Extending and combining single-molecule fluorescence methods to study site-specific recombination

Justin N. M. Pinkney

Wolfson College, University of Oxford



Submitted for the degree of Doctor of Philosophy, Michaelmas 2012

Disclaimer

I declare that the work in this thesis is mine alone, expect where indicated in the text, and as described below.

- Elements of figure 2.3 were adapted, with permission, from previously published images.
- Three-colour microscope design and construction was in collaboration with Stephan Uphoff.
- Holliday junction oligonucleotides were provided by Stephan Uphoff.
- Elements of figure 4.1 were adapted, with permission, from previously published images.
- DNA substrates were prepared by Pawel Zawadzki.
- Klenow fragment, nucleotides and associated buffers were provided by Geraint Evans and polymerisation experiments were performed in collaboration.
- Chapter 5 was adapted from my publication (Pinkney et al., 2012).
- DNA substrates, Cre proteins and mutants were prepared by Pawel Zawadzki who also assisted in experimental repeats.
- FCS experiments were performed by Jaroslaw Mazuryk.
- XerCD, FtsK, mutants and DNA substrates were prepared by Pawel Zawadzki who also assisted with repeats of recombination and mutant data.

Justin Pinkney

December 2012

Acknowledgements

I am hugely indebted to Achilles, not only for being kind enough to take me on after an early false start to my DPhil, but also for his continued advice, encouragement and guidance throughout my studies.

Clearly much of the work in this thesis couldn't have been achieved without the fruitful collaboration that has developed with the Sherratt lab in Oxford Biochemistry and my thanks go to Pawel for his huge dedication and hard work, and Dave and Lidia who both provided much sage advice and guidance.

Also, being in such a collaborative and friendly group I have to acknowledge all members of the Kapanidis lab, who have not only shared their knowledge and opinions but also made the past two and a half years very enjoyable. In particular I have to mention those I worked extensively with: Rob, for being my original 'lab daddy', Stephan, for being the 'wonder boy' and a great set-up building buddy, Javier for our work on rotavirus, and Geraint, for many fun times in the dark with C+C Music Factory (and staying angry). Of course all the members of downstairs office, both Kapanidians and Tuckerites helped keep me sane, as did Meat Loaf, Jaffa Ball, Coconut Monkey, and Charlie.

It should also go without saying that I owe many thanks to Ciara and my parents, but I will certainly mention them here. And to anyone that I've forgotten to mention who feels omitted, thanks!

Abstract

Förster resonance energy transfer (FRET) has become an important tool for studying biochemical reactions at the single-molecule level, despite its increasing maturity there is an on-going effort to improve and expand the technique. This thesis presents methods for extending conventional two-colour single-molecule FRET measurements; by expanding the range and applicability of single-molecule fluorescence methods a greater variety of biological reactions can be studied, in greater detail than previously possible.

To circumvent the complexities of multi-colour FRET measurements and extend the range of observable distances I developed and characterised a new single-molecule fluorescence method termed tethered fluorophore motion (TFM). TFM is based on the existing technique of tethered particle motion (TPM) which relies on Brownian motion of a particle, attached to a surface by DNA, to probe the effective length of the DNA tether. TFM takes this concept and applies it at the single-fluorophore level, allowing simultaneous measurement of other fluorescence observables such as FRET and protein induced fluorescence enhancement (PIFE).

Having developed TFM I combined it with FRET to study site-specific recombinase proteins at the single-molecule level, in greater detail than possible by either technique alone. Studying the model tyrosine recombinase Cre, I extend and clarify previous ensemble observations regarding the order of DNA strand exchange, as well as uncovering a previously unobserved complex conformation and molecular heterogeneity. Finally, I used TFM-FRET to study the more complex XerCD recombination system and its interaction with the DNA translocase FtsK. I made observations, for the first time, of synaptic complex formation and of recombination at the single-molecule level, and these suggest intriguing and unexpected intermediates in the recombination reaction. I also combine TFM with PIFE to investigate the mechanism of DNA looping by FtsK.

The introduction of TFM, and its combination with other fluorescence techniques, allows observation of complex protein-DNA interactions from a variety of perspectives and will help expand the repertoire and applicability of single-molecule biophysical experiments.

Contents

1 Introduction **1**

2 Single-molecule fluorescence **4**

2.1 Introduction . . . 4

2.2 Fluorescence . . . 4

2.3 Förster resonance energy transfer . . . 5

2.4 Detecting single molecules . . . 7

2.4.1 Imaging single molecules . . . 7

2.4.2 Total internal reflection microscopy . . . 9

2.5 Extending single-molecule FRET . . . 11

2.5.1 Alternating laser excitation . . . 12

2.5.2 Multi-colour FRET . . . 12

2.5.3 Combination instruments . . . 13

3 Three-colour FRET **15**

3.1 Introduction . . . 15

3.1.1 Three-colour FRET . . . 15

3.1.2 Holliday junctions . . . 17

3.2 Materials and Methods . . . 18

3.2.1 Microscope design and implementation . . . 18

3.2.2 Software . . . 20

3.2.3 Single-molecule experiments . . . 20

3.2.4 Data analysis . . . 20

3.3 Results . . . 23

3.4 Discussion . . . 25

3.4.1 Construction of a three-colour microscope . . . 25

3.4.2 Holliday junction dynamics . . . 25

3.4.3 Challenges of three-colour smFRET . . . 26

3.4.4 Advantages of three-colour ALEX . . . 26

3.5 Contributions 27

4 Tethered fluorophore motion **28**

4.1 Introduction 28

4.1.1 Tethered particle motion 29

4.1.2 Tethered fluorophore motion 30

4.2 Experimental Methods 31

4.2.1 DNA and protein preparation 31

4.2.2 Instrumentation 31

4.2.3 Single-molecule experiments 32

4.2.4 Data analysis 32

4.2.5 Klenow fragment experiments 33

4.3 Results 34

4.3.1 Concept of TFM 34

4.3.2 Comparing two DNA lengths 35

4.3.3 Resolution 36

4.3.4 Simulating TFM observables 39

4.3.5 DNA polymerisation by Klenow fragment 42

4.4 Discussion 45

4.5 Contributions 47

5 Cre-*loxP* recombination **48**

5.1 Introduction 48

5.1.1 Site-specific recombination 48

5.1.2 Cre recombinase 49

5.1.3 Single-molecule studies of site-specific recombination 51

5.2 Experimental methods 52

5.2.1 DNA and protein preparation 52

5.2.2 Instrumentation 52

5.2.3 Sample preparation 54

5.2.4 Data analysis 54

5.2.5 Corrected FRET and distance calculation 55

5.3 Results 55

5.3.1 TFM-FRET observations of Cre-*loxP* recombination 55

5.3.2 Quantifying complex formation frequency 57

5.3.3 Bottom-strand synaptic complexes are preferentially assembled 60

5.3.4 Non-productive complexes 61

5.3.5 Evidence for fast interconversions between Holliday junction isomers 62

5.3.6 The recombinant product synaptic complex is extremely stable 65
5.4 Discussion 65
5.4.1 Reaction paths and a novel intermediate 67
5.4.2 Rate-limiting step for productive complexes 68
5.4.3 Holliday junction dynamics 68
5.4.4 Stable recombinant synaptic complexes 68
5.5 Conclusion 69
5.6 Contributions 70

6 XerCD-*dif* recombination **71**
6.1 Introduction 71
6.1.1 XerCD recombinase 71
6.1.2 FtsK 74
6.1.3 Activation of XerCD recombination by FtsK 75
6.2 Experimental methods 76
6.2.1 DNA and protein preparation 76
6.2.2 XerCD recombination 78
6.2.3 Ftsk DNA looping experiments 78
6.2.4 Data analysis 78
6.3 Results 79
6.3.1 XerCD synaptic complex 79
6.3.2 XerCD recombination 80
6.3.3 XerCD mutants 83
6.3.4 Looping of DNA by FtsK 87
6.4 Discussion 94
6.4.1 XerCD synaptic complex has a similar overall structure to Cre 94
6.4.2 'XerC active' synaptic complexes are formed prior to transition to an 'XerD active' complex 95
6.4.3 DNA looping by FtsK is not caused by aggregates 95
6.4.4 Looping is initiated in response to stoppage of FtsK on DNA 96
6.4.5 Proposed mechanism for looping 96
6.5 Future work 98
6.5.1 XerCD recombination 98
6.5.2 FtsK function 100
6.6 Contributions 100

7 Conclusion **101**

A Corrected FRET and distance calculation **103**

A.1 FRET correction 103

A.2 Distance determination 104

B Publications and presentations arising from this thesis **106**

B.1 Publications 106

B.2 Presentations 106

Chapter 1

Introduction

The extraordinary complexity of the environment and interactions in a living cell makes the study of proteins in their native surroundings a daunting and challenging task. Divorcing proteins from their natural habitat through *in vitro* experiments affords researchers with many orders of magnitude of simplification in potential interactions and an exquisite level of control over the environment. This control and simplification allows the action of single types of proteins and their interaction with other biological molecules (such as nucleic acids and proteins) to be studied in great detail, although the caveat that comes with *in vitro* experiments, that protein behaviour may differ significantly when in the cell, must always be kept in mind.

Traditional *in vitro* assays have provided a wealth of biological insight into the behaviour of biomolecules, and recent developments in technology have allowed *in vitro* studies of proteins at high resolution (X-ray crystallography), under fine force manipulation (atomic force microscopy, optical and magnetic tweezers), and undergoing dynamic conformational changes at the single molecule level (single-molecule fluorescence). Refinement of single-molecule methods has led to increasing ease of implementation and avenues previously difficult to access by ensemble means are now opening, such as the observation of transient intermediates as well as temporal and inter-molecular heterogeneity, and the ability to directly manipulate proteins under study.

As individual techniques become increasingly refined, the push for ever more complex and powerful instrumentation continues, driven by a desire to address ever more challenging biological systems in increasing detail. Although single-molecule techniques afford impressive resolution and fine control (and these will no doubt continue to improve) they are often limited to the finely tuned range of their detection, thus a drive toward combination instruments is under way to provide a more 'holistic'

view of protein action *in vitro*. For example, correlating high resolution distance information with functional data in real time opens the possibility of addressing how a protein's structure correlates with its function.

In this thesis I develop two extensions to existing two-colour single-molecule Förster resonance energy transfer (FRET) experiments. The first adds a third (blue) fluorophore, to provide a total of three possible distance measurements within a single-molecule, as demonstrated previously by Lee et al. (2007a). Here I implement this in an alternating laser excitation (ALEX) scheme which removes some of the confounding experimental factors typically encountered in such systems. The second uses ideas from an existing single-molecule technique (tethered particle motion) and applies them at the single-fluorophore level, allowing large-scale (> 30 nm) conformational changes of DNA to be observed simultaneously with the small scales typically monitored by single-molecule FRET ($1 - 10$ nm). I apply this combined method to study the mechanism of two site-specific recombinases, clarifying and extending previous ensemble and single-molecule studies, as well as capturing previously unseen steps in the reaction and revealing new mechanistic details.

In chapter 2 I introduce the principles of fluorescence, FRET and single-molecule detection and discuss some recent advances and extensions to the simple case of two-colour FRET. In chapter 3, I review existing efforts to extend FRET measurements to three (or more) fluorophores and describe the design and construction of a three-colour ALEX microscope as well as demonstrating its use in measuring multiple distance changes in a DNA nanostructure. In chapter 4 I explain the principles of a new single-molecule fluoresence technique, tethered fluorophore motion (TFM), quantify its temporal and spatial resolution, as well as demonstrating its use to measure the rate of DNA synthesis at the single-molecule level. Chapters 5 and 6 relate to the use of TFM in combination with FRET to study two site-specific recombinases. In chapter 5 I give a brief introduction to site-specific recombinases in general, and demonstrate the use of TFM-FRET to study a model recombinase, Cre, expanding existing knowledge and gaining new insight into the mechanism of the reaction. Finally, in chapter 6 I apply TFM-FRET to study a more complex recombinase system, XerCD, revealing its action for the first time at the single-molecule level. I also combine TFM with another single-molecule fluorescence technique, protein induced fluorescence enhancement (PIFE), to investigate the mechanism and relevance of previous observations of DNA looping by FtsK.

The results presented in this thesis outline the development and use of new combination single-molecule fluorescence methods. In particular, the development of TFM and its combination with both FRET and PIFE, as well as fluorescence intensity measurements, provides a novel and useful

methodology for addressing complex biological systems. Further, I demonstrate its utility by using TFM to address important biological questions regarding the action of site-specific recombinases.

Chapter 2

Single-molecule fluorescence

2.1 Introduction

Here I briefly describe fluorescence and its application to the study of bio-molecular reactions by the measurement of Förster resonance energy transfer (FRET). I also introduce the concept and advantages of studying reactions at the single-molecule level, and give an outline of how this is typically implemented with respect to fluorescence, as well as some of the limitations of the technique and how extensions and modifications have attempted to address these.

2.2 Fluorescence

Fluorescence arises due to the radiative relaxation of a molecule from an excited electronic singlet state to the ground state, illustrated in figure 2.1, the energy lost by the molecule is conserved by emission of a photon, allowing the process to be observed.[1]

As illustrated in Figure 2.1, a molecule can absorb a photon if its energy is sufficient to raise the molecule to one of the many excited state singlet levels, from here the molecule will rapidly decay to the lowest singlet state via vibrational relaxation. Fluorescence emission will generally occur from the lowest vibrational level of the excited singlet state to a higher vibrational level of the ground state, which will then relax rapidly. This gives rise to the Stokes shift, the observation that the peak of absorption for a fluorophore is at a shorter wavelength than the peak in emission, figure 2.2.

[1]There also exists an excited triplet state to which the system can relax via inter-system crossing, emission from these to the singlet ground state is spin forbidden, thus extremely slow, and termed phosphorescence (Lakowicz, 2006). Fluorescence and phosphorescence are grouped together as luminescence, and when this arises due to excitation by an incident photon the process is known as photoluminescence.

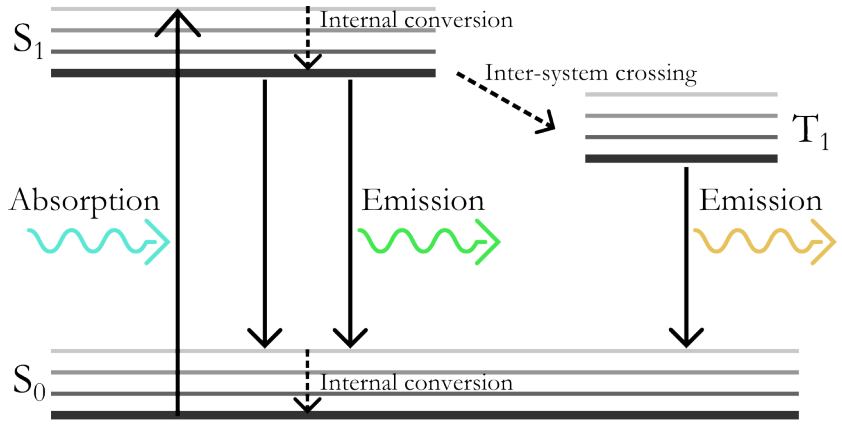


Figure 2.1: Jablonski diagram. Schematic of energy levels and transitions typical of fluorescence emission. S_0 is the singlet ground state, S_1 the excited singlet state, and T_1 the excited triplet state. The horizontal lines indicate different vibrational energy levels with the thickest indicating the lowest energy state for that level. Dashed arrows indicate non radiative transitions, solid arrows indicate radiative transitions.

This difference in absorption and emission wavelength proves useful experimentally as it means that excitation and emission light can be effectively separated.

2.3 Förster resonance energy transfer

A key physical phenomenon in translating fluorescence measurements into meaningful data from biological systems is Förster resonance energy transfer; FRET is the non-radiative transfer of energy from one fluorophore (donor) to another (acceptor) in close proximity (typically less than 10 nm). Energy is transferred via a dipole-dipole interaction and occurs at a rate,

$$k_{FRET} = \frac{1}{\tau_D}\left(\frac{R_0}{r}\right)^6, \tag{2.1}$$

where τ_D is the decay time of the donor in the absence of the acceptor, r is the distance between donor and acceptor, and R_0 is the Förster distance.

$$R_0 = \frac{9000\ln 10 Q_D \kappa^2 J}{128\pi^5 n^4 N_A} \tag{2.2}$$

$$J = \int f_D(\lambda)\epsilon_A(\lambda)\lambda^4 d\lambda, \tag{2.3}$$

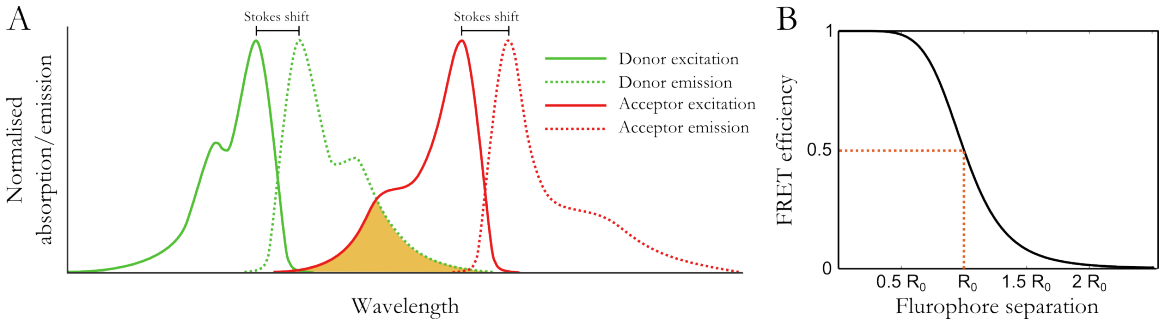


Figure 2.2: (A) Schematic representation of the donor and acceptor absorption and emission spectra, with spectral overlap integral highlighted in orange. (B) FRET efficiency dependence on inter-fluorophore separation, with the Förster distance, R_0, the distance at which FRET efficiency is 0.5, highlighted by the red dashed line.

Where κ^2 is the 'orientation factor' between the two fluorophores, Q_D is the quantum efficiency of the donor, n is the refractive index of the medium, N_A is Avagadro's constant and J is the normalized spectral overlap integral, see Figure 2.2, containing: the normalised donor emission, f_D, and the acceptor molar extinction coefficient ϵ_A, at a wavelength λ (Lakowicz, 2006).

FRET competes with the normal radiative decay process of the fluorophore, which occurs at a rate of $k_r = \frac{1}{\tau_D}$. Thus a FRET efficiency can be defined as

$$E = \frac{k_{FRET}}{k_{FRET} + k_r} \tag{2.4}$$

$$E = \frac{1}{1 + (\frac{r}{R_0})^6}. \tag{2.5}$$

As can be seen from equation 2.2, in the context of most biological experiments R_0, the distance at which $E = 0.5$, is a constant for a particular donor-acceptor pair, typically on the order of tens of angstroms (Gell et al., 2006). Thus the key, and most interesting, variable is the distance between fluorophores, r. This strong distance dependence enables FRET to detect distance changes occurring when fluorophores are within 1-10 nm (see Figure 2.2), an ideal range for the study of conformational dynamics of many proteins.

Several methods exist for determining the FRET efficiency from experimental observables (Lakowicz, 2006); during this work I calculate an uncorrected FRET efficiency from the observed fluorescence intensities by

$$E^* = \frac{F_{DA}}{F_{DA} + F_{DD}}. \tag{2.6}$$

The fluorescence intensities are written in the notation F_{XY} where X is the wavelength of excitation and Y denotes the emission channel (these are either donor, D, or acceptor A). Here I write E^* to denote that this has not been corrected for various experimental factors (such as spectral cross talk) so is not the true FRET efficiency. By performing appropriate controls and characterisation of the photophysical properties of the fluorophores involved, the value of E^* can readily be converted to an estimate of the true inter-fluorophore distance, see appendix A.

2.4 Detecting single molecules

In recent years advances in technology have allowed for the observation of fluorescence from single molecules with increasing ease and availability (Weiss, 1999; Roy et al., 2008). These observations of single biological molecules have opened entirely new possibilities for examination of biological processes due to the inherent advantages of observing individuals rather than the bulk. Single-molecule observations are free from the limitations of averaging, that necessarily occur at the ensemble level, this allows such methods to observe heterogeneity between different molecules, and temporal dynamics occurring within a single molecule, observations which are made difficult by ensemble averaging.

The key to successfully observing single fluorescent molecules is sufficient reduction of the illumination volume, such that background contributions (due to scattering from contaminants, out of focus light, and Raman scattering by water) fall below that of the signal from individual fluorophores. Single fluorophore measurements fall into two categories: those that observe molecules diffusing freely in solution, and those studying surface immobilized molecules. In this work I will be exclusively using total internal reflection (TIRF) microscopy a method to image single molecules attached to a surface.

2.4.1 Imaging single molecules

As light propagates through an optical imaging system, such as a microscope, the resolution attainable is limited by optical components, design and alignment, but more fundamentally, by the diffraction of light. The diffraction of light determines the smallest a point source can be focussed, and the resulting response of the system (i.e. the image that a point source forms at the detector) is named the point spread function (PSF). Thanks to the high quality optical components now readily available, the PSF of a typical single-molecule light microscope is generally diffraction limited, and its radius d,

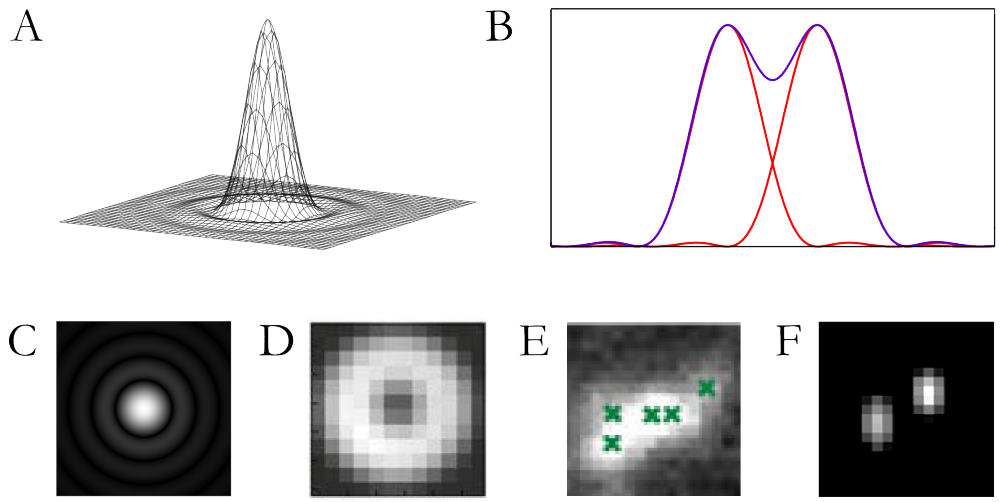


Figure 2.3: (A) Three dimensional render of the shape of the Airy disc due to imaging of a point source through a circular aperture. (B) Illustration of the Rayleigh criterion in one dimension, when the maximum of one Airy disc lies above the first minimum of another the two peaks can just be resolved, this is often taken to be the resolution limit of a conventional microscope. (C) Illustration of the intensity pattern expected due to an ideal diffraction limited PSF. Distortions away from the ideal PSF can be used to determine (D) the orientation of a fixed emitter, (E) the presence of multiple fluorophores, (F) the z-height of a molecule. Figure adapted from (A,C) Wikipedia: The free encyclopedia (2012), (D) Mortensen et al. (2010) (E) Holden et al. (2011) (F) Thompson et al. (2010)

and thus the microscopes ability to resolve objects, is governed by the Abbe diffraction limit,

$$d = \frac{\lambda}{2n\sin(\theta)}, \tag{2.7}$$

where λ is the wavelength of the light being focussed, n is the refractive index of the medium through which the objective lens images, and θ is maximum angle at which the objective lens can collect light ($n\sin(\theta)$ is the numerical aperture (NA) of the obejctive). This corresponds to a typical value of $200-250$ nm using a high NA objective, effectively blurring the image obtained, and limiting the resolution, figure 2.3.

Single fluorophores closely resemble point sources of light and the resulting PSF is generally taken to be a good approximation to an Airy disc; knowledge of the PSF resulting from a single fluorophore allows the position of the emitter to be determined with arbitrarily high precision[2]. This technique

[2]The localisation precision scales with the number of detected photons, N by $\sim \frac{1}{\sqrt{N}}$, and for typical single-molecule experiments is of the order 1 - 100 nm.

was introduced by Yildiz et al. (2003) to measure myosin steps with almost 1 nm precision, and the idea of high precision fluorophore localisation has since become widespread in biophysical studies.

The ability to localise single fluorophores with such a high degree of precisions, as well as the ability to control their activity using systems such as ROXS (Vogelsang et al., 2008), has paved the way for one class of approaches to exceeding the diffraction limit: localisation-based super-resolution (Huang et al., 2009). By spreading image acquisition temporally to ensure that any active fluorophores are sufficiently well separated that they can be individually localised, techniques such as stochastic optical reconstruction microscopy (STORM) are revolutionising light microscopy by regaining information lost in the conversion from a point source at the sample, to a PSF in the image.

This concept can be taken further than simply regaining the location of an emitter from its PSF. Knowledge of how various molecular properties of the emitter effect the observed PSF allows this information to be recapitulated merely from the shape of the molecular 'PSF'[3] on the camera. Information such as orientation (Mortensen et al., 2010), diffusion (Zareh et al., 2012), or the location of multiple, closely spaced, fluorophores (Holden et al., 2011) can be recovered from the PSF observed at the camera, figure 2.3. In fact the response of the PSF can deliberately be engineered (for example using extra optical elements or a spatial light modulator) to give a desired PSF change in response to parameters such as z-height (Pavani et al., 2009).

In Chapter 4 I introduce a method which utilises distortions in the PSF due to constrained diffusion of a fluorophore to detect changes in the properties of immobilised DNA molecules.

2.4.2 Total internal reflection microscopy

When light is incident on a boundary between two materials, the first with higher refractive index than the second, at a sufficiently oblique angle it will suffer total internal reflection. This can be derived from Snell's law,

$$n_1 \sin(\theta_i) = n_2 \sin(\theta_t), \tag{2.8}$$

where θ_i and θ_t are the angles of incidence and transmission (measured from the normal of the boundary), and n_1 and n_2 are the refractive indices of the two media of propagation. Thus total internal reflection occurs when θ_t is greater than 90°, this defines a critical angle above which total

[3]The resulting image of a molecule, such as a fluorophore, at the camera is not strictly the point spread function, as this is defined as the response of the system to an ideal point source emitter. Of course a fluorophore is not an ideal point, and under many conditions it does not closely resemble one, however, in general, the term PSF is often loosely applied to the image a fluorophore makes at the detector, and for convenience will be used as such throughout this thesis.

internal reflection must occur,

$$\theta_c = \arcsin\left(\frac{n_2}{n_1}\right) \tag{2.9}$$

Under these conditions an evanescent wave penetrates a small distance[4] into the second medium with a decay length,

$$I = I_0 \exp^{-\frac{z}{d}} \tag{2.10}$$

$$d = \frac{\lambda}{4\pi}\left(n_1^2 \sin^2(\theta) - n_2^2\right)^{-\frac{1}{2}} \tag{2.11}$$

typically in the range of 100 - 200 nm (Axelrod, 1981). In a TIRF microscope the glass coverslip and aqueous solution constitute the two materials and the incident laser light is coupled into the glass either via a prism (prism-type) or through a sufficiently high NA objective (objective-type), see figure 2.4. In prism-type TIRF a laser is introduced into a prism, coupled to the coverslip, at a sufficiently high angle, whereas in objective-type TIRF the laser beam must be focussed into the back focal plane of the objective and displaced perpendicular to the optical axis. Objective-type TIRF proves to be experimentally more facile to achieve, and is used exclusively throughout this work.

Single-molecule TIRF microscopy allows the simultaneous imaging of hundred of immobilised molecules with millisecond time resolution for acquisitions up to minutes long, this is made possible by the high speed and sensitivity of current EMCCD (and sCMOS) cameras, readily available high quality optical components, and extremely bright and stable organic fluorophores (Reck-Peterson et al., 2010). Single-molecule fluorescence TIRF observations were first made by Funatsu et al. (1994) to visualise the hydrolysis of ATP by single myosin motors in 1994, followed by the use of TIRF to measure single molecule FRET by Zhuang et al. (2000). Since then the single molecule FRET has been widely applied to single-molecule imaging both *in vivo* and *in vitro* (for examples see Kapanidis and Strick (2009)). The ability to collect extended histories of immobilised molecules, in a highly parallel fashion, makes TIRF an extremely appealing single-molecule imaging modality.

[4] As well as the reduction in illumination volume to within ~ 100 nm of the boundary, the use of TIRF can also provide a substantial enhancement to the intensity of the evanescent field as compared to the incident light (Axelrod et al., 1984).

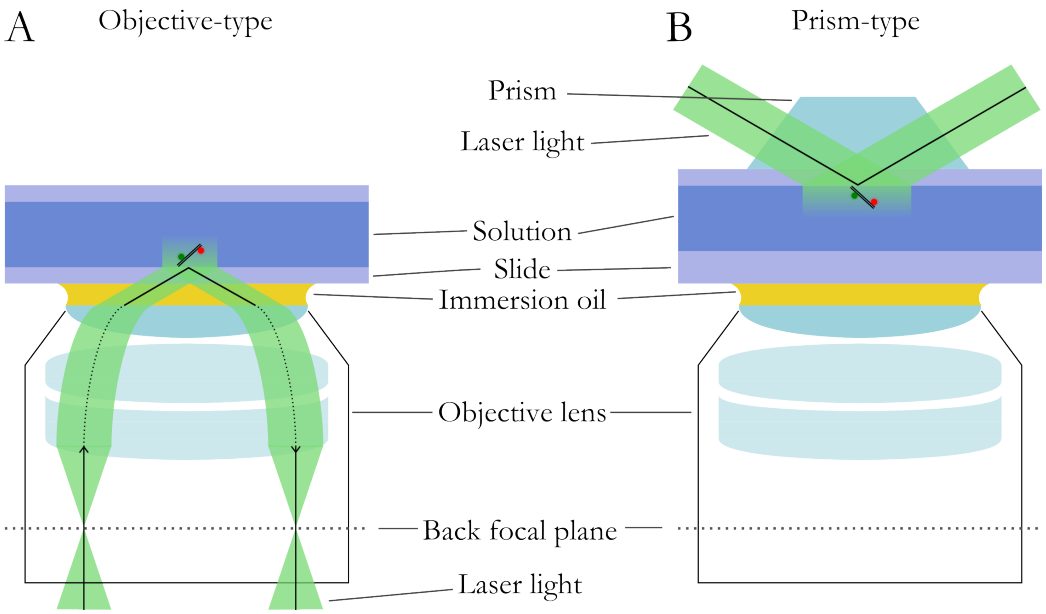


Figure 2.4: Schematic illustrations of (A) objective-type TIRF, where excitation light is introduced by the same objective used to collect the emission light, (B) prism-type TIRF, where a glass or quartz prism is used to introduce the excitation light at, or above, the critical angle.

2.5 Extending single-molecule FRET

Although the use of single-molecule FRET has shed light on many biochemical processes, the nature of the basic technique means that it comes with some inherent limitations. The measurement of FRET between a single donor and acceptor fluorophore provides information of the inter-fluorophore separation; correlating this single distance to complex conformational changes of biomolecules can be challenging without prior knowledge provided by techniques such as crystallography. The physical properties of the organic fluorophores typically used in single-molecule studies are never perfect and photo-physical effects such as blinking and bleaching can be confused with genuine signals arising from FRET fluctuations. Finally, the strong distance dependence of FRET in the range 1-10 nm, as well as being its strength, also limits many of the biological process that can be observed, and often correlating the observed conformational dynamics as measured by FRET, with other processes such as large-scale conformational changes and protein function are crucial in understanding the complete biological reaction, and how distance changes revealed by FRET correspond to progress through the reaction pathway. In order to overcome these limitations many adaptations and extensions to the basic two-colour single-molecule FRET method have been developed.

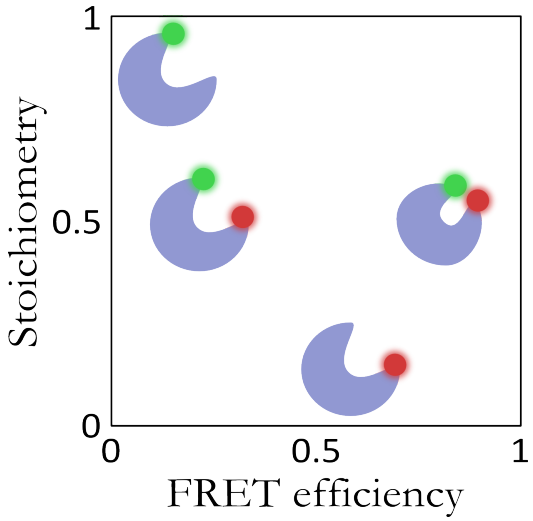


Figure 2.5: Schematic of a FRET efficiency-stoichiometry histogram (ES histogram). Donor only or acceptor only molecules have a stoichiometry value of 1 and 0 respectively so can be effectively separated from double labelled molecules along the y-axis. FRET efficiency along the x-axis can be used to detect distance changes or differences in the population of doubly labelled molecules.

2.5.1 Alternating laser excitation

An early addition to a typical FRET scheme was the use of alternating laser excitation (ALEX)(Kapanidis et al., 2004); rapid switching (nano- to milli-seconds, (Laurence et al., 2005; Müller et al., 2005; Margeat et al., 2006)) from donor to acceptor excitation wavelengths allows the state of the acceptor molecule to be probed, and thus collection of information usually inaccessible to single-laser FRET experiments. ALEX introduces a new ratiometric variable, stoichiometry, defined as:

$$S = \frac{F_{DA} + F_{DD}}{F_{DA} + F_{DD} + F_{AA}} \tag{2.12}$$

Directly probing the acceptor enables: molecular sorting to distinguish between acceptor only and low FRET molecules (figure 2.5); observation of acceptor photophysics, which may otherwise lead to misleading FRET efficiency signals (Heilemann et al., 2005; Chung et al., 2010; Di Fiori and Meller, 2010; Kalinin et al., 2010); and a facile method for extracting the information required to calculate accurate FRET efficiencies (Lee et al., 2005). In this work I further extend the use of ALEX to give information about the dynamic freedom of the acceptor fluorophore in chapter 4.

2.5.2 Multi-colour FRET

A logical extension to the typical two-colour FRET experiment is to add more, spectrally distinct, fluorophores in order to provide information on more than a single distance. This approach has the potential to examine more complex structural re-arrangements of proteins and also the possibility of using multiple distances to triangulate the positions of individual fluorophores (and thus the area of

a bio-molecule to which they are attached) in three dimensional space. Although the three-colour approach is potentially powerful, it also brings extra instrumental and experimental complications and has only begun to be applied to the study of protein-DNA interactions, as discussed further in Chapter 3. Recent advances in multiple distance FRET have included switchable FRET (Uphoff et al., 2010), using multiple identical photo-switchable acceptors to stochastically measure multiple FRET distances, as well as the extension of the technique to as many as four spectrally distinct fluorophores (Derocco et al., 2010; Lee et al., 2010a; Stein et al., 2011).

2.5.3 Combination instruments

As single-molecule techniques become more developed and their implementation more routine, the push to develop more advanced instruments, which can address a greater variety of biological questions continues in the emergence of several 'combination' single-molecule methods. These combination instruments put together the often disparate observation and manipulation capabilities of separate techniques to provide greater information and often a more complete view of biochemical reactions. The adaptability of single-molecule fluorescence has seen it combined in many contexts with other methodologies such as atomic force microscopy (AFM)(Cordes et al., 2010), optical (Ishijima et al., 1998; Lang et al., 2003; Comstock et al., 2011; Sirinakis et al., 2012; Lee et al., 2012b) and magnetic tweezers (del Rio et al., 2009), and electrical recordings (Borisenko et al., 2003; Heron et al., 2009; Harriss et al., 2011). However the added complexity of FRET measurements has seen fewer examples of its use in conjunction with other single-molecule techniques (Roy et al., 2008; Hohlbein et al., 2010); despite this, some pioneering studies have demonstrated the utility of combining the radically different (and often complimentary) distance scales and abilities of techniques such as AFM (He et al., 2012), optical and magnetic tweezers (Shroff et al., 2005) and single-channel recordings with FRET.

The combination of FRET with AFM and tweezers (optical and magnetic) allows the experimenter to directly manipulate molecules under investigation. Thus the response of conformational dynamics, as measured by FRET, can be correlated to the applied force and this has been used to explore the energy landscape of DNA junctions and hairpins (Hohng et al., 2007; Tarsa et al., 2007). Single-molecule FRET has also been demonstrated in conjunction with the progenitor of single-molecule techniques, single ion channel recordings, this has allowed the function of molecules under investigation to be observed, an important step toward directly correlating structure with function in real-time (Borisenko et al., 2003; Harms et al., 2003).

These complex combination methods pose significant challenges and typically require modification of a single-molecule fluorescence microscope to incorporate extra optical elements (optical tweezers) or entirely independent instrumentation (magnetic tweezers, AFM, etc). This complexity of implementation, as well as other limitations such as photo-bleaching caused by the optical tweezer trapping laser, or the inherently serial nature of some of the techniques, has meant that combination techniques (especially utilising FRET) have seen limited use, and have only recently been implemented by the most cutting edge biophysics laboratories. Despite this, the combination of single-molecule FRET with other techniques to provide more information about the function, force response, or global conformation of biomolecules, remains highly appealing. In Chapter 4 I introduce a method for simultaneously measuring large-scale conformational changes of DNA, simultaneous to single-molecule FRET measurements, which requires no additional instrumentation or adaptation to existing single-molecule fluorescence microscopes to implement.

Chapter 3

Three-colour FRET

3.1 Introduction

In order to expand the possible range of *in vitro* single molecule FRET experiments that could be performed, and the biological systems which could be addressed within the group, I designed and built a three-colour single-molecule FRET TIRF microscope equipped with ALEX, in collaboration with Stephan Uphoff. In this chapter I outline my contributions to the design and construction of the microscope and its validation by the observation of multiple co-ordinated distance changes in single DNA Holliday junctions.

3.1.1 Three-colour FRET

As mentioned in chapter 2, single-molecule two-colour FRET has seen a great variety of applications to the study of biological systems in recent years due to its sensitivity on distance scales relevant to macromolecules, and its relative ease of implementation (Roy et al., 2008). Although it has been widely used to gain insight into the mechanism of various biological processes, the inherent one dimensional nature of the distance measured, can limit interpretation of results and its potential applications. Given that FRET will occur between any two fluorophores with appropriate spectral properties, an obvious generalisation of the method is to extend its use to more than the typical single pair of fluorophores.

Multi-colour FRET schemes, in theory, allow the simultaneous measurement of many[1] inter-

[1] In general the number of distances that can be measured, in principle is the binomial coefficient $\binom{n}{k}$, where n is the total number of fluorophores and $k = 2$.

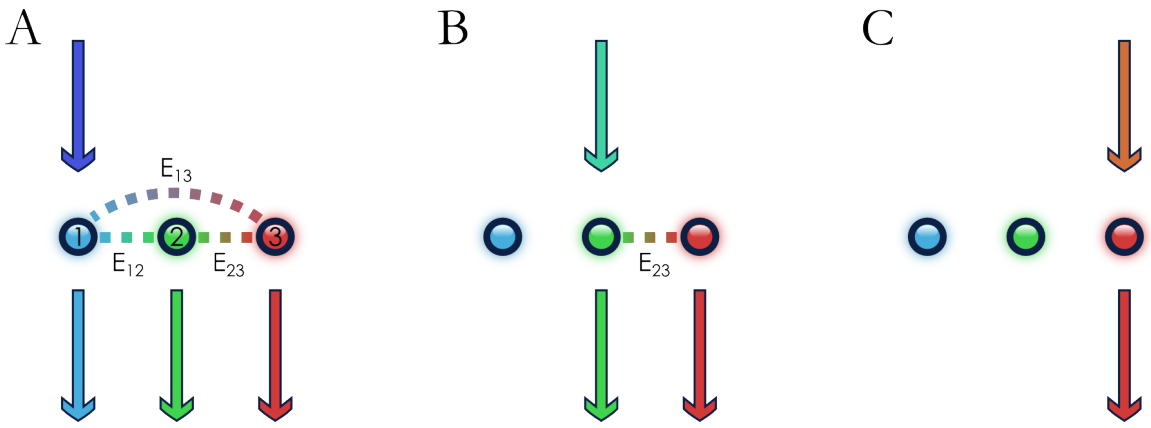


Figure 3.1: Schematic of the general FRET scheme, with three spectrally distinct fluorophores labelled, 1, 2, 3. (A) Excitation of the shortest wavelength fluorophore, 1, leads to three possible FRET pathways, E_{12}, E_{23}, and E_{13}. (B) Excitation of the middle fluorophore, 2, only permits one FRET pathway, E_{23}. (C) Direct excitation of the longest wavelength fluorophore does not lead to any FRET, but directly confirms the presence of fluorophore 3.

fluorophore distances and thus the triangulation and correlation of multiple moving components in a biomolecular reaction. In a general three-colour experiment, FRET can occur between each pair of fluorophores resulting in three measurable efficiencies, see figure 3.1. However, in order to for significant FRET to occur between all fluorophores, the spectral overlap, J, between each pair must be appreciable. This introduces the conflicting requirements of sufficient spectral overlap, with the need to adequately separate excitation from emission light. Further, experiments using a single laser to excite the shortest wavelength fluorophore (fluorophore 1, figure 3.1) cannot deconvolve FRET occurring directly from fluorophore 1 to 3 (E_{13}), to that occurring indirectly via fluorophore 2 (E_{12} followed by E_{23}). To overcome this difficulty several studies have employed the 'bifurcate' or 'alternative acceptor' scheme, where fluorophores and expected distances are chosen such that E_{23} is negligible, this scheme simplifies analysis and has been applied to the study of DNA Holliday junctions (Hohng et al., 2004), and several protein-DNA interactions (Roy et al., 2009; Munro et al., 2010; Ratzke et al., 2011).

Although this 'alternative acceptor' scheme simplifies the experimental set up required, it fails to take full advantage of the three potential FRET observables. To deconvolve the complex energy transfer paths possible in the general three-colour scheme it is necessary to excite fluorophore 2 directly to measure E_{23}, and from this the true value of E_{13} can be calculated, this 'semi-ALEX' (i.e. fewer excitation lasers, than spectrally distinct fluorophores) scheme has been used to measure three distances simultaneously within DNA Holliday junctions (Lee et al., 2010b).

Three-colour FRET experiments are particularly susceptible to the confounding effects of photophysical fluctuations of the fluorophores, and inability to distinguish between low FRET and unlabelled or bleached complexes (a problem exacerbated by the restrictions in the possible range of R_0 using three spectrally distinct fluorophores). A full three-colour ALEX scheme sequentially excites each of the fluorophores allowing, not only the detection of intensity fluctuations of fluorophore 3 and measurement of all possible FRET efficiencies, but also the virtual sorting of molecular species (Lee et al., 2007a). Due to imperfect labelling and photobleaching, there will always be a fraction of non-triply labelled molecules observed in experiments. By extending the concept of stoichiometry explained in chapter 2, to three separate ratiometric comparisons of fluorophore intensity (see equation 3.2.4) ALEX allows for filtering based on the presence of fluorophores, after the fact, allowing only those molecules of interest to be used for data analysis. The presence of non-fully labelled molecules also opens the possibility of performing the necessary controls and corrections for computing corrected FRET within a single experiment (Lee et al., 2005). Three-colour ALEX experiments have been applied to the study of DNA junctions (Lee et al., 2007b) and RNA polymerase translocation (Lee et al., 2007a).

The challenges of multiple wavelength excitation and detection have led to several instrumental design innovations including the use of small, laterally offset, broadband mirrors (Friedman et al., 2006), or an acousto-optical beam splitter (AOBS) (Ross et al., 2007) to introduce the excitation light and an expansion in types of fluorophores used in single-molecule experiments (Kim et al., 2010; Derocco et al., 2010). The use of three-colour FRET experiments to gain new insight into biological systems is slowly becoming more prevalent as the significant technical and experimental challenges are ameliorated. In fact, use of four spectrally distinct fluorophores has also recently been demonstrated (Lee et al., 2010a), however the further introduction of fluorphores, as well as increasing the amount of information obtainable, increases the complexity and difficulty in both acquiring and analysing data.

3.1.2 Holliday junctions

Holliday junctions are DNA junctions containing four duplex branches, comprised of four individual strands of DNA. Each strand base pairs with two partner strands, thus Holliday junctions can only form between appropriate sequences of DNA, which posses the correct complementarity to form the four-way structure. These junctions have been well studied at the single-molecule level (McKinney et al., 2002; Hohng et al., 2004, 2007), in part due to the ease of labelling and synthesis of the

molecules, but also due to the crucial role they play in recombination processes within cells. Holliday junctions are an important intermediate in homologous recombination, facilitating the exchange of complementary DNA between homologous strands (Alberts, 2008); their role in this process has been studied using a variety of single-molecule techniques, including some of the advanced combination instruments discussed in chapter 2. These studies have show that in the presence of di-valent ions, such as magnesium, Holliday junctions interconvert between two stacked conformers, with rates primarily dependant on di- and mono-valent ion concentrations (McKinney et al., 2002). These studies have led to Holliday junctions becoming the archetypal three-colour FRET reference sample (Hohng et al., 2004; Person et al., 2009; Lee et al., 2010b,a; Sobhy et al., 2011), as the conformational changes occurring during this isomerisation are at distances and timescales amenable to FRET measurement.

In this chapter I introduce a single-molecule TIRF microscope capable of implementing three-colour ALEX, and validate the microscope using synthetic DNA Holliday junctions. Later, the issue of Holliday junction isomerisation will recur as they are also a crucial intermediate of the site-specific recombination processes, studied in chapters 5 and 6 (Grindley et al., 2006).

3.2 Materials and Methods

3.2.1 Microscope design and implementation

In collaboration with Stephan Uphoff (Kapanidis group) I designed, constructed and aligned a three-colour ALEX microscope. The microscope was designed to be highly compact and a standard scheme for 'objective-type' TIRF (see chapter 2) was chosen due to the reduced alignment time required, the ease of working with conventional coverslips, and the accessibility provided to the sample from above, for buffer exchange, and the possiblity of adding further instrumentation (e.g. a DIC illumination source). Lasers and filters were chosen appropriate for observation of the red and green fluorophores already commonly used in the group (Cy3B, ATTO647N, Cy5) and also with typical blue organic fluorophores used in single-molecule experiments (ATTO488, Alexa488). Initial design concepts were outlined in computer aided design software (Sketchup, Google) and any custom component were machined in the research workshop (Oxford Physics).

A schematic of the excitation module is shown in figure 3.2. Three lasers (473 nm (Vortran Stradus), 543 nm (Laser 2000), and 640 nm (Cube Coherent)) are directed via broadband mirrors

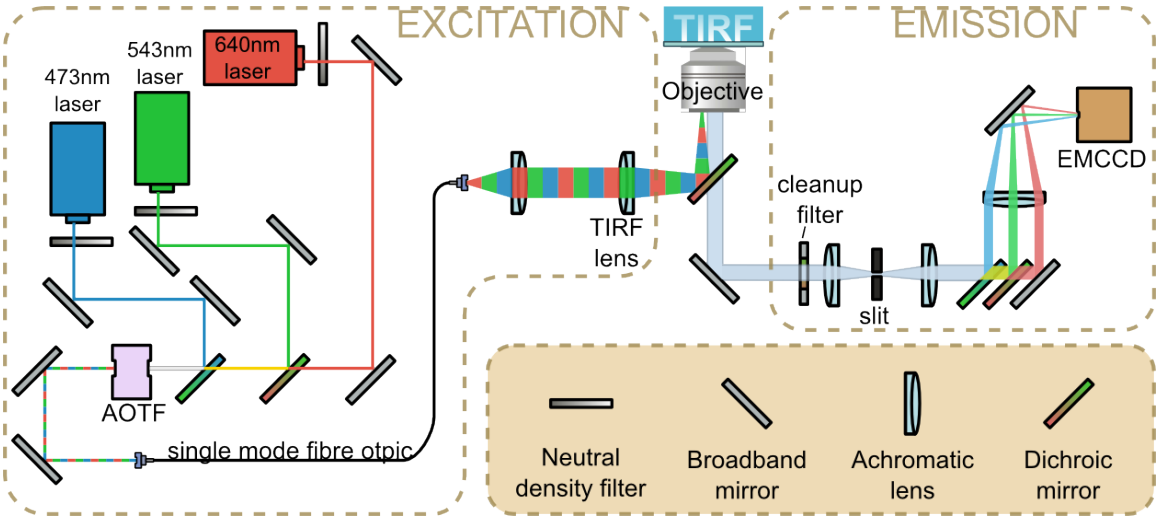


Figure 3.2: Illustration of the excitation and emission path of the three-colour FRET ALEX TIRF microscope 'Zugspitze'. The overall position of optical elements is shown as described in section 3.2.1

and combined via dichroic mirrors to pass through an acousto-optical tunable filter (AOTF). An AOTF uses a birefringent crystal which can diffract specific wavelengths of light depending on the acoustic frequency applied to it (via a piezoelectric device). The AOTF is used to achieve ALEX, both for wavelength selection, and intensity control of the laser light, the appropriate RF signals are generated using an RF driver (Gooch and Housego), which is controlled by a CompactDAQ board (National Instruments) run by custom written LabView software, see section 3.2.2. The AOTF deflects the selected beam/s into a single mode optical fibre (Schäfter + Kirchhoff), which outputs the light as a uniform gaussian beam, this is collimated and then focussed by the 'TIRF lens', off-axis, into the back focal plane of the objective as outlined in figure 3.2. The lateral position of the 'TIRF lens' can be altered to control the incident angle of the beam at the coverslip surface, and thus the depth of the evanescent illumination, see chapter 2.

Emitted light is collected by the same objective used to introduce the excitation light, passes through a dichroic mirror and then a cleanup filter (Chroma) to remove any remaining laser light. The emission light is then focussed on to a slit to crop the image for side-by-side imaging of all three emission channels on a single EMCCD chip (Andor iXon Ultra 897). So as to reduce both the complexity and loss of signal, a system to separate the emission light using a minimal number of optical components was designed. Two dichroic and one broadband mirror placed in close succession in the emission path serve to spatially segregate the emission into blue, green, and red channels. These are then directed and focussed onto the EMCCD. The camera is placed such that focus is

achieved at the same time as optimal TIRF illumination, and the magnification chosen (138x) was somewhat less than for the similar two-colour microscope already in the lab in order to partially compensate for the reduced field of view due to the three emission channels.

3.2.2 Software

An ALEX microscope requires that laser alternation periods are synchronised with camera frame acquisition. To achieve this a single cDAQ module (National Instruments) with both digital and analogue I/O boards was chosen as the control both the RF generator (and thus AOTF) and camera acquisition, due to its compactness and flexibility. I wrote custom LabView software to enable control of laser intensity, and to synchronise switching with camera frames, figure 3.4. The software was written so that the laser alternation scheme employed could easily be modified by the user, allowing facile switching between three or two colour modes of operation, and extra options to allow non-standard excitation schemes, or stroboscopic illumination were built into the software. The ease of use and flexibility of the software I implemented has seen its use extended to the existing two-colour microscope in the lab, figure 3.4.

3.2.3 Single-molecule experiments

Holliday junctions (HJs) were constructed according to Uphoff et al. (2010) where one of internally labelled Cy5 ('i-Cy5') was replaced with an Alexa Fluor 488 dye, figure 3.5. Strands X, R and H were kindly donated already labelled by Stephan Uphoff, and I performed labelling of strand B with Alexa Fluor 488. HJs were assembled by annealing all oligonucleotides at equimolar concentrations. HJs were tethered to a PEG-passivated glass coverslip using a biotin-neutravidin interaction, and were imaged in a buffer consisting of 50 mM Tris-HCl pH 7.5, 200mM $MgCl_2$, 2 mM Trolox, 100 μgml^{-1} BSA, 1 $mgml^{-1}$ glucose oxidase, 40 μgml^{-1} catalase and 1.4% (w/v) glucose. The total laser powers were measured to be 2 mW before entry into the objective and alternation times of 33 ms were used throughout.

3.2.4 Data analysis

All single-molecule fluorescence movies were analysed using a version of TwoTone (Holden et al., 2010) which Stephan Uphoff and I adapted to accept three-colour input. The positions of fluorophores



Figure 3.3: Photographs of the Zugspitze microscope.

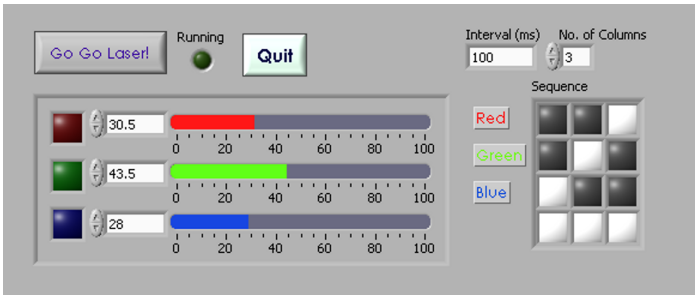


Figure 3.4: GUI for the custom written LabView software to control the three-colour microscope. The software synchronises alternation of lasers with the camera frames, as well as providing control of laser intensity, alternation pattern, and the triggering of the camera.

X strand	Cy3B - CCC AGT TGA GAG CTT GAT AGG G
B strand	CCC TAX CAA GCC GCT GCT ACG G
R strand	Cy5 - CCC ACC GCT CTT CTC AAC TGG G
H strand	Biotin - CCG TAG CAG CGA GAG CGG TGG G

Figure 3.5: DNA sequences used for construction of three-colour Holliday junctions, sequences are displayed 5'-3' with any modifications indicated. The red X marks the amino-C6-dT modification used for labelling of the Alexa Fluor 488 dye.

in the microscope images is determined by filtering the EMCCD output with a band pass filter and fitting gaussians to the resulting image. The positions of detected fluorophores in each channel are spatially associated using a calibration file, and are filtered on user specified grounds of eccentricity, size and nearest neighbour proximity. These filtered positions are then used to fit a gaussian profile to the observed PSF of indiviual fluorophores in each frame of the movie, and the area under this gaussian is taken to be the fluorescence intensity. From these intensities the ratiometric observables of FRET efficiency and stoichiometry are calculated as

$$E^*_{BG} = \frac{F_{BG}}{F_{BG} + F_{BB}} \tag{3.1}$$

$$E^*_{BR} = \frac{F_{BR}}{F_{BR} + F_{BB}} \tag{3.2}$$

$$E^*_{GR} = \frac{F_{GR}}{F_{GR} + F_{GG}} \tag{3.3}$$

$$\tag{3.4}$$

$$S_{BG} = \frac{F_{BB} + F_{BG} + F_{BR}}{F_{BB} + F_{BG} + F_{BR} + F_{GG} + F_{GR}} \tag{3.5}$$

$$S_{BR} = \frac{F_{BB} + F_{BG} + F_{BR}}{F_{BB} + F_{BG} + F_{BR} + F_{RR}} \tag{3.6}$$

$$S_{GR} = \frac{F_{GG} + F_{GR}}{F_{GG} + F_{GR} + F_{RR}} \tag{3.7}$$

$$\tag{3.8}$$

Using a similar notation to Chapter 2, where the fluorescence emission F_{XY} corresponds to excitation of the X and emission by Y fluorophore (where B, G, and R stand for blue, green, and red fluorophores)

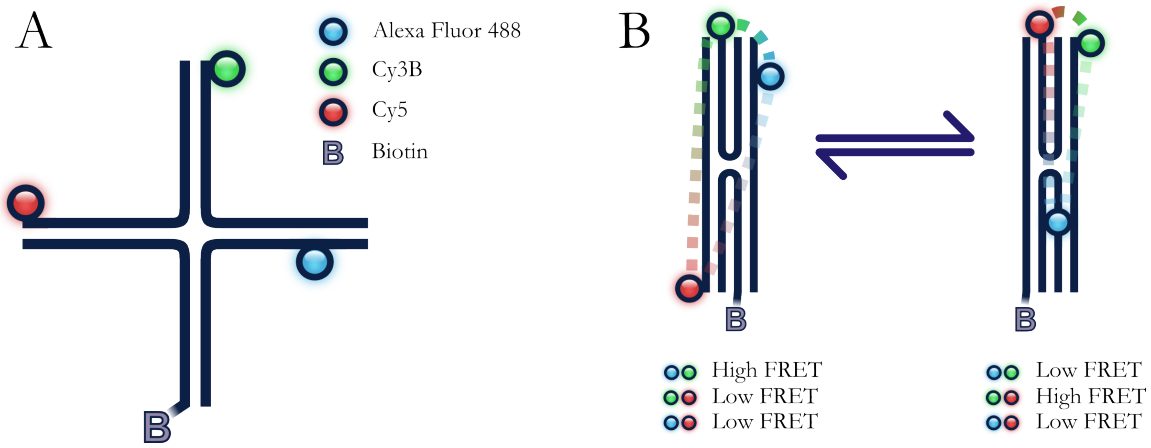


Figure 3.6: (A) Schematic of labelling positions of three-colour Holliday junctions. (B) Two stacked conformations expected to be adopted and the anticipated approximate FRET efficiencies between all fluorophores.

3.3 Results

To test and validate the new microscope set up DNA Holliday junctions were chosen as a test sample which would allow the measurement of multiple intra-molecular movements simultaneously by observing the three possible FRET paths. Four olignucleotide strands were used to form synthetic Holliday junctions, three of which were labelled with an organic fluorophore, and the fourth with a biotin moeity for surface attachment, figure 3.5. The labelling positions were chosen such that there would be measurable FRET between all fluorophores within the molecule and the two stacked isomers would each have distinct FRET signatures, figure 3.6.

After using the stoichiometry parameters to perform molecular sorting, so as to only inspect those molecules that contained three active fluorophores, clear intensity fluctuations could be observed for the majority of molecules. Calculation of uncorrected FRET efficiencies (E^*) showed the distinct anti-correlated changes in FRET from blue to green, and green to red fluorophores, figure 3.7. Although fluctuations were apparent in E_{BR}, due to the high level of background (partly due to auto-fluorescence from the objective in the red emission channel, under blue excitation), the small level of spectral overlap and the relatively minor distance changes expected between the fluorophores, FRET efficiency changes were less distinct and could not readily be separated into two states. Plotting a three dimensional FRET efficiency histogram, figure 3.8, demonstrates the two distinct conformation the Holliday junction can assume and also the sequence dependant bias to one conformer known to exist (Hohng et al., 2004; Uphoff et al., 2010).

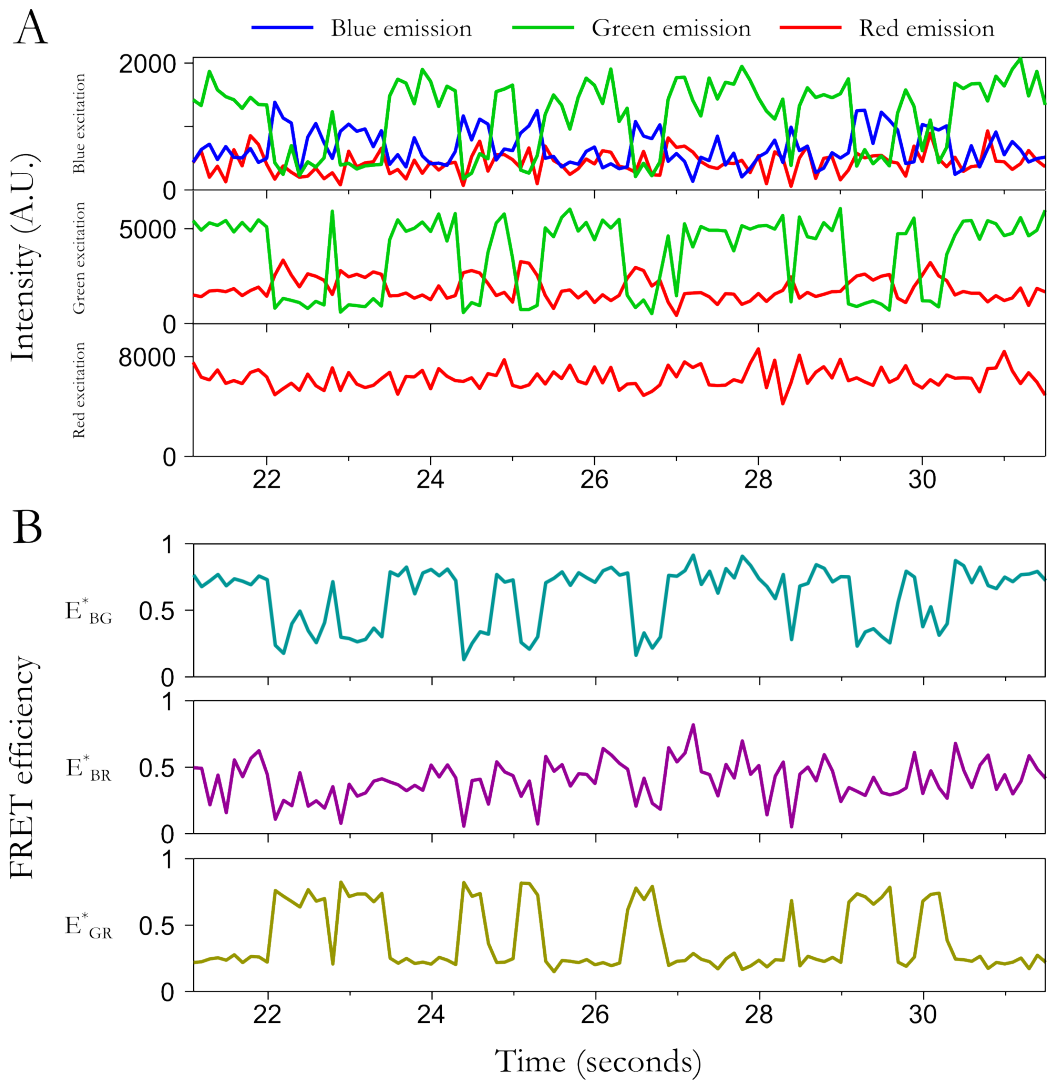


Figure 3.7: Example time-trace of single-molecule Holliday junction dynamics. (A) Emission intensity under blue, green, and red excitation, emission band is indicated by colour. Anti-correlated changes in fluorescence emission under one excitation wavelength indicates the occurrence of FRET; the stable emission under red excitation shows that the acceptor is not undergoing photo-physical fluctuations in emission intensity. (B) Uncorrected FRET efficiencies (E^*) between blue-green, blue-red, and green-red fluorophores. E^*_{BG} and E^*_{GR} show clear step changes in FRET efficiency as the Holliday junction isomerises. E^*_{BR} is not expected to show large changes in efficiency between the two conformers and extraneous background in this channel masks any subtle changes in FRET.

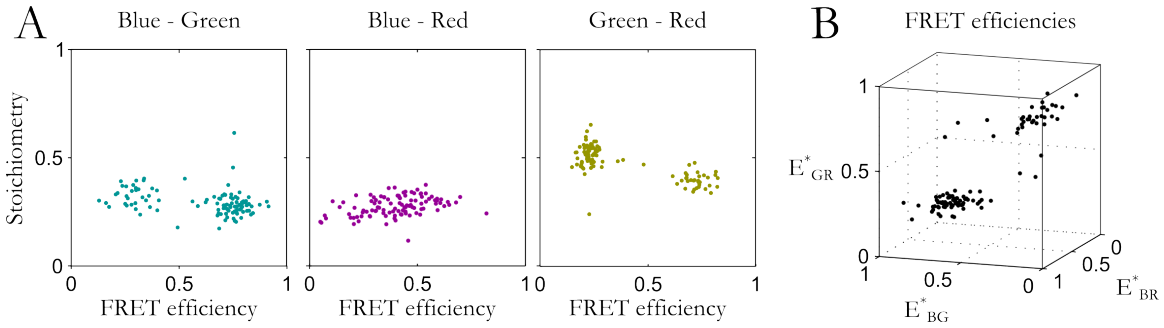


Figure 3.8: (A) ES scatter plot for each pair of fluorophores from the molecule in figure 3.7, showing the two distinct conformers detected by blue-green and green-red FRET. (B) Three dimensional scatter plot of the three FRET efficiencies showing the anti-correlated changes in FRET due to the co-ordinated motion of the Holliday junction arms.

3.4 Discussion

3.4.1 Construction of a three-colour microscope

In this chapter I have described the construction of a three-colour ALEX microscope capable of single-molecule FRET measurements. This microscope has seen extensive use for two-colour FRET measurements, however the complexity and difficulty of three-colour FRET assays has, as yet, limited the number of three-colour FRET measurements performed. These challenges are not unique to our instrument, but generally affect the utility of three-colour FRET in the field, see section 3.4.3. The flexibility of having three excitation and detection wavelengths, has also proved useful in performing super-resolution measurements in the group with a variety of fluorescent proteins, and to expand these capabilites, Stephan Uphoff and I have overseen the subsequent addition of several additional features to the microscope. We have added a 405 nm excitation laser, for photo-activation of fluorescent proteins, added the ability to interchange between the standard 543 nm laser, and a higher power 561 nm laser for increased temporal resolution, and along with Federico Garza de León (Kapanidis group) have added a Khöler illumination module, for bright field imaging.

3.4.2 Holliday junction dynamics

To validate the setup I demonstrated three-colour FRET measurements with the classic proof-of-principle molecule, DNA Holliday junctions (Hohng et al., 2004; Uphoff et al., 2010; Lee et al., 2010a). My results show the co-ordinated changes in two inter-fluorophore distances, between the blue to green, and green to red fluorophores. However, significant background in the F_B^R emission

channel, as well as the reduced spectral overlap between fluorophores makes reliable measurement of E^*_{BR} challenging in this experimental system. As previously observed I see transitions between the two stacked conformation with no visible open intermediate structure, and a bias to one stacked conformer (McKinney et al., 2002; Hohng et al., 2004)

3.4.3 Challenges of three-colour smFRET

Although clearly affording significant extra information, three-colour FRET measurements prove to be challenging to perform and interpret. One significant difficulty is quality of currently available short wavelength ('blue') organic dyes available, the best of which (Cy2, Alexa Fluor 488, ATTO 488) are still much less photostable and bright than comparable 'green' and 'red' fluorophores (Aitken et al., 2008; Lee et al., 2010b). This reduces the possible observation time at a given average photon emission count and, along with the reduced probability of observing fully labelled complexes (due to incomplete labelling), can make the capture of transient events more difficult. The presence of three fluorophores also increases the potential of errant photophysical fluctuations being exhibited by one of the dyes, thus rendering time traces difficult to interpret.

These difficulties are a general problem for three-colour FRET measurements and as organic dyes and the ability to control their behaviour improves, and auto-fluorescence of optical components is reduced, the ability to implement complex three-colour assays should subsequently improve. The specifics of the microscope introduced in this chapter also bring some difficulties. Primarily, the choice of spectral separation method, although reducing the number of components required, makes the alignment such that all three emission channel focusses coincide difficult, and due to the relatively off-axis path of the blue and red emission channel, introduces some astigmatism into the system. The placement of all three emission channels on one EMCCD chip, despite the reduction in magnification, reduces the number of simultaneously observable molecule in comparison to the previous two-colour system. Clearly both these factors could be eliminated by using a separate EMCCD camera for each emission channel, however the cost of such a set up proves prohibitive, and in practice the effect of the above prove to be relatively minor in routine single-molecule experiments.

3.4.4 Advantages of three-colour ALEX

The implementation of ALEX in this microscope overcomes some of the typical difficulties of three-colour FRET measurements, it allows virtual molecular sorting so as to eliminate the effect of partially

labelled molecules or complexes in analysis. This molecular sorting can also be used to separate out the required singly labelled molecules required to perform corrected FRET calculations without having to perform numerous separate control experiments (Lee et al., 2005). ALEX also helps to deconvolve the different possible FRET pathways, and doesn't require that all pairs exhibit significant FRET to be observed. In fact this last point highlights that ALEX can be used to simply detect the co-localistion of fluorophores in the absence of FRET, this suggests that even if three-colour FRET experiments may be challenging, three-colour ALEX allows for experiments to be envisioned wherein more reliable two-colour FRET between 'green' and 'red' fluorophores is monitored, while simultaneously observing the interaction, by co-localisation, of a third component in the reaction labelled with a 'blue' fluorophore.

3.5 Contributions

Microscope design and construction was performed in collaborations Stephan Uphoff, and Johannes Hohlbein provided much useful advice regarding design details. Stephan Uphoff also provided Holliday junction oligonucleotides.

Chapter 4

Tethered fluorophore motion

4.1 Introduction

In the previous chapter I outlined a three-colour implementation of single-molecule FRET. Adding a third colour to FRET experiments allows expansion of the number of distances simultaneously measurable, or can be used to detect the presence of an extra, fluorescently labelled, component in biological reactions. Despite the obvious advantages of three-colour FRET it comes with the significant experimental and instrumental challenges associated with such complex measurements; it is also limited to the fairly narrow dynamic range of distances addressable via FRET. In fact this is a general problem associated with single-molecule measurements, due to the inherent difficulties in making observations at single-molecule sensitivity the techniques used are highly specialised and typically tuned to a narrow range of length scales, or constraints. To this end experimental methods that combine multiple observables or techniques have recently begun to expand the scope of single-molecule experiments, to capture a wider range of information than available through a single technique alone, as outlined in chapter 2.5.3. Motivated by the desire to detect conformational changes of DNA outside the range of FRET measurements, I developed a method I term tethered fluorophore motion (TFM). In this chapter I introduce the principles and concept of TFM, an effect which occurs as a consequence of using long DNA substrates in a single-molecule imaging context, and demonstrate its ability to detect effective length changes of DNA molecules on the scale of hundreds to thousands of base pairs (bp) (30 - 300 nm). The following chapters (5 and 6) detail the combination TFM with FRET to study site-specific recombination in more detail than possible using either method individually.

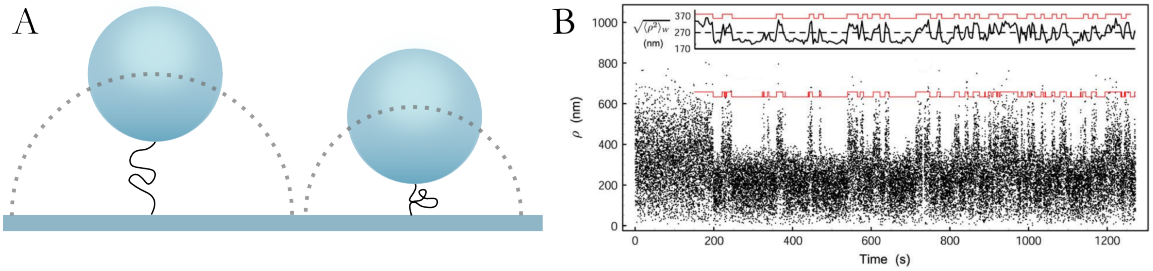


Figure 4.1: (A) Schematic illustration of TPM experiments. Beads tethered to a slide by a DNA are free to undergo constrained Brownian motion within the length of the tether, indicated by the dashed semi-circle. When the tether is shortened (e.g. by DNA looping) the range of bead motion is further constrained. (B) TPM data, the excursion of the bead centre from the tether location (ρ) is plotted for each frame, distinct looping events are evident where the nature of Brownian motion changes. For data analysis purposes the mean excursion during a sliding window is calculated (inset). Figure adapted from Beausang et al. (2007)

4.1.1 Tethered particle motion

The principle (and name) of TFM is based on the existing technique of tethered particle motion (TPM). TPM is a single-molecule method for observing large-scale conformational changes in polymers, typically DNA and its interaction with proteins. TPM experiments usually observe a long (> 400 bp) DNA molecule tethered to a surface, with a large (20 - 500 nm) bead attached to its surface-distal end (Schafer et al., 1991). Video microscopy of beads attached to microscope slides by single DNA tethers reveal the constrained Brownian motion which the bead undergoes, and properties of the underlying DNA molecule can be inferred from this motion, typically its apparent length (Nelson et al., 2006). TPM was first introduced in 1991 to monitor transcription (Schafer et al., 1991) and has since seen wide use to study a variety of protein-DNA interactions including: lac repressor induced DNA looping (Finzi and Gelles, 1995), restriction enzymes (Rusling et al., 2012), and site-specific recombinases (Fan, 2012), as well as studying fundamental aspects of polymer dynamics governing DNA (Brinkers et al., 2009; Manghi et al., 2010).

TPM contrasts to other techniques which use a tethered colloid, namely optical and magnetic tweezers, as those techniques inherently exert an external force on the DNA, controlled by the experimenter. This force allows tweezing techniques to similarly monitor the apparent length of the tether whilst also manipulating the molecule of interest. The force exerted (femto- to micro- newtons (Neuman and Nagy, 2008)) is a requisite of using these methods and although this manipulation ability often provides insight into the mechanical action of proteins, it is not always desirable. Stretching of DNA (implicit in exerting force to each end) can have significant effects on the activity of proteins

being studied (Bai et al., 2012), even abolishing the ability to carry out their biological roles (Forget and Kowalczykowski, 2012).

Although TPM does not explicitly exert a force on the attached colloid, the mere presence of the colloid can have a significant effect on the DNA tether. Two effects due to the presence of a large colloid have been shown to be important. The first is a volume-exclusion effect, wherein the constraints imposed upon the available DNA conformations, by steric clashes of the bead and surface, exert an effective stretching force on the DNA (10 - 100 fN for typical TPM experimental conditions (Segall et al., 2006)). The second is the effect that thermal motion of the bead has in dictating the rate at which local DNA conformations, and thus reaction intermediates, can be sampled, necessarily affecting reaction rates (Bai et al., 2012). The magnitude of such effects is dictated by the relative length of DNA and attached colloid, but both have been shown have an appreciable effect on the measured properties of biomolecular reactions (Segall et al., 2006; Milstein et al., 2011; Bai et al., 2012)

4.1.2 Tethered fluorophore motion

Although improvements in bead size and signal have enabled TPM measurements with a precision as high as 50 bp (17 nm) (Yin et al., 1994; Pouget et al., 2004), the nature of the method is still limited to the measurement of relatively large-scale changes in DNA length. This contrasts with the sensitivity and range of the single-molecule FRET measurements outlined in Chapters 2 and 3 (1 - 10 nm). Small intra-molecular movements within proteins often serve to mediate reactions involving conformational changes at much larger scales, particularly in the case of DNA processing enzymes. As such there is often the desire to monitor both the small-scale conformational changes occurring within nucleo-protein complexes, whilst simultaneously monitoring the overall conformation of the DNA to properly correlate structural changes in the protein with its DNA processing function. Despite addressing different and complementary length scales, the combination of TPM with FRET measurements is not trivial due to the strongly light-scattering nature of the bead used. To overcome this difficulty I developed a method to extend the principle of TPM, but using a single fluorophore as the reporter, rather than a large gold or polystyrene bead. This not only allows the simultaneous acquisition of single-molecule fluorescence measurements such as FRET or PIFE (Hwang et al., 2011), but it also avoids any perturbing effects associated with the attachment of a bead to the DNA tether.

In this chapter I demonstrate the principle of TFM, showing its ability to accurately measure

DNA lengths from 100 - 4000 base pairs. To assess the potential applications of TFM, I then address important questions regarding its spatial and temporal resolution and determine that under the typical conditions of single-molecule experiments, length changes down to 100 bp can be resolved, with time resolution comparable to that of TPM. To demonstrate the utility of TFM in performing biological measurements I observe polymerisation of DNA by Klenow Fragment, and measure the rate of polymerisation at the single-molecule level.

4.2 Experimental Methods

4.2.1 DNA and protein preparation

DNA substrates were provided by Pawel Zawadzki from the David Sherratt laboratory (Biochemistry Department, Oxford). Briefly, long DNA substrates were prepared via a PCR reaction using one fluorescently labelled, and one biotin labelled oligonucleotides as primers and a plasmid template using a Phusion High-Fidelity DNA polymerase (NEB). DNA sequences of the primers are shown in figure 4.2. Oligonucleotides were synthesized, labelled and HPLC purified by ATDBio ltd. After PCR reactions the product was gel-purified.

4.2.2 Instrumentation

Single-molecule TIRF experiments were performed on a custom built objective type TIRF microscope. A green (532nm Cobolt Samba) and red (635nm Cube Coherent) 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, numerical aperture 1.4, Olympus) 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. ALEX (Kapanidis et al., 2004) was implemented by directly modulating the lasers and all data was acquired using a 100Hz alternation rate, with excitation powers of 1mW for each laser. Fluorescence emission was collected by the objective and separated from the excitation light by a dichroic (545 nm/650 nm, Semrock) and cleanup filters (545 nm LP, Chroma; and 633/25 nm notch filter, Semrock). Emission signal was focussed on a rectangular slit to crop the image and then spectrally separated, using a dichroic (630 nm DRLP; Omega), into two emission channels which were focused side by side onto an EMCCD camera (Andor iXon 897). The EMCDD was set to an EM gain of 300, corresponding



Figure 4.2: A. DNA sequences and labelling positions used for initial PSF width investigation experiments. B. DNA primer-template construct used for KF polymerisation experiments. Sequence of DNA primers used in PCR are shown, location of Cy5 labelling is indicated by a star, long intervening sequences of DNA are not shown but indicated by dashed lines.

to an approximate real gain of 4.55 counts per photon. A CRIFF (ASI) autofocus system was used throughout this chapter to ensure focus stability over the course of data acquisition.

4.2.3 Single-molecule experiments

Biotinylated DNA was immobilised to the surface of a PEG passivated coverslip using biotin-neutravidin interactions, and sealed using a silicone gasket (Grace Bio-labs) and a second coverslip to seal the sample. Imaging was performed in a buffer consisting of 50 mM Tris-HCl (pH 7.5), 50 mM NaCl, 5 mM $MgCl_2$, 100 μgml^{-1} BSA, and 1 mM UV treated Trolox (Cordes et al., 2009). An enzymatic oxygen scavenging system consisting of 1 $mgml^{-1}$ glucose oxidase, 40 μgml^{-1} catalase and 1.4% (w/v) glucose was added just prior to sealing the sample before image acquisition.

4.2.4 Data analysis

Extraction of fluorescence intensity signals from microscope images was performed by the previously described twoTone software (Holden et al., 2010). An apparent FRET efficiency, E^*, was calculated

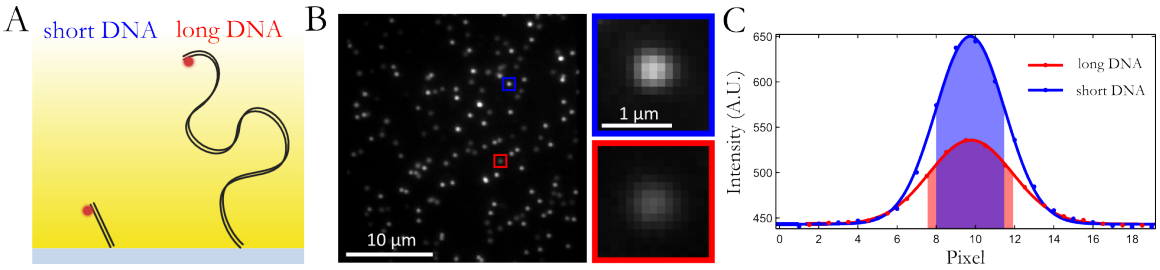


Figure 4.3: (A) Schematic representation of the principle of TFM, fluorophores attached to short DNAs are constrained to be close the tether point, whereas those on DNAs much longer than the persistence length spend significant amounts of time away from the tether point and slide surface. (B) Image from an EMCCD camera showing a mixture of long and short DNAs, an example of the PSFs arising from a short and long DNA are highlighted in blue and red respectively. (C) Single Gaussian fits to the pixel intensity values demonstrate the difference in photon count (area under curve) and PSF width (highlighted region).

from the extracted fluorescence emission as outlined in chapter 2. The acceptor PSF width was obtained from the mean width of the fitted elliptical Gaussian.

4.2.5 Klenow fragment experiments

Single-molecule experiments were performed as outlined above, but in a buffer containing 50 mM Tris pH 7.5, 10 mM $MgCl_2$, 1 mM DTT, 100 µgml^{-1} BSA, 1 mM Trolox, and the oxygen scavenging system described above. KF and nucleotides were added during data acquisition. Data was collected at 1 mW and 500 ms frames until no further polymerisation could be observed or the majority of fluorophores had photobleached. To extract single-molecule polymerisation rates I applied a 50 frame moving mean filter to the PSF width values (figure 4.10) and then divided the resulting trace into segments 25 frames long, to which a linear fit was performed, the gradient of which was used to obtain the polymerisation rate over the 12.5 second segment. The two window sizes were chosen to be as short as possible whilst still reducing obvious fluctuations due to noise. Any segment that contained intensities or PSF widths lying outside the normal range expected for a single fluorophore were excluded from the analysis.

4.3 Results

4.3.1 Concept of TFM

As explained in chapter 2, imaging a single immobile emitter (such as an organic fluorophore) onto a camera produces a point spread function (PSF). This PSF is much larger than the emitting molecule and arises due to the diffraction of light through the optical system depending on physical parameters of the fluorophore and the imaging equipment. Typical single-molecule imaging experiments observe fluorophores that are attached to proteins, nucleotides or antibodies, which themselves are generally immobilised to the slide surface using biotin-neutravidin (or similar) interactions. Often these conditions well approximate immobile emitters and thus the observed width of a PSF in such experiments generally corresponds to the diffraction limit . However in situations where the fluorophore cannot be considered to be immobilized (on scales > 10 nm), significant deviations can be observed in the width of PSF, dependant on the underlying motion.

DNA strands typically used in single-molecule experiments (< 100 bp) do not have significant end-to-end distances on the length scales relevant to optical microscopes. Thus, in the case of short, fluorescently labelled DNAs, immobilised to a slide surface, diffraction limited PSFs are typically observed. However, when the DNA is significantly longer than the persistence length[1] (~ 50 nm), diffusion of the free end, and thus the attached fluorophore, about the tether point has a significant effect on the observed PSF[2], figure 4.3. Dependant on the relative timescales of DNA motion and camera integration, changes to the PSF are manifest in one of two ways. If the DNA does not move appreciably during acquisition of a single camera frame, then individual snapshots of the fluorophore position will be observed, analogous to TPM measurements. However if the integration time is much longer than the relaxation time of the DNA (Nelson et al., 2006), an average of the fluorophore position will be observed and its diffusive motion will be apparent as a broadening of the PSF beyond the diffraction limit. Here I focus almost exclusively on the case where integration times are longer than the DNA relaxation time, and the justification for this is explained in section 4.3.3.

When using a total internal reflection (TIRF) microscope, typical for single-molecule measurements, in addition to the broadening in PSF, the mean position of the fluorophore above the coverslip surface

[1] Formally the persistence length is defined as the characteristic length scale over which correlations in the tangential direction between to points on a polymer decays. More informally it is used as a yardstick for the behaviour of a polymer. At contour lengths shorter than the persistence length, the polymer can be regarded to behave as an elastic rod, at lengths longer than the persistence length it increasingly behaves like a highly flexible string.

[2] As mentioned in chapter 2 this is no longer strictly a PSF, but will be referred to as such throughout this thesis for convenience.

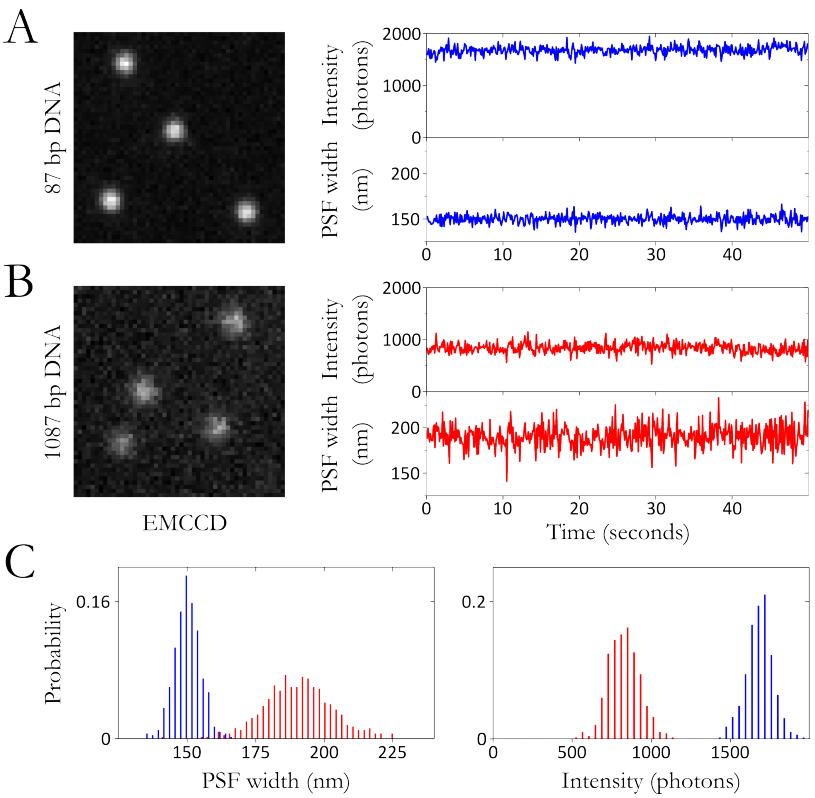


Figure 4.4: Example camera images with extracted intensity and PSF width signals from single molecules of (A) 87 bp and (B) 1087 bp DNA substrates. (C) Histograms of the single molecule TFM observables (PSF width and intensity) from the DNAs in (A) and (B) demonstrating the differences in mean value of each.

exposes it to a reduced average excitation intensity, due to the exponential fall off of the evanescent wave generated in TIRF (see section 2.4.2). This gives rise to an attenuated emission intensity as compared to a fluorophore tightly tethered to the surface, visible in figure 4.3.

4.3.2 Comparing two DNA lengths

To initially test the method and quantify the changes in PSF width and acceptor intensity due to differences in the length of labelled DNA, I separately acquired movies of short (87 bp) and long (1087 bp) double stranded DNA molecules. I observed many molecules in each field of view for 100 second movies, acquired at a frame rate of 100ms. Images from the camera clearly show differences in both the intensity and PSF width of the fluorophores, and extracting time-traces from individual DNAs using twoTone clearly demonstrates the stable single-molecule signals, figure 4.4A,B, with mean PSF values of ~ 150 nm and ~ 190 nm for short and long DNAs respectively, indicating a broadening of ~ 40 nm due to the diffusive motion of the long DNA, figure 4.4C. PSF widths are well defined and reproducible over the course of repeated experiments, and using the PEG-passivated slides typical in single-molecule experiments any transient association of the DNA end to the surface is seen in fewer than 1% of time traces.

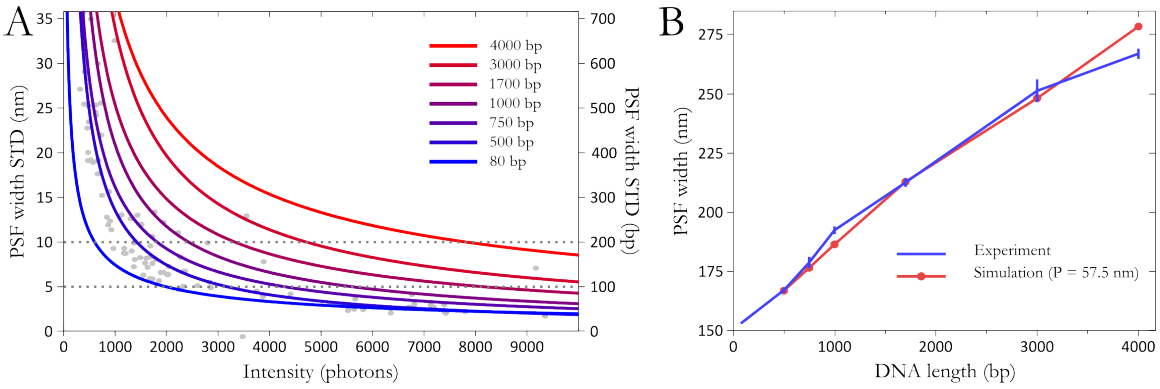


Figure 4.5: (A) Fits to the measured standard deviation of the PSF width from a single molecule against photon count plotted for each of the DNA lengths used in this chapter. Data points from the 500 bp DNA are shown in grey. Dashed lines indicate the required level of precision (and thus photon count) to resolve changes of 100 and 200 bp within 1000 bp molecules (B) Experimental measurements of the PSF width for DNAs 87 - 4000 bp, acquired with 1 second integration times, in blue, error bars indicate the standard deviation of three independent measurements for each sample. Simulated PSF width (see below) from 500 - 3000 bp (P = 575 nm) in red. Deviations between results and simulation at longer DNA lengths indicate that very long DNAs (> 3000 bp) do not adequately explore all accessible conformations during the integration time to fully average over the lateral position distribution.

There are also significant differences in intensities between DNAs, figure 4.4C, however this intensity difference is sensitive to the underlying single-molecule intensity of the fluorophore, the excitation pattern across the field of view and the precise incident angle of the excitation light. All these factors are subject to variation and thus absolute intensity differences between molecules can arise separate from DNA length changes. Although observations over time of single-molecules will not be subject to these confounding factors, the magnitude of observed intensity changes will still heavily depend on the depth of the evanescent illumination field (see section 2.4.2). In the rest of this chapter I focus on the PSF width which is not sensitive to such variable instrumental and alignment factors [3].

4.3.3 Resolution

Spatial resolution

The key to defining the spatial resolution of TFM, relevant to the nature of single-molecule experiments, is the precision to which the PSF width can be determined in a single frame from the EMCCD. This

[3] By appropriate adjustment of the angle of incident laser light, it is possible to adjust the penetration depth of the evanescent field so that the relative change in intensity is reduced, so as not to lose significant signal from longer DNA substrates.

precision naturally depends on the number of photons available, and I take the standard deviation (STD) of PSF measurements from a single-molecule time trace as an estimate of this precision. Determining the PSF width STD for many single molecule, for DNA lengths over the range of 80-4000 bp, and at varying levels of excitation intensity yields of plot of PSF width STD against mean photon count, for different lengths of DNA, figure 4.5A shows the power law fit to these data. Standard deviations as low as 5 nm (this corresponds to an effective length change of $\sim$ 100 bp for a 1000 bp DNA tether) can be obtained for photon counts of several thousand photons per frame, readily achieved under typical single-molecule experimental conditions (Roy et al., 2008).

For other types of single-molecule assays, those in which the desire is to compare substrate changes under varying conditions, rather than monitor the changes over time within a single DNA molecule, the key measure of resolution is instead the precision with which PSF width can be measured between experiments. To determine this sensitivity and obtain a calibration curve between the PSF width and DNA tether length, I measure the mean value of PSF for many single molecules of lengths between 80 - 4000 bp with a surface distal fluorophore, and use the STD of three independent repeats as an estimate of the precision. All lengths tested from 80 to 4000 bp are clearly distinguishable by PSF width, with STD consistently less than 10 nm, figure 4.5B. The response of PSF width to DNA length is approximately linear over the range of DNA lengths tested, with deviation from this trend occurring at lengths of 4000 bp where the approximation that the relaxation time of DNA is much less than the camera integration time no longer holds [4].

Temporal resolution

The time resolution for the reliable detection of a process such as loop formation (see Chapters 5 and 6) is determined by the magnitude of the effective length change upon loop formation and the number of photons emitted by the fluorophore in a single frame. However, time resolution in these experiments is fundamentally limited by the relaxation time of the DNA molecule under investigation. I have been working under the assumption that the integration time is much longer than the relaxation time of the DNA, this results in a broadened PSF centred above the tether point, see figure 4.6A, as the integration time is reduced, movement of the PSF centroid can be distinguished between subsequent frames, figure 4.6B, as the DNA can only partially explore its available conformations, effectively beginning to 'freeze' its motion. The effect of reduced integration times for a 1 kb substrate can be

[4] Although the 4kb substrate breaks this approximation it should be noted that its PSF width is still clearly resolvable from that of the 3kb DNA, although the magnitude of the difference is reduced.

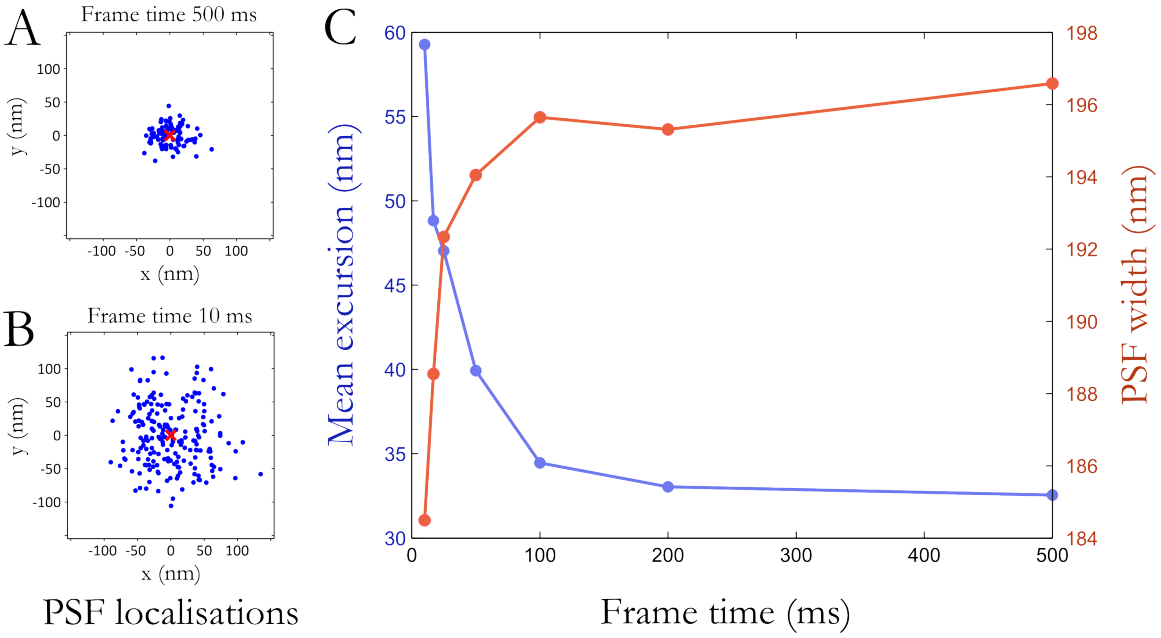


Figure 4.6: Effect of reduced integration time on the observed PSF width. Single-molecule localisations from sequential frames of a 50 frame movie of 1000 bp DNA tether, for integration times of (A) 500 ms and (B) 10 ms. Localisations (i.e. the PSF centroid) are shown in blue, and the mean of all localisations indicated by a red cross. (B) PSF width and mean excursion for 1000 bp DNAs at frame times from 10 - 500 ms, showing the clear decrease in PSF width and increase in mean excursion as the frame time becomes comparable to the DNA relaxation time. The value of the mean excursion for long integration times is significantly above that predicted by Mortensen et al. for the localisation precision (Mortensen et al., 2010) at the photon counts used, however it should be noted that their equation assumes freely rotating immobilised dipole sources, which is clearly not the case for TFM measurements.

seen clearly in figure 4.6C, as the frame time is reduced, the PSF width decreases, and the quantity typically calculated during TPM measurements, the mean excursion, increases. Despite the more rapidly acquired frames this does not afford any better temporal resolution as calculation of the mean excursion requires averaging over several frames and measuring the PSF width at longer integration times effectively performs this function 'in camera'. As the integration time used is analogous to the averaging time necessary for computing the mean excursion in TPM experiments, similar arguments can be applied to the possible time resolution as in Manghi et al. (2010). Thus temporal resolutions of 20 ms (Manghi et al., 2010) could readily be attainable provided that laser powers are adjusted to provide sufficient photons such that shot noise is not limiting.

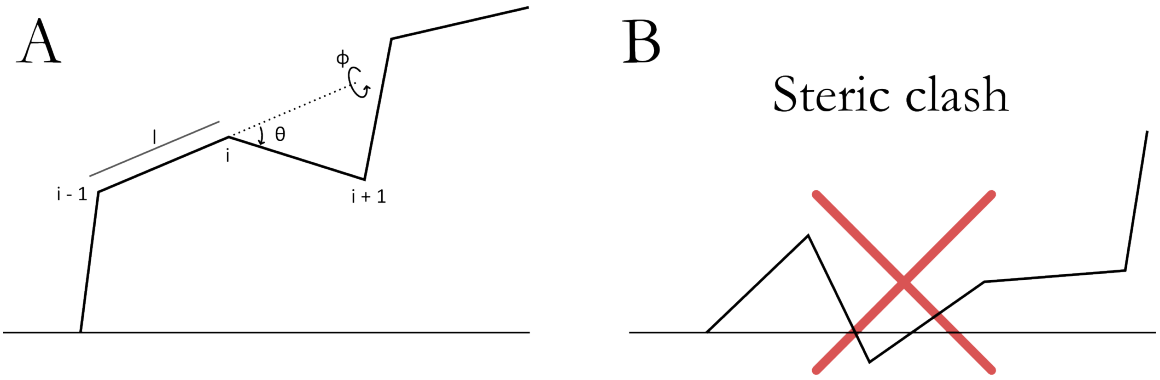


Figure 4.7: (A) Schematic of the discrete worm-like chain model employed. Chains are built segment by segment starting from the tether point (label), after the initial, randomly upward pointing segment, further segments of length l are extended along the vector of the previous segment rotated away from the axis parallel to the previous vector by θ and around this axis by ϕ. (B) any chains which intersect the $z = 0$ plane are discarded.

4.3.4 Simulating TFM observables

To ensure that the observed PSFs made sense from a physical model of DNA and to predict TFM observables for future experiments, I performed discrete Worm-Like Chain (WLC) simulations of the experimental system (similar to previous simulations of TPM (Nelson et al., 2006; Segall et al., 2006)) to obtain the static distribution of DNA conformations and relate these to the observed PSF width.

At length scales much shorter than the persistence length of DNA, double-stranded DNA molecules can readily be approximated as stiff rods (Doi and Edwards, 1988), however as the length of the DNA approaches and exceeds the persistence length, the inherent flexibility must be taken into account to capture the possible conformations adopted by the DNA molecule. A classic model for DNA at these length scales is the worm-like chain (also known as the Kratky-Porod model). In its simplest form the worm-like chain assumes a continuously flexible isotropic polymer chain with a quadratic bending energy (Doi and Edwards, 1988), and is known to provide a good approximation to many biologically relevant polymers including double stranded nucleic acids and unstructured polypeptides. The continuous worm like chain model can be further simplified to a discrete model, composed of rigid segments, each flexibly attached to those preceding and following it, with the same bending energy constraints as above. This provides a simulation framework that is computationally simple to implement, and although a considerable simplification, has been shown to adequately capture the behaviour of DNA in the current situation (Nelson et al., 2006; Segall et al., 2006).

WLC simulations were performed according to the scheme originally implemented by Hagerman

and Zimm (1981) and applied in the specific case of TPM by others (Nelson et al., 2006; Segall et al., 2006). Briefly, DNA molecules of length L are modelled as being composed of discrete rigid segments each of length $l = 3.2$ nm. Under the assumption that twisting energy and electrostatic and steric interactions between sections of the DNA can be ignored, then only the energy cost associated with bending need be considered to capture the flexibility of the DNA chain. To simulate a single conformation of a DNA chain, I sequentially assigned orientations to each segment in the chain relative to the previous one defined by the axial and azimuthal bend angles θ and ϕ. Assuming no preferred bending axis an isotropic bend model can be used, picking random values for ϕ from a uniform distribution between 0 and 2π, and picking values for θ from the distribution

$$Prob(\theta) = C \sin(\theta) \exp^{\frac{-P\theta^2}{2l}} \tag{4.1}$$

Where C is a normalisation constant and P is the persistence length (Hagerman and Zimm, 1981). To validate this model I generated many chain conformations (for a free DNA molecule) and calculated the mean end-to-end displacement, observing good agreement between simulation and the known analytical expression for the WLC. For simulations of DNAs tethered to a surface (the plane $z = 0$), the first linkage at the origin, of the DNA to the surface, is considered to be completely free (thus a random vector in the space $z > 0$). For each new segment of the chain I project the new segment in the direction of the previous segment, then apply rotation matrices corresponding to rotation by θ and ϕ. I then check the final position of the end point of the segment and if this lies below the plane $z = 0$ the entire chain is discarded, see figure 4.7.

To simulate the observed PSF width many ($> 10{,}000$) independent DNA chain simulations were generated to adequately sample from the space of accessible DNA conformations, figure 4.8A. From these simulated conformations, probability distributions of the fluorophore location in 3D space were constructed, figure 4.8B. The expected PSF width can then be computed as a numerical convolution between a weighted x-y spatial distribution (to take into account the varying intensity of illumination along the z-axis) and a Gaussian approximation to the diffraction limited PSF.

To fit the single free parameter in the simulations, the persistence length P, to the experimentally determined PSF widths in figure 4.5B, I systematically varied the persistence length over an experimentally measured range of P (Lu et al., 2002), and computed the mean square difference between the simulated result and the experimental values for the DNAs from 500-3000 bp[5]. The

[5]Shorter DNAs are not included as the discrete WLC model breaks down at very short DNA lengths i.e. when

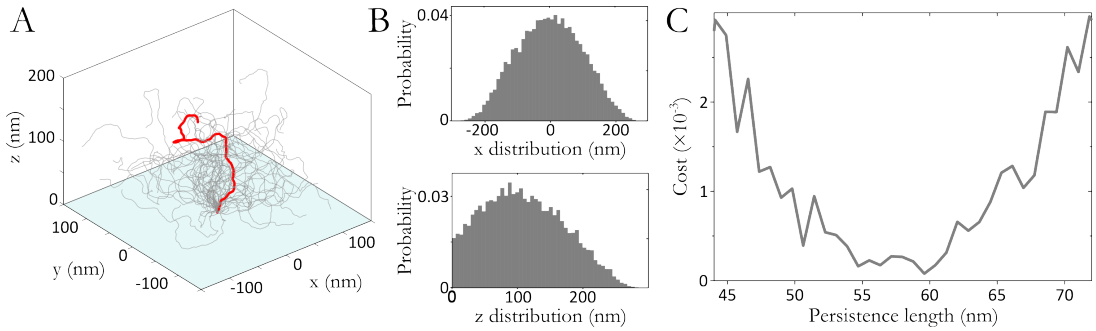


Figure 4.8: (A) Simulated tethered DNA configurations of length 1000 bp in grey, a single conformation is highlighted in red. (B) Lateral and axial distributions of fluorophore position from a simulated population of 1000 bp DNAs. (C) Fitting the persistence length input of the DNA simulation to experimental data, the optimal fit (see section 4.3.4) occurs at a value of $P \sim 57.5$ nm, the result of simulations using $P = 57.5$ nm is shown in figure 4.5.

optimal fit for persistence length occurs at $P = 57.5$ nm, figure 4.5B and 4.8C. The persistence length of double stranded DNA is not well defined for all buffer conditions and temperature, and varies with sequence, the fitted value of $P = 57.5$ nm broadly agrees with typically measured persistence lengths for double stranded DNA (Lu et al., 2002), but is somewhat larger than values expected in the given salt concentrations. These simulations do not consider electrostatic interactions or steric clashes of the DNA with itself (excluded volume effects), which could be responsible for the effective increase in predicted P. Assuming the true DNA persistence length is approximately 50 nm, these results indicate that the PSF observed is slightly broader than would be predicted solely due to the freedom of the fluorophore to diffuse in the x-y plane. One factor not considered in this simulation is the effect of defocussing as the fluorophore's mean position becomes further from the image plane (approximately coincident with the slide surface). Even with the high numerical aperture objective used in this study (NA = 1.4) this effect is not expected to be large due to the axially elongated nature of the PSF (Huang et al., 2008), but could account for the 'excess' PSF width observed in experiments. However the aim of these simulations is not to accurately model the fine details of DNA behaviour or imaging, but to give confidence in the predictable nature of PSF width variation with effective DNA tether length, and to allow the prediction of how this observable will vary using novel substrates, without having to perform laborious and costly controls (see below).

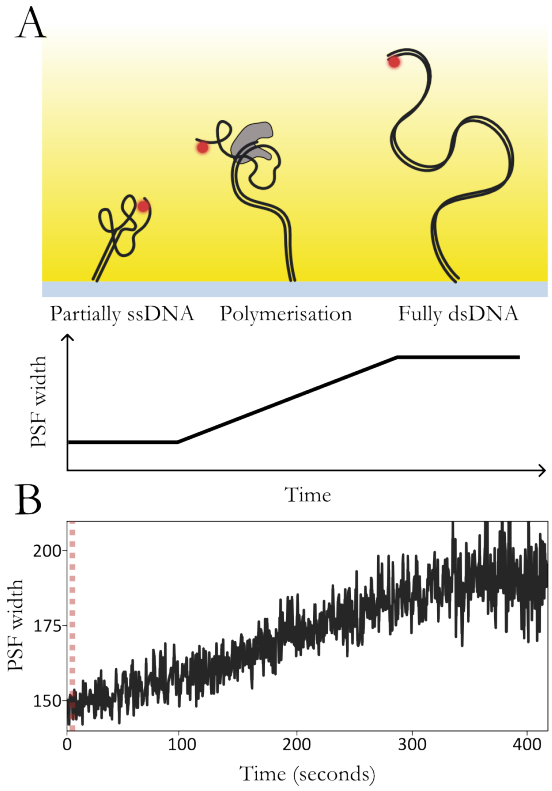


Figure 4.9: (A) Schematic representation of the KF polymerisation assay, initially the predominantly single stranded DNA substrate produces an observed PSF width very close to that of the diffraction limit, as the DNA is converted to double stranded during polymerisation the mean persistence length increases, broadening the PSF width until polymerisation is complete. (B) Representative time-trace of a single DNA undergoing complete polymerisation showing the increase in PSF width. The time of KF addition is indicated by the dashed red line.

4.3.5 DNA polymerisation by Klenow fragment

To demonstrate the utility of TFM in observing a biological process I applied it to the study of DNA polymerization by Klenow fragment (KF), a fragment of DNA polymerase I, lacking the 5′-3′ exonuclease domain, but retaining polymerase activity (Klenow and Henningsen, 1970). TFM was used to monitor the change in persistence length of a single DNA molecule, as it was converted from predominantly single to fully double stranded during polymerisation. The substrate DNA consisted of a short (55 bp) double stranded section (primer) which was tethered to the surface, followed by a long (1032 bases) single stranded region (template), figure 4.9.

To estimate the PSF width response to DNA polymerisation, I simulated DNA molecules at various stages of polymerisation, (i.e. 1087 bases long DNAs, with varying fractions of double stranded DNA), and using the optimal fitted persistence length for dsDNA ($P = 57.5$ nm) and a literature value of 1-3 nm for single stranded DNA (Tinland et al., 1997), and estimated the observed PSF width. These simulations indicated that during polymerisation the PSF observable would respond approximately linearly to the number of bases polymerised and that due to the very short persistence length of single stranded DNA the template region would not significantly contribute

$L \sim P$. Longer DNAs are not included as, for the above experimental data, the 4 kb substrate begins to break the assumption of well sampling accessible conformations as outlined in section 4.3.3

to the broadening of the PSF. Thus the width of the PSF effectively reflects the length of double stranded DNA in the molecule, figure 4.9,

Control experiments showed that the initial substrate had a PSF width ~ 160 nm, and the fully polymerised product had a width of ~ 190 nm[6]. Thus to convert from the rate of PSF change to the rate of polymerisation I used the linear equation

$$\frac{d(bases)}{dt} = 2.9 \times 10^3 \frac{d(PSF)}{dt}. \tag{4.2}$$

To measure the rate of polymerisation by KF I performed experiments in collaboration with Geraint Evans (Kapanidis group), we added saturating concentrations of both KF and nucleotides to the tethered primer-template DNA substrate and observe a gradual increase in PSF width of $\sim 60\%$ of molecules over the course of 10 minutes. A CRIFF autofocus systems was used to ensure focal stability over the course of the long acquisitions, and the stable nature of PSF widths from molecules which did not participate in the reaction, for whatever reason, confirmed its efficacy.

Data extraction

A smoothing filter was applied to the PSF width data and a linear fit was performed to 25 frame segments of the time trace, see section 4.2 and figure 4.10. Rates from all segments of all molecules from a single field of view were then binned into a histogram, figure 4.11. This rate histogram will contain both segments with a polymerisation rate of zero (i.e. before the start, or after completion of polymerisation, or any intervening pausing (Maier et al., 2000)) and those that correspond to the action of KF. In order to subtract the zero peak I assume measurement errors follow a normal distribution and fit a single Gaussian, with a mean fixed to zero, to all negative values of polymerisation rate. This assumes that negative values predominantly arise due to the error on measurement of non-polymerising segments, and that measurement error is small enough that those undergoing polymerisation do not significantly contribute. I then subtract this fitted 'zero peak' from the complete histogram to obtain an estimate for the distribution of non-zero polymerisation rates, figure 4.11.

We performed polymerisation experiments under a saturating nucleotide concentration (10μM) as well as a lower concentration (0.125 μM), and in their absence. For the saturating concentration

[6]This value is slightly smaller than that observed in the previous section using the same substrate, this change is due to the different buffer conditions used in the KF polymerisation experiments, see section 4.2

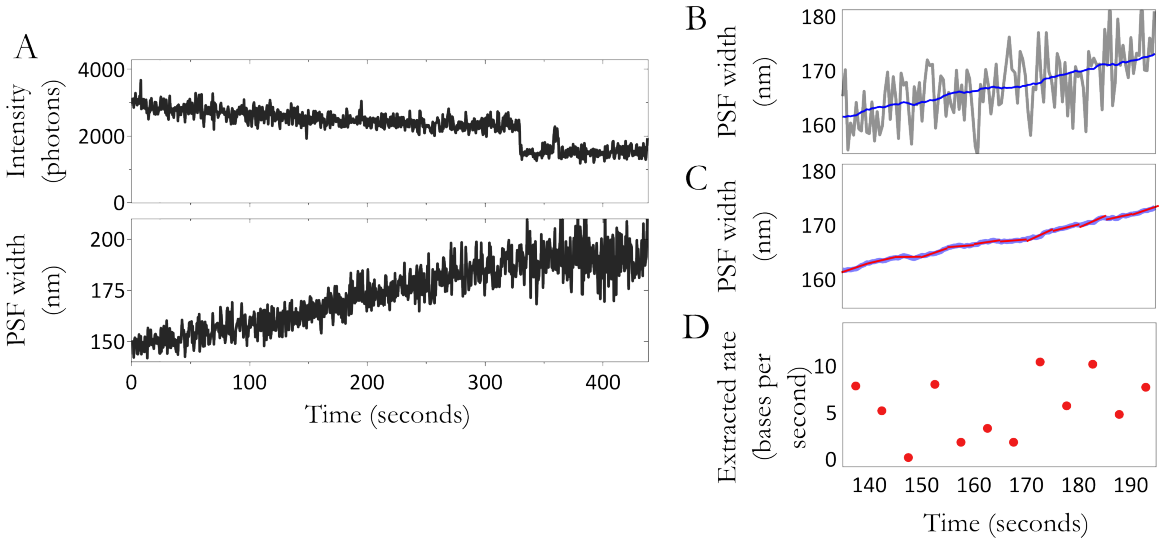


Figure 4.10: (A) Example time trace from a single DNA substrate undergoing polymerisation by KF, the intensity gradually decreases and PSF width increases as polymerisation proceeds. At time ~ 340 seconds, the PSF width reaches a value corresponding to complete polymerisation and does not increase further. This is accompanied by an abrupt decrease in intensity which I attribute to the change in fluorophore environment as KF reaches its proximity and the DNA in its locality is converted to double-stranded. Detail of the rate extraction procedure, (B) the PSF width (grey) is smoothed using a moving average filter (blue), (C) the result of which is linearly fitted in segments (red lines), (D) the gradient of fitted lines is then converted to a rate in bases per second, see section 4.2

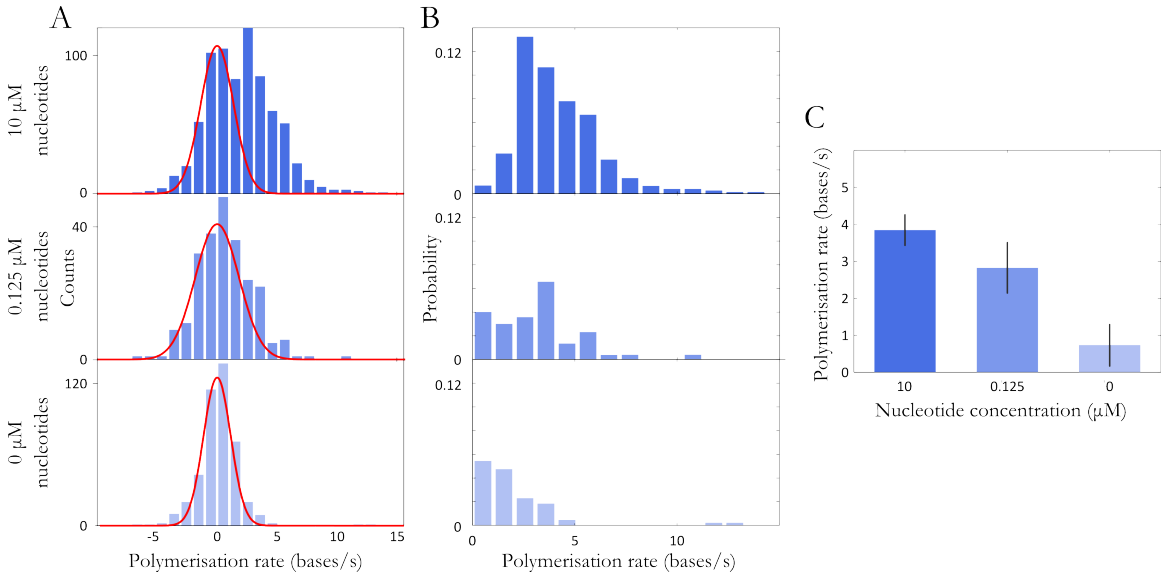


Figure 4.11: Results of rate extraction for varying concentrations of nucleotides. (A) Histograms of all extracted rate values from a single field of view. A single Gaussian, fixed at mean = 0, is fitted to negative values in order to estimate the size of the non-polymerising population (red line), this is then subtracted from the histogram to give the polymerisation rate histograms in (B) as explained in section 4.3.5. (C) Polymerisation rates compared at varying nucleotide concentrations, clearly showing the effect of KF on the DNA substrate when a high concentration of nucleotides is present, error bars indicate the standard error on the mean.

of nucleotides we measure a polymerisation rate of 3.8 ± 0.4 nucleotides per second, comparable to that previously measured by Maier et al. of 6 ± 1 nucleotides per second (Maier et al., 2000), and the two populations ('zero peak' and polymerising) are visible in the polymerisation rate histogram. However at low nucleotide concentration, the polymerising peak is not clearly visible, and the fitted 'zero peak' is significantly wider than in the high or no nucleotide case, indicating that due to the slow nature of polymerisation in this case we cannot accurately deconvolve the two populations, and making the obtained polymerisation rate unreliable.

4.4 Discussion

In this chapter I have introduced and characterised TFM, a novel single-molecule fluorescence method for measuring the diffusional characteristics of a single fluorophore tethered to a slide via a long DNA molecule, enabling inference about the underlying properties of the DNA. The use of a single fluorophore rather than a 20 - 500 nm diameter colloidal particle enables the technique to be used in

combination with a variety of single-molecule fluorescence methods available, such as FRET, PIFE, and co-localisation (Friedman and Gelles, 2012). The absence of a large colloid ensures that the motion of the DNA, rather than the bead, determines the dynamics of the system (Manghi et al., 2010; Bai et al., 2012) and also eliminates the entropic stretching force exerted on the DNA due to the excluded volume effect of the surface and bead (Segall et al., 2006). Sample preparation is more facile than that of traditional TPM due to requirement of only one non-covalent attachment chemistry (e.g. biotin-neutravidin). There is also no risk of multiply tethered particles which can affect TPM experiments and the technique is readily implementable on any existing single-molecule wide-field microscope. Despite these advantages TFM is limited in observation time to the lifetime of single fluorophores (up to several minutes), and signal to noise ratios as high as those possible from the strong scattering of light from colloidal gold (Lindner et al., 2011) are not generally attainable. However, under conditions typical for single-molecule experiments length resolutions down to 100 bp and temporal resolutions comparable to that of TPM (20 ms) should be readily achievable.

I demonstrated the use of TFM to measure the rate of DNA polymerisation by KF by monitoring gradual changes in the average persistence length of the molecule during the reaction. The rates obtained agree broadly with those from a similar single-molecule experiment performed by Maier et al. (2000). Differences in the rate measured here could well be due, not only to the variation of polymerisation rate and pause frequency with sequence (Schwartz and Quake, 2009), but probably more importantly, due the secondary structure of the large portion of single stranded DNA. Transient base-pairing or local hairpin formation is likely to slow the rate of KF polymerisation and this was suggested as a probable cause of the increase of polymerisation rate with applied stretching force observed by Maier et al. (2000). Further extensions to this work could be envisioned, although here the fluorophore is attached to the DNA, it would be entirely possible to use a labelled protein as the TFM reporter, enabling much longer observation times, as the turnover of protein binding to the DNA would continually replace bleached fluorophores. Although I have only used a single fluorophore in this case, the use of a second fluorophore attached to KF would allow direct measurement of the processivity of the enzyme, and a FRET scheme such as that employed by Santoso et al. (2010) could be used to observe conformational changes of the protein and directly correlate these to the progress of polymerisation.

In conclusion, the method of TFM which I have introdcued and characterised in this chapter can be used to combine the recent successes of TPM in studying processes involving DNA, such as looping, bending and synthesis, with the already established power of single-molecule FRET

measurements for elucidating mechanistic details of proteins involved in effecting such changes to DNA molecules. I hope that the method of TFM will expand the possibilities available to the single-molecule community and that advances in combination techniques such as TFM-FRET, combining local and global information at the level of single biological reactions, will provide new and useful insight into DNA-protein interactions. In subsequent chapters I go on to use TFM and FRET to address mechanistic questions regarding two site-specific recombinase systems in greater detail than previously possible, thanks to the use of the combined observables of TFM-FRET.

4.5 Contributions

Pawel Zawadzki prepared all DNA substrates (apart from the Cy3B labelling step), and KF polymerisation experiments were performed in conjunction with Geraint Evans.

Chapter 5

Cre-*loxP* recombination

5.1 Introduction

In the previous chapter I outlined the concepts of TFM, and demonstrated a proof of principle application to the observation of DNA polymerisation. The following two chapters use the combined TFM-FRET method to study mechanistic details of site-specific recombination reactions, and the combination of observable provides more information than could be obtained by either technique alone.

5.1.1 Site-specific recombination

Site-specific DNA recombination is the protein-mediated cleavage, exchange, and re-joining of DNA strands between two DNA duplexes containing a specific sequence. At its most basic site-specific recombination involves the binding of recombinase proteins to DNA sites, pairing of the two sites as a 'synaptic complex', exchange of DNA strands, and dissociation of the resulting synaptic complex to release the recombined product DNA. This process can lead to three distinct outcome depending on the relative orientation and location of the two sites: integration, excision or inversion, as illustrated in figure 5.1. Site-specific recombinases carry out a wide variety of roles within the cell including: integration and excision of viral genomes, resolution of chromosome or plasmid dimers, and regulation of gene expression (Grindley et al., 2006). Although all recombinases perform the same basic steps outlined above, two distinct families exist, named after the amino acid residue which performs cleavage of the scissile phosphate on the DNA backbone: tyrosine or serine, each with very different reaction mechanisms.

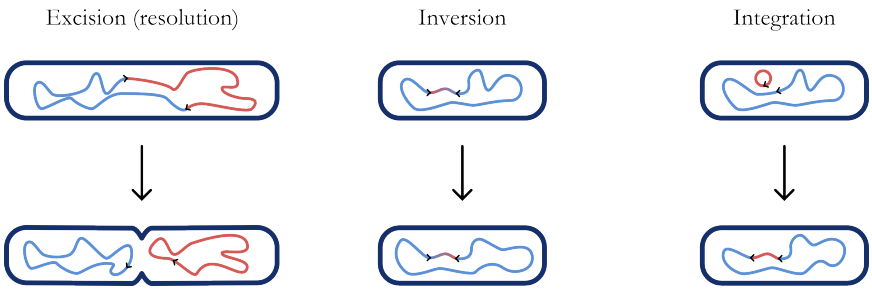


Figure 5.1: Schematic representation of the three possible outcomes of recombination. Recombination between two sites on a single molecule in an parallel orientation leads to excision. If these sites are in a anti-parallel arrangement the intervening DNA will be inverted. Recombination between two separate molecules leads to integration.

Tyrosine recombinases mediate sequential pairs of strand exchange between two DNA duplexes; the first pair of strand exchanges forms a Holliday junction (HJ) intermediate, whereas the second leads to resolution of the HJ to recombinant product (figure 5.2A). DNA cleavage requires no external energy factors, such as ATP, since the bond energy is stored as a covalent protein-DNA linkage during strand exchange.

Serine recombinases operate by a significantly different mechanism of strand-exchange where, again, four recombinase molecules bind to two DNA duplexes, here however, all four strands are cleaved simultaneously (Grindley et al., 2006). Strand-exchange between the two duplexes is then thought to occur by a unique 'molecular bearing' mechanism, wherein a large surface of hydrophobic protein-protein interactions, along the plane of cleavage, hold the complex together as strand-exchange is achieved by a 180° rotation of one half of the complex (Bai et al., 2011; Johnson and McLean, 2011). The successful rejoining of DNA ends is determined by correct base pairing of the small overhangs left by cleavage, and if these are not complementary between the exchanged duplex, the complex is free to continue to rotate to its original position (Bai et al., 2011).

In this thesis I will focus on the tyrosine recombinases, specifically two examples: Cre in this chapter and XerCD in chapter 6. Although I will not address the serine recombinases, the techniques presented here to study tyrosine recombinases, in principle, could equally well be applied to the study of the serine family.

5.1.2 Cre recombinase

One of the best studied tyrosine recombinases is Cre, a 38 kDa protein of the bacteriophage P1 of Escherichia coli. Cre catalyses cyclisation of the viral genome and resolution of plasmid dimers

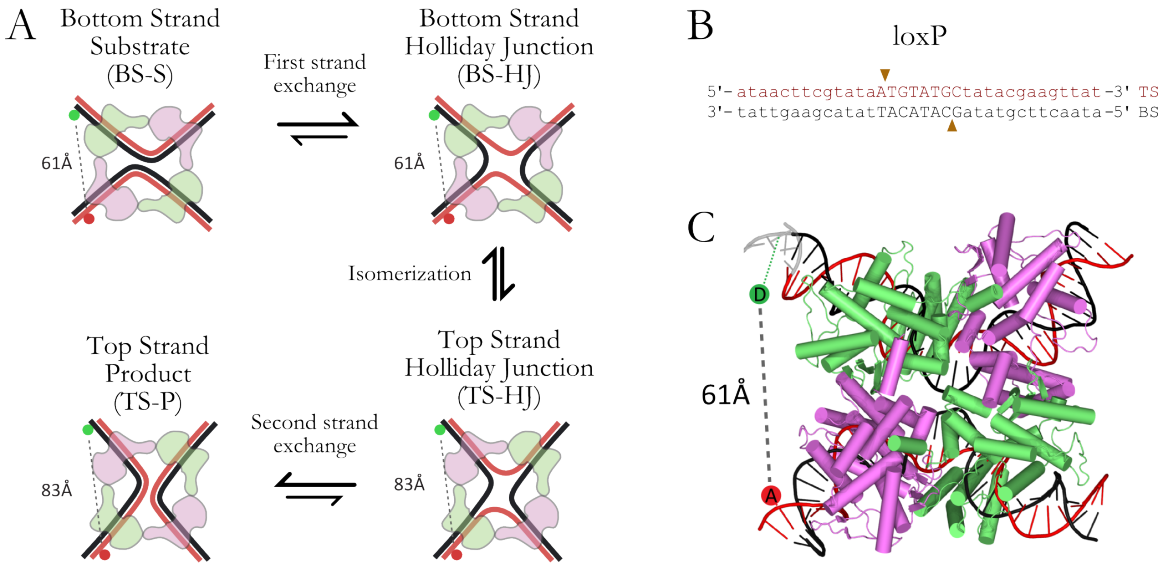


Figure 5.2: Cre-mediated site-specific recombination. (A) Schematic of site-specific recombination proceeding via the 'bottom-strand' synaptic complex (BS-S). After DNA cleavage by the active recombinases (green) within the BS-S complex, strand exchange forms a 'bottom-strand' HJ (BS-HJ). Isomerisation occurs between this and the 'top-strand' HJ (TS-HJ); after another round of cleavage and strand exchange a 'top-strand' product synaptic complex (TS-P) is formed. Positions of the donor and acceptor (green and red respectively) and inter-fluorophore distance measurements from crystal structure (PDB: 2HOI) are shown (see section 5.2.5). (B) Sequence of *loxP* site, Cre binding sites in lowercase, and asymmetric central region in uppercase, arrows indicate cleavage positions. Top and bottom strand nomenclature as in (Ghosh et al., 2005) (C) Crystal structure of Cre in BS synaptic complex with DNA (PDB:2HOI, (Ghosh et al., 2007)), DNA was extrapolated (grey) to the labelling site and mean fluorophore positions are shown as coloured circles.

during replication (Van Duyne, 2001). Cre mediates recombination between 34-bp DNA sequences named *loxP* (figure 5.2B). Unlike other recombinases (e.g., XerCD (Grainge et al., 2011) and λ Int (Mumm et al., 2006)), Cre does not require accessory proteins and has been well studied genetically, biochemically and structurally, consequently it serves as an excellent model system for understanding the mechanisms of tyrosine recombinases. Its simplicity and compatibility with eukaryotes has also made it a versatile tool for rearrangement and manipulation of genetic elements *in vivo* (Sauer, 1993).

Structural studies on Cre-*loxP* reaction intermediates have identified nucleoprotein complexes with remarkably similar overall structure (near-square planar) at each stage of recombination. This structure involves four Cre monomers, bound to two anti-parallel *loxP* sites (figure 5.2) connected through a network of protein-protein interactions; of these four monomers, only two are active at any time (green monomers, figure 5.2A). After the first pair of strand exchanges, only subtle structural changes are needed to activate the second pair of Cre monomers. It is widely believed that all steps of recombination occur within this nucleoprotein complex without major changes to its architecture (Van Duyne, 2001).

There is strong structural and biochemical evidence that Cre cleaves and exchanges DNA strands in an ordered fashion, with most studies finding recombination to be initiated with the 'bottom-strands' exchanged first (black DNA, figure 5.2A, B) (Hoess et al., 1987; Lee and Sadowski, 2003b; Ghosh et al., 2005). However, many of these studies relied on mutations to either Cre or the *loxP* site and there are conflicting reports on the order of strand exchange (Martin et al., 2002; Ennifar et al., 2003).

5.1.3 Single-molecule studies of site-specific recombination

Characterizing short-lived species using ensemble biochemical techniques, and relating them directly to the progress of reactions is challenging. To overcome the complications caused by ensemble-averaging (as outlined in Chapter 2) and observe site-specific recombination in real-time, various single-molecule techniques have been used to study recombination, including atomic force microscopy (Vetcher et al., 2006), tethered particle motion (TPM) (Mumm et al., 2006; Fan, 2012), and magnetic tweezers (Bai et al., 2011). In particular, the recent TPM work by Fan (2012) observed the overall progress of Cre mediated recombination reactions, measuring various association and dissociation rates, as well as noting the heterogeneous behaviour of complexes, but was unable to correlate these changes to nanometre-scale rearrangements within individual complexes.

To address such limitations, in this chapter I outline the combination of TFM with Förster resonance energy transfer (FRET) and simultaneously observe both large- and small-scale conformational changes of single DNA molecules, as they undergo recombination. Using TFM-FRET, Cre mediated recombination within individual nucleoprotein complexes was observed in real-time, and their reaction kinetics analysed. My results show that recombination is initiated predominantly by exchange of the 'bottom-strands' rather than 'top-strands', and that synaptic complex formation is reversible. In productive complexes I identify a rate-limiting step that occurs after HJ isomerisation and also observe two non-productive synaptic complexes, one of which is structurally distinct from the conformations seen in crystals. Finally, fast HJ dynamics in recombining complexes was inferred from the FRET data and single-molecule observations, complemented by fluorescence correlation spectroscopy (FCS) experiments, demonstrate the slow dissociation of the recombined product synaptic complex.

5.2 Experimental methods

5.2.1 DNA and protein preparation

Mutants of Cre and DNA substrates were provided by Pawel Zawadzki from the David Sherratt laboratory (Biochemistry Department, Oxford). Briefly, mutants were prepared with QuikChange Site-Directed Mutagenesis Kit (Stratagene) and tag-less proteins were purified according to Ghosh and Van Duyne (2002). After purification their identities were confirmed with mass spectroscopy, and activities checked by ensemble recombination assays performed by Pawel Zawadzki, figure 5.4. 1087 bp DNA substrates were prepared similarly to those in chapter 4, i.e. via a PCR reaction using two fluorescently labelled oligonucleotides as primers and a plasmid template containing directly repeated *loxP* sites separated by a 1 kb KmR gene (pRB10loxP). Forward and reversed oligonucleotides were labelled at with Cy5 and Cy3B, respectively. Oligonucleotides were synthesized and HPLC purified by ATDBio ltd. Cy5 labelling was performed by ATDBio, whereas I performed Cy3B labelling as previously described (Kapanidis, 2008).

5.2.2 Instrumentation

Single-molecule TIRF experiments were performed on the same microscope as described in chapter 4.

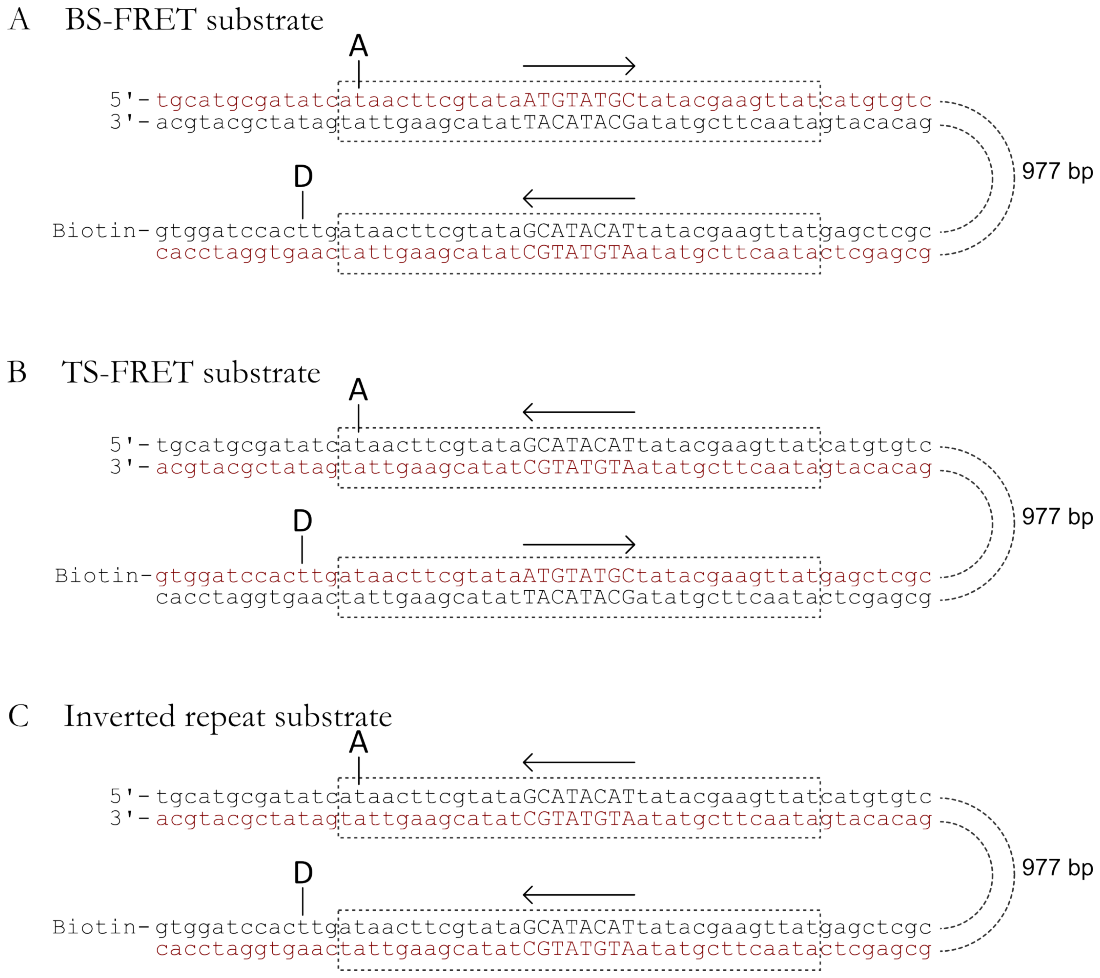


Figure 5.3: Sequences and labelling positions of DNA substrates used throughout this chapter. *loxP* sites are indicated by dashed boxes, and the orientation by the arrows above. Top and bottom strands are indicated in red and black respectively, the long sequence of DNA linking the sites is not show, and represented by dashed lines.

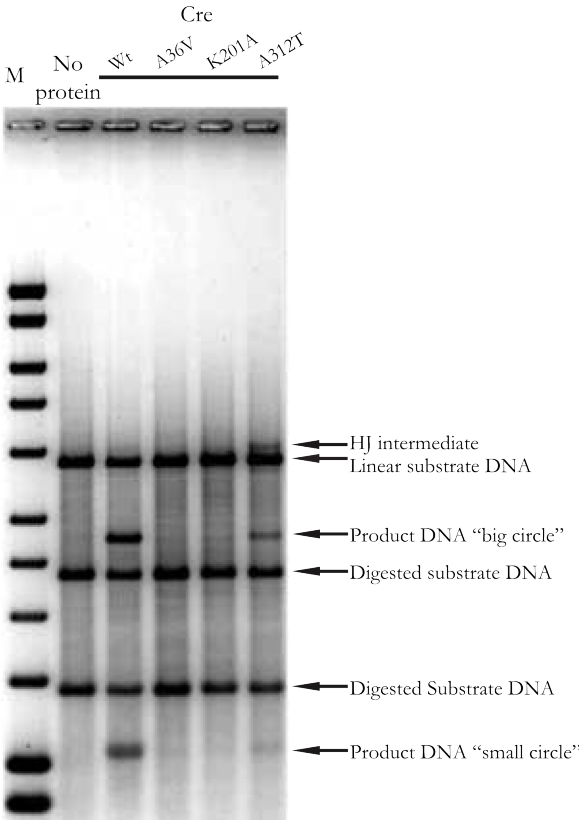


Figure 5.4: Gel electrophoresis of recombination reaction. A 3.7 kb plasmid carrying two directly repeated *loxP* sites was incubated with Cre for 10 min at 37°C. Products of recombination were expected to be two circles: big (2.7 kb) and small (1kb). These products were then (partially) digested to reveal linearised products in reaction with wt and A312T mutant but none when A36V or K201A was used. As expected A312T is less efficient in recombination and accumulates HJ intermediate.

5.2.3 Sample preparation

Biotinylated DNA was immobilised to the surface of a PEG-passivated coverslip using biotin-neutravidin interactions, and sealed using a silicone gasket (Grace Bio-labs) and a second coverslip as a lid. Imaging was performed in a buffer consisting of 50 mM Tris-HCl (pH 7.5), 50 mM NaCl, 5 mM $MgCl_2$, 100 μgml^{-1} BSA, and 1mM UV treated Trolox (Cordes et al., 2009). An enzymatic oxygen scavenging system consisting of 1 $mgml^{-1}$ glucose oxidase, 40 μgml^{-1} catalase and 1.4% (w/v) glucose was added just prior to sealing the sample before image acquisition. All experiments were performed at a room temperature of 21°C. In experiments where SDS was used to disrupt protein interactions it was added to the buffer at a concentration of 0.1%.

5.2.4 Data analysis

Extraction of fluorescence intensity signals from microscope images was performed by the previously described TwoTone software (Holden et al., 2010). An apparent FRET efficiency, E^*, was calculated from the extracted fluorescence emission, and the acceptor PSF width was obtained from the mean width of the fitted elliptical Gaussian, as outlined in Chapters 2 and 4.

Further analysis and quantification of extracted signals was performed with custom written MATLAB software. For analysis of recombination events intensity traces from individual molecules were manually inspected and classified. Transient FRET signals were defined on the basis of $E^* > 0.2$ along with the appropriate changes in PSF width and acceptor intensity. In the case of events which returned to substrate, short events (less than 2 frames) were discarded as these could not reliably be distinguished from fluctuations due to noise. However in the case of 'productive' events, unambiguous PSF and intensity signals allowed assignment of events which displayed short or no transient FRET. Any molecules whose fluorophores bleached within 10 seconds of a FRET event, or that exhibited confounding photophysical fluctuations were excluded from the analysis. Under the imaging conditions used, I measured the observation lifetimes (i.e the time until first the first long-lived off state) of Cy3B and Cy5 to be $\sim$ 400 s and 160 s respectively (fit to single exponential decay) and that $\sim$ 10% of molecules show errant photophysics from either fluorophore. Further, $\sim$ 5% of molecules were observed with two step photobleaching, or intensities corresponding to more than one fluorophore at a single localisation position, and these were discarded.

Histograms of E^* values were constructed by extracting the manually assigned FRET frames, discarding the first and last frames, and plotting the data points to a histogram, normalising each count by the length of the observed event.

5.2.5 Corrected FRET and distance calculation

To estimate the expected inter-fluorophore distance I extrapolate one DNA arm of the Cre-*loxP* synaptic complex structure (2HOI (Ghosh et al., 2007)) for the BS conformation, and 1NZB (Ennifar et al., 2003) for the TS conformation) to encompass the DNA labelling site. To account for the effect of linker length, I superimpose the sterically accessible region computed by Wozniak et al. (2008) and use the mean fluorophore positions to estimate their separation.

5.3 Results

5.3.1 TFM-FRET observations of Cre-*loxP* recombination

To observe recombination at the single-molecule level using TFM-FRET, 1087 bp long DNA substrates containing two *loxP* sites and labelled with a FRET donor and acceptor (figure 5.5) were observed

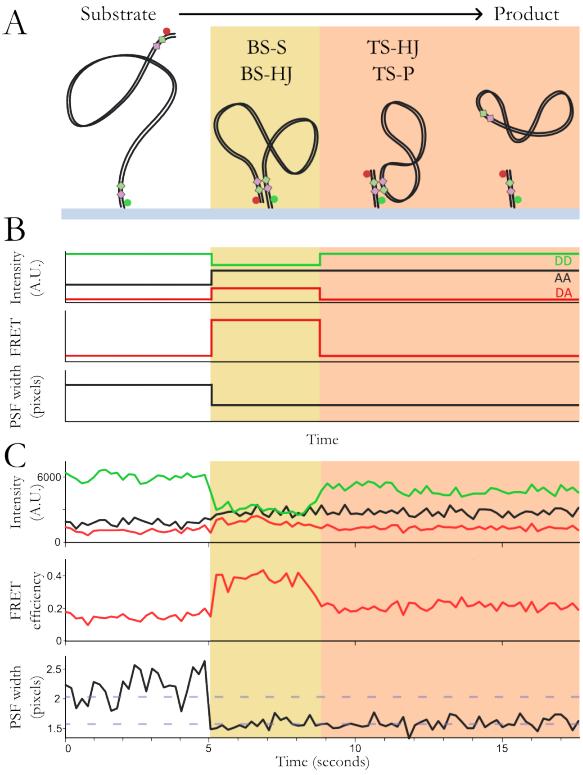


Figure 5.5: Single-molecule recombination observed in real-time. (A) Schematic of BS FRET substrate during 'bottom-strand' recombination. (B) Expected emission intensities of donor (DD; green) and acceptor under green excitation (DA; red), and acceptor under red excitation (AA; black). DD and DA are used to calculate apparent FRET efficiency (E^*) and the PSF width of the acceptor is obtained from AA images. The stages of reaction (figure 5.5A) giving rise to each of these states is indicated.(C) Representative time trace of a productive recombination event.

using TIRF microscopy. Since the fluorophores were attached at either end of the DNA substrate, the inter-fluorophore distance was well beyond the range of FRET.

Cre-mediated formation of a synaptic complex between the *loxP* sites dramatically reduced the inter-fluorophore distance (yellow panel, figure 5.5A). Donor and acceptor fluorophores were positioned such that formation of a Cre-*loxP* antiparallel synapsis (with the 'bottom-strands' in an active conformation for cleavage; BS-S, figure 5.2A) resulted in a significant FRET signal (apparent FRET efficiency, $E^* > 0.2$). Parallel Cre-*loxP* synaptic complexes, which are known to assemble, but not recombine (Van Duyne, 2001), were not expected to show significant FRET and were not addressed in this work. Due to structural similarities of the 'bottom-strand' synaptic complex (BS-S) and the 'bottom-strand' HJ (BS-HJ) (figure 5.2A, 5.5A), the FRET efficiency of these two states is expected to be identical. Any complexes following HJ isomerisation ('top-strand' HJ (TS-HJ), 'top-strand' product synaptic complex (TS-P), and the product DNA), were expected to show low FRET[1] ($E^* < 0.2$; pink panels, figure 5.5A). Consequently, the FRET efficiency could not distinguish directly between complexes that isomerise to a 'top strand' configuration (TS-HJ or TS-P), and

[1]The distance estimate from crystal structures of the TS-HJ and TS-P complex are both 83 Å, measuring the FRET value after the BS complex in 'productive' traces (i.e. the pink panels in figure 5.5A) results in $E^* = 0.18$, corresponding to a distance of 84 ± 11 Å. Although this is measurable over the ensemble of single molecules, I refer to this as a 'negligible' FRET efficiency due to its proximity to the FRET efficiency arising solely from spectral cross talk ($E^* = 0.13$), and consequently the difficulty in reliably differentiating this state in intensity traces from individual molecules.

those that dissociate back to the substrate DNA. To discriminate between the above possibilities FRET measurements are combined with TFM to simultaneously probe the effective length of the DNA. By using ALEX, FRET measurements are made by exciting the donor, followed by direct excitation of the acceptor to measure the TFM observable of PSF width.

To assess the efficiency of recombination on the surface, I incubated coverslips carrying immobilized 1087-bp long DNA substrate (BS FRET, figure 5.3) with Cre for 10 minutes. After washing the surface with SDS to disrupt protein-protein interactions, I identified DNA molecules which had undergone recombination; such molecules were expected to be short (61 bp), doubly labelled, and have a surface-proximal acceptor (i.e., a diffraction-limited PSF). 30% of the molecules showed a PSF width corresponding to a surface-proximal fluorophore, indicating that many molecules had undergone at least one pair of strand exchanges; this percentage is comparable to reactions performed in solution (figure 5.4, (Ringrose et al., 1998)).

To monitor individual recombination reactions in real-time, I acquired several 100-second movies during incubation of Cre with the surface-immobilized BS FRET substrate. Single-molecules showed transient FRET signals (dwells with $E* > 0.2$; see section 5.2.4) from 178 of 6548 observed DNA molecules (yellow panels, figure 5.5). The transient FRET signals were concurrent with a decrease in acceptor PSF width and were consistent with BS complex formation. The number of analysed events was limited by the restricted observation time (due to photobleaching) and by complex photophysics of some fluorophores (see section 5.2.4).

5.3.2 Quantifying complex formation frequency

To quantify the frequency of complex formation events I calculate the probability of complex formation per particle by dividing the total number of events observed by the initial number of substrate molecules in a given movie. To account for the varying length of movies, for example due to cessation caused by focal drift, I divide this probability by the mean movie length to obtain an approximate probability of complex formation, per particle, per unit time[2]. It is also worth noting that very short lived synaptic complexes are likely to be missed due to the temporal resolution of 100 ms, thus the probability values given in the case of synaptic complex formation are likely to be lower limits. I use this to compute an estimated probability of formation for BS complexes to be $P \sim 3.8 \times 10^{-4}$ s^{-1},

[2]This calculation assumes that all substrate molecules are available to participate in reactions, and that photobleaching is not a major factor over the timescales of observation. These assumptions likely do not hold for the experimental conditions used, so this probability cannot be used to accurately infer information about the absolute on-rate of complex formation. However, the above effects are equivalent over all experimental conditions so this value can reliably be used to compare between experiments to determine the relative frequency of complexes formed.

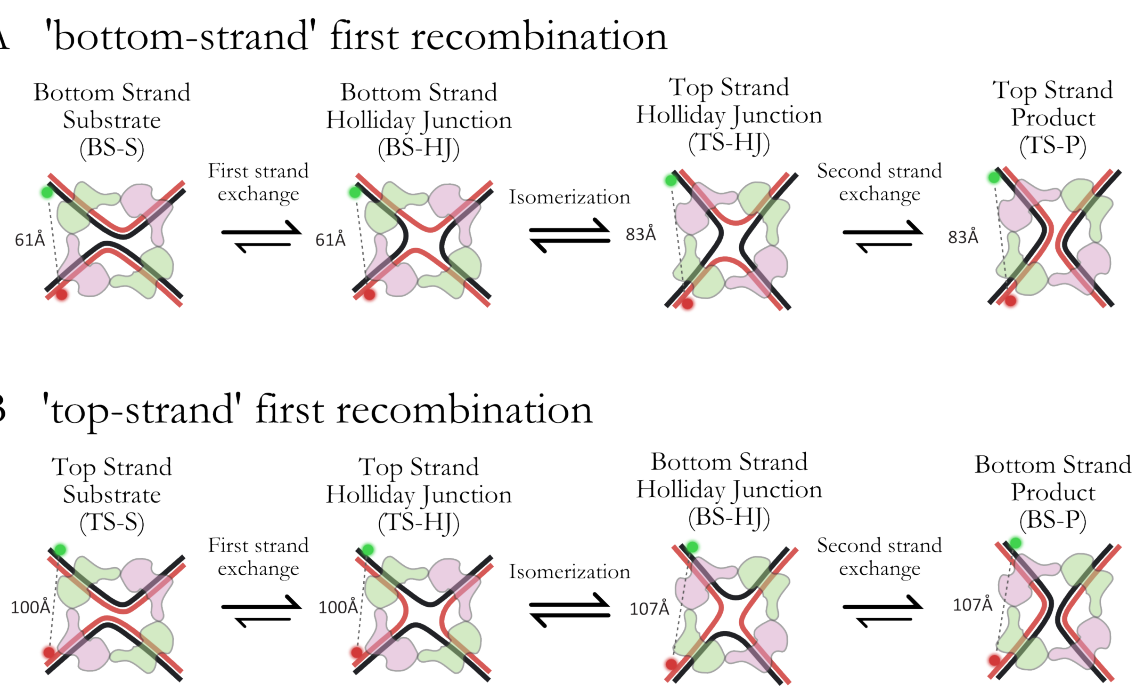


Figure 5.6: (Schematic of the two possible reaction paths for recombination: (A) 'bottome-strand' and (B) 'top-strand' first. The estimated fluorophore distances are shown for each complex in the reaction (using the BS-FRET substrate).

which can then be compared to the relative frequency at which other complexes form (see below).

Molecules undergoing BS complex formation were then classified according to the PSF width after loss of the transient FRET signal (i.e. a transition to $E^* < 0.2$). The majority of molecules (76%) showed a PSF that remained narrow after the decrease in FRET, representing complexes that progressed through HJ isomerisation, reaching either a TS complex (TS-HJ or TS-P) or dissociated recombinant product (pink panels, figure 5.5A). Previous studies have shown that Cre-*loxP* TS-HJs are efficiently resolved to recombinant product (Lee and Sadowski, 2001); thus it is likely that the majority of complexes after HJ isomerisation have undergone two pairs of strand exchange and formed a recombinant product, and are classified as 'productive complexes'. The remainder of molecules, where the PSF returns to that of free substrate DNA (left panel, figure 5.7B) represented BS complexes that dissociated back to the substrate (either having formed a BS-HJ intermediate or not), and are classified as 'non-productive complexes'.

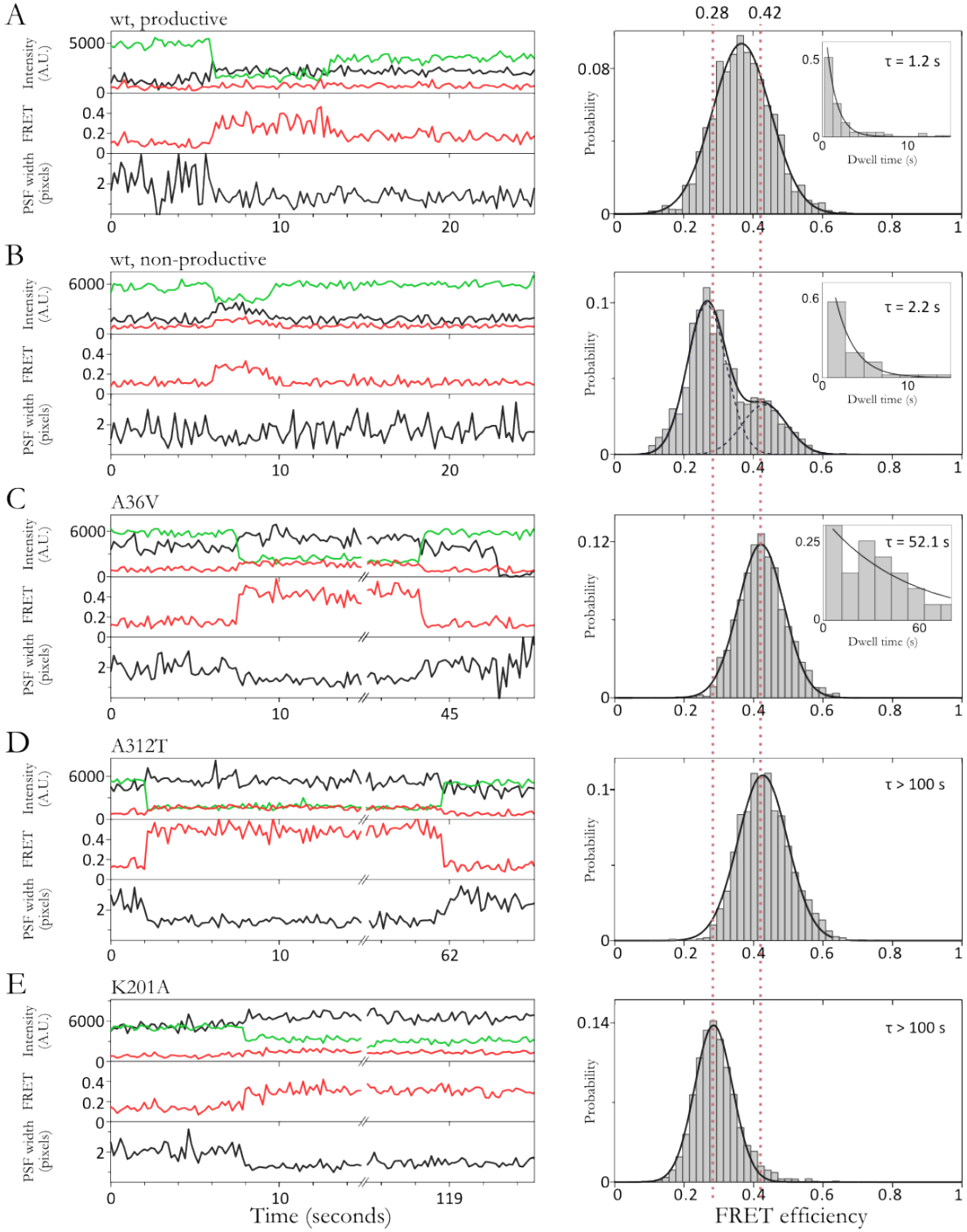


Figure 5.7: FRET efficiency histograms of protein-DNA complexes for wild-type Cre and mutants. Example time traces and population histograms for: (A) wt Cre productive events (91 molecules). (B) wt Cre non-productive synaptic complex events (48 molecules). (C) Cre A36V (27 molecules) (D) Cre A312T (25 molecules) (E) Cre K201A (86 molecules). Histograms plotted with E^* values from each frame of an event, counts added to the histogram were normalised by the length of the event in order to more accurately represent the population of molecules, histogram area was normalised to one (see section 5.2.4. Guidelines of $E^* = 0.42$ and 0.28, corresponding to the expected BS-HJ FRET efficiency and that measured for the K201A mutant respectively, are shown as red dashed lines. Dwell times of molecules in panels A, B, and C were fit to single exponentials (insets, mean lifetime, τ, shown). Panels D and E show very stable complexes and dwell times could not be measured over the duration of experiments.

5.3.3 Bottom-strand synaptic complexes are preferentially assembled

In the above experiment, events that showed PSF width changes without a concurrent FRET signal were also observed (42 of 6548). These events can be attributed to DNA molecules recombined faster than alternation time used in the experiments (200 ms), or to synaptic complexes which do not show a significant FRET in the BS FRET substrate (e.g. 'top-strand' synaptic complexes; figure 5.6).

To obtain the fraction of recombination events missed due to the temporal resolution of 200 ms, I extrapolated the dwell-time distribution of productive FRET events (right panel inset, figure. 5.7A), and estimated the fraction of missed events to be $\sim 16\%$ of all recombination events. As this value did not account for all events which did not exhibit FRET, a second DNA substrate was constructed (TS FRET; figure 5.3) that produced a FRET signal upon formation of TS-S (but none in the BS-S complex). Using this substrate, infrequent formation of TS-S was observed ($P \sim 4.55 \times 10^{-5} s^{-1}$; 9 of 1212 molecules), with $\sim 50\%$ proceeding through recombination. Moreover, most recombination events did not show FRET, indicating recombination through BS-S. These results demonstrate that BS synaptic complexes are formed preferentially over TS complexes by a factor of ~ 8 [3].

	Mean	Standard error of the mean	Standard deviation (SD)	Standard error of the SD
wt, productive	0.36	0.013	0.085	0.006
wt, non-productive (high)	0.43	0.015	0.065	0.008
wt, non-productive (low)	0.26	0.023	0.057	0.011
A36V	0.42	0.018	0.067	0.009
A312T	0.42	0.021	0.073	0.011
K201A	0.28	0.009	0.056	0.004

Table 5.1: Fit parameters of width and standard deviation for the Gaussian fits to histograms in figure 5.7.

	Corrected FRET	Standard deviation	Distance (Å)	Standard deviation
wt, productive	0.58	0.07	67	5
wt, non-productive (high)	0.65	0.06	64	5
wt, non-productive (low)	0.42	0.1	75	7
A36V	0.62	0.09	65	6
A312T	0.64	0.07	64	5
K201A	0.44	0.11	74	7

Table 5.2: Corrected FRET efficiencies and corresponding distances from the populations fitted to histograms in figure 5.7.

[3] I do not expect the intervening DNA (connecting the two *loxP* sites) to have a significant effect on the relative propensity of TS or BS synaptic complex formation due to the long (and therefore flexible) nature of this DNA, as well as the relatively subtle difference in DNA arrangement between the two complexes. However, complexes that require a more drastic change in DNA conformation (i.e. parallel complexes) could be affected by the length of the intervening DNA (these effects have previously seen in looping between two sites < 200 bp apart (Laurens et al., 2012)) and as such are not addressed in this work.

5.3.4 Non-productive complexes

Analysis of non-productive complexes (i.e. transient FRET, followed by return to a broad PSF; 48 molecules) revealed two distinct populations (right panel, figure 5.7B). Approximately 30% of complexes show $E^* = 0.43 \pm 0.015$ (mean of fitted Gaussian and standard error); after corrections (Lee et al., 2005), this FRET efficiency corresponds to an inter-fluorophore distance of 64 ± 5 Å, table 5.2 (see Appendix A). This distance agrees well with the distances from crystal structures of BS-S and BS-HJ (Ghosh et al., 2007), which predict a distance of $\sim$ 61 Å(figure 5.2C). The remaining non-productive complexes ($\sim$ 70%) correspond to a population with $E^* = 0.26 \pm 0.023$ and inter-fluorophore distance of 75 ± 7 Å. This distance is significantly longer[4] than that predicted from crystal structures (Ghosh et al., 2007), suggesting that, in solution, Cre can form complexes with a different architecture from the structure adopted in crystals.

To link the observed intermediates with stages of the reaction, I also tested three Cre mutants: A36V, shown to be partially defective in synapsis, but competent in HJ resolution (Wierzbicki et al., 1987; Lee and Sadowski, 2003a; Ghosh et al., 2007); A312T, known to be proficient in synapsis and first strand exchange, but inefficient in recombination, thus accumulating HJs during reactions (Hoess et al., 1987); and K201A, defective in DNA cleavage but proficient in synaptic complex formation (Ghosh et al., 2005). Pawel Zawadzki confirmed the defects of these mutants in ensemble recombination assays (figure 5.4), and I proceeded to used TFM-FRET to assess their ability to proceed through the reaction. I observed frequent synaptic complex formation with A36V ($P \sim 4.5 \times 10^{-4}\text{s}^{-1}, \tau \sim 52.1$ s), with $E^* = 0.42 \pm 0.018$ (figure 5.7C) that corresponds well to the higher-FRET species seen in the non-productive complexes of wt Cre, and in crystal structures (Ghosh et al., 2007). Despite forming synapses, the A36V mutant never completed recombination; since it is known that A36V can resolve HJs (Lee and Sadowski, 2003a), the main state formed by this mutant is most likely to be the BS-S (rather than the BS-HJ) complex.

The A312T mutant formed stable complexes (τ 100s; left panel, figure 5.7D) with a frequency ($P \sim 5.1 \times 10^{-4}$ s^{-1}) similar to A36V and $E^* = 0.42 \pm 0.021$ (right panel, figure 5.7D). Since the E^* value for A312T matches that of A36V and the existing BS-HJ crystal structure, it is likely that the HJ adopted is predominantly of the BS conformation.

Finally, the K201A mutant formed stable synaptic complexes ($P \sim 7.83 \times 10^{-4}$ s$^{-1}, \tau$ 100 s;

[4]Although the errors on the absolute distances calculated for the two observed populations of non-productive synaptic complexes would appear to make the difference non-significant, the error in the relative distance difference is in fact smaller, 11 ± 1 Å, assuming that the environment of the fluorophores does not significantly change between the two conformations.

left panel, figure 5.7E) with $E^* = 0.28 \pm 0.009$ (right panel, figure 5.7E) corresponding to an inter-fluorophore distance of 74 ± 7 Å. This value is surprising, since crystallographic studies have shown that K201A mutant interacts with *loxP* sites to form a synaptic complex that is almost identical to the canonical BS conformation (Ghosh et al., 2007). It has previously been suggested that Lys201 plays a role in organizing the un-synapsed Cre*loxP* complex into a cleavage-proficient conformation, which the K201A mutant may be unable to achieve (Lee and Sadowski, 2003a; Ghosh et al., 2005). Since it has been shown that crystallisation conditions can alter the behaviour of Cre mutants Y324F and R173K (Ghosh et al., 2007), I suggest that the discrepancy between my results and crystal structures could be due to crystal packing effects suppressing a structural defect of K201A. Intriguingly, the E^* of K201A complexes matches that of the low-FRET non-productive complexes observed with wt Cre (right panel figure 5.7B); although the equivalence or structure of the two cannot be concluded on the basis of a single FRET value, I suggest that they may both represent the same non-recombinogenic conformation, which could represent an out-of-plane twist of the HJ arms.

5.3.5 Evidence for fast interconversions between Holliday junction isomers

Analysing productive complexes (figure 5.7A) that exhibit more than 2 frames of transient FRET (91 molecules), I observed a FRET distribution with mean $E^* = 0.36 \pm 0.013$ (right panel, figure 5.7A), significantly (p-value < 0.01) lower than that observed for A312T and A36V ($E^* = 0.42$; right panels, figure 5.7C, D).

In addition, the width of the FRET distribution for productive complexes (standard deviation, $SD = 0.085 \pm 0.006$; right panel, figure 5.7A) is significantly greater than those observed for A312T and A36V ($SD \sim 0.07$; right panels, figure 5.7C, D; table 5.1). Since FRET distributions from ensembles of single molecules are broadened due to inter-molecule heterogeneity (Holden et al., 2010), simple comparisons of the observed width with that expected on the basis of the statistical nature of photon emission ('shot-noise') cannot be relied upon. However, since photon counts throughout these experiments are similar, it is likely that the excess FRET-width in the case of productive complexes is a result of structural dynamics that cause unresolved FRET fluctuations. Specifically, since the reaction requires HJ isomerisation before second strand exchange (figure 5.2A), I suggest that this broadening, along with the shift to a lower mean E^*, compared to A312T and A36V, reflects isomerisation of the BS-HJ ($E^* \sim 0.42$) to TS-HJ ($E^* \sim 0.18$) occurring reversibly at timescales

faster than the exposure time (100 ms).

FRET simulation

To further explore the proposed interconversion between HJ isomers and determine if rapid interconversion between two conformers could give rise to the observed FRET distributions, I performed Monte Carlo simulations to infer information about the underlying process (Santoso et al., 2010). I implemented a simulation of stochastic fluorophore emission from two states which were able to interconvert faster than the frame time of acquisition using a Markov chain approach. The two HJ states were assigned states with $E^* = 0.18$, and 0.43, for the TS-HJ and BS-HJ respectively, estimated from the 2HOI crystal structure (Ghosh et al., 2007). The fluorescence emission collected over a frame was then put through simulated background noise, and EM gain to obtain the final emission counts. The rate of fluorophore photon emission was tuned to produce distributions which matched the observed width arising from a static FRET standard. This gives rise to photon counts which are somewhat lower than those from single molecules in my experiments to reproduce the effect of broadening due to collating an ensemble of single molecules in the measured FRET distributions (Holden et al., 2010).

To obtain an estimate for the rates of inter-conversion which could recapitulate the observed distribution, I systematically varied the forward and backward rates of HJ interconversion (figure 5.8) and performed many simulated runs for each parameter pair, for a number of frames corresponding to those in the experimental data. I then fitted the output distribution using a single Gaussian and recorded the mean, variance, and coefficient of determination (R^2) of the fit. Assuming a normal distribution of parameter values for each model I then computed the likelihood of the model given the observed data (figure 5.8). By the definition of likelihood, the parameters which maximised this value would be those which are most likely to give rise to the data.

The ratio of interconversion rates effectively determines the mean of the distribution observed, and their magnitude dictates the width, with very slow rate resulting in two resolved populations, and fast rates leading to a width indistinguishable from that of a single state. From these results I estimate an equilibrium bias of 3:1 toward the BS-HJ, and forward and backward rates in the range of 20-40 s^{-1} and 50-100 s^{-1}, respectively, as these parameters best recapitulate the observed distributions, figure 5.8.

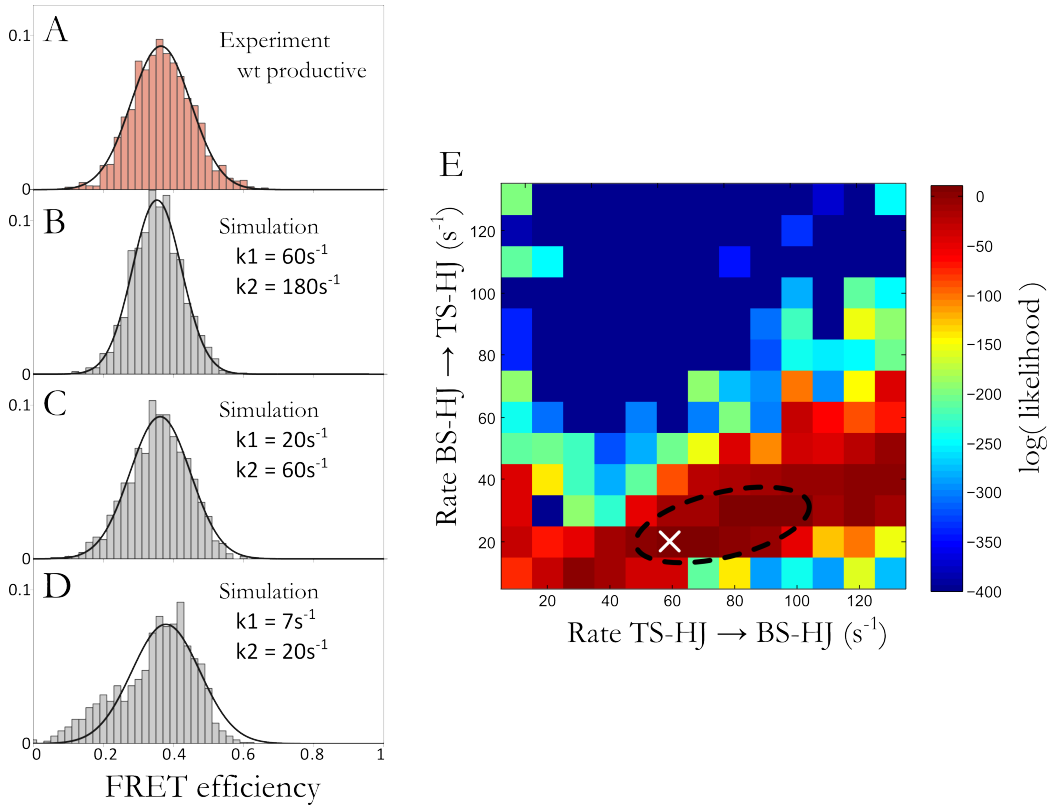


Figure 5.8: (A) Comparison of FRET efficiency distributions between a experimental wild type Cre productive complex distribution and, (B-D), simulated data with varying forward and backward rates of interconversion between the two Holliday Junction isomers. (B) Likelihood of the observed data given the distributions obtained from simulating many observations with a forward and backward rate of interconversion between Holliday Junction isomers, see section 5.3.5. The region of the graph giving the maximum likelihood values is outlined in black, and the pair of rates giving the maximum value is marked with a white cross.

5.3.6 The recombinant product synaptic complex is extremely stable

Using the BS-FRET substrate (figure 5.3A), I could not distinguish between the recombinant product complex (TS-P) and free recombinant DNA, as they have the same TFM-FRET signature (pink panels, figure 5.5A). To probe the stability of TS-P, I used a substrate with inverted *loxP* sites, figure 5.3; although recombination within this substrate inverts the DNA segment between *loxP* sites without producing a FRET signal, complex assembly and dissociation can be followed through changes in PSF width, figure 5.9. Using this substrate, I observed that complexes were stable for longer than the 2-minute acquisition time of experiments (figure 5.9), suggesting the TS-P complex is very stable.

The high stability of the TS-P complex made it difficult to characterise its dissociation constant, K_d. Collaborators at the Nanobiomedical Centre, A. Mickiewicz University, were able to corroborate my observation and measure the K_d of TS-P using a fluorescence correlation spectroscopy (FCS) assay (figure 5.10). Using FCS, they monitored changes in translational diffusion as the BS FRET substrate was converted, via recombination, from a long DNA substrate (1087 bp, diffusion coefficient, $D \sim 0.24 \times 10^{-6}$ cm^2 s^{-1}) to a short DNA product (61 bp; $D \sim 0.92 \sim 10^{-6}$ cm^2 s^{-1}) reflecting dissociation of the TS-P complex to free product DNA. To monitor TS-P dissociation after recombination, this change in diffusion coefficient was measured over time under indetical buffer conditions to my single-molecule experiments; in parallel, aliquots quenched with SDS (to disrupt protein interactions) were measured in order to evaluate progress of the reaction irrespective of TS-P dissociation, figure 5.10A.

Whereas SDS-quenched reactions displayed a rapid decrease in the fraction of substrate (similar to previous observations (Ringrose et al., 1998)), the native reaction showed a much slower decrease, indicating that TS-P dissociation occurs at a much slower timescale , figure 5.10B. Single-exponential fits give rates of 0.01 min^{-1} and 0.96 min^{-1} for native and SDS-stopped reactions respectively, indicating that the lifetime of the TS-P complex is ~ 75 minutes. These dissociation rates agree well with those derived from a kinetic model of Cre recombination (Ringrose et al., 1998) and previous observations of highly stable complexes (Hamilton and Abremski, 1984; Shoura et al., 2012).

5.4 Discussion

Cre mediated site-specific recombination demonstrates the utility of the combined TFM-FRET method in correlation global and local DNA rearrangements in addressing questions of genuine

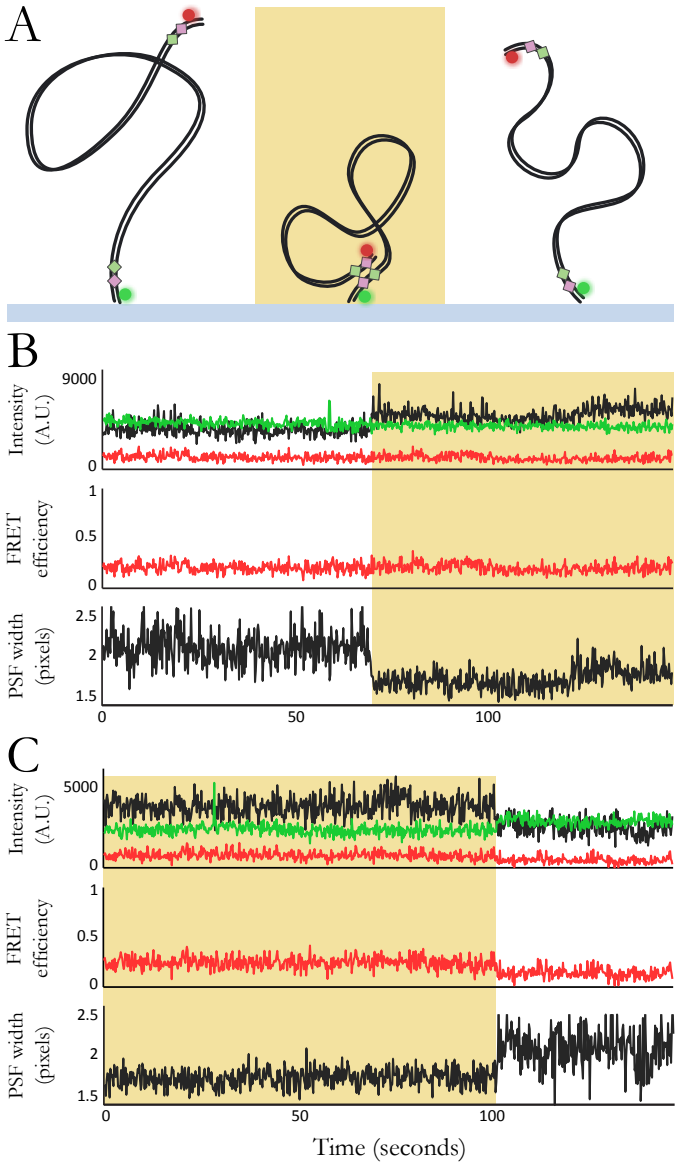


Figure 5.9: Using an alternative substrate to assess the stability of the recombinant product complex. (A) Schematic of substrate used with *loxP* sites in an inverted arrangement. BS-S complex formation leads to a decrease in PSF width but no FRET. Recombination results in the inversion of the DNA linking the two *loxP* sites of the substrate, and dissociation of the product complex leads to a return to high PSF. (B) Example time trace of a single substrate molecule undergoing complex formation. (C) Time trace from a separate molecule showing dissociation of a complex to a long DNA, which is attributed to dissociation of the recombinant product synaptic complex.

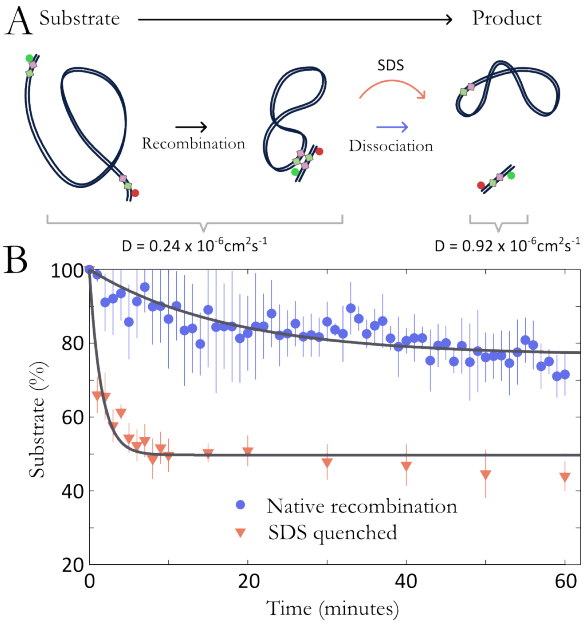


Figure 5.10: Stability of product synaptic complex. (A) Schematic of the fluorescence correlation spectroscopy (FCS) assay used to investigate kinetics of TS-P dissociation. Dissociation, which can either occur naturally or be induced by addition of SDS, leads to an increase in the diffusion coefficient of the labelled species. (B) Progress of native (blue) or SDS quenched (red) recombination reactions over time.

biological interest. The TFM observable of PSF width allows single reactions to be followed and the outcome, productive or otherwise, to be addressed. As an important advantage over previous TPM studies (Fan, 2012), the FRET efficiency allows the assignment of observed intermediates to conformations seen in crystal structures. Using TFM-FRET, I was able to identify conformational states formed during individual recombination reactions, and analyse their conversion kinetics. In addition to following complete reactions, I also identified molecules that progress only partially through recombination.

5.4.1 Reaction paths and a novel intermediate

The direct observation of intermediates allows new light to be shed on Cre-*loxP* recombination. Previous work has shown a preference for recombination via 'bottom-strand' exchange (Hoess et al., 1987; Lee and Sadowski, 2003b; Ghosh et al., 2005) and I confirmed that this preference arises because of an 8-fold bias toward BS synaptic complex formation, rather than major differences in recombination efficiency in BS and TS complexes (BS = 75%, TS $\sim$ 50%). I also showed that, alongside assembly of the canonical BS structure, wild type Cre (and K201A) can form, previously unobserved, non-productive complexes with a structure distinct from those captured in crystals. Although the role of such complexes is not clear, their existence needs to be considered in any ensemble kinetic studies of Cre recombination, and highlights the ability of FRET to examine distance differences on the nanometre scale.

5.4.2 Rate-limiting step for productive complexes

My results establish that, once formed, synaptic complexes proceed rapidly through the reaction. For productive complexes, I measured the lifetime of the transient FRET state of $E^* \sim 0.36$ (representing a mixture of BS-S, BS-HJ, and TS-HJ) to be ~ 1.2 s; this is the first measurement of its kind in Cre recombination. The single-exponential decay of the dwell times in this state points to a single rate-limiting step before further progress through the reaction. Since this FRET state reflects, in part, the interconversion between HJ complexes, I conclude that the rate-limiting step is concurrent with, or occurs after HJ isomerisation. This rate limiting step could be the second DNA strand exchange, or a conformational change that ends HJ isomerisation and/or changes the activity of the Cre monomers. This can be compared to a study on λ Int (which, in contrast to Cre, requires accessory proteins) that observed that a rate-limiting step occurred before, or concomitant with, HJ formation (Mumm et al., 2006).

5.4.3 Holliday junction dynamics

In productive complexes I infer the presence of previously undetected dynamics of the HJ intermediates, estimating isomerisation rates of $10 - 100$ s^{-1}. Intriguingly, these values are similar to rates observed for transitions between protein-free HJ conformers ($20 - 100$ s^{-1} (McKinney et al., 2002)), suggesting that, although Cre affects the structure of the HJs formed (Gopaul et al., 1998), it does not substantially affect the isomerisation rate. In contrast, no appreciable interconversion between HJ isomers is seen using the A312T mutant, indicating that this mutant favours the BS conformation, either due to a change in isomerisation rates between HJ isomers, or due to an alteration in HJ cleavage proficiencies. Although from published biochemical data, I do not expect that the A36V mutant will appreciably populate the HJ state (Lee and Sadowski, 2003a; Ghosh et al., 2007), it cannot be ruled out that if it is able to form a HJ, it may also be undergoing some biased HJ isomerisation (with the equilibrium being heavily shifted towards the BS-HJ form).

5.4.4 Stable recombinant synaptic complexes

After completion of recombination, single-molecule experiments suggest long-lived recombinant product synaptic complexes (TS-P) refractory to subsequent recombination, and this is confirmed and quantified by measurements by collaborators from A. Mickiewicz University. The measured

lifetime $\sim$ 75 minutes agrees with ensemble observations (Hamilton and Abremski, 1984; Shoura et al., 2012), but is significantly longer than that observed in a recent TPM study (Fan, 2012); however, it should be noted that the use of a 200-nm bead in the TPM work generates an effective stretching force ($\sim$ 30 fN) on the DNA (Segall et al., 2006), and this may have increased the rate of complex dissociation. Previous ensemble biochemical data had measured similar dissociation constants (K_d) for the synaptic complexes of wt Cre and K201A (14 and 9 nM, respectively (Ghosh et al., 2007)). I observe that wt Cre complexes quickly form the TS-P complex ($\tau \sim 1.2$ s, figure 5.7A), while those formed by K201A are very stable ($\tau > 100$ s). Unlike the ensemble assay, I was able to distinguish between BS-S and TS-P wt complexes, observing that BS-S is transient and TS-P is highly stable. My results indicate that the previous K_d measurements for wt Cre were dominated by the TS-P complex, rather than the BS-S complex (Ghosh et al., 2007). These results also show that the K_d of $>$ 500 nM for the A36V synaptic complex (Ghosh et al., 2007) does not indicate a deficiency in complex formation, but instead a $\sim$ 100-fold reduction in the stability of the complex, as compared to wt Cre TS-P. It is worth noting that, although TS-S and TS-P (i.e. TS complexes formed as either the substrate or product of the recombination reaction) could naively be assumed to be identical, my results indicate significant differences in the stability of each. This difference between complexes suggests an intriguing avenue for further study, which may shed further light on the mechanisms of regulation within Cre-*loxP* recombination.

5.5 Conclusion

In this chapter I have demonstrated the combination of the tethered fluorophore motion technique developed in Chapter 4 with single-molecule FRET measurements. In applying this TFM-FRET methodology to the Cre site-specific recombinase system I was able to build on previous single-molecule and ensemble studies of Cre. Crucially the combination of global (TFM) and local (FRET) structural information allows the study of Cre-*loxP* recombination in unprecedented detail, revealing the structural heterogeneity of complexes assembled and their relative propensities, as well as measuring kinetics associated with various steps in the reaction. The technique could generally be used to study the action of both families of site-specific recombinases (Mumm, 2012) and in the next chapter I extend this methodology to the study of a more complex, multi-protein, site-specific recombinase, XerCD, and its interactions with the motor protein FtsK.

5.6 Contributions

Pawel Zawadzki prepared all proteins and DNA substrates (apart from the Cy3B labelling step), as well as assisting in performing some of the numerous experimental repeats required for adequate statistics. All FCS experiments were performed by Jaroslaw Mazuryk, A. Mickiewicz University. David Sherratt and Lidia Arciszewska provided valuable advice and assistance in the biological interpretation of results.

Chapter 6

XerCD-*dif* recombination

6.1 Introduction

6.1.1 XerCD recombinase

XerC and XerD are a pair of closely related (Blakely et al., 1993) tyrosine site-specific recombinases of *Escherichia coli* that work in concert to help segregate chromosomal DNA and maintain plasmid copy number. The mechanism of reaction is thought to be very similar to that discussed in the previous chapter with respect to Cre. However in the case of XerC and XerD (XerCD), which each recognise and bind to adjacent halves of the recombination site, the process is more tightly regulated and depends both on numerous accessory proteins as well as the sequence of the recombination site. The basic biological role of XerCD is to unlink dimers of chromosomes or plasmids which are formed during the process of DNA replication. If homologous recombination produces an odd number of crossovers during replication, the two daughter chromosomes will be linked; clearly, unlinking of this chromosome dimer is crucial for successful cell division, similarly any plasmid dimers need to be separated in order for stable inheritance of plasmids. The unlinking of chromosome dimers is an important consideration for any organisms with a circular chromosome, and well conserved Xer-like recombinase sequences are prevalent in the chromosomes of a wide variety of bacteria (Subramanya et al., 1997).

Given its function, action of XerCD must be tightly regulated as indiscriminate recombination could lead to the formation of chromosome or plasmid dimers from monomers, the opposite of the desired outcome. To this end, and due to its multiple functions, the XerCD recombinase system is

much more complex than that of Cre, studied in the previous chapter; firstly it involves two distinct proteins, moreover it can act at three different recombination sites, all of which require the presence of accessory proteins.

Recombination at *cer*, and *psi*

Recombination by XerCD is known to occur at three distinct sequences: *dif*, *cer*, and *psi*. The sequences *cer*, and *psi* are responsible for directing the separation of plasmid dimers (plasmid pSC101 for *psi* (Arciszewska et al., 1997) and ColE1 for *cer* (Blakely et al., 1993)) produced during DNA replication. These short (28-30 bp) sequences are also flanked by $\sim$ 200 bp of DNA sequences to which accessory proteins bind to facilitate successful recombination, these proteins impose a 'topological filter' and ensure than only intra-molecular recombination occurs efficiently with a well defined topology (Colloms et al., 1997). The complex nature of recombination at *cer*, and *psi* as well as the requirement that this occurs on supercoiled plasmids makes single-molecule fluorescence studies of these a challenging prospect, here I will be focussing exclusively on Xer recombination at the chromosome encoded *dif* sequence.

Recombination at *dif*

The *dif* site is a 28 bp DNA sequence encoded in the *E. coli* chromosome, located near the terminus of replication (Blakely et al., 1993). In the presence of a chromosome dimer, XerCD can act to recombine the two *dif* sites located on the joined daughter chromosomes, thereby separating them such that cell division can be successfully completed, figure 6.1. In cells which are unable to use the Xer system, $\sim$ 15% of divisions are prevented due to inability to resolve a chromosome dimer (Grainge et al., 2011). XerCD show strong co-operative binding to *dif* ($K_d < 1$ nM (Blakely and Sherratt, 1996)) and synaptic complexes between two *dif* sites, each bound by XerCD can be formed. The synaptic complex formed by XerCD is assumed to have a similar structure to Cre (near-planar, pseudo-symmetric, figure 5.2C) although the XerCD synaptic complex has never been captured either in crystals of bulk assays. Without the action of the accessory protein, FtsK, on the complex recombination cannot occur.

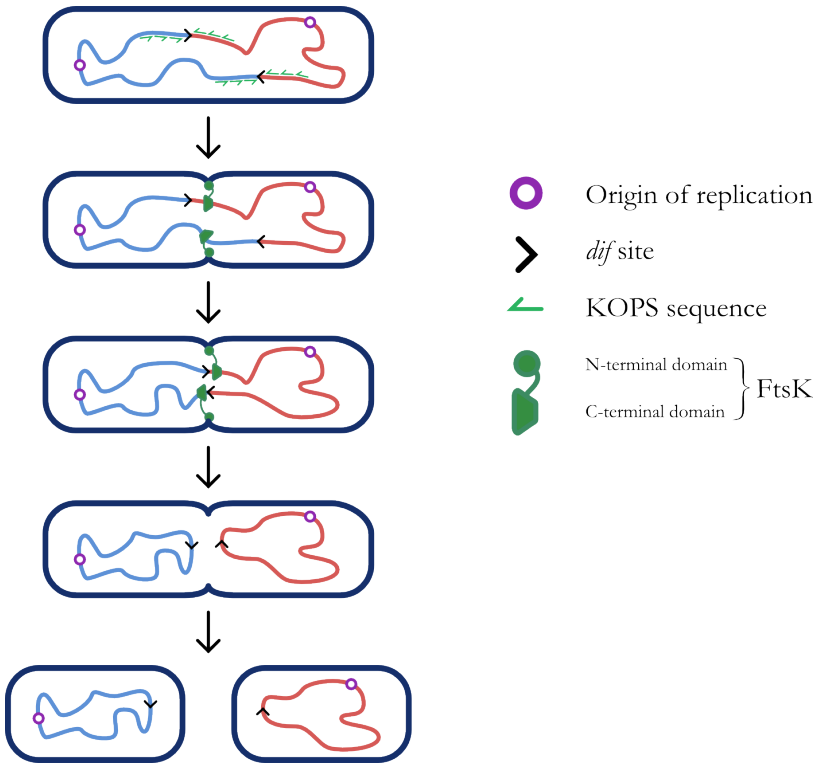


Figure 6.1: Schematic illustration of the process of chromosome dimer resolution. Chromosome dimers are resolved by XerCD at chromosomal *dif* sites located near the terminus of replication. The two *dif* sites on a chromosome dimer are each bound by XerCD and the septum located DNA translocase FtsK is thought to bring the distant sites into proximity for recombination, with its direction of action dictated by KOPS. The action of XerCD and FtsK promote recombination of the two *dif* sites leading to resolution of the chromosome dimer after which normal cell division can proceed. It should be noted that many details of the reaction (such as two independent FtsK molecules pumping DNA) shown in the illustration are actually unknown and only one possibility is shown here for clarity.

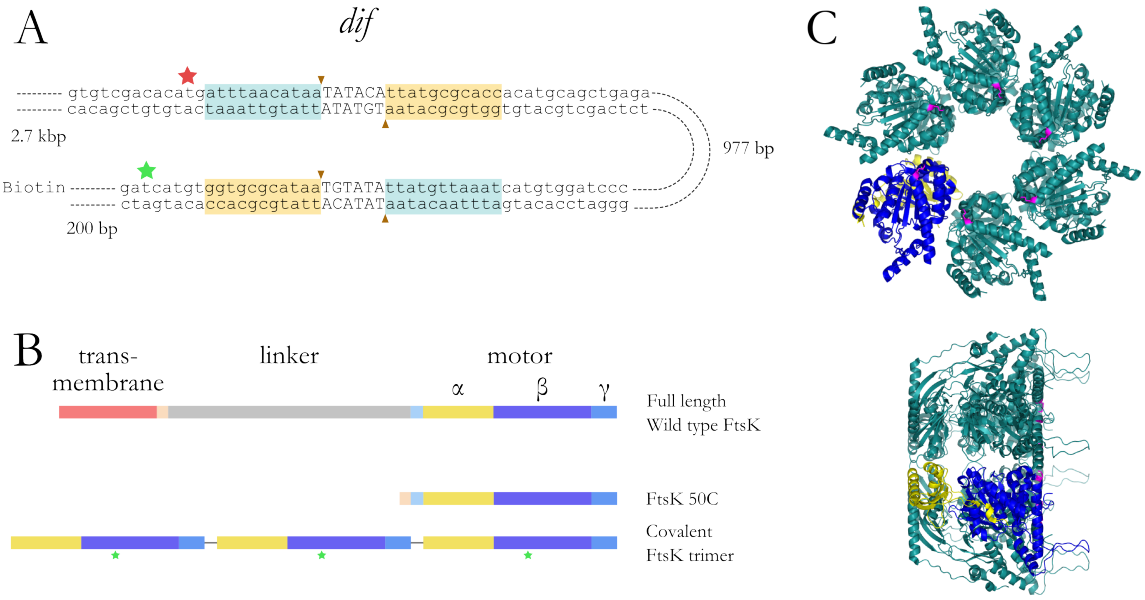


Figure 6.2: (A) Sequence of the DNA used for recombination experiments throughout this chapter. The *dif* sequence is shown with binding sites of XerC and XerD highlighted in blue and orange respectively. Locations of DNA cleavage are indicated by arrows, and the labelling positions of fluorophores are shown as stars. The long sections of DNA around the *dif* sites are not shown but are represented by dashed lines with lengths as indicated. (B) Schematic of the domains of FtsK and compositions of frequently used FtsK 50C derivative, and the covalent FtsK trimer used here. The approximate labelling position of the Cy3 Ftsk trimer is indicated by stars. (C) Front and side view of a hexamer modelled from the crystal structure of the *Pseudomonas aeruginosa* FtsK (PDB: 2IUU (Massey et al., 2006)). The α and β domains of one monomer are highlighted in yellow and blue respectively, the cysteine residue used for fluorophore labelling is shown in magenta.

6.1.2 FtsK

FtsK is a 1329 amino acid protein located in the septum of a dividing *E. coli* cell, it comprises of an N-terminal transmembrane domain, connected to the C-terminal intracellular DNA translocase domain via a long (600 residue) linker region figure 6.2. The N-terminal domain acts as an anchor in the bacterial septum, which in turn causes the ATP dependant translocase to effectively act as a DNA pump (Aussel et al., 2002). This translocase domain thus acts to position *dif* sites ready for recombination and also plays a crucial role in 'activating' XerCD so that the full recombination reaction can be carried out to resolve the chromosome dimer.

Translocation of FtsK

The translocase domain of FtsK consists of three subdomains, α and β comprise the motor, and γ has been implicated in activation of XerCD (Grainge et al., 2011) (see below). The αβ subdomains are known to assemble as hexameric complexes around DNA, and a crystollographic structure of this ring (in the absence of DNA) is shown in figure 6.2C. As FtsK is partially an integral membrane protein, studies on the action of the DNA translocase domain have focussed around one of two derivatives (illustrated in figure 6.2B): $FtsK_{50C}$ which comprises the αβγ C-terminal domains as well as a 50 residue sequence from the linker domain close to the N-terminus (this extra 50 residues was found to stabilise hexamer formation (Aussel et al., 2002)), more recently a covalently linked trimer of the αβγ domains was shown to be efficient in translocation and activation of recombination, without the aggregation problems associated with $FtsK_{50C}$ (Crozat et al., 2010).

Studies have been carried out both at the ensemble and single-molecule level as to the nature and regulation of the ATP dependant translocation by the FtsK motor domain (hereafter referred to simply as FtsK for convenience). Results have shown that FtsK translocates at speeds of up to 6 kbs^{-1} (Saleh et al., 2004; Pease et al., 2005; Ptacin et al., 2006; Lee et al., 2012a), can displace obstructions bound to the DNA (Chivers et al., 2010) (although it stops specifically at XerCD (Graham et al., 2010)), can undergo sudden changes in direction (Saleh et al., 2004; Pease et al., 2005; Ptacin et al., 2006; Lee et al., 2012a), and can loop the DNA substrate during translocation (Saleh et al., 2004, 2005; Pease et al., 2005; Ptacin et al., 2006; Bigot et al., 2006; Crozat and Grainge, 2010). There have also been several suggestions that FtsK hexamers can form higher order structures on DNA although their presence or relevance to conditions *in vivo* has yet to be determined (Massey et al., 2006; Pease et al., 2005; Saleh et al., 2004). FtsK also loads specifically and orientationally to KOPS (FtsK Orienting Polar Sequences), these short (8 bp) sequences, sensed by the γ region (Bigot et al., 2006), are asymmetrically distributed in the *E. coli* chromosome and are thought to direct FtsK translocation toward the *dif* site (Bigot et al., 2005; Ptacin et al., 2006; Lee et al., 2012a), figure 6.1.

6.1.3 Activation of XerCD recombination by FtsK

In the absence of FtsK XerCD can mediate formation of a complex in which XerC is active (CS-Syn, figure 6.3), and XerC is able to cleave DNA, leading to the formation of Holliday junctions (CS-HJs). However, without activation by FtsK, XerD remains inactive and these HJs can only be resolved

by XerC to the original substrate (Ferreira et al., 2003), figure 6.3 FtsK independent box. For recombination to complete, XerD must become active; it has been shown that the γ domain of FtsK is responsible for mediating this activity, and that its presence is sufficient to 'activate' the XerCD complex such that a DS-Syn complex is formed and XerD initially cleaves the DNA[1], followed by resolution of the HJ to recombinant product by XerC (Löwe et al., 2008; Grainge et al., 2011), figure 6.3, FtsK dependent box. This activation by γ is independant of FtsK translocation and the presence of the subdomain in solution, or covalently linked to either XerC or XerD is sufficient to promote recombination (Grainge et al., 2011). The translocation of FtsK does, however, have a role to play in controlling the topology of recombinant products; recombination reactions on a plasmid containing two *dif* sites have been shown to produce only unlinked products when recombination is activated by FtsK whereas those activated by γ alone result in the formation of catenated products (Ip et al., 2003; Grainge et al., 2011), although the mechanism involved is not well understood.

Details of the interaction and activation of XerCD by FtsK as yet remain unclear, precisely what this 'activation' entails, whether it acts on synapsed or unsynapsed complexes, how the catalytic activity of XerCD is controlled, and what influence translocation of FtsK plays, have yet to be determined. Even the state of FtsK during translocation, whether it operates as a single hexamer, or a higher order complex, and how it interacts with KOPS and roadblocks, and the nature of its directionality changes and relevance of the observed DNA looping are the subject of conflicting results and much debate. In this chapter I utilise the techniques developed and applied to the study of Cre-*loxP* recombination to the more complex situation of XerCD-*dif*.

6.2 Experimental methods

6.2.1 DNA and protein preparation

DNAs were prepared by PCR as described in chapters 4 and 5, as well as additional ligation steps to attach a trailing 2.8 kbp extension at the surface-distal DNA end, and 200 bp spacer between biotin and the first *dif* site, figure 6.2. All proteins were provided by the David Sherratt laboratory.

[1]The order of strand exchange for XerCD (i.e. XerC first or XerD first) are equivalent to the 'top strand' and 'bottom strand' first pathways discussed in chapter 5 as XerC and XerD act on separate strands of the DNA.

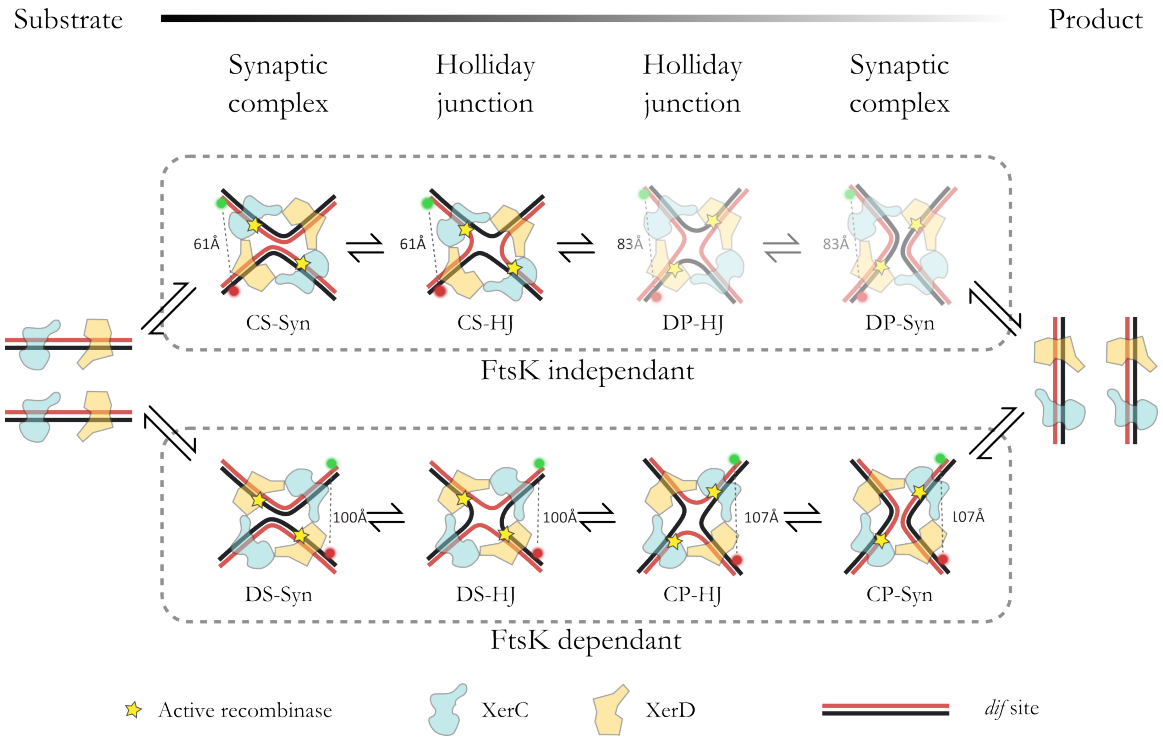


Figure 6.3: Schematic of the expected intermediate complexes of XerCD recombination at *dif*. XerC and XerD are shown as blue and orange respectively, and the active recombinase is indicated by a star. Fluorophore labelling positions and distances calculated from Cre crystal structures are shown. Complexes are labelled XY-Z, where X indicates the active recombinase (XerC or XerD), Y denotes whether the stage of recombination (S - substrate, or P - product), and Z describes the complex as synaptic (Syn) or a Holliday junction (HJ). The sequential nature of strand exchange in tyrosine family recombinases means that, in theory, two possible pathways to recombination exist, one in which XerC cleaves and exchanges DNA strands first, and on where XerD exchanges first. Previous studies have shown that in the absence of FtsK activation XerC is able to form synaptic complexes (CS-Syn) and cleave DNA to form a Holliday junction (CS-HJ), but that resolution by XerD is not permitted (indicated by faded complexes DP-HJ and DP-Syn). FtsK activation enables XerD to cleave and exchange strands to from the DS-HJ complex, followed by resolution of CP-HJ by XerC, and dissociation of CP-Syn leaves the recombined DNA.

6.2.2 XerCD recombination

Synapsis and recombination experiments were performed as for Cre in chapter 5, on the same microscope under identical buffer conditions. For synapsis experiments XerC and XerD were added at concentrations of 17 and 22 nM respectively. For recombination experiments 40 nM covalent FtsK trimer and 1 mM ATP were also added.

6.2.3 Ftsk DNA looping experiments

For FtsK looping experiments a 3 kbp DNA with a Cy5 internally labelled close to the surface-distal end was used (identical in sequence to the 3 kbp DNA used in chapter 4). For unlabelled FtsK looping experiments 40 nM covalent FtsK trimer, and 1 mM ATP were added to surface immobilised DNA under identical buffer conditions to above. Molecules were imaged under continuous 640 nm excitation at 3 mW, movies were acquired at a 33 ms frame rate using the CRIFF to ensure focal stability. For experiments with labelled FtsK, Cy3 covalent FtsK trimer was used at 40 nM, and molecules imaged using ALEX at a frame time of 100 ms and laser powers of 1 mW for both 532 nm and 640 nm excitation.

FtsK pulldown experiments were performed by first washing a neutravidin treated biotin-PEG slide with biotinylated His-tag directed antibody, followed by addition of the Cy3 covalent FtsK trimer at low concentration until the desired density of surface immobilised molecules was achieved. Without addition of the anitbody, very little immobilisation of the covalent trimer was observed at these concentrations.

6.2.4 Data analysis

Data extraction and analysis of synapsis and recombination events was performed under the same criteria as in chapter 5. Analysis of FtsK looping data was performed under similar criteria, traces that showed errant photophysical fluctuations, PSF widths outside normally expected limits, or multi-step photobleaching were discarded. For looping events, the start and end points were measured manually. FRET efficiency histograms were fit by a one dimensional Guassian, apart from distributions at low FRET which were fit with a Beta distribution (Gopich and Szabo, 2005).

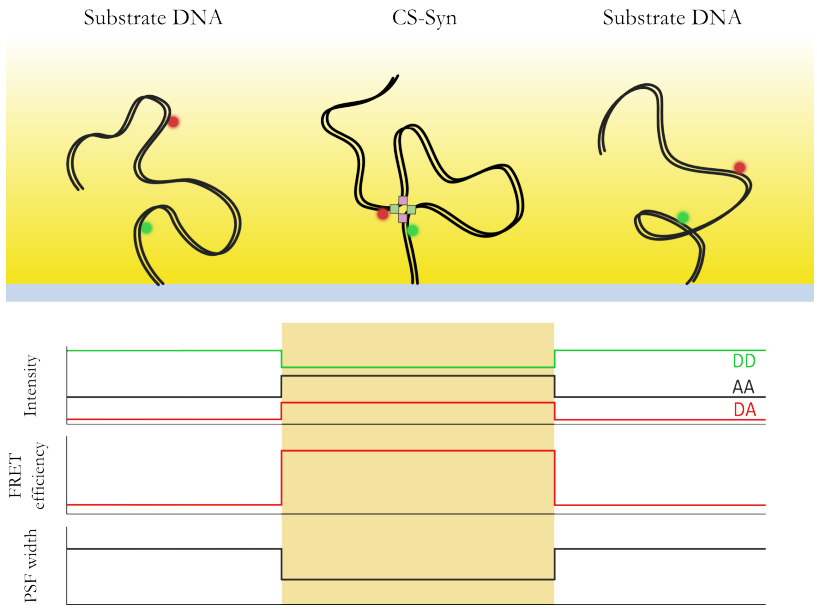


Figure 6.4: Schematic of synaptic complex formation experiment. Addition of XerCD to surface tethered DNA substrates lead to the reversible formation of CS-Syn complexes, expected to give both a decrease in PSF width and increase in FRET efficiency.

6.3 Results

6.3.1 XerCD synaptic complex

Throughout this chapter I employ a similar scheme of TFM-FRET to that used in chapter 5 to observe Cre-*loxP* recombination. Briefly, a long doubly labelled DNA (3000 bp) containing two *dif* sites, with a Cy3B donor, and Cy5 acceptor placed in close proximity, are bound to the slide surface using a biotin-neutravidin interaction. The two *dif* sites are located approximately 1 kbp apart, and the substrate DNA shows no FRET without addition of the proteins. Two important differences exist compared to the simple 1 kbp substrate used in the previous chapter. Firstly, after the second *dif* site there is a 2.8 kbp section of DNA designed to assist loading of FtsK, second, between the biotin and first *dif* site there is a 200 bp spacer to reduce steric clashes between the 2.8 kbp extension and the surface during complex formation. Addition of XerCD is expected to lead to formation of synaptic complexes between *dif* sites on a single DNA molecule and this will bring the Cy5 significantly closer to the surface, leading to a change in PSF width, figure 6.4. The fluorophores are placed such that, upon formation of an CS-Syn synaptic complex (i.e. one in which XerC is initially active), they are brought within range of FRET (according to distance estimates derived from the crystal structure of the closely related recombinase Cre). In this fluorophore configuration, other complexes such as the DS-Syn are expected to give much lower FRET efficiencies, see figure 6.3 .

When XerCD was added to surface immobilised DNAs repeated, reversible, occurrence of FRET was observed along with a decrease in PSF width, which I attribute to the formation of a CS-Syn

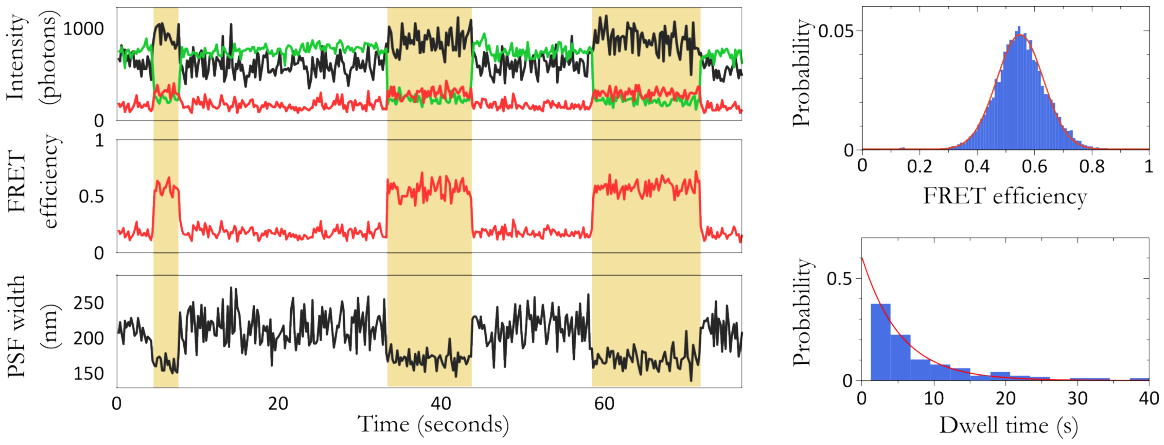


Figure 6.5: Synaptic complex formation by XerCD. (A) Example single-molecule time trace showing the repeated formation of an CS-Syn complex by a single DNA substrate molecule. Colours are as in figure 6.4 The FRET efficiency of complexes formed shows a single peak at $E^* = 0.55$, dwell times of synaptic complexes are fit well by a single exponential, with lifetime $\tau = 5.3$ s, indicating a single rate limiting step for dissociation of the complex. ($N = 196$ molecules)

complex in the absence of FtsK. Measuring the FRET efficiency during these events of $E^* = 0.55 \pm 0.01$ (mean of Gaussian fit and standard error of the mean) and following the procedure outlined in appendix A to convert to distance, gives an inter-fluorophore separation of 60 ± 5 Å. Plotting the dwell time of these synaptic events indicates a single rate limiting step, with an off-rate of $\tau \sim 5.4$ s.

Although other possible complexes, such as a DS-Syn are expected to show low FRET, figure 6.3, and would be difficult to distinguish from background by this criteria alone, monitoring PSF width allows their formation to be detected. On this basis my measurements indicate that the CS-Syn complex represents $> 95\%$ of those formed and that other possible conformations of the synaptic complex occur very infrequently.

6.3.2 XerCD recombination

When XerCD are incubated with surface immobilised DNAs in the presence of FtsK and ATP, recombination is able to occur. To assess the ability of molecules to recombine at the surface I performed end-point experiments similar to those described for Cre (section 5.3.1). Incubating the proteins for 10 minutes followed by an SDS wash to disrupt protein interactions, show that $\sim 75\%$ of DNA molecules had undergone recombination, indicated by a decreased PSF width corresponding to the DNA length expected for the product DNA.

Having established that XerCD recombination occurs efficiently at the slide surface, I attempted

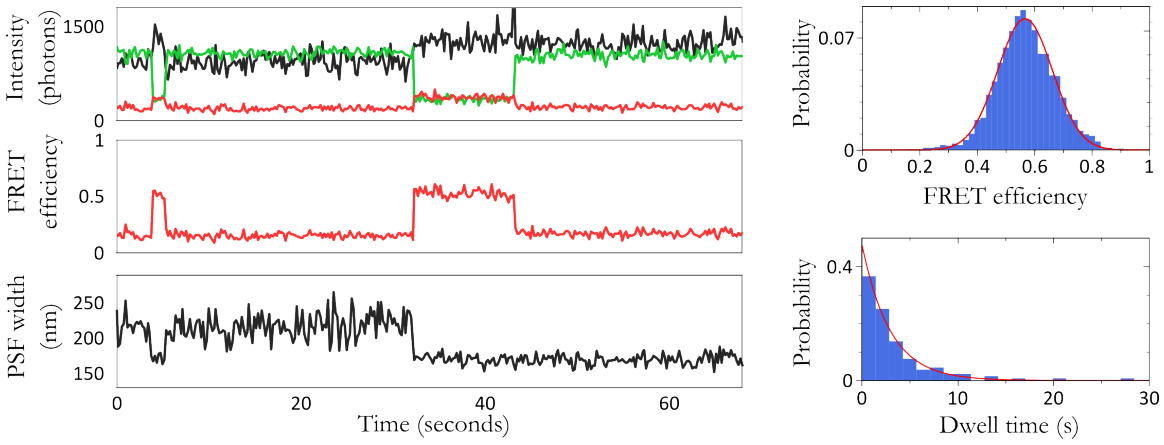


Figure 6.6: XerCD mediated DNA recombination. (A) Example single-molecule time trace showing reversible formation of an 'XerC active' synaptic complex at $\sim$ 5 seconds followed by successful recombination at $\sim$ 30 seconds. The FRET efficiency of the high FRET species formed prior to recombination shows a single peak at $E^* = 0.55$, and a dwell time fit by a single exponential, with lifetime $\tau = 2.8$ s($N = 131$ molecules)

to observe single recombination events in real time. When XerCD, FtsK, and ATP were incubated with surface immobilised DNA, XerC reversible synaptic complex formation could be observed as for the above experiment (i.e. in the absence of FtsK), presumably reflecting the behaviour of XerCD before FtsK had an effect on the complex. Other events were observed in which formation of a complex with $E^* = 0.55 \pm 0.01$, corresponding to the previously observed CS-Syn complex, was followed by a decrease in FRET whilst maintaining a narrow PSF width, the long lived nature of these low FRET, narrow PSF states were taken to be indicative of recombination. As noted above, treatment of the surface with SDS resulted in a majority of molecules with narrow PSF, demonstrating that this effect indeed occurs due to recombination rather than any extremely stable protein complex.

It is known from previous studies that, once activated by FtsK, XerCD proceeds through recombination from the DS-Syn complex (i.e XerD is initially active), followed by resolution of the HJ by XerC, figure 6.3 FtsK dependant box (Hallet et al., 1999; Aussel et al., 2002; Grainge et al., 2011). For the fluorophore labelling positions used in these experiments the FRET efficiency is not expected to be high during any of the intermediate complexes in recombination and from crystal structures of Cre I estimate distances of > 100 Åfor the DS-Syn and HJ and CP-Syn and HJ complexes, figure 6.3. Thus it is surprising to observe formation of a $E^* = 0.55$ state, presumably representing the CS-Syn complex, before further progress through recombination.

Plotting the FRET efficiency of these molecules after the disappearance of the $E^* = 0.55$ state,

gives a mean efficiency distinctly higher than that of the substrate DNA, table 6.1. Converting these to distance gives an inter-fluorophore separation of 123 Å, greater than that expected for any of the intermediate complexes during recombination, and likely arises from the single, doubly labelled, *dif* site after recombination. In a small number of traces a brief period of higher FRET can be observed immediately after the $E^* = 0.55$ state, figure 6.7, suggesting that the small residual FRET efficiency in either the DS-Syn/HJ or CP-Syn/HJ complexes can, in some cases, be distinguished. However the low FRET and transient nature of this state makes it difficult to capture. Plotting the FRET efficiency of only the first 5 frames after transition from the $E^* = 0.55$ state, gives a mean value of $E^* = 0.20$, equating to a distance of 105 Å, significantly higher than that measured long after recombination (table 6.1) and in relatively good agreement with distances predicted for the complexes along the FtsK dependant recombination pathway (DS-Syn/HJ, CP-Syn/HJ figure 6.3. Clearly this distances is at the limit of FRET sensitivity and thus distances calculated are highly sensitive to the accuracy of measured correction factors, making it difficult to unambiguously determine to which complex it corresponds.

	E^*	E^* SEM	Corrected FRET	Corrected FRET SEM	r (nm)	r SEM
Pre-recombination	0.16	0.009	0.00	0.182	-	-
Post-recombination	0.18	0.009	0.03	0.186	12.3	12.3
wt recombination SDS wash	0.17	0.007	-0.03	0.207	-	-
First 5 frames post-recombination	0.20	0.011	0.08	0.177	10.5	4.3
wt synaptic complex formation	0.55	0.013	0.72	0.049	6.0	0.5
wt recombination (high FRET)	0.55	0.017	0.72	0.045	6.0	0.4
XerC K172Q (major population)	0.21	0.014	0.08	0.227	10.6	5.8
XerD K172Q (major population)	0.52	0.020	0.68	0.067	6.2	0.5

Table 6.1: Uncorrected FRET (E^*), corrected FRET, and inter-fluorophore distance (r) obtained by fitting the different FRET populations observed in experiments, both the mean value and the standard error of the mean (SEM) are shown. Post-recombination complexes before and after SDS washing showed essentially no FRET. The first 5 frames of post-recombination events show a somewhat elevated FRET efficiency. The high FRET population observed during recombination is identical to that from molecule forming synaptic complexes. The major populations of the XerC K172Q and XerD K172Q mutants appear to reflect those seen for the first 5 frames post-recombination and the high FRET seen during recombination respectively. The slight reduction in FRET for the XerD K172Q mutant compared to the wt recombination population is likely due to the transient excursions to a lower FRET state biasing the mean FRET value in some camera frames to lower values.

These results suggest that, in the presence of FtsK and ATP, CS-Syn complex formation is followed by conversion to the DS-Syn complex[2], which then proceeds through recombination relatively rapidly, see figure 6.8 for a schematic illustration. My results show that all recombination events are preceded

[2]The reduction in FRET is not due to the cleavage and isomeristion from the CS-Syn/HJ complex (i.e. to form the DP-HJ/Syn figure 6.3), as this has been shown not to occur in ensemble studies (Grainge et al., 2011), and would also give rise to a higher FRET efficiency than that observed.

by formation of an CS-Syn complex, as those in which I do not observe FRET can be accounted for due to the time resolution of experiments (200 ms ALEX frame). In support of these conclusions, further work performed in the group, with an alternate arrangement of fluorophores, agree with the results presented here, see section 6.5.1 and figure 6.15.

6.3.3 XerCD mutants

To further investigate stages of the reaction pathway I used catalytic mutants of the XerC and XerD proteins to prevent the reaction from proceeding past known points. The mutants XerC K172Q and XerD K172Q are known to be unable to cleave DNA but able to participate in formation of synaptic complexes (Graham et al., 2010). When one of the wild type proteins is replaced by a K172Q mutant in the presence of FtsK and ATP the recombination reaction is prevented from forming HJs in the case of XerD K172Q, or prevented from resolving them for XerC K172Q, as illustrated in figures 6.10A and 6.9A.

Experiments were conducted with each mutant alone with the complementary wt protein (e.g. XerC K172Q and XerD wt), as well as both mutants together. Without FtsK or ATP, behaviour essentially matching that seen for the wt protein was observed, i.e. repeated reversible synaptic complex formation with $E^* = 0.55$. Experiments with both mutant proteins present also showed similar behaviour even in the presence of FtsK. In the presence of FtsK and ATP mixtures of wt and mutant proteins showed significantly different behaviour, as described below.

XerC K172Q

When activated by FtsK, a complex containing wt XerD and XerC K172Q is expected to be able to form a DS-Syn complex, perform cleavage and first strand exchange to form a DS-HJ, but should then be unable to resolve the HJ to the recombinant product. During single-molecule time-traces I observed two classes of events, those that resembled synaptic complex formation in the absence of FtsK and ATP, as above these are presumed to reflect the behaviour of the XerCD proteins before 'activation' by FtsK. The second class of events showed an initial excursion to a $E^* = 0.55$ state, followed by repeated transitions to a state with $E^* = 0.21$, during which the PSF width remained narrow, figure 6.9. These events were also seen to dissociate back to the substrate DNA (indicated by the increase in PSF), and this almost exclusively occurred via the $E^* = 0.55$ state.

Converting the E^* observed for the low FRET state to distance yields an interfluorophore

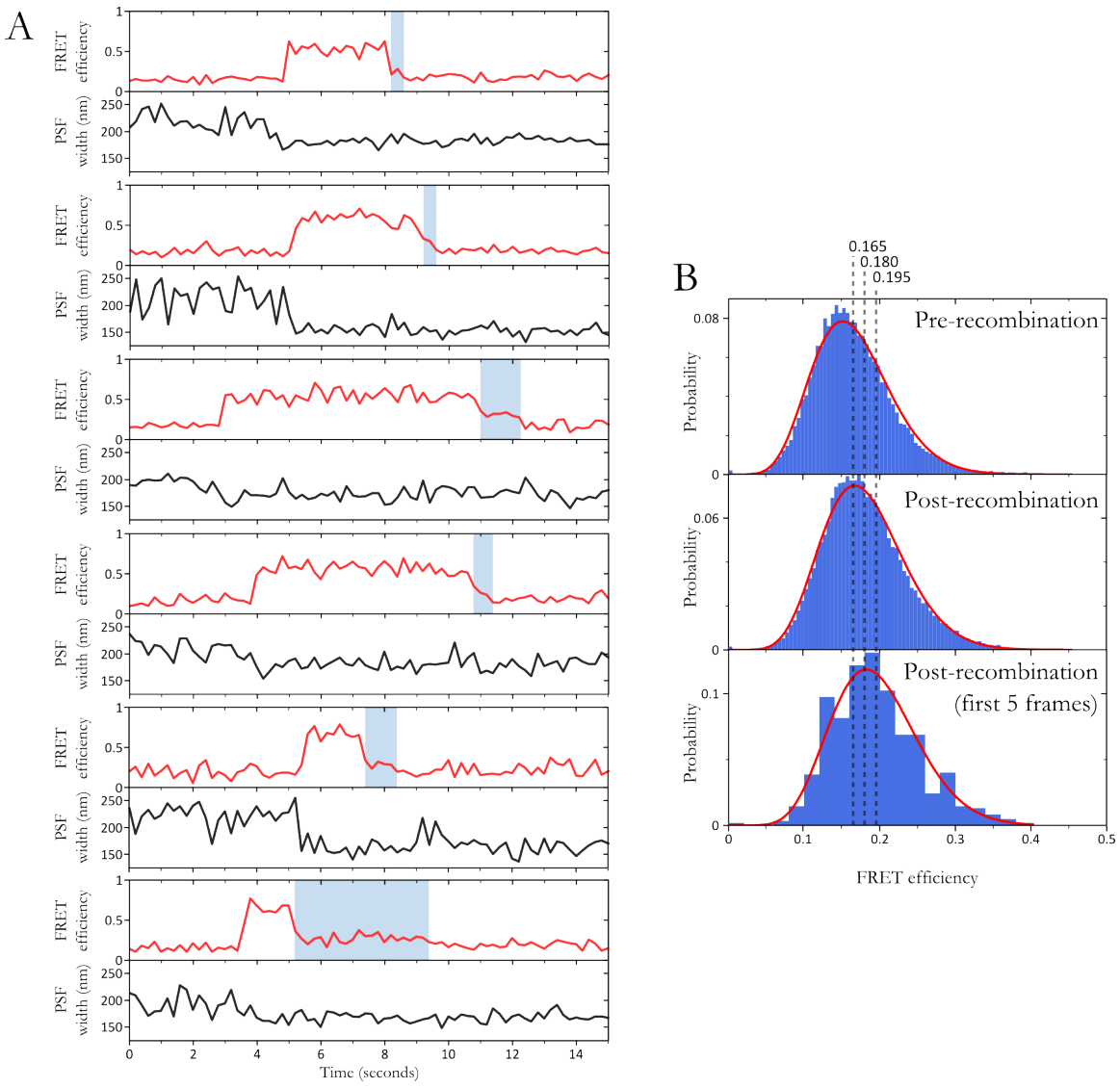


Figure 6.7: (A) Example FRET efficiency and PSF width time traces in which small periods of an intermediate FRET state can be observed between transition from the $E^* = 0.55$ to the post recombination state, highlighted in blue. (B) FRET efficiency histograms of the substrate, post recombination complex, and the first 5 frames after the $E^* = 0.55$ state. Histograms are fit with Beta distributions and the means are shown by grey dashed lines.

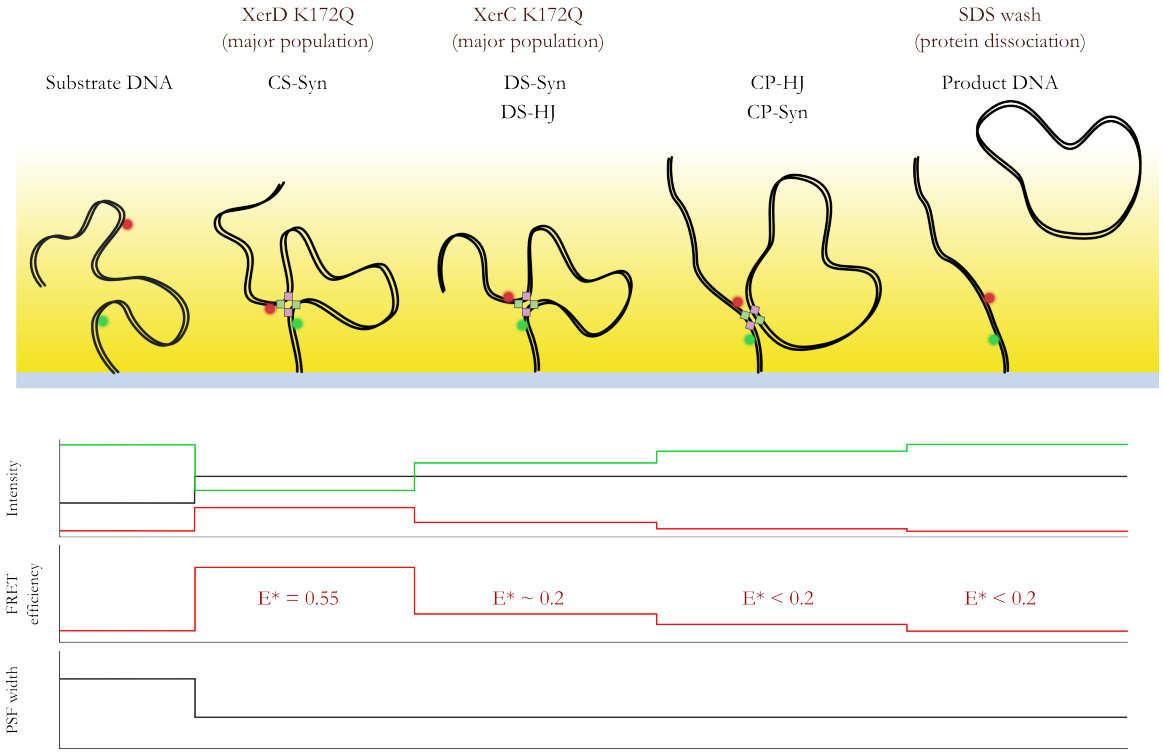


Figure 6.8: Schematic of a single DNA molecule undergoing FtsK dependant recombination and the expected single-molecule observables. Complex formation leads to a decrease in PSF, however in the 'XerD active' synaptic complex and Holliday junction the expected distance is 100 Å, figure 6.3, leading to a low FRET signal, which further decreases after isomerisation of the HJ to an 'XerC active' complex. After recombination, the unlabelled 1 kbp section of DNA between the fluorophores is excised and a permanent decrease in PSF width is observed.

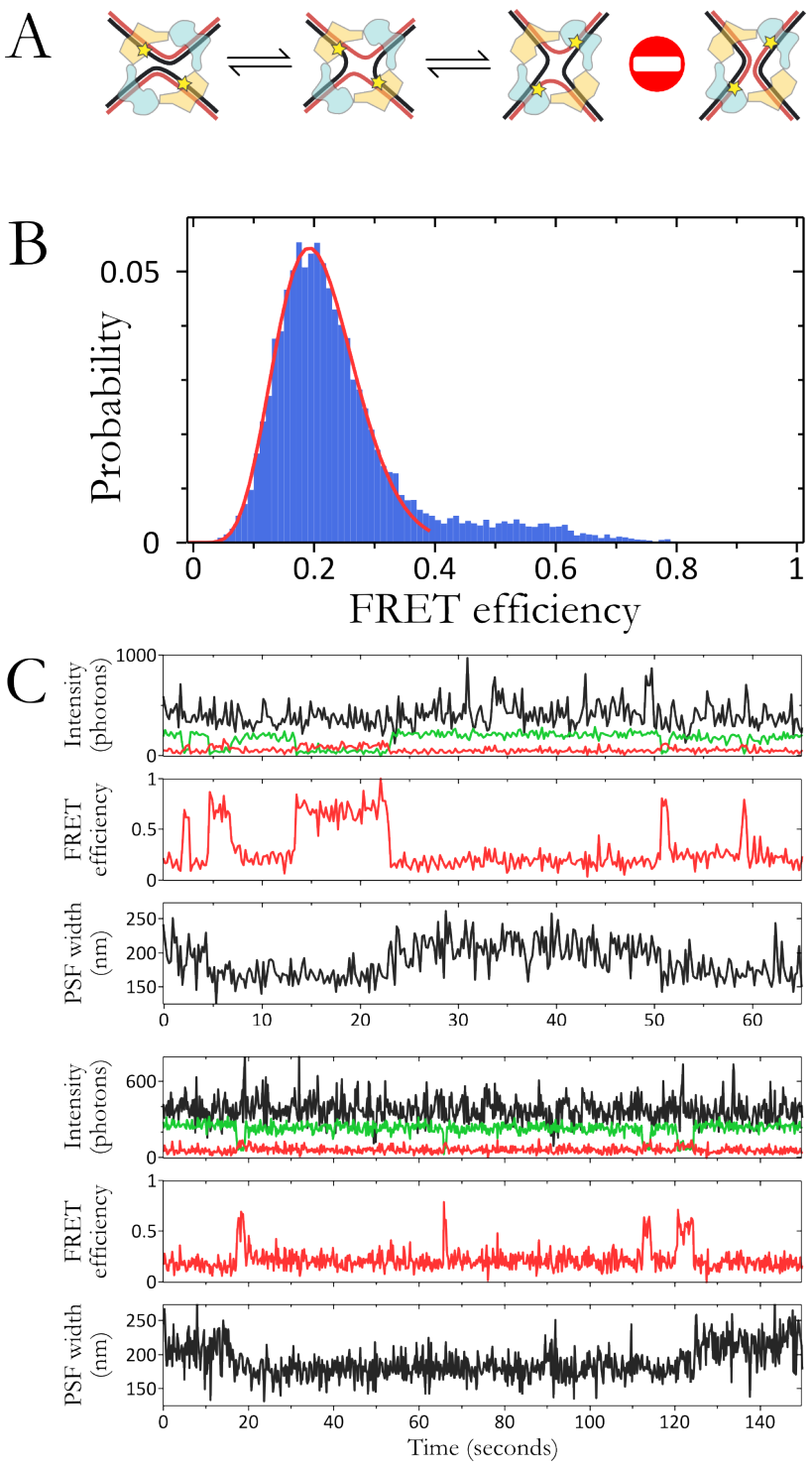


Figure 6.9: XerC K172Q result. (A) Schematic of where in the reaction pathway the mutant is unable to pass. (B) The FRET efficiency histogram of low PSF events shows a major peak at $E^* = 0.21$ with a minor population at higher FRET values (N = 82 molecules). A Beta distribution fit to $E^* < 0.4$ is shown in red. (C) Example single-molecule time traces show complex formation via the $E^* = 0.55$ state followed by reversible transition to the $E^* = 0.21$ state, return to substrate (indicated by an increase in PSF is almost always observed to occur from the $E^* = 0.55$ state.

separation of 106 ± 58 Å, agreeing with the distance expected for the DS-Syn/HJ and CP-Syn/HJ complexes, figure 6.3 FtsK dependant box.

XerD K172Q

I also performed similar experiments with wt XerC and XerD K172Q. In contrast to the above combination, this pair of proteins should be unable to form a DS-HJ by XerD cleavage, but would be expected to form DS-Syn complexes. Again, some events in single-molecule time traces showed FtsK mediated behaviour, whereupon formation of the $E^* = 0.55$ state was followed by repeated, rapid and transient excursions to a lower FRET state. These excursions occurred on a similar timescale to the camera frame time, so precisely determining the FRET value occupied is made challenging by the averaging effect of camera acquisition, however, longer dwells indicated that the FRET state occupied was the same as the low FRET state observed with the XerC K172Q mutant above.

Taken together, the above mutant data are consistent with the idea that both mutants are able to form DS-Syn complexes, as reflected in the $E^* = 0.21$ state. Formation of these synaptic complexes is dependant on the presence of FtsK and ATP and are seen to be transient from the XerD K172Q data, unless stabilised by formation of a DS-HJ when using XerC K172Q and wt XerD. A surprising observation with both these mutants is the ubiquity of the $E^* = 0.55$, representing the CS-Syn complex, which appears to be an obligatory step in forming the DS-Syn complex, or even in returning to the substrate DNA.

6.3.4 Looping of DNA by FtsK

Previous single-molecule studies of FtsK have observed looping of the DNA during FtsK translocation. The validity of these results has been questioned with suggestions that the observed looping arises due to protein aggregation or non-specific sticking to glass (or bead) surfaces (Crozat and Grainge, 2010). In this section I use TFM to confirm previous observations of ATP dependant looping of DNA by FtsK, and use TFM in combination with both protein induced fluorescent enhancement (PIFE) and fluorescently labelled FtsK to investigate the nature of the observed looping.

For the following experiments I used a 3 kbp DNA with a 5' attached biotin and an internally labelled Cy5, attached close to the surface-distal end. As introduced in chapter 4 the Cy5 acts as a reporter for the effective DNA length, which, in this case, undergoes length changes due to the looping of DNA by FtsK. It is important to note that for experiments with the unlabelled FtsK

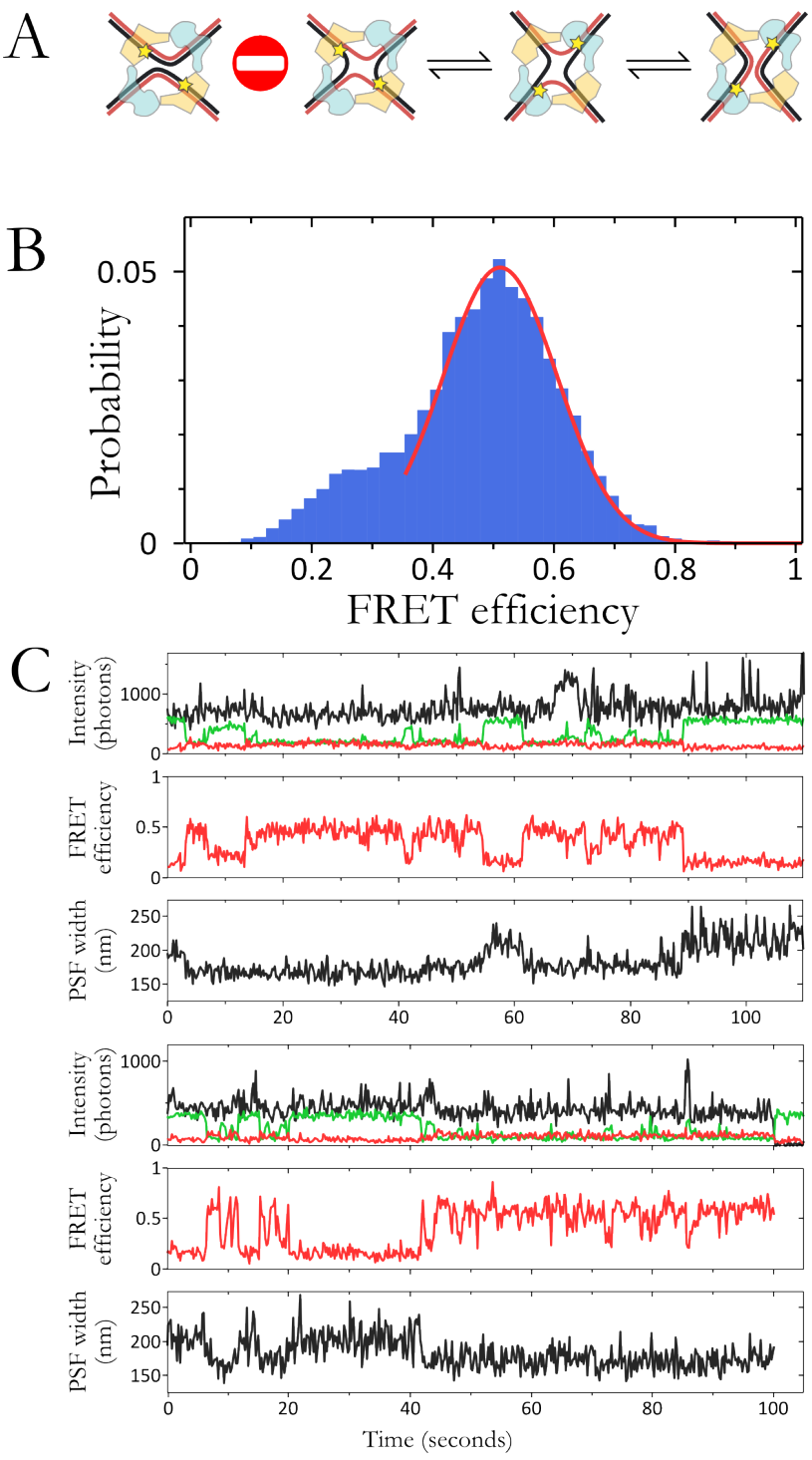


Figure 6.10: XerD K172Q result. (A) Schematic of where in the reaction pathway the mutant is unable to pass. (B) The FRET efficiency histrogram of low PSF events shows a major peak at $E^* = 0.55$ with a minor population at lower FRET values (N = 97 molecules). A Guassian fit to $E^* > 0.35$ is shown in red. (C) Example single-molecule time traces show formation of complexes with $E^* = 0.55$ during which repeated short lived excursions occur to a lower state.

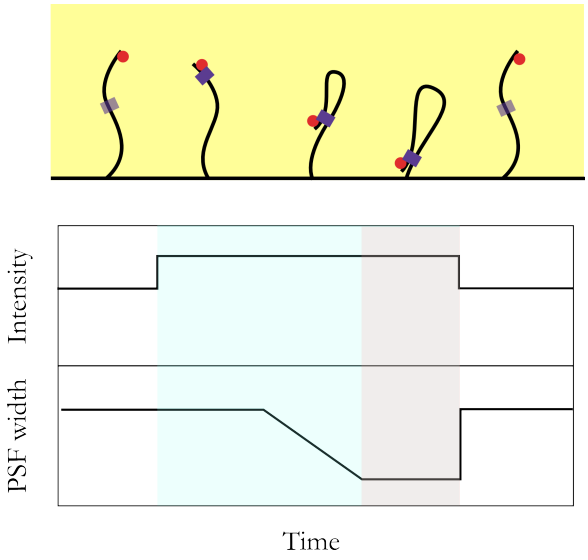


Figure 6.11: Schematic of unlabelled FtsK looping experiments, under the illumination conditions used, changes in DNA length do not lead to differences in fluorophore emission intensity. Changes in fluorophore intensity are only expected due to PIFE when FtsK is in close proximity to the fluorophore. The schematic time trace demonstrates one possible outcome of looping, FtsK binds the DNA, translocates to the Cy5 end, indicated by PIFE, translocation decreases the PSF width until the DNA is fully looped. Release of the DNA loop by FtsK produces an abrupt return to the original PSF width.

protein I adjusted the TIRF angle to maximise illumination of the 3 kbp substrate DNA. Using both long and short DNAs described in chapter 4, I confirmed that in this case there is no appreciable difference in intensity for different lengths of DNA, thus any changes in intensity of the acceptor corresponds to processes other than DNA length changes, such as PIFE.

At a frame time of 33 ms the 3 kbp substrate alone shows a PSF width of ~ 240 nm[3]. Addition of FtsK and ATP led to repeated, transient decreases in PSF width of the acceptor, indicating effective length changes of the DNA tether, figure 6.12, associated with FtsK mediated looping of the DNA. Addition of FtsK in the absence of ATP showed no change in PSF, confirming the ATP dependant nature of translocation and looping. In addition to excursions to low PSF states I also observed well defined step change increases in the observed intensity. Although all decreases in PSF width appeared to occur during increased Cy5 intensity, Cy5 intensity frequently increased without change in the PSF width. As described above, changes in fluorophore intensity are not due to differences in illumination arising from changes in DNA length under these conditions; along with the ATP and FtsK dependant nature of these intensity changes I suggest that they occur due to PIFE of the Cy5, caused by the translocation of FtsK to the fluorophore location. The FtsK crystal structure (figure 6.2C) suggests that FtsK tranlocates along DNA as a hexamer, with dsDNA passing through the central hole. This hole is wide enough to accommodate dsDNA, however the presence of a covalently attached fluorophore and linker to a DNA base would serve as a strong roadblock that FtsK would not be able to displace. The quantum yield of Cy5 is known to increase approximately

[3]This is less than that measured for the 3 kbp substrate in chapter 4 as here the 33 ms frame time is insufficient to fully average over all the accessible conformations of the DNA. The motion of the DNA can be observed as movement of the PSF centroid between subsequent frames. In order that this movement does not interfere with the width determination I utilise the free Gaussian PSF fitting available in TwoTone, see Experimental methods.

two-fold due to close proximity of proteins (Hwang et al., 2011), and assuming FtsK stalls at the fluorophore position such a PIFE effect would give rise to the observed increases in Cy5 fluorescence.

Rate of translocation

By using the combined method of TFM-PIFE I am able to measure both the rate and extent of loop extrusion as well as determine whether this originates at the surface-distal end, or otherwise. Previous studies have measured the translocation and looping rates for FtsK to be in the range 4-6 kbps^{-1} (Saleh et al., 2004; Pease et al., 2005; Ptacin et al., 2006; Lee et al., 2012a), using 33 ms frames I observed the progression of PSF width from $\sim$ 240 nm corresponding to the free 3 kbp DNA, to $\sim$ 150 nm, that of fully looped DNA. Measuring the start and end times for loop formation (or reversal) gives a mean rate of looping of 4.4 ± 0.1 kbps^{-1} (figure 6.13), in good agreement with previous studies under similar conditions.

Low PSF dwells occurred almost exclusively at $\sim$ 150 nm, corresponding to fully looped DNA, assuming FtsK binds non-specifically to this DNA substrate (which contains no KOPS) gives rise to two logical possibilities, either, FtsK is able to initiate looping from any location on the DNA, and thus must be able to translocate in both direction to form a full length loop, or looping is initiated from one end of the DNA molecule. Due to the unlabelled nature of the FtsK used here, the TFM-PIFE method only reports on the proximity of FtsK to the surface-distal DNA end or otherwise. My results show both cases in which FtsK initiates looping from the Cy5 DNA end (i.e. the PIFE is observed before the start of looping, figure 6.12B and those in which looping originates away from the Cy5, figure 6.12A. If looping were to originate from a non-specific location of the DNA, events in which this begins close enough to the Cy5 to induce PIFE should be relatively rare, however my results indicate that approximately 50% of events begin at the Cy5, suggesting that FtsK looping specifically originates from either DNA end.

Stoichiometery of the looping protein

There has previously been much discussion as to what the stoichiometery of the active FtsK motor is (Saleh et al., 2004; Pease et al., 2005; Massey et al., 2006). Although crystal structures suggest an elegant hexameric arrangement, the same study also demonstrated the formation of dimers of hexamers on DNA strands (Massey et al., 2006). Previous single-molecule studies of FtsK looping have variously suggested that loops are extruded by dimers of hexamers (Pease et al., 2005), or by

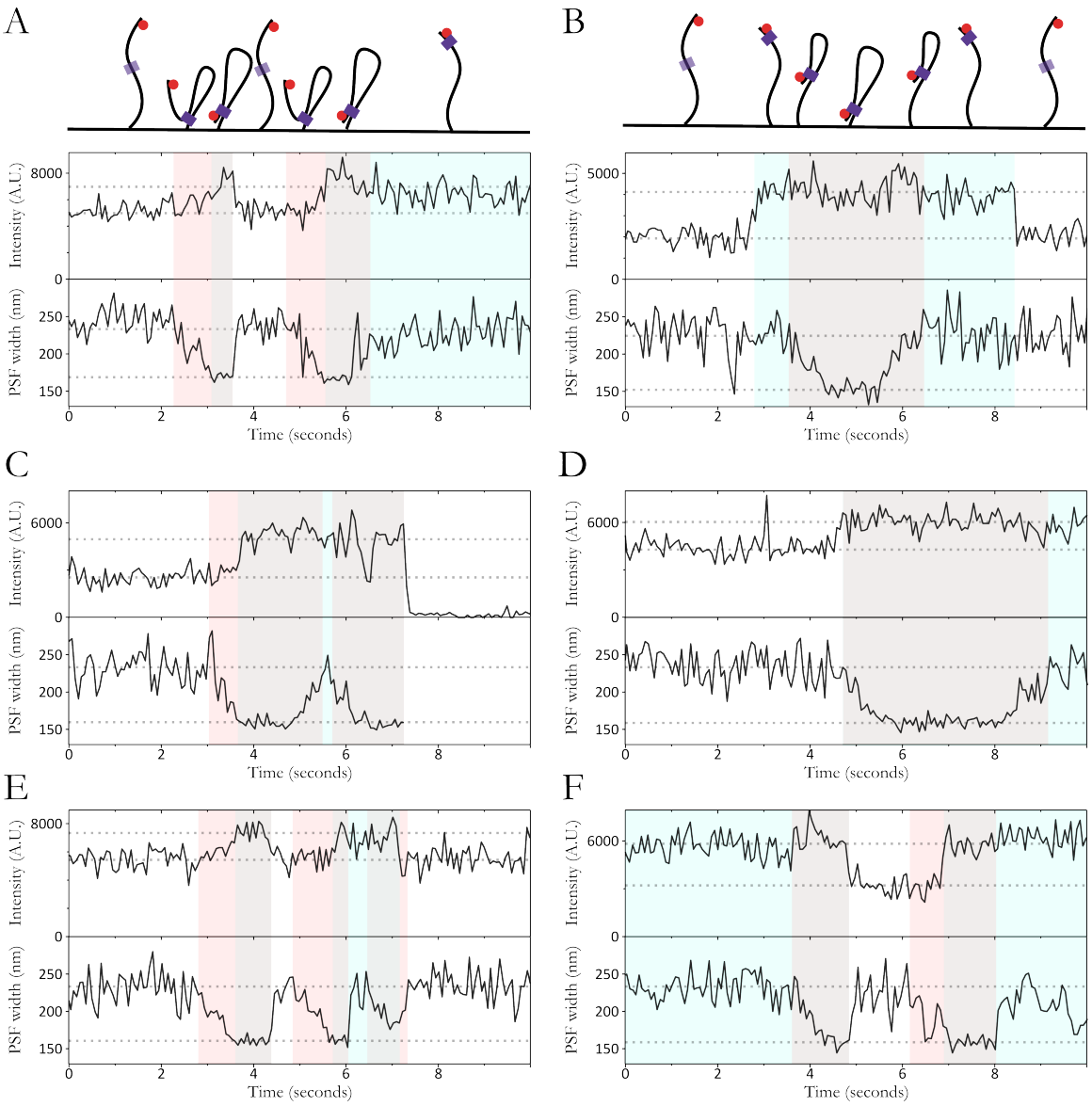


Figure 6.12: Example time-traces of unlabelled FtsK looping DNA. Increase in Cy5 intensity indicates presence of FtsK close to the fluorophore (highlighted in blue), periods of reduced PSF width are the result of DNA looping by FtsK (highlighted in red). A variety of looping outcomes in observed. (A) Trace in which looping is seen to originate away from the Cy5. (B) Trace in which FtsK reaches the Cy5 and subsequently commences DNA looping. (C-F) Further examples of DNA looping by FtsK.

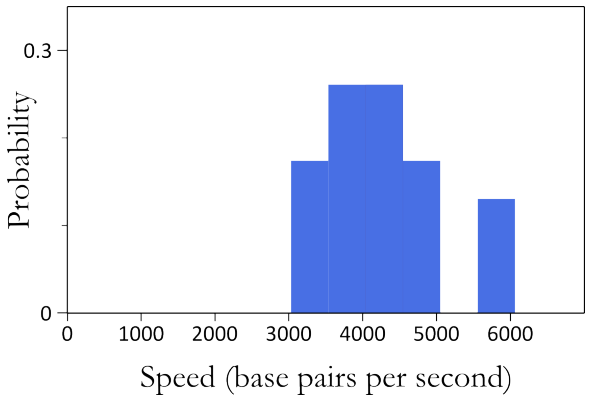


Figure 6.13: Rate of FtsK looping measured from 23 looping events. Only complete looping events from fully extended substrate to diffraction limited PSF (or vice-versa) are considered. These are assumed to correspond to 3000 bp of translocation thus the translocation speed can be calculated.

higher order complexes of FtsK (Saleh et al., 2004). These results have also been questioned due to the propensity of the FtsK 50C derivative (figure 6.2B) to aggregate, as well as the visualisation of such aggregates performing looping in experiments carried out by optical tweezers (Pease et al., 2005). In this section I use a fluorescently labelled FtsK to address the question of what the stoichiometry of the looping species is, by means of single-molecule fluorescence intensity measurements.

Cy3 labelled FtsK covalent trimers were produced by Pawel Zawadzki. Three surface accessible cysteine residues were introduced per covalent trimer (one for each FtsK monomer) and these were labelled with maleimide Cy3 fluorophores. The labelling efficiency was determined by absorption measurements to be approximately 0.6 per reactive cysteine (corresponding to, on average, 1.8 fluorophores per trimer). The activity of Cy3 labelled FtsK trimer (hereafter referred to as Cy3 FtsK) was determined by means of bulk biochemical plasmid resolution assays (performed by Pawel Zawadzki) and found to be comparable to that of the unlabelled FtsK.

Using the Cy3 FtsK, I performed essentially the same experiment as above, but using alternating laser excitation (ALEX) to observe emission from both the Cy3 FtsK, the Cy5 DNA, as well as any possible FRET that occurs between the two. An important difference in experimental conditions was that, due to the high concentration of labelled molecules in solution, the TIRF angle had to be adjusted, such that the penetration depth of the evanescent wave was as shallow as possible, to reduce the fluorescence background. This change in TIRF angle, as compared to the experiments with unlabelled FtsK above, meant that under these conditions I did expect a reduction of fluorescent intensity from fluorophores that were not close to the surface. The high background from Cy3 FtsK in solution, combined with the reduced intensity away from the surface made the sensitive measurement of PSF width, required for TFM, difficult for translocating Cy3 FtsK molecules.

Under the same looping conditions used for unlabelled FtsK experiments above, both DNA

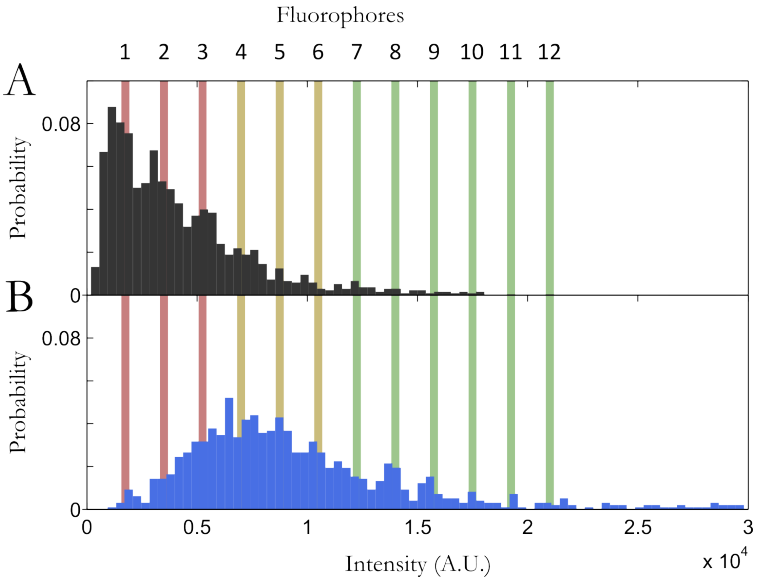


Figure 6.14: Comparison of intensities from surface immobilised and looping Cy3 FtsK. (A) Cy3 FtsK was immobilised using a biotinylated His-tag directed antibody. The histograms shows the frame by frame intensities of the first 10 frames from several movies. The intensity distribution does not follow that expected from the labelling efficiency due to photobleaching. (B) Frame by frame Cy3 FtsK intensity histogram during full looping events (i.e. Cy5 PSF width $\sim$ 150 nm).

binding and looping by Cy3 FtsK could be observed. Although binding of Cy3 FtsK could be observed without DNA looping, the intensity of the Cy3 signal was not a reliable indicator of the stoichiometry of the protein due to the variation in expected intensity with position along the DNA. During these Cy3 FtsK binding events I frequently observed PIFE of the Cy5 as well as FRET between the fluorophores, indicating that Cy3 FtsK was able to translocate toward the DNA end, stalling at Cy5, inducing PIFE and bringing the fluorophores within range of FRET[4].

To measure the stoichiometry of the looping protein I extract all Cy3 intensities that occur during a looping event, as indicated by the PSF width of Cy5, and compile these into a histogram, figure 6.14B. To compensate for any FRET that occurs in this looped state I use a value of $\gamma = 1.85$ for the Cy3-Cy5 pair (measured using a doubly labelled DNA standard) to calculate the total emission of Cy3 in the absence of Cy5,

$$F_{Dex} = F_{DD} + \frac{F_{DA}}{\gamma}. \tag{6.1}$$

To correlate this to the intensity corresponding to a single protein attached Cy3, I use the single-molecule pulldown technique (Jain et al., 2012). A biotinylated antibody, directed towards His-tags, is used to immobilise Cy3 FtsK molecules in the absence of DNA. Placing the observed intensities in a histogram shows clear peaks corresponding to the emission from integer numbers of Cy3 molecules attached to FtsK, figure 6.14A. These results indicate an expected intensitiy of $\sim$ 1800 counts for a single Cy3, this also corresponds well to intensities seen in the single-molecule time-traces, which can be confirmed to come from single Cy3 molecule due to single step photobleaching. The presence of

[4]This FRET efficiency cannot be used to measure any meaningful distance due to the multiple expected Cy3 fluorphores present on the FtsK molecule, as well as the uncertain position and environments of the dyes.

FtsK molecule containing intensities (and photobleaching steps) corresponding to more than three fluorophores also suggests that at least some fraction of FtsK trimers assemble into stable hexamers in solution.

The intensities observed from looping FtsK molecules indicate the presence of 3-6 fluorophores per looping molecule, suggesting that the looping species is a single hexamer of FtsK. The observed intensities clearly show that DNA looping by FtsK is not caused by protein aggregates, or higher order complexes of many FtsK hexamers as previously suggested (Saleh et al., 2004). Due to the labelling efficiency of FtsK, dimers of hexamers would be expected to show higher intensities, with on average 7.2 fluorophores per molecule, thus my results do not support the hypothesis that a dimer of hexamers is responsible for looping. There are a small number of molecules with intensities higher than 6 fluorophores, however these could be explained by the significant background in the Cy3 channel due to the high concentration of Cy3 FtsK in solution, or could also represent looping events in which further Cy3 FtsK molecule have associated elsewhere on the DNA.

6.4 Discussion

6.4.1 XerCD synaptic complex has a similar overall structure to Cre

In this chapter I have demonstrated the first direct observations of XerCD mediated synaptic complex formation. These complexes had not been observed by any previous means and my results confirm that CS-Syn complexes are formed frequently in solution, but never proceed to recombination in the absence of FtsK and ATP. The structure of these complexes appear to be almost identical to that observed for Cre recombinase, both by crystallographic methods as well as my own single-molecule FRET measurements in chapter 5. CS-Syn complexes appear to be the predominant species formed in solution, with DS-Syn complexes being formed very infrequently, if at all. Although I am unable to distinguish between CS-Syn and CS-HJ complexes my results indicate that the complex formed does not spend significant time in the post isomerisation DP-HJ, figure 6.3. Neither of these observations have been inferred from the extensive bulk biochemical studies of XerCD recombination, demonstrating the ability of single-molecule methods to obtain information about steps in the reaction pathway which induce changes that are exclusively conformational in nature. These observations also raise interesting questions about the nature of regulation of recombination by XerCD; what is the conformation of the unsynapsed XerCD-*dif* complex, and how is DS-Syn formation suppressed?

6.4.2 'XerC active' synaptic complexes are formed prior to transition to an 'XerD active' complex

Using TFM-FRET I am able to observe complete recombination reactions by XerCD in the presence of FtsK and ATP, these are the first single-molecule observations of such reactions and provide new insights into details of the FtsK activated recombination. Using wild-type proteins I observe that formation of CS-Syn complexes always precedes transition to the DS-Syn complex, followed by completion of recombination. Cleavage inactive mutant proteins show the expected transient formation of DS-Syn complexes, which are stabilised by formation of a DS-HJ when using the XerC K172Q mutant. These mutants further confirm that formation of the DS-Syn complex is preceded by initial formation of a CS-Syn complex, as is dissociation to the substrate DNA from the DS-Syn complex.

These data suggest that formation of the CS-Syn complex is an obligatory step in successful recombination. From existing crystal structures of Cre synaptic complexes, it is assumed that conversion from a CS-Syn to DS-Syn complex would require significant rearrangement of proteins and DNA, necessitating the breaking of large numbers of protein-protein interactions which maintain the complex. Thus it seems plausible that FtsK acts as a 'molecular remodelling' machine, perhaps using the energy of ATP hydrolysis to convert the CS-Syn to DS-Syn complex. The added complication of DNA looping by FtsK is also not addressed in these recombination experiments and it remains unclear how it may affect complex assembly.

6.4.3 DNA looping by FtsK is not caused by aggregates

Although it is known that FtsK acts as an ATP dependant translocase on double stranded DNA many of the details of the nature of this tranlocation remain unclear, spontaneous reversals of tranlocation direction have been observed on DNA curtains, and DNA loop extrusion (again with spontaneous reversals) has been observed with both optical and magnetic tweezers. These looping data have been questioned as to whether they originate due to aggregation of FtsK or non-specific sticking to the glass or bead surfaces. My observations once again demonstrate DNA loop extrusion by FtsK while translocating, and using fluorescently labelled FtsK I rule out the possibility that these arise due to the presence of large multi-protein aggregates.

It has also previously been suggested that FtsK translocates as a dimer of hexamers (Pease

et al., 2005), explaining both its ability to loop DNA as well as to undergo spontaneous reversals in direction, assuming the activity of the two hexamers is co-ordinated and in opposite directions. The co-ordinated action of two oppositely arranged FtsK motors would explain many of the results presented in this chapter, however, the use of Cy3 labelled FtsK, to observe the stoichiometry of FtsK during loop extrusion, indicates that the functional unit of looping is in fact a single hexamer. The use of Cy3 in this instance raises some questions as to how well the fluorescent intensity of immobilised FtsK can be assumed to correspond those involved in looping, as the quantum yield of Cy3 is known to be highly sensitive to the local environment of the fluorophore (Hwang et al., 2011). The photo-bleaching lifetime of Cy3 molecules on FtsK is also relatively short and as such the use of a dye less susceptible to PIFE and photo-bleaching, such as Cy3B, would give more confidence to the outcome of any future fluorescence intensity experiment.

6.4.4 Looping is initiated in response to stoppage of FtsK on DNA

If DNA looping were an artefact of non-specific sticking of FtsK, strong attachment to such diverse surfaces as, passivated glass, an internally labelled fluorophore, and magnetic and polystyrene beads, would all have to give rise to the same effect, which seems unlikely. The observation that looping is frequently initiated from the Cy5 DNA end, and results in complete looping of the DNA, suggests that these events are initiated exclusively from either extremity of the DNA molecule (i.e. the covalently attached Cy5, or the passivated slide surface). Thus I suggest that DNA looping occurs in response to stoppage of FtsK at a roadblock that it cannot pass. The biological significance of this looping (if any) remains unclear, however it is interesting to note that FtsK is able to displace, even strongly bound, roadblocks on double stranded DNA, but stops specifically at XerCD (Graham et al., 2010).

6.4.5 Proposed mechanism for looping

Taken together, the above results with both labelled and unlabelled FtsK, allows me to construct a hypothetical model for the looping process. FtsK is known to be able to translocate along DNA without looping (Pease et al. (2005); Lee et al. (2012a) and my own observations) thus I propose that FtsK binds non-specifically to the, KOPS lacking, 3 kbp substrate and commences translocation along the DNA. Upon reaching either the Cy5 fluorophore or the surface attached end of DNA, it is stalled and remains bound. In order for a single hexamer to extrude a loop of DNA, part of the

protein must remain bound to the DNA whilst the translocase domain is active. Thus stalling of FtsK at either roadblock must promote binding of part of FtsK to DNA, whilst the translocase domains are able to reverse their direction of operation, resulting in loop extrusion. Due to the relatively short nature of the DNA substrates used, FtsK generally translocates the full 3 kbp before stopping at the roadblock at the other end of DNA (although reversals before full looping are occasionally observed). At this point FtsK is able to either reverse translocation to reel in the loop, or dissociates from a single attachment point on the DNA, (or completely dissociates) to release the loop.

This model, although explaining the above data, is difficult to reconcile with existing knowledge of FtsK, although there exists a known DNA binding site separate to the translocase domain (γ domain), looping has been shown to occur in mutants lacking this domain ((Crozat et al., 2010) and my own preliminary data). Thus the single hexamer model presented above necessitates the existence of a second, strong, non-specific DNA binding domain, for which, currently, no candidates exist. It also raises the question of how FtsK is able to reverse translocation direction without releasing the loop, which would require a further DNA binding site, or the ability of a single hexamer motor to operate in both directions. It should be noted that a model in which a dimer of co-ordinated hexamers operates on DNA can also explain all the above data apart from apparent fluorescent intensity of the looping unit, to which strong caveats currently apply.

Although such a surprising mechanism begs further testing with complementary approaches, the data presented in this chapter clearly suggest an intriguing and complex model for loop formation by FtsK. It proposes that a single FtsK molecule can perform numerous functions: tranlocate along DNA, reverse direction without dissociating, bind to DNA during translocation to form loops, and activate XerCD recombination. How the *in vitro* looping caused by the FtsK derivatives studied here, relates to the *in vivo* functioning of wild type FtsK is not certain. It is possible that this effect plays no biological role, but the prevalence of looping in single-molecule studies of FtsK clearly merits further investigation as to what biological function this could perform; perhaps DNA looping plays an important role in bringing together distant *dif* sites and controlling the topology with which they recombine.

6.5 Future work

6.5.1 XerCD recombination

Since the completion of the work outlined in this chapter, further experiments have been performed on XerCD recombination, using alternative fluorophore positions, performed by Pawel Zawadzki and Peter May (Kapanidis group). This alternate arrangement of fluorophores produces higher FRET efficiencies and is arranged such that the two fluorophores are left on the non-surface attached DNA after recombination, thus simultaneous disappearance of both fluorophores occurs as the recombinant product synaptic complex dissociates. These experiments further support the conclusions of this chapter, and also demonstrates that unlike observations from Cre in chapter 5 the product of recombination is only stable for several seconds, figure 6.15.

Clearly one of the ultimate goals of studying XerCD recombination is to determine how FtsK interacts with the XerCD complex to activate recombination. Does FtsK act on a single *dif* site, or an already formed synaptic complex? Is the continued interaction of FtsK required to complete recombination or does it act as a molecular switch? Does DNA looping play a role in the mechanism of complex assembly or activation?

There could be a number of extensions to the work presented in this thesis to address these questions. Perhaps the most obvious is to use a three-colour experiment in combination with the present TFM-FRET assay to correlate the presence of a blue fluorescently labelled FtsK with the TFM-FRET observables during the reaction. Attempts I have made to perform such experiments have been hampered by the relative instability of current blue organic fluorophores, this creates difficulties in determining whether FtsK has dissociated, or merely photobleached. This issue could be addressed by utilising quantum dot labelled FtsK, as in Lee et al. (2012a), despite being much brighter and not susceptible to photobleaching, they undergo blinking and would make any potential FRET measurements difficult. An alternative approach might be to use the observed FtsK induced PIFE to probe the proximity of FtsK to Cy5, or a PIFE susceptible donor such as Cy3, but this technique inherently has an extremely short range and might perturb native action of FtsK by interaction with the fluorophore.

As described in the introduction, there exist other sites for recombination, which require their own set of accessory proteins for the reaction. Similar methods could be used to study those sites, although the extra requirement of a super-coiled substrate would add an extra level of complexity to

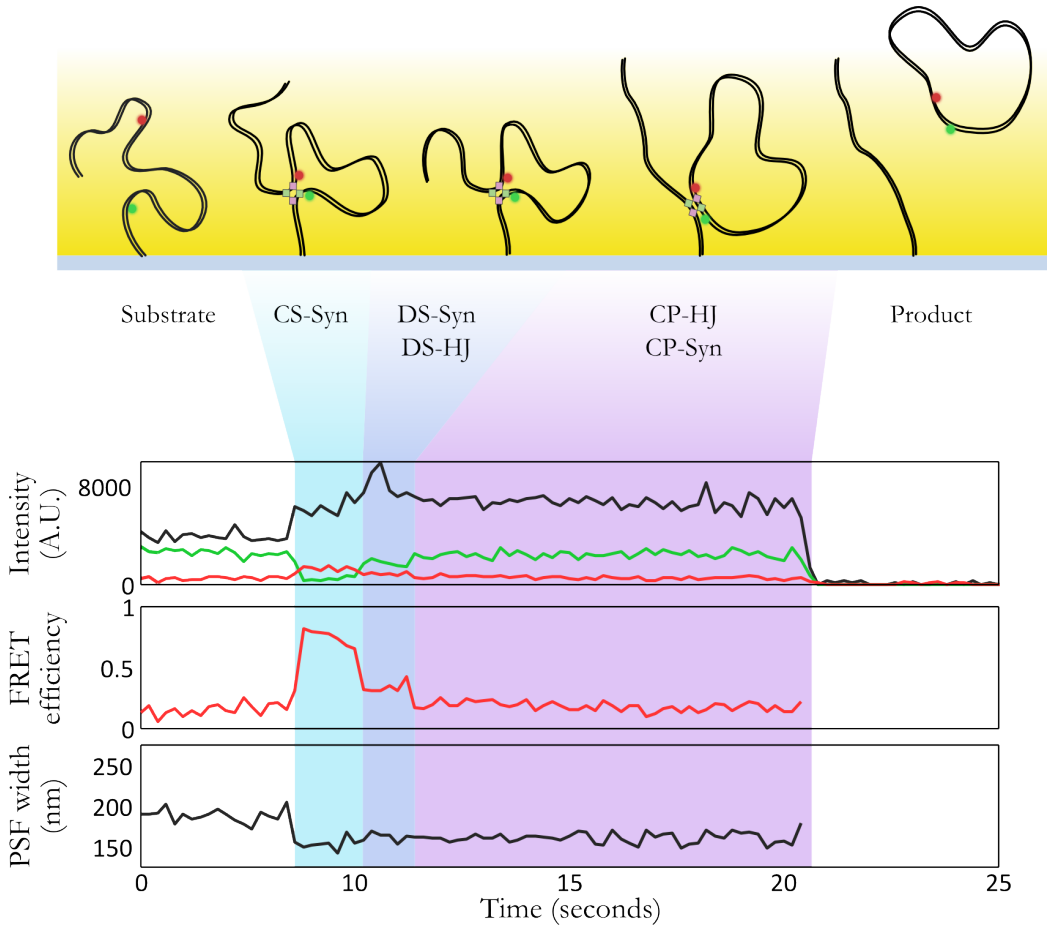


Figure 6.15: Schematic and example time trace of the behaviour seen with DNA substrates prepared since the completion of work in this thesis. Fluorophores are positioned such that release of the recombinant product DNA results in the simultaneous loss of both fluorophores and higher FRET throughout the reaction. These results concur with the conclusions of this chapter, that initial CS-Syn complex formation is followed by transition to a short lived DS-Syn/HJ complex, which then isomerises to the CP-Syn/HJ complex followed by subsequent dissociation of this product complex to release the excised DNA loop containing both fluorophores.

any such experiment.

6.5.2 FtsK function

Improvements and modifications to the experiments described in this chapter are currently being carried out in the group. Labelled FtsK constructs are being made with Cy3B rather than Cy3 to improve the fluorescent intensity measurements presented here. Efforts are being made to move to a system using covalent hexamers (rather than the trimers used here) as well as alternative labelling positions in an effort to address questions regarding the possible formation of dimers of heaxmers, as well as the direction of translocation.

Other approaches could include working at lower concentrations of labelled FtsK such that a deeper illumination could be used whilst keeping a relatively low background, so the location of FtsK along the DNA could be tracked by TFM. Alternatively using unlabelled FtsK a PIFE susceptible fluorophore could be placed at the surface-proximal DNA end, to determine whether DNA looping is indeed always initiated from one end. Experiments to determine whether looping can be initiated after stoppage at XerCD bound to a *dif* site would be an important step in demonstrating the biological relevance of the effect. Other TFM experiments could also be considered, such as the use of an immobilised FtsK and doubly labelled long DNAs, which may enable the simultaneous tracking of the position of both DNA ends.

Clearly many possible experiments could be designed to try and elucidate the evidently complex, and still largely unknown, action of DNA translocation and looping by FtsK, and I hope TFM will prove to be a useful tool in addressing such questions.

6.6 Contributions

DNA substrates (apart from Cy3B labelling step) and proteins were produced by Pawel Zawadzki (Oxford Biochemistry), who also assisted in performing repeats of the recombination experiments involving both wild-type and mutant proteins. David Sherratt and Lidia Arciszewska provided much useful advice on the design and interpretation of results.

Chapter 7

Conclusion

In this thesis I have presented two methods for extending conventional *in vitro* two-colour single-molecule FRET studies. Although providing many extra challenges the additional structural information afforded by three-colour FRET measurements are highly appealing for studying conformational dynamics of single proteins. The development and validation of a three-colour TIRF ALEX microscope in this thesis opens possibilities for three-colour FRET experiments to extend existing work in the group on DNA and RNA polymerases.

Tethered fluorophore motion (TFM) takes existing concepts from the well established tethered particle motion technique and applies them to the single-fluorophore level, providing the ability to measure effective length changes of DNA. In this thesis I develop and characterise this method, and demonstrate a proof-of-principle application to measure the rate of DNA replication by Klenow Fragment.

The primary advantage of TFM is its ability to be combined with other single-molecule fluorescence techniques; I used this capability to combine TFM and single-molecule FRET to study details of site-specific recombination by two tyrosine recombinases, Cre and XerCD.

Although recent experiments have shown Cre recombination at the single-molecule level, my use of TFM-FRET provides extra information on the structure of intermediate complexes formed. My results clarify existing data on the nature and stability of synaptic complexes, as well as uncovering previously unobserved structural and molecular heterogeneity. My experiments on XerCD recombination represent the first single-molecule observations of this complex, and highly regulated recombination system. My results capture formation of synaptic complexes that were previously only indirectly inferred from ensemble assays. I also show the FtsK activated recombination reaction,

and use catalytic mutants to demonstrate that this reaction likely proceeds via initial formation of the non-productive CS-Syn complex, which must be transformed into the active DS-Syn complex, presumably by the action of FtsK.

The nature of activation, and the mode of translocation of FtsK are still the subject of many open questions, to this end I used TFM to address questions regarding the looping of DNA by FtsK previously observed. The combination of TFM, both with protein induced fluorescence enhancement as well as single-molecule fluorescence intensity measurements allowed me to determine that FtsK looping is predominantly initiated from stoppage at a roadblock on DNA, and that this is not due to protein aggregates, but is instead likely to be due to the action of a single FtsK hexamer. Clearly the complex nature of FtsK translocation, as well as its interaction with XerCD to initiate recombination still begs further study, and new avenues and extensions of the work described in this thesis are suggested in section 6.5.1.

Although it suffers from shorter observation times than TPM, and lower resolution than (optical and magnetic) tweezers, the ability of TFM to be combined with a variety of single-molecule fluorescence techniques, as demonstrated throughout this thesis, make it a useful and flexible technique for obtaining a more 'holistic' view of protein action on DNA. It will continue to be used to address mechanistic questions regarding site-specific recombinases within the group and I hope that TFM (and its combination with other methods) can become a useful and widely applied addition to the single-molecule toolbox.

Appendix A

Corrected FRET and distance calculation

In Chapters 5 and 6 I convert from uncorrected FRET efficiencies, E^*, to distances using previously described methods, which I outline in brief here.

A.1 FRET correction

Following Kapanidis et al. (2004) I use the virtual molecular sorting abilities of ALEX to isolate molecules which contain only a single fluorophore in order to determine the required spectral cross-talk factors. To account for imperfect separation of donor and acceptor excitation/emission two effects must be taken into account: donor leakage into the acceptor emission channel, l, and direct excitation of the acceptor by the 532nm laser, d, these can be calculated as follows

$$l = \frac{F_{DA}}{F_{DD}} \text{ donor only molecules} \tag{A.1}$$

$$d = \frac{F_{DA}}{F_{AA}} \text{ acceptor only molecules,} \tag{A.2}$$

using the notation described in chapter 2. Using these correction factors, the uncorrected FRET efficiency (E^*) and stoichiometry (S^*) I calculate a cross-talk corrected 'proximity ratio', E_{PR},

$$E_{PR} = \frac{E^*S^* - d(1-S^*) - lS^*(1-E^*)}{E^*S^* + S^*(1-E^*) - d(1-S^*)}. \tag{A.3}$$

Apart from spectral cross-talk the quantum yields and relative detection efficiencies of the two

fluorophores need to be taken into account via the detection-correction factor, γ, where

$$\gamma = \frac{Q_A \eta_A}{Q_D \eta_D}. \tag{A.4}$$

Where Q and η are the quantum yield and detection efficiency of the fluorophores. As this is a ratiometric factor relating to the relative ability to detect photons from donor and acceptor, absolute determination of Q and η is not required. Instead I use the method introduced by Ha et al. (1999) to calculate γ from acceptor photobleaching events. Single-molecule traces in which the acceptor photobleaches during a FRET event are identified, the corresponding change in intensity of F_{DA} and F_{DD} can be used to calculate γ as follows,

$$\gamma = \frac{I_{DA}\ (\text{pre bleach}) - I_{DA}\ (\text{post bleach})}{I_{DD}\ (\text{post bleach}) - I_{DD}\ (\text{pre bleach})}. \tag{A.5}$$

Where the I represents the fluorescence intensity after correction for spectral cross-talk,

$$I_{DA} = (1 - lF_{DD} - dF_{AA}) \tag{A.6}$$

$$I_{DD} = (1 + lF_{DD}). \tag{A.7}$$

This procedure is performed for many (> 30) molecules to obtain an estimate for the mean value of γ which is then used as a correction factor for all molecules to obtain a corrected FRET efficiency given by

$$E_{Corrected} = \frac{E_{PR}}{\gamma - (\gamma - 1)E_{PR}}; \tag{A.8}$$

A.2 Distance determination

From the above corrected FRET efficiency, the inter-fluorophore distance, R, can be calculated via

$$E_{Corrected} = \frac{1}{1 + \left(\frac{r}{R_0}\right)^6}, \tag{A.9}$$

$$R_0 = 0.211 \left(\kappa^2 n^{-4} Q_D J\right)^{\frac{1}{6}} \tag{A.10}$$

$$J = \int f_D(\lambda)\eta_A(\lambda)\lambda^4 d\lambda. \tag{A.11}$$

Where n, the refractive index of the medium is 1.33, f_D is the normalised fluorescence emission of the donor, and ϵ_A is the extinction coefficient of the acceptor. I assume a literature value of 0.67 for Q_D, the quantum yield of Cy3B (Cooper et al., 2004), and that the dipole orientation factor, $\kappa^2 = 2/3$. To justify my assumption of rotational freedom of the fluorophores attached to DNA I measure their anisotropy values in conditions closely resembling single-molecule measurements, and use these to obtain an estimate for the possible range of true values for κ^2 (Dale et al., 1979). Emission spectra of Cy3B and absorption spectra of Cy5 were measured on a fluorometer (Photon Technology International) and used to calculate the spectral overlap J. Fluorometer measurements were carried out in a 50 µl cuvette with 10 nM of the DNA substrate, $>$ 50 nM protein, and the same buffers and oxygen scavenging system used in single molecule experiments.The correction factors and constants measured for use in this thesis are given in table A.1.

	Cre		Xer	
	mean	std	mean	std
d	0.057	0.051	0.037	0.04
l	0.075	0.054	0.165	0.05
gamma	0.28	0.05	0.3	0.04
Q	0.67	0.07	0.67	0.07
J	1.00E+16	1.00E+14	9.26E+15	9.26E+13
kappa	0.666667	0.24	0.666667	0.25

Table A.1: Parameters used to convert from E^* to inter-fluorophore distance.

Using the standard error of the measured E^* and S^* values, as well as estimates for the errors on the measurements of the parameter values described above, I perform standard propagation of errors to compute the standard error on the mean (SEM) for all distance measurements, as outlined in Uphoff et al. (2010).

Appendix B

Publications and presentations arising from this thesis

B.1 Publications

- Pinkney, J.N.M., Zawadzki, P., Mazuryk, J., Arciszewska, L.K., Sherratt, D.J., Kapanidis, A.N. (2012) Capturing reaction paths and intermediates in Cre-*loxP* recombination using single-molecule fluorescence. *Proceedings of the National Academy of Sciences of the United States of America*:20871-20876.

- Periz, J., Celma, C., Pinkney, J.N.M., Roy, P., Kapanidis, A.N., (2012) Rotavirus mRNAs are released by specialized channels in the double-layered viral capsid. *Proceedings of the National Academy of Sciences of the United States of America. Submitted.*

- Pinkney, J.N.M, Zawadazki, P., Evans, G.W., Sherratt, D.J, Kapanidis, A.N. (2012) Tethered fluorophore motion: correlating large-scale DNA conformational changes with single-molecule fluorescence observations. *Biophysical Journal. In preparation*

B.2 Presentations

- 2012 - 73rd Harden Conference - Machines on genes II, Oxford, UK

- 2012 - Biophysical Society 56th Annual Meeting, San Diego, California, USA

- 2012 - Biological Physics Seminar, Oxford, UK

- 2011 - Nano Meets Spectroscopy, National Physical Laboratory, London, UK
- 2011 - PicoQuant 17th International Workshop, Berlin, Germany *(Awarded best student talk)*

Bibliography

Aitken, C. E., Marshall, R. A., and Puglisi, J. D. (2008). An oxygen scavenging system for improvement of dye stability in single-molecule fluorescence experiments. *Biophysical journal*, 94(5):1826–35.

Alberts, B. (2008). *Molecular biology of the cell: Reference edition*. Garland Science.

Arciszewska, L. K., Grainge, I., and Sherratt, D. J. (1997). Action of site-specific recombinases XerC and XerD on tethered Holliday junctions. *The EMBO journal*, 16(12):3731–43.

Aussel, L., Barre, F.-X., Aroyo, M., Stasiak, A., Stasiak, A. Z., and Sherratt, D. (2002). FtsK Is a DNA Motor Protein that Activates Chromosome Dimer Resolution by Switching the Catalytic State of the XerC and XerD Recombinases. *Cell*, 108(2):195–205.

Axelrod, D. (1981). Cell-substrate contacts illuminated by total internal reflection fluorescence. *Journal of Cell Biology*, 89(1):141–145.

Axelrod, D., Burghardt, T. P., and Thompson, N. L. (1984). Total internal reflection fluorescence. *Annual review of biophysics and bioengineering*, 13:247–68.

Bai, H., Kath, J. E., Zörgiebel, F. M., Sun, M., Ghosh, P., Hatfull, G. F., Grindley, N. D. F., and Marko, J. F. (2012). Remote control of DNA-acting enzymes by varying the Brownian dynamics of a distant DNA end. *Proceedings of the National Academy of Sciences of the United States of America*, 109(41):16546–16551.

Bai, H., Sun, M., Ghosh, P., Hatfull, G. F., Grindley, N. D. F., and Marko, J. F. (2011). Single-molecule analysis reveals the molecular bearing mechanism of DNA strand exchange by a serine recombinase. *Proceedings of the National Academy of Sciences of the United States of America*, 108(18):7419–24.

Beausang, J. F., Zurla, C., Manzo, C., Dunlap, D., Finzi, L., and Nelson, P. C. (2007). DNA looping kinetics analyzed using diffusive hidden Markov model. *Biophysical journal*, 92(8):L64–6.

Bigot, S., Saleh, O. A., Cornet, F., Allemand, J.-F., and Barre, F.-X. (2006). Oriented loading of FtsK on KOPS. *Nature structural & molecular biology*, 13(11):1026–8.

Bigot, S., Saleh, O. A., Lesterlin, C., Pages, C., El Karoui, M., Dennis, C., Grigoriev, M., Allemand, J.-F., Barre, F.-X., and Cornet, F. (2005). KOPS: DNA motifs that control E. coli chromosome segregation by orienting the FtsK translocase. *The EMBO journal*, 24(21):3770–80.

Blakely, G., May, G., McCulloch, R., Arciszewska, L. K., Burke, M., Lovett, S. T., and Sherratt, D. J. (1993). Two related recombinases are required for site-specific recombination at dif and cer in E. coli K12. *Cell*, 75(2):351–361.

Blakely, G. and Sherratt, D. (1996). Determinants of selectivity in Xer site-specific recombination. *Genes & Development*, 10(6):762–773.

Borisenko, V., Lougheed, T., Hesse, J., Füreder-Kitzmüller, E., Fertig, N., Behrends, J. C., Woolley, G. A., and Schütz, G. J. (2003). Simultaneous optical and electrical recording of single gramicidin channels. *Biophysical journal*, 84(1):612–22.

Brinkers, S., Dietrich, H. R. C., de Groote, F. H., Young, I. T., and Rieger, B. (2009). The persistence length of double stranded DNA determined using dark field tethered particle motion. *The Journal of chemical physics*, 130(21):215105.

Chivers, C. E., Crozat, E., Chu, C., Moy, V. T., Sherratt, D. J., and Howarth, M. (2010). A streptavidin variant with slower biotin dissociation and increased mechanostability. *Nature methods*, 7(5):391–3.

Chung, H. S., Louis, J. M., and Eaton, W. A. (2010). Distinguishing between protein dynamics and dye photophysics in single-molecule FRET experiments. *Biophysical journal*, 98(4):696–706.

Colloms, S. D., Bath, J., and Sherratt, D. J. (1997). Topological Selectivity in Xer Site-Specific Recombination. *Cell*, 88(6):855–864.

Comstock, M. J., Ha, T., and Chemla, Y. R. (2011). Ultrahigh-resolution optical trap with single-fluorophore sensitivity. *Nature methods*, 8(4):335–40.

Cooper, M., Ebner, A., Briggs, M., Burrows, M., Gardner, N., Richardson, R., and West, R. (2004). Cy3B: Improving the Performance of Cyanine Dyes. *Journal of Fluorescence*, 14(2):145–150.

Cordes, T., Strackharn, M., Stahl, S. W., Summerer, W., Steinhauer, C., Forthmann, C., Puchner, E. M., Vogelsang, J., Gaub, H. E., and Tinnefeld, P. (2010). Resolving single-molecule assembled patterns with superresolution blink-microscopy. *Nano letters*, 10(2):645–51.

Cordes, T., Vogelsang, J., and Tinnefeld, P. (2009). On the mechanism of Trolox as antiblinking and antibleaching reagent. *Journal of the American Chemical Society*, 131(14):5018–9.

Crozat, E. and Grainge, I. (2010). FtsK DNA translocase: the fast motor that knows where it's going. *Chembiochem : a European journal of chemical biology*, 11(16):2232–43.

Crozat, E., Meglio, A., Allemand, J.-F., Chivers, C. E., Howarth, M., Vénien-Bryan, C., Grainge, I., and Sherratt, D. J. (2010). Separating speed and ability to displace roadblocks during DNA translocation by FtsK. *The EMBO journal*, 29(8):1423–33.

Dale, R. E., Eisinger, J., and Blumberg, W. E. (1979). The orientational freedom of molecular probes. The orientation factor in intramolecular energy transfer. *Biophysical Journal*, 26(2):161–93.

del Rio, A., Perez-Jimenez, R., Liu, R., Roca-Cusachs, P., Fernandez, J. M., and Sheetz, M. P. (2009). Stretching single talin rod molecules activates vinculin binding. *Science*, 323(5914):638–41.

Derocco, V., Anderson, T., Piehler, J., Erie, D., and Weninger, K. (2010). Four-color single-molecule fluorescence with noncovalent dye labeling to monitor dynamic multimolecular complexes. *BioTechniques*, 49(5):807–16.

Di Fiori, N. and Meller, A. (2010). The Effect of dye-dye interactions on the spatial resolution of single-molecule FRET measurements in nucleic acids. *Biophysical journal*, 98(10):2265–72.

Doi, M. and Edwards, S. (1988). *The theory of polymer dynamics*. Oxford University Press.

Ennifar, E., Meyer, J. E. W., Buchholz, F., Stewart, A. F., and Suck, D. (2003). Crystal structure of a wild-type Cre recombinase-loxP synapse reveals a novel spacer conformation suggesting an alternative mechanism for DNA cleavage activation. *Nucleic Acids Research*, 31(18):5449–60.

Fan, H.-F. (2012). Real-time single-molecule tethered particle motion experiments reveal the kinetics and mechanisms of Cre-mediated site-specific recombination. *Nucleic Acids Research*.

Ferreira, H., Butler-Cole, B., Burgin, A., Baker, R., Sherratt, D. J., and Arciszewska, L. K. (2003). Functional Analysis of the C-terminal Domains of the Site-specific Recombinases XerC and XerD. *Journal of Molecular Biology*, 330(1):15–27.

Finzi, L. and Gelles, J. (1995). Measurement of lactose repressor-mediated loop formation and breakdown in single DNA molecules. *Science*, 267(5196):378–80.

Forget, A. L. and Kowalczykowski, S. C. (2012). Single-molecule imaging of DNA pairing by RecA reveals a three-dimensional homology search. *Nature*, 482(7385):423–7.

Friedman, L. J., Chung, J., and Gelles, J. (2006). Viewing dynamic assembly of molecular complexes by multi-wavelength single-molecule fluorescence. *Biophysical journal*, 91(3):1023–31.

Friedman, L. J. and Gelles, J. (2012). Mechanism of transcription initiation at an activator-dependent promoter defined by single-molecule observation. *Cell*, 148(4):679–89.

Funatsu, T., Harada, Y., Tokunaga, M., Saito, K., and Yanagida, T. (1994). Imaging of single fluorescent molecules and individual ATP turnovers by single myosin molecules in aqueous solution. *Nature*, 374(6522):555–9.

Gell, C., Brockwell, D., and Smith, A. (2006). *Handbook of Single Molecule Fluorescence Spectroscopy*. Oxford University Press.

Ghosh, K., Guo, F., and Van Duyne, G. D. (2007). Synapsis of loxP sites by Cre recombinase. *Journal of Biological Chemistry*, 282(33):24004–16.

Ghosh, K., Lau, C. K., Gupta, K., and Van Duyne, G. D. (2005). Preferential synapsis of loxP sites drives ordered strand exchange in Cre-loxP site-specific recombination. *Nature Chemical Biology*, 1(5):275–82.

Ghosh, K. and Van Duyne, G. D. (2002). Cre loxP biochemistry. *Methods*, 28(3):374–383.

Gopaul, D. N., Guo, F., and Van Duyne, G. D. (1998). Structure of the Holliday junction intermediate in Cre-loxP site-specific recombination. *EMBO journal*, 17(14):4175–87.

Gopich, I. and Szabo, A. (2005). Theory of photon statistics in single-molecule Förster resonance energy transfer. *The Journal of chemical physics*, 122(1):14707.

Graham, J. E., Sivanathan, V., Sherratt, D. J., and Arciszewska, L. K. (2010). FtsK translocation on DNA stops at XerCD-dif. *Nucleic acids research*, 38(1):72–81.

Grainge, I., Lesterlin, C., and Sherratt, D. J. (2011). Activation of XerCD-dif recombination by the FtsK DNA translocase. *Nucleic Acids Research*, 39(12):5140–8.

Grindley, N. D. F., Whiteson, K. L., and Rice, P. A. (2006). Mechanisms of site-specific recombination. *Annual Review of Biochemistry*, 75:567–605.

Ha, T., Ting, A. Y., Liang, J., Caldwell, W. B., Deniz, A. A., Chemla, D. S., Schultz, P. G., and Weiss, S. (1999). Single-molecule fluorescence spectroscopy of enzyme conformational dynamics and cleavage mechanism. *Proceedings of the National Academy of Sciences of the United States of America*, 96(3):893–8.

Hagerman, P. J. and Zimm, B. H. (1981). Monte Carlo approach to the analysis of the rotational diffusion of wormlike chains. *Biopolymers*, 20(7):1481–1502.

Hallet, B., Arciszewska, L. K., and Sherratt, D. J. (1999). Reciprocal control of catalysis by the tyrosine recombinases XerC and XerD: an enzymatic switch in site-specific recombination. *Molecular cell*, 4(6):949–59.

Hamilton, D. L. and Abremski, K. (1984). Site-specific recombination by the bacteriophage P1 lox-Cre system. *Journal of Molecular Biology*, 178(2):481–486.

Harms, G. S., Orr, G., Montal, M., Thrall, B. D., Colson, S. D., and Lu, H. P. (2003). Probing conformational changes of gramicidin ion channels by single-molecule patch-clamp fluorescence microscopy. *Biophysical journal*, 85(3):1826–38.

Harriss, L. M., Cronin, B., Thompson, J. R., and Wallace, M. I. (2011). Imaging multiple conductance states in an alamethicin pore. *Journal of the American Chemical Society*, 133(37):14507–9.

He, Y., Lu, M., Cao, J., and Lu, H. P. (2012). Manipulating protein conformations by single-molecule AFM-FRET nanoscopy. *ACS nano*, 6(2):1221–9.

Heilemann, M., Margeat, E., Kasper, R., Sauer, M., and Tinnefeld, P. (2005). Carbocyanine dyes as efficient reversible single-molecule optical switch. *Journal of the American Chemical Society*, 127(11):3801–6.

Heron, A. J., Thompson, J. R., Cronin, B., Bayley, H., and Wallace, M. I. (2009). Simultaneous measurement of ionic current and fluorescence from single protein pores. *Journal of the American Chemical Society*, 131(5):1652–3.

Hoess, R., Wierzbicki, A., and Abremski, K. (1987). Isolation and Characterization of Intermediates in Site-Specific Recombination. *Proceedings of the National Academy of Sciences of the United States of America*, 84(19):6840–6844.

Hohlbein, J., Gryte, K., Heilemann, M., and Kapanidis, A. N. (2010). Surfing on a new wave of single-molecule fluorescence methods. *Physical biology*, 7(3):031001.

Hohng, S., Joo, C., and Ha, T. (2004). Single-molecule three-color FRET. *Biophysical journal*, 87(2):1328–37.

Hohng, S., Zhou, R., Nahas, M. K., Yu, J., Schulten, K., Lilley, D. M. J., and Ha, T. (2007). Fluorescence-force spectroscopy maps two-dimensional reaction landscape of the holliday junction. *Science*, 318(5848):279–83.

Holden, S. J., Uphoff, S., Hohlbein, J., Yadin, D., Le Reste, L., Britton, O. J., and Kapanidis, A. N. (2010). Defining the Limits of Single-Molecule FRET Resolution in TIRF Microscopy. *Biophysical Journal*, 99(9):3102–11.

Holden, S. J., Uphoff, S., and Kapanidis, A. N. (2011). DAOSTORM: an algorithm for high- density super-resolution microscopy. *Nature methods*, 8(4):279–80.

Huang, B., Bates, M., and Zhuang, X. (2009). Super-resolution fluorescence microscopy. *Annual review of biochemistry*, 78:993–1016.

Huang, B., Wang, W., Bates, M., and Zhuang, X. (2008). Three-dimensional super-resolution imaging by stochastic optical reconstruction microscopy. *Science*, 319(5864):810–3.

Hwang, H., Kim, H., and Myong, S. (2011). Protein induced fluorescence enhancement as a single molecule assay with short distance sensitivity. *Proceedings of the National Academy of Sciences of the United States of America*, 108(18):7414–7418.

Ip, S. C. Y., Bregu, M., Barre, F.-X., and Sherratt, D. J. (2003). Decatenation of DNA circles by FtsK-dependent Xer site-specific recombination. *The EMBO journal*, 22(23):6399–407.

Ishijima, A., Kojima, H., Funatsu, T., Tokunaga, M., Higuchi, H., Tanaka, H., and Yanagida, T. (1998). Simultaneous Observation of Individual ATPase and Mechanical Events by a Single Myosin Molecule during Interaction with Actin. *Cell*, 92(2):161–171.

Jain, A., Liu, R., Xiang, Y. K., and Ha, T. (2012). Single-molecule pull-down for studying protein interactions. *Nature protocols*, 7(3):445–52.

Johnson, R. C. and McLean, M. M. (2011). Recombining DNA by protein swivels. *Structure*, 19(6):751–3.

Kalinin, S., Sisamakis, E., Magennis, S. W., Felekyan, S., and Seidel, C. A. M. (2010). On the origin of broadening of single-molecule FRET efficiency distributions beyond shot noise limits. *The journal of physical chemistry. B*, 114(18):6197–206.

Kapanidis, A. N. (2008). Alternating-laser excitation of single molecules. In Selvin, P. R. and Ha, T., editors, *Single Molecule Techniques: A Laboratory Manual*, pages 85–119. Cold Spring Harbor Laboratory Press, 1st edition.

Kapanidis, A. N., Lee, N. K., Laurence, T. A., Doose, S., Margeat, E., and Weiss, S. (2004). Fluorescence-aided molecule sorting: analysis of structure and interactions by alternating-laser excitation of single molecules. *Proceedings of the National Academy of Sciences of the United States of America*, 101(24):8936–41.

Kapanidis, A. N. and Strick, T. (2009). Biology, one molecule at a time. *Trends in biochemical sciences*, 34(5):234–43.

Kim, S. H., Gunther, J. R., and Katzenellenbogen, J. A. (2010). Monitoring a coordinated exchange process in a four-component biological interaction system: development of a time-resolved terbium-based one-donor/three-acceptor multicolor FRET system. *Journal of the American Chemical Society*, 132(13):4685–92.

Klenow, H. and Henningsen, I. (1970). Selective elimination of the exonuclease activity of the deoxyribonucleic acid polymerase from Escherichia coli B by limited proteolysis. *Proceedings of the National Academy of Sciences of the United States of America*, 65(1):168–75.

Lakowicz, J. R. (2006). *Principles of fluorescence spectroscopy*. Springer.

Lang, M., Fordyce, P., and Block, S. (2003). Combined optical trapping and single-molecule fluorescence. *Journal of Biology*, 2(1):6.

Laurence, T. A., Kong, X., Jäger, M., and Weiss, S. (2005). Probing structural heterogeneities and fluctuations of nucleic acids and denatured proteins. *Proceedings of the National Academy of Sciences of the United States of America*, 102(48):17348–53.

Laurens, N., Rusling, D. A., Pernstich, C., Brouwer, I., Halford, S. E., and Wuite, G. J. L. (2012). DNA looping by FokI: the impact of twisting and bending rigidity on protein-induced looping dynamics. *Nucleic acids research*, 40(11):4988–97.

Lee, J., Lee, S., Ragunathan, K., Joo, C., Ha, T., and Hohng, S. (2010a). Single-Molecule Four-Color FRET. *Angewandte Chemie International Edition*, 49(51):9922–9925.

Lee, J. Y., Finkelstein, I. J., Crozat, E., Sherratt, D. J., and Greene, E. C. (2012a). Single-molecule imaging of DNA curtains reveals mechanisms of KOPS sequence targeting by the DNA translocase FtsK. *Proceedings of the National Academy of Sciences of the United States of America*, 109(17):6531–6.

Lee, J. Y., Wang, F., Fazio, T., Wind, S., and Greene, E. C. (2012b). Measuring intermolecular rupture forces with a combined TIRF-optical trap microscope and DNA curtains. *Biochemical and biophysical research communications*, 426(4):565–70.

Lee, L. and Sadowski, P. D. (2001). Directional resolution of synthetic Holliday structures by the Cre recombinase. *Journal of Biological Chemistry*, 276(33):31092–8.

Lee, L. and Sadowski, P. D. (2003a). Identification of Cre residues involved in synapsis, isomerization, and catalysis. *Journal of Biological Chemistry*, 278(38):36905–15.

Lee, L. and Sadowski, P. D. (2003b). Sequence of the loxP Site Determines the Order of Strand Exchange by the Cre Recombinase. *Journal of Molecular Biology*, 326(2):397–412.

Lee, N. K., Kapanidis, A. N., Koh, H. R., Korlann, Y., Ho, S. O., Kim, Y., Gassman, N., Kim, S. K., and Weiss, S. (2007a). Three-Color Alternating-Laser Excitation of Single Molecules: Monitoring Multiple Interactions and Distances. *Biophysical Journal*, 92:303–312.

Lee, N. K., Kapanidis, A. N., Wang, Y., Michalet, X., Mukhopadhyay, J., Ebright, R. H., and Weiss, S. (2005). Accurate FRET measurements within single diffusing biomolecules using alternating-laser excitation. *Biophysical Journal*, 88(4):2939–53.

Lee, N. K., Koh, H. R., Han, K. Y., and Kim, S. K. (2007b). Folding of 8-17 Deoxyribozyme Studied by Three-Color Alternating-Laser Excitation of Single Molecules. *Journal of the American Chemical Society*, 129(50):15526–15534.

Lee, S., Lee, J., and Hohng, S. (2010b). Single-molecule three-color FRET with both negligible spectral overlap and long observation time. *PloS one*, 5(8):9.

Lindner, M., Nir, G., Medalion, S., Dietrich, H., Rabin, Y., and Garini, Y. (2011). Force-free measurements of the conformations of DNA molecules tethered to a wall. *Physical Review E*, 83(1).

Löwe, J., Ellonen, A., Allen, M. D., Atkinson, C., Sherratt, D. J., and Grainge, I. (2008). Molecular mechanism of sequence-directed DNA loading and translocation by FtsK. *Molecular cell*, 31(4):498–509.

Lu, Y., Weers, B., and Stellwagen, N. C. (2002). DNA persistence length revisited. *Biopolymers*, 61(4):261–75.

Maier, B., Bensimon, D., and Croquette, V. (2000). Replication by a single DNA polymerase of a stretched single-stranded DNA. *Proceedings of the National Academy of Sciences of the United States of America*, 97(22):12002–7.

Manghi, M., Tardin, C., Baglio, J., Rousseau, P., Salomé, L., and Destainville, N. (2010). Probing DNA conformational changes with high temporal resolution by tethered particle motion. *Physical biology*, 7(4):046003.

Margeat, E., Kapanidis, A. N., Tinnefeld, P., Wang, Y., Mukhopadhyay, J., Ebright, R. H., and Weiss, S. (2006). Direct observation of abortive initiation and promoter escape within single immobilized transcription complexes. *Biophysical journal*, 90(4):1419–31.

Martin, S. S., Pulido, E., Chu, V. C., Lechner, T. S., and Baldwin, E. P. (2002). The order of strand exchanges in Cre-LoxP recombination and its basis suggested by the crystal structure of a Cre-LoxP Holliday junction complex. *Journal of Molecular Biology*, 319(1):107–27.

Massey, T. H., Mercogliano, C. P., Yates, J., Sherratt, D. J., and Löwe, J. (2006). Double-stranded DNA translocation: structure and mechanism of hexameric FtsK. *Molecular cell*, 23(4):457–69.

McKinney, S. A., Déclais, A.-C., Lilley, D. M. J., and Ha, T. (2002). Structural dynamics of individual Holliday junctions. *Nature Structural Biology*, 10(2):93–7.

Milstein, J. N., Chen, Y. F., and Meiners, J.-C. (2011). Bead size effects on protein-mediated DNA looping in tethered-particle motion experiments. *Biopolymers*, 95(2):144–50.

Mortensen, K. I., Churchman, L. S., Spudich, J. A., and Flyvbjerg, H. (2010). Optimized localization analysis for single-molecule tracking and super-resolution microscopy. *Nature methods*, 7(5):377–81.

Müller, B. K., Zaychikov, E., Bräuchle, C., and Lamb, D. C. (2005). Pulsed interleaved excitation. *Biophysical journal*, 89(5):3508–22.

Mumm, J. P. (2012). Single-molecule microscopy of Cre recombination. *Proceedings of the National Academy of Sciences of the United States of America*, pages 1218768110–.

Mumm, J. P., Landy, A., and Gelles, J. (2006). Viewing single lambda site-specific recombination events from start to finish. *EMBO journal*, 25(19):4586–95.

Munro, J. B., Altman, R. B., Tung, C.-S., Cate, J. H. D., Sanbonmatsu, K. Y., and Blanchard, S. C. (2010). Spontaneous formation of the unlocked state of the ribosome is a multistep process. *Proceedings of the National Academy of Sciences of the United States of America*, 107(2):709–14.

Nelson, P. C., Zurla, C., Brogioli, D., Beausang, J. F., Finzi, L., and Dunlap, D. (2006). Tethered particle motion as a diagnostic of DNA tether length. *Journal of Physical Chemistry B*, 110(34):17260–7.

Neuman, K. C. and Nagy, A. (2008). Single-molecule force spectroscopy: optical tweezers, magnetic tweezers and atomic force microscopy. *Nature methods*, 5(6):491–505.

Pavani, S. R. P., Thompson, M. A., Biteen, J. S., Lord, S. J., Liu, N., Twieg, R. J., Piestun, R., and Moerner, W. E. (2009). Three-dimensional, single-molecule fluorescence imaging beyond the diffraction limit by using a double-helix point spread function. *Proceedings of the National Academy of Sciences of the United States of America*, 106(9):2995–9.

Pease, P. J., Levy, O., Cost, G. J., Gore, J., Ptacin, J. L., Sherratt, D., Bustamante, C., and Cozzarelli, N. R. (2005). Sequence-directed DNA translocation by purified FtsK. *Science*, 307(5709):586–90.

Person, B., Stein, I. H., Steinhauer, C., Vogelsang, J., and Tinnefeld, P. (2009). Correlated movement and bending of nucleic acid structures visualized by multicolor single-molecule spectroscopy. *Chemphyschem : a European journal of chemical physics and physical chemistry*, 10(9-10):1455–60.

Pinkney, J. N. M., Zawadzki, P., Mazuryk, J., Arciszewska, L. K., Sherratt, D. J., and Kapanidis, A. N. (2012). Capturing reaction paths and intermediates in Cre-loxP recombination using single-molecule fluorescence. *Proceedings of the National Academy of Sciences of the United States of America*, pages 1211922109–.

Pouget, N., Dennis, C., Turlan, C., Grigoriev, M., Chandler, M., and Salomé, L. (2004). Single-particle tracking for DNA tether length monitoring. *Nucleic acids research*, 32(9):e73.

Ptacin, J. L., Nöllmann, M., Bustamante, C., and Cozzarelli, N. R. (2006). Identification of the FtsK sequence-recognition domain. *Nature structural & molecular biology*, 13(11):1023–5.

Ratzke, C., Berkemeier, F., and Hugel, T. (2011). Heat shock protein 90's mechanochemical cycle is dominated by thermal fluctuations. *Proceedings of the National Academy of Sciences of the United States of America*, 109(1):161–166.

Reck-Peterson, S. L., Derr, N. D., and Stuurman, N. (2010). Imaging single molecules using total internal reflection fluorescence microscopy (TIRFM). *Cold Spring Harbor protocols*, 2010(3):pdb.top73.

Ringrose, L., Lounnas, V., Ehrlich, L., Buchholz, F., Wade, R., and Stewart, A. F. (1998). Comparative kinetic analysis of Flp and Cre recombinases: mathematical models for DNA binding and recombination. *Journal of Molecular Biology*, 284(2):363–84.

Ross, J., Buschkamp, P., Fetting, D., Donnermeyer, A., Roth, C. M., and Tinnefeld, P. (2007). Multicolor single-molecule spectroscopy with alternating laser excitation for the investigation of interactions and dynamics. *The journal of physical chemistry. B*, 111(2):321–6.

Roy, R., Hohng, S., and Ha, T. (2008). A practical guide to single-molecule FRET. *Nature methods*, 5(6):507–16.

Roy, R., Kozlov, A. G., Lohman, T. M., and Ha, T. (2009). SSB protein diffusion on single-stranded DNA stimulates RecA filament formation. *Nature*, 461(7267):1092–7.

Rusling, D. A., Laurens, N., Pernstich, C., Wuite, G. J. L., and Halford, S. E. (2012). DNA looping by FokI: the impact of synapse geometry on loop topology at varied site orientations. *Nucleic acids research*, 40(11):4977–4987.

Saleh, O. A., Bigot, S., Barre, F.-X., and Allemand, J.-F. (2005). Analysis of DNA supercoil induction by FtsK indicates translocation without groove-tracking. *Nature structural & molecular biology*, 12(5):436–40.

Saleh, O. A., Pérals, C., Barre, F.-X., and Allemand, J.-F. (2004). Fast, DNA-sequence independent translocation by FtsK in a single-molecule experiment. *The EMBO journal*, 23(12):2430–9.

Santoso, Y., Joyce, C. M., Potapova, O., Le Reste, L., Hohlbein, J., Torella, J. P., Grindley, N. D. F., and Kapanidis, A. N. (2010). Conformational transitions in DNA polymerase I revealed by single-molecule FRET. *Proceedings of the National Academy of Sciences of the United States of America*, 107(2):715–20.

Sauer, B. (1993). Guide to Techniques in Mouse Development. *Methods in Enzymology*, 225:890–900.

Schafer, D. A., Gelles, J., Sheetz, M. P., and Landick, R. (1991). Transcription by single molecules of RNA polymerase observed by light microscopy. *Nature*, 352(6334):444–8.

Schwartz, J. J. and Quake, S. R. (2009). Single molecule measurement of the "speed limit" of DNA polymerase. *Proceedings of the National Academy of Sciences of the United States of America*, 106(48):20294–9.

Segall, D., Nelson, P., and Phillips, R. (2006). Volume-Exclusion Effects in Tethered-Particle Experiments: Bead Size Matters. *Physical Review Letters*, 96(8).

Shoura, M. J., Vetcher, A. A., Giovan, S. M., Bardai, F., Bharadwaj, A., Kessinger, M. R., and Levene, S. D. (2012). Measurements of DNA-loop formation via Cre-mediated recombination. *Nucleic Acids Research*.

Shroff, H., Reinhard, B. M., Siu, M., Agarwal, H., Spakowitz, A., and Liphardt, J. (2005). Biocompatible Force Sensor with Optical Readout and Dimensions of 6 nm. *Nano Letters*, 5(7):1509–1514.

Sirinakis, G., Ren, Y., Gao, Y., Xi, Z., and Zhang, Y. (2012). Combined versatile high-resolution optical tweezers and single-molecule fluorescence microscopy. *The Review of scientific instruments*, 83(9):093708.

Sobhy, M. A., Elshenawy, M. M., Takahashi, M., Whitman, B. H., Walter, N. G., and Hamdan, S. M. (2011). Versatile single-molecule multi-color excitation and detection fluorescence setup for studying biomolecular dynamics. *The Review of scientific instruments*, 82(11):113702.

Stein, I. H., Steinhauer, C., and Tinnefeld, P. (2011). Single-molecule four-color FRET visualizes energy-transfer paths on DNA origami. *Journal of the American Chemical Society*, 133(12):4193–5.

Subramanya, H. S., Arciszewska, L. K., Baker, R. A., Bird, L. E., Sherratt, D. J., and Wigley, D. B. (1997). Crystal structure of the site-specific recombinase, XerD. *The EMBO journal*, 16(17):5178–87.

Tarsa, P. B., Brau, R. R., Barch, M., Ferrer, J. M., Freyzon, Y., Matsudaira, P., and Lang, M. J. (2007). Detecting force-induced molecular transitions with fluorescence resonant energy transfer. *Angewandte Chemie (International ed. in English)*, 46(12):1999–2001.

Thompson, M. A., Lew, M. D., Badieirostami, M., and Moerner, W. E. (2010). Localizing and tracking single nanoscale emitters in three dimensions with high spatiotemporal resolution using a double-helix point spread function. *Nano letters*, 10(1):211–8.

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

Uphoff, S., Holden, S., Le Reste, L., Periz, J., Van De Linde, S., Heilemann, M., and Kapanidis, A. (2010). Monitoring multiple distances within a single molecule using switchable FRET. *Nature Methods*, 7:831–6.

Van Duyne, G. D. (2001). A structural view of Cre-loxp site-specific recombination. *Annual Review of Biophysics and Biomolecular Structure*, 30:87–104.

Vetcher, A. A., Lushnikov, A. Y., Navarra-Madsen, J., Scharein, R. G., Lyubchenko, Y. L., Darcy, I. K., and Levene, S. D. (2006). DNA topology and geometry in Flp and Cre recombination. *Journal of Molecular Biology*, 357(4):1089–104.

Vogelsang, J., Kasper, R., Steinhauer, C., Person, B., Heilemann, M., Sauer, M., and Tinnefeld, P. (2008). A reducing and oxidizing system minimizes photobleaching and blinking of fluorescent dyes. *Angewandte Chemie (International ed. in English)*, 47(29):5465–9.

Weiss, S. (1999). Fluorescence Spectroscopy of Single Biomolecules. *Science*, 283(5408):1676–1683.

Wierzbicki, A., Kendall, M., Abremski, K., and Hoess, R. (1987). A mutational analysis of the bacteriophage P1 recombinase Cre. *Journal of Molecular Biology*, 195(4):785–794.

Wikipedia: The free encyclopedia (2012). *Airy disk*. Wikimedia Foundation, Inc.

Wozniak, A. K., Schröder, G. F., Grubmüller, H., Seidel, C. A. M., and Oesterhelt, F. (2008). Single-molecule FRET measures bends and kinks in DNA. *Proceedings of the National Academy of Sciences of the United States of America*, 105(47):18337–42.

Yildiz, A., Forkey, J. N., McKinney, S. A., Ha, T., Goldman, Y. E., and Selvin, P. R. (2003). Myosin V walks hand-over-hand: single fluorophore imaging with 1.5-nm localization. *Science*, 300(5628):2061–5.

Yin, H., Landick, R., and Gelles, J. (1994). Tethered particle motion method for studying transcript elongation by a single RNA polymerase molecule. *Biophysical journal*, 67(6):2468–78.

Zareh, S. K., DeSantis, M. C., Kessler, J. M., Li, J.-L., and Wang, Y. M. (2012). Single-image diffusion coefficient measurements of proteins in free solution. *Biophysical journal*, 102(7):1685–91.

Zhuang, X., Bartley, L. E., Babcock, H. P., Russell, R., Ha, T., and Herschlag, D. (2000). A Single-Molecule Study of RNA Catalysis and Folding. *Science*, 288(5473):2048–2051.