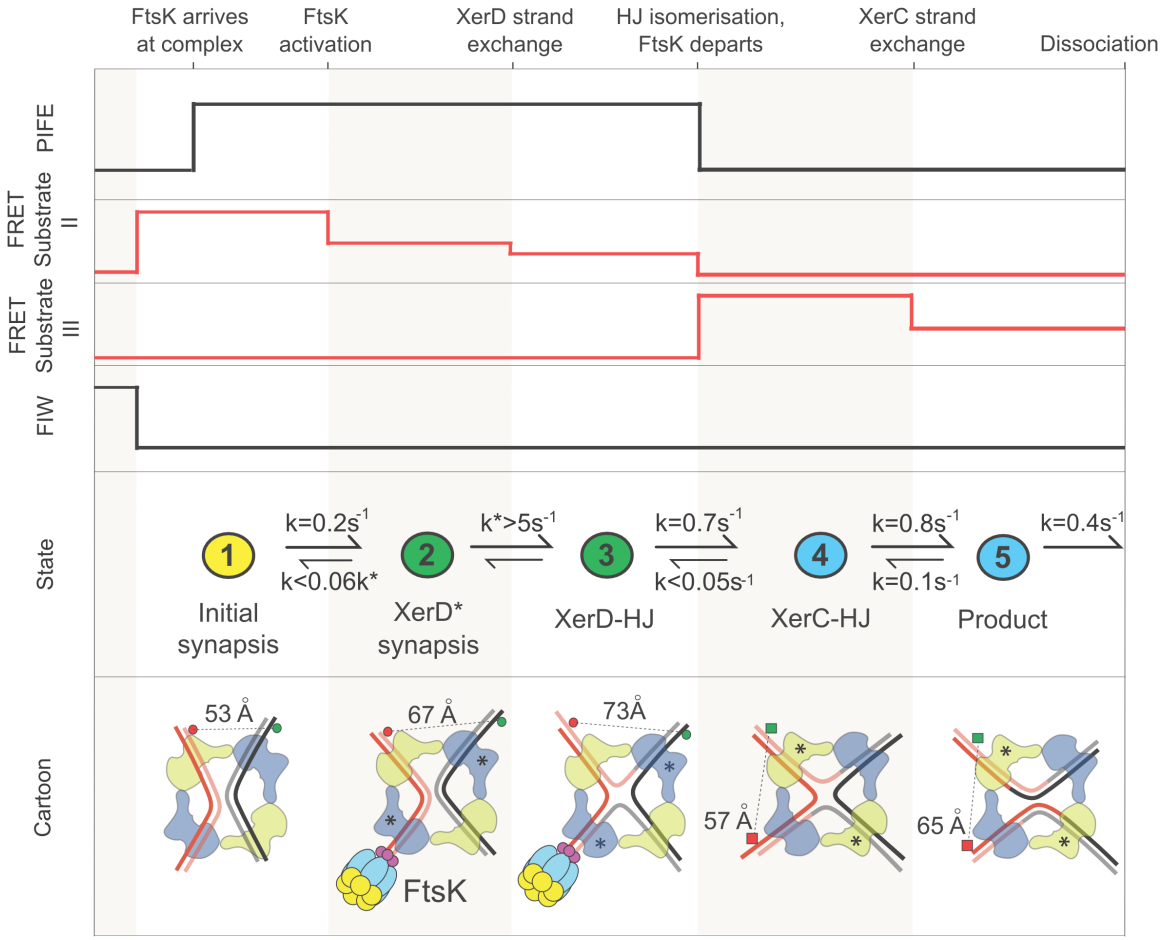

Figure 3.18: XerCD-*dif*-FtsK recombination pathway. Top and middle: observables in Substrate II and Substrate III, and rate constants deduced from FRET transitions. Bottom: proposed five-state pathway of recombination, with distances inferred from FRET.

is strong; even if the fluorophores were optimally aligned ($\kappa^2 = 4$), they would only be a distance of 7.5 ± 0.3 nm apart (with a reduced uncertainty because we have fixed κ^2), still too close to match the predictions using Cre-*loxP*.

The requirement for FtsK activation distinguishes XerCD-*dif* from Cre-*loxP* recombination. The observation that the initial synapsis is the substrate for activation argues against models where FtsK promotes XerD* synapsis formation, or remodels a complex that was initially poised for XerC catalysis [90]. Since the initial synapsis does not resemble a structure with XerC poised for catalysis, we suggest that XerC-active (XerC*) synapses form separately. The action of such XerC* synapses has been inferred from the ability of XerC to form and resolve HJs in vivo and in vitro [88, 90, 102]. Such XerC* complexes have not been observed in our assay, and may form too infrequently or too transiently for us to resolve. However, since reverse transitions were apparent from XerC-P to XerC-HJ in Substrate III we have evidence of the ability of XerC to form HJs ($k = 0.1\,\mathrm{s}^{-1}$). Given our observations, we predict that any XerC-first pathway is unlikely to be significant in vivo.

We note that in the Cre-*loxP* structure, the synaptic complex architecture is considered almost identical to the HJ intermediate that follows it in the pathway [85, 86, 87]. In XerCD-*dif*, our measurements support the presence of a small rearrangement between active synapses and their HJ intermediates (Figure 3.18), perhaps influencing the directionality of the reaction.

3.4.3 FtsK activates recombination

The PIFE observed in Substrate I and Substrate II could indicate the proximity of $\mathrm{FtsK_C}$, with the end of PIFE indicating a rearrangement of the synaptic complex or the dissociation of $\mathrm{FtsK_C}$. Since FtsK activates the conformational switch from initial synapsis to XerD*, and multiple rounds of this activation are apparent using Substrate II and $\mathrm{XerCD^{KQ}}$, FtsK potentially remains in the vicinity of the synaptic complex after it activates recombination. This may be consistent with the ability of XerCD-*dif* to perform the multiple rounds of

recombination expected to occur during decatenation [48].

3.4.4 Chromosome maintenance

Once replicated, *dif*-sites in vivo are likely to be recombinase bound and able to stably ($\tau = 8.6$ s in Substrate I) synapse without FtsK. This may serve as a mechanism to keep sister *dif*-sites close together, ready for FtsK activation, during potentially multiple rounds of recombinations to facilitate chromosome unlinking and decatenation. The spatial and temporal organisation of FtsK will limit the activity of XerCD-*dif*-FtsK to the mid-cell at late stages of the cell-cycle, safeguarding against inappropriate recombinations.

3.5 Conclusions

Here, I comment on the validity and applicability of the conclusions of my work.

3.5.1 Assignment of FRET states

The ability to distinguish of XerD*/XerD-HJ in Substrate II against the background due to cross-talks (Figure 3.8C), relies on the knowledge, from XerCD^{KQ} experiments (Figure 3.12B), that XerD* has approximately $E^* = 0.4$. Since the transition is fast and the background FRET distribution has a mean of $E^* = 0.20$ and a SD of 0.06, the segmentation of the transition from XerD*/XerD-HJ to XerC-HJ is less reliable than other segmentations. The conclusion that XerD* and XerD-HJ differ in apparent FRET, and that XerD* rapidly converts to XerD-HJ, is weakened by this segmentation uncertainty and the use of XerD^{KQ} to deduce that XerD* has $E^* = 0.44$ (since mutant proteins may have altered conformations).

Comparing $E^* = 0.56$ for XerC-HJ in wild-type complexes (Figure 3.13C) to $E^* = 0.50$ in complexes formed by XerC^{KQ}D (Figure 3.15), there is a difference in FRET efficiencies. Since XerD-HJ isomerisation to XerC-HJ happens rapidly (the lifetime of XerD*/XerD-HJ

in Substrate II was 1.3 ± 0.1 s in Figure 3.9), we attributed this discrepancy to rapid transitions to XerC-HJ within a single frame in the XerC^{KQ}D complexes that reduce the average FRET efficiency for the state assigned by HMM to XerC-HJ. This assignment of $E^* = 0.50$ informs the assignment of XerC-HJ to $E^* = 0.56$ and XerC-P to $E^* = 0.40$ in wild-type complexes. However, we are assuming that the use of mutant protein does not significantly alter the conformation of the nucleoprotein complexes.

Any misassignment of states later in the recombination pathway does not alter the conclusion that the FtsK dependent recombination pathway follows from the same synaptic structure that forms in the absence of FtsK, and that this initial synapsis does not resemble the structure of a synapsis ready for XerC-mediated strand exchange predicted using the Cre-*loxP* structure.

3.5.2 PIFE

The use of PIFE was not planned from the substrate design, however intensity changes consistent with it were was observed in our data. PIFE is a plausible explanation for our observations. It doesn't occur in all time traces suggesting that there might be an FtsK arrival orientation dependence. Since it is only sensitive to distance changes in the 1 nm to 2 nm range, we can't be sure that its disappearance really reflects the dissociation of FtsK from the synaptic complex.

3.5.3 Kinetics of recombination

Transitions were segmented manually and with HMM, introducing an additional uncertainty in any rate parameter. If transitions could be segmented to within a single frame, this introduces a 0.2 s additional uncertainty (reduced when averaging over hundreds of segmentations). This will slightly increase the uncertainty in quoted transition rates, but the uncertainty will still be dominated by the sampling error (0.1 s to 0.7 s) included in the quoted fit parameters. Despite this, since experiments were performed in vitro, caution

must be exercised when extrapolating to in vivo rates. The rates deduced provide a guide for the timescale of the reaction, but should not be considered definitive.

3.5.4 Accurate FRET

Following the procedure outlined in Section 3.2.8, FRET efficiencies were converted into distances. However, several assumptions are implicit in this conversion. We assume that measured FRET efficiencies correspond to particular individual structural states, rather than representing the mean of some rapidly (compared to the camera integration time) interconverting ensemble. Some artefacts will be introduced by manual and HMM segmentation of time traces, but these can be minimised by the discarding of frames either side of a transition as described in Section 3.2.7.

We assume that the local environment of the fluorophores measured in a different nucleoprotein conformation on a fluorometer is a good approximation to the local environment during the recombination reaction. We can confirm that gamma is consistent across states and substrates, but this gamma is measured from a transition (i.e. we could have different quantum yields before and after the transition, but are blind to it with the method used) and is a simple mean value with an uncertainty, rather than a distribution. Apparent FRET efficiencies, E^* , were recovered by fitting Gaussian distributions to all frames assigned to each state. The true distribution of FRET efficiencies may be better approximated by a beta-distribution. There is also a standard error in mean associated with quoting a mean apparent FRET for a state from a finite number of observations (which has been taken into account in our analysis).

The efficiency of energy transfer between fluorophores depends on their relative orientations, and the majority (90 %) of the uncertainty in distance in Table 3.3 can be attributed to uncertainty in the orientation factor, κ^2. One model of fluorophore motion has the fluorophore free to reorientate within a fixed cone [103]. If we assume that fluorophores can be considered free to reorientate during the excitation lifetime of the donor, then we

can use an average value for the orientation factor. Sensible limits on its value can be obtained from anisotropy measurements of singly labelled constructs containing donor or acceptor fluorophores [97, 103, 104] (Section 3.2.8). However, we are assuming that we are in the dynamic averaging regime. If we are in the static averaging regime, where dyes do not reorientate within their excitation lifetime, then we cannot use a mean value for κ^2 and must be careful of coupling between orientation and inter-dye distance [105]. Shorter linkers between biomolecule and dye seem to favour the static averaging regime, whereas in this work we use a relatively long, six-carbon, linker [105].

The effect of all these factors is to make it difficult to distinguish between structurally similar states with only a single FRET measurement. However, in this work we have three distances (one from each substrate) for each state in the recombination pathway and observe good agreement between the Cre-*loxP* model and our data across substrates and states for the final four states in the recombination pathway, strongly supporting a conserved structural basis of recombination between Cre-*loxP* and XerCD-*dif*. The new structural state, which acts as the substrate for FtsK activation, does not resemble a XerC* structure predicted using the Cre-*loxP* crystal structure, with a strong disagreement in Substrate II supported by the need to completely rearrange the nucleoprotein complex to move from a synapse with XerC ready for strand exchange to one with XerD ready.

There is a small difference in FRET values assigned between XerC-HJ and XerC-P, and between XerD* and XerD-HJ. This can be attributed to a small change in the distance between dyes after strand exchange, however, it may reflect a change in the relative orientations of the dye pair, or a change in local environment of at least one of the dyes. All of these factors reflect a general change in the nucleoprotein conformation. The distance in XerD* has been measured in a synapsis formed by XerCD^{KQ}, and hence we must be aware of the caveat that the structure may have been altered by the mutant protein.

3.6 Contributions

The work in this chapter has formed the basis of a publication [2]. Figures and text have been adapted from this publication, and were produced (for the publication) in collaboration with Pawel Zawadzki. The project was supervised, and the text in the publication edited, by David Sherratt, Lidia Arciszewska and Achillefs Kapanidis. They also provided much insight into the biological implications of our results. Pawel Zawadzki and Rachel Baker purified proteins and performed mutagenesis. Together with Pawel Zawadzki and Justin Pinkney, I performed TIRF experiments and produced the DNA substrates. Along with Pawel Zawadzki, I segmented and extracted FRET events from timetraces, and we both interpreted the data.

I was responsible for the work utilising accurate FRET to extract distances, and for the modelling of distances within the Cre-*loxP* structure. I was responsible for the interpretation and fitting of dwell times and for placing limits on reverse rates in the pathway.

Chapter 4

FtsK assembly, translocation and interactions with XerCD

4.1 Introduction

The previous chapter has described the pathway of XerCD-*dif* recombination, from the formation of the initial synapsis, its activation, and the formation and resolution of the Holliday junction intermediate, to the dissociation of the product. However, that work, using a single tethered fluorophore motion (TFM) reporter, focussed on the reaction from the point of view of the synapsis, and only inferred the proximity of FtsK indirectly, using protein induced fluoresence enhancement. The work in this chapter expands TFM in combination with Förster resonance energy transfer (FRET), using two spectrally distinct TFM reporters (one on $FtsK_C$ and one on DNA), to directly observe $FtsK_C$ as it assembles and translocates, and to correlate its behaviour, upon arrival at XerCD, with the progress of the recombination reaction.

4.1.1 FtsK

FtsK is a 1329 amino acid (aa) DNA translocase, which assembles at the division septum and functions in segregating sister chromosomes, in the late stages of the cell cycle, in a wide range of bacteria [49, 52, 53, 57], by activating site-specific recombination by XerCD at *dif* [90]. The two related tyrosine recombinases, XerC and XerD, bind the 28 bp *dif*-site located within the terminus region, *ter*, on the *Escherichia coli* chromosome. They unlink catenated chromosomes and resolve chromosome dimers formed by homologous recombination [45, 47, 48, 64, 106]. Independently of its role in activating XerCD-*dif* recombination, FtsK appears to play a direct role in the segregation of *ter* [57]. FtsK consists of three domains: an essential 179 aa N-terminal domain that anchors it to the division septum; a ~500 aa C-terminal motor domain; and a ~650 aa linker domain [52, 53, 54, 55]. The motor domain, $FtsK_C$, is composed of α, β and γ subdomains [56]. The α and β subdomains form a double-stranded DNA (dsDNA) translocase, belonging to the RecA family of ATPases [56]. The γ subdomain plays a regulatory role in the recognition of the FtsK orientating polar sequence (KOPS), that guides FtsK translocation towards the *dif*-site at *ter* [58, 59, 60], and in the activation of XerCD-*dif* recombination [90, 106, 107]. Activation of recombination requires direct interaction between the γ subdomain and XerD [61, 62].

4.1.2 Previous work

Previous studies of $FtsK_C$ assembly and translocation have utilised biochemical methods [108, 109] and single-molecule techniques, including magnetic and optical tweezers [58, 101, 110, 111], tethered particle motion (TPM) [58] and, more recently, DNA curtains [112, 113]. Optical/magnetic tweezers and TPM experiments have relied on loop extrusion by $FtsK_C$ to observe its action (looping by FtsK shortens the length between two DNA ends, hence displacing the bead used in TPM or optical/magnetic tweezers). Many of the single molecule assays involved the attachment of $FtsK_C$ to quantum dot (QD) labels

or used derivatives that were known to aggregate, thereby potentially confounding the interpretation of data, because multiple motors could be present in the region of analysis. DNA curtain experiments have revealed that $FtsK_C$ can push, evict and bypass proteins bound to DNA as it translocates [113]. However, $FtsK_C$ stops at least transiently and/or dissociates at XerCD bound to *dif* [108, 113]. Reversals in translocation direction have been observed to occur spontaneously [58, 101, 111, 112] and in response to XerCD-bound to *dif* [113]. The use of a fluorophore label, along with singly tethered DNA have allowed us to observe $FtsK_C$ without requiring any loop extrusion, and to observe its interaction with synaptic complexes of XerCD, where previous single-molecule work has only dealt with un-synapsed XerCD-*dif* [113]. Using this approach, we have determined that $FtsK_C$ assembles on DNA as a single hexamer, and begins translocating rapidly ($\sim$0.25 s), without extruding a loop of DNA. When it reached XerCD bound to *dif*, either in a synapsed or un-synapsed conformation, it resided briefly for 0.5 s to 1 s and then dissociated without any evidence of reversal. $FtsK_C$ activated recombination when it met synapsed XerCD-*dif* complexes, and then dissociated faster than the completion of recombination by XerCD.

4.2 Methods

We adapted TFM-FRET to follow the assembly, translocation and the behavior of $FtsK_C$ when it interacts with XerCD. We made a DNA substrate carrying two *dif*-sites separated by a 1 kb spacer (Figure 4.1A); the substrate was attached to the slide surface through a biotinylated 5' end, which was located 200 bp away from the surface-proximal *dif*-site; the second, surface-distal *dif*-site was flanked by a 2.8 kb DNA segment containing triple KOPS near its end. The KOPS was orientated towards the surface (Figure 4.1A). The DNA was labelled with Cy3B 1 bp away, in the direction of the free-end, from the XerD binding sequence of the surface-distal *dif*-site. TFM uses the width of the image of this single fluorophore to measure an effective distance along DNA; the fluorescence image

width (FIW) decreases as the effective tether length decreases. Synapsis formation between the two *dif*-sites, mediated by XerCD, was evident as a reduction in the green (Cy3B) FIW, and recombination was evident as a permanent reduction in green FIW (Figure 4.1A).

We used a covalent trimer of $FtsK_C$ [100], singly labelled with Cy5 at a surface cysteine introduced in the motor β subdomain of the middle monomer (Cy5-FtsK) (Section 4.2.2). The trimer concentration, 5 nM, was close to the ~40 nM in vivo monomer concentration [114]. The binding of $FtsK_C$ was apparent as an appearance of signal in the red (Cy5) channel and the position and behavior of the $FtsK_C$ was monitored via the red FIW (Figure 4.1B). To convert between red FIW and the position of Cy5-FtsK along DNA, we imaged double-stranded DNA (dsDNA), with lengths from 87 bp to 4000 bp, using our total internal reflection fluorescence (TIRF) microscope and standard conditions of 25 ms alternating laser excitation (ALEX) [37] (Figure 4.1C).

4.2.1 DNA preparation

The 4 kb DNA substrate (Figure 4.2) was prepared using polymerase chain reaction (PCR) with one fluorescently labelled oligonucleotide and a plasmid template containing directly repeated *dif*-sites separated by a 1 kb KmR gene cassette (pRB10) and Phusion High-Fidelity DNA polymerase (NEB, Ipswich, USA). Oligonucleotides were synthesised and high performance liquid chromatography purified by ATDBio Ltd (Southampton, UK). Cy3B labelling was performed as previously described [115]. After PCR, the product was digested with NcoI and ligated to a 200 bp fragment, produced by a separate PCR, which contained a 5' biotin moiety. A 2.8 kb tail was ligated, following a SalI digestion, and the 4 kb DNA produced was gel-purified.

4.2.2 Protein preparation

XerC, XerD and unlabelled $FtsK_C$ were purified according to established procedures [56, 92]. The mutant trimer of $FtsK_C$ was prepared using QuikChange site-directed mutagenesis kit

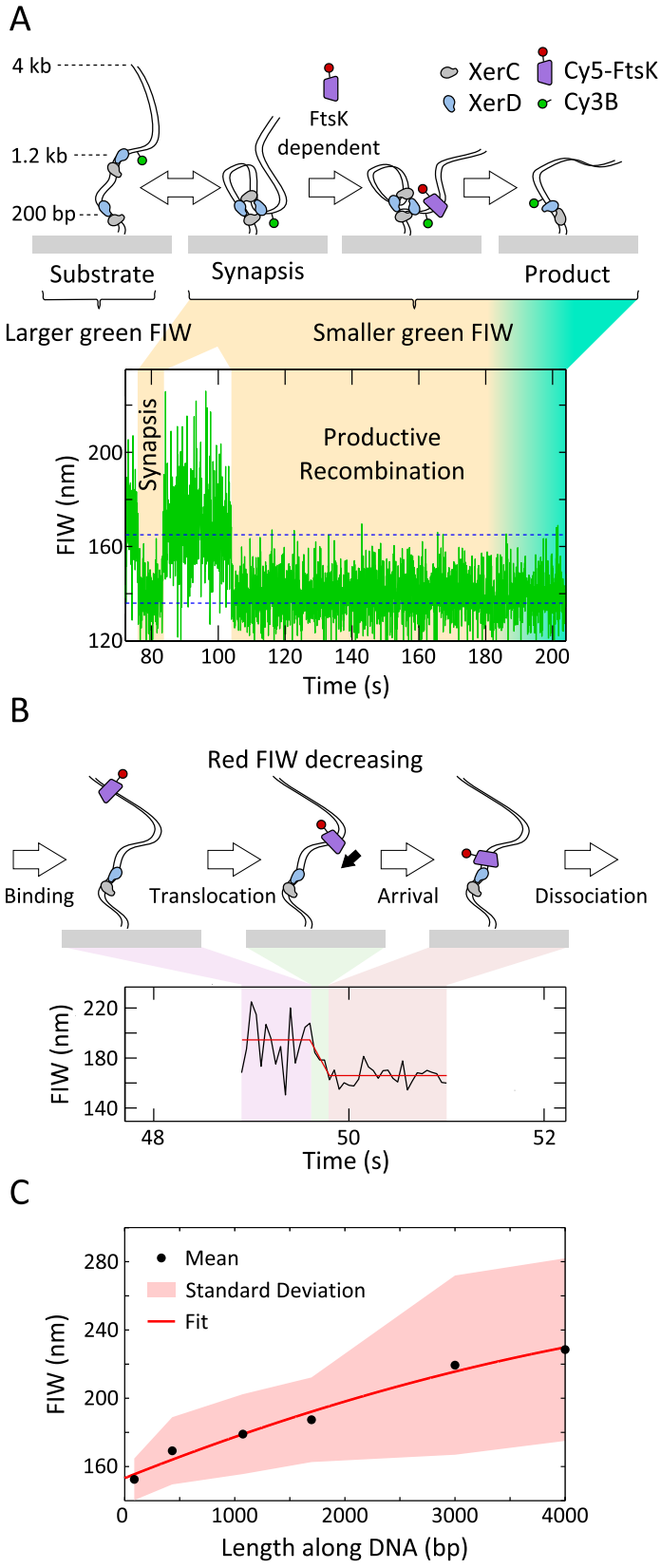

Figure 4.1: Experimental design. **(A)** Recombination is monitored using Cy3B. Synapsis transiently reduces FIW and recombination permanently reduces FIW. Blue dashed lines: population average FIWs for the un-synapsed and synapsed/product DNA (165 nm and 136 nm). **(B)** Cy5 monitors $FtsK_C$ behaviour. Translocation decreases or increases FIW. Red line: least-squares fit, used to segment translocation. **(C)** FIW calibration. The standard deviation from frame to frame over all the molecules is shown as the shaded region. Red line: quadratic least-squares fit. ($n > 50$ at each DNA length)

Figure 4.2: Sequence of DNA used throughout this study. Binding sites for XerC and XerD are indicated as boxes. The green 'Cy3B' indicates the attachment of a Cy3B fluorophore via a C6-linker.

(Stratagene, La Jolla, USA). We mutated the aspartic acid at position 953 in the middle subunit of the covalent trimer to a cysteine. The oligonucleotides used were as follows: forward 5'-ATCTGCGCGAAGTTTTGTGTAACGCCAAATTCCGC-3'; and reverse 5'-GCG GAATTTGGCGTTACACAAAACTTCGCGCAGAT-3'.

After purification, 1 mg of the covalent trimer was incubated for 20 min with 5 mM DTT. Next, the trimer was loaded onto a heparin column and washed extensively with buffer (25 mM Tris pH 7.5, 5 % glycerol, 100 nM NaCl) in order to remove DTT. Subsequently, Cy5 maleimide was dissolved in 50 µl of DMSO, and then 950 µl of buffer was added. This dye solution was injected into the column containing the trimer and, after 20 min incubation on ice, unbound dye was washed away using buffer and the labelled trimer was eluted in an NaCl gradient. Wild-type protein was reacted with Cy5-maleimide under the same conditions, and no Cy5 fluorescence was observed for the product, confirming that labelling targets the cysteine introduced at the surface of the protein.

The labelling efficiency of Cy5-FtsK was determined using a Cary UV-Vis spectrophotometer (Agilent, Santa Clara, USA) (Figure 4.3A). We used an absorption coefficient of $250\,000\,\mathrm{cm^{-1}M^{-1}}$ for Cy5 and $519\,870\,\mathrm{cm^{-1}M^{-1}}$ for our covalent trimer of $\mathrm{FtsK_C}$, determined from its primary sequence using the ExPasy ProtParam Tool [116], giving a labelling efficiency of $53 \pm 5\,\%$ (with the error quoted from the uncertainty in the protein absorption coefficient). To confirm that only one Cy5 label had been attached to

each $FtsK_C$ trimer, we pulled down Cy5-FtsK to our slide surface by its His-tag, using a biotinylated anti-His antibody (Qiagen, Venlo, Netherlands). We found that 86 % of the 215 molecules analysed showed single-step photobleaching (Figure 4.3B), with an photobleaching time of 28 s (Figure 4.3C), and that a minority displayed a photobleaching intermediate with an intensity $\sim$40 % of the intensity of a Cy5 attached to a short (87 bp) DNA at the same position in our field of view (Figure 4.3D). This confirms that we have, at most, one label per Cy5-FtsK. Pulling down Cy5-FtsK using the His-tag in the absence of an excess of unlabelled $FtsK_C$, we found that the vast majority of molecules corresponded to a single trimer of Cy5-FtsK, suggesting that $FtsK_C$ predominantly does not hexamerise in solution (Figure 4.3E).

4.2.3 Instrumentation

Single-molecule total internal reflection fluorescence (TIRF) experiments were performed on a custom-built objective-type TIRF microscope. A green (532 nm Cobolt Samba, Cobalt, Solna, Sweden) and red (635 nm Cube, Coherent, Santa Clara, USA) laser were combined using a dichroic mirror and coupled into a fibre optic cable. The output of the fibre was focussed into the back focal plane of the objective (100$\times$ oil-immersion, NA 1.4, $f/26.5$, UPlanSApo, Olympus, Tokyo, Japan) and displaced perpendicular to the optical axis such that laser light was incident at the slide-solution interface at greater than the critical angle, creating an evanescent excitation field. Alternating laser excitation, with each laser on alternately for 25 ms, was implemented using an acousto-optic modulator (Isomet, Cwmbran, UK) for the green laser and direct modulation of the red laser. Excitation powers were set to 2 mW for both lasers, which corresponds to an approximate power density of 0.8 µW/µm^2 incident on the 50 µm $\times$ 50 µm field of view. Fluorescence emission was collected by the objective and separated from the excitation light by a dichroic (545 nm/650 nm, Semrock Rochester, USA) and cleanup filters (545 nm long-pass, Chroma, Bellows Falls, USA, and 633/25 nm notch filter, Semrock). The emission signal

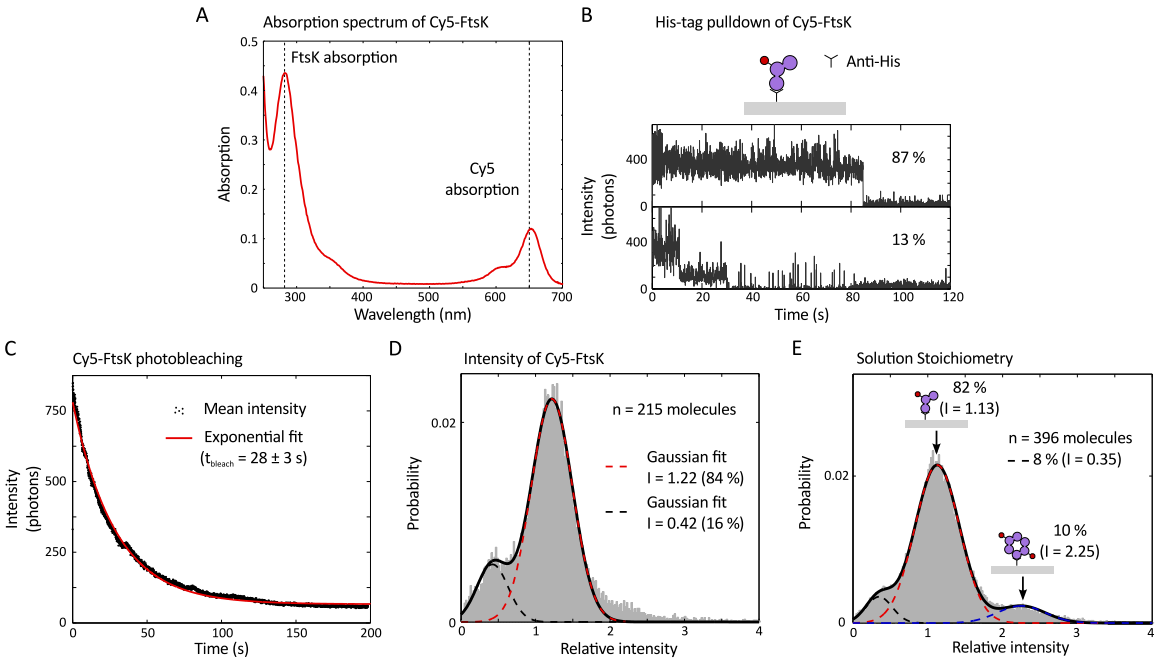

Figure 4.3: Cy5-FtsK. **(A)** The absorption spectrum of the labelled $FtsK_C$ was used to estimate the labelling efficiency. The absorption wavelengths used for Cy5 and for the protein are indicated as dashed lines. **(B)** Cy5-FtsK was pulled down to the surface, in the presence of 100-fold excess of unlabelled $FtsK_C$ trimer, using biotinylated anti-His antibody. The majority of time traces (87 %) showed single-step photobleaching, but a minority (13 %) displayed a photobleaching intermediate. **(C)** The photobleaching time for the surface immobilised Cy5-FtsK was 28 ± 3 s, much longer than any of the $FtsK_C$ timescales reported in our experiments. **(D)** Correcting for the uneven illumination across our field of view by imaging a short DNA fragment carrying a single Cy5 (Cy5-DNA), we produced a histogram of single-molecule intensities for Cy5-FtsK. We found that the majority (84 %) of frames of Cy5-FtsK, pulled down in the presence of a 100-fold excess of unlabelled $FtsK_C$, emission had an intensity 1.22 times higher than 87 bp Cy5-DNA imaged at the same position, and that a minority (16 %) of frames had a relative intensity of 0.42, corresponding to the intensity of the observed photobleaching intermediate. The intensity increase relative to DNA was consistent with a protein induced fluorescence enhancement, commonly observed when particular dyes, including Cy5, are in close proximity to a protein. The width parameters (standard deviations) of the fitted intensity distributions were 0.21 and 0.28 relative intensity units for the minor and major populations respectively. **(E)** Stoichiometry of Cy5-FtsK in solution. We used our His pulldown assay, without any excess of unlabelled $FtsK_C$, and found that only 10 % of molecules had an intensity corresponding to a doubly labelled hexamer of $FtsK_C$ (blue line). Given our labelling efficiency of $\sim$50 %, we predict that if $FtsK_C$ formed hexamers in solution, we would have seen 33 % of molecules with two Cy5 labels. Hence, we conclude that $FtsK_C$ predominantly existed, in solution, as trimers.

was focussed onto a rectangular slit to crop the image, and spectrally separated using a dichroic (630 nm long-pass, Omega, Brattleboro, USA) into two emission channels, which were focussed side-by-side onto an electron multiplying charge-coupled device camera (EMCCD, iXon 897, Andor, Belfast, UK). The EMCCD camera was set to an EM gain of 300, corresponding to an approximate real gain of 4.00 counts per photon. Each pixel on the camera corresponded to a 96 nm $\times$ 96 nm in the sample plane.

4.2.4 Sample preparation

DNA ($\sim$200 pM) was incubated for 60 s at the surface of a poly-ethylene glycol passivated coverslip, and attached to the coverslip through biotin-neutravidin interactions, which was sealed using a silicone gasket (Grace Bio-labs, Bend, OR) and a second coverslip. Imaging was performed in a buffer consisting of 50 mM Tris−HCl (pH 7.5), 50 mM NaCl, 5 mM $MgCl_2$, 100 µg/ml BSA, and 1 mM UV-treated Trolox. An enzymatic oxygen scavenging system consisting of 1 mg/ml glucose oxidase, 40 µg/ml catalase and 1.4 % (w/v) glucose was added prior to sealing the sample before image acquisition. ATP was used at a concentration of 1 mM, and XerC and XerD were used at a concentration of 2 µM.

4.2.5 Data analysis

Extraction of fluorescence intensity signals from microscope images was performed using previously described TwoTone software [42]. An apparent FRET efficiency was calculated from the extracted fluorescence emission:

$$E^* = \frac{DA}{DD + DA} \tag{4.1}$$

where DA is the fluorescence emission from the acceptor (Cy5) under donor excitation and DD is the emission from the donor (Cy3B). The fluorophore image width (FIW) was obtained from the mean width of the fitted elliptical Gaussian.

A two-dimensional projective transformation between our emission channels was generated using the image of 100 nm TetraSpeck Microspheres (Life Technologies Ltd, Paisley, UK) and TwoTone. Binding of Cy5-FtsK to DNA was automatically extracted using the following procedure. Candidate events started when the red intensity under red illumination (AA) was at least 250 photons per frame and ended when the intensity fell below 100 photons for three consecutive frames. Of these candidates, binding events were identified as having at least two frames in duration, a mean binding location in the FOV within 150 nm of the moving average position of a green signal (averaged over the preceding 2.4 s), and at least a third of individual frame localizations within 150 nm of the green moving average [117].

4.2.6 FRET event extraction

FRET was used to determine when Cy5-FtsK was within ∼10 nm of the distal *dif*-site. Since we were interested in the red intensity under red illumination (AA) signal during FRET events, we looked for events with a minimum duration of 2 frames (50 ms), because we could be sure of the location of Cy5-FtsK for the red frames that came between the first and last frame of a FRET event. To search for FRET events within binding events, we corrected the apparent FRET efficiency for leakage of the Cy3B signal into the red detection channel (parametrised by l) and for direct excitation of Cy5 by the green laser (parametrised by d) using:

$$\begin{aligned} E_{PR} &= \frac{DA'}{DA' + DD} \\ DA' &= DA - lDD - dAA \end{aligned} \tag{4.2}$$

The leakage and the direct excitation were determined by fitting the two-dimensional surface defined by:

$$DA_{\text{No FRET}} = lDD + dAA \tag{4.3}$$

using linear least squares in MATLAB. For our purposes, $DA_{\text{No FRET}}$, the DA signal when no FRET is occurring, was well approximated by our whole dataset, since FRET occurred in a vanishingly small proportion of frames. We recovered $l = 15\,\%$ and $d = 3\,\%$, consistent with previous measurements for these fluorophores on this microscope [2, 32, 118].

FRET events were considered to begin when E_{PR} was above 0.2 and end when E_{PR} dropped below 0.1 for more than one consecutive frame. Single frame events were discarded for the purpose of dwell time and stoichiometry analysis. The red intensity during FRET was the median intensity for the red frames that fell between the start and the end of the automatically extracted FRET event within a binding event.

4.2.7 Dwell time fitting

Dwell times were fit with the sum of two exponentials, with the model:

$$p(t) = \frac{a}{t_1}\exp\left(-\frac{t}{t_1}\right) + \frac{1-a}{t_2}\exp\left(-\frac{t}{t_2}\right) \tag{4.4}$$

Fits were performed using maximum likelihood estimation in MATLAB, and errors estimated as the standard deviation in each estimated parameter using bootstrap resampling with 100 resamples. The three-exponential fit, to all FtsK$_C$ dwell times on DNA, was performed using the same procedure with the model:

$$\begin{aligned} p(t) &= \frac{a_1}{t_1}\exp\left(-\frac{t}{t_1}\right) + \frac{a_2}{t_2}\exp\left(-\frac{t}{t_2}\right) + \frac{1-a_1-a_2}{t_3}\exp\left(-\frac{t}{t_3}\right) \\ a_1 + a_2 &\leq 1 \\ a_3 &= 1 - a_1 - a_2 \end{aligned} \tag{4.5}$$

Dwell times are plotted as cumulative distributions, rather than with histograms, because

cumulative distributions are free from binning artefacts and can be trivially displayed with one or more logarithmic axes, allowing processes on different timescales to be apparent on the same plot. In future work, I would suggest their continued use in preference to histograms.

4.2.8 Translocation segmentation

Translocation events were segmented using an ordinary least-squares fit to the red FIW, s, with the model:

$$\begin{aligned} s\left(t < t_{start}\right) &= s_{start} \\ s\left(t_{start} \leq t \leq t_{end}\right) &= \frac{s_{end} - s_{start}}{t_{end} - t_{start}}\left(t - t_{start}\right) + s_{start} \\ s\left(t > t_{end}\right) &= s_{end} \end{aligned} \tag{4.6}$$

where s_{start} and s_{end} are the FIW before and after translocation, and t_{start} and t_{end} are the start and end times of translocation, which are constrained to be exactly coincident with a frame for computational efficiency. When extracting image widths and intensities, the frames associated with t_{start} and t_{end} were discarded to mitigate the effect of the finite camera integration time. This model was chosen because previous models with only one translocating state well described triplex-displacement data [109], and because the shape of kymographs in studies utilising DNA curtains suggest translocation at a constant velocity [112]. Future work could relax this assumption by aligning the start and end of binding and looking at the average 'shape' of translocation events.

4.2.9 Likelihood of recombination

Recombination was apparent as a transition to a fixed, narrow green FIW. However, some proportion of timetraces that ended in a narrow FIW were due to synapsis formation (which did not proceed to recombination), with the transition back to a broad green FIW not

seen due to photobleaching or the end of a movie. Since synapsis has an exponentially distributed dwell time with a mean time of 9 ± 1 s (Chapter 3), and recombination leads to a permanent narrowing of FIW, 'true' recombinations can be distinguished with some degree of certainty based on the time for which the FIW has remained narrow.

We extracted the Cy3B bleaching time by fitting the mean DD intensity over time for all our experiments, and recovered a time of 320 ± 20 s. Using this bleaching time, the length of our movies (205 s), the lifetime of synapsis (8.6 ± 0.6 s, Figure 3.6B), and the observation that 50 % of our DNAs imaged in the presence of XerCD showed some synaptic activity, we determined a dwell time between synapsis formation of $\sim$190 s, by matching the fraction showing synaptic activity in a Markov-chain Monte Carlo simulation. Using this simulation of synapsis-only timetraces (where recombination was not possible), we found the dwell time for synaptic events that end in a narrow FIW (due to photobleaching or the end of a movie) was 9 ± 1 s (with the uncertainty dominated by the uncertainty in the synapsis lifetime).

Extracting the dwell times for events that end in a narrow FIW from our experimental data, and using our two-exponential fit (Figure 4.4), we find that $60 \pm 6\,\%$ of our dwell times come from a distribution with a dwell time of 13 ± 1 s, broadly consistent with the synaptic event lifetime determined above and that the remaining $40 \pm 6\,\%$ of the dwell times come from a distribution with a dwell time of 56 ± 5 s, which we attribute to 'true' recombination events. Given the relative proportions of the populations and their timescales, we determine that 90 % of timetraces that end in a narrow green FIW, which has been narrow for 43 s or longer, are 'true' recombinations.

4.2.10 Relative intensities

In our experiments, the detected fluorescence intensity from a single Cy5 fluorophore depended on its position within the microscope field of view (FOV), its binding position along DNA and what biomolecule it was attached to. The dependency on FOV position was

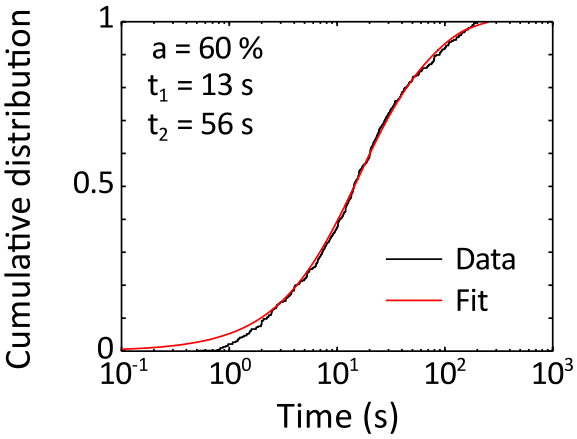

Figure 4.4: Dwell times of recombination-like events, as distinguished by a narrow green FIW at the end of data acquisition. The dwell times were fit by two exponentials, recovering that $60 \pm 6\,\%$ of dwells had a lifetime of 12.5 ± 1.3 s (consistent with a synapsis lifetime of around 9 s) and that the remaining $40 \pm 6\,\%$ of events had a lifetime of 56 ± 5 s. The longer events correspond to 'true' recombinations, and so 90 % of recombination-like events with a narrow FIW lasting 43 s or longer were recombinations ($n = 342$ recombination-like events).

due to the non-uniform excitation field across the FOV. We followed an established method for determining this dependency [118], modifying the procedure to use a reference standard DNA. Our reference DNA consisted of an internally labelled Cy5 separated from the 5' biotin moiety, used for surface attachment, by 87 bp. After setting laser powers and TIRF angle, we imaged many (∼100) FOVs of this Cy5-DNA standard. We excluded molecules that photobleached, or displayed photoblinking, by filtering those whose intensities deviated by more than twice the median absolute deviation for three or more frames. The mean intensity and the localization position were fit with a two-dimensional elliptical Gaussian, using non-linear least squares fitting in MATLAB, which recovered a predicted intensity for a single Cy5-DNA at any position within the FOV.

We corrected for the particular single-molecule intensity of Cy5-FtsK using our anti-His pulldown data. We extracted the frame-by-frame intensities of Cy5-FtsK pulled down to our surface using a biotinylated anti-His antibody, and fit two Gaussian distributions, using non-linear least squares in MATLAB (Figure 4.3D). We extracted a relative intensity for Cy5-FtsK 1.22 times higher than for the Cy5-DNA standard used. The minor population (∼16 % of frames) corresponds to the rare photobleaching intermediate observed in 13 % of pulldown time-traces (Figure 4.3B).

We also applied a correction for the evanescent TIRF field in our experiments, since the further along DNA that Cy5-FtsK bound, the dimmer it appeared. To estimate the TIRF

decay length in our experiments, we compared the mean intensity of synapsed and product DNAs to the mean intensity of un-synapsed DNAs. We found that synapsed and product DNA had $210 \pm 10\,\%$ of the brightness of un-synapsed DNA (with the error quoted as the standard deviation between sample wells). Assuming the position of the fluorophore above the coverslip is linearly proportional to its length along DNA (which can be considered a first-order approximation), this corresponds to a TIRF decay length in the green channel of 1400 bp, corresponding to a decay length of 1700 bp in the red channel, given that both green and red lasers are incident on the coverslip at the same angle, but they differ in wavelength (for a given incident angle, the TIRF decay length is inversely proportional to the wavelength used). Using this estimated red channel decay length, we predict that Cy5 attached to DNA at the position of the distal *dif*-site in a DNA in the un-synapsed configuration ($\sim$1300 bp) has 50 % of the intensity of the Cy5-DNA standard, and that Cy5 attached to DNA at the position of a synapsis, or the *dif*-site in the DNA product of recombination, ($\sim$230 bp) should have 92 % of the intensity of the Cy5-DNA standard. Using these three correction factors (for non-uniform illumination, the single-molecule intensity of Cy5-FtsK, and the TIRF field), we could predict the intensity of a single Cy5 attached to $\mathrm{FtsK_C}$ at any position within our FOV and bound anywhere along DNA.

4.2.11 Predicted intensity distributions

Given the labelling efficiency and the intensity of Cy5-FtsK, we predicted the range of intensities expected from a single hexamer of $\mathrm{FtsK_C}$ and a double hexamer of $\mathrm{FtsK_C}$. For an assembly of M trimers (two for a single hexamer and four for a double hexamer), with a labelling efficiency ϵ, the probability of observing a complex with n labels ($n > 0$) is:

$$p(n) = \frac{\epsilon^n \left(1-\epsilon\right)^{M-n} \frac{M!}{(M-n)!\,n!}}{\sum\limits_{1 \le l \le M} \epsilon (1-\epsilon)^{M-l} \frac{M!}{(M-l)!\,l!}} \tag{4.7}$$

Hence, the predicted intensity distribution is given by:

$$p(I) = \sum_{1 \leq n \leq M} \frac{p\,(n)}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(I - nI_1)^2}{2\sigma^2}\right) \tag{4.8}$$

The intensity I_1 refers to the predicted intensity for a single fluorophore ($I_1 = 1$ when we are expressing relative intensities), and σ is the width of the intensity distribution, which can be estimated from the intensity distribution for His-pulldwon Cy5-FtsK as 0.28 intensity units relative to the intensity of Cy5-DNA at the surface.

The average number of fluorophores per assembly is given by:

$$\langle n\,(M)\rangle = \sum_{1 \leq n \leq M} np(n) \tag{4.9}$$

For a labelling efficiency of $\sim$50 %, we get an average of 1.33 fluorophores per single hexamer and an average of 2.13 per double hexamer. Hence, we predict that a single hexamer would show a 33 % increase in intensity, on average, after the first labelled $FtsK_C$ trimer binds to DNA, and that a double hexamer would show a 110 % increase.

4.3 Results

The previous section has described the DNA substrate, the labelled trimer of $FtsK_C$, and the experimental and analytical tools used. In this section, I describe their use to understand the mechanism of $FtsK_C$ assembly and translocation, and its interaction with XerCD-*dif*.

4.3.1 Spectrally separate fluorophores report on DNA conformation and $FtsK_C$ behaviour

In the absence of ATP and XerCD, we observed transient binding events that showed no obvious translocation (Figure 4.5A); these binding events were distributed randomly along the DNA (Figure 4.5B). The $FtsK_C$ dwell time was fit using three exponentials (Figure 4.5C and Section 4.2.7), recovering a major time of 0.84 ± 0.07 s (61 ± 5 %),

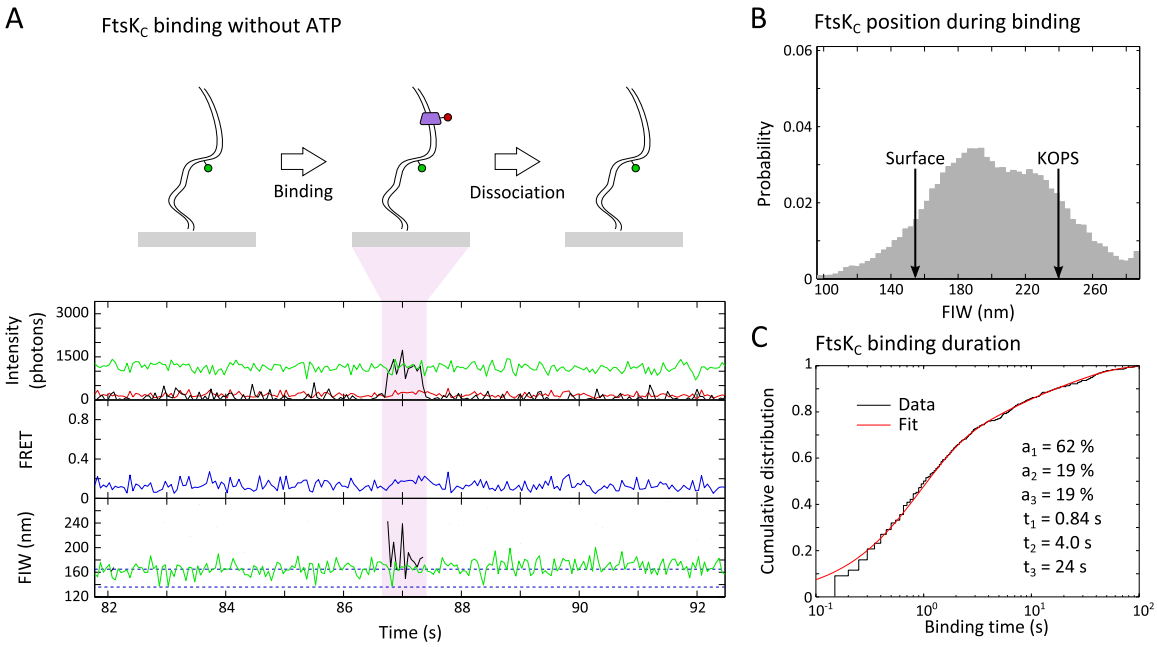

Figure 4.5: FtsK$_C$ binds DNA in the absence of ATP. **(A)** FtsK$_C$ binding is apparent as the appearance of red fluorescence. The top panel shows the intensities recorded in the green emission channel under green excitation (DD, green line), red emission under green excitation (DA, red line) and red emission under red excitation (AA, black line). The middle panel shows the apparent FRET between the Cy3B and Cy5. Under these conditions no FRET was observed. The bottom panel shows the FIW in the green channel under green excitation (green) and in the red channel under red excitation (black). **(B)** The FtsK$_C$ position during binding events is distributed along the DNA. The red FIW corresponding to binding at the surface, and binding at KOPS are indicated with arrows. ($n = 490$ binding events). **(C)** The binding time distribution of FtsKC is well fit by three exponential dwell times, which suggests that there are different types of binding event as FtsK$_C$ assembles on DNA.

and two minor times of 4.0 ± 0.9 s ($18 \pm 5\,\%$) and 24 ± 3 s ($19 \pm 3\,\%$). We attribute the shortest time to transient association of a single Cy5-FtsK with DNA, the intermediate time to formation of a hexamer from two Cy5-FtsK on DNA (since a hexamer would be more stable than a trimer), and the longest time to non-specific sticking to the surface of the slide and subsequent photobleaching (photobleaching time for Cy5 at the surface was 28 ± 3 s, Figure 4.3C). These non-specific sticking events were distinguished by their fixed narrow red FIW and were excluded from further analysis. Hence, around a quarter of DNA binding events had a dwell time that indicated hexamer assembly.

In the presence of ATP, in addition to transient binding events (Figure 4.6A), we observed events with a decreasing red FIW, consistent with Cy5-FtsK translocations along

DNA (Figure 4.6B, $n = 176$). Occasionally ($n = 7$), the red and green fluorescence disappeared simultaneously, which indicated that FtsK$_C$ had displaced the biotin-neutravidin interaction that tethered the DNA to the surface (Figure 4.6C). This rare displacement is consistent with previously described behavior (38). A minority of events ($n = 8$) showed a decrease in green FIW while FtsK$_C$ was bound to DNA, which could be interpreted as the sticking of FtsK$_C$ to the surface or translocation-induced looping (Figure 4.6D). These results establish that, in general, translocation does not induce DNA looping.

We used a least-squares (LS) fit to the red FIW (Section 4.2.8) to segment translocation events into three stages: assembly before translocation, translocation, and residence after translocation (Figure 4.6A and Figure 4.7A). The mean FIW before translocation, 220 nm, corresponded to the position of the KOPS sequences (FIW = 230 nm), and the mean position after translocation, 161 nm, was close to the surface (FIW = 155 nm), past the Cy3B (FIW = 183 nm) (Figure 4.7B and Figure 4.1C), indicating that Cy3B does not impede FtsK$_C$ translocation. This suggests that the translocating FtsK$_C$, that we observe, assembles at or near KOPS and translocates past Cy3B, reaching the surface attachment point of the DNA. FRET, which reports the proximity of Cy5-FtsK to Cy3B within ~10 nm ($R_0 = 6.5$ nm for these fluorophores, Section 3.2.8), was not observed under these conditions.

We determined the duration of FtsK$_C$ assembly, translocation and residence using the LS fit to the red FIW. We found that dwell time distributions were in better agreement with two independent exponential processes rather than one, as judged using the Bayesian information criterion (BIC) [119] (ΔBIC > 10 in all cases, indicating strong evidence in favour of a two exponential model). Hence, except where otherwise noted, we fit dwell time distributions with two exponentials, and restrict our discussion to the major population (Section 4.2.7 and Table 4.1). Averaging across all fits to dwell times, the major population accounts for ~90 % of dwells, and we suggest that the minor population is a fit to background events. Using this method, the assembly time before translocation was

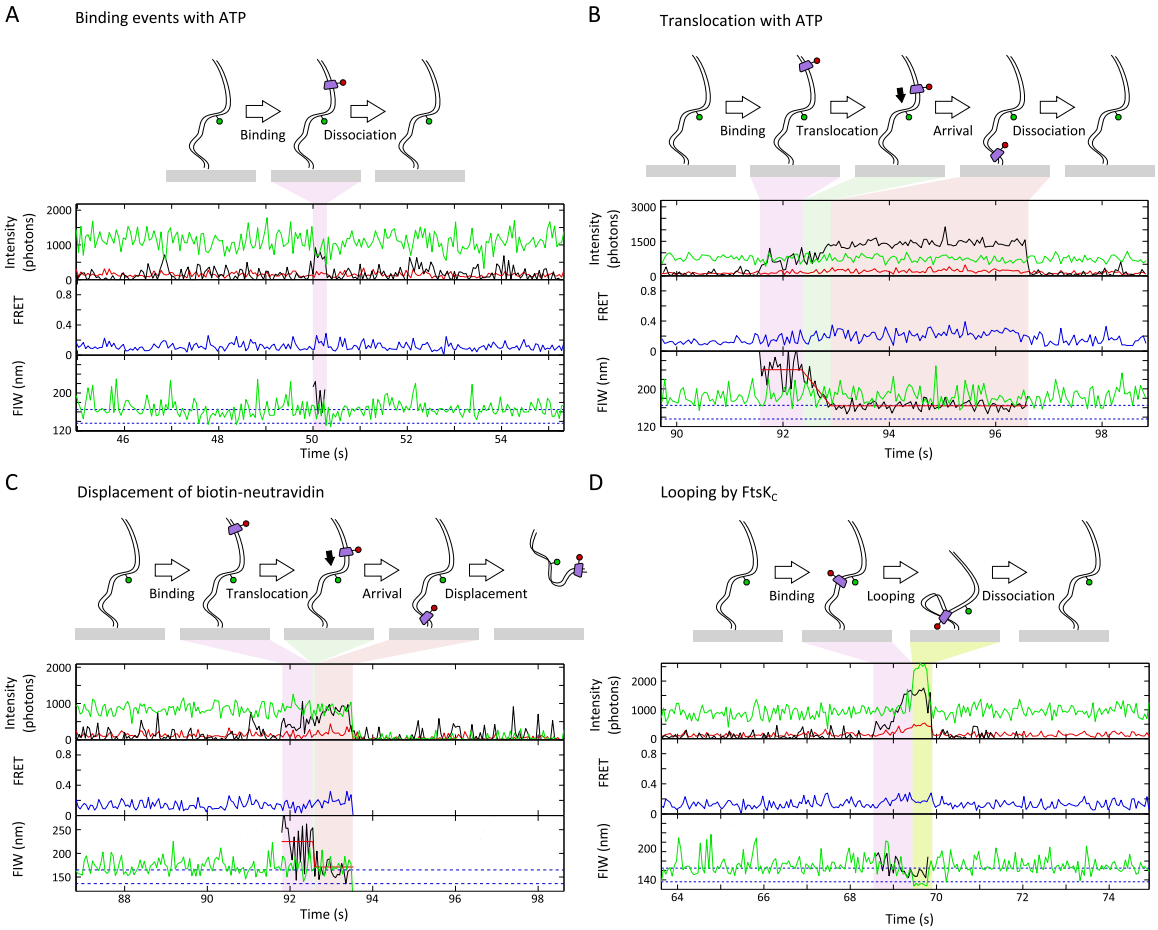

Figure 4.6: FtsK$_C$ binding and translocation in the presence of ATP. **(A)** Transient binding events with no obvious translocation are apparent in the presence of ATP, perhaps associated with incomplete FtsK$_C$ assembly or a too-short translocation distance to be resolved in our assay. **(B)** Translocation event in the presence of ATP. The same colour-coding is used throughout this work. ($n = 176$) **(C)** Displacement of biotin-neutravidin tethering interaction by FtsK$_C$ ($n = 7$). **(D)** Potential translocation dependent looping event apparent as a narrow green FIW at around 69.5 s. ($n = 8$ events with FtsK$_C$ present during a transient narrowing of green FIW in the absence of XerCD)

0.26 ± 0.12 s, and the residence time after translocation was 2.1 ± 0.2 s (Figure 4.7C). The mean translocation time was 0.33 ± 0.02 s ($\pm$ SEM), which corresponds to a translocation velocity of 12 ± 1 kb/s at 21 °C (given that translocation events originated at KOPS and ended close to the surface), which is a factor of two faster than previously measured at this temperature [111], but we note that our method did not require looping by $FtsK_C$ or the movement of a large bead with associated hydrodynamic drag to observe translocation. Hence, this velocity corresponds to the physiological translocation speed along unconstrained, flexible, 'naked' dsDNA.

4.3.2 Stoichiometry of translocating $FtsK_C$

Previous biochemical analysis has suggested that the translocating unit of $FtsK_C$ assembles on DNA in six steps to form single hexamers [109]. However, double ring structures that could correspond to head-to-head hexamers on DNA have been observed in electron micrographs of *Pseudomonas aeruginosa* $FtsK_C$ lacking the γ subdomain, and crystal structures of this protein contained dodecameric ring structures [56]. In order to directly assess the stoichiometry of the active unit of $FtsK_C$, we measured the intensities of Cy5-FtsK and counted the number of fluorophores present during assembly, and during interactions with XerCD. A single hexamer, made up of two labelled trimers, would have between one and two fluorophores attached, and a double hexamer between one and four.

For this analysis, we used the shorter, single *dif* containing, DNA fragments formed by XerCD recombination at the surface (Figure 4.8A), because all features of the DNA (*dif*-site, KOPS, and the Cy3B) were 1000 bp closer to the surface than in the original substrate and hence in a region of higher illumination intensity in the evanescent field of TIRF. This DNA recombinant product was identified as having a permanent narrow green FIW. Translocation events were segmented as before (Figure 4.8B and C). If $FtsK_C$ preassembled in solution and then bound DNA fully assembled, we would predict no change in intensity between $FtsK_C$ binding and the start of translocation, however, if $FtsK_C$

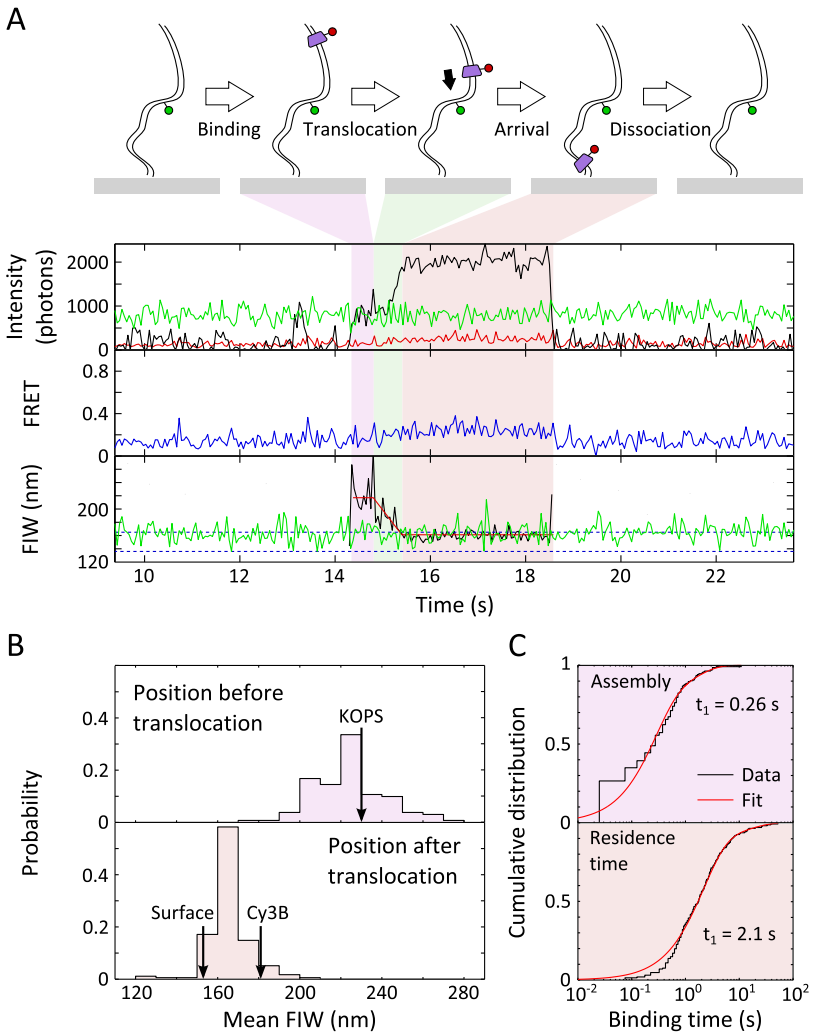

Figure 4.7: Translocation events in the absence of XerCD. **(A)** Translocation event. Top panel (intensity): green emission under green excitation (DD, green), red emission under green excitation (DA, red) and red emission under red excitation (AA, black). Middle panel: apparent FRET. Under these conditions no FRET was observed. Bottom panel: FIW in the green channel (green) and in the red channel (black), least-squares fit to the red FIW (red) used to segment the event into before (purple highlight), during (green highlight), and after (orange highlight) translocation. ($n = 179$ events) **(B)** Start and end positions of translocations. Black arrows: positions of features on the DNA, obtained using the FIW calibration (10 nm bin width). **(C)** Dwell times. Top: $FtsK_C$ assembly. Bottom: residence after translocation. Distributions fit using two exponential processes (Table 4.1 and Section 4.2.7).

assembled on DNA, the intensity should increase before translocation begins (Figure 4.9A). The mean intensities and positions during assembly, across all translocation events, were extracted (Figure 4.9B). At a time 0.5 s after the appearance of red fluorescence signalling the presence of $FtsK_C$, more than 66 % of translocations had begun (Figure 4.8C). We saw that on this timescale, there was a clear increase in the intensity of the red signal without any change in FIW (Figure 4.9B), which implies that there was a stepwise assembly of Cy5-FtsK before translocation. Given our labelling efficiency of $\sim$50 %, if $FtsK_C$ assembled into a single hexamer comprised of two trimers, we would anticipate an average increase of 33 %, and if it assembled into a double hexamer, we would anticipate an average increase of 110 % before translocation (Section 4.2.11). The observed 38.5 % increase in intensity during the 0.5 s after binding is consistent with assembly into a single hexamer (Figure 4.9B). The intensity increase suggests that $FtsK_C$ assembles on DNA, rather than in solution. Using biotinylated anti-His antibodies, we pulled Cy5-FtsK out of solution, and determined that, at 5 nM and in the absence of DNA, the predominant stoichiometry was a single trimer (Figure 4.3E), consistent with stepwise assembly on DNA before translocation.

We also determined the stoichiometry of $FtsK_C$ during its interaction with XerCD complexes. In the presence of Cy5-FtsK, ATP and XerCD, we observed FRET, signalling proximity between Cy5-FtsK and the Cy3B placed near the XerD binding site (Figure 4.8A and Figure 4.10). We used the Cy5 intensity during FRET to extract the number of Cy5 fluorophores attached to an $FtsK_C$ complex while it interacted with XerCD-*dif* (Section 4.2.6). This intensity was expressed relative to the calculated intensity for a single Cy5 at the same position (Section 4.2.10) [118]. Using this intensity for a single fluorophore, the labelling efficiency of Cy5-FtsK, and the width of the intensity distribution for a single fluorophore (Figure 4.3D), we predicted the intensity distribution for a single and a double hexamer (Section 4.2.11 and Figure 4.11). Across all DNA conformations, the data agrees much better with the single hexamer prediction (Figure 4.11). Hence, we conclude that, consistent with its assembly into a single hexamer before translocation,

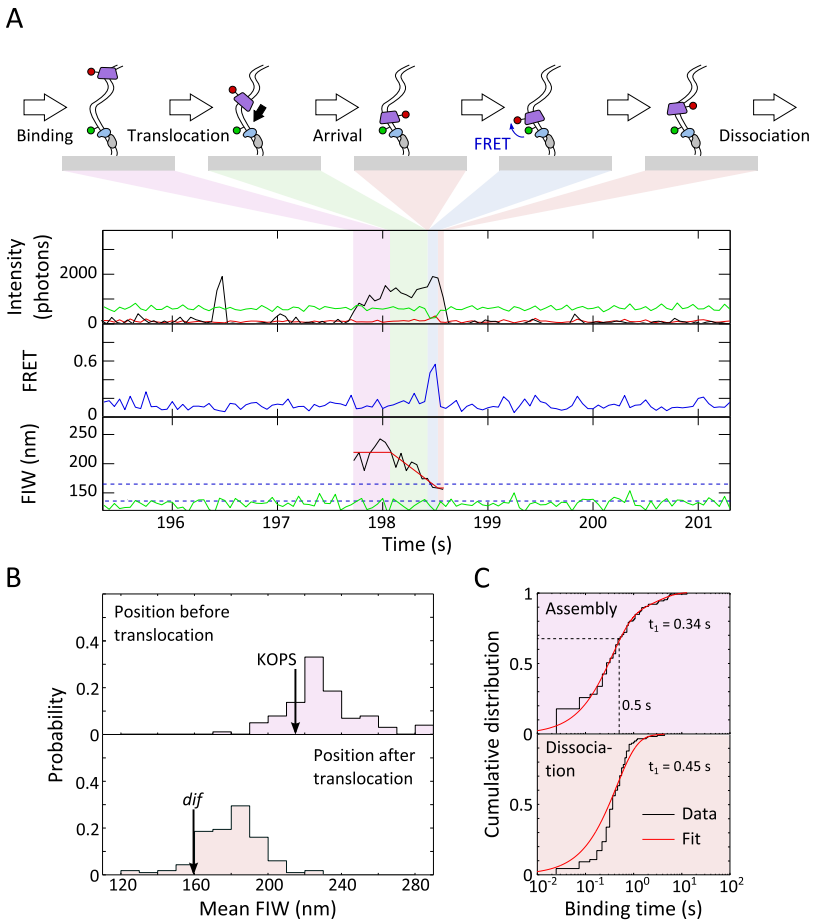

Figure 4.8: FtsK$_C$ translocates on the product of recombination. **(A)** The product of recombination is a single *dif*-site with a 2.7 kb tail, with the furthest part of the DNA 1 kb closer to the surface than the DNA substrate. Blue shading highlights FRET. ($n = 126$, with FRET apparent in 59) **(B)** Start and end positions of translocations. The mean starting position is 225 ± 2 nm, and the mean end position is 174 ± 2 nm (with the errors computed using bootstrapping with 1000 resamples; 10 nm bin width). **(C)** Dwell times before and after translocation. Distributions are fit using MATLAB (Table 4.1). At a time 0.5 s after binding (dashed line, top panel), more than 66 % of translocations have begun.

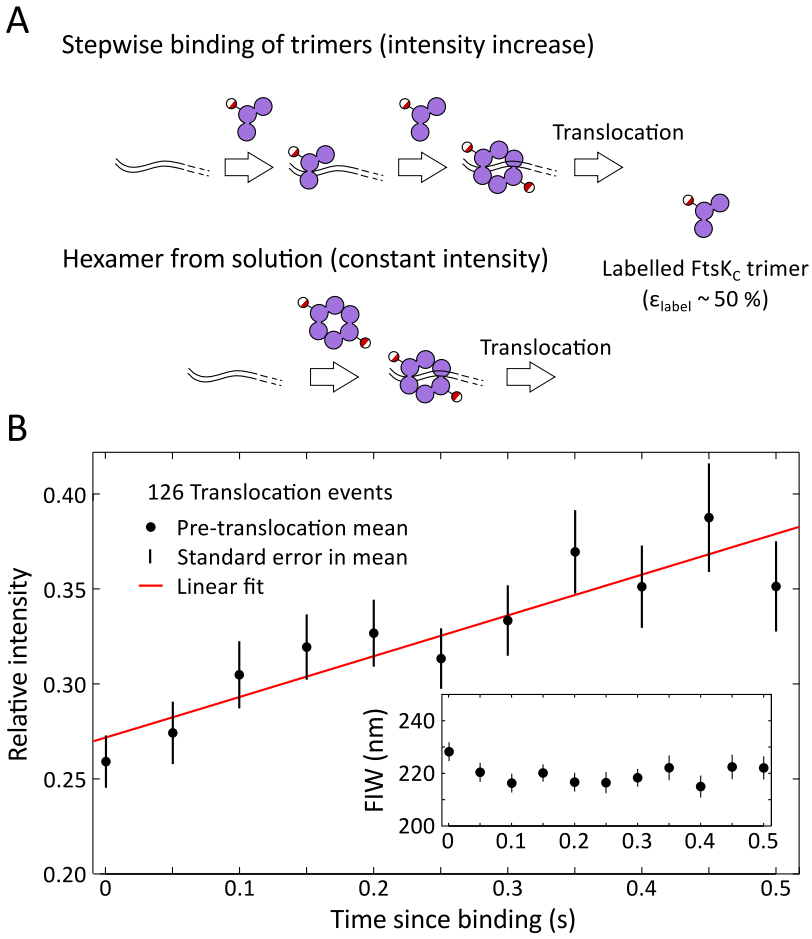

Figure 4.9: FtsK$_C$ stoichiometry. **(A)** Intensity changes before translocation. If FtsK$_C$ assembles in a step-wise fashion before translocation, we expect an increase in intensity before translocation starts; whereas, if fully assembled oligomers in solution bind and translocate there would be no increase. **(B)** Mean relative intensity for before translocation. A single Cy5 had a relative intensity of 0.2 to 0.4 between 2 kb to 3 kb along the DNA (Section 4.2.10). Error bars: $\pm$ SEM. Least squares fit (red) with intensity increase of $38.5 \pm 0.2\,\%$ in the 0.5 s since binding (by which time more than 66 % of translocations have begun). Inset panel: mean FIW for the same molecules.

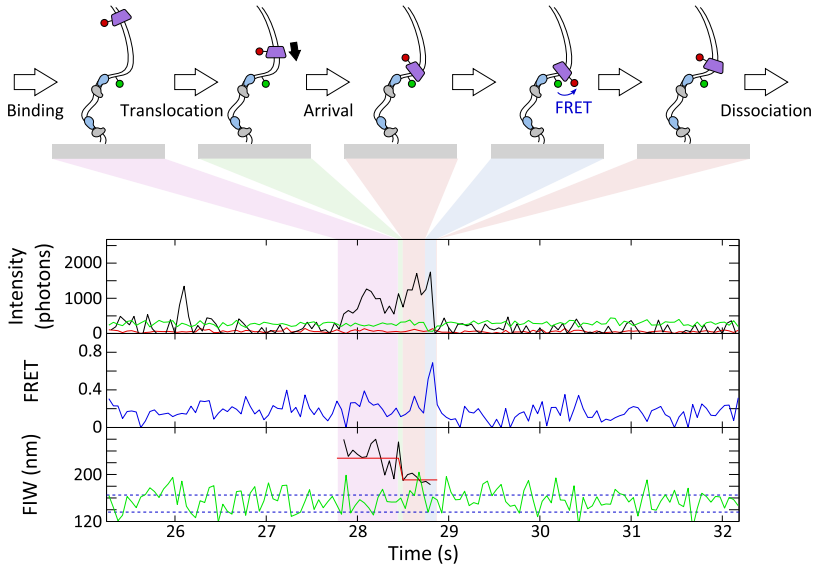

Figure 4.10: Translocation stops at un-synapsed XerCD-*dif*. Translocation. Blue shading highlights FRET. Translocation apparent in 342 events, of which 102 have at least a single frame of FRET accompanied by anti-correlated DD and DA changes.

$FtsK_C$ interacts with XerCD as a single hexamer.

4.3.3 Interactions betweeen translocating $FtsK_C$ and XerCD-*dif* complexes

FtsK activates site specific recombination by XerCD, and biochemical studies have shown that when $FtsK_C$ encounters XerD during translocation, it stops there [108]. Moreover, QD labelled $FtsK_C$ has been observed to pause and reverse translocation direction after meeting un-synapsed XerCD-*dif* on DNA curtains [113]. In this work, we use singly Cy5-labelled $FtsK_C$ and observe its interaction with XerCD-*dif* in both un-synapsed and synapsed configurations.

We imaged our double *dif* tethered DNA substrate in the presence of Cy5-FtsK, ATP and XerCD. We observed $FtsK_C$ translocations towards the surface when the DNA was in an un-synapsed conformation, as judged by the green FIW ($n = 342$, Figure 4.10 and Figure 4.12A). In 30 % of these events, at least one frame of FRET was apparent between the Cy3B attached to DNA and Cy5-FtsK, suggesting that $FtsK_C$ remained

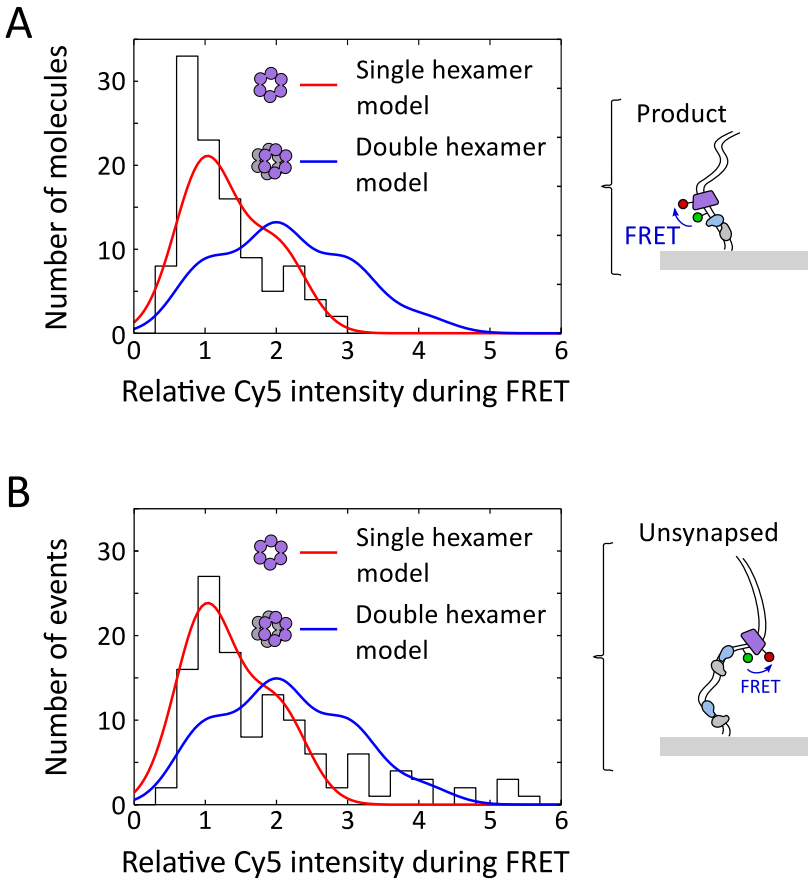

Figure 4.11: Stoichiometry during FRET. During FRET, the Cy5 attached to $FtsK_C$ must be within ∼10 nm of the Cy3B attached to the DNA, hence we know where in the evanescent field the Cy5 is (on average). The corrected intensity during FRET events for un-synapsed DNA conformations were extracted and histogrammed ($n = 122$). Given a labelling efficiency of 53 %, a correction for the distance along the DNA in our evanescent excitation field, and the width of the intensity distribution for single Cy5-FtsK pulled down to the surface, we plot the predicted intensity distribution for a single (red) and double (blue) hexamer of FtsKC. Using the sum of squared deviations (SSD) between the model and the data, we find that a single hexamer provides a better explanation of our data. **(A)** Product DNA conformation (right panel). ($n = 108$, single hexamer: SSD = 330, double hexamer: SSD = 1230) **(B)** Substrate DNA conformation (right panel). ($n = 122$, single hexamer: SSD = 230, double hexamer: SSD = 600)

within ∼10 nm of the *dif*-site for more than approximately the camera frame time (25 ms). Segmenting the translocation events as before, we observed that un-synapsed XerCD-*dif* acts as a roadblock for $FtsK_C$, with the mean FIW after translocation, 181 nm, matching the location of the surface-distal *dif*-site (FIW = 183 nm, Figure 4.12B). The dwell times for assembly and residence after translocation were fit (Figure 4.12C), recovering an assembly time of 0.25 ± 0.03 s, and a residence time of 0.50 ± 0.02 s. The assembly times were independent of the presence of XerCD as anticipated, but the residence time after translocation decreased fourfold when $FtsK_C$ encountered XerCD bound to un-synapsed *dif* (compared to when it encountered the biotinylated DNA end), suggesting that interactions between XerD and $FtsK_C$ promote $FtsK_C$ dissociation. This decreased residence time gives us confidence that, under these conditions, it reflects interactions with XerCD bound to *dif*. Only a small minority of translocations ($n = 1$) towards un-synapsed XerCD resulted in a reversal by $FtsK_C$, apparent as an increasing red FIW after residence. The lack of reversals in our data contradicts previous observations [113], where $FtsK_C$ reversed translocation direction after meeting XerCD bound to *dif*. The previous observations were made with QD labelled $FtsK_C$ trimers, where there was the possibility of multiple $FtsK_C$ trimers tethered to the same QD. This could permit loading of a second $FtsK_C$ motor during the residence time of the original motor at XerCD, and may explain the presence of reversals in the previous observations. The Cy5-FtsK used in this work had only one trimer of $FtsK_C$ per fluorescent label, and hence overcomes this limitation. In the presence of XerCD, we recovered a mean translocation time of 0.18 ± 0.02 s ($\pm$ SEM), which is shorter than the translocation time (0.33 ± 0.02 s) observed in the absence of XerCD, which, given the shorter translocation distance from KOPS to XerD, corresponds to a similar mean translocation speed of 15 ± 1 kb/s.

$FtsK_C$ activates XerCD recombination at *dif*, which was apparent as DNA molecules that transitioned to a persistent narrow green FIW (Figure 4.1A). Of these events, 90 % that had remained with a narrow FIW for at least 43 s had recombined (rather than being

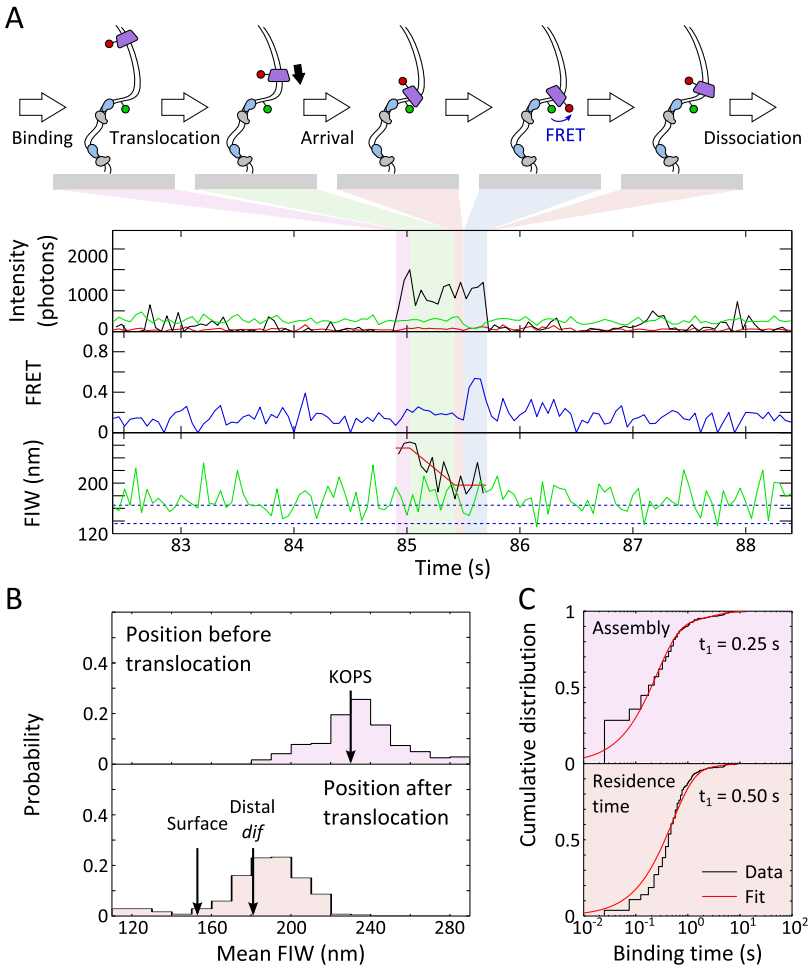

Figure 4.12: Translocation stops at un-synapsed XerCD-*dif*. **(A)** $FtsK_C$ interacts with XerCD after translocation: translocation towards single XerCD loaded *dif* (same conditions and similar event as in Figure 4.10; see Figure 4.10 legend for number of events). **(B)** Start and end positions of translocations. Mean starting position is 229 ± 1 nm; mean end position is 181 ± 1 nm ($\pm$ SEM; 10 nm bin width). **(C)** Dwell times before and after translocation. Distributions are fit in MATLAB (Table 4.1).

long-lived synapses) (Section 4.2.9 and Figure 4.13), and we define a recombination event as meeting this criterion ($n = 73$). In 50 % of recombination events ($n = 39$), FtsK$_C$ binding was apparent after synapsis formation, but not during (Figure 4.13A). In 8 % of recombination events ($n = 6$) obvious translocation (consisting of binding followed by a decreasing image width without any anomalous changes in intensity) accompanied recombination (Figure 4.13A). Obvious translocations did not accompany all recombination events because translocations may not have initiated at KOPS, and hence would have been too short to resolve in our assays, and because recombination could have been activated by the $\sim$25 % of FtsK$_C$ complexes that carry no fluorescent label. If translocation-induced looping (leading to formation of a synapsis) were a prerequisite for the activation of recombination, it would have been apparent as an FtsK$_C$ presence during synapsis formation. However, only 10 % of recombination events ($n = 8$) followed this pattern (Figure 4.13B), suggesting that translocation-induced looping between *dif*-sites was not necessary for the activation of recombination. This is consistent with the previous result that the formation of XerCD-dif synapsis is independent of FtsK$_C$ (Chapter 3).

FtsK can resolve catenated chromosomes, produced by replication [120]. We have previously suggested that FtsK may facilitate this decatenation by remaining in the vicinity of XerCD-*dif* after the activation of recombination and activating multiple rounds of recombination (Section 3.4.3 and [48]). Taking the residence time (during synapsis) of the first FtsK$_C$ to dissociate after the synapsis formation that leads to recombination, i.e. the best candidate for having activated the recombination, we recovered a residence time of 0.9 ± 0.2 s (Figure 4.14). This residence time is the total time from the first observation of binding (or synapsis formation) until FtsK$_C$ dissociates, and is an overestimate of the time FtsK$_C$ spends at XerCD-*dif*. Despite this, it is still shorter than the 1.6 s it takes XerCD to complete the recombination reaction (Figure 3.18). This suggests that FtsK$_C$ does not reside at XerCD long enough to activate multiple further rounds of recombination.

Not all translocations towards synapsed XerCD-*dif* resulted in recombination. We

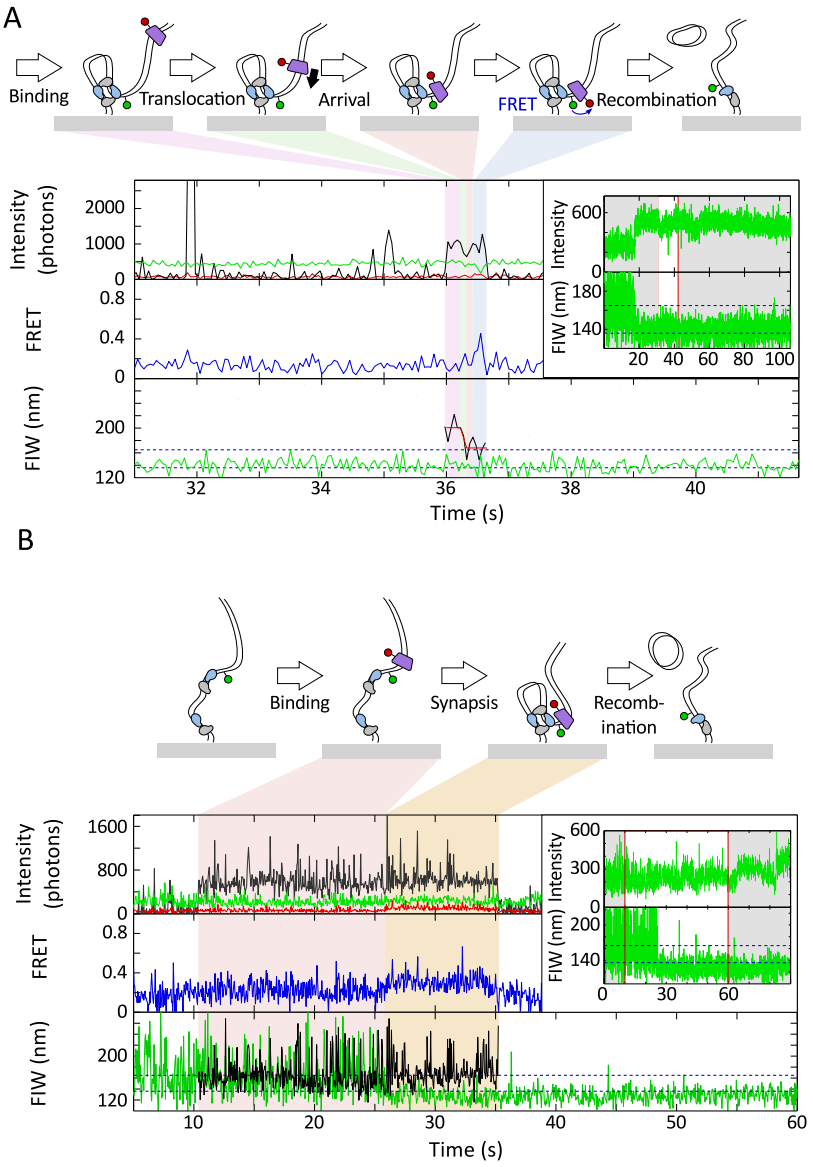

Figure 4.13: Activation of recombination by $FtsK_C$. **(A)** $FtsK_C$ arrives after synapsis formation leading to recombination. Successful recombination reactions result in a permanent narrow green FIW (inset panel). The highlighted region in the inset corresponds to the main figure. $FtsK_C$ binding is apparent after synapsis formation (but not during) in 39 out of 73, recombination events that meet the 90 % confidence criterion. **(B)** Recombination event with $FtsK_C$ present during synapsis formation. If $FtsK_C$ induced looping was required for recombination then recombination events would have predominately followed this pattern. $FtsK_C$ binding was apparent during synapsis formation in 8 out of 73 recombination events.

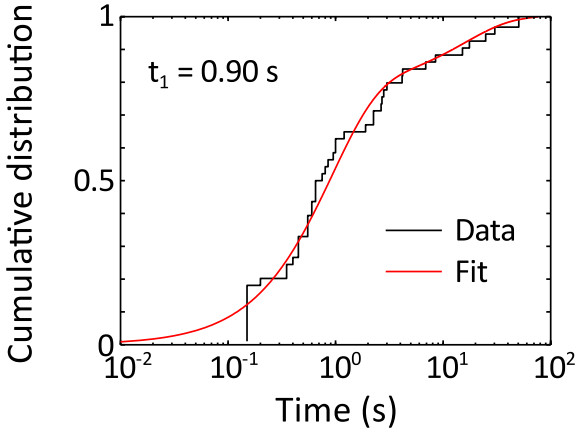

Figure 4.14: Dwell time of first $FtsK_C$ to dissociate after synapsis formation in recombining events. ($n = 47$ recombination events with $FtsK_C$ binding; fit described in Section 4.2.7).

analysed these non-productive translocation events, distinguished by a transiently narrowing green FIW (Figure 4.15A and B). Of 65 events, around a quarter showed $FtsK_C$ present during synapsis disassembly (Figure 4.15A). Again, XerCD acted as a roadblock for $FtsK_C$, with the end of translocation approximately matching the position of the synapsis (Figure 4.15C). FRET was apparent after translocation in 46 % of events. For all these translocations towards a synapsis, the assembly time was 0.18 ± 0.40 s, consistent with the assembly time on un-synapsed DNA, and the post-translocation residence time was 0.65 ± 0.07 s (Figure 4.15D), again consistent with the residence time at un-synapsed XerCD dif (Figure 4.12C), and with the residence time of $FtsK_C$ during productive recombination events (0.9 ± 0.2 s). From this, we infer that $FtsK_C$ does not reside for a significantly longer or shorter time depending on whether it encounters a single XerCD-*dif* or a XerCD-*dif* synaptic complex, or depending on whether it has activated recombination.

FRET may represent a transient association between Cy5-FtsK and XerCD-*dif* complexes. The automatic extraction of FRET events (Section 4.2.6) required a minimum duration of 0.1 s (2 frames), and so an upper bound on the duration of this association was estimated (Figure 4.16). We find that FRET events had a dwell time of 0.30 ± 0.04 s and 0.23 ± 0.02 s for DNA in the substrate and product configurations respectively. The dwell after FRET (before dissociation) was 0.09 ± 0.02 s and 0.07 ± 0.01 s for DNA in the substrate and product configurations. These dwells, between the end of FRET and dissociation, are shorter than the mean translocation time of 0.18 ± 0.02 s, suggesting there is no time for $FtsK_C$ to reverse and translocate after leaving the proximity of XerCD, consistent with the absence of an increasing FIW before dissociation, confirming that $FtsK_C$ dissociates, rather than reverses, after meeting XerCD-*dif*.

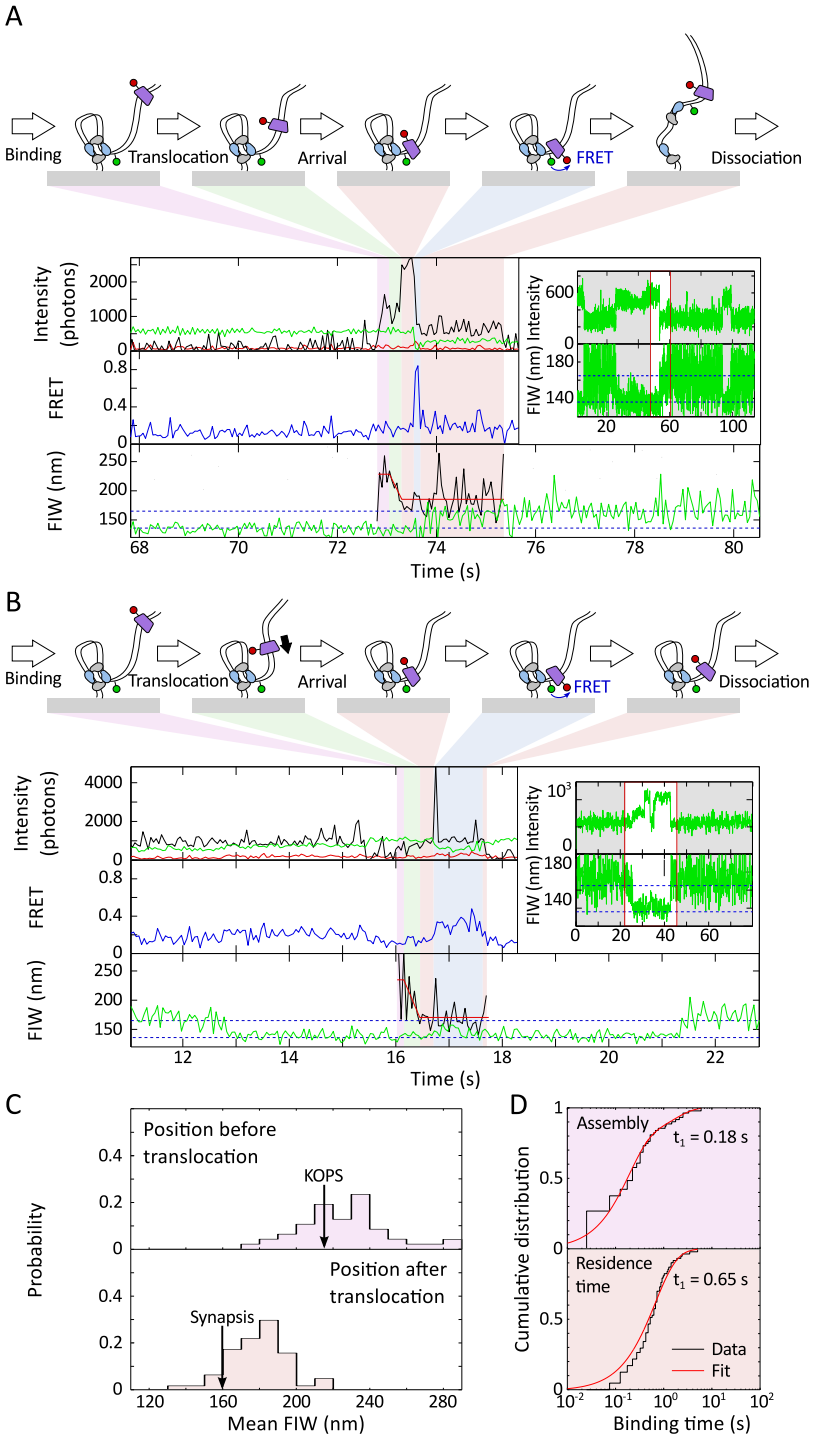

Figure 4.15: Arrival of $FtsK_C$ at XerCD-*dif* synapses ($n = 65$). **(A)** $FtsK_C$ present during synapsis disassembly. A translocating $FtsK_C$ meets an XerCD-*dif* synapsis, which disassembles while $FtsK_C$ remains bound to DNA. The red and green intensities simultaneously decrease at around 73.5 s, consistent with $FtsK_C$ remaining at the distal *dif*-site as the synapsis disassembles. ($n = 17$, of which 10 show FRET) **(B)** $FtsK_C$ arrives at a synapsis and dissociates before the synapsis disassembles. Synapsis formation and disassembly is apparent as a transiently narrowing FIW (inset panel). ($n = 48$, of which 20 show FRET) **(C)** Start and end positions of translocations. The mean starting position is 221 ± 4 nm, and the mean end position is 175 ± 2 nm ($\pm$ SEM; 10 nm bin width) **(D)** Dwell times before and after translocation. Distributions are fit using MATLAB (Table 4.1).

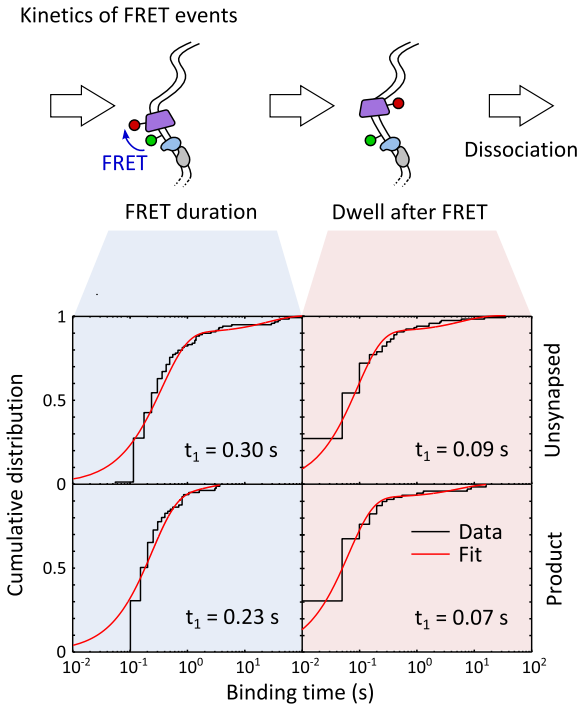

Figure 4.16: FRET reports on the proximity of $FtsK_C$ to the XerCD bound to XerCD-*dif*. The dwell time distribution for automatically extracted FRET events (lasting a minimum of two frames), for un-synapsed (top panels, $n = 119$) and for product (bottom panels, $n = 81$) DNA, was fit using two exponentials (Table 4.1). The dwell times during and after FRET are similar for both DNA conformations, and the short time between the end of FRET and the disappearance of red fluorescence suggests that $FtsK_C$ dissociates after meeting XerCD bound to *dif*, rather than reversing translocation direction.

4.4 Discussion

4.4.1 $FtsK_C$ assembles stepwise on DNA into a single hexamer

In this work, using labelled $FtsK_C$ trimers, we have addressed the question of whether $FtsK_C$ functions as a single or double hexamer. Double hexamers of $FtsK_C$ have been observed using electron microscopy and in crystal structures of $FtsK_C$, and have been implicated in the explanation of translocation reversals, and the extrusion of DNA loops, in optical and magnetic tweezers single-molecule experiments [56, 101, 111]. Our precise fluorescence measurements unequivocally show that $FtsK_C$ assembles on DNA into a single hexamer, rather than by forming a solution hexamer that then loads onto DNA, and subsequently translocates as a single hexamer. Moreover, our data show that $FtsK_C$ activates XerCD-*dif* recombination as a single hexamer. We therefore conclude that double hexamers may not be physiologically relevant. Consistent with our observations, previous biochemical analysis using FtsK monomers led to the proposal that $FtsK_C$ assembles on DNA from monomers in a stepwise fashion [109]. These results contrast with the demonstration that *Bacillus*

subtilis SpoIIIE, an FtsK ortholog, forms hexamers in the absence of DNA at comparable concentrations to those used here [121, 122].

As revealed by a structural analysis, three FtsK γ subdomains bind a KOPS sequence [123]. Hence, it seems likely that interaction of a single FtsK trimer with KOPS will allow the engagement of the three γ domains with DNA, supporting the physiological relevance of the transient association time, 0.84 ± 0.07 s, which we determined for a single trimer and DNA. The onset of hexamer translocation happens on a similar timescale ($\sim$0.25 s, Figure 4.17A), indicating that assembly of a second trimer into a functional hexamer and subsequent translocation is rapid, and will be dependent on the local concentration of $FtsK_C$ in the cell. A short initial association may serve to prevent $FtsK_C$ assembling and translocating on DNA when its local concentration is low, which is the case during most of the cell cycle. In vivo, FtsK translocation occurs only when its concentration is high, at an invaginating septum [114, 124],

Previous TPM and optical trapping experiments have required loop extrusion by $FtsK_C$ to observe its action at the single molecule level. In some of these experiments, large visible $FtsK_C$ aggregates were observed to translocate at the same rate as species observed using loop extrusion in optical trapping [101]. It has been suggested that loop extrusion arises from multiple points of contact on the same $FtsK_C$, and that reversals by $FtsK_C$ require that it exists as more than a 'single unidirectional motor' [64, 101, 111]. Our work establishes that $FtsK_C$ assembles into a single hexamer before translocating, does not usually extrude loops of DNA, and dissociates after it encounters XerCD-*dif*. Hence, we suggest that the loop extrusion and reversals observed in TPM [58] and optical trapping experiments [101, 110] may have been introduced by the surface or by protein aggregates.

4.4.2 $FtsK_C$ resides briefly at XerCD-*dif* before dissociating

When $FtsK_C$ encountered XerCD bound to *dif*, it resided there with a lifetime of 0.5 s to 1 s, irrespective of whether it encountered an un-synapsed or synapsed XerCD-*dif* complex

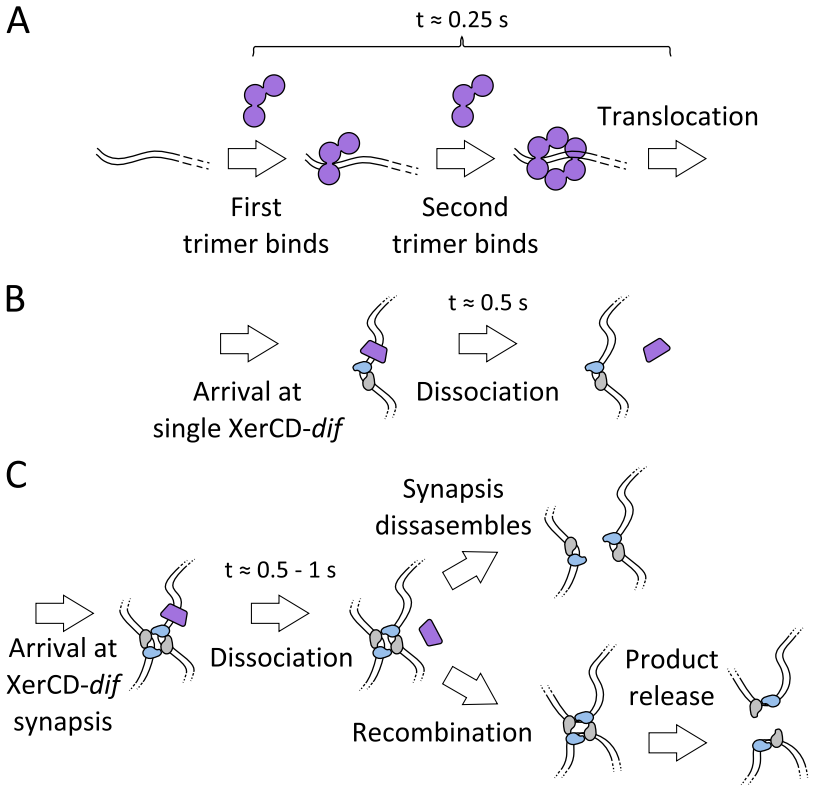

Figure 4.17: $FtsK_C$ assembly, stoichiometry, translocation and activation of recombination. **(A)** $FtsK_C$ assembles stepwise on DNA into a single hexamer which begins translocating around 0.25 s after the first binding. **(B)** When $FtsK_C$ arrives at a single *dif*-site loaded with XerCD it resides for around 0.5 s and then dissociates. **(C)** $FtsK_C$ arrives at a preformed XerCD-*dif* synapsis, resides for 0.5 s to 1 s and then dissociates. Frequently, the synapsis then disassembles, but occasionally recombination is activated.

and independently of whether the encounter led to recombination (Figure 4.17B and C). This pause time is consistent with that observed following collisions of QD labelled $FtsK_C$ trimers with XerCD-*dif* or XerD-*dif* in DNA curtain studies [113]. However, the outcome of the encounters we observed was different. While 80 % of encounters of QD labelled $FtsK_C$ with XerCD-*dif* led to reversals of translocation direction, in our experiments, more than 99 % of encounters of Cy5-FtsK led to dissociation. A possible explanation is that the Cy3B label, positioned 1 bp away from the surface distal XerD binding site, influenced the behavior of $FtsK_C$ in our experiments. However, in the absence of XerCD, the translocation of Cy5-FtsK was not impeded by the Cy3B linkage on DNA, and when Cy5-FtsK encountered another obstruction, the biotin-neutravidin used for surface attachment, we did not observe reversals. In the work described in Chapter 3, the activation of recombination was no less efficient when a fluorophore was placed in the position used here. It is also possible that our assay would not be able to resolve a rapid translocation away from XerCD-*dif*, however, using the time between the disappearance of FRET and the disappearance of fluorescence marking the disassociation of Cy5-FtsK, we have established that the average time of 0.08 ± 0.02 s that $FtsK_C$ appeared to remain on DNA after leaving the XerCD-*dif* complex was too short for it to have translocated off the 'free end' of our DNA (which would have taken 0.18 ± 0.02 s at the measured translocation velocity). Since, in our experiments, we had no more than one trimer of $FtsK_C$ per Cy5 fluorophore, we suggest that QD labelled $FtsK_C$ could have had other $FtsK_C$ trimers, attached to the same fluorescent probe, which might have assembled on DNA during residence at XerD. Since assembly before translocation took $\sim$0.25 s, and since the residence time at XerCD-*dif* was $\sim$0.5 s, there is sufficient time for the assembly of a second hexamer, attached to the same QD, while the first resides at XerCD-*dif*.

4.4.3 FtsK$_C$ activates a single round of recombination before dissociating

Using single colour TFM-FRET, in Chapter 3, we previously determined that FtsK$_C$ activates recombination by inducing the remodeling of pre-existing XerCD-*dif* synapses. In those experiments, we deduced the fate of FtsK$_C$ after activation by following the disappearance of PIFE, which resulted from the close proximity of FtsK$_C$ to the recombining complex. The PIFE signal was only observed up to isomerisation of the HJ, which was generated by XerD-mediated DNA strand exchanges, which initiate FtsK-dependent recombination. This disappearance of PIFE, however, can not determine FtsK$_C$ dissociation from the complex, since the HJ isomerization could also lead to a change in conformation, which in turn could lead to the end of PIFE, since PIFE is only sensitive to distance changes in the 1 nm to 2 nm range [78]. Here, using a labelled trimer of FtsK$_C$, we have followed precisely the presence of FtsK$_C$ at the recombining complex. Our results show that predominantly, FtsK$_C$ dissociated 0.5 s to 1 s after arriving at XerCD-*dif*. Since HJ formation and resolution by XerCD takes an average of 1.6 s (Figure 3.18), it seems unlikely that FtsK can activate subsequent rounds of recombination without dissociating, rebinding DNA and encountering the XerCD-*dif* complex de novo. Therefore, multiple rounds of recombination would have to be mediated by multiple arrivals of FtsK$_C$ hexamers at XerCD-*dif*. The rapid dissociation of FtsK$_C$, observed once it encounters XerCD-*dif*, may prevent an FtsK$_C$ hexamer from blocking access to XerCD-*dif* during these subsequent activation attempts by other FtsK$_C$ hexamers.

The vast majority of the events where FtsK successfully activated XerCD-*dif* recombination were not accompanied by any evidence of translocation-mediated looping. This provides more evidence to support the conclusion that FtsK remodels existing initial synapses, rather than extruding a loop and forming an active synapsis or otherwise promoting active synapse assembly as suggested by TPM experiments utilizing a purified γ subdomain [64, 125, 126].

Most frequently, when $FtsK_C$ encountered an XerCD-*dif* synapsis, recombination was not activated. If FtsK behaves in a similar fashion in vivo, then this would serve as regulatory mechanism to help ensure it resolves, rather than produces, chromosome dimers. Only when a synapsis reformed several times, because chromosome topology was preventing segregation, would recombination, on average, be activated.

4.4.4 Two-colour TFM-FRET follows two effective lengths along the same DNA

The extension of TFM, using two spectrally distinct fluorophores to simultaneously track two positions along the same DNA, has allowed us to correlate large scale DNA conformation with the behavior of $FtsK_C$ as it assembled, translocated and interacted with XerCD-*dif*. This method, performed on a standard-fluorescence microscope, provides a blueprint for single-molecule experiments with increasing bandwidth, allowing the study of increasingly complicated protein-nucleic acid systems. This work has encoded information in the intensities, widths and positions of the images of two fluorophores, realizing the possibility of six simultaneous fluorescence observables (increased further with the use of ALEX). TFM-FRET can be extended with the introduction of a third spectral channel [22, 24] and a third excitation laser, which would allow up to three simultaneous FRET distances, co-localisation of three binding partners and the simultaneous measurement of three distances along the same DNA molecule.

4.5 Conclusions

Here, I comment on the validity and applicability of the conclusions of my work.

Translocation velocity will be dependent on the particular protein preparation used and the particular ATP concentration, so the measured translocation velocity must be considered with these caveats. Fits to dwell times appear to overestimate the data for

$FtsK_C$ dissociation at short dwell times. This suggests that the dissocation process is not single exponential and may involve a multi-step mechanism as $FtsK_C$ recognises XerCD-*dif* and attempts to activate recombination. There is also the second exponential process in the dwell time fits and we are operating in vitro, hence, the rate parameters Table 4.1 should not be considered definitive, but are useful for giving relative timescales and order of magnitude estimates of the rates in vivo. There is the confounding effect of the photobleaching intermediate (Figure 4.3B), but we note that the photobleaching lifetime is much longer than any measured dwell and so should only be rarely present during data acquisition (Figure 4.3C).

Considered individually, the evidence for $FtsK_C$ functioning as a single hexamer from the intensity increase before translocation (Figure 4.9) and the distribution of intensities during FRET (Figure 4.11) is relatively weak, however, taken together combined with the lack of reversals, these data support a self consistent picture of a single hexamer that is capable of uni-directional translocation. Intensity correction relies on the assumption that the Cy5-DNA (used for the intensity calibration) and Cy5-FtsK are at the same position on the surface despite different surface attachment strategies. It also relies on a linear approximation for the relationship between Cy5 position and DNA length, however the correction is small for the product DNA, and hence the stoichiometry data from this conformation is more reliable (Figure 4.11A). We note that the predicted intensity distributions are not fits to data, and have no free parameters, strengthening the support for a single hexamer model. A Gaussian distribution of intensities is plausible given the intensity distribution from surface immobilised Cy5-FtsK (Figure 4.3D).

The FRET events (Figure 4.16) could be attributed to a specific orientation of $FtsK_C$ with respect to the synaptic complex, but since they do not always accompany arrivals at XerCD-*dif* (I assume in part due to their short duration), I have not commented on a physiological interpretation of them. Finally, using binding as a proxy for $FtsK_C$ activation of XerCD-*dif* recombination weakens the conclusion that FtsK dwells for only 0.5 s to 1 s

after arrival. However, it is evident from the data that FtsK usually does not dwell for tens of seconds, suggesting that many rounds of recombination do not get activated by a single $FtsK_C$.

4.6 Contributions

The work in this chapter has formed the basis of a publication [3]. Figures and text have been adapted from this publication, and were produced for the publication by myself. The project was supervised, and the text in the publication edited, by David Sherratt, Lidia Arciszewska and Achillefs Kapanidis. Pawel Zawadzki purified proteins and performed mutagenesis. I performed all experiments and produced the DNA substrate. I segmented and extracted events from timetraces, interpreted the data, and devised analysis methods (except where procedures have been cited) and performed all of the steps in the analysis.

4.7 Fits to dwell times

The fits to dwell times, described in Section 4.2.7, are presented in Table 4.1.

Table 4.1: Maximum likelihood fits to dwell times presented in figures

Description	a (%)	Δa (%)	t_1 (s)	Δt_1 (s)	t_2 (s)	Δt_2 (s)
Translocation on substrate without XerCD						
Assembly	80	+20, -30	0.26	0.12	1.4	+1.3, -0.6
Residence	88	5	2.1	0.2	14	4
Translocation toward single *dif*						
Assembly	90	6	0.25	0.03	2.7	0.7
Residence	95	2	0.50	0.02	3.5	0.6
Translocation toward synapsis						
Assembly	77	12	0.18	0.04	1.6	0.6
Residence	92	9	0.50	0.08	1.9	0.5
Translocation on product of recombination						
Assembly	82	8	0.34	0.05	2.9	0.9
Residence	97	2	0.45	0.03	2.2	0.5
Recombination						
Narrow FIW duration	60	6	13	1	56	5
$FtsK_C$ dwell	78	9	0.9	0.2	0.9	0.2
FRET duration						
Substrate	90	3	0.30	0.04	17	6
Product	91	4	0.23	0.02	1.7	0.4
Dwell after FRET before dissociation						
Substrate	90	5	0.09	0.02	6	4
Product	92	4	0.07	0.01	5	2

Errors are estimated as the standard deviation of the parameter fit, using bootstrapping with 100 resamples.

Chapter 5

Increasing the bandwidth of TFM-FRET

5.1 Introduction

Having characterised tethered fluorophore motion (TFM) in Chapter 2, used it in combination with Förster resonance energy transfer (FRET) and protein induced fluorescence enhancement in Chapter 3 and expanded it to utilise two spectrally distinct TFM reporters in Chapter 4, here, I describe the design, construction and evaluation of a new, three-colour, fluorescence microscope, intended to permit the continued expansion of TFM as a method for studying increasingly complicated protein-nucleic acid interactions.

5.1.1 Three-colour FRET

The use of a third spectral channel on a fluorescence microscope permits the observation of another binding partner in a reaction and the simultaneous measurement of two additional FRET efficiencies, allowing two more intermolecular distances to be probed. Three-colour FRET has been used to study the folding of enzymes, such as 8-17 deoxyribozyme [127], to simultaneously watch nuclotide binding and conformational changes by heat shock

protein 90 [128], to observe whether two structural changes are coupled during translocation of substrate through the ribosome [129], and to watch the movement of single-stranded DNA binding proteins as they migrate along single-stranded DNA [130].

Early three-colour microscopes utilised Cy3-Cy5-Cy5.5, with Cy5 and Cy5.5 acting as alternative acceptors [24]. However, since the emission peak wavelengths for Cy5 and Cy5.5 only differ by 24 nm (Amersham Biosciences, Piscataway, NJ), there was a large crosstalk between the two acceptor channels [24] and it is impossible to choose an excitation wavelength that can selectively excite only one of the acceptors. More recently, the use of a blue spectral channel, utilising dyes such as Cy2, Alexa 488 and Atto 488, has allowed three-colour ALEX with reduced spectral overlap between channels [25]. In spite of this, dyes suitable for the blue channel have relatively short photobleaching lifetimes [131], red-shifted states, and, in the case of one of the most photostable, Alexa 488, multistate behaviour [132]. Hence, to overcome these limitations, three-colour FRET utilising green, red and near-infrared (NIR) channels has been developed, and implemented using confocal and prism-type TIRF microscopy with a semi-ALEX scheme utilising 532 nm and 633 nm lasers [133]. NIR dyes offer good photobleaching lifetimes at reasonable signal-to-noise ratios (SNR, defined as the magnitude of a signal divided by its standard deviation) (e.g. Cy7-COT, a stabilised Cy7 derivative can be imaged for $\sim$180 s at SNR $>$ 7 [134]). NIR dyes also offer improved spectral separation from red dyes and green dyes, compared to blue dyes (Figure 5.1).

5.2 Microscope design

In order to be able to image single biomolecules, labelled with a combination of three spectrally distinct fluorescent dyes, I designed and constructed a three-colour ALEX-FRET microscope. As a guiding principle, I aimed to make the use of the microscope as user-friendly as possible, with all controls integrated in a single graphical user interface (GUI),

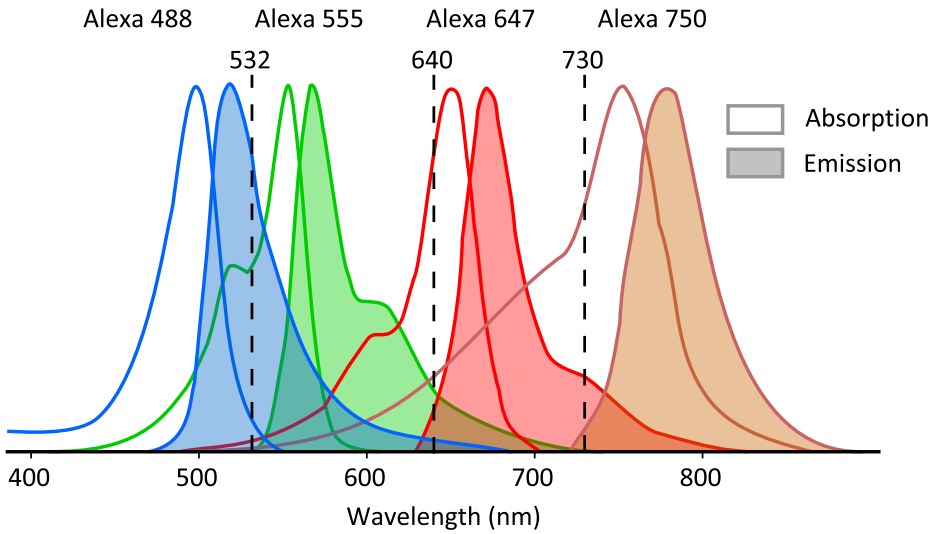

Figure 5.1: Fluorescence spectra of typical blue, green, red and NIR dyes. Data taken from the Life Technologies website (http://www.lifetechnologies.com). Dashed lines indicate excitation wavelengths chosen.

and with all controls and features accessible without the need to physically interact with any components. I planned that all imaging optics would be fully enclosed so that a user would be able to operate the microscope in a well-lit laboratory and to protect components from accidental misalignment. I chose to implement TIRF using the objective-type, rather than the prism-type, configuration to ensure that sample preparation and loading, and microscope operation was as convenient as possible and consistent with other microscopes in my research group (Figure 5.2). Appendix B contains technical drawings of components manufactured for use on the microscope.

5.2.1 Components

A green (532 nm, MGL-III-532-300 mW, CNI, Changchun, China), red (640 nm, Coherent Cube, Coherent, Santa Clara, USA) and NIR laser (730 nm, PMT40 (730-50) G36/A436, PTI, Alexander, USA) were combined using two dichroic mirrors (552 nm long-pass, Semrock, Rochester, USA, and 685 nm long-pass, Chroma, Bellows Falls, USA) and coupled into a single mode optical fibre (Schäfter und Kirchoff, Hamburg, Germany). Laser line filters (532 nm, 640 nm and 730 nm, Thorlabs, Newton, USA) were placed in

each laser beam path to ensure that only the desired wavelength was coupled into the fibre. ALEX was implemented by directly modulating the red and NIR lasers, and using an acousto-optic modulator (AOM) (Isomet, Springfield, USA) to modulate the green laser. Synchronisation was achieved using using a timing board with digital and analogue outputs (PCIe-6351, National Instruments, Austin, USA).

The output of the fibre was collimated using cylindrical lenses, focussed into the back focal plane of the objective (60× PlanApo N NA 1.45 $f/26.5$, Olympus, Tokyo, Japan), and displaced perpendicular to the optical axis such that laser light was incident at the slide-solution interface at greater than the critical angle, creating an evanescent illumination field. Fluorescence light was collected by the objective and separated from the excitation light by a dichroic (405 nm/488 nm/532 nm/640 nm/730 nm, Chroma) and a cleanup filter (405 nm/488 nm/532 nm/640 nm/730 nm, Chroma). Emission signal was focussed on a rectangular silt to crop the image and spectrally separated, using dichroic mirrors (647 nm long-pass and 760 nm long-pass, Chroma), into three emission channels and focussed side by side on an electron multiplying charge coupled device (EMCCD) camera (iXon 3, Andor, Belfast, UK).

Lenses in the emission path were anti-reflection coated (650 nm-1050 nm, Thorlabs). Correction lenses were inserted in the red ($f = 10\,000$ mm, anti-reflection coated 640 nm-730 nm, Eksma Optics, Vilnius, Lithuania) and NIR channels ($f = 4000$ mm, anti-reflection coated 730 nm-900 nm, Eksma Optics) to correct for chromatic aberration between channels. Each pixel on the EMCCD corresponded to a 162 nm × 162 nm region in the imaging plane. An autofocus system was implemented by placing a 70:30 beamsplitter (Thorlabs) between the fibre collimation lenses and the TIRF lens, and almost focussing the back reflection of the excitation lasers onto a complementary metal-oxide semiconductor (CMOS) camera (Thorlabs).

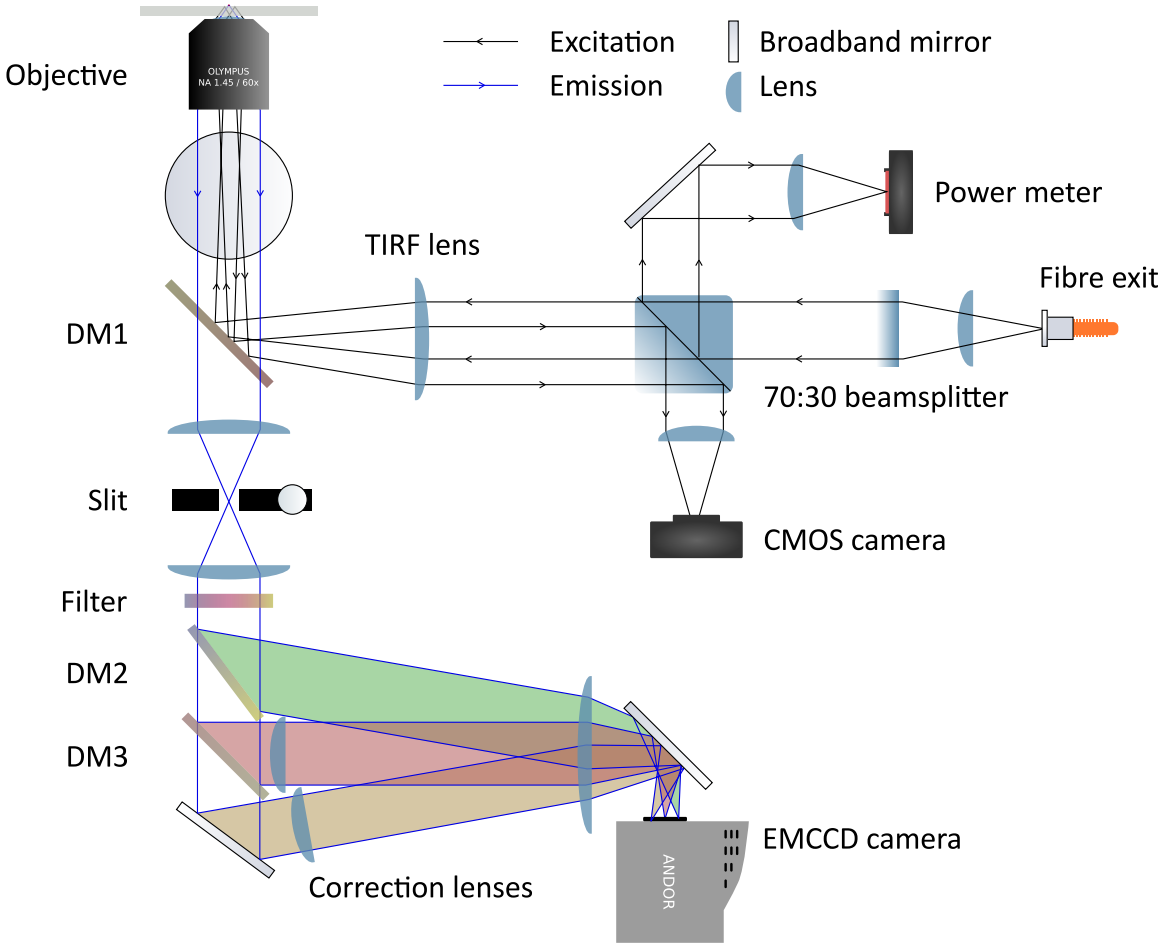

Figure 5.2: Layout of components. DM1 is a custom 405 nm/488 nm/532 nm/640 nm/730 nm dichroic mirror, and the filter used matches it. DM2 is a 650 nm long-pass dichroic and DM3 is a 760 nm long-pass dichroic. The two lenses after the fibre exit are elliptical and together shape and collimate the beam into an ellipse with a 2:1 axis ratio. The beamsplitter and the CMOS camera form a closed-loop automatic focus system that maintains a fixed distance between the objective and the sample. Correction lenses were introduced in the red and NIR channels to correct for chromatic aberration in the imaging path.

5.2.2 Features

The microscope extends ALEX by introducing a third excitation laser, with a wavelength of 730 nm, chosen to directly excite NIR fluorophores, such as Cy7 and Alexa 750, whilst not exciting red or green fluorophores (Figure 5.1). This feature allows the sorting of molecules based on their stoichiometry, i.e. which fluorophore labels they carry. Low FRET in all channels can be distinguished from an absent FRET acceptor (due to photobleaching or incomplete labelling) by directly exciting the FRET acceptor to check its presence. This extends the range of usable FRET. It also permits TFM in all three spectral channels, since TFM relies on the direct excitation of the TFM reporter fluorophore.

The microscope has a total magnification of 99×, which corresponds to a pixel size of 162 nm in the imaging plane. In the red channel, given an approximate emission wavelength of 700 nm and an NA of 1.45, a diffraction limited point emitter would have an image width of 101 nm. Hence, the chosen pixel size is of the order of the image width, which roughly optimises the signal to noise ratio in single-particle tracking fluorescence [39]. As discussed in Chapter 2, for DNA lengths around 1 kb, the resolution of TFM is relatively insensitive to pixel size, up a pixel size of around 300 nm. Therefore, the chosen pixel size represents a good compromise between TFM resolution, and size of field of view (FOV), which determines the maximum number of molecules which can be imaged simultaneously.

The red laser output power was controlled via RS-232 serial port. In order to reduce the number of moving parts, and to permit the inclusion of green laser illumination intensity control from within the GUI, the green laser was modulated using an AOM and an analogue radio-frequency (RF) driver. An AOM consists of a piezoelectric substrate that oscillates at high frequency (80 MHz), producing sound waves in an attached optical medium. Light is diffracted by the periodic refractive index modulation, and the first order of the resulting diffraction pattern can be coupled into an optical fibre. By changing the RF driving amplitude, the proportion of the diffracted power in the first order can be controlled, with a rise time of ∼10 ns, allowing ALEX modulation of the green laser, and control of its

intensity, to be implemented with user control via software.

Sample position was controlled using three closed-loop piezoelectric stepper motors (LPS-45, PI miCos, Eschbach, Germany), mounted with their directions of travel orthogonal to each other using invar (a nickel-iron alloy) mounts (Figures 5.3 and B.1 to B.3). Invar is known for its ultra-low coefficient of thermal expansion; hence, it was chosen for the stage mounts and the objective lens holder to minimise the effect of temperature fluctuations on stability. A separate closed-loop piezoelectric stepper motor was used for mounting the TIRF lens and DM1 (Figure 5.2) using aluminium mounts (Figures B.4 and B.5). This allowed the TIRF angle to be controlled precisely from the GUI, without the need to physically interact with any components within the microscope enclosure. It also permitted precise re-adjustment of TIRF angle to a previously set value. This is useful when illuminating long DNAs, where a shallower TIRF field increases the average excitation intensity experienced by the fluorophore. In addition to remote adjustment of the TIRF angle, the microscope design utilised a beamsplitter, introduced after the fibre collimation optics in the excitation path (Figure 5.2), to offer remote measurement of the illumination intensity without the need to interact with any components. So the microscope could be used with ambient lights on in the laboratory, I designed and constructed a light-tight enclosure using a modular rail system (Thorlabs), and a light-tight objective and sample enclosure by sheet-bending aluminium (Figure 5.3). All data in this chapter were taken with normal room lighting conditions.

5.2.3 Automatic focus

To maintain the same focal position, I implemented an automatic focussing feedback system, which relied on maintaining the axial position of the back reflection (from the glass-sample medium interface) of the illumination lasers. This was necessary because even a small ($\sim$1 %) misalignment of the sample, from the horizontal, introduces a significant ($\sim$1 µm) defocussing when the sample is translated to a new FOV (full FOV = 80 µm).

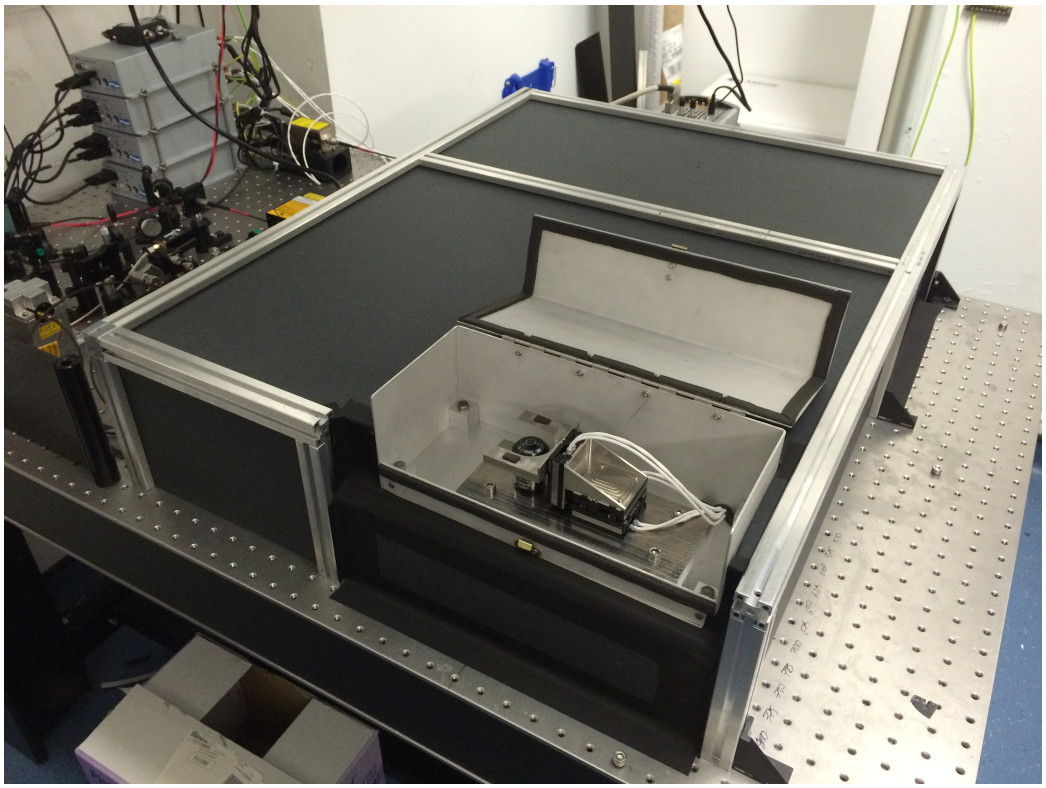

Figure 5.3: A photograph of the microscope enclosure. A custom built, light-tight, enclosure permits the use of the microscope in a brightly lit room. The sample holder and objective are visible in the middle of the image. The piezoelectric translation stages were mounted to the sample holder and to the objective with custom designed invar mounts, used because of its ultra-low coefficient of thermal expansion.

The introduction of autofocus means that a user can translate the sample without needing to adjust the focus. It should also make the microscope robust to accidental interference with the sample or stage, such as during sample loading, and limit the amount of long-term drift in the focal plane.

A CMOS camera, behind a beamsplitter, was used to monitor the position of the back reflection (Figure 5.2). If the distance between the sample and the objective changed, the back reflection would translate left or right (as seen by the camera). By comparing the centroid position of the back reflection, averaged for more than 0.5 s, ensuring it sampled a whole number of ALEX cycles (since each of the lasers had a slightly different back reflection position, likely due to chromatic aberration), with a reference centroid position, the stage was moved an appropriate amount to maintain a constant centroid position. To set a reference focus, once illumination conditions have been set, a user finds the focal position they desire, and then clicks a button on the GUI. A reference (using the same averaging scheme as during operation) is taken, and then a calibration between the change in centroid position and focus stage position is generated by displacing the sample by 500 nm, waiting for twice the averaging period, and calculating the change in centroid position. When the system is engaged, the TIRF control and the focus stage control are not accessible, since these would reposition the back reflection.

5.2.4 Correcting chromatic aberration

Since the design of the microscope uses one EMCCD camera for all three emission channels, and because the emission wavelength varies from ∼550 nm to ∼900 nm across the channels, there is some defocussing introduced between them. To overcome this defocussing, I introduced correction lenses into two of the three channels (Figure 5.2). Given that, without correction lenses, the red and NIR channels were defocussed by 0.6 µm and 1.2 µm (in the sample plane), relative to the green channel focus, I calculated the focal length required for the correction in each channel. The thin lens equation is as follows:

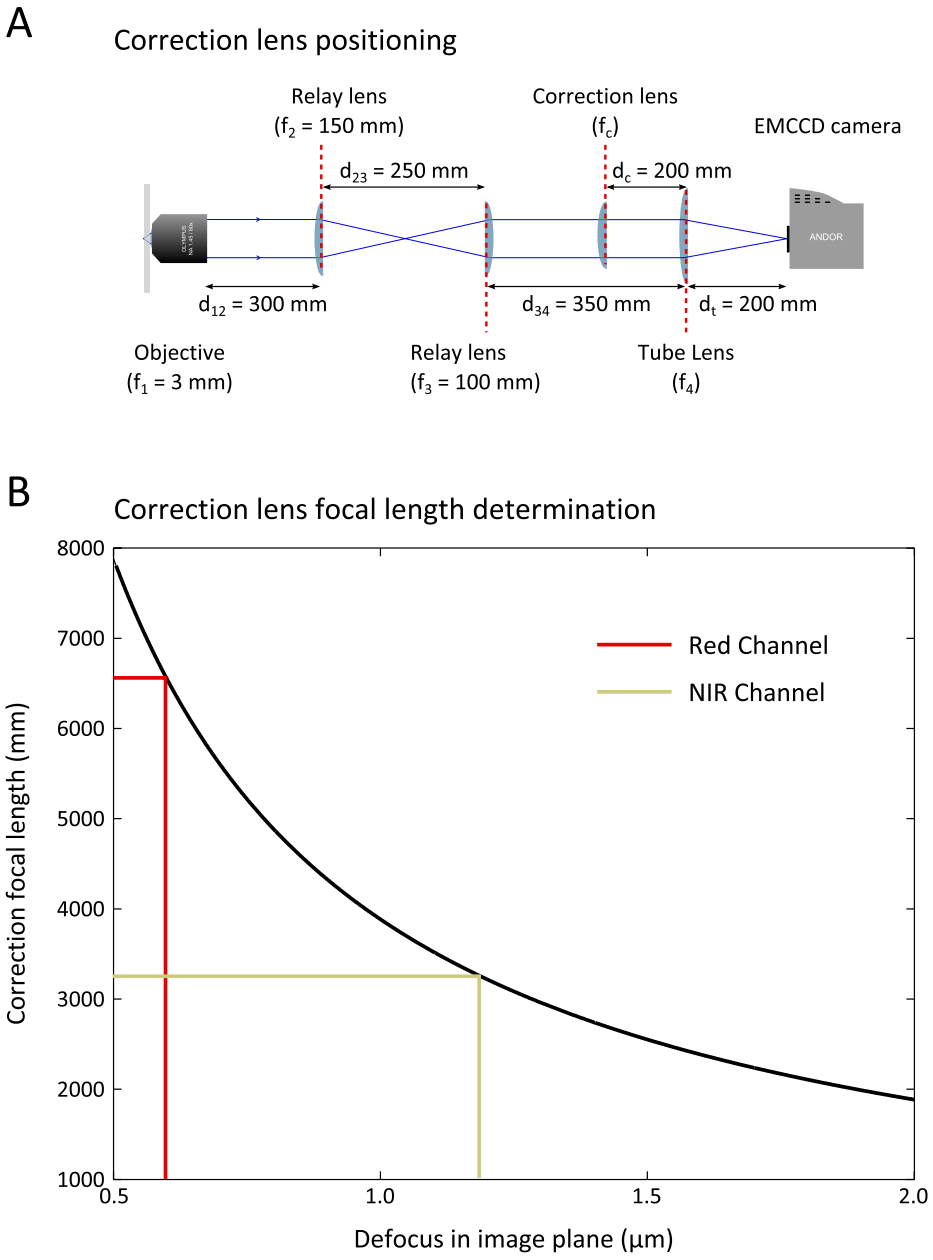

Figure 5.4: Correction of chromatic aberration. **(A)** Lens positions and focal lengths in the imaging path. The tube lens is positioned 200 mm from the camera chip. A correction lens was added to the red and NIR channels between the second relay lens and the camera tube lens to focus all channels to the same plane. **(B)** Choosing the correct correction focal length. The appropriate correction lens for each channel can be determined using Equation (5.8)). Before introducing correction lenses, in the sample plane, the red channel focus was 0.6 µm above the green channel, and the NIR channel focus was 1.2 µm in the same direction.

$$\frac{1}{u} + \frac{1}{v} = \frac{1}{f} \tag{5.1}$$

where u is the distance from the lens to the object, v is the distance from the lens to the image and f is the focal length of the lens. If a positive u points to the left, then a positive v points to the right and vice versa. I define f_1, f_2, f_3, and f_4 to be the focal lengths of the objective lens, the first and second relay lenses, and the tube lens, and d_{12}, d_{23}, and d_{34} to be the distances between each pair of lenses (Figure 5.4A). I assume that the aberration that leads to defocussing can be treated as deviation of f_4 from its correct value (for that particular spectral channel). Hence, using Equation (5.1) and writing the distance from the objective lens to the focal plane in the sample, u_1, as:

$$u_1 = f_1 + \Delta u_1 \tag{5.2}$$

where Δu_1 is the defocussing in the sample plane, we can write that the distance from the objective to the image it forms, v_1, is given by:

$$v_1 = \left(\frac{1}{f_1} - \frac{1}{u_1}\right)^{-1} \tag{5.3}$$

Simliarly, since $u_2 = d_{12} - v_1$, $u_3 = d_{23} - v_2$, and $u_4 = d_{34} - v_3$, we can write:

$$v_2 = \left(\frac{1}{f_2} - \frac{1}{d_{12} - v_1}\right)^{-1} \tag{5.4}$$

$$v_3 = \left(\frac{1}{f_3} - \frac{1}{d_{23} - v_2}\right)^{-1} \tag{5.5}$$

$$f_4 = \left(\frac{1}{d_t} + \frac{1}{d_{34} - v_3}\right)^{-1} \tag{5.6}$$

where d_t is the distance from the tube lens to the camera chip. By substituting

Equations (5.2) to (5.5) into Equation (5.6), we can generate an expression for $f_4(\Delta u_1)$, the effective tube lens focal length that describes a channel defocussed in the image plane by an amount, Δu_1. Since we want all the channels to have the same effective tube lens focal length, we can add correction lenses to the defocussed channels to produce a compound tube lens with an effective back focal length (BFL) equal to d_t. For two thin lenses separated by a distance, d_c, with focal lengths f_c and f_4, the BFL is given by:

$$\text{BFL} = \frac{f_4(d_c - f_c)}{d_c - (f_c + f_4)} \tag{5.7}$$

Setting the BFL equal to d_t, and substituting in for $f_4(\Delta u_1)$, we can rearrange Equation (5.7) to get:

$$f_c = \frac{f_4(\Delta u_1)(d_t + d_c) - d_t d_c}{f_4(\Delta u_1) - d_t} \tag{5.8}$$

This expression can be used to find the approximate focal length of lens needed to correct for a given amount of defocussing (Figure 5.4B). For the microscope described in this chapter, approximate focal lengths of 6500 mm and 3100 mm were suggested by Equation (5.8) for the red and NIR channels respectively. The experimentally determined optimum focal lengths were 10 000 mm and 4000 mm, close to the suggested values. The difference can be accounted for by the approximations made in deriving Equation (5.8), which were: assuming all the lenses were thin, and assuming the aberration can be described entirely by a shift in the tube lens focal length. Equation (5.8) is useful when determining whether it is possible to buy lenses with appropriate focal lengths, and for determining a good starting point when testing correction lenses.

5.3 Software

In this section I follow the naming convention of referring to the green dye as the donor (D), the red dye as the transmitter (T), the NIR dye as the acceptor (A), and similarly for their corresponding illumination laser [135].

5.3.1 Microscope control

The microscope is controlled from a single GUI in MATLAB[1], written in an object-oriented manner, and interfacing with many components using manufacturer software development kits (SDK) via C/C++ MATLAB executable files (called MEX files). Each of the pieces of hardware (e.g. green laser, EMCCD camera, power meter) corresponds to an object in the software. Each object is responsible for communication with the appropriate piece of hardware through its methods (e.g. the power meter object can query the current power meter reading). The use of object-oriented, rather than functional, programming should facilitate any changes to hardware made over the life cycle of the microscope: any replacement hardware just needs a software object that fulfils the same functions as the software object of the hardware it replaces. Although MATLAB is typically single-threaded, spooling to memory by the autofocus camera, and spooling to disk by the EMCCD is handled by separate threads spawned by the manufacturer SDKs. This overcomes the limitation of computational power imposed by single-threaded operation, and allows simultaneous updating of the main camera display, autofocus camera display, implementation of the autofocus feedback loop, and the spooling to disk of acquired data - all from the same user interface (Figures 5.5 and 5.6).

Upon instantiation, the microscope controller object initiates communication with the PCIe card used for synchronisation and control of the AOM, the PCI card used for communication with the EMCCD, the serial port controller for the red laser and the USB

[1] github.com/petermay

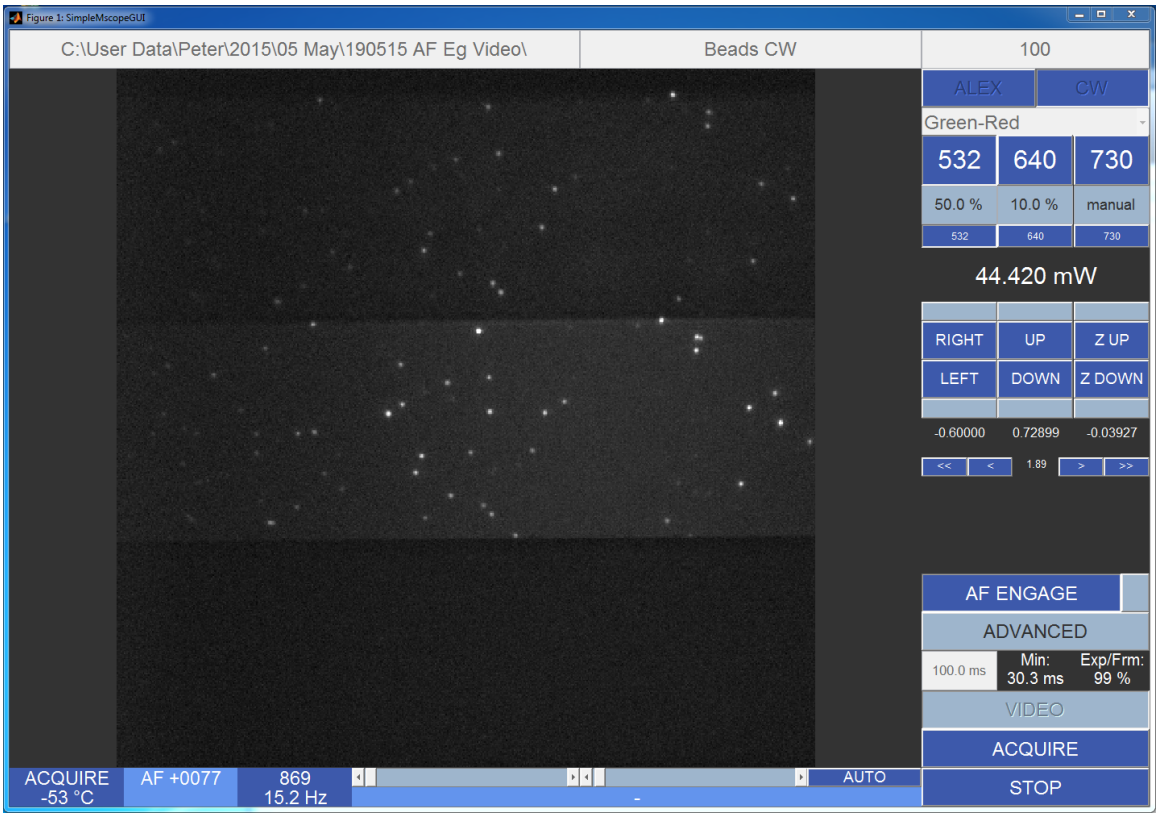

Figure 5.5: Microscope control GUI. The controls for laser power, alternation pattern, sample and TIRF lens positioning, and camera control have been integrated into one MATLAB GUI. The software displays the most recent image, the camera status and temperature, and the current power meter reading. The displayed image is of fluorescent beads, imaged using continuous wave 532 nm illumination. Advanced controls are launched in a separate window, and include EMCCD gain, camera temperature and cropping of the active EMCCD region to allow increased acquisition speeds.

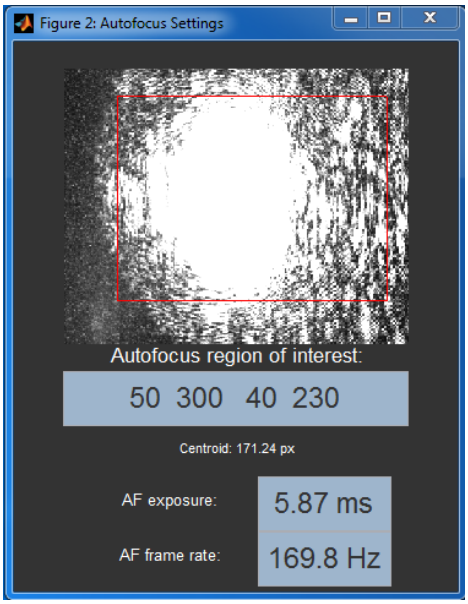

Figure 5.6: Autofocus control GUI. The displayed image is from the back reflection of the green laser from a fluorescent beads sample (in air). The average centroid, of the pixels within the region of interest, of the most recent ∼0.5 s of autofocus camera images, is maintained constant by repositioning the sample height in a closed-loop feedback system.

controllers for the four piezoelectric stages, the CMOS camera, and the power meter. It displays the last image captured by the camera and user-interface controls for every aspect of microscope function (with the exception of NIR laser output power, which is controlled using a neutral density filter). The autofocus panel is launched from the main GUI and displays the current centroid position of the back reflection and the most recently acquired image by the autofocus camera. The autofocus camera display offers a simple method of rough focus alignment: when the sample is close to in-focus, the back reflection is near the centre of the autofocus display.

5.3.2 Image registration

Having acquired three-colour fluorescence movies, we need to extract single-molecule trajectories from them. For this purpose, I wrote a suite of analysis software that automates the registration of images, analysis of movies and the display of single molecule timetraces.

The first step in image analysis is registration between the three image channels (green,

red and NIR, or D, T and A). The image of fluorescent beads, dried sparsely onto a coverslip, which fluoresce brightly in the green and red channels, can be used as a reference standard. To see the beads in the NIR channel, I used a long (∼200 ms) frame time and high (∼15 mW) red laser power. The tail of the beads' red emission (>760 nm) leaked into the NIR channel, and the image of this red leakage can be used to find the position of the beads in the NIR channel.

A user loads a movie, sets the image limits for each channel and then runs an algorithm that generates a projective two-dimensional transformation (Figure 5.7), defined as:

$$\begin{pmatrix} u & v & 1 \end{pmatrix} = \begin{pmatrix} x & y & 1 \end{pmatrix} \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \tag{5.9}$$

where x and y are coordinates in the original channel, u and v are coordinates in the new channel, and a, b, c, d, e, f, g, h, and i are coefficients to be determined. The algorithm used is as follows:

1. Generate an averaged image over all frames, and use this for all further steps.

2. Find peaks in the image (corresponding to beads) in each channel, by looping over all pixels in all channels and comparing pixels in their local background, defined as the mean pixel value the background radius (3 px) away, marking peaks as being both at least 5 digital units (DU) higher than the local background, and the highest pixel value within their background radius.

3. Refine peak localisations by fitting a Gaussian profile using ordinary least-squares estimation.

4. For the red channel, build a cost function by convolving a Gaussian of width 5 px with an array of delta-functions centred at each bead localisation.

5. Starting with the identity transformation ($a = e = i = 1$, other elements $= 0$), map the bead localisations from either of the other channels into the red channel (this implicitly matches the top right hand corner of both channels to each other, from the user-defined channel segmentation, as an initial guess).

6. Compute the maximand for this channel by summing the cost function evaluated at each of the transformed bead localisations. Maximise this by varying the coefficients of the transformation matrix, subject to the constraint that Equation (5.9) is valid.

7. Cease maximisation when a local maximum is found, or 1000 iterations are performed.

8. Repeat steps 4 to 7, using a decreasing width when generating the cost function (4 px, 3 px, 2 px, 1 px, 0.5 px) and using the transform generated from the previous optimisation as an initial guess for the current round.

9. Repeat for the other channel to generate two transformations in total: one from green to red, and the other from NIR to red.

This algorithm is robust to beads not appearing in all channels, since for the final round of optimisation, only localisations within $\sim$0.5 px of one in the other channel contribute significantly to the cost function. The effect of the transformation is displayed for the user to verify its correct operation, with histograms showing the nearest neighbour distance after the two transformations are applied (Figure 5.7, position error $<$0.5 px for both transformations). To transform coordinates from NIR to green, the NIR to red transformation has to be applied, followed by the green to red transformation in reverse.

5.3.3 Analysis in parallel on a CPU

The next steps in analysing three-colour single molecule data are to localise single molecules within movies, link them between channels, and fit their images to recover their trajectories (Figure 5.8). I used the same constrained Levenberg-Marquardt optimisation as used

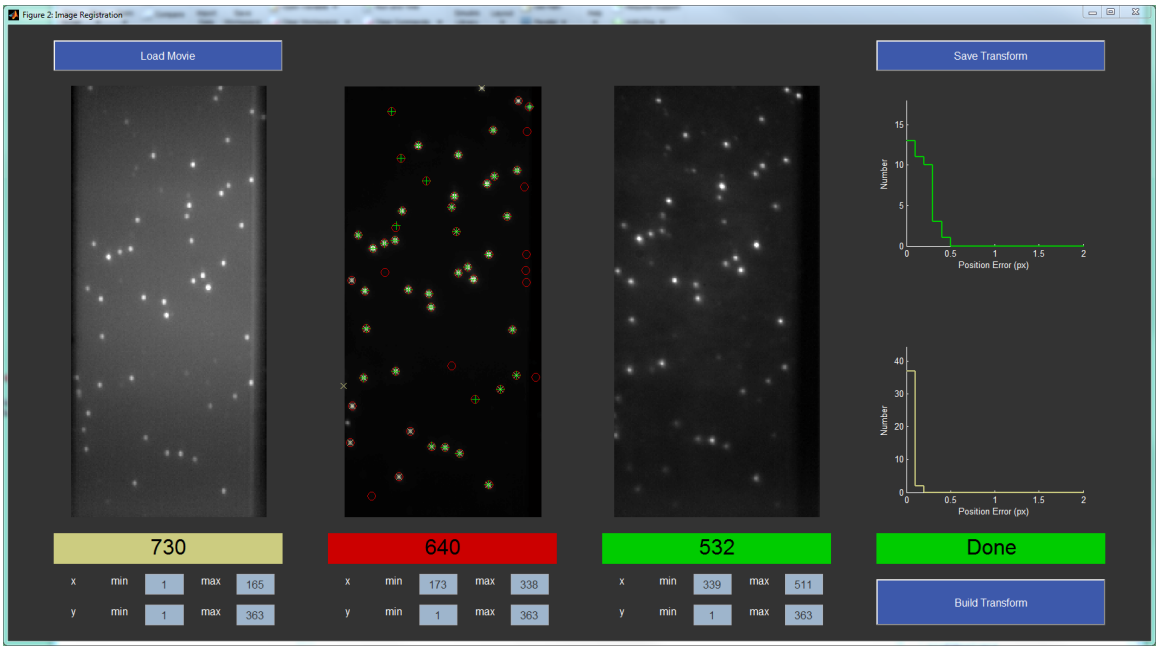

Figure 5.7: Image registration GUI. The two projective two-dimensional transformations between the three channels can be estimated using an image of fluorescent beads. Channel segmentation is defined manually and is used as an initial guess for the registration algorithm. The distance from each red localisation to its nearest neighbour in the green and NIR channels is plotted as a histogram and can be used to evaluate the accuracy of the transformation.

by TwoTone, an existing two-colour single molecule TIRF analysis package used in my research group [42], and recompiled it to run under 64 bit MATLAB. This was chosen because the algorithm and optimisation settings had already been validated in detail.

A user can load a transform (which contains the channel segmentation defined when it was created), and a movie, and then set the particle detection parameters, the channel linking parameters, and the analysis settings. Once these are set, the display can be updated and the localisations in each channels and the linking can be visually inspected and their correctness confirmed.

Movies can be analysed in parallel; this was achieved using the parallel computing toolbox in MATLAB. Using this toolbox, a parallel cluster (e.g. the local cluster consisting of 4 CPUs on a desktop computer) can be created, and jobs containing tasks to evaluate can be executed by the cluster. Jobs containing one task, defined by a static method with arguments specifying the movie to analyse, the analysis settings to use and the directory to

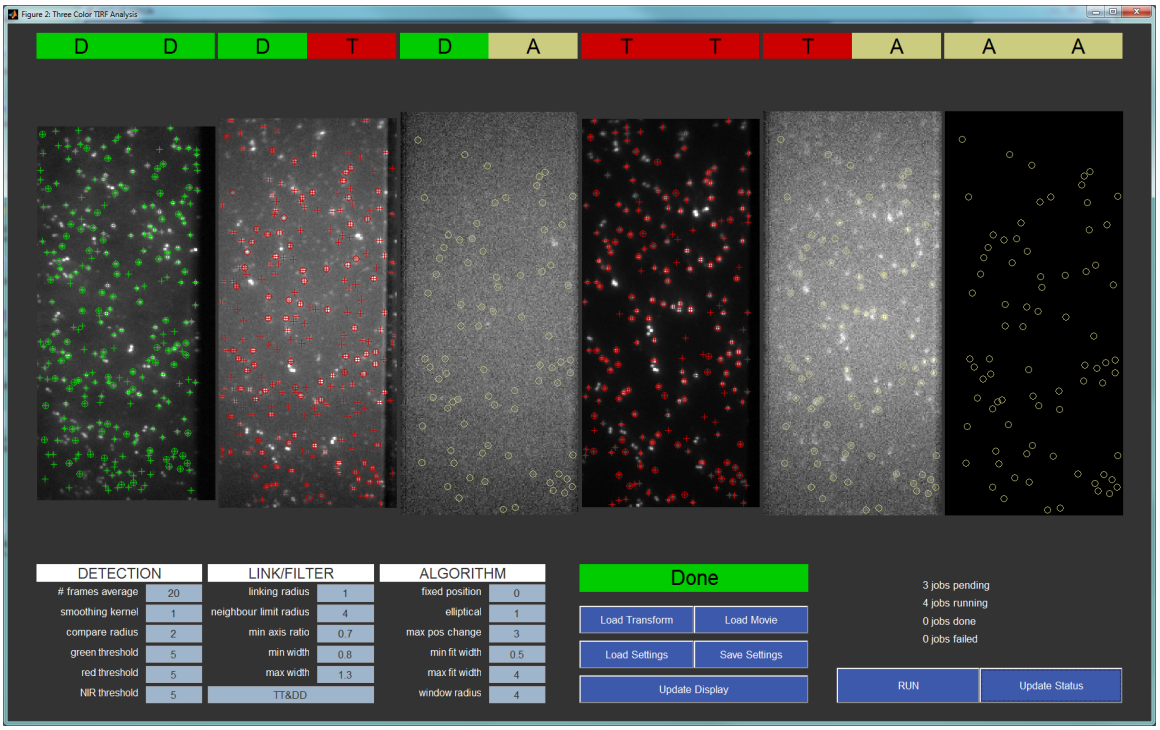

Figure 5.8: Image analysis GUI. Images are fit using constrained Levenberg-Marquardt ordinary least-squares regression in MATLAB. Peaks in images corresponding to single molecules are found using a peak-finding algorithm, then fit with an elliptical Gaussian. Molecules are filtered based on width, nearest neighbours and their axis ratio to exclude cases when multiple labelled molecules are attached too close to each other on the surface.

save to, are queued by the cluster, and are analysed independently by each worker (CPU). This method takes advantage of the independence of analysis tasks, and allows a user to define multiple analysis schemes, to be queued up for different experimental conditions, in the same analysis session.

5.3.4 Trajectory viewer

The final step in the basic analysis of three-colour TIRF movies is the viewing of the timetraces associated with each molecule. To do this, and to allow basic interactive filtering, compiling, and annotation of datasets, I wrote a simple trajectory viewer (Figure 5.9). It allows a user to choose which channels (DD, DA, DT, TT, TA, AA) to display, to navigate through the extracted region around the particle in each frame, and to navigate through all the particles in a movie. When loading analysed movie files, mulitple can be selected and then saved as a compiled dataset. In addition to the main, trajectory viewer, window, a small notebook window is opened with the option to tag particles based on up to 7 user-defined criteria. When an analysis file is saved, this information is stored, and retrieved on loading.

5.4 Evaluation

So far in this chapter, I have described the design and construction of a new three-colour TIRF microscope. I will now present some example data taken using it, and evaluate its capabilities with reference to an existing two-colour microscope in my research group.

5.4.1 Autofocus

To confirm that the autofocus feedback system can maintain a fixed focal position when the sample is defocussed as it is translated in a direction axial to the optical axis, I imaged fluorescent beads under 640 nm illumination. After manually finding an appropriate

Figure 5.9: Timetrace viewer GUI. Once analysed, three-colour single-molecule timetraces can be displayed. The user can select which channels to display and can view the region of the movie corresponding to each frame in each channel.

focus, the sample stage was translated by 5 FOV and a video was acquired (Figure 5.10). Even with this substantial defocussing, within 6 s, the system returned the sample to an appropriate focus. This confirms the suitablility of the autofocus system for general use.

5.4.2 Single molecule experiments

DNA (∼200 pM) was incubated for 60 s at the surface of a polyethylene glycol passivated coverslip, attaching by biotin-neutravidin interactions, and sealed using a silicone gasket (Grace Bio-labs, Bend, USA) and a second coverslip. Imaging was performed in a buffer consisting of 50 mM Tris-HCl (pH 7.5), 50 mM NaCl, 5 mM $MgCl_2$, 100 µg/ml BSA, and 1 mM UV-treated Trolox. An enzymatic oxygen scavenging system consisting of 1 mg/ml glucose oxidase, 40 µg/ml catalase and 1.4 % (w/v) glucose was added prior to sealing the sample before image acquisition.

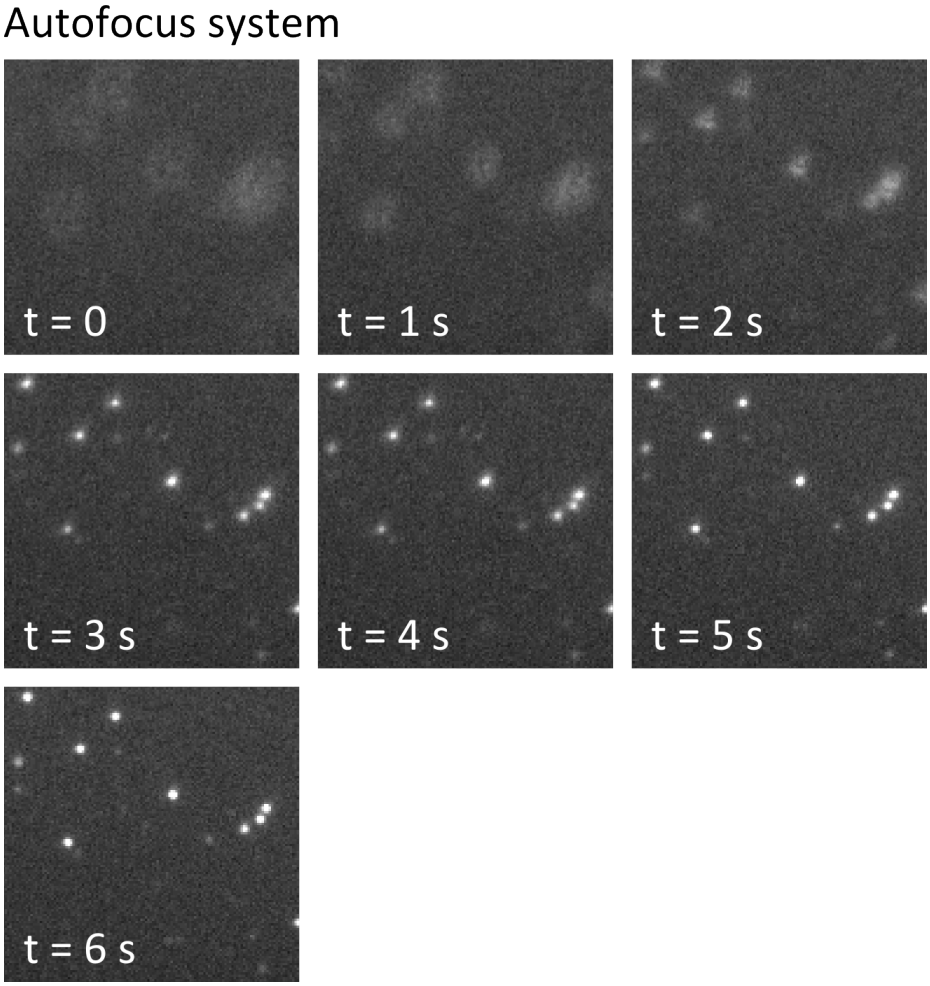

Figure 5.10: Autofocus example. A fluorescent beads sample was loaded and manually brought to focus, then the sample was repositioned axially by 400 µm, corresponding to 5 FOV. Since the sample holder is slightly inclined to the horizontal this introduces a defocus. Over the course of ∼6 s the feedback system returns the sample to the original focal plane.

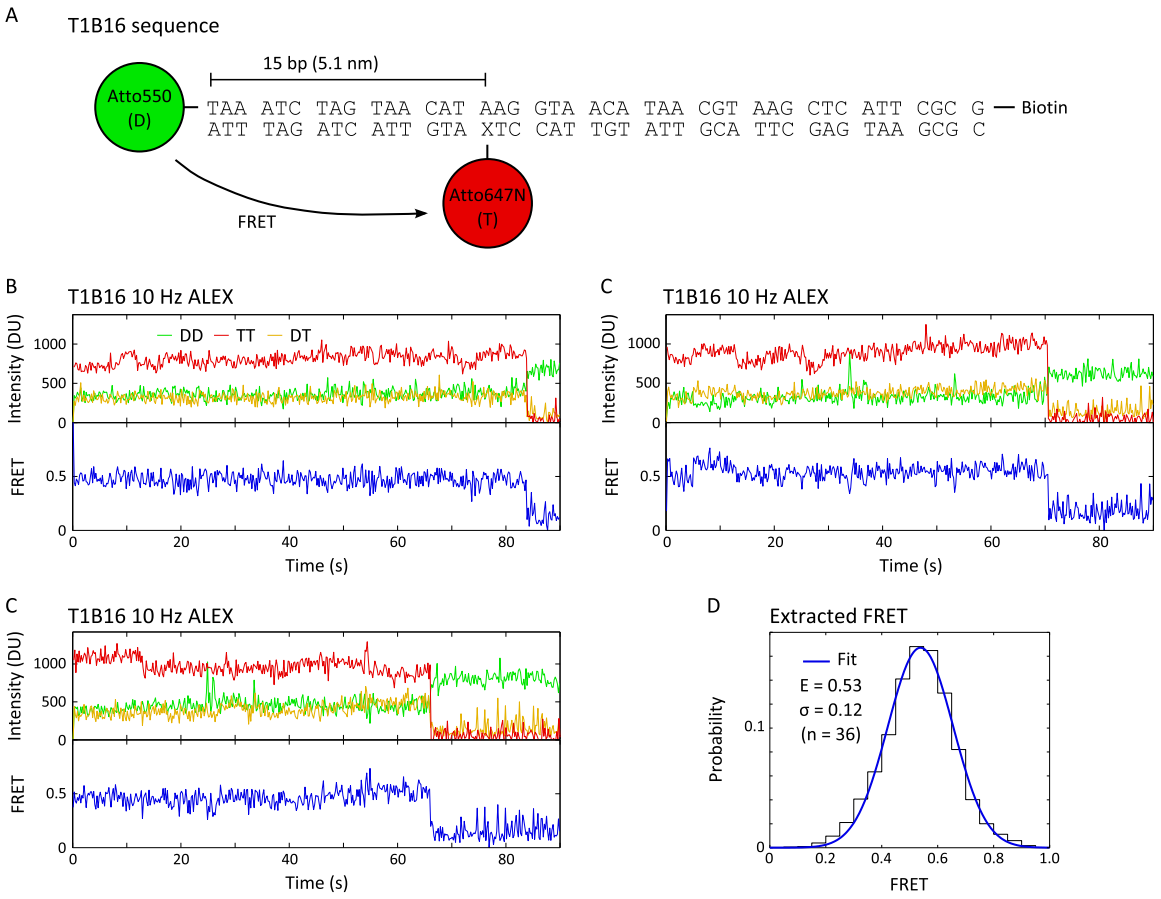

Figure 5.11: Green-red FRET. **(A)** Sequence of green-red FRET standard used. **(B)** Example timetrace showing: green emission under green excitation (DD), red emission under red excitation (TT) and red emission under green excitation (DT). Photobleaching of Atto 647N is apparent at around 85 s and an anti correlated change in DD and DT confirm the presence of FRET. **(C)** A second example timetrace, showing photobleaching at around 70 s. **(D)** A third example timetrace. Data taken using 10 Hz green-red ALEX. **(E)** Histogram of FRET from each frame of molecules that did not photobleach during the 90 s acquisition. A Gaussian is fit using ordinary least-squares regression in MATLAB to recover the position and width of the distribution.

5.4.3 Green and red channels

The experiments presented so far in this thesis have relied on the green and red fluorescence channels for all their single molecule observables. To validate that the addition of a third, NIR, channel in the microscope design had not disproportionately affected the operation of the green and red channels, I imaged a FRET, and a TFM, standard (Figures 5.11 and 5.12).

The DNA FRET standard used was labelled with Atto 550 (donor, D) and Atto 647N

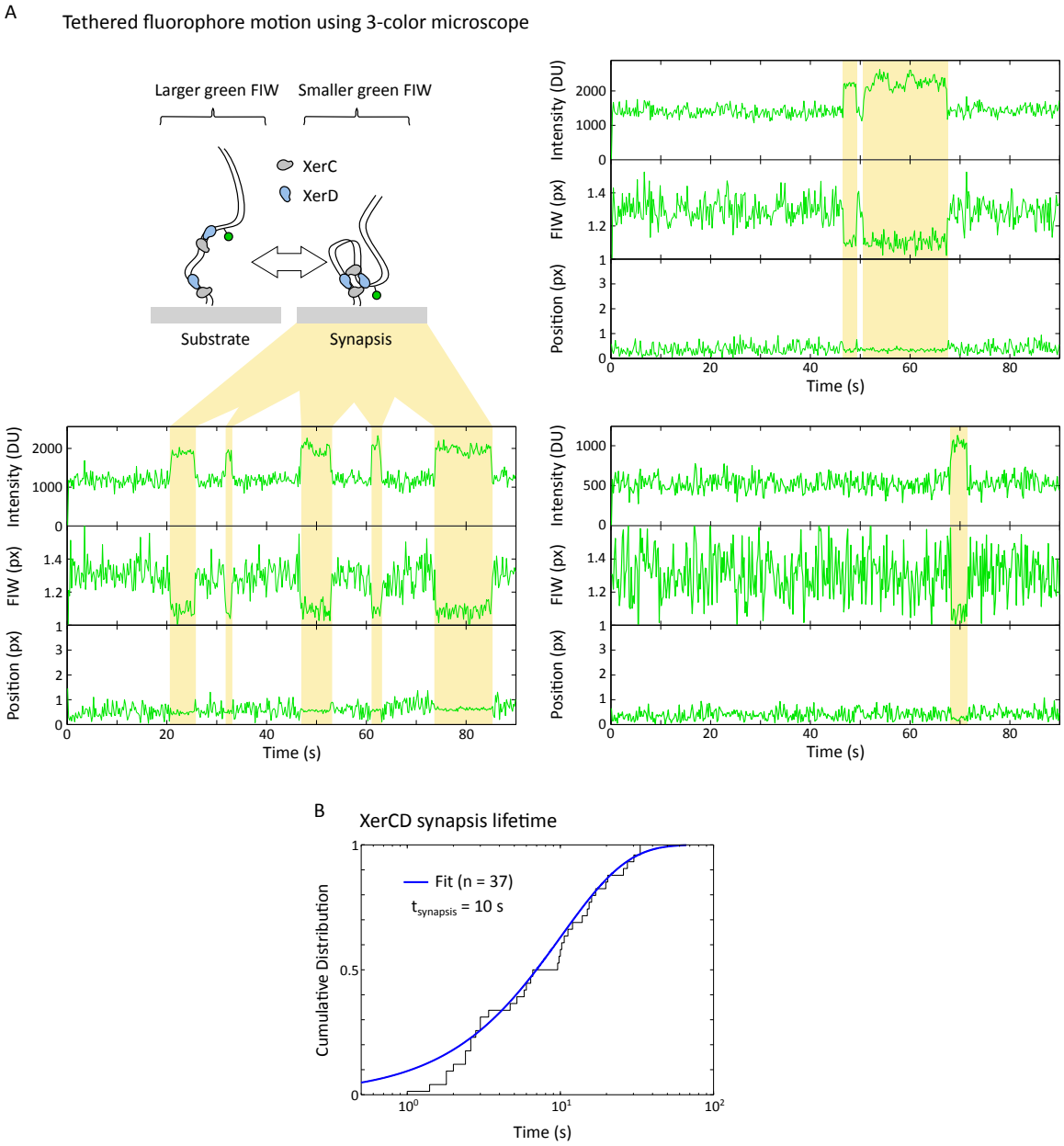

Figure 5.12: TFM example. **(A)** Cartoon and example traces. The top panels show the green intensity under green excitation, the middle panels show the FIW and the bottom panel shows the position of the localisation relative to the mean position over the first 10 frames. Synapsis, mediated by XerCD, is evident as a decreased FIW (yellow highlight). The associated intensity increase is due to the evanescent excitation field. Data taken at 5 Hz with 100 ms stroboscopic illumination. **(B)** Cumulative distribution of the lifetime of XerCD mediated synapsis. Synaptic events were manually segmented using the FIW. Maximum likelihood fitting in MATLAB was used to fit a single exponential distribution of dwell times.

(transmitter, T), red and green fluorophores (Figure 5.11A). This pair of fluorophores has an R_0 of 6.5 nm (ATTO-TEC GmbH, Siegen, Germany). The same FRET standard, labelled with Cy3B and Atto 647N, has a similar R_0 of 6.2 nm, and shows an apparent FRET efficiency of 0.53 when imaged using Halfdome (an existing two-colour FRET microscope) [42]. The timetraces obtained had relatively stable fluorescence and showed anti-correlated changes in DD and DT upon Atto 647N photobleaching, confirming the presence of FRET (Figure 5.11B-D). Selecting only portions of timetraces where both donor and transmitter were not photobleached, a histogram of apparent FRET in each frame was produced, and fit with a Gaussian (Figure 5.11D). The fit recovers that the FRET distribution is centred at 0.53, consistent with the apparent FRET obtained on Halfdome for the related DNA standard, confirming that the inclusion of a NIR channel has not significantly compromised the FRET value observed.

To confirm the suitability of the microscope design for TFM, I imaged the DNA substrate used in Chapter 4 in the presence of XerCD (2 µM) (Figure 5.12A). Synapsis was clearly evident as a reduction in FIW, accompanied by an increase in intensity due to the evanescent TIRF field. The dwell time for synapses was extracted manually and fit with a single exponential (Figure 5.12B), extracting a synapsis lifetime of 10 s, consistent with the lifetime obtained in Chapter 3. Again, this confirms that the three-colour design has not significantly compromised the ability to perform TFM.

5.4.4 NIR channel

Having looked at the performance of the microscope in the red and green channels, I now turn to the third, NIR, spectral channel. Cy7 was chosen as the fluorophore in this channel because previous NIR experiments have used it successfully [133]. The DNA was labelled with Cy5 (transmitter, T) and Cy7 (acceptor, A) (Figure 5.13A), and was immobilised to a slide surface and imaged using 10 Hz red-NIR ALEX. Stable fluorescence traces were obtained and the presence of FRET confirmed by the anti-correlated change in TT and

TA intensity upon Cy7 photobleaching (Figure 5.13B to D).

To check the crosstalks, l and d (defined in Section 3.2.8), between the NIR channel and the red channel, portions of timetraces with either a bleached transmitter or bleached acceptor were used. The leakage, l, was $5.2 \pm 0.5\,\%$ (with the error quoted as the standard deviation from 10 molecules), and the direct excitation, d, was $9 \pm 2\,\%$. The crosstalks were smaller than between the red and green channel on the two-colour microscope used in Chapters 2 to 4 (Section 3.2.8). The relative detection factor, γ, was 0.25 ± 0.06, with transmitter excitations more likely to be detected, due to the reduced quantum efficiency of the EMCCD in the NIR channel ($\sim$50 % of the efficiency in the red channel) and the lower transmittance of the objective at longer wavelengths.

For the pair of dyes used, Cy5 and Cy7, R_0, is 6.2 nm [133]. The dyes were separated by 12 bp (4.1 nm), plus their six-carbon linkers (roughly increasing the distance by 0.7 nm each), giving a total separation of $\sim$5.4 nm. Hence, the predicted FRET efficiency was 0.70. Extracting the apparent FRET efficiency from portions of timetraces where Cy5 and Cy7 have not photobleached, an efficiency of 0.40 was recovered (Figure 5.13). Using the correction factors, the accurate FRET value was 0.67 ± 0.07 (with the error quoted as the standard deviation from 10 molecules), very close to that predicted. The use of this NIR FRET standard confirms that the microscope offers the ability to resolve FRET between red and NIR fluorophores to a experimentally useful extent and with a useable observation lifetime in the NIR channel (46 s, Figure 5.13E).

5.4.5 Resolution

To compare the characteristics of the microscope to an existing two-colour microscope, Halfdome, I imaged fluorophore labelled DNA standards in each channel. The signal to noise (SNR), defined as the ratio of the mean intensity to the standard deviation of the intensity distribution from a single fluorophore before photoblinking or photobleaching, was calculated along with the photobleaching lifetime (Figure 5.14). Cy5 was imaged on both

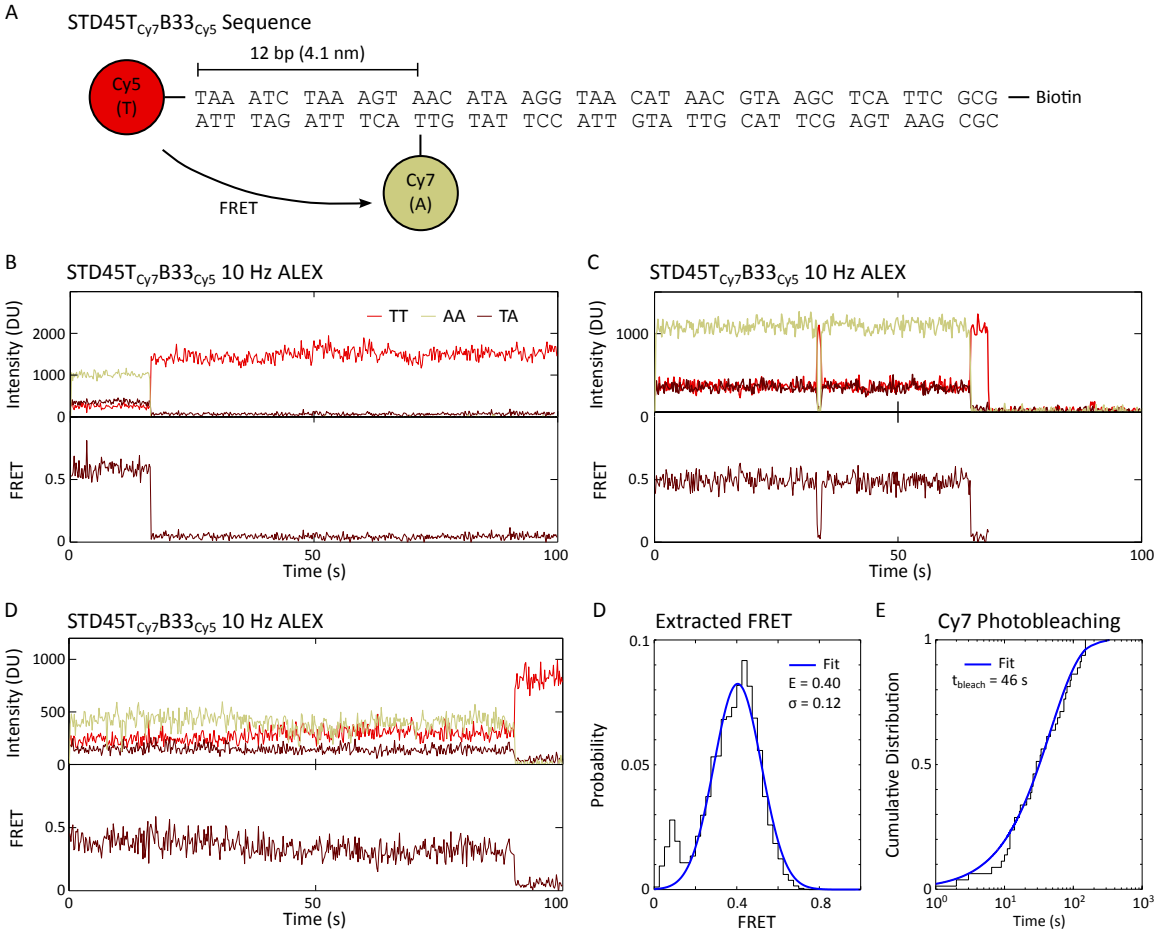

Figure 5.13: Red-NIR FRET. **(A)** Sequence of red-NIR FRET standard used. **(B)** Example timetrace showing: red emission under red excitation (TT), NIR emission under NIR excitation (AA) and NIR emission under red excitation (TA). Photobleaching of Cy7 is apparent at around 15 s and an anti correlated change in TT and TA confirm the presence of FRET. **(C)** A second example timetrace, showing photobleaching at around 70 s, and blinking of Cy7 at around 40 s **(D)** A third example timetrace. Data taken using 10 Hz red-NIR ALEX. **(E)** Histogram of FRET before photobleaching ($n = 43$ molecules). A Gaussian is fit using ordinary least-squares regression in MATLAB to recover the position and width of the distribution. **(F)** Cumulative distribution of photobleaching time ($n = 43$ molecules). A single exponential is fit using maximum likelihood estimation in MATLAB to recover the photobleaching time.

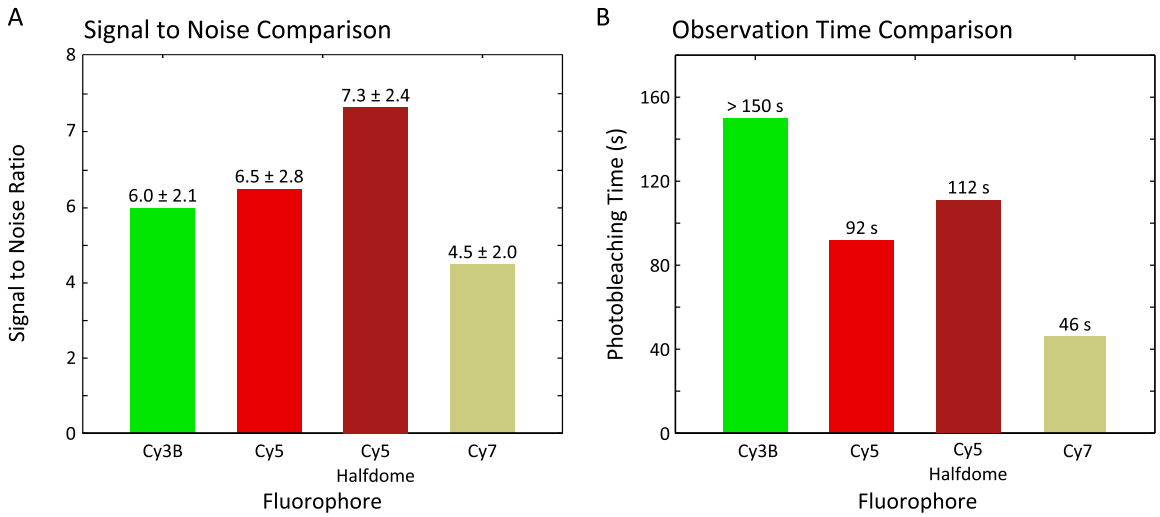

Figure 5.14: Signal to noise ratio and photobleaching time comparison. **(A)** The signal to noise ratio for three dyes (Cy3B, Cy5 and Cy7) in the three channels of the microscope were determined and compared to the signal to noise of Cy5 on an existing two-colour TIRF microscope (Halfdome) **(B)** The photobleaching lifetime at the same intensity for the dyes, in the presence of an enzymatic oxygen scavenging system consisting of 1 mg/ml glucose oxidase and 1 mM UV-treated Trolox.

the microscope described in this chapter and on Halfdome. The SNR and photobleaching lifetime (during 10 Hz ALEX) were reduced $\sim$ 15 % relative to Halfdome, but still represent reasonable values, useful for single-molecule fluorescence experiments. The reductions can be accounted for by the small cropping of the Cy5 emission spectrum, by the image splitting dichroic, in order to fit the NIR channel. At approximately the same SNR, Cy3B has a significantly longer photobleaching lifetime ($>$150 s, compared to 92 s) than Cy5, since Cy3B is a more recently developed fluorophore that has been rigidified to enhance its photophysical properties [28].

The NIR channel offers a SNR of 4.5 with an observation lifetime of 46 s using 10 Hz red-NIR ALEX. This lifetime is until a photobleaching or photoblinking event, and is during FRET, using the NIR FRET standard from the previous section. This lifetime, and SNR, are clearly sufficient for simple colocalisation, and for FRET experiments, and are possible without significantly compromising the green or red spectral channels.

5.5 Discussion

Three-colour single-molecule fluorescence offers the ability to simultaneously watch three spectrally separate labels attached to biomolecules as they perform their function. Three-colour TFM-FRET offers the ability to watch three effective lengths along the same DNA, and to simultaneously measure up to three FRET distances within any complexes formed. To realise these capabilities, I have designed, constructed and evaluated a three-colour fluorescence microscope. It offers almost the same performance as an existing two-colour microscope, in the green and red spectral channels, and adds a third, NIR, channel, with its own spectrally distinct excitation laser, which offers a relatively long observation time at a useable signal to noise ratio.

The microscope offers convenient operation, with a fully enclosed emission path and control of laser power and alternation pattern, TIRF angle and sample position, and EMCCD camera operation, all from within the same user interface. To ensure that the three spectral channels were brought to the same focus on the same EMCCD camera, I have derived an expression that facilitates choosing an appropriate focal length correction lens for each channel, and have implemented such a scheme on the microscope. Since a small misalignment of the sample holder with respect to the horizontal introduces a defocussing as the sample is translated to new fields of view, I have implemented a cost-effective ($\pounds 500$) automatic focus correction system.

In addition to this, I have also developed a suite of three-colour single molecule analysis software, suitable for FRET, colocalisation and TFM experiments. I have developed a simple robust algorithm for image registration that requires minimal user interaction, and have implemented parallel analysis of acquired data. I hope that this software will be useful for future users of the microscope described here, and any other three-colour microscopes.

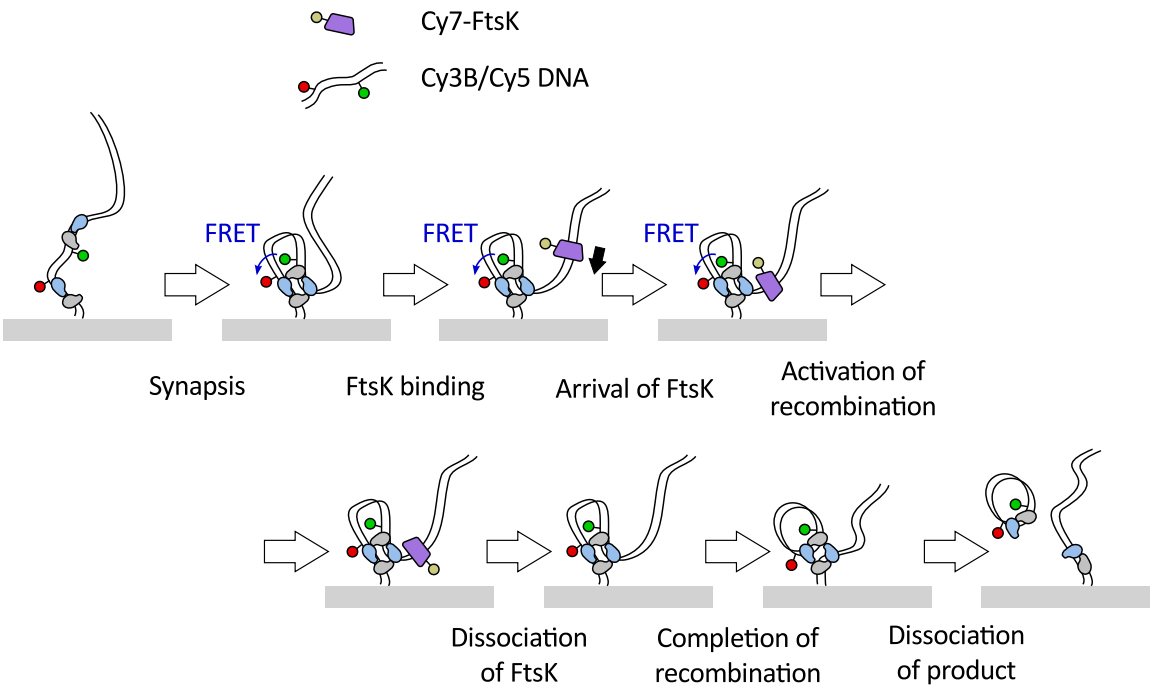

Figure 5.15: Three-colour XerCD-*dif*-FtsK experiments. Two labels on the DNA follow synaptic complex assembly and the formation, isomerisation and resolution of the Holliday junction intermediate. A third, spectrally distinct label that can be directly excited in a three-colour ALEX scheme, can follow the assembly, translocation, arrival and dissociation of FtsK.

5.5.1 XerCD-*dif*-FtsK experimental design

The microscope will be useful for the continued study of XerCD-*dif* recombination. Two labels on DNA can monitor the progress of recombination and the dissociation of product (Figure 5.15), and a spectrally distinct label on FtsK can monitor its behaviour. The precise timing of FtsK dissociation can be used to infer if there is mechanistic feedback between the recombining XerCD-*dif* complex and the activating FtsK. The conclusion in Chapter 4, that FtsK does not remain at XerCD-*dif* long enough to activate subsequent rounds of recombination, can be directly tested.

5.5.2 Future improvements

Although Cy7 offers good photostability and photobleaching lifetime, newer, modified Cy7-derivatives with improved properties can be used in its place. Sulfo-Cy7 is an improved

analog of Cy7, with a 20 % increase in quantum yield, and better photostability (Lumiprobe, Hannover, Germany). Other stabilised Cy7 derivatives have been shown to offer imaging for ∼180 s at SNR > 7 [134].

One possible expansion to the microscope configuration would be the addition of a fourth, blue, spectral channel, if reduced physical dye size, or further increased observation bandwidth was desirable [26]. This would be a relatively simple addition, since the dichroic mirror, which separates incident lasers and emitted fluorescence (Figure 5.2, DM1), and the laser filter (Figure 5.2, Filter) are already suitable for 488 nm illumination. A second expansion, facilitated by the use of objective TIRF, would be the addition of a white-light illumination source above the objective, which would allow the use of the NIR spectral channel for live-cell imaging. This would take advantage of the low cellular autofluorescence in the NIR channel. Since the R_0 for Cy3-Cy7 is 3.8 nm [133], this dye pair could offer the chance to monitor shorter distance changes than the commonly used green-red dye pairs ($R_0 \sim 6.5$ nm), extending the range of applicability of FRET.

Finally, since the analysis software already utilises parallel execution on a local cluster (consisting of the multiple cores of a single local processor), the expansion to analysis on a large centrally managed cluster consisting of many (> 4) independent workers should be simple. Once such a cluster was set up, the cluster profile would have to be set in MATLAB, and some care taken to make sure that the files to analyse were accessible and properly addressed from the point of view of each worker. This would allow remote analysis of collected data and would continue to realise the potential of parallel computing for increasing the speed of data analysis.

5.6 Contributions

With the exception of the laser mounts and the image splitting dichroic mount (Figures B.8, B.9 and B.11), which I machined, custom components were machined by the Clarendon

Workshop (Oxford, UK). Proteins used for single-molecule experiments were provided by Pawel Zawadzki. All other work is my own.

Chapter 6

Conclusion

Chapter 1 has briefly introduced single-molecule techniques and described the basic principles of single-molecule total internal reflection fluorescence (TIRF) microscopy. Central to the studies in this work has been the technique of tethered fluorophore motion (TFM). It uses a single fluorophore, in place of the bead in tethered particle motion (TPM), to monitor an effective length along a DNA molecule. TFM can be combined with existing fluorescence techniques, and hence allows additional information to be gathered simultaneously about a biological complex under investigation.

Chapter 2 has given a firm theoretical grounding to TFM and has placed its resolution and experimental considerations in context, with reference to the established technique of TPM. TFM has been used to extract a polymerisation rate for the Klenow fragment of DNA polymerase I and to extract a synaptic lifetime for a mutant of the well-studied Cre recombinase, combining TFM and FRET, using only a single excitation laser. This latter case demonstrates the simplicity of TFM-FRET for use as a combination single-molecule technique.

Chapters 3 and 4 have applied and extended TFM to study the XerCD-*dif*-FtsK chromosome maintenance and segregation machinery. The work in Chapter 3 has shown that XerCD-*dif* recombination has five distinct structural states, the first of which does not

resemble any predicted using the mechanism of Cre-*loxP*, a related tyrosine recombinase, and has determined the kinetics of each of the distinct steps in the pathway.

Previous studies have shown that FtsK may form either a single or double hexamer on DNA [56], that it might be capable of reversals in translocation direction, both spontaneously and in response to a collision with XerD [58, 101, 111, 112, 113], and that it possibly extrudes loops of DNA as it translocates [58, 101, 110, 111]. However, these previous studies may have been confounded by protein aggregation, or multiple attachment of proteins to fluorescent labels. In the work in Chapter 4 we set out to measure the stoichiometry of the translocating unit of FtsK and determine its behaviour when it interacts with XerCD-*dif*, in both synapsed and un-synapsed conformations. We produced a labelled trimer of the motor domain of FtsK, which carried no more than one protein per fluorescent label, and used its intensity to measure stoichiometry, ensuring aggregates were not present. Our results demonstrate that FtsK assembles into a single hexamer on DNA and begins translocating shortly after it completes assembly. When it arrives at XerCD-*dif*, it dissociates rapidly (0.5 s to 1 s), rather than reversing, irrespective of whether XerCD-*dif* is synapsed or un-synapsed. When activating recombination, FtsK seems to dissociate faster than the completion of recombination, as judged by PIFE (Chapter 3) and by direct monitoring of fluorescent FtsK (Chapter 4).

These observations suggest a simple model of XerCD-*dif*-FtsK function. Once replication is almost complete, and two *dif*-sites are present within the cell, initial synapses form and dissociate without any strand exchange occuring. Since the transient association time of FtsK is short (< 1 s, Chapter 3), FtsK will only form the translocating unit on DNA when its local concentration is high, at an invaginating septum [114, 124]. When hexamers do form on DNA, they will translocate preferentially towards *dif*, due to the presence of KOPS sequences [58, 59, 60], without extruding loops of DNA. These translocating hexamers will arrive at XerCD-*dif* and dissociate. If XerCD-*dif* synapses form regularly, because the chromosome topology is preventing segregation (due to dimerisation

or catenation), then eventually one of the colliding FtsK hexamers will activate site-specific recombination, leading to the resolution of the sister chromosomes, one recombination at a time. This simple mechanism, with FtsK tending to break apart, or not activate, synapses in the majority of collisions, may be sufficient to bias the mechanism towards simplifying chromosome topology.

Future work, utilising the three-colour TIRF microscope, whose design and construction were detailed in Chapter 5, could combine the approaches of Chapters 3 and 4, with two fluorophores used to follow each step in the XerCD-*dif* recombination pathway and a third label on FtsK, to relate this to the precise timing of FtsK arrival and departure. Three-colour work could also be used to elucidate the role of the γ subdomain of FtsK in the recognition of KOPS and the activation of recombination, using one label to track DNA conformation, and two to measure a precise distance within a functioning FtsK hexamer.

Appendix A

Publications and presentations arising from this thesis

A.1 Publications

1. P. F. May, J. N. Pinkney, P. Zawadzki, G. W. Evans, D. J. Sherratt, and A. N. Kapanidis, "Tethered Fluorophore Motion: Studying Large DNA Conformational Changes by Single-fluorophore Imaging," *Biophysical Journal*, vol. 107, pp. 1205–16, Sept. 2014

2. P. Zawadzki, P. F. J. May, R. A. Baker, J. N. M. Pinkney, A. N. Kapanidis, D. J. Sherratt, and L. K. Arciszewska, "Conformational transitions during FtsK translocase activation of individual XerCD-dif recombination complexes.," *Proceeding of the National Academy of Sciences of the United States of America*, vol. 110, pp. 17302–7, Oct. 2013

3. P. F. J. May, P. Zawadzki, D. J. Sherratt, A. N. Kapanidis, and L. K. Arciszewska, "Assembly, translocation, and activation of XerCD-dif recombination by FtsK translocase analyzed in real-time by FRET and two-color tethered fluorophore motion.," *Proceeding of the National Academy of Sciences of the United States of America*, vol. 112, pp. E5133–5141, Aug. 2015

A.2 Presentations

1. "Combining TPM and single-molecule fluorescence to study the FtsK translocase" (poster and talk) at the *7th Oxford Photonics Day* held on the 20$^{\text{th}}$ April 2015 in Oxford

2. "Combining TPM and single-molecule fluorescence to study the FtsK translocase" (poster) at the *58th Annual Meeting of the Biophysical Society* held on the 7$^{\text{th}}$ - 11$^{\text{th}}$ February 2015 in Baltimore, Maryland

3. "Tethered fluorophore motion studies of bacterial site-specific recombination" (poster) at the *2nd International workshop on The Biology and Physics of Bacterial Genomes*, held on the 16th - 18th June 2014 in Birmingham

4. "Mulitparameter tethered fluorophore motion studies of DNA recombination" (talk) at the *19th International Workshop on Single Molecule Spectroscopy and Ultrasensitive Analysis in the Life Sciences* held on the 3rd - 6th October 2013 in Berlin

Appendix B

Technical drawings

In Chapter 5, I described the design, construction and evaluation of a new three-colour fluorescence microscope. As part of that work, I designed several components used to precisely align and anchor components and the designs of these components are presented here. All scales are in mm.

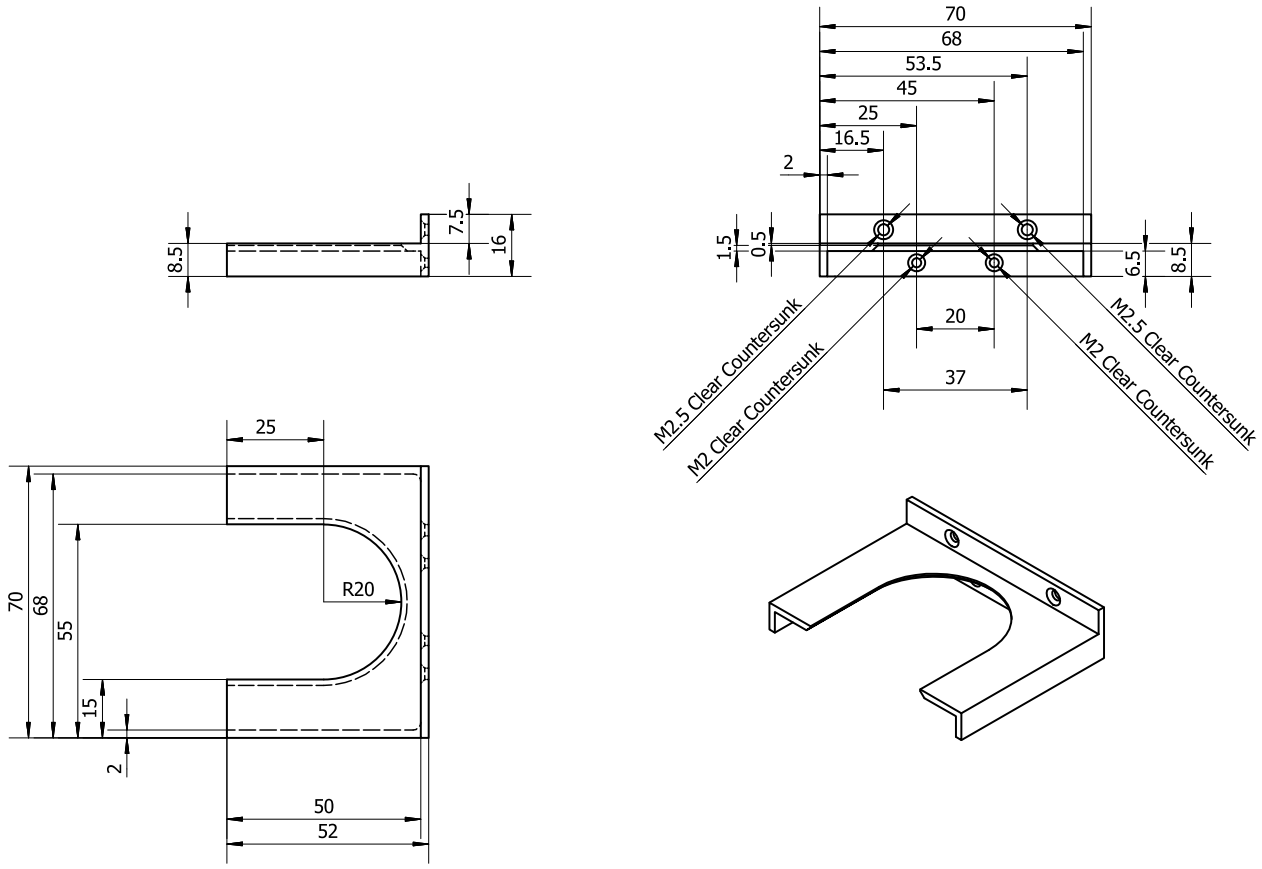

Figure B.1: Technical drawing of the slide holder mount for the piezoelectric stepper stage. Material: invar.

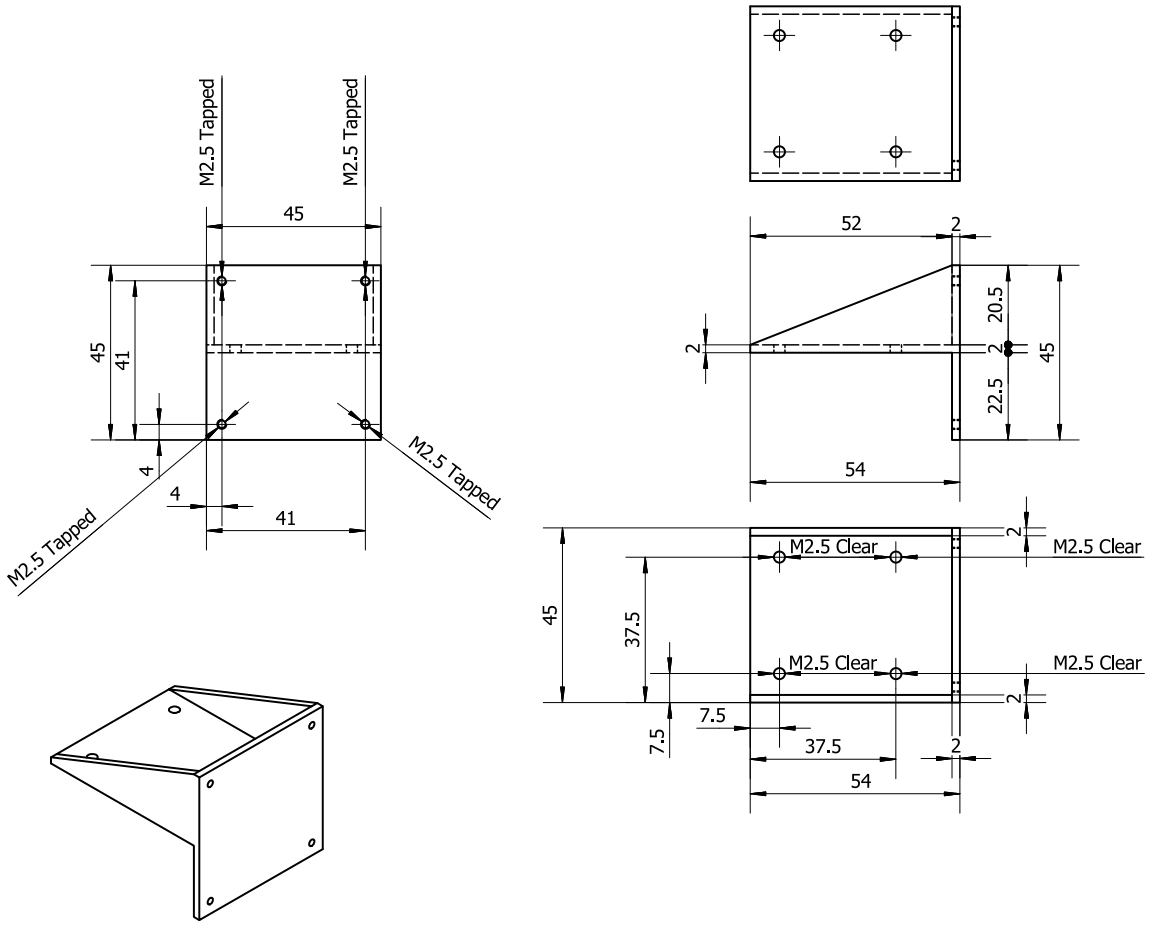

Figure B.2: Technical drawing of the corner mount for the piezoelectric stepper stages. Material: invar.

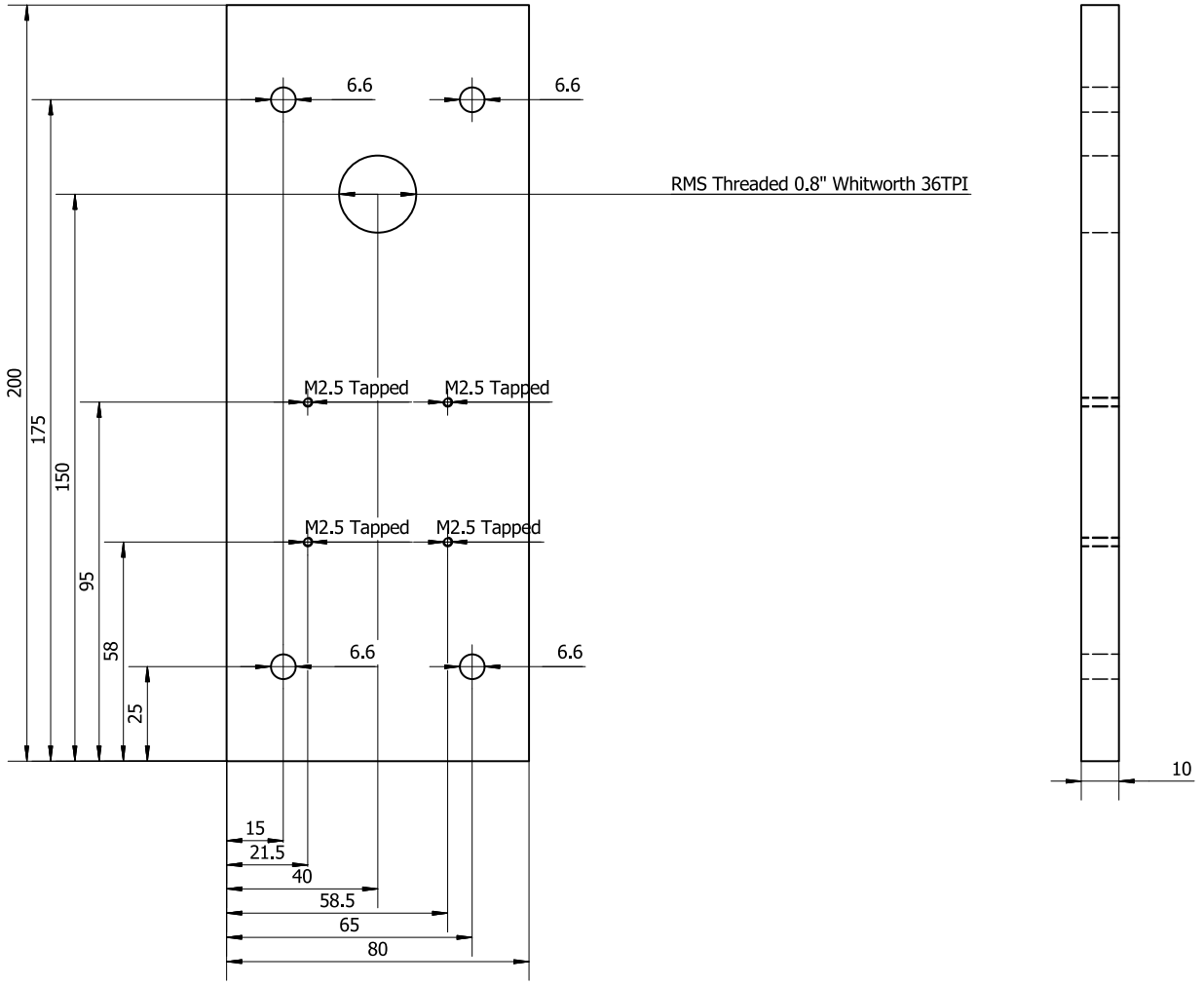

Figure B.3: Technical drawing of the baseplate for the piezoelectric stepper stages. Material: invar.

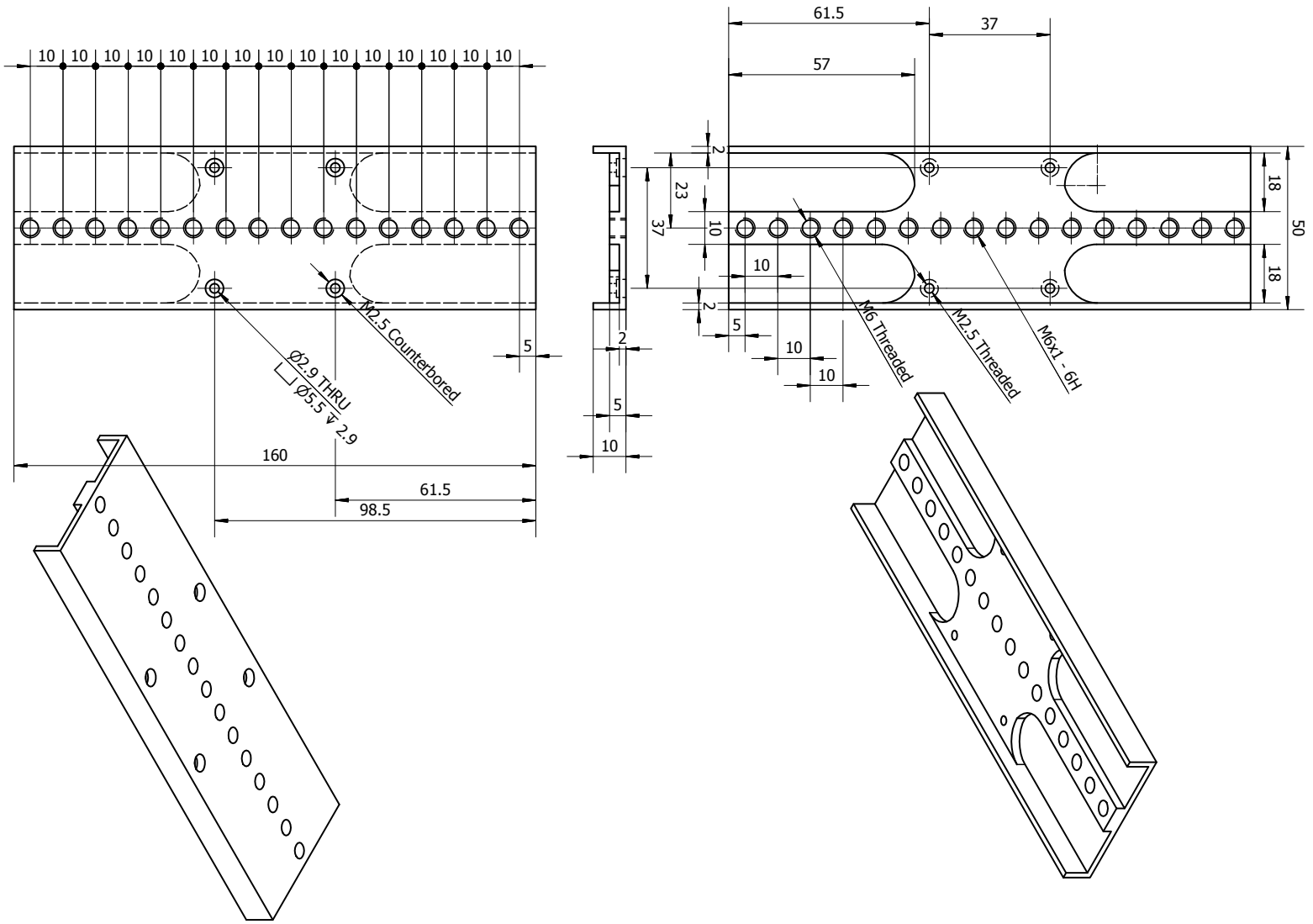

Figure B.4: Technical drawing of the mount for the TIRF lens. Material: aluminium.

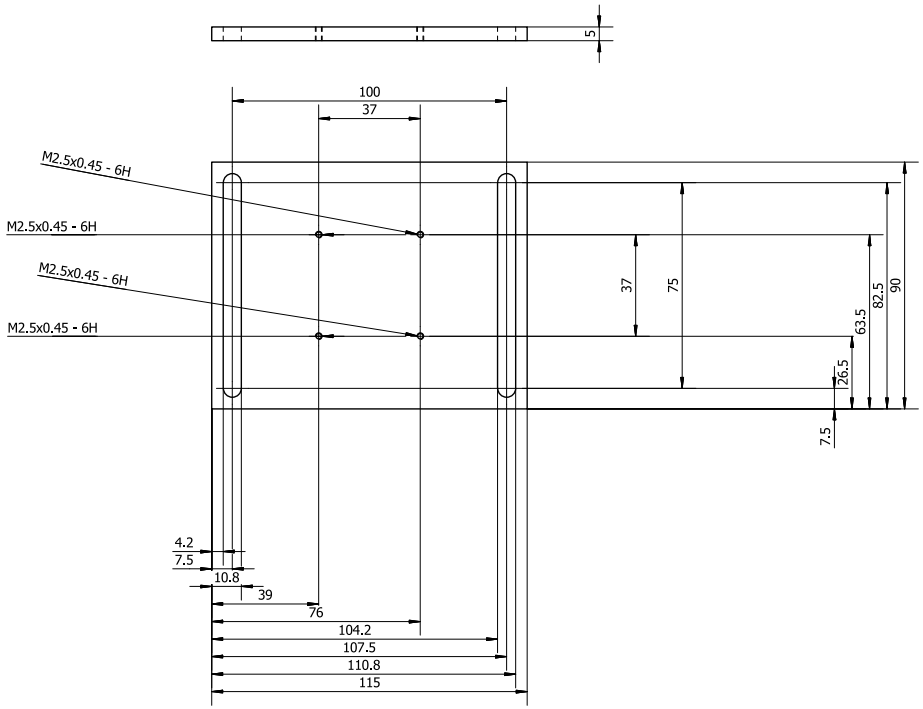

Figure B.5: Technical drawing of the baseplate for TIRF lens stage. Material: aluminium.

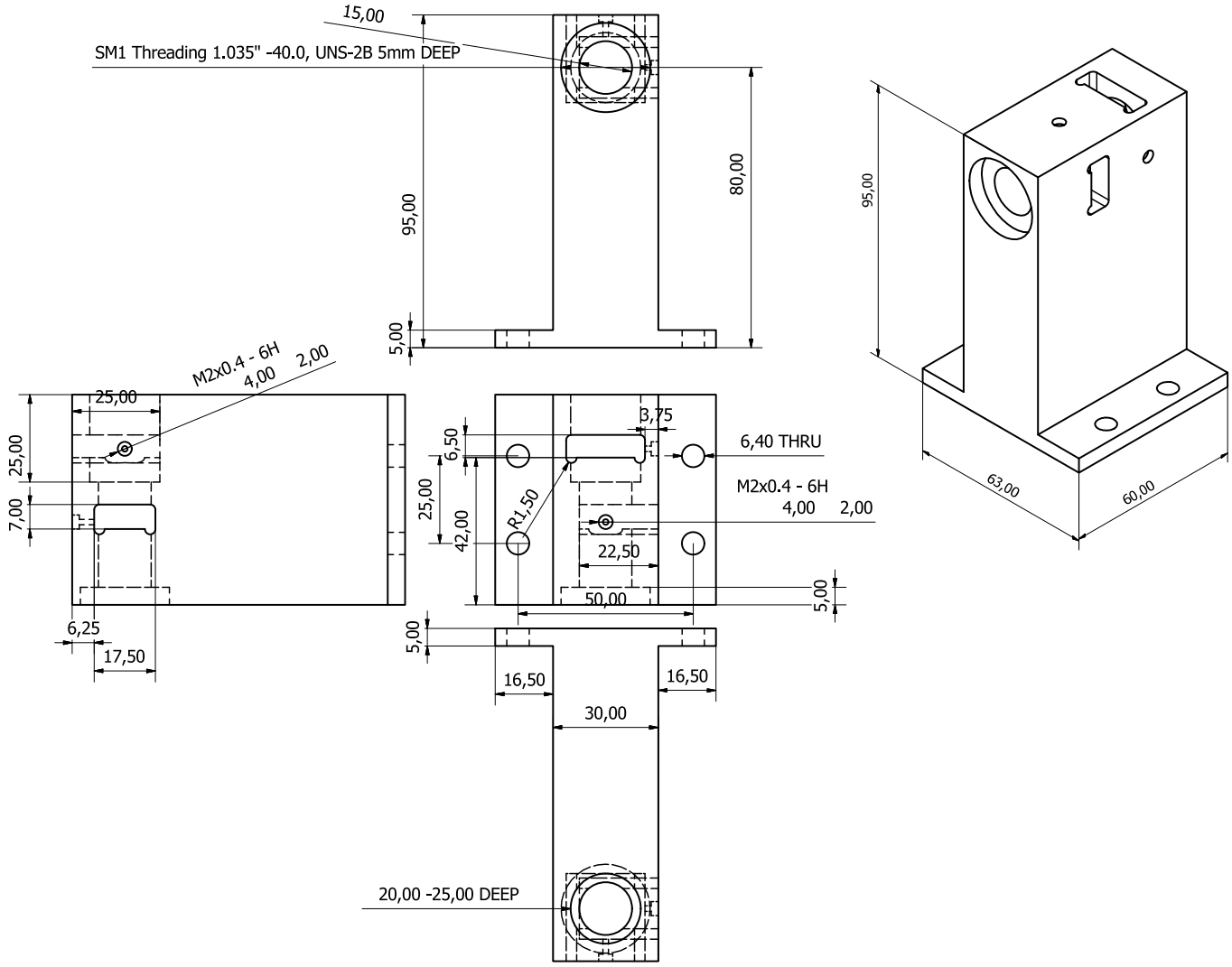

Figure B.6: Technical drawing of the fibre launch, which holds two cylindrical lenses. Material: aluminium.

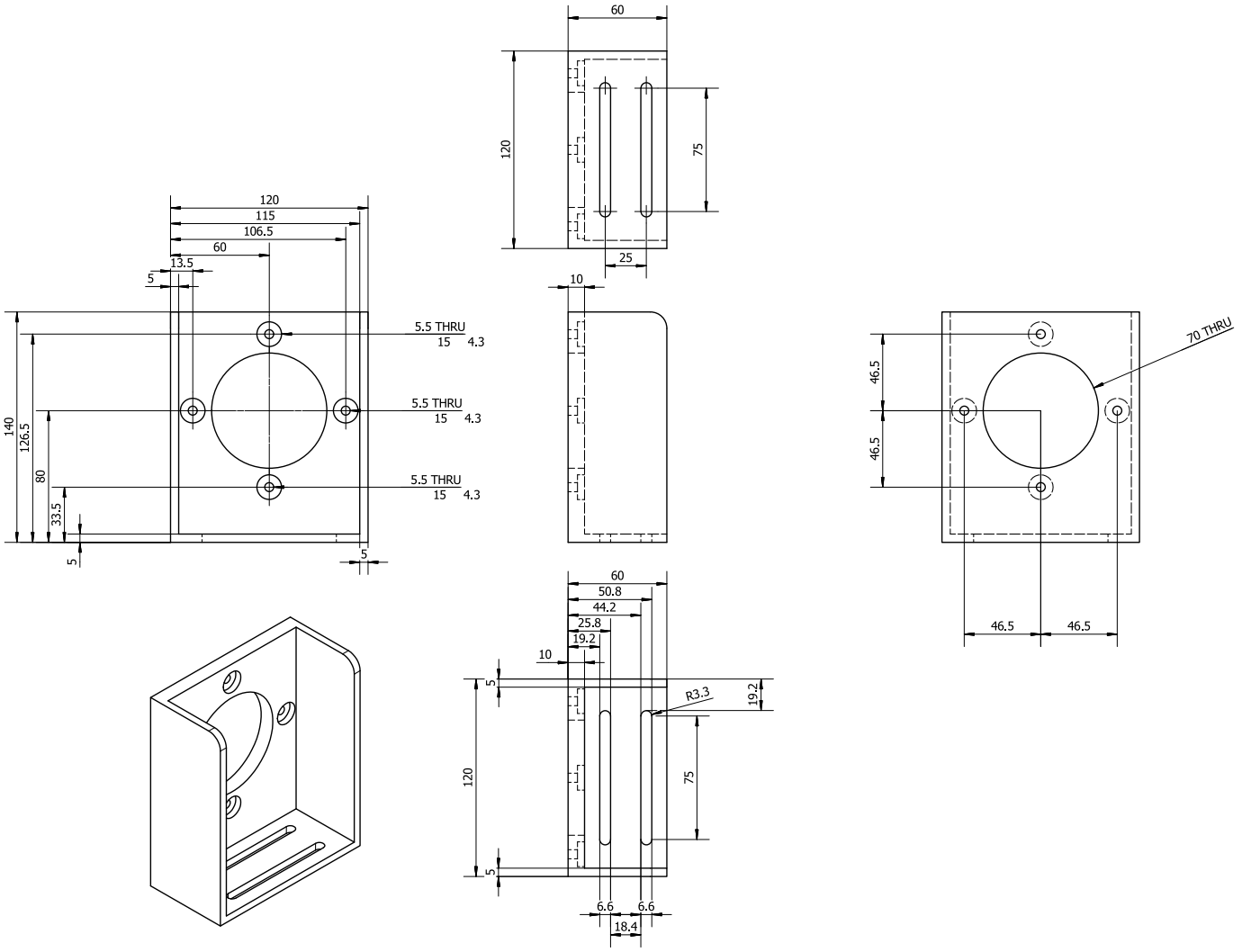

Figure B.7: Technical drawing of the camera mount. Material: aluminium.

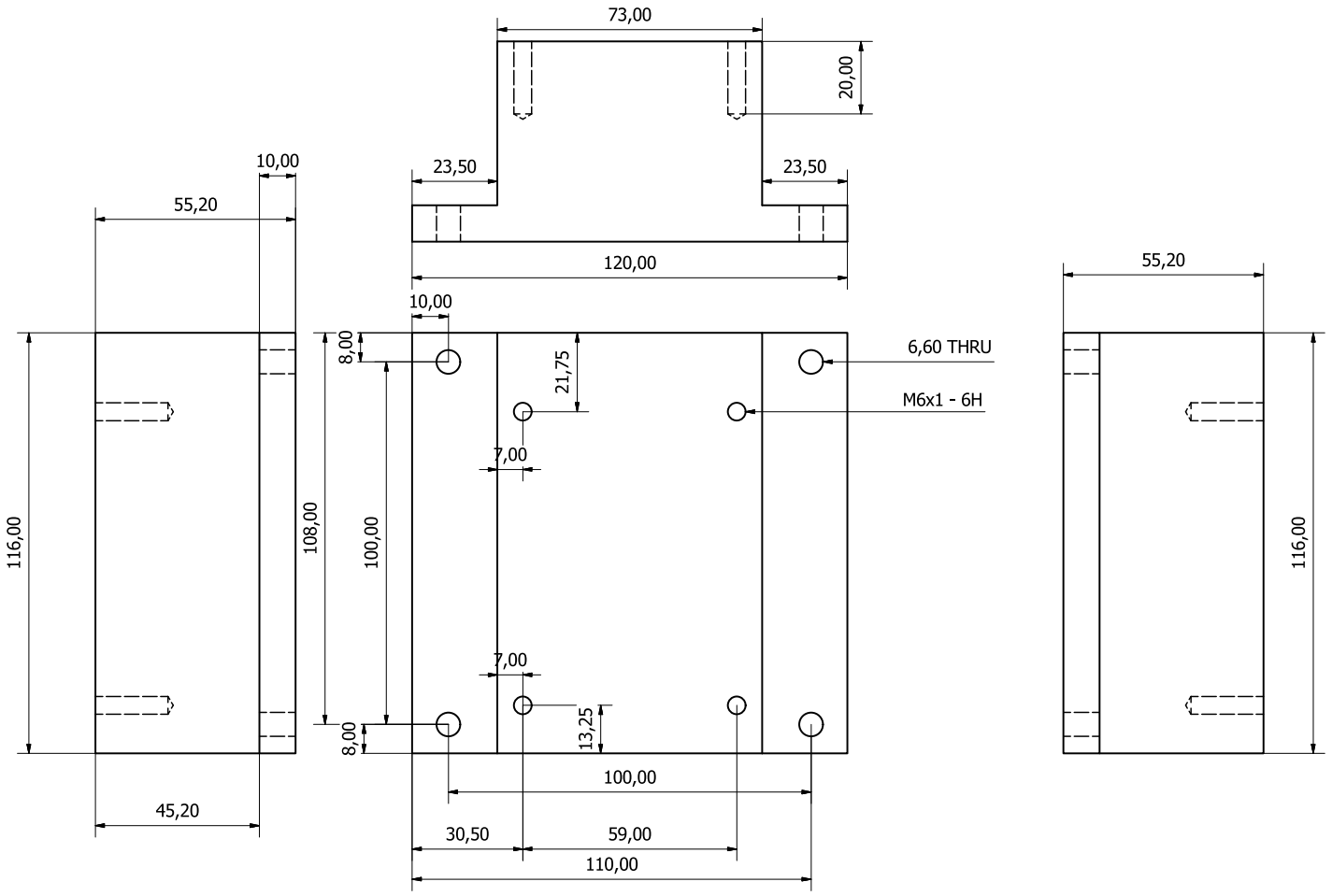

Figure B.8: Technical drawing of the green laser mount. Material: aluminium.

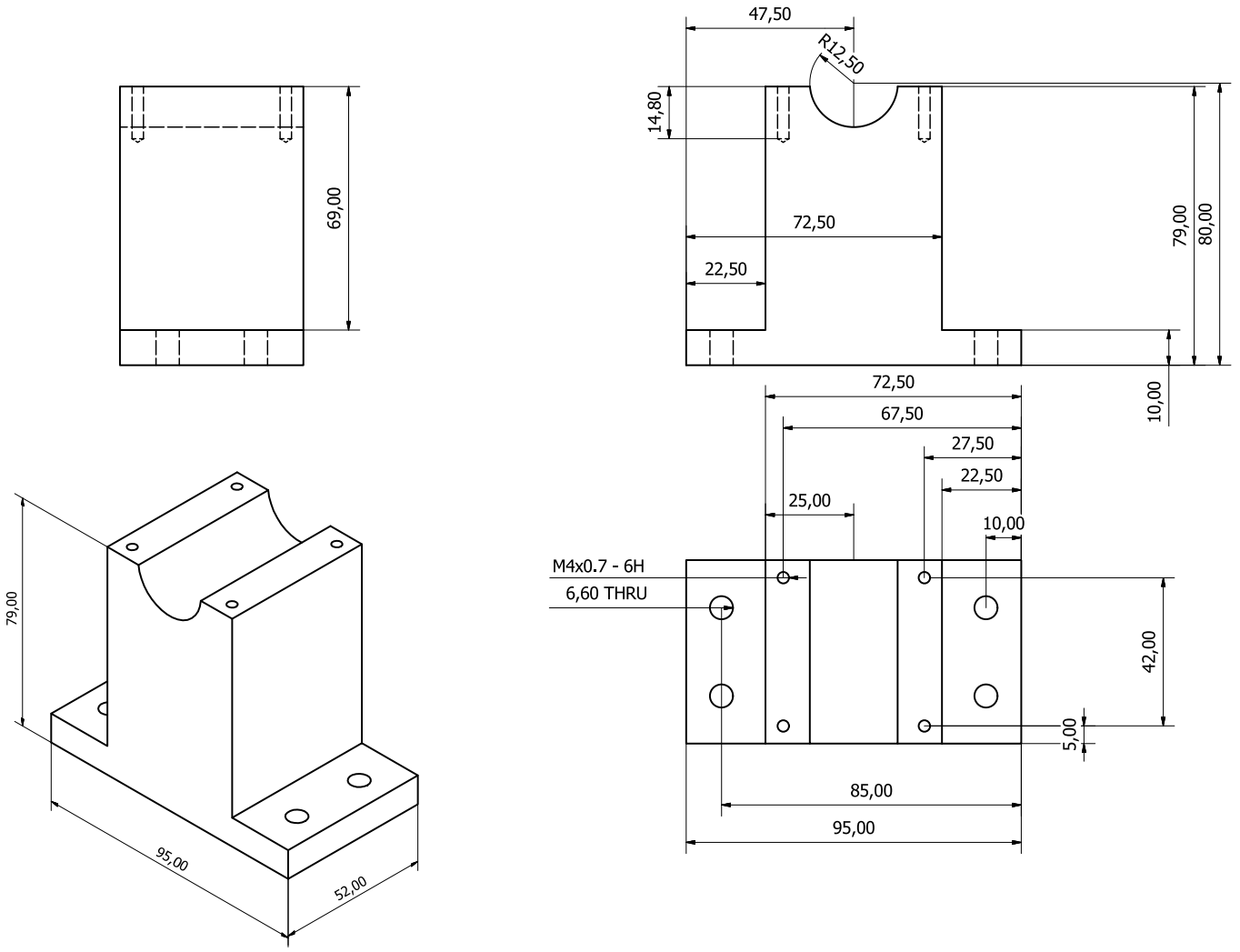

Figure B.9: Technical drawing of the NIR laser mount. Material: aluminium.

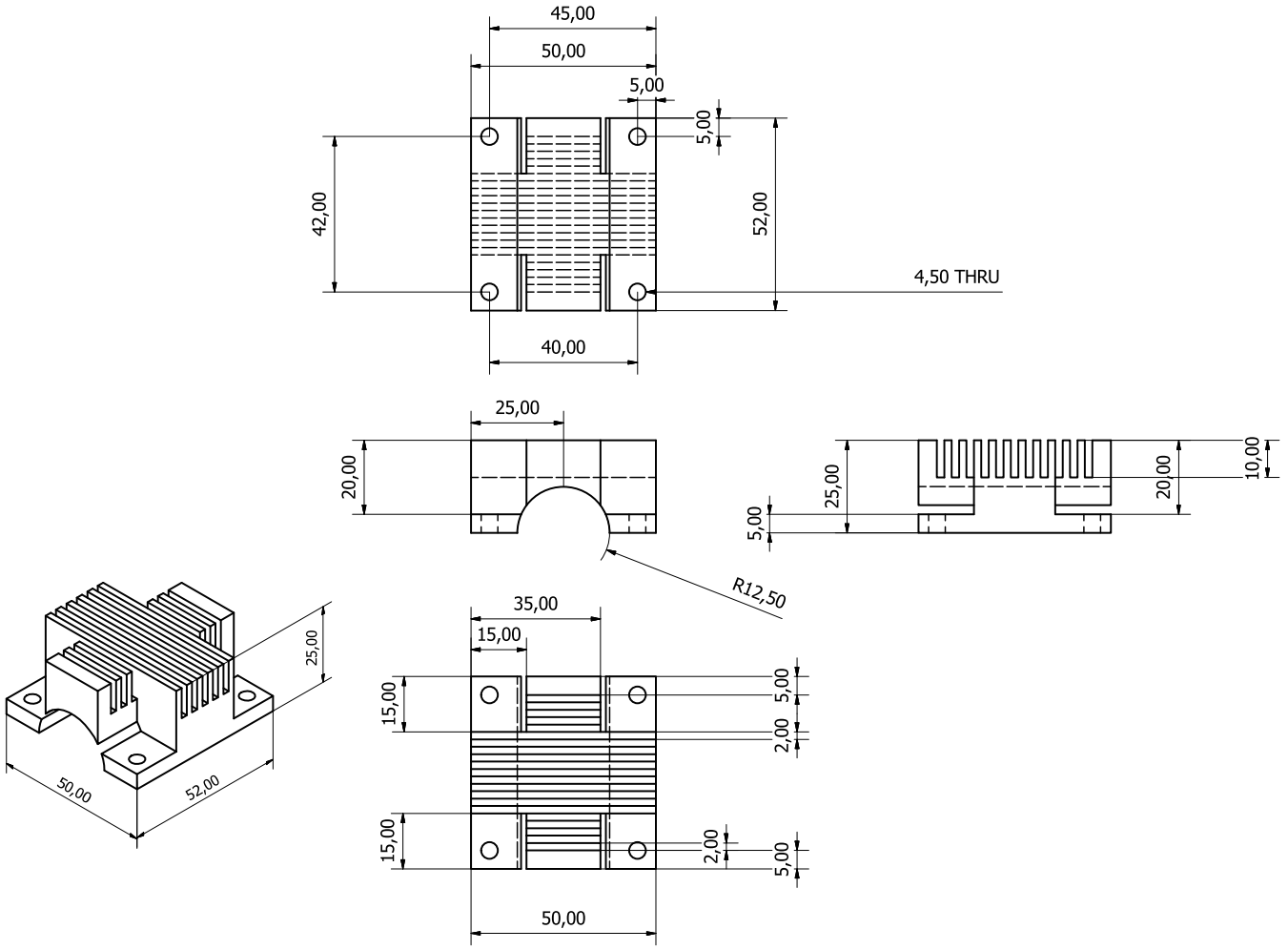

Figure B.10: Technical drawing of the NIR laser heatsink. Material: aluminium.

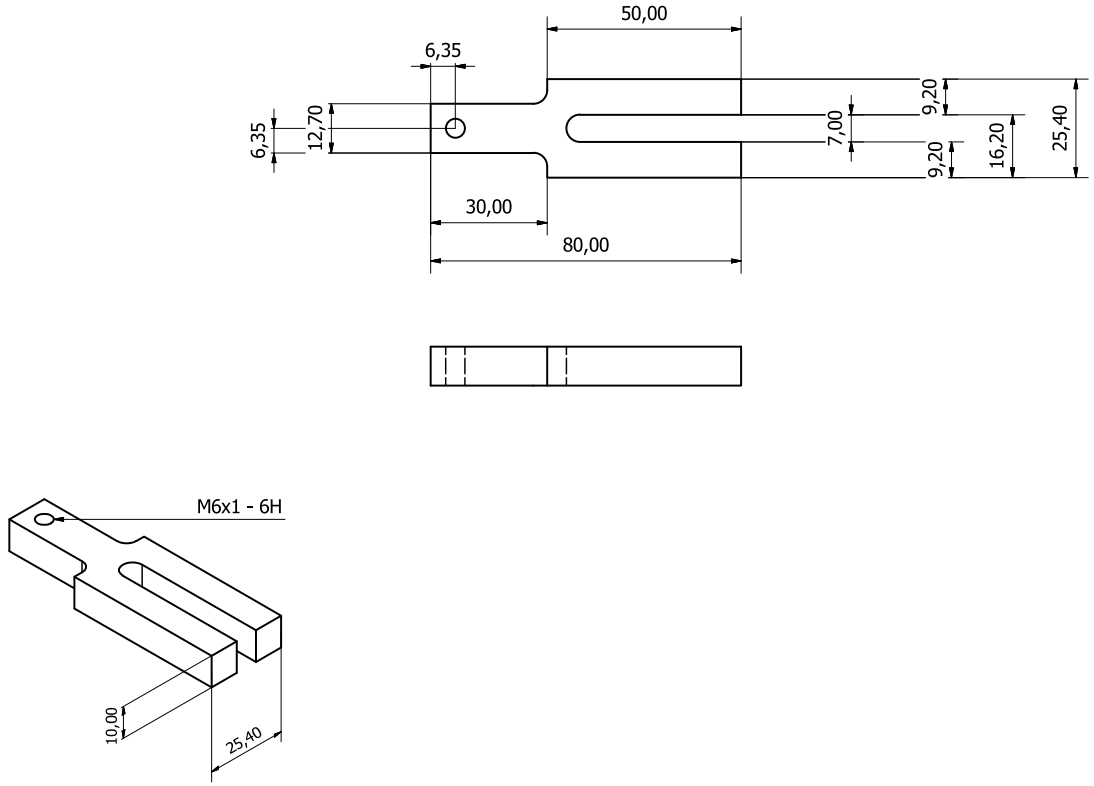

Figure B.11: Technical drawing of the image splitting dichroic mount. Material: aluminium.

References

[1] P. F. May, J. N. Pinkney, P. Zawadzki, G. W. Evans, D. J. Sherratt, and A. N. Kapanidis, "Tethered Fluorophore Motion: Studying Large DNA Conformational Changes by Single-fluorophore Imaging," *Biophysical Journal*, vol. 107, pp. 1205–16, Sept. 2014.

[2] P. Zawadzki, P. F. J. May, R. A. Baker, J. N. M. Pinkney, A. N. Kapanidis, D. J. Sherratt, and L. K. Arciszewska, "Conformational transitions during FtsK translocase activation of individual XerCD-dif recombination complexes.," *Proceeding of the National Academy of Sciences of the United States of America*, vol. 110, pp. 17302–7, Oct. 2013.

[3] P. F. J. May, P. Zawadzki, D. J. Sherratt, A. N. Kapanidis, and L. K. Arciszewska, "Assembly, translocation, and activation of XerCD-dif recombination by FtsK translocase analyzed in real-time by FRET and two-color tethered fluorophore motion.," *Proceeding of the National Academy of Sciences of the United States of America*, vol. 112, pp. E5133–5141, Aug. 2015.

[4] A. N. Kapanidis and T. Strick, "Biology, one molecule at a time.," *Trends in biochemical Sciences*, vol. 34, pp. 234–243, 2009.

[5] S. Weiss, "Fluorescence Spectroscopy of Single Biomolecules," *Science*, vol. 283, pp. 1676–1683, Mar. 1999.

[6] A. Ashkin, J. M. Dziedzic, J. E. Bjorkholm, and S. Chu, "Observation of a single-beam gradient force optical trap for dielectric particles.," *Optics Letters*, vol. 11, no. 5, p. 288, 1986.

[7] T. R. Strick, J. F. Allemand, D. Bensimon, A. Bensimon, and V. Croquette, "The elasticity of a single supercoiled DNA molecule.," *Science*, vol. 271, no. 5257, pp. 1835–1837, 1996.

[8] D. A. Schafer, J. Gelles, M. P. Sheetz, and R. Landick, "Transcription of single molecules of RNA polymerase observed by light microscopy," *Nature*, vol. 352, pp. 444–448, 1991.

[9] G. Binnig, C. Quate, and C. Gerber, "Atomic force microscope," *Physical review letters*, vol. 56, no. 9, pp. 930–933, 1986.

[10] T. Ha, T. Enderle, D. F. Ogletree, D. S. Chemla, P. R. Selvin, and S. Weiss, "Probing the interaction between two single molecules: fluorescence resonance energy transfer between a single donor and a single acceptor.," *Proceeding of the National Academy of Sciences of the United States of America*, vol. 93, no. 13, pp. 6264–6268, 1996.

[11] U. Kubitscheck, O. Kückmann, T. Kues, and R. Peters, "Imaging and tracking of single GFP molecules in solution.," *Biophysical Journal*, vol. 78, no. 4, pp. 2170–2179, 2000.

[12] M. Goulian and S. M. Simon, "Tracking single proteins within cells.," *Biophysical Journal*, vol. 79, no. 4, pp. 2188–2198, 2000.

[13] P. R. Selvin and T. Ha, "The New Era of Biology In Singulo," in *Single-Molecule Techniques: A Laboratory Manual* (P. R. Selvin and T. Ha, eds.), ch. 1, pp. 1–2, New York: Cold Spring Harbor Laboratory Press, 2008.

[14] P. C. Nelson, C. Zurla, D. Brogioli, J. F. Beausang, L. Finzi, and D. Dunlap, "Tethered particle motion as a diagnostic of DNA tether length.," *The journal of physical chemistry. B*, vol. 110, pp. 17260–7, Aug. 2006.

[15] L. Finzi and J. Gelles, "Measurement of lactose repressor-mediated loop formation and breakdown in single DNA molecules.," *Science*, vol. 267, pp. 378–380, 1995.

[16] D. a. Rusling, N. Laurens, C. Pernstich, G. J. L. Wuite, and S. E. Halford, "DNA looping by FokI: the impact of synapse geometry on loop topology at varied site orientations.," *Nucleic Acids Research*, vol. 40, pp. 4977–87, 2012.

[17] S. Brinkers, H. R. C. Dietrich, F. H. de Groote, I. T. Young, and B. Rieger, "The persistence length of double stranded DNA determined using dark field tethered particle motion.," *The Journal of chemical physics*, vol. 130, p. 215105, June 2009.

[18] M. Manghi, C. Tardin, J. Baglio, P. Rousseau, L. Salomé, and N. Destainville, "Probing DNA conformational changes with high temporal resolution by tethered particle motion.," *Physical biology*, vol. 7, p. 046003, 2010.

[19] H. Yin, R. Landick, and J. Gelles, "Tethered particle motion method for studying transcript elongation by a single RNA polymerase molecule.," *Biophysical Journal*, vol. 67, pp. 2468–78, Dec. 1994.

[20] D. E. Segall, P. C. Nelson, and R. Phillips, "Volume-exclusion effects in tethered-particle experiments: bead size matters.," *Physical review letters*, vol. 96, p. 088306, 2006.

[21] H. Bai, J. E. Kath, F. M. Zörgiebel, M. Sun, P. Ghosh, G. F. Hatfull, N. D. F. Grindley, and J. F. Marko, "Remote control of DNA-acting enzymes by varying the Brownian dynamics of a distant DNA end," *Proceedings of the National Academy of Sciences*, vol. 109, pp. 16546–16551, 2012.

[22] S. Hohng, R. Zhou, M. K. Nahas, J. Yu, K. Schulten, D. M. J. Lilley, and T. Ha, "Fluorescence-force spectroscopy maps two-dimensional reaction landscape of the holliday junction.," *Science*, vol. 318, pp. 279–83, Oct. 2007.

[23] G. S. Harms, G. Orr, M. Montal, B. D. Thrall, S. D. Colson, and H. P. Lu, "Probing conformational changes of gramicidin ion channels by single-molecule patch-clamp fluorescence microscopy.," *Biophysical Journal*, vol. 85, pp. 1826–38, Sept. 2003.

[24] S. Hohng, C. Joo, and T. Ha, "Single-molecule three-color FRET.," *Biophysical Journal*, vol. 87, pp. 1328–37, Aug. 2004.

[25] N. K. Lee, A. N. Kapanidis, H. R. Koh, Y. Korlann, S. O. Ho, Y. Kim, N. Gassman, S. K. Kim, and S. Weiss, "Three-color alternating-laser excitation of single molecules: monitoring multiple interactions and distances.," *Biophysical Journal*, vol. 92, pp. 303–12, Jan. 2007.

[26] J. Lee, S. Lee, K. Ragunathan, C. Joo, T. Ha, and S. Hohng, "Single-molecule four-color FRET.," *Angewandte Chemie (International ed. in English)*, vol. 49, pp. 9922–5, Dec. 2010.

[27] L. A. Ernst, R. K. Gupta, R. B. Mujumdar, and A. S. Waggoner, "Cyanine dye labeling reagents for sulfhydryl groups.," *Cytometry*, vol. 10, pp. 3–10, Jan. 1989.

[28] M. Cooper, A. Ebner, M. Briggs, M. Burrows, N. Gardner, R. Richardson, and R. West, "Cy3B (TM): Improving the performance of cyanine dyes," *Journal of Fluorescence*, vol. 14, no. 2, pp. 145–150, 2004.

[29] I. Rasnik, S. A. McKinney, and T. Ha, "Nonblinking and long-lasting single-molecule fluorescence imaging.," *Nature methods*, vol. 3, no. 11, pp. 891–893, 2006.

[30] R. E. Benesch and R. Benesch, "Enzymatic Removal of Oxygen for Polarography and Related Methods," *Science*, vol. 118, pp. 447–448, Oct. 1953.

[31] T. Cordes, J. Vogelsang, and P. Tinnefeld, "On the mechanism of trolox as antiblinking and antibleaching reagent," *Journal of the American Chemical Society*, vol. 131, no. 14, pp. 5018–5019, 2009.

[32] J. N. M. Pinkney, P. Zawadzki, J. Mazuryk, L. K. Arciszewska, D. J. Sherratt, and A. N. Kapanidis, "Capturing reaction paths and intermediates in Cre-loxP recombination using single-molecule fluorescence.," *Proceeding of the National Academy of Sciences of the United States of America*, vol. 109, pp. 1–6, Nov. 2012.

[33] S. Sindbert, S. Kalinin, H. Nguyen, A. Kienzler, L. Clima, W. Bannwarth, B. Appel, S. Müller, and C. a. M. Seidel, "Accurate distance determination of nucleic acids via Förster resonance energy transfer: implications of dye linker length and rigidity.," *Journal of the American Chemical Society*, vol. 133, pp. 2463–80, Mar. 2011.

[34] T. Förster, "Delocalized Excitation and Excitation Transfer," in *Modern Quantum Chemistry. Istanbul Lectures. Part III: Action of Light and Organic Crystals* (O. Sinanoglu, ed.), pp. 93–137, New York and London: Academic Press, 1965.

[35] J. R. Lakowicz, *Principles of Fluorescence Spectroscopy*. London: Kluwer Academic/Plenum, 2nd ed., 1999.

[36] N. K. Lee, A. N. Kapanidis, Y. Wang, X. Michalet, J. Mukhopadhyay, R. H. Ebright, and S. Weiss, "Accurate FRET measurements within single diffusing biomolecules using alternating-laser excitation.," *Biophysical Journal*, vol. 88, pp. 2939–53, Apr. 2005.

[37] A. N. Kapanidis, N. K. Lee, T. a. Laurence, S. Doose, E. Margeat, and S. Weiss, "Fluorescence-aided molecule sorting: analysis of structure and interactions by alternating-laser excitation of single molecules.," *Proceeding of the National Academy of Sciences of the United States of America*, vol. 101, pp. 8936–41, June 2004.

[38] D. Axelrod, "Cell-substrate contacts illuminated by total internal reflection fluorescence," *Journal of Cell Biology*, vol. 89, no. 1, pp. 141–145, 1981.

[39] R. E. Thompson, D. R. Larson, and W. W. Webb, "Precise nanometer localization analysis for individual fluorescent probes.," *Biophysical Journal*, vol. 82, pp. 2775–83, May 2002.

[40] K. I. Mortensen, L. S. Churchman, J. a. Spudich, and H. Flyvbjerg, "Optimized localization analysis for single-molecule tracking and super- resolution microscopy," *Nature methods*, vol. 7, pp. 377–81, May 2010.

[41] A. V. Abraham, S. Ram, J. Chao, E. S. Ward, and R. J. Ober, "Quantitative study of single molecule location estimation techniques.," *Optics express*, vol. 17, pp. 23352–73, Dec. 2009.

[42] S. J. Holden, S. Uphoff, J. Hohlbein, D. Yadin, L. L. Reste, O. J. Britton, A. N. Kapanidis, M. Carlo, and L. Le Reste, "Defining the limits of single-molecule FRET resolution in TIRF microscopy.," *Biophysical Journal*, vol. 99, pp. 3102–11, Nov. 2010.

[43] M. C. DeSantis, S. H. DeCenzo, J.-L. Li, and Y. M. Wang, "Precision analysis for standard deviation measurements of immobile single fluorescent molecule images.," *Optics express*, vol. 18, pp. 6563–76, Mar. 2010.

[44] N. D. F. Grindley, K. L. Whiteson, and P. A. Rice, "Mechanisms of site-specific recombination.," *Annual review of biochemistry*, vol. 75, pp. 567–605, Jan. 2006.

[45] B. Hallet, "Transposition and site-specific recombination: adapting DNA cut-and-paste mechanisms to a variety of genetic rearrangements," *FEMS Microbiology Reviews*, vol. 21, pp. 157–178, Sept. 1997.

[46] A. Akopian and W. Marshall Stark, "Site-specific DNA recombinases as instruments for genomic surgery.," *Advances in genetics*, vol. 55, pp. 1–23, Jan. 2005.

[47] D. J. Sherratt, L. K. Arciszewska, G. Blakely, S. Colloms, K. Grant, N. Leslie, and R. McCulloch, "Site-specific recombination and circular chromosome segregation.," *Philosophical transactions of the Royal Society of London. Series B, Biological Sciences*, vol. 347, pp. 37–42, Jan. 1995.

[48] I. Grainge, M. Bregu, M. Vazquez, V. Sivanathan, S. C. Y. Ip, and D. J. Sherratt, "Unlinking chromosome catenanes in vivo by site-specific recombination.," *The EMBO journal*, vol. 26, pp. 4228–38, Oct. 2007.

[49] S. Bigot, V. Sivanathan, C. Possoz, F.-X. Barre, and F. Cornet, "FtsK, a literate chromosome segregation machine.," *Molecular microbiology*, vol. 64, pp. 1434–41, June 2007.

[50] D. J. Sherratt, L. K. Arciszewska, E. Crozat, J. E. Graham, and I. Grainge, "The Escherichia coli DNA translocase FtsK.," *Biochemical Society transactions*, vol. 38, no. 2, pp. 395–398, 2010.

[51] G. D. Van Duyne, "A structural view of cre-loxp site-specific recombination.," *Annual review of biophysics and biomolecular structure*, vol. 30, pp. 87–104, 2001.

[52] F.-X. Barre, "FtsK and SpoIIIE: the tale of the conserved tails.," *Molecular microbiology*, vol. 66, pp. 1051–5, Dec. 2007.

[53] C. Kaimer and P. L. Graumann, "Players between the worlds: multifunctional DNA translocases.," *Current opinion in microbiology*, vol. 14, pp. 719–25, Dec. 2011.

[54] G. C. Draper, N. McLennan, K. Begg, M. Masters, and W. D. Donachie, "Only the N-terminal domain of FtsK functions in cell division.," *Journal of bacteriology*, vol. 180, pp. 4621–7, Sept. 1998.

[55] X. C. Yu, E. K. Weihe, and W. Margolin, "Role of the C terminus of FtsK in Escherichia coli chromosome segregation.," *Journal of bacteriology*, vol. 180, pp. 6424–8, Dec. 1998.

[56] T. H. Massey, C. P. Mercogliano, J. Yates, D. J. Sherratt, and J. Löwe, "Double-stranded DNA translocation: structure and mechanism of hexameric FtsK.," *Molecular cell*, vol. 23, pp. 457–69, Aug. 2006.

[57] M. Stouf, J.-C. Meile, and F. Cornet, "FtsK actively segregates sister chromosomes in Escherichia coli.," *Proceeding of the National Academy of Sciences of the United States of America*, vol. 110, pp. 11157–62, July 2013.

[58] J. L. Ptacin, M. Nöllmann, C. Bustamante, and N. R. Cozzarelli, "Identification of the FtsK sequence-recognition domain.," *Nature structural & molecular biology*, vol. 13, pp. 1023–5, Nov. 2006.

[59] S. Bigot, O. A. Saleh, C. Lesterlin, C. Pages, M. El Karoui, C. Dennis, M. Grigoriev, J.-F. Allemand, F.-X. Barre, and F. Cornet, "KOPS: DNA motifs that control E. coli chromosome segregation by orienting the FtsK translocase.," *The EMBO journal*, vol. 24, pp. 3770–80, Nov. 2005.

[60] V. Sivanathan, M. D. Allen, C. de Bekker, R. Baker, L. K. Arciszewska, S. M. Freund, M. Bycroft, J. Löwe, and D. J. Sherratt, "The FtsK gamma domain directs oriented DNA translocation by interacting with KOPS.," *Nature structural & molecular biology*, vol. 13, pp. 965–72, Nov. 2006.

[61] T. H. Massey, L. Aussel, F.-X. Barre, and D. J. Sherratt, "Asymmetric activation of Xer site-specific recombination by FtsK.," *EMBO reports*, vol. 5, pp. 399–404, Apr. 2004.

[62] J. Yates, I. Zhekov, R. Baker, B. r. Eklund, D. J. Sherratt, and L. K. Arciszewska, "Dissection of a functional interaction between the DNA translocase, FtsK, and the XerD recombinase.," *Molecular microbiology*, vol. 59, pp. 1754–66, Mar. 2006.

[63] G. Blakely, G. May, and R. McCulloch, "Two related recombinases are required for site-specific recombination at dif and cer in E. coli K12," *Cell*, vol. 75, pp. 351–361, 1993.

[64] S. C. Y. Ip, M. Bregu, F.-X. Barre, D. J. Sherratt, E. Barre, and D. J. Sherratt, "Decatenation of DNA circles by FtsK-dependent Xer site-specific recombination.," *The EMBO journal*, vol. 22, pp. 6399–407, Dec. 2003.

[65] H. Klenow and I. Henningsen, "Selective elimination of the exonuclease activity of the deoxyribonucleic acid polymerase from Escherichia coli B by limited proteolysis.," *Proceeding of the National Academy of Sciences of the United States of America*, vol. 65, pp. 168–175, Jan. 1970.

[66] K. Ghosh, F. Guo, and G. D. Van Duyne, "Synapsis of loxP sites by Cre recombinase.," *The Journal of biological chemistry*, vol. 282, pp. 24004–24016, 2007.

[67] R. Roy, S. Hohng, and T. Ha, "A practical guide to single-molecule FRET.," *Nature methods*, vol. 5, pp. 507–16, June 2008.

[68] Y. Santoso, C. M. Joyce, O. Potapova, L. Le Reste, J. Hohlbein, J. P. Torella, N. D. F. Grindley, and A. N. Kapanidis, "Conformational transitions in DNA polymerase I revealed by single-molecule FRET.," *Proceeding of the National Academy of Sciences of the United States of America*, vol. 107, pp. 715–20, Jan. 2010.

[69] P. J. Hagerman and B. H. Zimm, "Monte Carlo approach to the analysis of the rotational diffusion of wormlike chains," *Biopolymers*, vol. 20, pp. 1481–1502, 1981.

[70] B. Tinland, A. Pluen, J. Sturm, G. Weill, and I. C. Sadron-cnrs universite, "Persistence Length of Single-Stranded DNA," *Macromolecules*, vol. 30, pp. 5763–5765, 1997.

[71] W. Lukosz, "Light emission by magnetic and electric dipoles close to a plane dielectric interface III Radiation patterns of dipoles with arbitrary orientation," *Journal of the Optical Society of America*, vol. 69, p. 1495, Nov. 1979.

[72] H. Lee, R. M. Venable, A. D. Mackerell, and R. W. Pastor, "Molecular dynamics studies of polyethylene oxide and polyethylene glycol: hydrodynamic radius and shape anisotropy.," *Biophysical Journal*, vol. 95, pp. 1590–9, Aug. 2008.

[73] M. H. Ulbrich and E. Y. Isacoff, "Subunit counting in membrane-bound proteins.," *Nature methods*, vol. 4, pp. 319–21, Apr. 2007.

[74] S. A. McKinney, C. Joo, and T. Ha, "Analysis of single-molecule FRET trajectories using hidden Markov modeling.," *Biophysical Journal*, vol. 91, pp. 1941–1951, 2006.

[75] J. E. Bronson, J. Fei, J. M. Hofman, R. L. Gonzalez, and C. H. Wiggins, "Learning rates and states from biophysical time series: a Bayesian approach to model selection," *Biophysical Journal*, vol. 97, pp. 3196–205, Dec. 2009.

[76] Y. Lu, B. Weers, and N. C. Stellwagen, "DNA persistence length revisited.," *Biopolymers*, vol. 61, no. 4, pp. 261–75, 2002.

[77] N. Pouget, C. Dennis, C. Turlan, M. Grigoriev, M. Chandler, and L. Salomé, "Single-particle tracking for DNA tether length monitoring.," *Nucleic Acids Research*, vol. 32, p. e73, 2004.

[78] H. Hwang, H. Kim, and S. Myong, "Protein induced fluorescence enhancement as a single molecule assay with short distance sensitivity," *Proceeding of the National Academy of Sciences of the United States of America*, vol. 108, no. 18, pp. 7414–8, 2011.

[79] B. Maier, D. Bensimon, and V. Croquette, "Replication by a single DNA polymerase of a stretched single-stranded DNA.," *Proceeding of the National Academy of Sciences of the United States of America*, vol. 97, pp. 12002–12007, 2000.

[80] J. J. Schwartz and S. R. Quake, "Single molecule measurement of the "speed limit" of DNA polymerase.," *Proceeding of the National Academy of Sciences of the United States of America*, vol. 106, pp. 20294–9, Dec. 2009.

[81] J. J. Champoux, "DNA topoisomerases: structure, function, and mechanism.," *Annual review of biochemistry*, vol. 70, pp. 369–413, Jan. 2001.

[82] M. T. J. van Loenhout, M. V. de Grunt, and C. Dekker, "Dynamics of DNA supercoils.," *Science*, vol. 338, pp. 94–7, Oct. 2012.

[83] T. T. Perkins, R. V. Dalal, P. G. Mitsis, and S. M. Block, "Sequence-dependent pausing of single lambda exonuclease molecules.," *Science*, vol. 301, pp. 1914–8, Oct. 2003.

[84] S. E. Nunes-Düby, H. J. Kwon, R. S. Tirumalai, T. Ellenberger, and A. Landy, "Similarities and differences among 105 members of the Int family of site-specific recombinases," *Nucleic Acids Research*, vol. 26, no. 2, pp. 391–406, 1998.

[85] F. Guo, D. N. Gopaul, and G. D. van Duyne, "Structure of Cre recombinase complexed with DNA in a site-specific recombination synapse.," *Nature*, vol. 389, no. 6646, pp. 40–46, 1997.

[86] D. N. Gopaul, F. Guo, and G. D. Van Duyne, "Structure of the Holliday junction intermediate in Cre-loxP site-specific recombination," *EMBO journal*, vol. 17, no. 14, pp. 4175–87, 1998.

[87] F. Guo, D. N. Gopaul, and G. D. Van Duyne, "Asymmetric DNA bending in the Cre-loxP site-specific recombination synapse.," *Proceeding of the National Academy of Sciences of the United States of America*, vol. 96, no. 13, pp. 7143–7148, 1999.

[88] F.-X. Barre, "FtsK functions in the processing of a Holliday junction intermediate during bacterial chromosome segregation," *Genes & Development*, vol. 14, pp. 2976–2988, Dec. 2000.

[89] I. Grainge, C. Lesterlin, and D. J. Sherratt, "Activation of XerCD-dif recombination by the FtsK DNA translocase.," *Nucleic Acids Research*, vol. 39, pp. 5140–5148, 2011.

[90] L. Aussel, F. X. Barre, M. Aroyo, A. Stasiak, A. Z. Stasiak, and D. Sherratt, "FtsK is a DNA motor protein that activates chromosome dimer resolution by switching the catalytic state of the XerC and XerD recombinases.," *Cell*, vol. 108, pp. 195–205, Jan. 2002.

[91] S. Doose, H. Neuweiler, and M. Sauer, "A close look at fluorescence quenching of organic dyes by tryptophan.," *Chemphyschem : a European journal of chemical physics and physical chemistry*, vol. 6, pp. 2277–85, Nov. 2005.

[92] H. Ferreira, D. Sherratt, and L. Arciszewska, "Switching catalytic activity in the XerCD site-specific recombination machine.," *Journal of molecular biology*, vol. 312, pp. 45–57, Sept. 2001.

[93] S. Jung and R. M. Dickson, "Hidden Markov analysis of short single molecule intensity trajectories," *Journal of Physical Chemistry B*, vol. 113, no. 42, pp. 13886–13890, 2009.

[94] S. Uphoff, K. Gryte, G. Evans, and A. N. Kapanidis, "Improved temporal resolution and linked hidden Markov modeling for switchable single-molecule FRET," *ChemPhysChem*, vol. 12, no. 3, pp. 571–579, 2011.

[95] L. Le Reste, J. Hohlbein, K. Gryte, and A. N. Kapanidis, "Characterization of dark quencher chromophores as nonfluorescent acceptors for single-molecule FRET," *Biophysical Journal*, vol. 102, no. 11, pp. 2658–2668, 2012.

[96] T. Ha, A. Y. Ting, J. Liang, W. B. Caldwell, A. A. Deniz, D. S. Chemla, P. G. Schultz, and S. Weiss, "Single-molecule fluorescence spectroscopy of enzyme conformational dynamics and cleavage mechanism.," *Proceeding of the National Academy of Sciences of the United States of America*, vol. 96, no. 3, pp. 893–898, 1999.

[97] A. Muschielok, J. Andrecka, A. Jawhari, F. Brückner, P. Cramer, and J. Michaelis, "A nano-positioning system for macromolecular structural analysis.," *Nature methods*, vol. 5, no. 11, pp. 965–971, 2008.

[98] S. Uphoff, S. J. Holden, L. Le Reste, J. Periz, S. van de Linde, M. Heilemann, and A. N. Kapanidis, "Monitoring multiple distances within a single molecule using switchable FRET.," *Nature methods*, vol. 7, no. 10, pp. 831–836, 2010.

[99] I. Rasnik, S. Myong, W. Cheng, T. M. Lohman, and T. Ha, "DNA-binding Orientation and Domain Conformation of the E. coli Rep Helicase Monomer Bound to a Partial Duplex Junction: Single-molecule Studies of Fluorescently Labeled Enzymes," *Journal of Molecular Biology*, vol. 336, no. 2, pp. 395–408, 2004.

[100] E. Crozat, A. Meglio, J.-F. Allemand, C. E. Chivers, M. Howarth, C. Vénien-Bryan, I. Grainge, and D. J. Sherratt, "Separating speed and ability to displace roadblocks

during DNA translocation by FtsK.," *The EMBO journal*, vol. 29, pp. 1423–33, Apr. 2010.

[101] P. J. Pease, O. Levy, G. J. Cost, J. Gore, J. L. Ptacin, D. Sherratt, C. Bustamante, and N. R. Cozzarelli, "Sequence-directed DNA translocation by purified FtsK.," *Science*, vol. 307, pp. 586–90, Jan. 2005.

[102] L. K. Arciszewska, R. A. Baker, B. Hallet, and D. J. Sherratt, "Coordinated control of XerC and XerD catalytic activities during Holliday junction resolution.," *Journal of molecular biology*, vol. 299, no. 2, pp. 391–403, 2000.

[103] R. E. Dale, J. Eisinger, and W. E. Blumberg, "The orientational freedom of molecular probes. The orientation factor in intramolecular energy transfer.," *Biophysical Journal*, vol. 26, pp. 161–93, May 1979.

[104] B. W. van der Meer, "Kappa-squared: from nuisance to new sense.," *Journal of biotechnology*, vol. 82, pp. 181–96, Jan. 2002.

[105] D. B. VanBeek, M. C. Zwier, J. M. Shorb, and B. P. Krueger, "Fretting about FRET: Correlation between κ and R," *Biophysical Journal*, vol. 92, pp. 4168–4178, June 2007.

[106] F.-X. Barre and D. J. Sherratt, "Chromosome dimer resolution," in *The Bacterial Chromosome* (N. P. Higgins, ed.), pp. 513–24, Washington, DC: American Society for Microbiology Press, 2005.

[107] S. Bigot, J. Corre, J.-M. Louarn, F. Cornet, and F.-X. Barre, "FtsK activities in Xer recombination, DNA mobilization and cell division involve overlapping and separate domains of the protein.," *Molecular microbiology*, vol. 54, pp. 876–86, Nov. 2004.

[108] J. E. Graham, V. Sivanathan, D. J. Sherratt, and L. K. Arciszewska, "FtsK translocation on DNA stops at XerCD-dif.," *Nucleic Acids Research*, vol. 38, pp. 72–81, Jan. 2010.

[109] J. E. Graham, D. J. Sherratt, and M. D. Szczelkun, "Sequence-specific assembly of FtsK hexamers establishes directional translocation on DNA.," *Proceeding of the National Academy of Sciences of the United States of America*, vol. 107, pp. 20263–8, Nov. 2010.

[110] O. Levy, J. L. Ptacin, P. J. Pease, J. Gore, M. B. Eisen, C. Bustamante, and N. R. Cozzarelli, "Identification of oligonucleotide sequences that direct the movement of the Escherichia coli FtsK translocase.," *Proceeding of the National Academy of Sciences of the United States of America*, vol. 102, pp. 17618–23, Dec. 2005.

[111] O. a. Saleh, C. Pérals, F.-X. Barre, and J.-F. Allemand, "Fast, DNA-sequence independent translocation by FtsK in a single-molecule experiment.," *The EMBO journal*, vol. 23, pp. 2430–9, June 2004.

[112] J. Y. Lee, I. J. Finkelstein, E. Crozat, D. J. Sherratt, and E. C. Greene, "Single-molecule imaging of DNA curtains reveals mechanisms of KOPS sequence targeting by the DNA translocase FtsK.," *Proceeding of the National Academy of Sciences of the United States of America*, vol. 109, pp. 6531–6, Apr. 2012.

[113] J. Y. Lee, I. J. Finkelstein, L. K. Arciszewska, D. J. Sherratt, and E. C. Greene, "Single-Molecule Imaging of FtsK Translocation Reveals Mechanistic Features of Protein-Protein Collisions on DNA.," *Molecular cell*, vol. 54, pp. 832–43, Apr. 2014.

[114] P. Bisicchia, B. Steel, M. H. Mariam Debela, J. Löwe, and D. Sherratt, "The N-terminal membrane-spanning domain of the Escherichia coli DNA translocase FtsK hexamerizes at midcell.," *mBio*, vol. 4, pp. e00800–13, Jan. 2013.

[115] A. N. Kapanidis, "Alternating-laser excitation of single molecules," in *Single-Molecule Techniques: A Laboratory Manual* (P. R. Selvin and T. Ha, eds.), pp. 85–119, Cold Spring Harbour, NY: Cold Spring Harbor Laboratory Press, 2008.

[116] E. Gasteiger, C. Hoogland, A. Gattiker, S. Duvaud, M. R. Wilkins, R. D. Appel, and A. Bairoch, "Protein Identification and Analysis Tools on the ExPASY Server," in *The Proteomics Protocols Handbook* (J. M. Walker, ed.), pp. 571–607, New York: Humana Press, 2005.

[117] G. W. Evans, J. Hohlbein, T. Craggs, L. Aigrain, and A. N. Kapanidis, "Real-time single-molecule studies of the motions of DNA polymerase fingers illuminate DNA synthesis mechanisms.," *Nucleic Acids Research*, vol. 43, pp. 5998–6008, July 2015.

[118] J. Periz, C. Celma, B. Jing, J. N. M. Pinkney, P. Roy, and A. N. Kapanidis, "Rotavirus mRNAS are released by transcript-specific channels in the double-layered viral capsid," *Proceeding of the National Academy of Sciences of the United States of America*, vol. 110, pp. 12042–7, July 2013.

[119] R. E. Kass and A. E. Raftery, "Bayes Factors," *Journal of the American Statistical Association*, vol. 90, pp. 773–95, Feb. 1995.

[120] K. Shimokawa, K. Ishihara, I. Grainge, D. J. Sherratt, and M. Vazquez, "FtsK-dependent XerCD-dif recombination unlinks replication catenanes in a stepwise manner.," *Proceeding of the National Academy of Sciences of the United States of America*, vol. 110, pp. 20906–11, Dec. 2013.

[121] D. I. Cattoni, S. Thakur, C. Godefroy, A. Le Gall, J. Lai-Kee-Him, P.-E. Milhiet, P. Bron, and M. Nöllmann, "Structure and DNA-binding properties of the Bacillus subtilis SpoIIIE DNA translocase revealed by single-molecule and electron microscopies.," *Nucleic Acids Research*, vol. 42, pp. 2624–36, Feb. 2014.

[122] D. I. Cattoni, O. Chara, C. Godefroy, E. Margeat, S. Trigueros, P.-E. Milhiet, and M. Nöllmann, "SpoIIIE mechanism of directional translocation involves target search coupled to sequence-dependent motor stimulation.," *EMBO reports*, vol. 14, pp. 473–9, May 2013.

[123] J. Löwe, A. Ellonen, M. D. Allen, C. Atkinson, D. J. Sherratt, and I. Grainge, "Molecular mechanism of sequence-directed DNA loading and translocation by FtsK.," *Molecular cell*, vol. 31, pp. 498–509, Aug. 2008.

[124] C. A. Hale and P. A. J. de Boer, "ZipA Is Required for Recruitment of FtsK, FtsQ, FtsL, and FtsN to the Septal Ring in Escherichia coli," *Journal of Bacteriology*, vol. 184, pp. 2552–6, May 2002.

[125] C. T. Diagne, M. Salhi, E. Crozat, L. Salomé, F. Cornet, P. Rousseau, and C. Tardin, "TPM analyses reveal that FtsK contributes both to the assembly and the activation of the XerCD-dif recombination synapse," *Nucleic Acids Research*, vol. 42, no. 3, pp. 1721–32, 2014.

[126] M. Besprozvannaya and B. M. Burton, "Do the same traffic rules apply? Directional chromosome segregation by SpoIIIE and FtsK," *Molecular Microbiology*, vol. 93, no. July, pp. 599–608, 2014.

[127] N. K. Lee, H. R. Koh, K. Y. Han, and S. K. Kim, "Folding of 8-17 deoxyribozyme studied by three-color alternating-laser excitation of single molecules.," *Journal of the American Chemical Society*, vol. 129, pp. 15526–34, Dec. 2007.

[128] C. Ratzke, F. Berkemeier, and T. Hugel, "Heat shock protein 90's mechanochemical cycle is dominated by thermal fluctuations.," *Proceeding of the National Academy of Sciences of the United States of America*, vol. 109, pp. 161–6, Jan. 2012.

[129] J. B. Munro, R. B. Altman, C.-S. Tung, J. H. D. Cate, K. Y. Sanbonmatsu, and S. C. Blanchard, "Spontaneous formation of the unlocked state of the ribosome is a multistep process.," *Proceeding of the National Academy of Sciences of the United States of America*, vol. 107, pp. 709–14, Jan. 2010.

[130] R. Roy, A. G. Kozlov, T. M. Lohman, and T. Ha, "SSB protein diffusion on single-stranded DNA stimulates RecA filament formation.," *Nature*, vol. 461, pp. 1092–7, Oct. 2009.

[131] C. E. Aitken, R. A. Marshall, and J. D. Puglisi, "An oxygen scavenging system for improvement of dye stability in single-molecule fluorescence experiments.," *Biophysical Journal*, vol. 94, pp. 1826–35, Mar. 2008.

[132] H. S. Chung, J. M. Louis, and W. A. Eaton, "Distinguishing between protein dynamics and dye photophysics in single-molecule FRET experiments.," *Biophysical Journal*, vol. 98, pp. 696–706, Feb. 2010.

[133] S. Lee, J. Lee, and S. Hohng, "Single-molecule three-color FRET with both negligible spectral overlap and long observation time.," *PloS one*, vol. 5, p. e12270, Jan. 2010.

[134] R. B. Altman, Q. Zheng, Z. Zhou, D. S. Terry, J. D. Warren, and S. C. Blanchard, "Enhanced photostability of cyanine fluorophores across the visible spectrum.," *Nature methods*, vol. 9, pp. 428–9, May 2012.

[135] J. Ross, P. Buschkamp, D. Fetting, A. Donnermeyer, C. M. Roth, and P. Tinnefeld, "Multicolor single-molecule spectroscopy with alternating laser excitation for the investigation of interactions and dynamics.," *The journal of physical chemistry. B*, vol. 111, pp. 321–6, Jan. 2007.